# Un Curso de Álgebra Lineal

*con notación asociativa, su implementación en Python*

Versión: 16 de noviembre de 2026

**Marcos Bujosa**

Este libro ha sido escrito con LaTeX.

Puede encontrar la última versión de este libro en: https://github.com/mbujosab/CursoDeAlgebraLineal

Copyright © 2008–2026 Marcos Bujosa
Licencia: Creative Commons Reconocimiento-CompartirIgual 4.0 Internacional



Algunos derechos reservados. Esta obra está bajo una licencia Creative Commons Reconocimiento-CompartirIgual 4.0 Internacional. Para ver una copia de esta licencia, visite http://creativecommons.org/licenses/by-sa/4.0/ o envíe una carta a Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

"La notación lo es todo"

Charles F. Van Loan

# Prefacio

Este libro está escrito para estudiantes que cursan la asignatura de Álgebra Lineal por primera vez. No requiere conocimientos previos de la materia. Todos los resultados de álgebra lineal están demostrados a partir de las propiedades de los números reales y complejos.[1]

### El punto de vista del libro

Muchas personas perciben las matemáticas como una colección de procedimientos para resolver problemas tipo. Este libro parte de otro punto de vista: las matemáticas son un *lenguaje formal*. Un lenguaje formal consiste en un conjunto de símbolos y unas reglas para manipularlos, y hacer matemáticas es esencialmente manipular expresiones siguiendo esas reglas. Así entiendo la frase de Charles F. Van Loan que encabeza este prefacio.[2]

A las expresiones les damos interpretaciones. Las interpretaciones son importantes porque ayudan a pensar. Pero una misma expresión admite interpretaciones distintas. Por ejemplo, el Teorema de Pitágoras se emplea con distancias euclídeas, con números complejos o con la varianza de una variable aleatoria. En todos esos casos las expresiones y las reglas son las mismas, y por eso las reglas son la base común de todas las interpretaciones.

También es frecuente oír que las matemáticas están en todas las cosas. No comparto esa opinión. Comparto la de Eduardo Sáenz de Cabezón: *"las matemáticas son el lenguaje en el que nosotros leemos el mundo"*.[3] Las matemáticas son el lenguaje con el que describimos el mundo, y esas descripciones se llaman *modelos*. El lenguaje y el mundo que describimos con él son cosas distintas.

Por todo ello, el libro presta especial atención a la notación y a la manipulación de los símbolos. El libro también incluye interpretaciones geométricas, como la representación gráfica de un vector o la proyección de un vector sobre un subespacio. Aparecen a partir de la parte dedicada a la ortogonalidad, y la exposición no se basa en ellas.

### En qué se diferencia de otros manuales

El libro se diferencia de otros manuales en dos aspectos.

El primero es la notación. La notación del libro distingue los números, los vectores y las matrices. Por ejemplo, en este libro no se transpone ningún vector. Además, la notación busca que se pueda operar con los símbolos. La escritura habitual con subíndices y sumatorios describe las operaciones que hay que hacer con los números, pero no destaca las relaciones entre los vectores y las matrices. Aquí las operaciones con los números se deducen operando con los símbolos de los vectores y de las matrices. Algunas reglas de escritura no son las habituales. Espero que la claridad de la notación compense el coste de aprender una escritura distinta.

La notación está implementada en un lenguaje de programación: el módulo NAcAL de Python.[4] Las expresiones del libro se escriben casi igual en el ordenador. Con ellas el ordenador opera con vectores y

[1] La única excepción es el *Teorema Fundamental del Álgebra*, cuya demostración está fuera del alcance de este curso.

[2] La frase original es "*Notation is everything*". La tomo de Higham (1998, p. 15), que la emplea como epígrafe y la atribuye a *FFTs and the Sparse Factorization Idea* (1992).

[3] En su charla "Las matemáticas nos hacen más libres y menos manipulables", minuto 6:17.

[4] https://github.com/mbujosab/nacallib

matrices, resuelve sistemas de ecuaciones, calcula determinantes y diagonaliza matrices. Lo hace aplicando las mismas reglas que el libro.

El segundo son los procedimientos. El *método de eliminación* y las *transformaciones elementales* son las herramientas principales del curso. Muchos manuales siguen ese modelo (Cullen, 1990; Larson et al., 2004; Lay, 2007; Poole, 2004; Strang, 2007, 2003; Arvesú Carballo et al., 2005). La diferencia es que esos manuales aplican la eliminación a las filas de la matriz y este libro la aplica a las columnas. Un sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ tiene solución cuando el vector $\boldsymbol{b}$ es una combinación lineal de las columnas de $\mathbf{A}$. La eliminación por columnas busca esas combinaciones lineales de manera directa. Con un mismo procedimiento se calcula la inversa de una matriz y se resuelven sistemas de ecuaciones. La eliminación por columnas también se emplea para operar con subespacios de $\mathbb{R}^n$ y para calcular determinantes. Solo la diagonalización requiere transformar a la vez las filas y las columnas.

Estas decisiones se justifican en la Introducción. La Introducción está dirigida a quien ya conoce el álgebra lineal. El estudiante que empieza el curso puede pasar directamente a la Lección 1.

**Cómo está organizado**

Este libro nació en la asignatura Matemáticas II del Grado en Economía de la Universidad Complutense de Madrid. Mi experiencia es que los estudiantes de economía no suelen tener una especial afinidad por las matemáticas. Cuando se sienten a gusto con ellas es porque conocen una serie de algoritmos que les dan seguridad, pues es reconfortante saber que si se siguen los pasos correctamente se llega a la solución de ciertos problemas. Apoyándome en esto, la exposición de este libro es fundamentalmente algorítmica.[5] Los conceptos más abstractos se introducen a partir de los algoritmos.

Muchas demostraciones son sencillas y están propuestas como ejercicios. Con frecuencia el enunciado ofrece pistas sobre los pasos de la demostración. Esas demostraciones aparecen en la sección de soluciones a los ejercicios.[6] Así el texto es más breve y el estudiante puede demostrar los resultados por su cuenta.

Algunas lecciones terminan con una sección titulada "Generalizaciones o visiones alternativas (★)". Estas secciones tratan cuestiones más abstractas. Se pueden omitir sin que las lecciones siguientes dejen de entenderse. Espero que sirvan de referencia cuando el estudiante encuentre en otras asignaturas espacios distintos de $\mathbb{R}^n$.

Algunas lecciones tienen un apéndice. Los apéndices no se deben pasar por alto, aunque la lectura de algunos se puede posponer. Los apéndices de las lecciones 3 y 4 no son necesarios para esas lecciones, pero sí para lecciones posteriores. El apéndice de la Lección 6 es necesario para esa misma lección.

**Cómo usar este libro en un curso**

Este libro está dividido en lecciones. Para aprovechar las clases es IMPRESCINDIBLE leer con antelación la lección correspondiente, y el profesor puede dar por supuesto que usted lo ha hecho. En clase se da una exposición general de la lección, y los detalles debe prepararlos usted estudiando las secciones correspondientes.

Hay tres materiales de apoyo. El Apéndice A contiene un resumen de cada lección (**en elaboración**). Las transparencias[7] sirven de guion al profesor en el aula. Si las lleva a clase podrá atender a la explicación en lugar de copiar la pizarra. Los vídeos[8] exponen el contenido de las lecciones (**en elaboración**; de momento solo los de la primera lección). En algunos vídeos las interpretaciones geométricas tienen más peso que en el libro. Recomiendo ver además las clases de Grant Sanderson en su canal 3blue1brown.

Al final del libro están las soluciones de los ejercicios. No las mire hasta haber encontrado su propia solución. Consultar la solución de otra persona sin haber resuelto antes el ejercicio sirve de muy poco. El

[5] El desarrollo en paralelo del módulo NAcAL para Python me ha ayudado en este sentido.

[6] En la versión en `pdf` el número de cada ejercicio es un enlace a su solución. Al final de cada solución hay un cuadrado que es un enlace al enunciado.

[7] https://github.com/mbujosab/MatematicasII/tree/main/Esp

[8] https://www.youtube.com/playlist?list=PLA94O6zqDj0VREglVLepuCd0d8F0d-rnL

aprendizaje es una tarea activa.

El estudio individual o en tándem es imprescindible. Atender a las explicaciones en clase no es suficiente. Mirar las transparencias tampoco lo es. Es necesario estudiar el contenido del libro. Si alguna parte del libro no se entiende, comuníquemelo.[9] Intentaré redactarla de manera que se entienda mejor.

**Resolución de ejercicios con el ordenador.** Además de resolver los ejercicios con lápiz y papel, puede resolverlos con el ordenador empleando el módulo NAcAL de Python. El módulo está pensado para usarlo con los Notebooks de Jupyter de https://github.com/mbujosab/nacal-Jupyter-Notebooks. Aprenderá mucho si experimenta con el ordenador. No se desanime si al principio le cuesta.

El módulo se instala ejecutando `pip3 install nacal` desde un terminal.[10] El código fuente y la documentación están en

https://github.com/mbujosab/nacallib

También puede usar el módulo sin instalar nada. Para ello abra los Notebooks de Jupyter en su navegador desde la página

https://mbujosab.github.io/nacal-Jupyter-Notebooks/

que enlaza con los Notebooks de cada lección.

[9] En persona o por GitHub en https://github.com/mbujosab/CursoDeAlgebraLineal/issues
[10] Previamente debe tener instalado Python 3 en su PC.

# Índice general

**Introducción al libro** **xiii**
A quién se dirige esta introducción . . . xiii
La notación . . . xiv
Los procedimientos . . . xxviii
Ejemplos de uso de NAcAL . . . xxxii
Para terminar . . . xxxiv

**I Álgebra matricial** **1**

**1. Vectores y matrices (notación y operaciones)** **3**
1.1. Vectores . . . 3
1.1.1. Definiciones de algunos vectores especiales . . . 5
1.2. Sumas de vectores y productos de vectores por escalares . . . 6
1.2.1. Propiedades . . . 8
1.3. Vectores como representación de datos . . . 10
1.4. Matrices . . . 11
1.4.1. Filas de una matriz . . . 13
1.4.2. Columnas de una matriz . . . 13
1.4.3. Resumen de la notación para seleccionar componentes, filas y columnas . . . 15
1.4.4. Extensión de la notación vectorial y reglas de reescritura . . . 16
1.4.5. La transposición (o trasposición) . . . 17
1.4.6. Tipos de matrices . . . 18
1.5. Suma de matrices y producto de matrices por escalares . . . 20
1.5.1. Operaciones componente a componente . . . 21
1.5.2. Propiedades . . . 22
1.5.3. Operaciones fila a fila . . . 23
1.5.4. La transposición es un operador lineal . . . 23
1.6. Extensión de la notación matricial y las reglas de reescritura . . . 24
*Apendices a la lección* . . . 25
1.A. Subvectores mediante selección de una lista de índices . . . 25

**2. Combinaciones lineales** **27**
2.1. Producto punto . . . 27
2.2. Matriz por vector . . . 28
2.2.1. Combinación lineal de vectores . . . 28
2.2.2. Producto de una matriz por un vector . . . 29
2.2.3. Propiedades del producto matriz por vector . . . 30
2.3. Vector por matriz . . . 31
2.3.1. Propiedades del producto de un vector por una matriz . . . 32

**3. Multiplicación matricial** **35**
3.1. Producto de matrices . . . 35
3.1.1. Otras dos formas de calcular el producto de matrices . . . 37
Cálculo del producto de matrices componente a componente (filas por columnas) . . . 37
Cálculo del producto de matrices operando con las filas . . . 37
3.2. Nuevas reglas de reescritura. . . . 38
Transpuesta de un producto . . . 38
*Apendices a la lección* . . . 39
3.A. Productos de matrices triangulares . . . 39
3.B. Submatrices mediante selección de una lista de índices . . . 39
3.C. Producto matricial como suma de productos de submatrices . . . 40
3.C.1. Producto matricial expresado como una suma de matrices . . . 43
3.C.2. Submatriz de un producto como suma de productos de submatrices . . . 44
Matrices partidas en bloques . . . 44
3.C.3. Matriz por vector como suma de submatrices por subvectores . . . 45
3.D. Propiedades de la traza . . . 45
3.E. Elemento neutro multiplicativo universal . . . 45
3.F. Potencias de una matriz cuadrada . . . 46
3.F.1. Potencia cero de una matriz cuadrada . . . 46

**II Transformaciones elementales, métodos de eliminación y matriz inversa** **49**

**4. Transformaciones elementales y métodos de eliminación** **51**
4.1. Transformaciones y matrices elementales . . . 51
4.1.1. Transformaciones elementales (y matrices elementales) de Tipo I . . . 51
4.1.2. Transformaciones elementales (y matrices elementales) de Tipo II . . . 53
4.1.3. Transformación de las columnas de una matriz . . . 54
4.1.4. Sobre las matrices elementales y su notación . . . 54
4.1.5. Secuencias de transformaciones elementales. Parte I . . . 56
4.1.6. Intercambios . . . 57
4.1.7. Permutaciones . . . 58
4.2. Eliminación por columnas . . . 59
4.2.1. Método de eliminación (o eliminación de "izquierda a derecha") . . . 59
4.2.2. Método de eliminación gaussiana . . . 60
4.2.3. Método de eliminación Gauss-Jordan . . . 61
*Apendices a la lección* . . . 62
4.A. Transformaciones elementales de las filas . . . 62
4.A.1. Matriz escalonada por filas . . . 65
4.B. Transformación elemental "*espejo*" de otra transformación . . . 65

**5. Inversa de una matriz. Rango de una matriz** **69**
5.1. Matrices invertibles . . . 69
5.1.1. Inversa de las matrices (y las transformaciones) elementales. . . . 70
5.1.2. Matrices preescalonadas con inversa. . . . 72
5.1.3. Matrices invertibles . . . 73
5.1.4. Secuencias de transformaciones elementales. Parte II . . . 76
5.1.5. Inversa de una matriz triangular . . . 77
5.2. Rango de una matriz . . . 79
5.2.1. Algunas propiedades del rango de una matriz . . . 80
*Apendices a la lección* . . . 81
5.A. *"Espejo"* de la inversa de una secuencia de transformaciones elementales . . . 82

**III Subespacios y resolución de sistemas de ecuaciones lineales** **83**

**6. Espacios vectoriales y funciones lineales** **85**
6.1. Espacios vectoriales . . . 85
6.1.1. Algunas propiedades . . . 87
6.2. Funciones lineales . . . 88
6.2.1. Ejemplos de funciones lineales que ya hemos usado . . . 88
6.2.2. Algunas propiedades de las funciones lineales . . . 89
6.3. Creación de nuevos espacios vectoriales . . . 90
6.3.1. Subespacios . . . 90
Intersección de subespacios . . . 91
6.3.2. Espacios vectoriales de productos cartesianos . . . 91
6.3.3. Espacios vectoriales de funciones . . . 92
Ejemplos de espacios vectoriales que probablemente encontrará en otras asignaturas . . . 93
6.3.4. Subespacio de funciones lineales . . . 93
6.4. Generalizaciones o visiones alternativas (★) . . . 93
6.4.1. Subespacios y funciones lineales . . . 93
*Apendices a la lección* . . . 95
6.A. Funciones . . . 95
6.A.1. Notación . . . 96
6.A.2. Invertibilidad . . . 97
6.A.3. Composición de funciones . . . 97
6.A.4. Restricción de una función . . . 98
6.A.5. Funciones inyectivas, sobreyectivas y biyectivas . . . 98
6.A.6. Los sistemas (listas ordenadas) son funciones . . . 98

**7. Resolviendo Ax = 0** **101**
7.1. Sistemas de ecuaciones lineales . . . 101
7.1.1. Notación matricial de los sistemas de ecuaciones lineales . . . 102
7.2. Sistemas de ecuaciones lineales homogéneos . . . 102
7.3. Espacio nulo de una matriz $\mathcal{N}(\mathbf{A})$ . . . 103
7.4. Resolución de un Sistema Homogéneo por eliminación . . . 104

**8. Resolviendo Ax = b** **109**
8.1. Eliminación sobre la matriz ampliada . . . 109
8.1.1. El Teorema de Rouché-Frobenius . . . 113
8.2. Espacio columna de una matriz . . . 114
El espacio columna y la eliminación . . . 114
8.3. Generalizaciones o visiones alternativas (★) . . . 115

**9. Independencia, base y dimensión** **117**
9.1. Combinaciones lineales de vectores de $\mathcal{V}$ . . . 117
9.1.1. Extensión de la notación matricial a los espacios vectoriales . . . 117
Producto de un sistema de $n$ vectores de $\mathcal{V}$ por un vector de $\mathbb{R}^n$ . . . 117
Sistema de $n$ vectores de $\mathcal{V}$ por una matriz de $\mathbb{R}^{n\times p}$ . . . 118
9.2. Sistemas generadores . . . 119
9.3. Transformaciones elementales sobre sistemas de vectores . . . 121
Sistemas equivalentes . . . 121
9.4. Sistemas linealmente dependientes y sistemas linealmente independientes . . . 122
9.5. Bases y dimensión . . . 122
9.5.1. Eliminación "de izquierda a derecha" y sistemas "acoplados" de vectores . . . 124
Encontrando dos bases de $\mathcal{C}(\mathbf{A})$ y una base de $\mathcal{N}(\mathbf{A})$ . . . 126

9.6. Generalizaciones o visiones alternativas (★) . . . 128
Conjunto generador de un subespacio . . . 128
Sistemas de vectores y funciones lineales . . . 129
Sistemas linealmente independientes y coordenadas . . . 130
*Apendices a la lección* . . . 130

**10. Los cuatro subespacios fundamentales de una matriz A** **131**
10.1. Suma de subespacios . . . 131
10.1.1. Suma directa y subespacios suplementarios . . . 133
10.2. Los cuatro subespacios fundamentales de una matriz **A** . . . 134
10.2.1. El espacio fila . . . 134
10.2.2. El espacio nulo por la izquierda . . . 135
10.2.3. El gran esquema . . . 135
10.3. Encontrando bases para los espacios $\mathcal{C}\left(\mathbf{A}^{\intercal}\right)$ y $\mathcal{N}\left(\mathbf{A}^{\intercal}\right)$ . . . 137
Encontrando una base para el espacio fila de **A** . . . 138
Encontrando una base para el espacio nulo por la izquierda. . . . 139
10.4. Ejercicios . . . 139

**IV Ortogonalidad** **141**

**11. Vectores ortogonales. Subespacios ortogonales** **143**
11.1. Geometría . . . 143
11.1.1. Longitud de un vector en $\mathbb{R}^2$ y en $\mathbb{R}^3$ . . . 143
11.1.2. Ángulo formado por dos vectores tanto en $\mathbb{R}^2$ como en $\mathbb{R}^3$ . . . 144
11.1.3. Generalizando al espacio euclídeo $\mathbb{R}^n$ . . . 145
11.2. Subespacios ortogonales en $\mathbb{R}^n$ . . . 148
11.2.1. Ortogonalidad de los 4 subespacios fundamentales de **A** . . . 149
El método de eliminación como generador de bases del complemento ortogonal . . . 150
11.3. Generalizaciones o visiones alternativas (★) . . . 151
11.3.1. Productos escalares en $\mathcal{V}$ . . . 151
Ejemplos de productos escalares en $\mathbb{R}^n$ . . . 151
Ejemplos de productos escalares en otros subespacios . . . 152
11.3.2. Generalizando los conceptos geométricos a otros espacios vectoriales . . . 153
11.3.3. Subespacios ortogonales en espacios vectoriales abstractos . . . 154
11.3.4. Espacios pre-Hilbert . . . 154
Complementos ortogonales . . . 154
Un subespacio $\mathcal{H}$ de dimensión infinita y el conjunto $\mathcal{H}^{\perp}$ no siempre son suplementarios 156
*Apendices a la lección* . . . 157

**12. Ecuaciones cartesianas y paramétricas** **159**
12.1. De las ecuaciones paramétricas a las ecuaciones cartesianas . . . 159
12.1.1. El camino recorrido hasta ahora: de las ecuaciones cartesianas a las paramétricas . . . 159
12.1.2. Recorriendo el camino inverso: de las ecuaciones paramétricas a las cartesianas . . . 159
12.2. Puntos, rectas, planos e hiperplanos de $\mathbb{R}^n$. Espacios afines . . . 161
12.3. Escogiendo ecuaciones paramétricas (despejando variables) . . . 162

**13. Proyecciones sobre subespacios** **163**
13.1. Proyección ortogonal y mínima distancia . . . 163
13.1.1. Sistema de ecuaciones normales . . . 164
13.1.2. Alternativa a un sistema de ecuaciones lineales sin solución . . . 164
13.2. Expresión matricial de la proyección ortogonal . . . 165
13.2.1. Matrices proyección . . . 165

13.3. Bases ortogonales y proyección ortogonal. Coeficientes de Fourier . . . 166
13.4. Generalizaciones o visiones alternativas (★) . . . 166
Funciones idempotentes . . . 167
Funciones autoadjuntas . . . 167
13.4.1. Proyección ortogonal . . . 167
Sobre algunos subespacios de dimensión infinita no existe la proyección ortogonal . . . 169
13.4.2. Bases ortogonales y proyección ortogonal. Coeficientes de Fourier . . . 170
13.5. Ejercicios . . . 170

**V Determinantes 173**

**14. Propiedades de los determinantes 175**
14.1. Función determinante y función volumen . . . 175
14.1.1. Tres propiedades de la función volumen de un paralelogramo . . . 175
14.1.2. Las tres primeras propiedades (P-1 a P-3) . . . 176
La relación entre la función *Volumen* y la función *Determinante* . . . 176
14.2. Seis propiedades del determinante . . . 177
14.2.1. Determinante de una matriz con una columna de ceros . . . 177
14.2.2. Determinantes de matrices elementales . . . 177
14.2.3. Sucesión de transformaciones elementales por columnas . . . 178
Intercambio de columnas. Propiedad antisimétrica . . . 178
14.2.4. Matriz singular, matriz inversa y producto de matrices . . . 178
Determinante de matrices singulares y determinante de matrices inversas . . . 178
Determinante del producto. . . . 179
14.2.5. Determinante de la matriz transpuesta. . . . 179

**15. Fórmulas para el cálculo del determinante y aplicaciones 181**
15.1. Determinante de matrices triangulares . . . 181
15.2. Cálculo del determinante por eliminación Gaussiana . . . 182
15.3. Determinante de matrices diagonales por bloques . . . 182
15.4. Fórmulas sencillas para matrices de orden menor que 4 . . . 183
15.5. No hay fórmulas sencillas para matrices de orden mayor a 3 . . . 184
15.6. Expansión de Laplace (desarrollo por cofactores). . . . 184
15.6.1. Propiedad multilineal . . . 185
15.6.2. Menores y cofactores . . . 186
Nueva notación y definición de menores y cofactores . . . 186
15.6.3. Desarrollo del determinante por cofactores (Expansión de Laplace). . . . 187
15.7. Aplicación de los determinantes . . . 188
15.7.1. Regla de Cramer para la resolución de un sistema de ecuaciones . . . 188
15.7.2. Cálculo de la inversa de una matriz . . . 188

**VI Autovalores y autovectores. Diagonalización y formas cuadráticas 191**

**16. Autovalores y autovectores 193**
16.1. Cálculo de los autovalores y los autovectores . . . 193
16.2. Autovalores y autovectores de potencias de una matriz . . . 196

**17. Diagonalización en bloques triangulares por semejanza 199**
17.1. Matrices semejantes . . . 199
17.2. Diagonalización por bloques triangulares . . . 201
17.2.1. Autovalores, determinante y traza de matrices semejantes . . . 207

17.2.2. Autovectores . . . 207
17.3. Matrices diagonalizables . . . 208

**18. Diagonalización ortogonal de matrices simétricas** **209**
18.1. Método de Gram-Schmidt . . . 209
18.2. Matrices ortogonales . . . 210
18.3. Nota sobre la conjugación de números complejos . . . 211
18.4. Diagonalización de matrices simétricas . . . 211

**19. Diagonalización por congruencia de matrices simétricas. Formas cuadráticas** **215**
19.1. Formas cuadráticas y matrices definidas positivas . . . 215
19.1.1. Formas cuadráticas reales . . . 215
19.1.2. Matrices definidas positivas . . . 215
19.1.3. Algunos tipos de formas cuadráticas . . . 217
19.2. Diagonalización de matrices simétricas por congruencia . . . 217
19.2.1. Ley de inercia . . . 220
19.3. Completar el cuadrado para clasificar las formas cuadráticas . . . 221
*Apendices a la lección* . . . 222
19.A. Factorización de Cholesky de una matriz definida positiva . . . 222
19.B. Productos escalares en $\mathbb{R}^n$ . . . 224

**A. Resumen de los temas por lecciones** **225**
A.1. Resumen del Tema 1 . . . 225

**Soluciones a los ejercicios** **231**
Soluciones a los Ejercicios . . . 231

**Glosario** **273**

**Símbolos** **275**

**Indice alfabético** **279**

# Introducción al libro

## A quién se dirige esta introducción

Esta introducción está dirigida a quien ya conoce el álgebra lineal, y en particular a los profesores. El estudiante que empieza el curso puede pasar directamente a la Lección 1.

El libro se aparta de los manuales habituales en dos aspectos: la notación y el modo de aplicar el método de eliminación. Esta introducción explica los motivos de las dos decisiones. La sección siguiente trata la notación y su implementación en Python. Después se tratan los procedimientos. La introducción acaba con cuatro ejemplos de uso de NAcAL y una sección final.

Las matemáticas son un lenguaje en el que se han establecido ciertos convenios, y un convenio se puede cambiar cuando hay motivos para hacerlo. Lo habitual es que los autores reproduzcamos la terminología, la notación y la forma de exposición que hemos recibido. Pero el contenido de una materia y la manera de exponerlo son cuestiones distintas.

Algunos convenios de los manuales de álgebra lineal no ayudan a la exposición, aunque darse cuenta de ello no es inmediato. He llegado a esa conclusión por tres vías: mi experiencia como docente, la discusión con otros profesores y el trabajo de programar el módulo NAcAL. Para implementar la notación, los objetos y los algoritmos del curso hay que definir sin ambigüedad cada objeto y cada operación, y ese trabajo me mostró qué cambios convenía introducir.

Se puede pensar que el rigor y una notación más precisa añaden dificultad para quien empieza. Godement (1974, prefacio, pp. 19–20) sostiene lo contrario:

> *Muchas personas, y sobre todo las que se limitan a utilizar las Matemáticas, pretenden que cuando se escribe para principiantes es inútil, o incluso perjudicial, tratar de hacer gala de un gran rigor, de demostrarlo todo, de introducir nociones demasiado generales, de utilizar una terminología estrictamente definida y desprovista de frases floridas. Si estas gentes tuviesen razón resultaría que, a diferencia de los matemáticos profesionales y contra lo que es de sentido común, los principiantes comprenden tanto mejor un texto cuanto peor escrito esté.*[11]

Comparto esa opinión, porque una notación clara y precisa permite al estudiante adquirir confianza en lo que hace.

[11] La edición en inglés (Godement, 1968, p. 15) dice: *"Many people—in particular the majority of those whose attitude to mathematics is purely* utilitarian*—are of the opinion that, in textbooks designed to be read by beginners, it is useless or even dangerous to put great emphasis on rigour, to prove every statement, to introduce notions of great generality or to use carefully and strictly defined terminology. If this view were correct it would imply that, contrary to all professional mathematicians and to common sense, the worst-written textbooks would be those most easily understood by the beginner."*

# La notación

*Una buena notación debe ser inequívoca, expresiva y fácil de recordar; debe evitar los segundos significados perjudiciales y aprovechar los segundos significados útiles; el orden y la conexión de los signos deben sugerir el orden y la conexión de las cosas.*[12]

George Polya, *How to Solve It* (1957, p. 136)

En esta sección explico la notación del libro y los motivos que tengo para adoptarla. La tesis es la siguiente. Creo que la notación habitual resulta cómoda porque el producto de matrices es asociativo, y que para aprovechar esa propiedad trata los números y los vectores como matrices. La notación de este libro conserva la asociatividad y distingue los tres tipos de objetos. El argumento tiene cuatro partes.

1. Muchas operaciones del álgebra lineal se pueden realizar con productos de matrices. El producto de matrices es asociativo y eso simplifica la escritura.
2. La notación habitual aprovecha este hecho, y para lograrlo trata los números y los vectores como si fueran matrices.
3. Con la introducción de dos operadores en la notación se conservan las ventajas del producto de matrices sin confundir números, vectores y matrices.
4. La notación resultante permite deducir resultados manipulando símbolos y se puede implementar en un lenguaje de programación.

## Muchas operaciones se pueden realizar con productos de matrices

Indicaré tres ejemplos.

**Ejemplo 1: Selección de componentes.** Multipliquemos una matriz $\mathbf{A}$ por la derecha por una matriz cuya única columna es la columna $j$-ésima de la matriz identidad. El resultado es una *matriz* cuya única columna es la columna $j$-ésima de $\mathbf{A}$. Multipliquemos además por la izquierda por una matriz cuya única fila es la fila $i$-ésima de la matriz identidad. El resultado es la *matriz* cuya única componente es $a_{ij}$:

$$\text{si } \mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, \quad \text{entonces} \quad \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \end{bmatrix}. \tag{1}$$

**Ejemplo 2: Transformaciones elementales.** Restemos a la segunda fila de $\mathbf{A}$ el doble de la primera. Esta operación se puede realizar en dos pasos. Primero se aplica la misma operación a la matriz identidad, y el resultado es una matriz elemental. Después se multiplica la matriz elemental por $\mathbf{A}$:

$$\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & -2 \end{bmatrix}. \tag{2}$$

Si la matriz elemental multiplica por la derecha, la operación se aplica a las columnas de $\mathbf{A}$.

**Ejemplo 3: Producto punto.** Tomemos dos vectores de $\mathbb{R}^n$. Multipliquemos la matriz cuya única fila es el primer vector por la matriz cuya única columna es el segundo. El resultado es una *matriz* cuya única componente es el producto punto de ambos vectores:

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \end{bmatrix}. \tag{3}$$

El producto de matrices es asociativo. Por tanto, cuando una operación se realiza con un producto, los paréntesis son innecesarios. Por ejemplo, llamemos $\mathbf{E}$ a la matriz columna del Ejemplo 1. La igualdad

[12] Traducción propia. El original dice: *"A good notation should be unambiguous, pregnant, easy to remember; it should avoid harmful second meanings, and take advantage of useful second meanings; the order and connection of signs should suggest the order and connection of things."*

$(\mathbf{AB})\mathbf{E} = \mathbf{A}(\mathbf{BE})$ dice que la columna $j$-ésima de $\mathbf{AB}$ es el producto de $\mathbf{A}$ por la columna $j$-ésima de $\mathbf{B}$. El producto de matrices también es distributivo respecto de la suma. La igualdad $(\mathbf{A}+\mathbf{B})\mathbf{E} = \mathbf{AE}+\mathbf{BE}$ dice que la columna $j$-ésima de una suma es la suma de las columnas $j$-ésimas.

Estos ejemplos ayudan a entender el planteamiento de Matlab (*Matrix Laboratory*). Si todos los objetos son matrices, una gran cantidad de operaciones se realizan únicamente con el producto de matrices. Y como el producto es asociativo, las expresiones se pueden escribir sin paréntesis.

## En la notación habitual todos los objetos acaban siendo matrices

El producto de dos matrices siempre es una matriz. Por tanto, los resultados de los tres ejemplos anteriores son matrices. El primer resultado de (1) es una matriz con una columna, no un vector. Los resultados de (1) y (3) que tienen una única componente son matrices de orden 1, no números.

Sin embargo, los manuales de álgebra lineal definen operaciones entre tres tipos de objetos: los *números*, los *vectores* de $\mathbb{R}^n$ y las *matrices*. El producto de un número por un vector es un vector. El producto punto de dos vectores es un número. El producto de una matriz por un vector es un vector. El producto de dos matrices es una matriz. Los nombres de las operaciones indican que hay tres tipos de objetos.

La notación habitual resuelve esta discrepancia con dos identificaciones: las matrices con una única fila o una única columna se tratan como vectores y las matrices de orden 1 se tratan como números.

Mostraré estas identificaciones en dos manuales. Los he elegido porque son excelentes y porque están entre los más utilizados. Mi objeción no se dirige a estos libros, sino a una convención que comparten casi todos los textos de la materia.

Strang (2016, p. 2) presenta un vector de $\mathbb{R}^2$ como un par de números escrito en columna. En la página 4 indica lo siguiente:[13]

> *El vector fila* $[1\ \ 1\ \ -1]$ *es absolutamente diferente, aunque tiene las mismas tres componentes. Ese vector fila 1 por 3 es la "transpuesta" del vector columna* $\boldsymbol{v}$ *de 3 por 1.*

Así pues, en ese texto un vector de $\mathbb{R}^3$ tiene tres filas y una columna. En la página 70 se añade que un vector columna se puede considerar una matriz con una única columna.[14] Aquí un vector se trata como una matriz con una única columna.

En la página 11 se calcula el producto punto de los *vectores* $(4,2)$ y $(-1,2)$. El texto destaca que este producto es cero y lo escribe así:

$$\begin{bmatrix}4\\2\end{bmatrix}\cdot\begin{bmatrix}-1\\2\end{bmatrix} = -4+4 = 0.$$

El resultado es un número. La frase siguiente del texto lo confirma: *"En matemáticas, el cero es siempre un número especial"*.[15] En la página 17 se indica que es más frecuente escribir el producto punto como el producto de una fila por una columna. Con los vectores anteriores esa escritura es:

$$\begin{bmatrix}4 & 2\end{bmatrix}\begin{bmatrix}-1\\2\end{bmatrix}.$$

Esta segunda expresión es un producto de dos *matrices* y su resultado es la matriz [0]. En la página 71 se dice que el producto de una fila por una columna es una matriz 1 por 1 y que *"ese único número es el producto punto"*.[16] Aquí una matriz de orden 1 se trata como un número.

Poole (2004, p. 4) indica que *"con frecuencia es conveniente utilizar vectores columna en lugar de (o además de) vectores renglón"*. En la página 133 define una matriz como *"un arreglo rectangular de números"*.

[13] Traducción propia. El original dice: *"The row vector* $[1\ \ 1\ \ -1]$ *is absolutely different, even though it has the same three components. That 1 by 3 row vector is the "transpose" of the 3 by 1 column vector* $\boldsymbol{v}$*."*

[14] El original dice: *"We could even regard a column vector as a matrix with only one column (so* $n = 1$*)."*

[15] Traducción propia. El original dice: *"In mathematics, zero is always a special number."*

[16] Traducción propia. El original dice: *"Then 1 by n multiplies n by 1. The result will be 1 by 1. That single number is the 'dot product'."*

En la página 134 añade que *"una matriz de $1 \times m$ se conoce como una matriz renglón (o vector renglón) y una matriz de $n \times 1$ se conoce como una matriz columna (o vector columna)"*. Aquí una matriz con una única fila o una única columna recibe también el nombre de vector. Una nota al pie de esa página dice: *"Técnicamente, existe una distinción entre matrices renglón/columna y vectores, pero no abundaremos en esta distinción"*. La nota indica que una matriz columna y un vector no son exactamente lo mismo, pero no indica en qué consiste la diferencia. Y el texto emplea los dos nombres para el mismo objeto.

Con estas identificaciones el estudiante se puede hacer varias preguntas.

- Según la cita de Strang, un vector fila y un vector columna pueden tener las mismas componentes en el mismo orden y ser diferentes. Entonces un vector no queda determinado por sus componentes. ¿Qué más lo determina?
- ¿Los vectores de $\mathbb{R}^n$ son vectores fila o son vectores columna?
- En la expresión $\boldsymbol{y}^\intercal \mathbf{A}$ se transpone $\boldsymbol{y}$. La transposición es una operación entre matrices. ¿El símbolo $\boldsymbol{y}$ representa un vector o una matriz?
- El producto de dos matrices se define como una matriz. ¿Puede ser también un vector o un número?
- Si los números son matrices de orden 1, ¿una matriz es un arreglo rectangular de matrices?

La primera pregunta tiene respuesta en un curso más avanzado. Allí un vector columna representa un vector y un vector fila representa una forma lineal, y los dos pertenecen a espacios distintos. Los manuales citados no dan esa respuesta, porque no corresponde a un primer curso. En $\mathbb{R}^n$ con el producto punto esa distinción no es necesaria. Este libro la sustituye por la posición del vector respecto de la matriz: a la izquierda de la matriz el vector opera con sus filas y a la derecha opera con sus columnas.

Hay además un caso en el que la identificación obliga a cambiar el tipo de un objeto en mitad de una fórmula. Sean $\boldsymbol{a}$, $\boldsymbol{b}$ y $\boldsymbol{c}$ tres matrices columna con $n$ filas, y sea $n > 1$. El producto $\boldsymbol{a}^\intercal\boldsymbol{b}$ es una matriz de orden 1. El producto $\boldsymbol{c}(\boldsymbol{a}^\intercal\boldsymbol{b})$ está definido, porque $\boldsymbol{c}$ tiene una única columna. Pero el producto $(\boldsymbol{a}^\intercal\boldsymbol{b})\boldsymbol{c}$ no está definido como producto de matrices, porque $\boldsymbol{c}$ tiene $n$ filas. Para escribir $(\boldsymbol{a}^\intercal\boldsymbol{b})\boldsymbol{c}$ hay que tratar la matriz $\boldsymbol{a}^\intercal\boldsymbol{b}$ como un número. La notación no indica ese cambio.

Este cambio aparece en fórmulas habituales. Strang (2016, p. 209) escribe la proyección $\boldsymbol{p}$ de $\boldsymbol{b}$ sobre la recta que pasa por $\boldsymbol{a}$ con el cociente a la izquierda:

$$\boldsymbol{p} = \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}\,\boldsymbol{a}.$$

Después indica que la matriz se ve mejor si el número está a la derecha de $\boldsymbol{a}$,[17] y escribe:

$$\boldsymbol{p} = \boldsymbol{a}\,\widehat{x} = \boldsymbol{a}\,\frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}} = P\boldsymbol{b} \qquad \text{con} \qquad P = \frac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}.$$

En la primera fórmula $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}$ es un número, porque el producto de una matriz de orden 1 por la matriz columna $\boldsymbol{a}$ no está definido. En la segunda $\boldsymbol{a}^{\mathrm{T}}$ es un factor de la matriz $\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}$. Entre una fórmula y la otra $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}$ cambia de lado, y solo un número puede cambiar de lado.

Un lector con experiencia responde a estas preguntas por el contexto, y por ese motivo se suele considerar que las identificaciones son inofensivas. Yo creo que tienen un coste, por tres razones. La primera es que el estudiante todavía no dispone de ese contexto. La segunda es que una buena notación debe ser inequívoca. Polya (1957, p. 136) lo formula así: *"Es inadmisible que el mismo símbolo denote dos objetos diferentes en un mismo problema"*.[18] La tercera es que un programa de ordenador no puede decidir por el contexto. Para

[17] El original dice: *"You can see the matrix better if the number $\widehat{x}$ is on the right side of $\boldsymbol{a}$"*. El número $\widehat{x}$ es el cociente $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}$.

[18] Traducción propia. El original dice: *"It is inadmissible that the same symbol denote two different objects in the same inquiry."*

programar las operaciones hay que establecer de qué tipo es cada objeto. Vuelvo sobre esta cuestión en la página xxv.

**Los nombres de las filas y de las columnas.** La notación habitual necesita nombres para las filas y las columnas de una matriz. Lo usual es emplear letras con subíndices. Poole (2004, p. 134) denota las columnas de la matriz $A$ con $\mathbf{a}_1$, $\mathbf{a}_2$, ... y denota sus filas con $\mathbf{A}_1$, $\mathbf{A}_2$, ... Con estos nombres, en la página 156 demuestra que $A(B+C) = AB + AC$ del modo siguiente:

$$\begin{aligned}[A(B+C)]_{ij} &= \mathbf{A}_i \cdot (\mathbf{b}_j + \mathbf{c}_j) \\ &= \mathbf{A}_i \cdot \mathbf{b}_j + \mathbf{A}_i \cdot \mathbf{c}_j \\ &= (AB)_{ij} + (AC)_{ij} \\ &= (AB + AC)_{ij}\end{aligned} \tag{4}$$

La demostración es correcta. Se apoya en el modo de calcular la componente $ij$ del producto, que es el producto punto de la fila $i$-ésima de $A$ por la columna $j$-ésima de $B$. El resultado es una igualdad entre matrices y en ella solo intervienen objetos de un tipo. Aun así, la demostración necesita tres tipos de letra y dos clases de subíndices.

Strang (2016, p. 75) describe una forma de calcular el producto de matrices que multiplica columnas por filas:

$$AB = \begin{bmatrix} | & & | \\ \mathbf{a}_1 & \cdots & \mathbf{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} - & \mathbf{b}_1 & - \\ & \vdots & \\ - & \mathbf{b}_n & - \end{bmatrix} = \begin{bmatrix} \\ \mathbf{a}_1\mathbf{b}_1 + \cdots + \mathbf{a}_n\mathbf{b}_n \\ \\ \end{bmatrix}. \tag{5}$$

En esta ecuación $\mathbf{a}_k$ es una columna y $\mathbf{b}_k$ es una fila, pero la letra con subíndice no distingue las filas de las columnas. Por eso la ecuación emplea un recurso visual: los guiones verticales y horizontales indican cuál es cada una. Sin los guiones hay que indicar con palabras cuál es cada una. Para la factorización $LU$ la descripción emplea palabras para las filas y letras con subíndices para las componentes de la tercera fila de $L$ (Strang, 2016, p. 99):

$$(\text{Row 3 of } A) = \ell_{31}(\text{Row 1 of } U) + \ell_{32}(\text{Row 2 of } U) + 1(\text{Row 3 of } U). \tag{6}$$

Las tres escrituras son correctas, pero ninguna facilita las operaciones con las expresiones.

**La definición del producto de matrices.** La definición más frecuente es la que dan Larson et al. (2004, p. 49):

> *Si $A = [a_{ij}]$ es una matriz $m \times n$ y $B = [b_{ij}]$ es una matriz $n \times p$, el producto $AB$ es la matriz $m \times p$*
>
> $$AB = [c_{ij}]$$
>
> *donde*
>
> $$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{in}b_{nj}. \tag{7}$$

Esta definición indica cómo se calcula cada componente del producto. Cada componente es una suma de productos, y esa suma es el producto punto de una fila de $A$ por una columna de $B$. La definición solo indica eso. No indica qué relación hay entre las columnas de $AB$ y las columnas de $A$, ni qué relación hay entre las filas de $AB$ y las filas de $B$.

Una definición que describe un modo de calcular conduce a demostraciones que repiten el cálculo. Sirve de ejemplo la propiedad $(AB)^T = B^T A^T$. Larson et al. (2004, pp. 67–68) la enuncian, la comprueban con un ejemplo numérico y proponen su demostración como ejercicio. Con la definición anterior la demostración es la siguiente. Por una parte,

$$\left((AB)^T\right)_{ij} = (AB)_{ji} = \sum_{k=1}^{n} a_{jk}b_{ki}, \tag{8}$$

y por otra

$$(B^T A^T)_{ij} = \sum_{k=1}^{n} (B^T)_{ik}(A^T)_{kj} = \sum_{k=1}^{n} b_{ki}a_{jk} = \sum_{k=1}^{n} a_{jk}b_{ki}. \tag{9}$$

La segunda parte exige decidir qué componentes de $A$ y de $B$ intervienen en cada producto punto cuando las matrices están transpuestas.

En resumen, la notación habitual consigue la asociatividad tratando todos los objetos como matrices, y sus definiciones describen modos de calcular. La notación de este libro pretende cuatro cosas: conservar la asociatividad, distinguir los tres tipos de objetos, definir las operaciones por la relación entre los objetos y ser operativa. Con *operativa* quiero decir que todas las relaciones entre los objetos se escriben con símbolos y que se opera con esos símbolos, como se opera con las expresiones de un lenguaje de programación.

## Tres tipos de objetos y dos operadores

Los números, los vectores y las matrices son objetos distintos. Un vector es una lista ordenada de números. Una matriz es una lista ordenada de vectores con el mismo número de componentes. La escritura los distingue.

Los vectores $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}\ldots$ se escriben encerrando la lista de números entre paréntesis:

$$\boldsymbol{a} = (1, 2, 3, ) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Cuando el vector está escrito en horizontal, se escribe una coma detrás de cada componente. Escribir la lista en horizontal o en vertical no cambia el vector.

Las matrices $\mathsf{A}$, $\mathsf{B}$, $\mathsf{C}\ldots$ se escriben entre corchetes. La expresión $\left[\boldsymbol{a}; \boldsymbol{b}; \boldsymbol{c}; \right]$ es una lista de vectores de $\mathbb{R}^m$ entre corchetes. Denota la matriz cuyas columnas son dichos vectores. Una matriz con una única columna no es un vector:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \neq \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

En particular, la notación distingue tres objetos que la notación habitual no distingue: el número 3, el vector $(3, )$ de $\mathbb{R}^1$ y la matriz [3] de orden 1. El vector $(3, )$ es una lista con un único número. La matriz [3] es una lista con un único vector. La coma final distingue el vector $(3, )$ del número tres escrito entre paréntesis, (3). Así, el resultado de la Ecuación (3) es la matriz [3], y el producto punto de los dos vectores es el número 3. Esta distinción no es usual. El motivo para adoptarla se explica en la sección sobre la implementación de la notación en Python (página xxv).

En este libro no existen vectores fila ni vectores columna. Las palabras *fila* y *columna* siempre se refieren a una fila o a una columna de una matriz. Las filas y las columnas de una matriz son vectores. Por tanto, son listas ordenadas de números. Pero sí hay *matrices fila* y *matrices columna*. Son las matrices con una única fila y las matrices con una única columna. La transposición es una operación sobre matrices, no sobre vectores. El producto punto de dos vectores se escribe $\boldsymbol{a} \cdot \boldsymbol{b}$ y el producto de un vector por una matriz se escribe $\boldsymbol{a}\mathsf{B}$.

**Introducción de dos operadores en la notación.** Los Ejemplos 1 y 2 emplean productos de matrices para seleccionar componentes y para aplicar transformaciones elementales. La notación de este libro realiza esas dos operaciones con dos operadores. El resultado de cada operador es un objeto del tipo que corresponde: un número, un vector o una matriz.

El primer operador es el selector " $|$ ". Junto al selector se escribe el índice de la componente seleccionada. Escrito a la derecha de una matriz selecciona una columna y escrito a la izquierda selecciona una fila. Con la matriz $\mathsf{A}$ del Ejemplo 1:

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}_{|2} = \begin{pmatrix} 3 \\ 4 \end{pmatrix} \qquad \text{y} \qquad {}_{1|}\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix} = \begin{pmatrix} 1, & 3, \end{pmatrix}.$$

Los dos resultados son vectores. El primer producto de la Ecuación (1) obtiene los números 3 y 4, pero su resultado es una matriz columna. En general, $\mathsf{A}_{|j}$ es la columna $j$-ésima de $\mathsf{A}$ y ${}_{i|}\mathsf{A}$ es la fila $i$-ésima de $\mathsf{A}$.

Escrito junto a un vector, el selector selecciona una componente y el resultado es un número. Es el mismo número si el selector se escribe a la derecha o a la izquierda del vector:

$$ {}_{i|}\boldsymbol{a} \;=\; \boldsymbol{a}_{|i} \;=\; a_i. $$

El selector separa dos usos del subíndice que la notación habitual escribe igual. Con $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ se denotan tres vectores. Con $\boldsymbol{a}_{|1}$, $\boldsymbol{a}_{|2}$, $\boldsymbol{a}_{|3}$ se denotan las tres primeras componentes del vector $\boldsymbol{a}$. La componente $k$-ésima del vector $\boldsymbol{a}_h$ es $\boldsymbol{a}_{h|k}$.

Las filas y las columnas de una matriz son vectores y por tanto se puede seleccionar una componente de una fila o de una columna. Seleccionar la componente $j$-ésima de la fila $i$-ésima da el mismo número que seleccionar la componente $i$-ésima de la columna $j$-ésima:

$$ \big({}_{i|}\mathsf{A}\big)_{|j} = {}_{i|}\big(\mathsf{A}_{|j}\big). \tag{10} $$

Esta regla es análoga a la asociatividad del producto de matrices. Por eso la expresión se escribe sin paréntesis:

$$ {}_{i|}\mathsf{A}_{|j} \;=\; a_{ij}. $$

Con la matriz del Ejemplo 1 se obtiene

$$ {}_{1|}\begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}_{|2} = 3. $$

El resultado es el número 3. El segundo producto de la Ecuación (1) obtiene la matriz [3].

Cuando el argumento del selector es una lista de índices, el resultado es la submatriz formada por las filas o las columnas de la lista. Así, $\mathsf{A}_{|(2,4,)}$ es la submatriz formada por las columnas 2 y 4 de $\mathsf{A}$, y ${}_{(1,3,)|}\mathsf{A}$ es la submatriz formada por las filas 1 y 3. En particular, $\mathsf{A}_{|(j,)}$ es una matriz columna y $\mathsf{A}_{|j}$ es un vector.

Las dos selecciones se pueden aplicar a la vez. Seleccionar primero las filas y después las columnas da la misma submatriz que hacerlo en el orden contrario, y por eso la expresión también se escribe sin paréntesis. Por ejemplo, la selección de las filas 1 y 3 y de las columnas 2 y 4 de la matriz siguiente es una submatriz cuyas componentes son todas iguales a uno:

$$ {}_{(1,3,)|}\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}_{|(2,4,)} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. $$

El segundo operador es "$\boldsymbol{\tau}$" y realiza transformaciones elementales. Escrito a la derecha de una matriz transforma sus columnas y escrito a la izquierda transforma sus filas. En los dos casos el resultado es una matriz. Bajo el operador se escribe entre corchetes una abreviatura de la transformación. Cuando el operador está a la derecha de la matriz las abreviaturas son tres:

$\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}$ suma $\lambda$ veces la columna $i$-ésima a la columna $j$-ésima;

$\underset{[(\lambda)\boldsymbol{j}]}{\boldsymbol{\tau}}$ multiplica por $\lambda$ la columna $j$-ésima;

$\underset{[\boldsymbol{i}\rightleftharpoons\boldsymbol{j}]}{\boldsymbol{\tau}}$ intercambia las columnas $i$-ésima y $j$-ésima.

Cuando el operador está a la izquierda de la matriz, las mismas abreviaturas se refieren a las filas. El intercambio tiene una abreviatura propia, pero en este libro no es una transformación elemental sino una secuencia de transformaciones elementales.

La transformación del Ejemplo 2 resta a la segunda fila de $\mathbf{A}$ el doble de la primera. Con el operador se escribe así:

$$\underset{[(-2)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\begin{bmatrix}1 & 3\\ 2 & 4\end{bmatrix} = \begin{bmatrix}1 & 3\\ 0 & -2\end{bmatrix}.$$

El resultado es la matriz de la Ecuación (2).

En la notación habitual las transformaciones elementales se describen de dos maneras. La primera es una anotación junto a las matrices. Poole (2004, p. 69) la escribe sobre una flecha entre dos matrices: la anotación $R_2 - 2R_1$ indica que a la segunda fila se le resta el doble de la primera. Larson et al. (2004, p. 17) la escriben a la derecha de la matriz: la anotación $R_3 + (-2)R_1 \to R_3$ indica que a la tercera fila se le suma $-2$ veces la primera. La anotación indica la operación realizada, pero no forma parte de ninguna expresión algebraica y por tanto no se puede operar con ella. La segunda manera es el producto por una matriz elemental, como en el Ejemplo 2. Con el producto sí se puede operar, pero el nombre de una matriz elemental no indica qué transformación realiza. Para saberlo hay que mirar sus componentes.

Además, el orden de la matriz elemental depende de la matriz que se transforma. Supongamos que una matriz tiene 10 filas y 15 columnas. La matriz elemental que transforma sus filas es de orden 10 y la que transforma sus columnas es de orden 15.

El operador "$\boldsymbol{\tau}$" reúne las dos cosas. La abreviatura $\underset{[(-2)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}$ indica la operación realizada, igual que la anotación de los manuales. El operador forma parte de las expresiones y cumple reglas, igual que el producto por la matriz elemental. Y la abreviatura es la misma para cualquier matriz.

Cuando no hace falta describir la transformación se escribe solo el operador: $\mathbf{A}_{\boldsymbol{\tau}}$ es el resultado de transformar las columnas de $\mathbf{A}$ y ${}_{\boldsymbol{\tau}}\mathbf{A}$ es el resultado de transformar sus filas. Una secuencia de $k$ transformaciones se escribe $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ cuando se aplica a las columnas y ${}_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{A}$ cuando se aplica a las filas.

Por ejemplo, sea $\mathbf{A}$ una matriz de orden 7. La expresión

$$\mathbf{A}\underset{[(-2)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}\underset{[(5)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[\mathbf{1}\rightleftharpoons\mathbf{4}]}{\boldsymbol{\tau}}\underset{[(1)\mathbf{3}+\mathbf{7}]}{\boldsymbol{\tau}}$$

indica una secuencia de cuatro transformaciones de las columnas de $\mathbf{A}$. Las transformaciones se aplican en el orden en que están escritas. Primero se resta a la tercera columna el doble de la primera. Después se multiplica por 5 la segunda columna. Después se intercambian las columnas primera y cuarta. Por último se suma la tercera columna a la séptima. Con matrices elementales la misma secuencia es el producto $\mathbf{A}\mathbf{E}_1\mathbf{E}_2\mathbf{E}_3\mathbf{E}_4$.[19] Los nombres $\mathbf{E}_1$, …, $\mathbf{E}_4$ no indican qué transformaciones son, y para indicarlo hay que escribir las cuatro matrices de orden 7 o describirlas con palabras. La expresión de este libro indica las cuatro transformaciones y se puede operar con ella.

En las transformaciones de las filas la secuencia se lee de derecha a izquierda. Es el orden de un producto de matrices a la izquierda de la matriz. Por eso los índices de las transformaciones de las filas se escriben en orden decreciente: el índice indica el orden de aplicación. En los dos casos se aplica primero la transformación más próxima a la matriz.

Apliquemos a una matriz una secuencia de $p$ transformaciones de las filas y una secuencia de $q$ transformaciones de las columnas. Las transformaciones de las columnas se marcan con un asterisco para distinguirlas de las transformaciones de las filas. Transformar primero las filas y después las columnas da la misma matriz que hacerlo en el orden contrario:

$$\left({}_{\boldsymbol{\tau}_p\cdots\boldsymbol{\tau}_1}\mathbf{A}\right)_{\boldsymbol{\tau}_1^*\cdots\boldsymbol{\tau}_q^*} = {}_{\boldsymbol{\tau}_p\cdots\boldsymbol{\tau}_1}\left(\mathbf{A}_{\boldsymbol{\tau}_1^*\cdots\boldsymbol{\tau}_q^*}\right). \tag{11}$$

[19] La matriz $\mathbf{E}_3$ corresponde a un intercambio. Los manuales que emplean matrices elementales consideran elemental esa matriz. En este libro el intercambio no es una transformación elemental.

Esta regla también es análoga a la asociatividad del producto de matrices. Por eso la expresión se escribe sin paréntesis:

$$_{\boldsymbol{\tau}_p\cdots\boldsymbol{\tau}_1}\mathbf{A}_{\boldsymbol{\tau}_1^*\cdots\boldsymbol{\tau}_q^*}.$$

El operador "$\boldsymbol{\tau}$" también transforma las componentes de un vector y el resultado es un vector. Es el mismo vector si el operador se escribe a la derecha o a la izquierda:

$$_{\boldsymbol{\tau}}\boldsymbol{a} \;=\; \boldsymbol{a}_{\boldsymbol{\tau}}.$$

Con un vector las abreviaturas se refieren a sus componentes. En el primero de los dos ejemplos siguientes se multiplica por 5 la segunda componente. En el segundo se suma a la tercera componente el doble de la primera:

$$\underset{[(5)\mathbf{2}]}{\boldsymbol{\tau}}\big(1,\quad 1,\quad 1,\big) = \big(1,\quad 5,\quad 1,\big) \qquad \text{y} \qquad \big(1,\quad 1,\quad 1,\big)\underset{[(2)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}} = \big(1,\quad 1,\quad 3,\big).$$

**Las propiedades del producto se conservan.** En la presentación de los operadores han aparecido dos reglas que permiten omitir paréntesis: la regla del selector de la Ecuación (10) y la regla de las transformaciones elementales de la Ecuación (11). Las dos cumplen la misma función que la asociatividad del producto de matrices, $(\mathbf{AB})\mathbf{E} = \mathbf{A}(\mathbf{BE})$.

Los operadores cumplen otras reglas. Seleccionar una columna de una suma de matrices da la suma de las columnas seleccionadas. Transformar las columnas de una suma de matrices da la suma de las matrices transformadas:

$$\big(\mathbf{A}+\mathbf{B}\big)_{|j} = \mathbf{A}_{|j} + \mathbf{B}_{|j} \qquad \text{y} \qquad (\mathbf{A}+\mathbf{B})_{\boldsymbol{\tau}} = \mathbf{A}_{\boldsymbol{\tau}} + \mathbf{B}_{\boldsymbol{\tau}}.$$

Estas dos reglas corresponden a la distributividad del producto de matrices, $(\mathbf{A}+\mathbf{B})\mathbf{E} = \mathbf{AE} + \mathbf{BE}$.

Seleccionar una columna del producto de un número por una matriz da el producto del número por la columna seleccionada. Transformar las columnas del producto de un número por una matriz da el producto del número por la matriz transformada:

$$\big(\lambda\mathbf{A}\big)_{|j} = \lambda\big(\mathbf{A}_{|j}\big) \qquad \text{y} \qquad (\lambda\mathbf{A})_{\boldsymbol{\tau}} = \lambda(\mathbf{A}_{\boldsymbol{\tau}}).$$

Estas dos reglas corresponden a la igualdad $(\lambda\mathbf{A})\mathbf{E} = \lambda(\mathbf{AE})$. Las cuatro reglas juntas indican que los dos operadores son lineales. Las mismas reglas se cumplen cuando los operadores se escriben a la izquierda.

Por último, la transposición cambia de lado los operadores:

$$_{i|}(\mathbf{A}^\mathsf{T}) = \mathbf{A}_{|i} \qquad \text{y} \qquad (_{\boldsymbol{\tau}}\mathbf{A})^\mathsf{T} = (\mathbf{A}^\mathsf{T})_{\boldsymbol{\tau}}.$$

Con una secuencia de transformaciones la transposición además invierte el orden de la secuencia.

En el libro el producto de matrices se define después de presentar el selector " | ", y por tanto las reglas del selector no se demuestran a partir del producto. Algunas reglas son definiciones. Por ejemplo, la suma de matrices se define con la regla $\big(\mathbf{A}+\mathbf{B}\big)_{|j} = \mathbf{A}_{|j} + \mathbf{B}_{|j}$. Las demás reglas del selector se demuestran a partir de las definiciones. El operador "$\boldsymbol{\tau}$" se presenta después del producto de matrices.

Las reglas tienen la misma forma que las propiedades del producto de matrices por un motivo: según los Ejemplos 1 y 2, cada operador obtiene los mismos números que un producto por una matriz.

## El producto de matrices y la asociatividad

El producto matriz por vector, $\mathbf{A}\boldsymbol{b}$, se define como la combinación lineal de las columnas de $\mathbf{A}$ cuyos coeficientes son las componentes de $\boldsymbol{b}$:

$$\mathbf{A}\boldsymbol{b} = b_1(\mathbf{A}_{|1}) + \cdots + b_n(\mathbf{A}_{|n}).$$

Es la definición que emplea Strang (2016, p. 23). Los demás productos de este libro se definen del mismo modo o a partir de esta definición. La definición no describe un modo de calcular, sino que dice qué es el objeto $\mathbf{A}\boldsymbol{b}$: un vector que es combinación lineal de las columnas de $\mathbf{A}$. La definición habitual con sumatorios, $(A\mathbf{b})_i = \sum_j a_{ij}b_j$, describe el cálculo de cada componente, pero no destaca esa relación entre el vector y las columnas de la matriz. El modo de calcular se deduce de la definición del producto $\mathbf{A}\boldsymbol{b}$ y de la linealidad del selector:

$$\big({}_{i|}(\mathbf{A}\boldsymbol{b})\big) \;=\; b_1\big({}_{i|}\mathbf{A}_{|1}\big) + \cdots + b_n\big({}_{i|}\mathbf{A}_{|n}\big) \;=\; \big({}_{i|}\mathbf{A}\big)\cdot\boldsymbol{b}.$$

La componente $i$-ésima de $\mathbf{A}\boldsymbol{b}$ es el producto punto de la fila $i$-ésima de $\mathbf{A}$ por $\boldsymbol{b}$.

La igualdad entre el primer miembro y el último se obtiene desplazando el paréntesis de la expresión ${}_{i|}\big(\mathbf{A}\boldsymbol{b}\big)$. Al desplazarlo quedan juntos la fila ${}_{i|}\mathbf{A}$ y el vector $\boldsymbol{b}$, y los dos son vectores. El producto de dos vectores es el producto punto, y en este libro se escribe con un punto. Por eso el punto aparece al desplazar el paréntesis. La expresión sin paréntesis, ${}_{i|}\mathbf{A}\boldsymbol{b}$, ya indica cómo se calcula cada componente del producto.

El producto vector por matriz, $\boldsymbol{a}\mathbf{B}$, se define como la combinación lineal de las filas de $\mathbf{B}$ cuyos coeficientes son las componentes de $\boldsymbol{a}$:

$$\boldsymbol{a}\mathbf{B} = a_1({}_{1|}\mathbf{B}) + \cdots + a_m({}_{m|}\mathbf{B}).$$

El modo de calcular se deduce del mismo modo:

$$\big(\boldsymbol{a}\mathbf{B}\big)_{|j} \;=\; a_1\big({}_{1|}\mathbf{B}_{|j}\big) + \cdots + a_m\big({}_{m|}\mathbf{B}_{|j}\big) \;=\; \boldsymbol{a}\cdot\big(\mathbf{B}_{|j}\big).$$

Al desplazar el paréntesis quedan juntos el vector $\boldsymbol{a}$ y la columna $\mathbf{B}_{|j}$. Los dos son vectores, y el punto aparece de nuevo. La expresión sin paréntesis es $\boldsymbol{a}\mathbf{B}_{|j}$.

De la definición del producto $\mathbf{A}\boldsymbol{b}$ y de la definición del producto $\boldsymbol{a}\mathbf{B}$ se deduce que la transposición también cambia de lado al vector: $\boxed{\boldsymbol{a}\mathbf{B} = (\mathbf{B}^\intercal)\boldsymbol{a}}$.

El producto de matrices, $\mathbf{AB}$, se define como la matriz cuya columna $j$-ésima es el producto de $\mathbf{A}$ por la columna $j$-ésima de $\mathbf{B}$:

$$\big(\mathbf{AB}\big)_{|j} = \mathbf{A}\big(\mathbf{B}_{|j}\big). \tag{12}$$

Strang (2016, p. 72) presenta esta igualdad como una de sus cuatro maneras de multiplicar matrices. La Ecuación (12) es la igualdad $\big(\mathbf{AB}\big)\mathbf{E} = \mathbf{A}\big(\mathbf{BE}\big)$ del comienzo de esta sección. Allí era una consecuencia de la asociatividad del producto. Aquí es la definición del producto, y la asociatividad del producto de matrices se demuestra a partir de ella.

De la definición se deducen otras dos igualdades:

$${}_{i|}\big(\mathbf{AB}\big) = \big({}_{i|}\mathbf{A}\big)\mathbf{B} \qquad \text{y} \qquad {}_{i|}\big(\big(\mathbf{AB}\big)_{|j}\big) = \big({}_{i|}\mathbf{A}\big)\cdot\big(\mathbf{B}_{|j}\big).$$

Las tres igualdades permiten escribir la expresión ${}_{i|}\mathbf{AB}_{|j}$ sin paréntesis. Cada forma de colocar los paréntesis es una lectura correcta de la expresión:

| | |
|---|---|
| $\big(\mathbf{AB}\big)_{\|j} = \mathbf{A}\big(\mathbf{B}_{\|j}\big)$ | Las columnas de $\mathbf{AB}$ son combinaciones lineales de las columnas de $\mathbf{A}$. |
| ${}_{i\|}\big(\mathbf{AB}\big) = \big({}_{i\|}\mathbf{A}\big)\mathbf{B}$ | Las filas de $\mathbf{AB}$ son combinaciones lineales de las filas de $\mathbf{B}$. |
| ${}_{i\|}\mathbf{AB}_{\|j} = \big({}_{i\|}\mathbf{A}\big)\cdot\big(\mathbf{B}_{\|j}\big)$ | Cada componente es el producto punto de una fila de $\mathbf{A}$ por una columna de $\mathbf{B}$. |

*Una expresión sin paréntesis admite varias lecturas, y todas son correctas.*

El único signo que cambia es el punto, que se escribe cuando quedan juntos dos vectores.

La tercera lectura es el modo de calcular de la definición de Larson de la Ecuación (7). Las dos primeras lecturas son relaciones entre las matrices $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{AB}$, y la Ecuación (7) no muestra esas relaciones. Strang

(2016, pp. 71–72) describe las tres lecturas como tres maneras distintas de multiplicar matrices. Aquí las tres se obtienen de una única expresión.

Polya (1957, pp. 138–139) llama más *expresiva* a la notación que expresa más que otra. Su ejemplo es una fórmula de la que se deducen consecuencias sin mirar la figura. La expresión ${}_{i|}\mathbf{AB}_{|j}$ es expresiva en ese sentido, porque las tres lecturas anteriores se obtienen colocando los paréntesis en la expresión.

Con las transformaciones elementales ocurre lo mismo:

$$(\mathbf{AB})_{\boldsymbol{\tau}} = \mathbf{A}(\mathbf{B}_{\boldsymbol{\tau}}), \qquad {}_{\boldsymbol{\tau}}(\mathbf{AB}) = ({}_{\boldsymbol{\tau}}\mathbf{A})\mathbf{B} \qquad \text{y} \qquad {}_{\boldsymbol{\tau}}(\mathbf{AB})_{\boldsymbol{\tau}^*} = ({}_{\boldsymbol{\tau}}\mathbf{A})(\mathbf{B}_{\boldsymbol{\tau}^*}).$$

Transformar las columnas de $\mathbf{AB}$ es transformar las columnas de $\mathbf{B}$. Transformar las filas de $\mathbf{AB}$ es transformar las filas de $\mathbf{A}$. Transformar las filas y las columnas de $\mathbf{AB}$ es transformar las filas de $\mathbf{A}$ y las columnas de $\mathbf{B}$. Las tres igualdades son paralelas a las tres lecturas de la expresión ${}_{i|}\mathbf{AB}_{|j}$. Los métodos de eliminación del libro se apoyan en las dos primeras igualdades.

De la primera igualdad se deduce un resultado conocido. Basta sustituir $\mathbf{B}$ por la matriz identidad:

$$\mathbf{A}_{\boldsymbol{\tau}} \;=\; (\mathbf{AI})_{\boldsymbol{\tau}} \;=\; \mathbf{A}(\mathbf{I}_{\boldsymbol{\tau}}).$$

Transformar las columnas de $\mathbf{A}$ da la misma matriz que multiplicar $\mathbf{A}$ por la matriz $\mathbf{I}_{\boldsymbol{\tau}}$. La matriz $\mathbf{I}_{\boldsymbol{\tau}}$ es la matriz elemental de la transformación. Larson et al. (2004, p. 87) y Poole (2004, p. 170) enuncian este resultado para las filas sin demostración. Aquí se deduce en una línea. Por tanto, las matrices elementales no desaparecen con esta notación: se obtienen cuando hacen falta. El Cuadro 1 emplea esta igualdad.

## Las ecuaciones anteriores con esta notación

Reescribiré ahora el producto punto del Ejemplo 3 y dos ecuaciones de los manuales citados.

- El producto punto de la Ecuación (3):

  $$\begin{pmatrix}1, & 1,\end{pmatrix} \cdot \begin{pmatrix}2, & 1,\end{pmatrix} = 3.$$

  Es una operación entre dos vectores y su resultado es un número. No interviene ninguna matriz.

- El producto como suma de columnas por filas de la Ecuación (5):

  $$\mathbf{AB} = \sum_{k=1}^{n} \big(\mathbf{A}_{|(k,)}\big)\big({}_{(k,)|}\mathbf{B}\big).$$

  El argumento de cada selector es una lista con un único índice. Por tanto, $\big(\mathbf{A}_{|(k,)}\big)$ es una matriz columna y $\big({}_{(k,)|}\mathbf{B}\big)$ es una matriz fila. La notación indica que cada sumando es el producto de una matriz columna por una matriz fila sin recurrir a los guiones de la Ecuación (5).

- La Ecuación (6) sobre la factorización $LU$:

  $${}_{3|}\mathbf{A} \;=\; {}_{3|}\mathbf{LU} \;=\; \big({}_{3|}\mathbf{L}\big)\mathbf{U}.$$

  La tercera fila de $\mathbf{A}$ es el producto del vector ${}_{3|}\mathbf{L}$ por la matriz $\mathbf{U}$. Por tanto, es una combinación lineal de las filas de $\mathbf{U}$ cuyos coeficientes son las componentes de la tercera fila de $\mathbf{L}$.

**Tres demostraciones.** En las tres demostraciones siguientes cada igualdad aplica una única regla.

La demostración de $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$ es:

$$\begin{aligned}
\Big(\mathbf{A}\big(\mathbf{BC}\big)\Big)_{|j} &= \mathbf{A}\Big(\big(\mathbf{BC}\big)_{|j}\Big) && \text{definición del producto de matrices}\\
&= \mathbf{A}\Big(\mathbf{B}\big(\mathbf{C}_{|j}\big)\Big) && \text{definición del producto de matrices}\\
&= \Big(\mathbf{AB}\Big)\big(\mathbf{C}_{|j}\big) && \text{propiedad } \mathbf{A}(\mathbf{B}\boldsymbol{c}) = (\mathbf{AB})\boldsymbol{c}\\
&= \Big(\big(\mathbf{AB}\big)\mathbf{C}\Big)_{|j} && \text{definición del producto de matrices.}
\end{aligned}$$

El tercer paso aplica una propiedad del producto de una matriz por un vector, que el libro demuestra antes con la linealidad de ese producto. El argumento no es nuevo. Strang (2016, p. 82) propone la misma idea en un problema: demostrar la igualdad para una columna con la linealidad y aplicarla a todas las columnas de la tercera matriz. En la página 73 indica que la propiedad distributiva también se demuestra columna a columna. Lo que cambia aquí es la escritura, porque las cuatro igualdades anteriores contienen la demostración completa.

Varios manuales evitan la demostración por componentes de la asociatividad. Larson et al. (2004, p. 62) demuestran por componentes la propiedad distributiva y dejan la asociativa como ejercicio. Strang (2016, p. 73) dice que la demostración directa es algo incómoda.[20] Poole (2004, p. 214) la demuestra con la composición de transformaciones lineales y se refiere a *"la complejidad de la notación involucrada en una demostración 'por elementos'"*. Fraleigh y Beauregard (1989, p. 30) indican que la asociatividad *"es algo difícil de demostrar a partir de la definición"* y aplazan la demostración. Otros manuales sí la hacen. Cullen (1990, p. 17) y Abadir y Magnus (2005, p. 25) la demuestran con sumatorios dobles para matrices de cualquier orden, y Herstein y Winter (1989, pp. 16–17) para matrices de orden 2. Con la notación de la Ecuación (7) esa demostración es la siguiente:

$$\big[A(BC)\big]_{ij} = \sum_k a_{ik}\Big(\sum_l b_{kl}c_{lj}\Big) = \sum_k\sum_l a_{ik}b_{kl}c_{lj} = \sum_l\Big(\sum_k a_{ik}b_{kl}\Big)c_{lj} = \big[(AB)C\big]_{ij}.$$

Emplea dos sumatorios y cuatro índices, y en ella hay que intercambiar el orden de los sumatorios y seguir qué índice recorre cada uno. En la demostración anterior no interviene ningún índice de suma. La suma queda en la propiedad del producto de una matriz por un vector, que se demuestra una vez y se usa después sin índices.

La demostración de $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$ es:

$$\begin{aligned}
\Big(\mathbf{A}(\mathbf{B}+\mathbf{C})\Big)_{|j} &= \mathbf{A}\Big((\mathbf{B}+\mathbf{C})_{|j}\Big) && \text{definición del producto de matrices}\\
&= \mathbf{A}\Big(\mathbf{B}_{|j}+\mathbf{C}_{|j}\Big) && \text{definición de la suma de matrices}\\
&= \mathbf{A}\big(\mathbf{B}_{|j}\big)+\mathbf{A}\big(\mathbf{C}_{|j}\big) && \text{linealidad del producto matriz por vector}\\
&= \big(\mathbf{AB}\big)_{|j}+\big(\mathbf{AC}\big)_{|j} && \text{definición del producto de matrices}\\
&= \big(\mathbf{AB}+\mathbf{AC}\big)_{|j} && \text{definición de la suma de matrices.}
\end{aligned}$$

Tiene un paso más que la demostración (4), y en las dos cada igualdad aplica una regla con nombre. En esta propiedad las dos notaciones quedan a la par. La diferencia es que aquí solo aparecen las tres matrices y el selector.

La demostración de $\big(\mathbf{AB}\big)^{\mathsf{T}} = (\mathbf{B}^{\mathsf{T}})(\mathbf{A}^{\mathsf{T}})$ es:

$$\begin{aligned}
\big(\mathbf{AB}\big)^{\mathsf{T}}{}_{|j} &= {}_{j|}\big(\mathbf{AB}\big) && \text{la transposición cambia de lado el selector}\\
&= \big({}_{j|}\mathbf{A}\big)\mathbf{B} && \text{la fila de un producto}\\
&= \Big((\mathbf{A}^{\mathsf{T}})_{|j}\Big)\mathbf{B} && \text{la transposición cambia de lado el selector}\\
&= (\mathbf{B}^{\mathsf{T}})\Big((\mathbf{A}^{\mathsf{T}})_{|j}\Big) && \text{la transposición cambia de lado el vector}\\
&= \Big((\mathbf{B}^{\mathsf{T}})(\mathbf{A}^{\mathsf{T}})\Big)_{|j} && \text{definición del producto de matrices.}
\end{aligned}$$

Compárese con las ecuaciones (8) y (9). Esta demostración tampoco es más breve, pero no calcula ninguna componente del producto y no emplea ningún índice de suma. Cada paso aplica una regla de reescritura. En la demostración con sumatorios hay que decidir qué componentes intervienen cuando las matrices están transpuestas.

[20] El original dice: *"The direct proof is sort of awkward (Problem 37) but this law is extremely useful."*

## La implementación en Python: el módulo NAcAL

He implementado la notación en el módulo NAcAL para Python. El código es una transcripción de las definiciones del libro. Dos ejemplos lo muestran.

**Ejemplo 1: La suma de dos vectores de $\mathbb{R}^n$.** La suma se define en la Lección 1 como el vector cuya componente $i$-ésima es la suma de las componentes $i$-ésimas:

$$(\boldsymbol{a}+\boldsymbol{b})_{|i} = \boldsymbol{a}_{|i} + \boldsymbol{b}_{|i} \quad \text{para} \quad i = 1:n,$$

y su implementación en NAcAL es, en esencia:

```
Vector([ (self|i) + (other|i) for i in range(1, len(self)+1) ])
```

donde `self` es $\boldsymbol{a}$, `other` es $\boldsymbol{b}$ y `len(self)` es el número de componentes de $\boldsymbol{a}$.[21]

**Ejemplo 2: El producto de matrices.** El producto se define en la Lección 3 como la matriz cuya columna $j$-ésima es el producto de la primera matriz por la columna $j$-ésima de la segunda:

$$(\mathbf{AX})_{|j} = \mathbf{A}(\mathbf{X}_{|j}) \quad \text{para} \quad j = 1:n,$$

y su implementación en NAcAL es, en esencia:

```
Matrix([ self*(x|j) for j in range(1, x.n+1) ])
```

donde `self` es $\mathbf{A}$, `x` es $\mathbf{X}$ y `x.n` es el número de columnas de $\mathbf{X}$.

**El número 3, el vector $(3,)$ y la matriz $[3]$.** Cuando se escribe a mano se puede trabajar sin distinguir estos tres objetos. El lector resuelve por el contexto cuál es cada uno, pero un programa no puede hacerlo. En NAcAL los vectores y las matrices son objetos de dos clases distintas: `Vector` y `Matrix`. Por tanto, para cada operación hay que establecer de qué clase es el resultado. Por ejemplo, el producto punto de dos vectores es un número, la selección de una columna de una matriz es un vector y el producto de dos matrices es una matriz. Con estas reglas el número 3, el vector $(3,)$ y la matriz $[3]$ son resultados de operaciones distintas, y el programa los trata como objetos distintos. Una manera de no distinguirlos es que todos los objetos sean matrices, como en Matlab. Eso es lo que esta notación quiere evitar. En Matlab una matriz de orden 1 actúa además como un número, y así se puede calcular `(a'*b)*c`. Esa regla es el cambio de tipo descrito en la página xvi. Con ella el producto deja de ser asociativo, porque `(a'*b)*c` está definido y `a'*(b*c)` no lo está.

La biblioteca NumPy de Python también lo evita. NumPy es una biblioteca de cálculo numérico muy extendida. En NumPy una lista de números y una tabla con una única columna son objetos distintos, como aquí un vector y una matriz columna. La expresión `A[:, 1]` es una lista de números y la expresión `A[:, [1]]` es una tabla con una única columna. Del mismo modo, $\mathbf{A}_{|j}$ es un vector y $\mathbf{A}_{|(j,)}$ es una matriz columna. Y el producto `a @ b` de dos listas de números es un número, sin transponer nada. Por tanto, un programa de cálculo numérico no necesita que todos los objetos sean matrices.

La distinción se puede formalizar. Un convenio habitual define los espacios $\mathbb{R}^n$ de manera recursiva a partir de $\mathbb{R}$:

$$\begin{cases} \mathbb{R}^1 = \mathbb{R} \\ \mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n \end{cases}.$$

Con este convenio un vector de $\mathbb{R}^1$ es un número. Este libro sigue un convenio distinto, que parte de $\mathbb{R}^0$:

$$\begin{cases} \mathbb{R}^0 = \{\emptyset\} \\ \mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n \end{cases},$$

[21] Python cuenta empezando en cero: `range(n)` recorre $n$ índices, de 0 a $n-1$. Del mismo modo, `range(1, n)` va de 1 a $n-1$. Para recorrer los índices de 1 a $n$ hay que escribir `range(1, n+1)`. Lo mismo ocurre con `x.n+1` en el ejemplo siguiente.

donde $\emptyset$ denota el conjunto vacío. Aquí el conjunto vacío es la lista sin componentes. Con este convenio $\mathbb{R}^1 = \mathbb{R} \times \{\emptyset\} = \{(\alpha, \emptyset) \mid \alpha \in \mathbb{R}\}$. Un vector de $\mathbb{R}^1$ es un par formado por un número y la lista vacía, y por tanto no es un número.

En general, un vector de $\mathbb{R}^{n+1}$ es un par $(v_1, \boldsymbol{w})$ formado por un número $v_1$ y un vector $\boldsymbol{w}$ de $\mathbb{R}^n$. El selector se define de manera recursiva sobre esos pares:

$$\begin{cases} (v_1, \boldsymbol{w})_{|1} = v_1 \\ (v_1, \boldsymbol{w})_{|k+1} = \boldsymbol{w}_{|k} \end{cases}.$$

La escritura $(1, 2, 3, )$ abrevia el par $(1, (2, (3, \emptyset)))$. La coma final es la coma que separa la última componente de la lista vacía.[22]

Una matriz es una lista de vectores. Por tanto, la matriz $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ no es una lista de tres números. Es la lista de tres vectores de $\mathbb{R}^1$: $\big[(1, );\ (2, );\ (3, );\big]$.

La coma final tampoco es una invención del libro, sino la escritura de Python. En Python `(3)` es el número 3 entre paréntesis y `(3,)` es una tupla con un único elemento. Una tupla es una lista ordenada, y la tupla vacía se escribe `()`. La coincidencia está en la escritura. Python no guarda una tupla como un par formado por su primer elemento y el resto.

El código de los dos ejemplos anteriores emplea la misma barra vertical que el libro y es tan breve como las definiciones. Las expresiones del libro se escriben casi igual en el ordenador, y el ordenador las evalúa. Un programa solo puede evaluar una operación si está definida para los tipos de sus operandos. Por tanto, el módulo obliga a que cada operación tenga un resultado de un tipo determinado.

Los algoritmos del libro también están transcritos en el módulo. Las demostraciones por inducción sobre la eliminación y la diagonalización son constructivas: cada paso de la demostración indica una operación con la matriz. El código de NAcAL ejecuta esas mismas operaciones en el mismo orden.

## La notación y los criterios de Polya

La cita de Polya que encabeza esta sección enumera varios criterios. Repaso ahora cómo los cumple la notación del libro.

*Inequívoca.* Los números, los vectores y las matrices se escriben de manera distinta y ningún símbolo denota objetos de dos tipos.

*Expresiva.* De la expresión ${}_{i|}\mathbf{AB}_{|j}$ se obtienen tres lecturas colocando los paréntesis.

*Fácil de recordar.* Una única regla de posición sirve para los dos operadores, para el producto y para la transposición. A la izquierda de una matriz se actúa sobre sus filas y a la derecha se actúa sobre sus columnas. La transposición cambia de lado.

*Sin segundos significados perjudiciales.* En la notación habitual el subíndice tiene varios significados. En $\mathbf{a}_1$, $\mathbf{a}_2$ distingue los vectores de una lista o indica las columnas de la matriz $A$. En $a_1$, $a_2$ indica las componentes del vector $\mathbf{a}$. En $I_3$ indica en algunos textos el orden de la matriz identidad. Las columnas de una matriz y las componentes de un vector solo se diferencian por el tipo de letra. Con el selector cada significado se escribe de manera distinta: $\boldsymbol{a}_1$ y $\boldsymbol{a}_2$ son dos vectores de una lista, $\boldsymbol{a}_{|1}$ y $\boldsymbol{a}_{|2}$ son dos componentes del vector $\boldsymbol{a}$, y $\mathbf{A}_{|1}$ y $\mathbf{A}_{|2}$ son dos columnas de la matriz $\mathbf{A}$. La barra vertical tiene otros usos en matemáticas: indica "tal que" en la definición de un conjunto y también denota el valor absoluto y el determinante. El selector se distingue de ellos porque se escribe a la altura de los subíndices y junto a un índice.

*Con segundos significados útiles.* Un operador escrito junto a una matriz recuerda a un factor de un producto. Esa asociación es correcta, porque los operadores cumplen las mismas reglas que un factor: las

[22] Esta construcción es la de las listas de los lenguajes de programación Lisp y Haskell. En ellos una lista no vacía es un par formado por su primer elemento y la lista de los elementos restantes.

reglas del apartado "Las propiedades del producto se conservan". Quien aplica a los operadores esas reglas del producto de matrices no se equivoca.

*El orden y la conexión de los signos sugieren el orden y la conexión de las cosas.* En la igualdad $\left(\mathbf{AB}\right)_{|j} = \mathbf{A}\left(\mathbf{B}_{|j}\right)$ el selector queda escrito junto al factor sobre el que actúa. En la expresión ${}_{i|}\mathbf{AB}_{|j}$ la fila de $\mathbf{A}$ y la columna de $\mathbf{B}$ quedan escritas en el orden en que se multiplican.

## Los costes de una notación distinta

Adoptar una notación que no es la habitual tiene costes. Los indico a continuación.

- El lector que consulte otros textos debe pasar de una notación a otra. El Cuadro 1 recoge las equivalencias principales.
- La coma detrás de la última componente de un vector y el selector a la izquierda de una matriz son escrituras nuevas y hay que acostumbrarse a ellas.
- El estudiante que continúe con otras asignaturas tendrá que aprender la notación habitual.
- El profesor no puede apoyar sus clases en los manuales habituales. Ese fue el motivo por el que escribí este libro. Para escribir sus propios materiales con esta notación en LaTeX puede descargar un paquete de macros desde https://github.com/mbujosab/nacal-latex-package.

Polya (1957, p. 141) advierte de este coste: *"aprender una notación nueva es una carga para la memoria. El estudiante inteligente se niega a asumir la carga si no ve ninguna compensación".*[23] Creo que las ventajas compensan estos costes. El lector puede juzgarlo con las tres demostraciones de esta sección y con las del resto del libro.

| Objeto | Este libro | Notación habitual |
|---|---|---|
| Vector escrito en horizontal | $(1,2,3,)$ | $(1,2,3)$ |
| Componente $i$-ésima de un vector | $\boldsymbol{a}_{\|i}$ o ${}_{i\|}\boldsymbol{a}$ | $a_i$ |
| Columna $j$-ésima de una matriz | $\mathbf{A}_{\|j}$ | $\mathbf{a}_j$, $A_{*j}$, `A(:,j)` |
| Fila $i$-ésima de una matriz | ${}_{i\|}\mathbf{A}$ | $\mathbf{a}_i^T$, $A_{i*}$, `A(i,:)` |
| Componente de la fila $i$ y la columna $j$ | ${}_{i\|}\mathbf{A}_{\|j}$ | $a_{ij}$ |
| Submatriz con las columnas 2 y 4 | $\mathbf{A}_{\|(2,4,)}$ | `A(:,[2 4])` |
| Producto punto de dos vectores | $\boldsymbol{a}\cdot\boldsymbol{b}$ | $\mathbf{a}^T\mathbf{b}$ |
| Producto de un vector por una matriz | $\boldsymbol{a}\mathbf{B}$ | $\mathbf{a}^T B$ |
| Forma cuadrática | $\boldsymbol{x}\mathbf{A}\boldsymbol{x}$ | $\mathbf{x}^T A\mathbf{x}$ |
| Una transformación elemental de las columnas | $\mathbf{A}_{\boldsymbol{\tau}}$ | $AE$ |
| Una transformación elemental de las filas | ${}_{\boldsymbol{\tau}}\mathbf{A}$ | $EA$ |
| Secuencia de transformaciones de las columnas | $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ | $AE_1\cdots E_k$ |
| Secuencia de transformaciones de las filas | ${}_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{A}$ | $E_k\cdots E_1A$ |

Cuadro 1: Equivalencias entre la notación del libro y la notación habitual. En las cuatro últimas filas $E$ y $E_i$ son matrices elementales. En las transformaciones de las columnas $E_i = \mathbf{I}_{\boldsymbol{\tau}_i}$. En las transformaciones de las filas $E_i = {}_{\boldsymbol{\tau}_i}\mathbf{I}$.

[23] Traducción propia. El original dice: *"to learn a new notation is a burden for the memory. The intelligent student refuses to assume the burden if he does not see any compensation for it."*

# Los procedimientos

El algoritmo más usado en los manuales de álgebra lineal es el método de eliminación. Los manuales lo aplican a las filas de la matriz y emplean dos variantes: la eliminación gaussiana escalona la matriz y la eliminación Gauss-Jordan además la reduce.

Este libro aplica la eliminación a las columnas. Además distingue tres métodos de eliminación y solo emplea en cada problema los pasos que ese problema requiere. En esta sección explico los motivos.

## La eliminación por columnas

Consideremos el siguiente sistema de ecuaciones lineales:

$$\begin{cases} \quad 2y+2z+6w=0 \\ 2x+2y \quad -2w=2 \end{cases}.$$

**La resolución por filas.** Esta resolución tiene dos partes. La primera parte es la eliminación. Con la eliminación Gauss-Jordan se escalona y se reduce la matriz ampliada, y el resultado es la matriz ampliada de otro sistema con el mismo conjunto de soluciones:

$$\left[\begin{array}{cccc|c} 0 & 2 & 2 & 6 & 0 \\ 2 & 2 & 0 & -2 & 2 \end{array}\right] \xrightarrow[[\mathbf{1}\rightleftharpoons\mathbf{2}]]{\tau} \left[\begin{array}{cccc|c} 2 & 2 & 0 & -2 & 2 \\ 0 & 2 & 2 & 6 & 0 \end{array}\right] \xrightarrow[[(-1)\mathbf{2}+\mathbf{1}]]{\tau} \left[\begin{array}{cccc|c} 2 & 0 & -2 & -8 & 2 \\ 0 & 2 & 2 & 6 & 0 \end{array}\right] \xrightarrow[\substack{[(\frac{1}{2})\mathbf{1}]\\[(\frac{1}{2})\mathbf{2}]}]{\tau} \left[\begin{array}{cccc|c} 1 & 0 & -1 & -4 & 1 \\ 0 & 1 & 1 & 3 & 0 \end{array}\right].$$

La segunda parte es la lectura de la solución. Hay que identificar las variables pivote y las variables libres, y después hay que despejar las variables pivote:

$$\begin{cases} x \quad = 1+z+4w \\ \quad y=-z-3w \end{cases}.$$

Por último hay que dar valores a las variables libres. Con $z=0$ y $w=0$ se obtiene una solución particular, $\boldsymbol{x}_p = \begin{pmatrix} 1, & 0, & 0, & 0, \end{pmatrix}$. Con $z=1$ y $w=0$ se obtiene una solución del sistema homogéneo, $\boldsymbol{x}_a = \begin{pmatrix} 1, & -1, & 1, & 0, \end{pmatrix}$. Con $z=0$ y $w=1$ se obtiene otra, $\boldsymbol{x}_b = \begin{pmatrix} 4, & -3, & 0, & 1, \end{pmatrix}$. El conjunto de soluciones es

$$\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists a,b \in \mathbb{R},\ \boldsymbol{v} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + a \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + b \begin{pmatrix} 4 \\ -3 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Si se emplea la eliminación gaussiana, la matriz no queda reducida y la segunda parte requiere además la sustitución hacia atrás.

**La resolución por columnas.** Esta resolución tiene una única parte. Se parte de la matriz ampliada de la resolución por filas con un cambio: su última columna es el opuesto del vector del lado derecho, $-\boldsymbol{b}$. El motivo del cambio de signo se explica más abajo. Esta matriz se amplía hacia abajo: debajo de ella se escribe una matriz identidad con el mismo número de columnas. En el ejemplo la matriz identidad es de orden 5 y el resultado es una matriz con siete filas y cinco columnas.

La eliminación consiste en anular las componentes que están a la derecha de los pivotes de la matriz ampliada original. El pivote de una columna es su primera componente no nula. La eliminación es secuencial: primero se anulan las componentes que están a la derecha del pivote de la primera columna, después las que están a la derecha del pivote de la segunda columna, y así hasta la última columna. Para anular esas componentes se aplican transformaciones elementales a las columnas de la matriz completa, de modo que cada transformación actúa sobre las siete filas a la vez. Sobre cada flecha se indican las transformaciones

aplicadas:

$$\left[\begin{array}{ccccc|c} 0&2&2&6&0&\\ 2&2&0&-2&-2&\\ \hline 1&0&0&0&0&\\ 0&1&0&0&0&\\ 0&0&1&0&0&\\ 0&0&0&1&0&\\ \hline 0&0&0&0&1& \end{array}\right] \xrightarrow[\substack{[(1)\mathbf{1}+\mathbf{5}]\\ [(1)\mathbf{1}+\mathbf{4}]\\ [(-1)\mathbf{1}+\mathbf{2}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccccc|c} 0&2&2&6&0&\\ 2&0&0&0&0&\\ \hline 1&-1&0&1&1&\\ 0&1&0&0&0&\\ 0&0&1&0&0&\\ 0&0&0&1&0&\\ \hline 0&0&0&0&1& \end{array}\right] \xrightarrow[\substack{[(-3)\mathbf{2}+\mathbf{4}]\\ [(-1)\mathbf{2}+\mathbf{3}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccccc|c} 0&2&0&0&0&\\ 2&0&0&0&0&\\ \hline 1&-1&1&4&1&\\ 0&1&-1&-3&0&\\ 0&0&1&0&0&\\ 0&0&0&1&0&\\ \hline 0&0&0&0&1& \end{array}\right]$$

Las rectas no forman parte de la matriz y solo ayudan a leerla. La recta horizontal superior separa la matriz ampliada original de la matriz identidad. La recta vertical separa la última columna, como en la resolución por filas. La recta horizontal inferior separa la última fila. El motivo de esta separación se explica más abajo. Las cuatro filas que quedan entre las dos rectas horizontales corresponden a las cuatro incógnitas.

La solución se lee en la matriz final, en las cuatro filas que quedan entre las dos rectas horizontales. En la matriz final son nulas la última columna de la matriz ampliada y las columnas tercera y cuarta. Bajo la última columna está una solución particular. Bajo las columnas tercera y cuarta están dos soluciones del sistema homogéneo. Esas dos soluciones forman una base del conjunto de soluciones del sistema homogéneo. Obsérvese que los tres vectores coinciden con los vectores $\boldsymbol{x}_p$, $\boldsymbol{x}_a$ y $\boldsymbol{x}_b$ obtenidos con la eliminación por filas. No hay que identificar variables libres ni despejar.

Llamo *pre-escalonar* una matriz a anular las componentes que están a la derecha de cada pivote. Una matriz está *pre-escalonada* cuando las componentes a la derecha de cada pivote son nulas. Una matriz está *escalonada* cuando está pre-escalonada, sus pivotes describen una escalera descendente y sus columnas nulas están a la derecha de las columnas no nulas. Para resolver un sistema de ecuaciones no es necesario escalonar la matriz, porque basta pre-escalonarla.

Las dos eliminaciones operan con objetos distintos. Un sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ tiene solución cuando $\boldsymbol{b}$ es una combinación lineal de las columnas de $\mathbf{A}$, y cada solución contiene los coeficientes de una de esas combinaciones. La eliminación por columnas opera con las columnas de $\mathbf{A}$. Por tanto, trabaja en el espacio columna y busca esas combinaciones de manera directa. La eliminación por filas opera con las ecuaciones. Cada ecuación es una forma lineal sobre $\mathbb{R}^n$, de modo que la eliminación por filas trabaja en el espacio dual. Su resultado es otro sistema de ecuaciones con las mismas soluciones.

Esta diferencia tiene una consecuencia. La eliminación por filas obtiene en cada paso un sistema de ecuaciones equivalente al de partida. Con la eliminación por columnas no ocurre así: las matrices de los pasos intermedios no corresponden a sistemas equivalentes, y las soluciones encontradas solo lo son del sistema original. El papel de la matriz identidad se explica a continuación.

**La ventaja didáctica principal.** La ventaja principal no es el ahorro de pasos. Es que un mismo procedimiento resuelve tres problemas: invertir una matriz, resolver un sistema homogéneo y resolver un sistema no homogéneo. En los tres se escribe una matriz identidad debajo de la matriz, se aplica la eliminación a las columnas y el resultado se lee en la parte inferior. Para la inversa hay que continuar la eliminación hasta reducir la matriz. Con la eliminación por filas la inversa también se lee en la matriz final, pero en los sistemas de ecuaciones hay que identificar después las variables libres y despejar, y ese paso no forma parte de la eliminación.

Además una misma explicación justifica el procedimiento en los tres problemas. El estudiante que entiende por qué se obtiene la inversa ya puede entender los otros dos. La explicación emplea las igualdades siguientes:

$$\left[\frac{\mathbf{A}}{\mathbf{I}}\right]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \left[\frac{\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}{\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}\right] = \left[\frac{\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)}{\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}\right] \qquad \text{y} \qquad \left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)_{|j} = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)_{|j}.$$

La primera igualdad de la izquierda indica que transformar las columnas de la matriz completa da el mismo resultado que aplicar las mismas transformaciones a su parte superior y a su parte inferior. La parte superior

procede de la matriz original y la parte inferior es la matriz identidad. Se cumple porque una transformación de las columnas actúa sobre todas las filas de la matriz.

La segunda igualdad de la izquierda emplea la igualdad $\mathbf{A}_{\boldsymbol{\tau}} = \mathbf{A}(\mathbf{I}_{\boldsymbol{\tau}})$ de la sección anterior. Esa igualdad se aplica a cada transformación de la secuencia. El resultado indica el papel de la matriz identidad: en cada paso la matriz de arriba es el producto de $\mathbf{A}$ por la matriz de abajo.

La igualdad de la derecha dice lo mismo columna a columna. Cada columna de la matriz de arriba es el producto de $\mathbf{A}$ por la columna correspondiente de la matriz de abajo. Se deduce de la definición del producto de matrices. De estas igualdades se deduce el resultado en los tres problemas.

*El sistema homogéneo.* Bajo cada columna nula de la matriz de arriba hay una solución de $\mathbf{A}\boldsymbol{x} = \mathbf{0}$. Esas soluciones son linealmente independientes porque la matriz de abajo es invertible: se obtiene de la matriz identidad con transformaciones elementales. La Proposición 7.4.2 demuestra que toda solución del sistema es combinación lineal de ellas.

*La inversa.* Supongamos que las transformaciones convierten $\mathbf{A}$ en la matriz identidad. Entonces

$$\mathbf{A}\big(\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}\big) = \mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} = \mathbf{I}.$$

Por tanto, la matriz de abajo es la inversa de $\mathbf{A}$.

*El sistema no homogéneo.* El sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ se convierte en un sistema homogéneo con una incógnita más:

$$\left[\begin{array}{c|c} \mathbf{A} & -\boldsymbol{b} \end{array}\right] \begin{pmatrix} \boldsymbol{x} \\ 1 \end{pmatrix} = \mathbf{A}\boldsymbol{x} - \boldsymbol{b} = \mathbf{0}.$$

Por eso la última columna de la matriz ampliada es el opuesto de $\boldsymbol{b}$. La barra vertical y la raya horizontal de esta igualdad son la recta vertical y la recta horizontal inferior del ejemplo. La última fila se separa porque corresponde al 1 del vector y no a una incógnita. Las soluciones de $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ se obtienen de las soluciones del sistema homogéneo cuya última componente es 1. En el ejemplo, bajo la última columna de la matriz final hay un vector cuya última componente es 1. Sus cuatro primeras componentes son la solución particular.

El razonamiento es completamente algebraico y es nuevo para el estudiante. Los estudiantes llegan a la universidad con procedimientos aprendidos en la enseñanza secundaria y con frecuencia no saben por qué funcionan. Con un procedimiento distinto pretendo que el estudiante no aplique un procedimiento aprendido sin entenderlo. La advertencia del comienzo de la Lección 7 trata esta cuestión.

La eliminación por columnas tiene otra ventaja, que afecta a la escritura de las secuencias de transformaciones. Una secuencia de transformaciones de las columnas se escribe a la derecha de la matriz y se lee de izquierda a derecha, en el orden en que se aplican las transformaciones. Con la eliminación por filas las transformaciones se acumulan a la izquierda de la matriz y la secuencia se lee de derecha a izquierda. Ocurre lo mismo con el producto de matrices elementales $E_k \cdots E_1 A$ de los manuales. Este orden no depende de la notación, porque es una consecuencia de operar con las filas. La diferencia solo aparece en estas expresiones. Una cadena de matrices unidas por flechas se lee de izquierda a derecha en los dos casos.

Al pre-escalonar una matriz, cada columna se modifica solo con las columnas que están a su izquierda. Este modo de proceder tiene dos consecuencias: la matriz que aparece debajo es triangular superior unitaria, y las $j$ primeras columnas de la matriz pre-escalonada generan el mismo subespacio que las $j$ primeras columnas de la matriz de coeficientes, para cualquier $j$. De estos dos hechos se deducen de manera sencilla varios resultados.

- Si la columna $j$-ésima de la matriz pre-escalonada es nula, la columna $j$-ésima de la matriz de coeficientes es combinación lineal de las columnas que están a su izquierda. La columna que queda debajo contiene los coeficientes de una combinación lineal nula de las $j$ primeras columnas de la matriz de coeficientes, y de esa combinación se despeja la columna $j$-ésima.

- Las columnas no nulas de la matriz pre-escalonada forman una base del espacio columna de la matriz de coeficientes.
- Con más generalidad, las columnas no nulas que hay entre las $j$ primeras columnas de la matriz pre-escalonada forman una base del subespacio generado por las $j$ primeras columnas de la matriz de coeficientes.

Estos resultados requieren que no se intercambien columnas. Un resultado más abstracto está en la Sección 9.5.1.

**Los costes.** La eliminación por columnas también tiene costes. La matriz con la que se opera es mayor, porque incluye una matriz identidad: para resolver un sistema de $m$ ecuaciones con $n$ incógnitas tiene $m+n+1$ filas en lugar de $m$. Los demás textos y la mayoría de los programas aplican la eliminación a las filas. Y se pierde una lectura propia de la eliminación por filas: cada paso produce un sistema de ecuaciones equivalente al anterior. Creo que disponer de un mismo procedimiento y de una misma explicación para los tres problemas compensa estos costes.

## Tres métodos de eliminación

El libro distingue tres métodos de eliminación. Cada uno añade pasos al anterior.

1. La *eliminación de izquierda a derecha* pre-escalona la matriz. Es el paso fundamental y basta para la mayoría de los problemas del curso.
2. La *eliminación gaussiana* escalona la matriz. Es la que emplea el libro para calcular un determinante (Sección 15.2).
3. La *eliminación Gauss-Jordan* escalona y reduce la matriz. Es necesaria para invertir una matriz (Sección 5.1.4).

Solo la diagonalización requiere combinar la eliminación por filas y la eliminación por columnas. La diagonalización por semejanza aplica la eliminación gaussiana a las columnas y las transformaciones inversas a las filas (Sección 17.2). La diagonalización por congruencia de una matriz simétrica escalona por filas y por columnas (Sección 19.2).

Esta distinción indica al estudiante qué pasos son necesarios en cada problema y por qué. Los manuales no distinguen el pre-escalonado, porque su eliminación no se detiene hasta obtener una matriz escalonada. Por eso no estudian qué se puede obtener de una matriz que solo está pre-escalonada.

## Las transformaciones elementales como base de los algoritmos

El curso trata con detalle las transformaciones elementales, y ese detalle permite distinguir los tres métodos de eliminación. Hay dos decisiones sobre las transformaciones elementales que conviene señalar.

La primera es que el libro solo define dos tipos de transformaciones elementales: una suma a un vector un múltiplo de otro y la otra multiplica un vector por un número distinto de cero. El intercambio de dos vectores no se define como una transformación elemental, porque se obtiene con una secuencia de transformaciones elementales de esos dos tipos.

La segunda es que el libro distingue los intercambios de las permutaciones y emplea una escritura distinta para cada cosa. Una permutación es una secuencia de intercambios. La distinción aporta precisión. Por ejemplo, un intercambio de dos columnas cambia el signo del determinante, y de una permutación no se sabe si lo cambia sin contar sus intercambios.

El estudio de la relación entre las transformaciones de las filas y las transformaciones de las columnas permite exponer dos algoritmos que no suelen aparecer en los manuales. Los describo a continuación.

**La diagonalización por bloques triangulares.** Toda matriz cuadrada es diagonalizable por bloques triangulares cuando se trabaja con números complejos (Teorema 17.2.1). La demostración describe un algoritmo que emplea transformaciones elementales. El algoritmo requiere conocer los autovalores de la matriz,

como los demás métodos de diagonalización por semejanza. Se aplica a cualquier matriz cuadrada. Cuando la matriz es diagonalizable, el algoritmo la diagonaliza.

Del algoritmo se deducen varios resultados conocidos: la suma de los autovalores es igual a la traza, el producto de los autovalores es igual al determinante, y una matriz es diagonalizable si y solo si la multiplicidad geométrica y la multiplicidad algebraica de cada autovalor coinciden. El cálculo con lápiz y papel es largo, pero el algoritmo está implementado en el módulo NAcAL, como los demás algoritmos del curso.

**La diagonalización por congruencia de matrices simétricas.** Este algoritmo se expone junto con la Ley de inercia de Sylvester en la Sección 19.2. Sirve para clasificar formas cuadráticas.

Los manuales suelen clasificar las formas cuadráticas de dos maneras. La primera emplea los signos de los autovalores. Para obtenerlos hay que calcular el polinomio característico y después hallar sus raíces o aplicar un criterio sobre los signos de sus coeficientes. La segunda manera es el criterio de Sylvester. Con los menores principales superiores el criterio decide si una matriz es definida, pero para decidir si es semidefinida hay que calcular todos los menores principales.

La diagonalización por congruencia obtiene los signos con la eliminación. No requiere calcular el polinomio característico ni ningún determinante, y se aplica del mismo modo a matrices simétricas de cualquier orden.

## Ejemplos de uso de NAcAL

El módulo NAcAL trabaja con vectores de $\mathbb{R}^n$, con matrices, con transformaciones elementales, con subespacios de $\mathbb{R}^n$ y con subespacios afines de $\mathbb{R}^n$. También resuelve sistemas de ecuaciones cuyos vectores son matrices.

El módulo implementa el álgebra matricial y los tres métodos de eliminación, por filas y por columnas. Con ellos invierte matrices, resuelve sistemas de ecuaciones y calcula determinantes.[24] También diagonaliza por bloques triangulares, diagonaliza ortogonalmente matrices simétricas y diagonaliza matrices simétricas por congruencia.

El módulo opera con subespacios y con subespacios afines de $\mathbb{R}^n$. Obtiene sus ecuaciones paramétricas y cartesianas, su intersección y su complemento ortogonal. También obtiene la suma de dos subespacios. Comprueba la inclusión y la pertenencia de un vector.

Además, el módulo representa los objetos con LaTeX, muestra el desarrollo de muchos cálculos y trabaja con símbolos. Siguen cuatro ejemplos.

*Ejemplo 1.* Cálculo de la inversa de $\mathbf{B} = \begin{bmatrix} 2 & \sqrt{3} \\ 0 & \frac{1}{\pi} \end{bmatrix}$ por eliminación Gauss-Jordan:

Módulo NAcAL para Python

```
B = Matrix([[2,sympy.sqrt(3)],[0,fracc(1,sympy.pi)]])
InvMat(B,1)
```

$$\left[\begin{array}{cc|cc} 2 & \sqrt{3} & 1 & 0 \\ 0 & \frac{1}{\pi} & 0 & 1 \end{array}\right] \xrightarrow{\left[\left(-\frac{\sqrt{3}}{2}\right)\mathbf{1}+\mathbf{2}\right]} \left[\begin{array}{cc|cc} 2 & 0 & 1 & -\frac{\sqrt{3}}{2} \\ 0 & \frac{1}{\pi} & 0 & 1 \end{array}\right] \xrightarrow[[(\pi)\mathbf{2}]]{\left[\left(\frac{1}{2}\right)\mathbf{1}\right]} \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & -\frac{\sqrt{3}\pi}{2} \\ 0 & 1 & 0 & \pi \end{array}\right]$$

*Por tanto* $\mathbf{B}^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}\pi}{2} \\ 0 & \pi \end{bmatrix}$.

[24] También está implementado el cálculo recursivo mediante la expansión de Laplace.

*Ejemplo 2.* Resolución de un sistema de ecuaciones lineales:

Módulo NAcAL para Python

```
A = Matrix( [ [2,-2,0], [-1,3,4*sympy.sqrt(7)] ] );
b = Vector( [2*sympy.pi, sympy.exp(1)] )
SEL( A, b, 1 )     # resuelve el Sistema de Ecuaciones Lineales
```

$$\left[\begin{array}{ccc|c} 2 & -2 & 0 & -2\pi \\ -1 & 3 & 4\sqrt{7} & -e \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[[(\pi)\mathbf{1}+\mathbf{4}]]{\overset{\tau}{[(1)\mathbf{1}+\mathbf{2}]}} \left[\begin{array}{ccc|c} 2 & 0 & 0 & 0 \\ -1 & 2 & 4\sqrt{7} & -\pi-e \\ \hline 1 & 1 & 0 & \pi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\left[\left(\frac{e+\pi}{2}\right)\mathbf{2}+\mathbf{4}\right]]{\overset{\tau}{[(-2\sqrt{7})\mathbf{2}+\mathbf{3}]}} \left[\begin{array}{ccc|c} 2 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ \hline 1 & 1 & -2\sqrt{7} & \frac{e+3\pi}{2} \\ 0 & 1 & -2\sqrt{7} & \frac{e+\pi}{2} \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

$$\left\{ \boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^1,\ \boldsymbol{v} = \begin{pmatrix} \frac{e+3\pi}{2} \\ \frac{e+\pi}{2} \\ 0 \end{pmatrix} + \begin{bmatrix} -2\sqrt{7} \\ -2\sqrt{7} \\ 1 \end{bmatrix} \boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \begin{bmatrix} -1 & 1 & 0 \\ \frac{\sqrt{7}}{14} & 0 & 1 \end{bmatrix} \boldsymbol{v} = \begin{pmatrix} -\pi \\ \frac{\sqrt{7}(e+3\pi)}{28} \end{pmatrix} \right\}$$

*El conjunto solución es un espacio afín. Se representa con unas ecuaciones paramétricas a la izquierda y con unas cartesianas a la derecha.*

*Ejemplo 3.* Diagonalización de una matriz simétrica por congruencia:

Módulo NAcAL para Python

```
f = sympy.symbols(r'f(\lambda)')
A = Matrix([[0,f],[f, sympy.exp(-2)]])
DiagonalizaC(A,1)
```

$$\begin{bmatrix} 0 & f(\lambda) \\ f(\lambda) & e^{-2} \end{bmatrix} \xrightarrow[\overset{\tau}{[\mathbf{1} \rightleftharpoons \mathbf{2}]}]{[\mathbf{1} \rightleftharpoons \mathbf{2}]} \begin{bmatrix} e^{-2} & f(\lambda) \\ f(\lambda) & 0 \end{bmatrix} \xrightarrow[\overset{\tau}{[(-f(\lambda)e^2)\mathbf{1}+\mathbf{2}]}]{[(-f(\lambda)e^2)\mathbf{1}+\mathbf{2}]} \begin{bmatrix} e^{-2} & 0 \\ 0 & -f(\lambda)^2 e^2 \end{bmatrix}$$

*Las transformaciones representadas por encima de las flechas son transformaciones de las columnas. Las que aparecen por debajo son transformaciones de las filas.*

*Ejemplo 4.* Resolución de un sistema de ecuaciones cuyos vectores son matrices. Se buscan las combinaciones lineales de $\mathbf{I}$ y $\mathbf{A}$ que son iguales a $\mathbf{A}^2$:

Módulo NAcAL para Python

```
a, b, c, d = sympy.symbols('a b c d')
A = Matrix( [ [a,b], [c,d] ] )
M = Sistema( [ A**0, A ] )     # lista con las matrices I y A
SEL( M, A**2, 1 )
```

$$\left[\begin{array}{cc|cc|c} \multicolumn{2}{c|}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} & \multicolumn{2}{c|}{\begin{bmatrix} a & b \\ c & d \end{bmatrix}} & \begin{bmatrix} -a^2-bc & -b(a+d) \\ -c(a+d) & -bc-d^2 \end{bmatrix} \\ \hline 1 & 0 & \multicolumn{2}{c|}{} & 0 \\ 0 & 1 & \multicolumn{2}{c|}{} & 0 \\ \hline 0 & 0 & \multicolumn{2}{c|}{} & 1 \end{array}\right] \xrightarrow[[(a^2+bc)\mathbf{1}+\mathbf{3}]]{\overset{\tau}{[(-a)\mathbf{1}+\mathbf{2}]}} \left[\begin{array}{cc|cc|c} \multicolumn{2}{c|}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} & \multicolumn{2}{c|}{\begin{bmatrix} 0 & b \\ c & -a+d \end{bmatrix}} & \begin{bmatrix} 0 & -b(a+d) \\ -c(a+d) & (a-d)(a+d) \end{bmatrix} \\ \hline 1 & -a & \multicolumn{2}{c|}{} & a^2+bc \\ 0 & 1 & \multicolumn{2}{c|}{} & 0 \\ \hline 0 & 0 & \multicolumn{2}{c|}{} & 1 \end{array}\right]$$

$$\xrightarrow{[(a+d)\mathbf{2}+\mathbf{3}]} \left[\begin{array}{cc|cc|c} \multicolumn{2}{c|}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} & \multicolumn{2}{c|}{\begin{bmatrix} 0 & b \\ c & -a+d \end{bmatrix}} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \\ \hline 1 & -a & \multicolumn{2}{c|}{} & -ad+bc \\ 0 & 1 & \multicolumn{2}{c|}{} & a+d \\ \hline 0 & 0 & \multicolumn{2}{c|}{} & 1 \end{array}\right]$$

*La solución $(bc-ad,\ a+d,)$ indica que $\mathbf{A}^2 = (bc-ad)\mathbf{I} + (a+d)\mathbf{A}$ para cualquier matriz $\mathbf{A}$ de orden 2. El procedimiento es el del Ejemplo 2. Solo cambia que las columnas son matrices.*

Estos ejemplos muestran que NAcAL no solo realiza los cálculos. Además los muestra paso a paso. Trabaja

con números y con símbolos, y emplea la notación de este libro.

Para poder visualizar la representación en LaTeX se debe trabajar con NAcAL en un Notebook de Jupyter, o usar el paquete PythonTeX y llamar a NAcAL desde el propio código LaTeX del documento.

## Para terminar

La notación y la eliminación por columnas son elecciones independientes. Con esta notación se puede exponer la eliminación por filas con el mismo rigor, y la eliminación por columnas se puede exponer con la notación habitual. El profesor que solo encuentre convincente una de las dos puede adoptarla sin la otra.

De las dos, la notación es la que afecta a todo el libro. Sus expresiones se manejan como las de un lenguaje de programación: el tipo de los operandos determina el resultado de cada operador y cada paso aplica una regla escrita. El módulo NAcAL lo muestra, porque su código transcribe las definiciones del libro y sus algoritmos siguen las demostraciones.

He indicado los costes de ambas elecciones. Polya (1957, p. 140) pide que la notación se elija con cuidado y que haya alguna buena razón para la elección.[25] Esta introducción expone mis razones. La valoración corresponde al lector, y las lecciones que siguen le permiten hacerla.

[25] El original dice: *"we should choose our notation carefully, and have some good reason for our choice."*

# Parte I

# Álgebra Matricial

# Vectores y matrices (notación y operaciones)

En matemáticas, es crucial comprender la definición de cada objeto. Empecemos con uno que usaremos mucho:

**Definición 1.1.** *Un* sistema[1] *es una lista* ***ordenada*** *de objetos.*[2]

Dos sistemas son distintos si las listas son diferentes; así, los siguientes sistemas de números son diferentes:

$$[1;\ 2;\ 1;] \neq [1;\ 2;] \neq [2;\ 1;] \neq [2;\ 1;\ 2;] \neq [2;\ 1;\ 1;] \neq [2;\ 1;\ 1;\ 1;\ 1;\ 2;].$$

Un sistema no es un conjunto (o colección) de objetos. En un conjunto no hay orden en la disposición de sus elementos y, de hecho, si todo elemento del conjunto $A$ está en el conjunto $B$, y todo elemento de $B$ está en el conjunto $A$, se dice que los conjuntos $A$ y $B$ son iguales; por tanto, los siguientes conjuntos son iguales:

$$\{1,\ 2,\ 1\} = \{1,\ 2\} = \{2,\ 1\} = \{2,\ 1,\ 2\} = \{2,\ 1,\ 1\} = \{2,\ 1,\ 1,\ 1,\ 1,\ 2\}.$$

Distinguiremos conjuntos y sistemas del siguiente modo: por una parte, encerramos los elementos de un conjunto entre llaves y separados por comas. Por otra parte, encerramos los elementos de un *sistema* entre corchetes y pondremos un "punto y coma" después de cada elemento. Por ejemplo: $[a;\ B;\ \mathbf{C}; \ldots;]$.

En este curso trataremos con dos importantes tipos de sistemas: los vectores de $\mathbb{R}^n$ (listas de números) y las matrices (listas de vectores de $\mathbb{R}^n$).

## 1.1. Vectores de $\mathbb{R}^n$

Denotamos con $\mathbb{R}$ al *conjunto de números reales*. Ejemplos de números reales son: $7$, $\frac{-3}{11}$, $\sqrt{2}$ o $-\pi$. Con frecuencia, también llamaremos *escalares* a los números reales.

**Definición 1.2.** *Llamamos* vector *de* $\mathbb{R}^n$ *a un sistema (o lista ordenada) de* $n$ *números reales.*

El símbolo "$\mathbb{R}$" nos indica que el sistema está compuesto por números *reales*, y el superíndice "$n$" nos indica la cantidad de números que hay en la lista.

Recuerde que los sistemas están ordenados. Consecuentemente, podemos identificar los elementos de la lista por la posición que ocupan: hay un primer elemento, un segundo elemento, etc. La forma de escribir la lista no es importante siempre que quede claro quién es el primer elemento, quién el segundo, etc. Por eso podemos escribir un *mismo* vector tanto en horizontal como en vertical:[3]

$$(\pi,\ -1,\ 0,) \qquad \text{o} \qquad \begin{pmatrix} \pi \\ -1 \\ 0 \end{pmatrix}.$$

Ambas son representaciones del vector de $\mathbb{R}^3$ cuya primera componente es $\pi$, la segunda $-1$ y la tercera $0$.

[1] Una definición más formal aparecerá en la página 98

[2] Con "objeto" siempre nos referiremos a *objetos matemáticos*: números, funciones, ecuaciones, variables, etc.

[3] Otros libros distinguen entre vectores *fila* y vectores *columna*; pero aquí, dada la Definición 1.2, no cabe tal distinción.

**Notación.** Los vectores de $\mathbb{R}^n$ tienen una notación específica, y la hemos usado en el ejemplo anterior:[4] $(a,\ b,\ c, \ldots,)$. Con ella escribimos la correspondiente lista de números entre ***paréntesis*** (poniendo *una coma detrás de cada elemento* cuando escribimos el vector en horizontal).[5]

Denotamos los vectores de $\mathbb{R}^n$ con letras minúsculas en negrita cursiva: $\boldsymbol{a}$, $\boldsymbol{b}$,..., $\boldsymbol{y}$, $\boldsymbol{z}$. Y mencionamos sus componentes con idéntica letra pero en fuente cursiva y con un subíndice que indica la posición de la componente dentro del vector. Por ejemplo:

$$\boldsymbol{a} = \big(a_1, \quad a_2, \quad a_3, \quad a_4, \quad a_5, \quad a_6,\big)\,;$$

es un vector de $\mathbb{R}^6$ cuya tercera componente es $a_3$. Para indicar que la segunda componente de $\boldsymbol{b} = \big(0, \quad 4, \quad 6,\big)$ es un cuatro, escribiremos $b_2 = 4$.

Decimos que dos vectores $\boldsymbol{x}$ e $\boldsymbol{y}$ de $\mathbb{R}^n$ son iguales si lo son sus correspondientes listas ordenadas de componentes, es decir, si y solo si: $x_i = y_i$ para $i = 1 : n$ (i.e. si son iguales componente a componente).[6] Por ejemplo, con los números 1, 3 y 5, podemos formar dos vectores *distintos* del siguiente modo:

$$\boldsymbol{x} = \big(1, \quad 3, \quad 5,\big) \quad \text{y} \quad \boldsymbol{z} = \big(3, \quad 5, \quad 1,\big)\,.$$

Recuerde que dos vectores son iguales solo si lo son sus respectivas listas. Por eso $\boldsymbol{x}$ y $\boldsymbol{z}$ son distintos ($\boldsymbol{x} \neq \boldsymbol{z}$).

### El operador selector de componentes de un sistema

También denotaremos la $i$-ésima componente de un vector $\boldsymbol{a}$ con un subíndice en el que aparecen tanto el símbolo " | " como el correspondiente índice de la componente: $\boldsymbol{a}_{|i}$. Es decir, el operador selector " | " *actúa* sobre $\boldsymbol{a}$ para seleccionar su $i$-ésima componente. Consecuentemente, las siguientes formas de denotar la tercera componente de $\boldsymbol{b}$ son equivalentes:

$$\boldsymbol{b}_{|3} \ \equiv \ b_3.$$

El selector también puede actuar por la izquierda del vector, con el mismo resultado: ${}_{i|}\boldsymbol{a} = \boldsymbol{a}_{|i}$. Lo usaremos a partir de la Sección 1.4.2.

### Uso de NAcAL

☞ Durante el curso utilizaremos el módulo NAcAL para Python. Así podremos experimentar en el ordenador con los objetos y procedimientos que vayamos viendo a lo largo de este curso de Álgebra Lineal. Dicho módulo es una implementación literal de los objetos y procedimientos que veremos en el libro.

Para utilizar el módulo en Python primero hay que importarlo

Módulo NAcAL para Python

```
from nacal import *  # importamos el módulo en memoria
```

Una vez importado, con las dos líneas siguientes definimos el vector `b` y lo mostramos en un terminal:

terminal de Python

```
>>> from nacal import *
>>> b = Vector( [0, 4, 6] ) # Vector b cuya lista de componentes es 0, 4 y 6
>>> b
Vector([0, 4, 6])
```

[4] Nótese que con la notación genérica de sistema, también podemos escribir dicho vector con $[\pi;\ -1;\ 0;]$, es decir, con una lista entre corchetes con sus elementos seguidos por puntos y comas (como en los ejemplos dados tras la Definición 1.1). Dicho de otra manera, aunque $[\pi;\ -1;\ 0;] = (\pi,\ -1,\ 0,)$, siempre usaremos la segunda expresión para denotar un vector.

[5] Así $(2,)$ es un *vector* de $\mathbb{R}^1$ (es decir, una lista con un único elemento); y $(2)$ es un *número* entre paréntesis.

[6] Con "$p : q$" denotamos la secuencia $p, p+1, \ldots, q$, es decir, la lista ordenada de los números naturales que van de $p$ a $q$: $[k \in \mathbb{N} \mid p \leq k \leq q]$. Así, con $i = 1 : n$ indicamos que el índice $i$ recorre los números naturales de 1 a $n$, es decir $1, 2, \ldots, n$. Cuando el primer número, $p$, sea mayor que el segundo, $q$, entenderemos que "$p : q$" es una lista vacía: [ ].

donde el resultado aparece en gris en el anterior recuadro de `terminal de Python`; y si ejecutamos el código en un Notebook de Jupyter obtenemos:

$$\begin{pmatrix}0\\4\\6\end{pmatrix}$$

Para obtener la tercera componente de $\boldsymbol{b}$ tecleamos:

Módulo NAcAL para Python

```
b|3
```

y tanto en un terminal de Python como en un Notebook de Jupyter obtendremos el número 6.

☞ Hemos definido un vector de $\mathbb{R}^n$ como un sistema de $n$ números reales. Al encerrar una lista de números entre paréntesis, denotamos un vector (cuyas componentes son los números encerrados en el orden en el que aparecen en la lista dentro del paréntesis). Por ejemplo

$$\boldsymbol{a} = (1,\ 4,\ 9,\ 2,).$$

Para seleccionar una componente de la lista, usamos el operador "|" y el índice de la componente que queremos seleccionar. Por ejemplo, la componente $a_3$ (la tercera componente de $\boldsymbol{a}$) es:

$$\boldsymbol{a}_{|3} = 9.$$

### 1.1.1. Definiciones de algunos vectores especiales

**Definición 1.3.** *Llamamos* vector nulo *(o vector cero) de* $\mathbb{R}^n$ *al vector cuyas n componentes son cero, y lo denotamos con* $\mathbf{0}$. *Por ejemplo, los vectores nulos de* $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$ *y* $\mathbb{R}^7$ *son respectivamente:*

$$(0,)\,,\quad \begin{pmatrix}0\\0\end{pmatrix},\quad \begin{pmatrix}0\\0\\0\end{pmatrix}\quad y\quad (0,\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ \ 0,\ \ 0,)\,.$$

¡Nótese que estos cuatro vectores nulos **no son iguales**, pues no tienen el mismo número de componentes (y por tanto, sus correspondientes listas ordenadas no son iguales)!

Módulo NAcAL para Python

```
V0( 3 )       # vector nulo de tres componentes
```

$$\begin{pmatrix}0\\0\\0\end{pmatrix}$$

**Definición 1.4.** *Sea el vector* $\boldsymbol{a}$. *Llamamos* vector opuesto *de* $\boldsymbol{a}$ *al vector* $-\boldsymbol{a}$, *que es el vector cuyas componentes son las opuestas de las de* $\boldsymbol{a}$; *es decir,* $(-\boldsymbol{a})_{|i} = -(\boldsymbol{a}_{|i})$. *Por ejemplo:*

$$\text{si } \boldsymbol{a} = \begin{pmatrix}a_1\\\vdots\\a_n\end{pmatrix} \text{ entonces } -\boldsymbol{a} = \begin{pmatrix}-a_1\\\vdots\\-a_n\end{pmatrix};\ \text{y si } \boldsymbol{c} = (1,\ -3\pi,\ \sqrt{2},) \text{ entonces } -\boldsymbol{c} = (-1,\ 3\pi,\ -\sqrt{2},).$$

Módulo NAcAL para Python

```
-b              # vector opuesto de b
```

$$\begin{pmatrix} 0 \\ -4 \\ -6 \end{pmatrix}$$

- Enlace al Notebook de Jupyter para trabajar con lo expuesto en esta Sección 1.1.

## 1.2. Sumas de vectores y productos de vectores por escalares

Las dos operaciones fundamentales con vectores son la suma y el producto de un vector por un escalar.

**Definición 1.5** (suma de vectores)**.** *La suma, $\boldsymbol{a}+\boldsymbol{b}$, de dos vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ de $\mathbb{R}^n$ es el vector de $\mathbb{R}^n$ cuya componente $i$-ésima es la suma de las componentes $\boldsymbol{a}_{|i}$ y $\boldsymbol{b}_{|i}$. Es decir, sumamos componente a componente:*

$$\boxed{(\boldsymbol{a}+\boldsymbol{b})_{|i} = \boldsymbol{a}_{|i} + \boldsymbol{b}_{|i}} \qquad \text{para } \ i = 1:n.$$

Por tanto:

$$\boldsymbol{a}+\boldsymbol{b} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1+b_1 \\ \vdots \\ a_n+b_n \end{pmatrix}.$$

*Observación.* Nótese que el operador " $|i$ " es *distributivo*[7] para la suma: $(\boldsymbol{a}+\boldsymbol{b})_{|i} = \boldsymbol{a}_{|i} + \boldsymbol{b}_{|i}$.

**Definición 1.6** (producto de un vector por un escalar)**.** *El producto, $\lambda\boldsymbol{b}$, de un vector $\boldsymbol{b}$ de $\mathbb{R}^n$ por un escalar $\lambda$ es el vector de $\mathbb{R}^n$ cuya componente $i$-ésima es el producto de $\lambda$ por $\boldsymbol{b}_{|i}$. Es decir, multiplicamos cada componente de $\boldsymbol{b}$ por $\lambda$:*

$$\boxed{(\lambda\boldsymbol{b})_{|i} = \lambda(\boldsymbol{b}_{|i})} \qquad \text{para } \ i = 1:n.$$

Por tanto

$$\lambda\boldsymbol{b} = \lambda \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} \lambda b_1 \\ \vdots \\ \lambda b_n \end{pmatrix}.$$

Al vector $\lambda\boldsymbol{b}$ lo denominamos *múltiplo* de $\boldsymbol{b}$. Al multiplicar cualquier vector por cero se obtiene el vector nulo: $0\boldsymbol{b} = \boldsymbol{0}$.

La operación de multiplicar un vector por un número también se denomina "escalar el vector", es decir, cambiarlo de escala. Por ejemplo, si tiene un vector con datos de salarios en miles de euros, expresar esos datos en euros es escalar los datos, y se logra multiplicando por 1000.

*Observación.* El operador "|" es *asociativo*[8] para el producto por escalares: $(\lambda\boldsymbol{b})_{|i} = \lambda(\boldsymbol{b}_{|i})$. Fíjese cómo la región encerrada por el paréntesis *se desplaza* en la expresión anterior... en un lado de la igualdad "$\lambda$" está dentro del paréntesis y "$|i$ " fuera, ocurriendo lo contrario al otro lado de la igualdad. El orden en que se realizan las operaciones resulta ser irrelevante cuando podemos "desplazar" los paréntesis manteniendo la igualdad; por tanto, podemos omitir los paréntesis.[9] Consecuentemente podemos escribir $\lambda\boldsymbol{b}_{|i}$.

**Definición 1.7** (Operador lineal)**.** *Un operador*[10] *distributivo respecto a la suma y asociativo respecto al producto por escalares se dice que es lineal.*

[7] De manera similar a la propiedad distributiva del producto de números respecto de la suma: $a(b+c) = ab+ac$.
[8] De manera similar a la propiedad asociativa del producto entre escalares: $(ab)c = a(bc)$.
[9] De manera similar a cuando escribimos el producto de tres escalares: $abc$.
[10] Un operador es una regla que transforma un objeto en otro. El selector transforma un vector en un número. Más adelante veremos operadores que transforman una matriz en otra matriz.

Así pues, el operador "$|i$" (que selecciona la componente $i$-ésima de un vector) es un operador lineal.

☞ La definición de las operaciones de suma de vectores y producto de un vector por un escalar convierten al operador " $|i$ " en un **operador lineal**; es decir, *distributivo* respecto a la suma:

$$(\boldsymbol{a}+\boldsymbol{b})_{|i} = \boldsymbol{a}_{|i} + \boldsymbol{b}_{|i}$$

y *asociativo* respecto al producto por *escalares*:

$$\big(\lambda \boldsymbol{b}\big)_{|i} = \lambda\big(\boldsymbol{b}_{|i}\big).$$

**La omisión de los paréntesis y del símbolo del producto en Python** En un texto, la expresión $2\boldsymbol{b}$ significa "multiplicar por 2 el vector $\boldsymbol{b}$", es decir

$$2\boldsymbol{b} \;\equiv\; 2\cdot\boldsymbol{b}$$

donde el punto denota el producto. Los humanos estamos acostumbrados a *omitir* el "punto" de la operación producto (así logramos expresiones más compactas). No obstante, debe recordar que Python no funciona como un humano. El intérprete de Python requiere la presencia *explícita* del símbolo de la operación producto (el asterisco "`*`"):

Módulo NAcAL para Python

```
2*b          # multiplica el Vector b por 2  (también podemos escribir b*2 )
```

$$\begin{pmatrix} 0 \\ 8 \\ 12 \end{pmatrix}$$

El intérprete de Python tiene establecidas una serie de precedencias entre operadores. En el siguiente cuadro de operadores, los de arriba preceden a (*se ejecutan antes que*) los de abajo:

**Precedencia en las operaciones**

| Símbolo en Python | Operación asociada en NAcAL |
|---|---|
| `()`, `[]`, `{}` | |
| `**` | exponenciación |
| `~` | transposición |
| `*`, `/` | multiplicación, división |
| `+`, `-` | suma, resta |
| `&` | transformación elemental |
| `\|` | selector |

Cuadro 1.1: Precedencia de operadores (detalles aquí: 6.16 *Orden de evaluación* y 6.17 *Prioridad de operador*)

Fíjese en que Python ejecuta antes la operación *producto* (`*`) que la selección de una componente (`|`). Por tanto, en Python podemos producir errores si no escribimos con suficiente cuidado las expresiones (hay que recordar que las expresiones se ejecutan de izquierda a derecha, pero con unas reglas de precedencia estrictas).

Dado que la precedencia no es fácil de recordar, para lograr mayor claridad en el código (*y asegurarnos de que realizamos las operaciones en el orden deseado*) **es recomendable usar los paréntesis**: lo que haya dentro de los paréntesis será lo primero en ser evaluado.

EJERCICIO 1. ¿Cómo evalúa el módulo NAcAL las siguientes expresiones: `(3*b)|2`; `3*(b|2)` y `3*b|2` ?

EJERCICIO 2. ¿Cuál es la primera operación realizada en cada una de estas expresiones? `b+b*2` y `b+b|2`.

EJERCICIO 3. ¿Qué problema puede tener la expresión `b+b|2` para nosotros?

Moraleja: lo más conveniente es poner paréntesis tanto al escribir con lápiz y papel como al escribir en Python. Con ello evitamos ambigüedades (en el primer caso) y sorpresas (en el segundo).

### 1.2.1. Propiedades de la suma y del producto por escalares

☞ Explotando las propiedades de la notación (junto con algunas propiedades de los *números reales*) vamos a demostrar ocho propiedades que verifican la *suma de vectores* y el *producto de un vector por un escalar* (Proposición 1.2.1).

Para empezar, recordemos algunas propiedades (¡no todas![11]) de los números reales que seguro usted conoce y que ha manejado desde la escuela...

**Recordatorio de algunas propiedades del conjunto de numeros reales**

Sean $a$, $b$ y $c$ números reales. Entonces:

1. $a+b=b+a$ (conmutativa de la suma).
2. $a+(b+c)=(a+b)+c$ (asociativa de la suma).
3. $a+0=a$ (elemento neutro de la suma).
4. $a+(-a)=0$ (opuesto).
5. $ab=ba$ (conmutativa del producto).
6. $a(b+c)=ab+ac$ (distributiva).
7. $a(bc)=(ab)c$ (asociativa del producto).
8. $1a=a$ (elemento neutro del producto).

La cuestión que queremos abordar ahora es: ¿también se verificarán estas propiedades cuando operamos con vectores? En concreto, hagámonos las siguientes ocho preguntas (puede investigar con el Notebook de esta sección):

1. ¿Es verdad que
$$\begin{pmatrix}1\\2\\3\end{pmatrix}+\begin{pmatrix}4\\5\\6\end{pmatrix}=\begin{pmatrix}4\\5\\6\end{pmatrix}+\begin{pmatrix}1\\2\\3\end{pmatrix}?$$
¿Es cierto solo para los vectores del ejemplo? O ¿es una regla general para cualquier par de vectores?

2. ¿Es verdad que
$$\begin{pmatrix}1\\2\end{pmatrix}+\left(\begin{pmatrix}4\\5\end{pmatrix}+\begin{pmatrix}1\\0\end{pmatrix}\right)=\left(\begin{pmatrix}1\\2\end{pmatrix}+\begin{pmatrix}4\\5\end{pmatrix}\right)+\begin{pmatrix}1\\0\end{pmatrix}?$$
Los dos grandes paréntesis indican que primero hay que sumar los vectores que contienen. Luego, al vector resultante se le debe sumar el vector que hay fuera del gran paréntesis.

¿Es cierto solo para los vectores del ejemplo? O ¿es una regla general para cualquier trío de vectores?

3. ¿Es verdad que
$$\begin{pmatrix}4\\5\\6\end{pmatrix}+\begin{pmatrix}0\\0\\0\end{pmatrix}=\begin{pmatrix}4\\5\\6\end{pmatrix}?$$
¿Es cierto solo para ese vector y el vector nulo? O ¿es una regla general para cualquier vector más el vector nulo?

[11] Los números reales tienen más propiedades. Por ejemplo, todo número real distinto de cero tiene inverso, y los números reales están ordenados. La lista contiene solo las propiedades que usaremos una y otra vez. Cuando usemos alguna otra, lo avisaremos.

4. ¿Es verdad que
$$\begin{pmatrix}4\\5\\6\end{pmatrix}+\begin{pmatrix}-4\\-5\\-6\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}?$$
¿Es cierto solo para el vector del ejemplo? O ¿es una regla general para cualquier vector y su opuesto?

5. ¿Es verdad que
$$2\left(\begin{pmatrix}1\\-1\end{pmatrix}+\begin{pmatrix}4\\5\end{pmatrix}\right)=2\begin{pmatrix}1\\-1\end{pmatrix}+2\begin{pmatrix}4\\5\end{pmatrix}?$$
(recuerde que primero se realizan las operaciones dentro del gran paréntesis) ¿Es una regla general para cualquier par de vectores? ¿Y si cambiamos el escalar 2 por otro número?

6. ¿Es verdad que
$$(2+3)\begin{pmatrix}1\\2\\3\end{pmatrix}=2\begin{pmatrix}1\\2\\3\end{pmatrix}+3\begin{pmatrix}1\\2\\3\end{pmatrix}?$$
(recuerde que primero se realizan las operaciones dentro del paréntesis) ¿Es una regla general para cualquier vector? ¿Y si cambiamos los escalares 2 y 3 por otros números?

7. ¿Es verdad que
$$2\left(3\begin{pmatrix}1\\2\end{pmatrix}\right)=(2\cdot 3)\begin{pmatrix}1\\2\end{pmatrix}?$$
(recuerde, primero las operaciones dentro de los paréntesis). Hágase las mismas preguntas de arriba.

8. ¿Es verdad que si $a=1$
$$a\begin{pmatrix}1\\2\\3\end{pmatrix}=\begin{pmatrix}1\\2\\3\end{pmatrix}?$$
¿Es una regla general para cualquier vector? ¿Y si cambiamos el escalar 1 por otro número?

En el Ejercicio 4 se le pide que demuestre las ocho propiedades que acabamos de revisar[12]. Su solución aparece al final del libro ¡pero usted debe resolverlo por su cuenta y sin mirar! Si "cotillea" previamente la solución, el ejercicio le servirá de muy poco.

Para resolver el ejercicio únicamente necesita aplicar las reglas de notación vistas más arriba (definiciones de *suma* y *producto por escalares* de la Página 6); y cuando tenga que operar con las componentes de los vectores (que son números reales) aplique la propiedad de las operaciones entre números reales que necesite (las tiene en el recordatorio de la Página 8).

Ejercicio 4. Demuestre las propiedades de la siguiente proposición:

**Proposición 1.2.1** (Propiedades de las operaciones entre vectores)**.** *Para cualesquiera vectores* $\boldsymbol{a}$*,* $\boldsymbol{b}$ *y* $\boldsymbol{c}$ *con el mismo número de componentes, y para cualesquiera escalares (números)* $\lambda$ *y* $\varphi$*, se verifica que:*

[12]Se le propone como ejercicio una ¡demostración! Las demostraciones son imprescindibles en matemáticas. Es IMPOSIBLE que adquiera una buena formación matemática si no realiza las demostraciones. Solo se adquiere una verdadera comprensión al realizar las demostraciones. Las que se le propondrán en el curso son muy sencillas (como las de este primer ejercicio) o son "casi" una repetición de otras ya vistas, o las pistas que se le dan describen los pasos necesarios para dar con la demostración.
Aunque las demostraciones sean sencillas, es posible que le cueste escribirlas al principio (quizá sea la primera vez que intenta demostrar algo... y la primera vez suele resultar difícil). No se desanime; *es importante lograr demostrar proposiciones sencillas* (si trabaja el curso adecuadamente, cuando termine habrá adquirido una formación fundamental para un buen economista).

1. $\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$.
2. $\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})=(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}$.
3. $\boldsymbol{a}+\boldsymbol{0}=\boldsymbol{a}$.
4. $\boldsymbol{a}+(-\boldsymbol{a})=\boldsymbol{0}$.
5. $\lambda(\boldsymbol{a}+\boldsymbol{b})=\lambda\boldsymbol{a}+\lambda\boldsymbol{b}$.
6. $(\lambda+\varphi)\boldsymbol{a}=\lambda\boldsymbol{a}+\varphi\boldsymbol{a}$.
7. $\lambda(\varphi\boldsymbol{a})=(\lambda\varphi)\boldsymbol{a}$.
8. $1\boldsymbol{a}=\boldsymbol{a}$.

*Pista.* Dispone de dos estrategias. La primera se asemeja a la forma de operar mostrada en las ocho preguntas de más arriba (por ello es posible que le parezca más sencilla). La segunda estrategia explota las propiedades de *linealidad* del operador "|". Es importante que se familiarice con este segundo modo de trabajar: es más abstracto pero mucho más rápido y versátil (pues la demostración solo se centra en aplicar las propiedades a la componente $i$-ésima).

**Estrategia 1** Como estas propiedades provienen de los escalares (los números reales), la estrategia es operar con las componentes dentro de los vectores. Puede escribir sus vectores o como filas o como columnas (en ambos casos la demostración es idéntica).

**Estrategia 2** Recuerde que dos vectores $\boldsymbol{v}$ y $\boldsymbol{w}$ son iguales si son idénticos componente a componente: $\boldsymbol{v}_{|i}=\boldsymbol{w}_{|i}$. Use las propiedades del operador " $|i$ " siempre que pueda; es decir, tenga en cuenta que son intercambiables las expresiones:

- $(\boldsymbol{a}+\boldsymbol{b})_{|i}$ y $\boldsymbol{a}_{|i}+\boldsymbol{b}_{|i}$ (propiedad distributiva de " $|i$ ").
- $\big(\lambda\boldsymbol{a}\big)_{|i}$ y $\lambda\big(\boldsymbol{a}_{|i}\big)$ (propiedad asociativa de " $|i$ ").

Además, recuerde que $\boldsymbol{a}_{|i}$, $\boldsymbol{b}_{|i}$ y $\boldsymbol{c}_{|i}$ son números reales; por tanto, puede usar las propiedades de las operaciones entre números reales del mismo modo que en la primera estrategia.

Para que la demostración sea clara, escriba primero la componente $i$-ésima del lado izquierdo de la igualdad y construya a partir de ella una cadena de igualdades en la que cada paso sustituye una expresión por otra equivalente aplicando una sola regla: una de las dos propiedades de " $|i$ ", una definición (la del vector nulo o la del vector opuesto) o una sola de las ocho propiedades de los números reales. La cadena termina cuando llega a la componente $i$-ésima del lado derecho.

**Nota 1.** *Estas ocho propiedades nos permitirán definir el* espacio vectorial $\mathbb{R}^n$ *en la Lección 6.*

**Sumas de varios vectores y resta de vectores** La propiedad 2 dice que el resultado de sumar tres vectores no depende de cómo se agrupen. Por tanto, podemos omitir los paréntesis y escribir $\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c}$. Lo mismo ocurre con las sumas de más de tres vectores.

El opuesto es el múltiplo que se obtiene al multiplicar por $-1$. Multiplicar un vector por $-1$ cambia el signo de todas sus componentes,[13] así que $(-1)\boldsymbol{a}=-\boldsymbol{a}$.

Restar un vector es sumar su opuesto. Escribimos $\boldsymbol{a}-\boldsymbol{b}$ para denotar $\boldsymbol{a}+(-\boldsymbol{b})$. Por tanto, $\big(\boldsymbol{a}-\boldsymbol{b}\big)_{|i}=\boldsymbol{a}_{|i}-\boldsymbol{b}_{|i}$. Del mismo modo, $\boldsymbol{a}-\lambda\boldsymbol{b}$ denota $\boldsymbol{a}+\big(-(\lambda\boldsymbol{b})\big)$, que es igual a $\boldsymbol{a}+(-\lambda)\boldsymbol{b}$, pues $-(\lambda\boldsymbol{b})=(-1)(\lambda\boldsymbol{b})=(-\lambda)\boldsymbol{b}$ por la propiedad 7.

- Enlace al Notebook de Jupyter para trabajar con lo expuesto en esta Sección 1.2.

## 1.3. Vectores como representación de datos

Dado que los vectores son un *sistema* de números, podemos interpretar cada componente en función de la posición que ocupa. Por ejemplo, imagine que tiene datos del año 2003 sobre la tasa de crecimiento interanual del PIB de cinco países: España (1), Francia (2), Alemania (3), Reino Unido (4) y EE. UU. (5);

[13] Cada componente queda multiplicada por $-1$, y para cualquier número real $x$ se verifica que $(-1)x=-x$. Es otra propiedad de los números reales que seguro conoce, aunque no esté entre las ocho del recordatorio de la Página 8.

entonces podemos interpretar las distintas componentes del siguiente vector en función de su posición en la lista:

$$\boldsymbol{d}_{2003} = \begin{pmatrix} 2.9 \\ 0.9 \\ -0.2 \\ 2.5 \\ 1.8 \end{pmatrix}.$$

Puesto que el número 1.8 está en la quinta posición, sabemos que corresponde al quinto país, y por lo tanto sabemos que EE. UU. creció a una tasa del 1.8 % en el año 2003.[14] ¿Qué país no creció en el año 2003?

*En algunos manuales verá una interpretación de los vectores como segmentos con módulo, dirección y sentido. Esta es una interpretación frecuente en aplicaciones físicas, pero poco natural para los economistas. Para un economista, la interpretación más natural es que las componentes de los vectores de $\mathbb{R}^n$ son datos (precios, cantidades, número de parados, tasas de inflación, etc.) y la posición de cada componente indica a quién corresponde el dato (como en el ejemplo anterior). Esta es la interpretación que se usa en Econometría.*

- Enlace al Notebook de Jupyter de la Sección 1.2, que termina con un apartado sobre lo expuesto en esta sección.

☞ Ahora pasamos a definir las matrices como un sistema de $n$ vectores de $\mathbb{R}^m$ (es decir, una lista ordenada de listas ordenadas). La siguiente introducción así lo ilustra.

## 1.4. Matrices de $\mathbb{R}^{m\times n}$

Imagine que además tiene datos anuales de los citados cinco países durante el período: 2003 a 2006.

$$\boldsymbol{d}_{2003} = \begin{pmatrix} 2.9 \\ 0.9 \\ -0.2 \\ 2.5 \\ 1.8 \end{pmatrix}, \quad \boldsymbol{d}_{2004} = \begin{pmatrix} 3.1 \\ 2.0 \\ 1.1 \\ 3.2 \\ 3.1 \end{pmatrix}, \quad \boldsymbol{d}_{2005} = \begin{pmatrix} 3.3 \\ 1.6 \\ 0.7 \\ 1.9 \\ 3.7 \end{pmatrix}, \quad \boldsymbol{d}_{2006} = \begin{pmatrix} 3.3 \\ 2.0 \\ 0.9 \\ 1.5 \\ 3.5 \end{pmatrix}.$$

Imagine que queremos crear un registro $\mathbf{M}$ con los datos de crecimiento de los cinco países en los años 2003–2006. Podemos hacerlo con una *matriz*...

Una matriz es una lista ordenada de vectores *con el mismo número de componentes* (es decir, ¡una lista de listas de números!); la definición precisa es la Definición 1.8, un poco más abajo. De este modo, podemos definir la matriz $\mathbf{M}$ como el siguiente *sistema* con los cuatro vectores de datos económicos mencionados más arriba:

$$\mathbf{M} = \left[\boldsymbol{d}_{2003};\quad \boldsymbol{d}_{2004};\quad \boldsymbol{d}_{2005};\quad \boldsymbol{d}_{2006};\right].$$

Los vectores se pueden escribir tanto en vertical como en horizontal, pero si organizamos horizontalmente toda la información disponible, su escritura arrojaría una lista larga y difícil de leer:

$$\Big[(2.9,\ 0.9,\ -0.2,\ 2.5,\ 1.8,)\,;\quad (3.1,\ 2.0,\ 1.1,\ 3.2,\ 3.1,)\,;\quad (3.3,\ 1.6,\ 0.7,\ 1.9,\ 3.7,)\,;\quad (3.3,\ 2.0,\ 0.9,\ 1.5,\ 3.5,)\,;\Big].$$

Aunque sabemos que las cuartas componentes de cada vector corresponden al Reino Unido, esta disposición dificulta la lectura.[15] Para evitar este problema, organizaremos los datos correspondientes al año 2003 en

[14] Fuente: Informe ICAE sobre la economía española (octubre de 2006). Instituto Complutense de Análisis Económico.

[15] Un ordenador no tendría ningún problema con esta representación, pero los humanos sí lo tenemos... ¡Imagine una lista de 18 vectores con 27 componentes cada uno! Algo similar ocurriría si escribiéramos una lista vertical de vectores en columna.

una primera columna, los del año 2004 en la segunda, los del año 2005 en la tercera y los del año 2006 en la cuarta. De este modo, obtendremos el siguiente arreglo rectangular de datos:

$$\mathbf{M} = \begin{bmatrix} 2.9 & 3.1 & 3.3 & 3.3 \\ 0.9 & 2.0 & 1.6 & 2.0 \\ -0.2 & 1.1 & 0.7 & 0.9 \\ 2.5 & 3.2 & 1.9 & 1.5 \\ 1.8 & 3.1 & 3.7 & 3.5 \end{bmatrix}.$$

Ahora, cada columna corresponde a un año y cada fila a un país... ¡Mucho más fácil de leer!

**Definición 1.8.** *Llamamos* matriz *de* $\mathbb{R}^{m\times n}$ *a un sistema de* $n$ *vectores de* $\mathbb{R}^m$.

$$\begin{bmatrix}\boldsymbol{a};\ \boldsymbol{b};\ \ldots\ \boldsymbol{v};\end{bmatrix} \qquad \text{donde } \boldsymbol{a}, \boldsymbol{b}, \ldots \boldsymbol{v} \text{ son } n \text{ vectores de } \mathbb{R}^m.$$

*Cada vector del sistema se escribe verticalmente y se denomina* columna *de la matriz.*

Con las matrices usaremos indistintamente tanto la representación genérica de los sistemas, como otra que mantiene los corchetes pero omite los puntos y comas detrás de cada columna:

$$\begin{bmatrix}\boldsymbol{a}\ \boldsymbol{b}\ldots\ \boldsymbol{v}\end{bmatrix}.$$

Consecuentemente, tenemos el siguiente convenio de notación: encerrando un vector entre corchetes, escribimos una matriz con una única columna (*matriz columna*). Por ejemplo, si $\boldsymbol{x} = \begin{pmatrix}1, & 5, & 9,\end{pmatrix}$, entonces

$$\begin{bmatrix}\boldsymbol{x}\end{bmatrix} = \begin{bmatrix}1\\5\\9\end{bmatrix}.$$

Al encerrar entre corchetes una lista de vectores (todos con el mismo número de componentes), denotamos la matriz cuyas columnas son dichos vectores en la misma disposición en la que aparecen en la lista. Así, si

$$\boldsymbol{a} = \begin{pmatrix}1, & 0, & 3, & 0,\end{pmatrix}, \quad \boldsymbol{f} = \begin{pmatrix}2, & 2, & 2, & 2,\end{pmatrix}, \quad \text{y} \quad \mathbf{0} = \begin{pmatrix}0, & 0, & 0, & 0,\end{pmatrix},$$

entonces:

$$\begin{bmatrix}\boldsymbol{a} & \boldsymbol{f} & \mathbf{0} & \boldsymbol{f}\end{bmatrix} = \left[\begin{pmatrix}1\\0\\3\\0\end{pmatrix}; \begin{pmatrix}2\\2\\2\\2\end{pmatrix}; \begin{pmatrix}0\\0\\0\\0\end{pmatrix}; \begin{pmatrix}2\\2\\2\\2\end{pmatrix};\right] = \begin{bmatrix}1 & 2 & 0 & 2\\0 & 2 & 0 & 2\\3 & 2 & 0 & 2\\0 & 2 & 0 & 2\end{bmatrix}.$$

**Definición 1.9.** *El* orden *de una matriz de* $\mathbb{R}^{m\times n}$ *está compuesto por dos números: el primero* $(m)$ *es el número de componentes de cada columna, que es el número de filas que se ven al escribir la matriz; y el segundo* $(n)$ *es el número de columnas. El orden se suele expresar como "*$m \times n$*" (y se lee "*$m$ *por* $n$*").*

En el ejemplo con datos económicos, el orden de $\mathbf{M}$ es 5 por 4. Recuerde que *primero se indica el número de filas y luego el de columnas.* Frecuentemente expresaremos el orden de la matriz bajo su nombre: $\mathbf{M}_{5\times 4}$.

Los vectores de la lista que forma una matriz tienen nombre propio: son sus *columnas*. Reservamos la palabra *componente* para los números: llamamos *componentes* (o elementos) de una matriz a las componentes de sus columnas. Al igual que en el caso de los vectores, identificamos las componentes de la matriz por la posición que ocupan. De esta manera, podemos decir que la única componente negativa de $\mathbf{M}$ está ubicada en la tercera fila, primera columna.

**Notación.** Denotamos las matrices con letras mayúsculas, sin cursiva y en negrita: $\mathbf{A}, \mathbf{B}, \ldots, \mathbf{Y}, \mathbf{Z}$; y sus componentes genéricas con idéntica letra, en minúscula cursiva, y con dos subíndices que indican la posición que ocupa dentro de la matriz (comenzando siempre por la fila, y luego por la columna). Consecuentemente

$a_{ij}$ es la componente de $\mathbf{A}$ de la fila $i$-ésima y columna $j$-ésima; es decir, la componente $i$-ésima de la columna $j$-ésima. Otra forma de denotar dicha componente es:

$$a_{ij} \;\equiv\; \mathrm{elem}_{ij}(\mathbf{A}).$$

Así, escribimos la matriz genérica $\mathbf{A}$ de orden 2 por 4 como:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix};$$

y de la matriz

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 7 \end{bmatrix}$$

podemos afirmar que su orden es $2 \times 3$, que su componente $b_{21}$ es cero y que su componente $b_{23}$ es 7.

Nótese que usamos *un único* índice para identificar las componentes de un vector, pero necesitamos *dos* para identificar las componentes de una matriz. Así, al referirnos a la componente $i$-ésima de $\boldsymbol{x}$ escribimos:

$$x_i \qquad \text{o} \qquad \boldsymbol{x}_{|i};$$

pero si queremos referirnos a la componente de la fila $i$-ésima y columna $j$-ésima de $\mathbf{A}$ escribimos:

$$\mathrm{elem}_{ij}(\mathbf{A}) \qquad \text{o} \qquad a_{ij}.$$

¿Y cómo haremos para denotar toda una fila, toda una columna o una componente con nuestro operador selector? Veámoslo.

### 1.4.1. Filas de una matriz

La fila $i$-ésima de $\mathbf{A}$ (de orden $m \times n$) es el *vector* de $\mathbb{R}^n$ cuyas componentes son $a_{ik}$, donde el índice $k$ recorre todas las columnas ($k = 1 : n$) y que denotamos como:

$$\mathrm{fila}_i\,(\mathbf{A}) \quad \equiv \quad {}_{i|}\mathbf{A} \quad = \big(a_{i1}, \ldots, a_{in},\big) \;\in \mathbb{R}^n$$

(que por ser un vector de $\mathbb{R}^n$ se puede escribir tanto en horizontal como en vertical).

**El operador selector de filas:** Escribiendo $i\,|$ como subíndice a la *izquierda* de una matriz, seleccionamos su *fila* $i$-ésima.

Por ejemplo, si $\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 7 \end{bmatrix}$, entonces ${}_{1|}\mathbf{B} = \big(1, \quad 2, \quad 3,\big) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

Módulo NAcAL para Python

```
1|B          # primera fila de B
```

$$\big(1, \quad 2, \quad 3,\big)$$

### 1.4.2. Columnas de una matriz

Las columnas de $\mathbf{A}$ (de orden $m \times n$) son los vectores de su lista (Definición 1.8). La columna $j$-ésima es el *vector* de $\mathbb{R}^m$ que ocupa la posición $j$. Sus componentes son $a_{kj}$, donde el índice $k$ recorre todas las filas ($k = 1 : m$). Denotaremos dicha columna como $\mathrm{col}_j\,(\mathbf{A})$ o bien como $\mathbf{A}_{|j}$:

$$\mathrm{col}_j\,(\mathbf{A}) \quad \equiv \quad \mathbf{A}_{|j} \quad = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \mathbb{R}^m,$$

(que por ser un vector de $\mathbb{R}^m$ se puede escribir tanto en horizontal como en vertical).

**El operador selector de columnas:** Escribiendo $\,|\,j\,$ como subíndice a la *derecha* de una matriz, seleccionamos su *columna* $j$-ésima.

Así, si $\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 7 \end{bmatrix}$; entonces $\mathbf{B}_{|2} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$.

Módulo NAcAL para Python

```
B|2          # segunda columna de B
```

¿Y cómo escribir la matriz de orden $2\times 1$ cuya única columna es la segunda columna de $\mathbf{B}$? Basta encerrar entre "corchetes" dicha columna:

$$\left[\mathbf{B}_{|2}\right] = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Módulo NAcAL para Python

```
Matrix( [ B|2 ] ) # Matriz cuya única columna es la segunda columna de B
```

$$\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Nótese que $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ es un *vector* (una lista de dos números, que se puede escribir en vertical o en horizontal), pero que $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ es una *matriz* cuya lista contiene una única *columna*. Vectores y matrices son objetos matemáticos distintos (lista de números *vs* lista de vectores) y gracias a estas reglas de notación serán fácilmente distinguibles: los vectores se encierran entre "paréntesis" y las matrices entre "corchetes", de manera que $\boldsymbol{a} = \big(1,\quad 2,\quad 3,\big)$ es un vector y $\mathbf{A} = \big[1\quad 2\quad 3\big]$ es una matriz (y por tanto **son objetos distintos**).[16]

Con la notación que hemos establecido, podemos describir una matriz mediante sus columnas:

$$\underset{m\times n}{\mathbf{A}} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \underbrace{\big[\mathbf{A}_{|1};\quad \mathbf{A}_{|2};\quad \dots\quad \mathbf{A}_{|n};\big]}_{\text{lista de columnas}};\quad \text{donde}\quad \mathbf{A}_{|j} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Diremos que dos matrices del mismo orden son iguales si y solo si lo son sus correspondientes listas de vectores; es decir: $\mathbf{A} = \mathbf{B}$ si y solo si $\mathbf{A}_{|j} = \mathbf{B}_{|j}$ para $j = 1:n$ (i.e. si son iguales columna a columna). Dos matrices de distinto orden no son iguales.

Como las columnas son vectores, y dos vectores son iguales si lo son sus componentes, se deduce inmediatamente que dos matrices del mismo orden son iguales si y solo si lo son componente a componente; es decir, $\mathbf{A} = \mathbf{B}$ si y solo si $a_{ij} = b_{ij}$ para $i = 1:m$ y $j = 1:n$.

**Selección de una componente:** Cada componente de una matriz está a la vez en una fila y en una columna. En la matriz $\mathbf{B}$ del ejemplo anterior, el 7 es la tercera componente de la segunda fila y la segunda componente de la tercera columna; así que lo podemos obtener de dos maneras: seleccionando primero la fila y después su tercera componente, o seleccionando primero la columna y después su segunda componente:

$$\big({}_{2|}\mathbf{B}\big)_{|3} = \big(0,\quad -1,\quad 7,\big)_{|3} = 7 \qquad \text{y} \qquad {}_{2|}\big(\mathbf{B}_{|3}\big) = {}_{2|}\begin{pmatrix} 3 \\ 7 \end{pmatrix} = 7.$$

[16] **Nota para profesores.** La definición formal de $\mathbb{R}^n$ que distingue un número de un vector de $\mathbb{R}^1$ se explica en la introducción (página xxv).

En el segundo caso el selector actúa por la izquierda de un *vector*: como dijimos en la Sección 1.1, el operador "|" puede operar por cualquier lado de un vector, con el mismo resultado:

$$ {}_{i|}\boldsymbol{a} = \boldsymbol{a}_{|i} = a_i. $$

Con este convenio, y puesto que tanto las filas como las columnas de una matriz son vectores, al seleccionar las *componentes* de $\underset{m\times n}{\mathbf{A}}$ nos dará igual operar dos veces por el mismo lado o una vez por cada lado:

$$ {}_{i|}\big(\mathbf{A}_{|j}\big) = \big(\mathbf{A}_{|j}\big)_{|i} \qquad \text{y} \qquad \big({}_{i|}\mathbf{A}\big)_{|j} = {}_{j|}\big({}_{i|}\mathbf{A}\big). $$

Y como ambos caminos llevan a la misma componente de $\mathbf{A}$ (la de la fila $i$ y la columna $j$), da igual seleccionar primero la fila $i$ y después su componente $j$-ésima, o primero la columna $j$ y después su componente $i$-ésima: podemos desplazar los paréntesis sin cambiar el resultado. Es decir, la selección de componentes de una matriz es *asociativa* (como el producto por escalares), y por eso podemos prescindir de los paréntesis y escribir sencillamente

$$ {}_{i|}\mathbf{A}_{|j} = {}_{i|}\big(\mathbf{A}_{|j}\big) = \big({}_{i|}\mathbf{A}\big)_{|j}. $$

Así, en el ejemplo, ${}_{2|}\mathbf{B}_{|3} = 7$.

terminal de Python

```
>>> a = Vector( [1, 0] )
>>> b = Vector( [2,-1] )
>>> c = Vector( [3, 7] )
>>> B = Matrix( [a, b, c] ) # Matrix cuya lista de vectores es a, b, c
>>> 2|B|3                   # componente de B situada en la segunda fila, tercera columna
7
```

### 1.4.3. Resumen de la notación para seleccionar componentes, filas y columnas

Disponemos de varias alternativas para denotar tanto las componentes, como las filas y columnas de $\underset{m\times n}{\mathbf{A}}$.

**Notaciones alternativas**

| Componente $ij$ (escalar) | Fila $i$-ésima (vector de $\mathbb{R}^n$) | Columna $j$-ésima (vector de $\mathbb{R}^m$) |
|---|---|---|
| $a_{ij}$ | | |
| ${}_{i\|}\mathbf{A}_{\|j}$ | ${}_{i\|}\mathbf{A}$ | $\mathbf{A}_{\|j}$ |
| $\mathrm{elem}_{ij}(\mathbf{A})$ | $\mathrm{fila}_i(\mathbf{A})$ | $\mathrm{col}_j(\mathbf{A})$ |

En el ejemplo con datos económicos, ${}_{1|}\mathbf{M}$ (primera fila) corresponde a los datos de España y ${}_{5|}\mathbf{M}$ (quinta fila) a los datos de EE. UU.; y, por otra parte, $\mathbf{M}_{|3}$ (tercera columna) corresponde a los datos del año 2005.

☞ *La notación de las matrices tiene paralelismos con la de los vectores;* ***pero con importantes diferencias.***

Una lista de vectores encerrada entre *corchetes* denota una matriz, cuyas *columnas* son los vectores de la lista en el orden en que aparecen en la lista dentro del corchete. Por ejemplo, si

$$ \boldsymbol{a} = \big(1,\ 2,\big), \quad \boldsymbol{b} = \big(1,\ 0,\big), \quad \boldsymbol{c} = \big(9,\ 2,\big); \quad \text{entonces} \quad \mathbf{A} = \big[\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c}\big] = \begin{bmatrix} 1 & 1 & 9 \\ 2 & 0 & 2 \end{bmatrix}. $$

Cuando una matriz se escribe como una tabla de números, sin los puntos y comas, NO hay libertad para escribir horizontal o verticalmente la lista que hemos encerrado (es una importante diferencia con los vectores).

Los subíndices "$i\,|$" (por la izquierda) y "$|\,j$" (por la derecha) operan sobre la matriz, pero su comportamiento es distinto. El subíndice "$i\,|$" a la izquierda de la matriz selecciona la *fila* $i$-ésima y el subíndice "$|\,j$" a la derecha selecciona la *columna* $j$-ésima (tanto las filas como las columnas son *vectores*):

$$\mathbf{A}_{|3} = \begin{pmatrix} 9 \\ 2 \end{pmatrix}, \qquad {}_{2|}\mathbf{A} = \begin{pmatrix} 2, & 0, & 2, \end{pmatrix}.$$

Así, $\mathbf{A}_{|3}$ es un *vector* formado por las componentes de la tercera columna de $\mathbf{A}$; pero $\left[\mathbf{A}_{|3}\right]$ es una *matriz columna* cuya única columna es igual a la tercera columna de $\mathbf{A}$.

$$\mathbf{A}_{|3} = \boldsymbol{c} = \begin{pmatrix} 9, & 2, \end{pmatrix} = \begin{pmatrix} 9 \\ 2 \end{pmatrix} \quad \neq \quad \left[\mathbf{A}_{|3}\right] = \left[\boldsymbol{c}\right] = \begin{bmatrix} 9 \\ 2 \end{bmatrix} \quad \neq \quad \begin{bmatrix} 9 & 2 \end{bmatrix}.$$

Cuando se emplean dos subíndices (uno por cada lado), se selecciona la *componente* situada en la fila $i$-ésima (indicada por el índice de la izquierda) y columna $j$-ésima (indicada por el índice de la derecha)

$${}_{1|}\mathbf{A}_{|3} = a_{13} = 9.$$

### 1.4.4. Extensión de la notación vectorial y reglas de reescritura

En la Sección 1.1 aceptamos que "$|$" opere también por la izquierda de un *vector*, con el mismo resultado: ${}_{i|}\boldsymbol{a} = \boldsymbol{a}_{|i}$. Como consecuencia, el selector por la izquierda también es lineal:

$${}_{i|}(\boldsymbol{a}+\boldsymbol{b}) = (\boldsymbol{a}+\boldsymbol{b})_{|i} = \boldsymbol{a}_{|i} + \boldsymbol{b}_{|i} = {}_{i|}\boldsymbol{a} + {}_{i|}\boldsymbol{b} \qquad \text{y} \qquad {}_{i|}(\lambda\boldsymbol{b}) = (\lambda\boldsymbol{b})_{|i} = \lambda(\boldsymbol{b}_{|i}) = \lambda\Big({}_{i|}\boldsymbol{b}\Big).$$

(En las dos cadenas, el primer paso y el último son el convenio que acabamos de recordar; el paso central es la linealidad del selector por la derecha.)

Para no perder la *asociatividad* de la notación cuando "$|$" opera por la izquierda de los *vectores*, también aceptaremos que un escalar $\lambda$ aparezca multiplicando por la derecha de un vector $\boldsymbol{b}$; la expresión $\boldsymbol{b}\lambda$ denota, por definición, el producto de $\boldsymbol{b}$ por $\lambda$:[17]

$$\boxed{\boldsymbol{b}\lambda = \lambda\boldsymbol{b}}$$

No es una operación nueva, sino un convenio de notación: otra manera de escribir el mismo producto. Con el escalar a la derecha, la asociatividad del selector por la izquierda también se puede escribir "desplazando los paréntesis":

$${}_{i|}(\boldsymbol{b}\lambda) = {}_{i|}(\lambda\boldsymbol{b}) = \lambda\Big({}_{i|}\boldsymbol{b}\Big) = \Big({}_{i|}\boldsymbol{b}\Big)\lambda.$$

(El primer paso es el convenio que acabamos de adoptar; el segundo, que el selector por la izquierda es asociativo respecto al producto por escalares; y el tercero, la propiedad conmutativa del producto de números reales, pues ${}_{i|}\boldsymbol{b}$ y $\lambda$ son números.) Y lo mismo ocurre con el selector por la derecha:

$$(\boldsymbol{b}\lambda)_{|i} = (\lambda\boldsymbol{b})_{|i} = \lambda(\boldsymbol{b}_{|i}) = (\boldsymbol{b}_{|i})\lambda.$$

Esto da lugar (jugando con multiplicar por la derecha o seleccionar por la izquierda) a dos nuevas reglas de reescritura:

[17] El **modo habitual de escribir** este producto es "$2\boldsymbol{a}$", con el coeficiente primero. Aquí aceptaremos también "$\boldsymbol{a}2$" para mantener la asociatividad por la izquierda, pero tenga en cuenta que no es una escritura habitual.

☞ Primera. - "*podemos desplazar los paréntesis*" para sacar un símbolo por un lado e introducir otro por el otro lado:

$$(\lambda\boldsymbol{b})_{|i} \quad = \quad \lambda(\boldsymbol{b}_{|i}) \qquad = \qquad \left({}_{i|}\boldsymbol{b}\right)\lambda \quad = \quad {}_{i|}(\boldsymbol{b}\lambda).$$

Segunda. - "*podemos intercambiar las posiciones del escalar y el selector*" sacando uno de ellos fuera del paréntesis y metiendo el otro:

$$(\boldsymbol{b}_{|i})\lambda \quad = \quad (\boldsymbol{b}\lambda)_{|i} \qquad \text{y} \qquad \lambda\left({}_{i|}\boldsymbol{b}\right) \quad = \quad {}_{i|}(\lambda\boldsymbol{b}).$$

Puesto que las ocho expresiones de más arriba arrojan el mismo resultado (aunque el orden de ejecución de las operaciones difiera entre ellas) podemos omitir los paréntesis y escribir sencillamente $\lambda\boldsymbol{b}_{|i}$ o ${}_{i|}\boldsymbol{b}\lambda$.

No obstante, en las demostraciones siempre usaremos expresiones con paréntesis, pues ayudan a identificar la propiedad empleada en cada paso de la demostración.

Ejercicio 5. ¿Cómo evalúa el módulo NAcAL las siguientes expresiones si `b` es un vector de $\mathbb{R}^3$? (Revise el Cuadro 1.1 en la página 7)

(a) `3*b|1` (b) `3*1|b`
(c) `1*3|b` (d) `1|b*3`

### 1.4.5. La transposición (o trasposición)

Encerrando $\boldsymbol{a}$ entre corchetes denotamos una matriz cuya única columna es $\boldsymbol{a}$. No tenemos una escritura parecida para la matriz cuya única fila es $\boldsymbol{a}$. La obtendremos con ayuda de una operación adicional: la *transposición*, `~`.

Tomemos la matriz $\mathbf{B}$ de los ejemplos anteriores y formemos otra matriz con dos vectores: el primero con los números de la primera fila de $\mathbf{B}$, y el segundo con los de la segunda:

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 7 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} {}_{1|}\mathbf{B}; & {}_{2|}\mathbf{B}; \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 7 \end{bmatrix}.$$

Convertir las filas de una matriz en las columnas de otra se llama *transponer*: la matriz que hemos formado es la *transpuesta* de $\mathbf{B}$, y la denotamos $\mathbf{B}^\intercal$. Fíjese en que $\mathbf{B}$ es de orden $2 \times 3$ y $\mathbf{B}^\intercal$ es de orden $3 \times 2$.

**Definición 1.10** (Matriz transpuesta (o traspuesta))**.** *Sea* $\mathbf{A}$ *de orden* $m$ *por* $n$*, la matriz transpuesta de* $\mathbf{A}$*, que denotamos con* $\mathbf{A}^\intercal$*, es la matriz de orden* $n \times m$ *cuyas* $m$ *columnas son iguales a las* $m$ *filas de* $\mathbf{A}$*:*

$$(\mathbf{A}^\intercal)_{|i} = {}_{i|}\mathbf{A}; \qquad i = 1 : m.$$

Es decir, cada columna de $\mathbf{A}^\intercal$ es la fila de $\mathbf{A}$ con el mismo índice: la transpuesta de $\underset{m\times n}{\mathbf{A}}$ es la matriz cuyas columnas son las filas de $\mathbf{A}$ escritas en vertical:

$$\mathbf{A}^\intercal = \begin{bmatrix} {}_{1|}\mathbf{A}; & \dots & {}_{m|}\mathbf{A}; \end{bmatrix}.$$

Módulo NAcAL para Python

```
~B          # Transpuesta de la matriz B
```

$$\begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 7 \end{bmatrix}$$

Escribimos una *matriz columna* (matriz con una única columna) poniendo un vector entre corchetes; ahora podemos escribir una *matriz fila* (matriz con una única fila) transponiendo una matriz columna; por ejemplo, $\left(\begin{bmatrix}1\\0\\5\\1\end{bmatrix}\right)^{\mathsf{T}} = \begin{bmatrix}1 & 0 & 5 & 1\end{bmatrix}$.

Así $\left[{}_{i|}\mathbf{A}\right]^{\mathsf{T}}$ es una matriz cuya única fila es igual a la fila $i$-ésima de $\mathbf{A}$; y si

$$\mathbf{B} = \begin{bmatrix}1 & 2 & 3\\ 0 & -1 & 7\end{bmatrix}; \quad \text{entonces} \quad \begin{cases}\left[{}_{1|}\mathbf{B}\right]^{\mathsf{T}} = \begin{bmatrix}1 & 2 & 3\end{bmatrix}\\ \left[\mathbf{B}_{|3}\right]^{\mathsf{T}} = \begin{bmatrix}3 & 7\end{bmatrix}\end{cases}.$$

Módulo NAcAL para Python

```
~Matrix( [1|B] )    # Matriz cuya única fila es la primera fila de B
```

$$\begin{bmatrix}1 & 2 & 3\end{bmatrix}$$

Módulo NAcAL para Python

```
~Matrix( [B|3] )    # Matriz cuya única fila es la tercera columna de B
```

$$\begin{bmatrix}3 & 7\end{bmatrix}$$

Ejercicio 6. Demuestre las reglas de reescritura de los apartados (a) y (b), y después la propiedad de la transposición del apartado (c):

(a) La transposición intercambia los índices de las componentes de la matriz: ${}_{i|}(\mathbf{A}^{\mathsf{T}})_{|j} = {}_{j|}\mathbf{A}_{|i}$.

(b) Si $\underset{m\times n}{\mathbf{A}}$ y $1 \le i \le n$ entonces ${}_{i|}(\mathbf{A}^{\mathsf{T}}) = \mathbf{A}_{|i}$.

(c) $(\mathbf{A}^{\mathsf{T}})^{\mathsf{T}} = \mathbf{A}$.

☞ Pese a que las expresiones ${}_{i|}(\mathbf{A}^{\mathsf{T}})$, $(\mathbf{A}^{\mathsf{T}})_{|j}$ y ${}_{i|}(\mathbf{A}^{\mathsf{T}})_{|j}$ solo tienen una interpretación posible incluso si se quitaran los paréntesis (puesto que *la transposición es una operación sobre matrices* y no sobre vectores) siempre escribiremos los paréntesis. Así la notación es más clara.

Fíjese además en los apartados (a) y (b) del ejercicio anterior: "*puedo transponer (o quitar la transposición) si a la vez cambio de lado los subíndices*". Las usaremos como *reglas de reescritura*. Una regla de reescritura es una igualdad que empleamos para sustituir una expresión por otra. Unas reglas son definiciones o convenios de notación, y no se demuestran: por ejemplo, la definición de la suma de vectores. Las demás se deducen de las definiciones, y esa deducción es una *demostración* aunque sea breve. *Comprobar* es otra cosa: es verificar una igualdad en un caso concreto, y eso no la demuestra para todos los casos.

### 1.4.6. Tipos de matrices

☞ Al igual que hicimos con los vectores, daremos nombre a distintos tipos de matrices según su forma (cuadradas o rectangulares), la disposición de sus componentes (nulas, diagonales, identidad) o su relación con su transpuesta (simétricas). Una misma matriz puede ser de varios tipos a la vez. *Preste atención al uso de la notación.*

**Definición 1.11.** *Decimos que una matriz es cuadrada cuando tiene el mismo número de filas que de columnas. Es decir, una matriz cuadrada es de orden $n \times n$.*

Cuando tenemos una matriz cuadrada de orden $n$ por $n$ abreviamos y decimos sencillamente que es de orden $n$. Así, si decimos que una matriz es *de orden $n$* (y nada más) estamos indicando que es cuadrada.

**Definición 1.12.** *A las matrices que no son cuadradas las denominamos rectangulares.*

Recuerde: al expresar el orden de matrices rectangulares es necesario indicar el número de filas y columnas.

Las componentes de una matriz cuyos dos índices son iguales (${}_{1|}\mathbf{A}_{|1}$, ${}_{2|}\mathbf{A}_{|2}$, etc.) forman su *diagonal principal*.

**Definición 1.13.** *Decimos que una matriz $\mathbf{D}$ es diagonal cuando las componentes fuera de la diagonal principal son nulas; es decir, cuando ${}_{i|}\mathbf{D}_{|j} = 0$ siempre que $i \neq j$.*

Por ejemplo

$$\mathbf{D} = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{mm} \end{bmatrix}; \quad \mathbf{F} = \begin{bmatrix} f_{11} & 0 & \cdots & 0 & 0 \\ 0 & f_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & f_{mm} & 0 \end{bmatrix}; \quad \text{y} \quad \mathbf{G} = \begin{bmatrix} g_{11} & 0 & \cdots & 0 \\ 0 & g_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g_{nn} \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Fíjese que de las tres matrices diagonales, solo $\mathbf{D}$ es cuadrada. Tenga en cuenta además que las componentes de la diagonal principal pueden tomar cualquier valor.

**Definición 1.14.** *Llamamos matriz nula (o matriz cero) de orden $m \times n$, $\mathbf{0}$, a la matriz cuyas $n$ columnas son vectores nulos de $\mathbb{R}^m$ : $\mathbf{0}_{|j} = \mathbf{0} \in \mathbb{R}^m$ para $j = 1 : n$.*

Módulo NAcAL para Python

```
MO(2,4)        # matriz nula de orden 2 por 4
```

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Todas las componentes de una matriz nula son cero. Por tanto, toda matriz nula $\mathbf{0}$ es diagonal.

**Definición 1.15.** *Llamamos matriz identidad (que denotamos con $\mathbf{I}$) a una matriz cuadrada y diagonal cuyas componentes en la diagonal principal son unos. Es decir,* ${}_{i|}\mathbf{I}_{|j} = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases}$.

Por ejemplo, las matrices identidad de órdenes 2 y 1 son: $\underset{2\times 2}{\mathbf{I}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$; $\underset{1\times 1}{\mathbf{I}} = \begin{bmatrix} 1 \end{bmatrix}$.

Módulo NAcAL para Python

```
I(2)           # matriz identidad de orden 2
```

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

**Definición 1.16** (Matriz simétrica)**.** *Decimos que la matriz $\mathbf{A}$ es simétrica cuando $\mathbf{A} = \mathbf{A}^\intercal$.*

Si $\mathbf{A}$ es de orden $m \times n$, su transpuesta $\mathbf{A}^\intercal$ es de orden $n \times m$. Como dos matrices solo pueden ser iguales si son del mismo orden, ha de ser $m = n$; es decir, toda matriz simétrica es cuadrada. Además, como $\mathbf{A}_{|j} = (\mathbf{A}^\intercal)_{|j} = {}_{j|}\mathbf{A}$, la "*columna*" $j$-ésima es igual a la "*fila*" $j$-ésima; por tanto, las componentes $i$-ésimas de dichos vectores son iguales: ${}_{i|}\mathbf{A}_{|j} = {}_{j|}\mathbf{A}_{|i}$.

*Toda matriz diagonal y cuadrada es simétrica*, ya que por una parte

$$ {}_{i|}\mathbf{D}_{|j} = \begin{cases} 0 & (i \neq j) \\ d_{ii} & (i = j) \end{cases}, \quad \text{y por otra} \quad {}_{i|}(\mathbf{D}^\intercal)_{|j} = {}_{j|}\mathbf{D}_{|i} = \begin{cases} 0 & (i \neq j) \\ d_{ii} & (i = j) \end{cases}; \quad \text{por tanto} \quad \mathbf{D}^\intercal = \mathbf{D}. $$

En particular, las matrices identidad son simétricas, pues son cuadradas y diagonales.

- Enlace al Notebook de Jupyter para trabajar con lo expuesto en esta Sección.

## 1.5. Suma de matrices y producto de matrices por escalares

Las columnas de una matriz son vectores, y con los vectores ya sabemos operar. Por eso definimos las operaciones con matrices a partir de las operaciones con sus columnas.

**Definición 1.17** (suma de matrices)**.** *Se define la suma de dos matrices* $\mathbf{A}$ *y* $\mathbf{B}$*, de orden* $m \times n$*, como la matriz del mismo orden cuya columna* $j$*-ésima es la suma de la columna* $j$*-ésima de* $\mathbf{A}$ *y la* $j$*-ésima de* $\mathbf{B}$*:*

$$ \boxed{(\mathbf{A}+\mathbf{B})_{|j} = \mathbf{A}_{|j} + \mathbf{B}_{|j}} \qquad \textit{para} \;\; j = 1:n, $$

*es decir,* $\mathbf{A}+\mathbf{B} = \left[(\mathbf{A}_{|1}+\mathbf{B}_{|1}); \quad \dots \quad (\mathbf{A}_{|n}+\mathbf{B}_{|n});\right]$.

**Definición 1.18** (producto de una matriz por un escalar)**.** *Se define el producto de una matriz* $\underset{m\times n}{\mathbf{A}}$ *por un escalar* $\lambda$ *como la matriz resultante de multiplicar las columnas de* $\mathbf{A}$ *por el escalar* $\lambda$*:*

$$ \boxed{(\lambda\mathbf{A})_{|j} = \lambda(\mathbf{A}_{|j})} \qquad \textit{para} \;\; j = 1:n, $$

*es decir,* $\lambda\mathbf{A} = \left[\lambda(\mathbf{A}_{|1}); \quad \dots \quad \lambda(\mathbf{A}_{|n});\right]$.

**Definición 1.19.** *Para cualquier* $\lambda$*, decimos que* $\lambda\mathbf{A}$ *es* un múltiplo *de* $\mathbf{A}$.

Por ejemplo, tomemos dos matrices de orden $2 \times 3$ que ya han aparecido en la lección:

$$ \mathbf{A} = \begin{bmatrix} 1 & 1 & 9 \\ 2 & 0 & 2 \end{bmatrix} \qquad \text{y} \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 7 \end{bmatrix}. $$

Su suma es la matriz cuyas columnas son las sumas de las columnas de $\mathbf{A}$ y de $\mathbf{B}$; y el múltiplo $2\mathbf{B}$ es la matriz cuyas columnas son las de $\mathbf{B}$ multiplicadas por 2:

$$ \mathbf{A}+\mathbf{B} = \left[\begin{pmatrix}1\\2\end{pmatrix}+\begin{pmatrix}1\\0\end{pmatrix}; \quad \begin{pmatrix}1\\0\end{pmatrix}+\begin{pmatrix}2\\-1\end{pmatrix}; \quad \begin{pmatrix}9\\2\end{pmatrix}+\begin{pmatrix}3\\7\end{pmatrix};\right] = \begin{bmatrix}2 & 3 & 12\\ 2 & -1 & 9\end{bmatrix}; $$

$$ 2\mathbf{B} = \left[2\begin{pmatrix}1\\0\end{pmatrix}; \quad 2\begin{pmatrix}2\\-1\end{pmatrix}; \quad 2\begin{pmatrix}3\\7\end{pmatrix};\right] = \begin{bmatrix}2 & 4 & 6\\ 0 & -2 & 14\end{bmatrix}. $$

Módulo NAcAL para Python

```
A + B          # suma de las matrices A y B
```

$$ \begin{bmatrix} 2 & 3 & 12 \\ 2 & -1 & 9 \end{bmatrix} $$

Módulo NAcAL para Python

```
2*B            # múltiplo de B
```

$$\begin{bmatrix} 2 & 4 & 6 \\ 0 & -2 & 14 \end{bmatrix}$$

☞ Nuevamente, las definiciones de la suma y del producto por escalares convierten al operador "$|j$" (que selecciona la columna $j$-ésima de una matriz) en un **operador lineal**; es decir, *distributivo* respecto a la suma de matrices

$$\big(\mathbf{A}+\mathbf{B}\big)_{|j} = \mathbf{A}_{|j} + \mathbf{B}_{|j}$$

y *asociativo* respecto al producto de una matriz por un *escalar*

$$\big(\lambda\mathbf{A}\big)_{|j} = \lambda\big(\mathbf{A}_{|j}\big).$$

Todo lo que sigue se deduce de estas dos definiciones "por columnas", y en este orden: que las matrices también se suman y se multiplican por un escalar *componente a componente*; las ocho propiedades de estas operaciones; que también se puede operar *fila a fila*; y, por último, que la transposición es un operador lineal.

### 1.5.1. Operaciones componente a componente

En el ejemplo anterior, cada componente de $\mathbf{A}+\mathbf{B}$ ha resultado ser la suma de las componentes de $\mathbf{A}$ y de $\mathbf{B}$ situadas en la misma posición; y cada componente de $2\mathbf{B}$, el doble de la correspondiente componente de $\mathbf{B}$. De hecho, la mayoría de los manuales definen la suma y el producto por escalares componente a componente. Definen la suma de dos matrices $\mathbf{A}$ y $\mathbf{B}$ del mismo orden como la matriz que resulta de sumar cada componente de $\mathbf{A}$ con la correspondiente componente de $\mathbf{B}$. Y definen el producto de una matriz $\mathbf{A}$ por un escalar $\lambda$ como la matriz que resulta de multiplicar cada componente de $\mathbf{A}$ por $\lambda$. Con nuestras definiciones esto no es una definición, sino una consecuencia:

Ejercicio 7. Demuestre que las definiciones 1.17 y 1.18 implican que las operaciones de suma y producto por escalares se pueden calcular componente a componente; es decir, que:

$$_{i|}(\mathbf{A}+\mathbf{B})_{|j} = {}_{i|}\mathbf{A}_{|j} + {}_{i|}\mathbf{B}_{|j}, \qquad \text{y} \qquad {}_{i|}(\lambda\mathbf{A})_{|j} = \lambda\Big({}_{i|}\mathbf{A}_{|j}\Big); \qquad \text{con } i = 1:m, \quad \text{y} \quad j = 1:n.$$

*Pista.* La componente ${}_{i|}\mathbf{A}_{|j}$ es la componente $i$-ésima de la columna $j$-ésima: ${}_{i|}\big(\mathbf{A}_{|j}\big)$. Seleccione primero la columna (ahí se aplican las definiciones) y después la componente de esa columna, que es un vector.

O, empleando expresiones menos compactas,

$$\mathbf{A}+\mathbf{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} (a_{11}+b_{11}) & (a_{12}+b_{12}) & \cdots & (a_{1n}+b_{1n}) \\ \vdots & \vdots & \ddots & \vdots \\ (a_{m1}+b_{m1}) & (a_{m2}+b_{m2}) & \cdots & (a_{mn}+b_{mn}) \end{bmatrix};$$

y

$$\lambda\mathbf{A} = \lambda \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \lambda(a_{11}) & \lambda(a_{12}) & \cdots & \lambda(a_{1n}) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda(a_{m1}) & \lambda(a_{m2}) & \cdots & \lambda(a_{mn}) \end{bmatrix}.$$

### 1.5.2. Propiedades de la suma de matrices y del producto por escalares

Para enunciar las propiedades nos falta una definición, análoga a la del vector opuesto.

**Definición 1.20.** *Sea la matriz* $\mathbf{A}$*. Llamamos* matriz opuesta *de* $\mathbf{A}$ *a la matriz* $-\mathbf{A}$*, que es la matriz cuyas columnas son las opuestas de las de* $\mathbf{A}$*; es decir,* $\big(-\mathbf{A}\big)_{|j} = -\big(\mathbf{A}_{|j}\big)$*.*

Módulo NAcAL para Python

```
-B          # matriz opuesta de B
```

$$\begin{bmatrix} -1 & -2 & -3 \\ 0 & 1 & -7 \end{bmatrix}$$

*Observación.* Vimos que $(-1)\boldsymbol{a} = -\boldsymbol{a}$. Del mismo modo, multiplicar una matriz por $-1$ cambia de signo todas sus columnas: $\big((-1)\mathbf{A}\big)_{|j} = (-1)\big(\mathbf{A}_{|j}\big) = -\big(\mathbf{A}_{|j}\big) = \big(-\mathbf{A}\big)_{|j}$; por tanto, $(-1)\mathbf{A} = -\mathbf{A}$. Y, como con los vectores, restar una matriz es sumar su opuesta: escribimos $\mathbf{A} - \mathbf{B}$ para denotar $\mathbf{A} + (-\mathbf{B})$.

La suma de matrices y el producto de matrices por escalares verifican propiedades análogas a las de la Proposición 1.2.1, y se demuestran del mismo modo que en el Ejercicio 4.

☞ Es la misma demostración, con vectores donde antes había números. Un vector es una lista de números y una matriz es una lista de vectores; por eso las propiedades de los vectores las demostramos seleccionando una componente (un número) con el selector "$|i$", y las de las matrices las demostraremos seleccionando una columna (un vector) con el selector "$|j$". Las reglas que se usan son las mismas en los dos casos:

| | **Vectores** | **Matrices** |
|---|---|---|
| selector distributivo respecto a la suma | $\big(\boldsymbol{a}+\boldsymbol{b}\big)_{\|i} = \boldsymbol{a}_{\|i} + \boldsymbol{b}_{\|i}$ | $\big(\mathbf{A}+\mathbf{B}\big)_{\|j} = \mathbf{A}_{\|j} + \mathbf{B}_{\|j}$ |
| y asociativo respecto al producto por escalares | $\big(\lambda\boldsymbol{a}\big)_{\|i} = \lambda\big(\boldsymbol{a}_{\|i}\big)$ | $\big(\lambda\mathbf{A}\big)_{\|j} = \lambda\big(\mathbf{A}_{\|j}\big)$ |
| nulo | $\mathbf{0}_{\|i} = 0$ toda componente es cero | $\mathbf{0}_{\|j} = \mathbf{0}$ toda columna es un vector nulo |
| opuesto | $\big(-\boldsymbol{a}\big)_{\|i} = -\big(\boldsymbol{a}_{\|i}\big)$ | $\big(-\mathbf{A}\big)_{\|j} = -\big(\mathbf{A}_{\|j}\big)$ |
| y después se aplican las propiedades de | los números reales | los vectores |

Las demostraciones de las ocho propiedades de los vectores solo usaron dos cosas: las cuatro reglas de la columna de la izquierda y las propiedades de los números reales. Las matrices cumplen las mismas cuatro reglas con el selector de columnas (columna de la derecha de la tabla), y lo que se selecciona son vectores, que verifican las ocho propiedades que ya hemos demostrado. Por eso las matrices tienen las mismas ocho propiedades, y no hace falta pensar de nuevo las demostraciones: basta repetir cada una sustituyendo cada regla de la columna de la izquierda por la de su derecha.

Esa es la ventaja del álgebra: lo que se demuestra usando solo unas reglas de manipulación vale para cualquier objeto que cumpla esas reglas. (En la Lección 6 daremos nombre a los conjuntos en los que se verifican estas ocho propiedades; véase la Nota 2.)

Ejercicio 8. Demuestre las propiedades de la siguiente proposición.

**Proposición 1.5.1** (Propiedades de las operaciones entre matrices)**.** *Para cualesquiera matrices* $\mathbf{A}$*,* $\mathbf{B}$ *y* $\mathbf{C}$

*de idéntico orden y para cualesquiera escalares $\lambda$ y $\varphi$, se verifica que:*

1. $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$.
2. $\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}$.
3. $\mathbf{A}+\mathbf{0}=\mathbf{A}$.
4. $\mathbf{A}+(-\mathbf{A})=\mathbf{0}$.
5. $\lambda(\mathbf{A}+\mathbf{B})=\lambda\mathbf{A}+\lambda\mathbf{B}$.
6. $(\lambda+\varphi)\mathbf{A}=\lambda\mathbf{A}+\varphi\mathbf{A}$.
7. $\lambda(\varphi\mathbf{A})=(\lambda\varphi)\mathbf{A}$.
8. $1\mathbf{A}=\mathbf{A}$.

*Pista.* Recuerde que dos matrices del mismo orden son iguales si lo son columna a columna. Escriba la columna $j$-ésima del lado izquierdo de la igualdad y construya a partir de ella una cadena de igualdades en la que cada paso aplica una sola regla: la linealidad del operador "$|j$" (es decir, que es distributivo respecto a la suma de matrices y asociativo respecto al producto por escalares), una definición (la de matriz nula o la de matriz opuesta) o una de las propiedades de las operaciones con vectores de la Proposición 1.2.1 (ya que $\mathbf{A}_{|j}$, $\mathbf{B}_{|j}$ y $\mathbf{C}_{|j}$ son vectores de $\mathbb{R}^m$). La cadena termina cuando llega a la columna $j$-ésima del lado derecho.

**Nota 2.** *Estas ocho propiedades (de manera similar a las ocho propiedades de las operaciones con vectores de la Proposición 1.2.1) nos permitirán definir en la Lección 6 el espacio vectorial de matrices $\mathbb{R}^{m\times n}$.*

Como con los vectores, la propiedad 2 permite escribir sin paréntesis la suma de varias matrices: $\mathbf{A}+\mathbf{B}+\mathbf{C}$.

### 1.5.3. Operaciones fila a fila

Hemos definido las operaciones por columnas. De las definiciones se deduce que también se puede operar por filas:

Ejercicio 9. Demuestre que las definiciones 1.17 y 1.18 implican que se puede operar "por filas", es decir:

$$ {}_{i|}(\mathbf{A}+\mathbf{B}) = {}_{i|}\mathbf{A} + {}_{i|}\mathbf{B}, \qquad \text{y} \qquad {}_{i|}(\lambda\mathbf{A}) = \lambda({}_{i|}\mathbf{A}); \qquad \text{donde } \mathbf{A} \text{ y } \mathbf{B} \text{ son de orden } m\times n, \text{ e } i = 1:m. $$

*Pista.* Las filas de una matriz son vectores, y dos vectores son iguales si lo son componente a componente. Seleccione la componente $j$-ésima de la fila del lado izquierdo y cambie el orden de selección para seleccionar primero la columna: ahí se aplican las definiciones de las operaciones con matrices.

☞ Por tanto, el operador selector de *filas*, "$i\,|$", también es **lineal**; es decir, *distributivo* respecto a la suma de matrices

$$ {}_{i|}(\mathbf{A}+\mathbf{B}) = {}_{i|}\mathbf{A} + {}_{i|}\mathbf{B} $$

y *asociativo* respecto al producto de una matriz por un *escalar*

$$ {}_{i|}(\lambda\mathbf{A}) = \lambda\big({}_{i|}\mathbf{A}\big). $$

### 1.5.4. La transposición es un operador lineal

Las columnas de $\mathbf{A}^\intercal$ son las filas de $\mathbf{A}$. Como ya sabemos que se puede operar por filas, podemos demostrar las propiedades de linealidad de la transposición.[18]

Ejercicio 10. Demuestre las siguientes proposiciones.

(a) **Proposición 1.5.2** (Transpuesta de un múltiplo)**.** *Sea* $\underset{m\times n}{\mathbf{A}}$ *entonces* $(\lambda\mathbf{A})^\intercal = \lambda(\mathbf{A}^\intercal)$.

(b) **Proposición 1.5.3** (Transpuesta de una suma)**.** *Sean* $\underset{m\times n}{\mathbf{A}}$ *y* $\underset{m\times n}{\mathbf{B}}$ *entonces* $(\mathbf{A}+\mathbf{B})^\intercal = \mathbf{A}^\intercal + \mathbf{B}^\intercal$.

[18] De la transposición ya solo nos falta demostrar que $(\mathbf{AB})^\intercal = (\mathbf{B}^\intercal)(\mathbf{A}^\intercal)$, pero para ello antes tenemos que ver las definiciones de producto de un vector por una matriz $\boldsymbol{a}\mathbf{B}$ y producto de matrices $\mathbf{AB}$.

☞ Así, el operador transposición es un **operador lineal**.

- Enlace al Notebook de Jupyter para trabajar con lo expuesto en esta Sección.

## 1.6. Extensión de la notación matricial y las reglas de reescritura

Con los vectores aceptamos que los escalares pudieran multiplicar también por la derecha (Sección 1.4.4). Con las matrices haremos lo mismo.

**Definición 1.21** (escalar a la derecha de una matriz)**.** *Aceptaremos que un escalar $\lambda$ aparezca multiplicando por la derecha de una matriz $\mathbf{A}$; la expresión $\mathbf{A}\lambda$ denota, por definición, el producto de $\mathbf{A}$ por $\lambda$:*

$$\boxed{\mathbf{A}\lambda = \lambda\mathbf{A}}$$

Esta definición no introduce una operación nueva: fija un convenio de notación, otra manera de escribir el mismo producto.[19] Sirve para lo mismo que con los vectores: para no perder la *asociatividad* de la notación cuando el selector opera por la izquierda. En efecto, con el escalar a la derecha, la linealidad del selector de filas también se puede escribir "desplazando los paréntesis":

$$_{i|}(\mathbf{A}\lambda) = {}_{i|}(\lambda\mathbf{A}) = \lambda({}_{i|}\mathbf{A}) = ({}_{i|}\mathbf{A})\lambda.$$

(El primer paso es el convenio que acabamos de adoptar; el segundo, que el selector de filas es asociativo respecto al producto por escalares; y el tercero, el convenio análogo para los vectores, pues ${}_{i|}\mathbf{A}$ es un vector.) Y lo mismo ocurre con las columnas:

$$(\mathbf{A}\lambda)_{|j} = (\lambda\mathbf{A})_{|j} = \lambda(\mathbf{A}_{|j}) = (\mathbf{A}_{|j})\lambda.$$

Así hemos demostrado que disponemos de reglas de reescritura similares a las que obtuvimos en el caso de los vectores.

☞ **Reglas de reescritura:** como **los dos selectores (el de columnas y el de filas) son lineales**, podemos

- "*distribuir* el operador" entre sumandos

$$(\mathbf{A}+\mathbf{B})_{|j} = \mathbf{A}_{|j} + \mathbf{B}_{|j} \qquad \text{y} \qquad {}_{i|}(\mathbf{A}+\mathbf{B}) = {}_{i|}\mathbf{A} + {}_{i|}\mathbf{B};$$

- "*desplazar los paréntesis*" para sacar un símbolo por un lado e introducir otro por el otro lado:

$$(\lambda\mathbf{A})_{|j} = \lambda(\mathbf{A}_{|j}) \qquad \text{y} \qquad ({}_{i|}\mathbf{A})\lambda = {}_{i|}(\mathbf{A}\lambda).$$

Y, además, también podemos

- "*intercambiar las posiciones del escalar y el selector*" dentro y fuera de los paréntesis:

$$(\mathbf{A}_{|j})\lambda = (\mathbf{A}\lambda)_{|j} \qquad \text{y} \qquad {}_{i|}(\lambda\mathbf{A}) = \lambda({}_{i|}\mathbf{A}).$$

En las dos últimas líneas del recuadro, las dos igualdades de la izquierda arrojan el mismo resultado (la columna $j$-ésima de $\mathbf{A}$ multiplicada por $\lambda$), y las dos de la derecha también (la fila $i$-ésima de $\mathbf{A}$ multiplicada

[19] Como en el caso de los vectores, tenga en cuenta que no es una escritura aceptada generalmente: el modo habitual de escribir este producto es "$2\mathbf{A}$", con el coeficiente primero.

por $\lambda$), pese al distinto orden de ejecución de las operaciones. Por eso podemos omitir los paréntesis y escribir tanto $\lambda\mathbf{A}_{|j}$ como ${}_{i|}\mathbf{A}\lambda$. En aras de la claridad, *en las demostraciones siempre usaremos los paréntesis para destacar qué regla o propiedad se aplica en cada paso*, pero en otras ocasiones es práctico omitirlos si no son necesarios (¡esa es la ventaja de la **notación asociativa**!).[20]

El mismo convenio, aplicado a la Proposición 1.5.2 ("transpuesta de un múltiplo"), permite demostrar otra regla de reescritura, que intercambia las posiciones del escalar y el símbolo de transposición:

$$(\mathbf{A}\lambda)^\intercal = (\lambda\mathbf{A})^\intercal = \lambda(\mathbf{A}^\intercal) = (\mathbf{A}^\intercal)\lambda.$$

(El primer paso y el último son el convenio, aplicado a $\mathbf{A}$ y a $\mathbf{A}^\intercal$; el segundo es la proposición.)

☞

**Reglas de reescritura:** como **el operador transposición es lineal**, podemos

- "*distribuir* el operador" en una suma

$$(\mathbf{A}+\mathbf{B})^\intercal = \mathbf{A}^\intercal + \mathbf{B}^\intercal;$$

- "*desplazar los paréntesis*" para sacar un símbolo por un lado e introducir otro por el otro lado:

$$(\lambda\mathbf{A})^\intercal = \lambda(\mathbf{A}^\intercal) \qquad \text{(y que podemos escribir sencillamente como } \lambda\mathbf{A}^\intercal).$$

Y, además, también podemos

- "*intercambiar las posiciones del escalar y el operador transposición*" dentro y fuera de los paréntesis:

$$(\mathbf{A}^\intercal)\lambda = (\mathbf{A}\lambda)^\intercal.$$

# *Apéndices a la lección*

## 1.A. Subvectores mediante selección de una lista de índices

Este apéndice no se necesita en esta lección. La selección con una lista de índices se usa a partir de la Lección 5 y se extiende a las matrices en el apéndice de la Lección 3 (Sección 3.B).

Por conveniencia, vamos a extender la notación para permitir que el operador selector seleccione las componentes indicadas en una lista de índices.

Sea $\boldsymbol{b}$ un vector de $\mathbb{R}^n$ y sea $\boldsymbol{\alpha}$ una lista de índices menores o iguales a $n$. Definimos $\boldsymbol{b}_{|\boldsymbol{\alpha}}$ como el vector (de tantas componentes como *elementos en* $\boldsymbol{\alpha}$) cuyas componentes son las indicadas en la lista $\boldsymbol{\alpha}$. Es decir,

$$\text{si } \boldsymbol{b}\in\mathbb{R}^n, \text{ entonces } \boldsymbol{b}_{|(i,j,\ldots,p,)} = (b_i, b_j, \ldots, b_p,); \quad \text{para } i,j,\ldots,p \le n.$$

Por ejemplo, si $\boldsymbol{b} = (7,0,\pi,\sqrt{2},)$ y $\boldsymbol{\alpha} = (3,3,1,)$, entonces $\boldsymbol{b}_{|\boldsymbol{\alpha}}$ es el vector de $\mathbb{R}^3$ cuyas dos primeras componentes son iguales a la tercera de $\boldsymbol{b}$ y su última componente es igual a la primera componente de $\boldsymbol{b}$, es decir

$$\boldsymbol{b}_{|\boldsymbol{\alpha}} = (\pi,\pi,7,).$$

Evidentemente, para todo vector $\boldsymbol{b}\in\mathbb{R}^n$ se verifica que $\boldsymbol{b} = \boldsymbol{b}_{|(1:n)}$.

[20] Aunque debe recordar que, dada la precedencia establecida en la ejecución de las operaciones en Python, programando con el ordenador siempre es conveniente poner los paréntesis.

**Hiperenlaces a los Notebooks de Jupyter en mybinder.org correspondientes a esta lección.**

- Notebook Sección 1.1
- Notebook Secciones 1.2 y 1.3
- Notebook Sección 1.4
- Notebook Sección 1.5

Todos los Notebooks de Jupyter están enlazados desde la página:

- https://mbujosab.github.io/nacal-Jupyter-Notebooks/

# Combinaciones lineales

En esta lección se definen tres productos: el producto punto de dos vectores, el producto de una matriz por un vector y el producto de un vector por una matriz. Los dos últimos son combinaciones lineales de las columnas y de las filas de la matriz. El producto punto va primero porque con él se escriben las componentes de los otros dos.

## 2.1. Producto punto (o producto escalar usual en $\mathbb{R}^n$)

Definamos el producto punto entre dos vectores de $\mathbb{R}^n$:

**Definición 2.1.** *El producto punto* $\boldsymbol{a} \cdot \boldsymbol{b}$ (o producto escalar usual en $\mathbb{R}^n$) *de dos vectores* $\boldsymbol{a}$ *y* $\boldsymbol{b}$ *es*[1]

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1b_1 + a_2b_2 + a_3b_3 + \cdots + a_nb_n \;=\; \sum_{i=1}^{n} a_ib_i \;=\; \sum_{i=1}^{n} \boldsymbol{a}_{|i}\boldsymbol{b}_{|i}.$$

*Ejemplo* 1.

$$\begin{pmatrix}1\\2\\3\end{pmatrix} \cdot \begin{pmatrix}1\\-1\\2\end{pmatrix} = (1\cdot 1) + (2\cdot(-1)) + (3\cdot 2) = (1) + (-2) + (6) = 5$$

Módulo NAcAL para Python

```
a = Vector( [1, 2, 3] )
b = Vector( [1,-1, 2] )
a * b          # producto punto entre los vectores a y b
```

5

Este producto será importante al definir la ortogonalidad en $\mathbb{R}^n$ en la parte dedicada a la ortogonalidad. No obstante, lo introducimos en esta lección para poder aprovechar algunas propiedades de la notación. Con ellas podremos escribir demostraciones mucho más "limpias" (utilizando productos punto en lugar de sumatorios) y con pasos más sencillos (empleando repetidamente la linealidad tanto del producto punto como del operador selector).

### Propiedades del producto punto

El producto punto (o producto escalar usual de $\mathbb{R}^n$) tiene las cuatro propiedades del siguiente ejercicio. En la Lección 11 serán los axiomas que definen un producto escalar.

Ejercicio 11. Demuestre que, para cualesquiera vectores $\boldsymbol{x}$, $\boldsymbol{y}$ y $\boldsymbol{z}$ de $\mathbb{R}^n$ y para cualquier escalar $\lambda$, el producto punto cumple las siguientes propiedades:

(a) **Simetría:** $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{y} \cdot \boldsymbol{x}$

[1] El símbolo $\sum$ abrevia una suma: $\sum_{i=1}^{n} a_ib_i$ significa $a_1b_1 + a_2b_2 + \cdots + a_nb_n$.

(b) **Linealidad respecto al primer argumento:**
   1. $(\lambda \boldsymbol{x}) \cdot \boldsymbol{y} = \lambda(\boldsymbol{x} \cdot \boldsymbol{y})$
   2. $(\boldsymbol{x} + \boldsymbol{y}) \cdot \boldsymbol{z} = \boldsymbol{x} \cdot \boldsymbol{z} + \boldsymbol{y} \cdot \boldsymbol{z}$

(c) **Positivo:** $\boldsymbol{x} \cdot \boldsymbol{x} \geq 0$

(d) **Definido:** $\boldsymbol{x} \cdot \boldsymbol{x} = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$.

Una operación con dos argumentos es lineal en uno de ellos si, con el otro fijo, es distributiva respecto a la suma y asociativa respecto al producto por escalares. Fíjese que, dado que el producto punto es simétrico, también es lineal en el segundo argumento:

$$\begin{aligned} \boldsymbol{x} \cdot (\lambda \boldsymbol{y}) = (\lambda \boldsymbol{y}) \cdot \boldsymbol{x} &= \lambda(\boldsymbol{y} \cdot \boldsymbol{x}) = \lambda(\boldsymbol{x} \cdot \boldsymbol{y}); \\ \boldsymbol{x} \cdot (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{y} + \boldsymbol{z}) \cdot \boldsymbol{x} &= \boldsymbol{y} \cdot \boldsymbol{x} + \boldsymbol{z} \cdot \boldsymbol{x} = \boldsymbol{x} \cdot \boldsymbol{y} + \boldsymbol{x} \cdot \boldsymbol{z} \end{aligned}$$

Por tanto, el producto punto es lineal en sus dos argumentos; y en consecuencia $\boldsymbol{x} \cdot (\lambda \boldsymbol{y}) = (\lambda \boldsymbol{x}) \cdot \boldsymbol{y}$.

**Caso especial 1** (Producto punto de un vector por una fila o columna de $\mathbf{I}$)**.** *Al multiplicar un vector de $\mathbb{R}^n$ por la fila (o columna) $i$-ésima de la matriz identidad de orden $n$, se selecciona la componente $i$-ésima del vector,* $\left({}_{i|}\mathbf{I}\right) \cdot \boldsymbol{x} = {}_{i|}\boldsymbol{x} = \boldsymbol{x}_{|i} = \boldsymbol{x} \cdot \left(\mathbf{I}_{|i}\right)$ :

$$\left({}_{i|}\mathbf{I}\right) \cdot \boldsymbol{x} = 0x_1 + 0x_2 + \cdots + 1x_i + \cdots + 0x_n = x_i = \boldsymbol{x} \cdot \left(\mathbf{I}_{|i}\right).$$

## 2.2. Producto de una matriz por un vector (a su derecha)

### 2.2.1. Combinación lineal de vectores

Hay una operación *muy importante* (¡la más importante de todas las que veremos en este curso!): la suma de múltiplos de vectores. Por ejemplo

$$3\boldsymbol{a} + \boldsymbol{b} - 7\boldsymbol{c} + 2\boldsymbol{d},$$

donde 3, 1, $-7$ y 2 son los coeficientes. Es una operación tan importante que tiene nombre propio:

**Definición 2.2** (Combinación lineal de vectores de $\mathbb{R}^n$)**.** *Sean $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_p$ vectores de $\mathbb{R}^n$. Llamamos combinación lineal a cualquier suma de múltiplos de dichos vectores:*

$$a_1\boldsymbol{b}_1 + a_2\boldsymbol{b}_2 + \cdots + a_p\boldsymbol{b}_p,$$

*donde los números "$a_i$" son los coeficientes de la combinación lineal.*

Ejercicio 12. Sean los vectores

$$\boldsymbol{x} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}; \quad \boldsymbol{y} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}; \quad \boldsymbol{z} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}.$$

Compruebe que la combinación lineal $2\boldsymbol{x} + 2\boldsymbol{y} + 2\boldsymbol{z}$ es el vector nulo.

*Ejemplo* 2. Nótese que todo vector $\boldsymbol{a}$ de $\mathbb{R}^n$ es combinación lineal de las columnas de $\underset{n \times n}{\mathbf{I}}$, pues

$$\left(\mathbf{I}_{|1}\right)a_1 + \cdots + \left(\mathbf{I}_{|n}\right)a_n = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} a_1 + \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} a_2 + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} a_n = \begin{pmatrix} a_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_2 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \boldsymbol{a}.$$

### 2.2.2. Producto de una matriz por un vector

El producto de una matriz por un vector a su derecha, $\mathbf{A}\boldsymbol{b}$, es una combinación lineal de las columnas de la matriz. Para recordarlo, *escribiremos el vector de la derecha en forma de columna*. Llamaremos a esta operación *combinación lineal de las columnas de* $\mathbf{A}$ o *producto de una matriz por un vector a su derecha*.

**Definición 2.3** (matriz por un vector a su derecha)**.** *El producto de* $\underset{m\times n}{\mathbf{A}}$ *por un vector* $\boldsymbol{b}\in\mathbb{R}^n$ *a su derecha es la combinación lineal de las* $n$ *columnas de* $\mathbf{A}$ *cuyos coeficientes son las componentes de* $\boldsymbol{b}$*:*

$$\mathbf{A}\boldsymbol{b}=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \dots & a_{mn}\end{bmatrix}\begin{pmatrix} b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}\equiv\begin{pmatrix} a_{11}\\ a_{21}\\ \vdots\\ a_{m1}\end{pmatrix}b_1+\begin{pmatrix} a_{12}\\ a_{22}\\ \vdots\\ a_{m2}\end{pmatrix}b_2+\cdots+\begin{pmatrix} a_{1n}\\ a_{2n}\\ \vdots\\ a_{mn}\end{pmatrix}b_n. \tag{2.1}$$

O expresado de forma más compacta: $\mathbf{A}\boldsymbol{b}=\left(\mathbf{A}_{|1}\right)b_1+\left(\mathbf{A}_{|2}\right)b_2+\cdots+\left(\mathbf{A}_{|n}\right)b_n$.

Escribimos los coeficientes a la derecha de las columnas para conservar el orden de los factores de $\mathbf{A}\boldsymbol{b}$. Como $\boldsymbol{b}\lambda=\lambda\boldsymbol{b}$, es lo mismo que $b_1\left(\mathbf{A}_{|1}\right)+b_2\left(\mathbf{A}_{|2}\right)+\cdots+b_n\left(\mathbf{A}_{|n}\right)$.

Fíjese en que el número de columnas de la matriz $\mathbf{A}$ *debe ser igual* al número de componentes del vector $\boldsymbol{b}$ (¡una componente de $\boldsymbol{b}$ por cada columna de $\mathbf{A}$ para poder expresar la suma de múltiplos de las columnas!)

Módulo NAcAL para Python

```
A = Matrix( [ Vector([1,1,1]), Vector([1,0,1]), Vector([-2,2,1]) ] )
b = Vector( [1,2,-3] )
A*b
```

$$\mathbf{A}\boldsymbol{b}=\begin{bmatrix} 1 & 1 & -2\\ 1 & 0 & 2\\ 1 & 1 & 1\end{bmatrix}\begin{pmatrix} 1\\ 2\\ -3\end{pmatrix}=1\begin{pmatrix} 1\\ 1\\ 1\end{pmatrix}+2\begin{pmatrix} 1\\ 0\\ 1\end{pmatrix}-3\begin{pmatrix} -2\\ 2\\ 1\end{pmatrix}=\begin{pmatrix} 9\\ -5\\ 0\end{pmatrix}.$$

Ejercicio 13. Escriba como producto de una matriz por un vector la combinación lineal del ejercicio 12.

Y ahora veamos un par de casos especialmente sencillos:

**Caso especial 2** (Producto de una matriz identidad por un vector)**.** *Ahora podemos expresar el producto del Ejemplo 2 en la página anterior de manera mucho más compacta*

$$\mathbf{I}\boldsymbol{a}=\boldsymbol{a}.$$

(es decir, todo vector de $\mathbb{R}^n$ es combinación lineal de las columnas de la matriz identidad de orden $n$)

**Caso especial 3** (Producto de una matriz por una columna de $\mathbf{I}$)**.** *Por otra parte, de la Definición 2.3 se deduce que*

$$\mathbf{A}\left(\mathbf{I}_{|j}\right)=\mathbf{A}_{|j}$$

*es decir, multiplicar* $\mathbf{A}$ *por la columna* $j$*-ésima de la identidad es lo mismo que seleccionar su columna* $j$*-ésima. Lo demostrará en el ejercicio siguiente.*

Ejercicio 14. Demuestre que $\mathbf{A}\left(\mathbf{I}_{|j}\right)=\mathbf{A}_{|j}$.

Pues bien, hay una estrecha relación entre el producto $\mathbf{A}\boldsymbol{b}$ y los productos punto de $\boldsymbol{b}$ y las filas de $\mathbf{A}$:

Ejercicio 15. Demuestre que la componente $i$-ésima del vector $\mathbf{A}\boldsymbol{b}$ es el producto punto $\left({}_{i|}\mathbf{A}\right)\cdot\boldsymbol{b}$:

$$\boxed{{}_{i|}\left(\mathbf{A}\boldsymbol{b}\right)=\left({}_{i|}\mathbf{A}\right)\cdot\boldsymbol{b}\qquad\text{donde } i=1:m.}$$

Si $\mathbf{A}$ tiene una sola fila, $\mathbf{A}\boldsymbol{b}$ es un vector de $\mathbb{R}^1$: $\big(\big({}_{1|}\mathbf{A}\big)\cdot\boldsymbol{b},\big)$. No es el número $\big({}_{1|}\mathbf{A}\big)\cdot\boldsymbol{b}$.

Módulo NAcAL para Python

```
      2|(A*b) == (2|A)*b    # ¿son iguales?
```

```
True
```

☞ Hemos visto que la operación **MATRIZ** (de orden $m\times n$) por un ***vector*** (de $n$ componentes) es un *vector* formado por una suma de múltiplos de las $n$ columnas:

$$\mathbf{A}\boldsymbol{b} \;=\; (\mathbf{A}_{|1})b_1 + (\mathbf{A}_{|2})b_2 + \cdots + (\mathbf{A}_{|n})b_n \quad \in \mathbb{R}^m.$$

A un vector que es suma de múltiplos de vectores se le llama combinación lineal. Por tanto $\mathbf{A}\boldsymbol{b}$ es una *combinación lineal de las columnas* de $\mathbf{A}$.

También hemos visto que ${}_{i|}(\mathbf{A}\boldsymbol{b}) = ({}_{i|}\mathbf{A})\cdot\boldsymbol{b}$, que $\mathbf{I}\boldsymbol{a}=\boldsymbol{a}$ y que $\mathbf{A}(\mathbf{I}_{|j}) = \mathbf{A}_{|j}$.

### 2.2.3. Propiedades del producto de una matriz por un vector a su derecha

El producto de una matriz por un vector posee dos importantes propiedades: las *propiedades de linealidad en el vector*. La primera respecto a la suma: $\mathbf{A}(\boldsymbol{b}+\boldsymbol{c}) = \mathbf{A}\boldsymbol{b}+\mathbf{A}\boldsymbol{c}$, y la segunda respecto al producto por escalares: $\mathbf{A}(\lambda\boldsymbol{b}) = \lambda(\mathbf{A}\boldsymbol{b})$. Veámoslo.

Ejercicio 16. Demuestre las siguientes proposiciones (inténtelo primero sin mirar las pistas).

(a) **Proposición 2.2.1.** *Sea la matriz* $\underset{m\times n}{\mathbf{A}}$ *y los vectores* $\boldsymbol{b}$ *y* $\boldsymbol{c}$ *de* $\mathbb{R}^n$*, entonces:* $\mathbf{A}(\boldsymbol{b}+\boldsymbol{c}) = \mathbf{A}\boldsymbol{b}+\mathbf{A}\boldsymbol{c}$.

*Pista.* Demuestre que ambos vectores son iguales componente a componente, es decir, que ${}_{i|}(\mathbf{A}(\boldsymbol{b}+\boldsymbol{c}))$ es igual a ${}_{i|}(\mathbf{A}\boldsymbol{b}+\mathbf{A}\boldsymbol{c})$. Para ello, emplee que cada componente es un producto punto y que los productos punto también son lineales en el segundo argumento. Use también que el operador selector es lineal.

(b) **Proposición 2.2.2.** *Sea la matriz* $\underset{m\times n}{\mathbf{A}}$*, el vector* $\boldsymbol{b}$ *de* $\mathbb{R}^n$ *y el escalar* $\lambda$*, entonces:* $\mathbf{A}(\lambda\boldsymbol{b}) = \lambda(\mathbf{A}\boldsymbol{b})$.

*Pista.* Siga la misma estrategia que en el apartado anterior.

☞ El producto de una matriz por un vector es un **operador lineal** en el vector; es decir, *distributivo* respecto a la suma

$$\mathbf{A}(\boldsymbol{b}+\boldsymbol{c}) = \mathbf{A}\boldsymbol{b}+\mathbf{A}\boldsymbol{c}.$$

y *asociativo* respecto al producto por escalares

$$\mathbf{A}(\lambda\boldsymbol{b}) = \lambda(\mathbf{A}\boldsymbol{b}).$$

Con el escalar a la derecha, la Proposición 2.2.2 se escribe "desplazando los paréntesis":

$$\boxed{\mathbf{A}(\boldsymbol{b}\lambda) = (\mathbf{A}\boldsymbol{b})\lambda}$$

En efecto, $\mathbf{A}(\boldsymbol{b}\lambda) = \mathbf{A}(\lambda\boldsymbol{b}) = \lambda(\mathbf{A}\boldsymbol{b}) = (\mathbf{A}\boldsymbol{b})\lambda$. El primer paso y el tercero son el convenio de escribir el escalar a la derecha de un vector; el segundo es la Proposición 2.2.2.

Ejercicio 17. Demuestre el siguiente corolario (inténtelo primero sin mirar las pistas).

**Corolario 2.2.3.** *Sea la matriz* $\underset{p\times m}{\mathbf{A}}$ *, la matriz* $\underset{m\times n}{\mathbf{B}}$ *, y el vector* $\boldsymbol{c}$ *de* $\mathbb{R}^n$*, entonces:*

$$\mathbf{A}(\mathbf{B}\boldsymbol{c}) = \Big[\mathbf{A}(\mathbf{B}_{|1}); \quad \dots \quad \mathbf{A}(\mathbf{B}_{|n});\Big]\,\boldsymbol{c}.$$

*Pista.* Comience con la expresión de la izquierda. Exprese el producto dentro de paréntesis como una combinación lineal de columnas (Definición 2.3) y aplique las propiedades de linealidad (proposiciones 2.2.1 y 2.2.2) y de nuevo la Definición 2.3 para obtener la expresión de la derecha de la igualdad.

Para terminar la sección, añadamos dos proposiciones más; la primera se parece a la Proposición 2.2.2, pero no es igual (fíjese en la posición de los paréntesis del lado derecho de las igualdades).

Ejercicio 18. Demuestre las siguientes proposiciones (inténtelo sin mirar las pistas).

(a) **Proposición 2.2.4.** *Sea la matriz* $\underset{m\times n}{\mathbf{A}}$ *, el vector* $\boldsymbol{b}$ *de* $\mathbb{R}^n$ *y el escalar* $\lambda$*, entonces:* $\mathbf{A}(\lambda\boldsymbol{b}) = (\lambda\mathbf{A})\boldsymbol{b}$.

*Pista.* Recuerde que, como el producto punto es lineal en sus dos argumentos, $\boldsymbol{x}\cdot(\lambda\boldsymbol{y}) = (\lambda\boldsymbol{x})\cdot\boldsymbol{y}$.

(b) **Proposición 2.2.5.** *Sean las matrices* $\underset{m\times n}{\mathbf{A}}$ *y* $\underset{m\times n}{\mathbf{B}}$ *y el vector* $\boldsymbol{c}$ *de* $\mathbb{R}^n$*, entonces:* $(\mathbf{A}+\mathbf{B})\boldsymbol{c} = \mathbf{A}\boldsymbol{c} + \mathbf{B}\boldsymbol{c}$.

**Más reglas de reescritura.**

De las proposiciones 2.2.2 y 2.2.4 se deduce que

$$(\lambda\mathbf{A})\boldsymbol{b} = \mathbf{A}(\lambda\boldsymbol{b}) = \lambda(\mathbf{A}\boldsymbol{b})$$

y, por tanto, no son necesarios los paréntesis. Así que, sin ninguna ambigüedad, podemos escribir:

$$\lambda\mathbf{A}\boldsymbol{b} = \mathbf{A}\lambda\boldsymbol{b}.$$

Observe, además, que (como aceptamos que el escalar pueda multiplicar a los vectores y matrices por la derecha) el escalar puede aparecer en cualquier lugar; lo único que se mantiene invariante es la posición relativa del vector respecto de la matriz, pues siempre está a su derecha: $\lambda\mathbf{A}\boldsymbol{b} = \mathbf{A}\lambda\boldsymbol{b} = \mathbf{A}\boldsymbol{b}\lambda$.

☞ Hemos visto varias propiedades que cumple la operación **MATRIZ** por ***vector***

- **Lineal en el vector** $\begin{cases}\mathbf{A}(\boldsymbol{b}+\boldsymbol{c}) = \mathbf{A}\boldsymbol{b} + \mathbf{A}\boldsymbol{c}\\ \\ \mathbf{A}(\lambda\boldsymbol{b}) = \lambda(\mathbf{A}\boldsymbol{b})\end{cases}$
- **Lineal en la matriz** $\begin{cases}(\mathbf{A}+\mathbf{B})\boldsymbol{c} = \mathbf{A}\boldsymbol{c} + \mathbf{B}\boldsymbol{c}\\ \\ (\lambda\mathbf{A})\boldsymbol{b} = \lambda(\mathbf{A}\boldsymbol{b})\end{cases}$
- $\mathbf{A}(\mathbf{B}\boldsymbol{c}) = \Big[\mathbf{A}(\mathbf{B}_{|1}); \quad \dots \quad \mathbf{A}(\mathbf{B}_{|n});\Big]\,\boldsymbol{c}$

## 2.3. Producto de un vector por una matriz (vector a la izquierda)

☞ Hasta aquí hemos visto que una matriz por un vector ($\mathbf{A}\boldsymbol{b}$) es una combinación lineal de las *columnas*.

A continuación, veremos que un vector por una matriz ($\boldsymbol{a}\mathbf{B}$) es una combinación lineal de las *filas*

**Definición 2.4** (producto de un vector por una matriz)**.** *El producto de un vector* $\boldsymbol{a}\in\mathbb{R}^m$ *por una matriz* $\mathbf{B}$ *de orden* $m\times n$ *a su derecha es la combinación lineal de las* $m$ *filas de* $\mathbf{B}$ *cuyos coeficientes son las componentes de* $\boldsymbol{a}$*:*

$$\boldsymbol{a}\mathbf{B} = a_1\big({}_{1|}\mathbf{B}\big) + \cdots + a_m\big({}_{m|}\mathbf{B}\big).$$

*Las filas de* $\mathbf{B}$ *son vectores de* $\mathbb{R}^n$*. Por tanto, el producto* $\boldsymbol{a}\mathbf{B}$ *es un vector de* $\mathbb{R}^n$*.*

Para recordar que operamos sobre las filas de $\mathbf{B}$, usualmente escribiremos $\boldsymbol{a}$ en horizontal:

$$\boldsymbol{a}\mathbf{B} = \begin{pmatrix} a_1, & a_2, & \dots & a_m, \end{pmatrix} \begin{bmatrix} b_{11} & \dots & b_{1n} \\ b_{21} & \dots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{bmatrix} = a_1 \begin{pmatrix} {}_{1\vert}\mathbf{B} \end{pmatrix} + \cdots + a_m \begin{pmatrix} {}_{m\vert}\mathbf{B} \end{pmatrix} \quad \in \mathbb{R}^n.$$

Módulo NAcAL para Python

```
b*A
```

$$\begin{pmatrix} 0, & -2, & -1, \end{pmatrix}$$

$$\boldsymbol{b}\mathbf{A} = \begin{pmatrix} 1, & 2, & -3, \end{pmatrix} \begin{bmatrix} 1 & 1 & -2 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} = 1\begin{pmatrix} 1, & 1, & -2, \end{pmatrix} + 2\begin{pmatrix} 1, & 0, & 2, \end{pmatrix} - 3\begin{pmatrix} 1, & 1, & 1, \end{pmatrix} = \begin{pmatrix} 0, & -2, & -1, \end{pmatrix}.$$

Las filas de $\mathbf{B}$ son las columnas de $\mathbf{B}^\intercal$. Por eso el producto de un vector por una matriz es igual a un producto de una matriz por un vector a su derecha:

$$\boxed{\boldsymbol{a}\mathbf{B} = (\mathbf{B}^\intercal)\boldsymbol{a}}$$

Se demuestra con las dos definiciones:

$$\begin{aligned} (\mathbf{B}^\intercal)\boldsymbol{a} &= \left((\mathbf{B}^\intercal)_{\vert 1}\right)a_1 + \cdots + \left((\mathbf{B}^\intercal)_{\vert m}\right)a_m \\ &= \begin{pmatrix} {}_{1\vert}\mathbf{B} \end{pmatrix}a_1 + \cdots + \begin{pmatrix} {}_{m\vert}\mathbf{B} \end{pmatrix}a_m \\ &= a_1\begin{pmatrix} {}_{1\vert}\mathbf{B} \end{pmatrix} + \cdots + a_m\begin{pmatrix} {}_{m\vert}\mathbf{B} \end{pmatrix} \;=\; \boldsymbol{a}\mathbf{B}. \end{aligned}$$

El primer paso es la Definición 2.3; el segundo, la definición de matriz transpuesta; el tercero, que el escalar se puede escribir a la derecha de un vector (Lección 1); y el cuarto, la Definición 2.4. Con esta igualdad, las propiedades del producto $\boldsymbol{a}\mathbf{B}$ se deducen de las propiedades ya demostradas del producto $\mathbf{A}\boldsymbol{b}$.

Cada componente $j$-ésima del vector $\boldsymbol{a}\mathbf{B}$ es el producto punto entre $\boldsymbol{a}$ y la columna $j$-ésima de $\mathbf{B}$:

$$\left(\boldsymbol{a}\mathbf{B}\right)_{\vert j} = \left((\mathbf{B}^\intercal)\boldsymbol{a}\right)_{\vert j} = {}_{j\vert}\left((\mathbf{B}^\intercal)\boldsymbol{a}\right) = \left({}_{j\vert}(\mathbf{B}^\intercal)\right)\cdot\boldsymbol{a} = \left(\mathbf{B}_{\vert j}\right)\cdot\boldsymbol{a} = \boldsymbol{a}\cdot\left(\mathbf{B}_{\vert j}\right), \quad \text{donde } j = 1:n.$$

El primer paso es la igualdad anterior; el segundo, que las componentes de un vector se pueden seleccionar por los dos lados; el tercero, el Ejercicio 15; el cuarto, que la fila $j$-ésima de $\mathbf{B}^\intercal$ es la columna $j$-ésima de $\mathbf{B}$ (Lección 1); y el quinto, que el producto punto es simétrico.

Módulo NAcAL para Python

```
print( (b*A)|3 == b*(A|3) )
b*(A|3)
```

```
True
```

$$-1$$

### 2.3.1. Propiedades del producto de un vector por una matriz

Ejercicio 19. Demuestre las siguientes propiedades del producto $\boldsymbol{a}\mathbf{B}$. (En cada solución, el paso marcado con $(*)$ aplica la proposición que cita la pista.)

(a) Sea el vector $\boldsymbol{a}$ de $m$ componentes y la matriz $\underset{m\times m}{\mathbf{I}}$, entonces: $\boxed{\boldsymbol{a}\mathbf{I} = \boldsymbol{a}.}$

*Pista.* Recuerde que $\mathbf{I}$ es simétrica: $\mathbf{I}^\intercal = \mathbf{I}$.

(b) Sea $\big({}_{i|}\mathbf{I}\big)$ la fila $i$-ésima de $\underset{m\times m}{\mathbf{I}}$ y la matriz $\underset{m\times n}{\mathbf{A}}$, entonces: $\boxed{\big({}_{i|}\mathbf{I}\big)\mathbf{A} = {}_{i|}\mathbf{A}.}$

*Pista.* Recuerde el caso especial 3 en la página 29 y que como $\mathbf{I}$ es simétrica, $\mathbf{I}_{|i} = {}_{i|}\mathbf{I}$.

(c) Sean $\boldsymbol{a}$ y $\boldsymbol{b}$ vectores de $\mathbb{R}^m$ y sea la matriz $\underset{m\times n}{\mathbf{C}}$, entonces: $\boxed{(\boldsymbol{a}+\boldsymbol{b})\mathbf{C} = \boldsymbol{a}\mathbf{C} + \boldsymbol{b}\mathbf{C}}$

*Pista.* Recuerde la Proposición 2.2.1 y que $\boldsymbol{v}\mathbf{A} = (\mathbf{A}^\intercal)\boldsymbol{v}$.

(d) Sea el escalar $\lambda$, el vector $\boldsymbol{a}$ de $\mathbb{R}^m$ y sea la matriz $\underset{m\times n}{\mathbf{B}}$, entonces: $\boxed{(\lambda\boldsymbol{a})\mathbf{B} = \lambda(\boldsymbol{a}\mathbf{B})}$

*Pista.* Recuerde la Proposición 2.2.2

(e) Sean $\boldsymbol{a}$ un vector de $\mathbb{R}^m$ y $\lambda$ un escalar; y sea la matriz $\underset{m\times n}{\mathbf{B}}$, entonces: $\boxed{\boldsymbol{a}(\lambda\mathbf{B}) = (\lambda\boldsymbol{a})\mathbf{B}.}$

*Pista.* Recuerde la Proposición 2.2.4

(f) Sean $\boldsymbol{a}$ un vector de $\mathbb{R}^m$, y las matrices $\mathbf{B}$ y $\mathbf{C}$ de orden $m$ por $n$, entonces: $\boxed{\boldsymbol{a}(\mathbf{B}+\mathbf{C}) = \boldsymbol{a}\mathbf{B} + \boldsymbol{a}\mathbf{C}.}$

*Pista.* Recuerde la Proposición 2.2.5.

De los apartados (d) y (e) se deduce que $\boldsymbol{a}(\lambda\mathbf{B}) = (\lambda\boldsymbol{a})\mathbf{B} = \lambda(\boldsymbol{a}\mathbf{B})$.

☞ El producto de un ***vector*** (de $m$ componentes) por una **MATRIZ** (de orden $m\times n$) es una combinación lineal de las $m$ filas de $\mathbf{B}$:

$$\boldsymbol{a}\mathbf{B} \;=\; a_1\big({}_{1|}\mathbf{B}\big) + \cdots + a_m\big({}_{m|}\mathbf{B}\big) \quad \in \mathbb{R}^n.$$

Cada componente del vector $\boldsymbol{a}\mathbf{B}$ se puede expresar como un producto punto:

$$\big(\boldsymbol{a}\mathbf{B}\big)_{|j} = \boldsymbol{a}\cdot\big(\mathbf{B}_{|j}\big), \quad \text{donde } j = 1:n.$$

El producto $\boldsymbol{a}\mathbf{B}$ tiene propiedades análogas a las del producto $\mathbf{A}\boldsymbol{b}$.

- $\boldsymbol{a}\mathbf{I} = \boldsymbol{a}$
- $\big({}_{i|}\mathbf{I}\big)\mathbf{A} = {}_{i|}\mathbf{A}$;

**Propiedades de linealidad en el vector**

- $(\boldsymbol{a}+\boldsymbol{b})\mathbf{C} = \boldsymbol{a}\mathbf{C} + \boldsymbol{b}\mathbf{C}$
- $(\lambda\boldsymbol{a})\mathbf{B} = \lambda(\boldsymbol{a}\mathbf{B})$

**Propiedades de linealidad en la matriz**

- $\boldsymbol{a}(\mathbf{B}+\mathbf{C}) = \boldsymbol{a}\mathbf{B} + \boldsymbol{a}\mathbf{C}$
- $\boldsymbol{a}(\lambda\mathbf{B}) = \lambda(\boldsymbol{a}\mathbf{B})$

# Multiplicación matricial

## 3.1. Producto de matrices

**Definición 3.1** (producto de matrices (por **columnas**))**.** *Sean* $\underset{m\times p}{\mathbf{A}}$ *y* $\underset{p\times n}{\mathbf{B}}$ *, definimos la matriz producto,* $\mathbf{AB}$*, como aquella matriz cuya columna* $j$*-ésima es el producto de* $\mathbf{A}$ *por la columna* $j$*-ésima de* $\mathbf{B}$*:*

$$\boxed{(\mathbf{AB})_{|j} = \mathbf{A}(\mathbf{B}_{|j})}.$$

Es decir, las columnas de $\mathbf{AB}$ son combinaciones lineales de las columnas de $\mathbf{A}$ (y los coeficientes de cada combinación son los componentes de cada una de las columnas de $\mathbf{B}$).

$$\mathbf{AB} \equiv \left[\mathbf{A}(\mathbf{B}_{|1}); \quad \dots \quad \mathbf{A}(\mathbf{B}_{|n});\right].$$

*Ejemplo* 3. Si $\mathbf{A} = \begin{bmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 1 \end{bmatrix}$ y $\mathbf{B} = \begin{bmatrix} -2 & 1 \\ 1 & 1 \\ 1 & -2 \end{bmatrix}$ entonces

$$\mathbf{AB} = \left[\mathbf{A}(\mathbf{B}_{|1}); \quad \mathbf{A}(\mathbf{B}_{|2});\right] = \left[\mathbf{A}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}; \quad \mathbf{A}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix};\right] = \left[\begin{pmatrix} -3 \\ -3 \\ 6 \end{pmatrix}; \quad \begin{pmatrix} 6 \\ -3 \\ -3 \end{pmatrix};\right] = \begin{bmatrix} -3 & 6 \\ -3 & -3 \\ 6 & -3 \end{bmatrix};$$

nótese que calculamos multiplicando $\mathbf{A}$ por cada columna de $\mathbf{B}$.[1]

Deduzcamos algunas propiedades que verifica el producto de matrices...

Ejercicio 20. Demuestre las siguientes proposiciones.

(a) **Proposición 3.1.1.** *Sean* $\underset{m\times p}{\mathbf{A}}$ *y* $\underset{p\times n}{\mathbf{B}}$ *, y un vector* $\boldsymbol{c}$ *de* $\mathbb{R}^n$*, entonces* $\mathbf{A}(\mathbf{B}\boldsymbol{c}) = (\mathbf{AB})\boldsymbol{c}$.

(b) **Proposición 3.1.2** (El producto de matrices es asociativo)**.** *Sean las matrices* $\underset{m\times p}{\mathbf{A}}$ *,* $\underset{p\times q}{\mathbf{B}}$ *y* $\underset{q\times n}{\mathbf{C}}$ *, entonces*

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}.$$

**Nota importante respecto al operador " $|\,j$ ", el producto de matrices y la asociatividad:** La definición de producto matricial mantiene la *asociatividad* del operador " $|\,j$ ", POR LO QUE NO NECESITAMOS PARÉNTESIS. Así, $\mathbf{AB}_{|j}$ se refiere indistintamente a la columna $j$-ésima del producto $(\mathbf{AB})_{|j}$, o al producto de $\mathbf{A}$ por $\mathbf{B}_{|j}$ (i.e., la columna $j$-ésima), $\mathbf{A}(\mathbf{B}_{|j})$. Y puesto que $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$, podemos escribir simplemente $\mathbf{ABC}$.

[1] Aunque conozca otro método para calcular el producto, no lo emplee por el momento. AHORA ES MUY IMPORTANTE que se acostumbre a considerar las columnas de $\mathbf{AB}$ como combinaciones de las columnas de $\mathbf{A}$.

☞ Hemos definido el producto $(\mathbf{AB})$ de las matrices $\underset{m\times p}{\mathbf{A}}$ y $\underset{p\times n}{\mathbf{B}}$ como aquella matriz cuya columna $j$-ésima es el producto de $\mathbf{A}$ por la columna $j$-ésima de $\mathbf{B}$

$$(\mathbf{AB})_{|j} = \mathbf{A}(\mathbf{B})_{|j}.$$

Por tanto, las columnas de $\mathbf{AB}$ son combinaciones lineales de las columnas de $\mathbf{A}$.

¡De nuevo nótese la asociatividad de la notación!

El producto solo está definido si la primera matriz tiene tantas columnas como filas tiene la segunda.

Ejercicio 21. Demuestre las siguientes proposiciones.

(a) **Proposición 3.1.3.** *Sean las matrices* $\underset{m\times p}{\mathbf{A}}$, $\underset{m\times p}{\mathbf{B}}$ *y* $\underset{p\times n}{\mathbf{C}}$, *entonces* $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$.

(b) **Proposición 3.1.4.** *Sean las matrices* $\underset{m\times p}{\mathbf{A}}$, $\underset{p\times n}{\mathbf{B}}$ *y* $\underset{p\times n}{\mathbf{C}}$, *entonces* $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$.

(c) **Proposición 3.1.5.** *Sean las matrices* $\underset{m\times p}{\mathbf{A}}$, $\underset{p\times n}{\mathbf{B}}$ *y el escalar* $\lambda$, *entonces* $\mathbf{A}(\lambda\mathbf{B}) = \lambda(\mathbf{AB}) = (\lambda\mathbf{A})\mathbf{B}$.

*Pista.* Recuerde que $\mathbf{A}(\lambda\boldsymbol{b}) \begin{cases} = \lambda(\mathbf{A}\boldsymbol{b}) \\ \\ = (\lambda\mathbf{A})\boldsymbol{b} \end{cases}$ ; (Proposiciones 2.2.2 y 2.2.4 de la Página 30).

(d) **Proposición 3.1.6.** *Sean la matriz identidad* $\mathbf{I}$ *de orden* $m$ *y la matriz* $\underset{m\times n}{\mathbf{A}}$, *entonces* $\mathbf{IA} = \mathbf{A}$.

*Pista.* Recuerde que $\mathbf{I}\boldsymbol{a} = \boldsymbol{a}$ (Caso Especial 2 en la página 29).

(e) **Proposición 3.1.7.** *Sean la matriz* $\underset{m\times n}{\mathbf{A}}$ *y la matriz identidad* $\mathbf{I}$ *de orden* $n$, *entonces* $\mathbf{AI} = \mathbf{A}$.

*Pista.* Recuerde que $\mathbf{A}(\mathbf{I}_{|j}) = \mathbf{A}_{|j}$ (Caso Especial 3 en la página 29).

**Muy muy importante:** hay que subrayar que no todas las propiedades de las matrices replican las propiedades de los números que usted conoce: por ejemplo, en general $\mathbf{AB}$ es distinto de $\mathbf{BA}$. Por ejemplo:[2]

$$\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 3\end{bmatrix} \neq \begin{bmatrix}1 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}.$$

Por otro lado: entre los escalares, si $\lambda^2 = \lambda$ entonces $\lambda$ es necesariamente 0 o 1; pero con las matrices $\mathbf{A}^2 = \mathbf{A}$ no implica que la matriz $\mathbf{A}$ sea cero $\mathbf{0}$ o la identidad $\mathbf{I}$. Por ejemplo, si $\mathbf{A} = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$ entonces $\mathbf{A}^2 = \mathbf{A}$.

En los escalares, si $\lambda^2 = 0$ entonces $\lambda$ es necesariamente 0, pero tampoco esto es cierto con las matrices. Por ejemplo, para $\mathbf{B} = \begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}$

$$\mathbf{B}^2 = \begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} = \mathbf{0}.$$

[2] para que ambos productos estén definidos es necesario que ambas matrices sean cuadradas y del mismo orden.

☞ El producto de matrices verifica las siguientes propiedades:

- $\mathbf{A}(\mathbf{B}c) = (\mathbf{AB})c$
- $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
- $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$
- $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$
- $\mathbf{A}(\lambda\mathbf{B}) = \lambda(\mathbf{AB}) = (\lambda\mathbf{A})\mathbf{B}$
- $\mathbf{IA} = \mathbf{A}$
- $\mathbf{AI} = \mathbf{A}$

Pese a las dos últimas propiedades, en general $\mathbf{AB} \neq \mathbf{BA}$.

Por tanto...

☞ El producto de matrices es un **operador lineal**; pues es *distributivo* respecto a la suma de matrices

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}.$$

y *asociativo* respecto al producto de una matriz por un escalar

$$\mathbf{A}(\lambda\mathbf{B}) = \lambda(\mathbf{AB}).$$

Y también es lineal por la izquierda, pues

$$(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC} \qquad \text{y} \qquad (\lambda\mathbf{A})\mathbf{B} = \lambda(\mathbf{AB}).$$

### 3.1.1. Otras dos formas de calcular el producto de matrices

Aplicando las propiedades que ya conocemos, es sencillo deducir otras formas de calcular el producto. Veamos dos nuevas (la primera seguramente le resultará familiar...).

**Cálculo del producto de matrices componente a componente (filas por columnas)**

EJERCICIO 22. Demuestre la siguiente proposición:

**Proposición 3.1.8** (Cálculo del producto componente a componente)**.** *Sean las matrices* $\underset{m\times p}{\mathbf{A}}$ , $\underset{p\times n}{\mathbf{B}}$ *, entonces:*

$$\boxed{{}_{i|}(\mathbf{AB})_{|j} = ({}_{i|}\mathbf{A}) \cdot (\mathbf{B}_{|j})}.$$

Verifíquelo con el producto del Ejemplo 3 en la página 35 (hágalo a mano y con el módulo NAcAL).

Pero hay más formas de calcular el producto...

**Cálculo del producto de matrices operando con las filas**

EJERCICIO 23. Demuestre la siguiente proposición:

**Proposición 3.1.9.** *Sean* $\underset{m\times p}{\mathbf{A}}$ *y* $\underset{p\times n}{\mathbf{B}}$ *, entonces:* $\boxed{{}_{i|}(\mathbf{AB}) = ({}_{i|}\mathbf{A})\mathbf{B}}$.

*Pista.* Recuerde que $\boldsymbol{a} \cdot (\mathbf{B}_{|j}) = (\boldsymbol{a}\mathbf{B})_{|j}$, para $j = 1:n$.

Verifíquelo con el producto del Ejemplo 3 en la página 35 (hágalo a mano y con el módulo NAcAL).

Por tanto, las filas de $\mathbf{AB}$ son combinaciones lineales de las filas de $\mathbf{B}$.

☞ Explotando el producto punto, hemos obtenido otras dos formas de calcular $\mathbf{AB}$.

La primera alternativa calcula el producto componente a componente usando el *producto punto*.

$$\boxed{{}_{i|}(\mathbf{AB})_{|j} = ({}_{i|}\mathbf{A}) \cdot (\mathbf{B}_{|j})}.$$

de manera que los elementos de $\mathbf{AB}$ son el producto punto de una fila de $\mathbf{A}$ y una columna de $\mathbf{B}$.

La segunda alternativa calcula el producto fila a fila:

$$\boxed{{}_{i|}(\mathbf{AB}) = ({}_{i|}\mathbf{A})\mathbf{B}},$$

de manera que las filas de $\mathbf{AB}$ son combinaciones lineales de las filas de $\mathbf{B}$.

## 3.2. Nuevas reglas de reescritura.

Realmente no es necesario escribir el "punto" en las siguientes expresiones

$$({}_{i|}\mathbf{A}) \cdot \boldsymbol{b}\,, \qquad \boldsymbol{a} \cdot (\mathbf{B}_{|j})\,, \qquad ({}_{i|}\mathbf{A}) \cdot (\mathbf{B}_{|j})$$

(como tampoco lo es en el producto $2 \cdot \boldsymbol{x}$). Si omitimos el "punto" recuperamos la asociatividad

$${}_{i|}(\mathbf{A}\boldsymbol{b}) = ({}_{i|}\mathbf{A})\boldsymbol{b}$$

$$(\boldsymbol{a}\mathbf{B})_{|j} = \boldsymbol{a}(\mathbf{B}_{|j})$$

$${}_{i|}(\mathbf{AB})_{|j} = {}_{i|}\Big((\mathbf{AB})_{|j}\Big) = {}_{i|}\Big(\mathbf{A}(\mathbf{B}_{|j})\Big) = ({}_{i|}\mathbf{A})(\mathbf{B}_{|j}).$$

Por tanto, podemos escribir sin ambigüedad las siguientes expresiones:

$$\boxed{{}_{i|}\mathbf{A}\boldsymbol{b} \quad ; \quad \boldsymbol{a}\mathbf{B}_{|j} \quad ; \quad {}_{i|}\mathbf{AB}_{|j}}$$

pues da el mismo resultado si se interpretan como selección de un componente o como un producto punto.

**Transpuesta de un producto**

Para finalizar esta lección, demuestre la siguiente propiedad de la transposición que usaremos a menudo:

Ejercicio 24. Recordando que $(\mathbf{A}^\intercal)_{|j} = {}_{j|}\mathbf{A}$ y que $\boldsymbol{a}\mathbf{B} = (\mathbf{B}^\intercal)\boldsymbol{a}$; demuestre que

**Proposición 3.2.1.** *Sean* $\underset{m\times p}{\mathbf{A}}$ *y* $\underset{p\times n}{\mathbf{B}}$ *, entonces:*

$$\boxed{(\mathbf{AB})^\intercal = (\mathbf{B}^\intercal)(\mathbf{A}^\intercal)}. \tag{3.1}$$

Por tanto, para toda matriz $\mathbf{A}$, se verifica que las matrices $(\mathbf{A}^\intercal)\mathbf{A}$ y $\mathbf{A}(\mathbf{A}^\intercal)$ son simétricas:

$$\Big((\mathbf{A}^\intercal)\mathbf{A}\Big)^\intercal = (\mathbf{A}^\intercal)\Big((\mathbf{A}^\intercal)^\intercal\Big) = (\mathbf{A}^\intercal)\mathbf{A}; \qquad \text{y} \qquad \Big(\mathbf{A}(\mathbf{A}^\intercal)\Big)^\intercal = \Big((\mathbf{A}^\intercal)^\intercal\Big)(\mathbf{A}^\intercal) = \mathbf{A}(\mathbf{A}^\intercal).$$

Nótese que la transposición no es asociativa respecto al producto por una *matriz*, ya que $(\mathbf{AB})^\intercal \neq \mathbf{A}(\mathbf{B}^\intercal)$; por tanto, el uso de los paréntesis cobra importancia, pues la expresión $\mathbf{AB}^\intercal$ es ambigua. Lo usual es interpretar que la transposición solo afecta a $\mathbf{B}$ en la expresión anterior, pero es preferible escribir $\mathbf{A}(\mathbf{B}^\intercal)$.

# *Apéndices a la lección*

## 3.A. Productos de matrices triangulares

**Definición 3.2.** *Decimos que una matriz* $\mathbf{L}$ *es triangular inferior cuando todos los componentes por encima de la diagonal principal son nulos, es decir, si* $l_{ij} = 0$ *cuando* $i < j$.

Por ejemplo, las matrices

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} b_{11} & 0 & \cdots & 0 & 0 \\ b_{21} & b_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mp} & 0 \end{bmatrix}; \quad \text{y} \quad \mathbf{C} = \begin{bmatrix} c_{11} & 0 & \cdots & 0 \\ c_{21} & c_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ : & : & \cdots & c_{(m-1)n} \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix}.$$

Las tres son triangulares inferiores, aunque solo $\mathbf{A}$ es cuadrada. Fíjese que las componentes de la diagonal principal y por debajo pueden tomar cualquier valor (por eso el nombre de "inferior")... ¡Por tanto, una matriz nula $\mathbf{0}$ siempre es triangular inferior... pero también es triangular superior!

**Definición 3.3.** *Decimos que una matriz* $\mathbf{U}$ *es triangular superior cuando todos los componentes por debajo de la diagonal principal son nulos, es decir, si* $u_{ij} = 0$ *cuando* $i > j$.

Por tanto, una matriz es triangular superior si su transpuesta es triangular inferior.

**Definición 3.4.** *Una matriz triangular es unitaria si los elementos de la diagonal son iguales a uno.*

Las matrices *diagonales* son simultáneamente triangulares superiores e inferiores. Las matrices identidad son *cuadradas*, simultáneamente triangulares superiores e inferiores, y además *unitarias*.

Ejercicio 25. Demuestre las siguientes proposiciones.

(a) **Proposición 3.A.1.** *El producto de dos matrices triangulares superiores es triangular superior.*
(b) **Proposición 3.A.2.** *Si* $\mathbf{A}$ *es triangular superior unitaria y* $\mathbf{B}$ *es triangular superior,* $\mathbf{AB}$ *es triangular superior y mantiene la diagonal de* $\mathbf{B}$.
(c) **Proposición 3.A.3.** *Si* $\mathbf{A}$ *es triangular superior y* $\mathbf{B}$ *es triangular superior unitaria,* $\mathbf{AB}$ *es triangular superior y mantiene la diagonal de* $\mathbf{A}$.

Como el producto de matrices triangulares superiores unitarias mantiene la diagonal (formada por unos), tenemos el siguiente

**Corolario 3.A.4.** *El producto de dos matrices triangulares superiores unitarias es también triangular superior unitaria.*

Resultados análogos son ciertos con matrices triangulares inferiores; basta transponer para demostrarlo.

## 3.B. Submatrices mediante selección de una lista de índices

Este apéndice extiende a las matrices la selección con una lista de índices del apéndice de la Lección 1 (Sección 1.A).

Sea $\mathbf{I}$ la matriz identidad de orden $n$ y sea $\boldsymbol{\alpha}$ una lista de índices menores o iguales a $n$, por ejemplo, $n = 4$ y $\boldsymbol{\alpha} = (2, 3, 4, )$. Definimos $\mathbf{I}_{|\boldsymbol{\alpha}}$ como la matriz (de $n$ filas y tantas columnas como *elementos en* $\boldsymbol{\alpha}$) cuyas columnas son las columnas de $\underset{n\times n}{\mathbf{I}}$ indicadas en la lista $\boldsymbol{\alpha}$. Por ejemplo

$$\mathbf{I}_{|(2,3,4,)} = \begin{bmatrix} \mathbf{I}_{|2} & \mathbf{I}_{|3} & \mathbf{I}_{|4} \end{bmatrix} \neq \begin{bmatrix} \mathbf{I}_{|4} & \mathbf{I}_{|2} & \mathbf{I}_{|3} \end{bmatrix} = \mathbf{I}_{|(4,2,3,)}.$$

Y definimos ${}_{\boldsymbol{\alpha}|}\mathbf{I}$ como: ${}_{\boldsymbol{\alpha}|}\mathbf{I} = \big(\mathbf{I}_{|\boldsymbol{\alpha}}\big)^{\mathsf{T}}$.

**Selección de una submatriz por filas o por columnas** Ahora, con ayuda del producto de matrices, podemos definir dos nuevas operaciones. Sea $\boldsymbol{\alpha}$ una lista de índices, entonces podemos seleccionar una lista de columnas (o filas) de una matriz $\mathbf{A}$ para construir una submatriz:

$$\mathbf{A}_{|\boldsymbol{\alpha}} = \mathbf{A}\big(\mathbf{I}_{|\boldsymbol{\alpha}}\big), \qquad \text{y} \qquad {}_{\boldsymbol{\alpha}|}\mathbf{A} = \big({}_{\boldsymbol{\alpha}|}\mathbf{I}\big)\mathbf{A}.$$

Consideremos de nuevo la matriz $\mathbf{M}$ de la Página 11. La submatriz con las filas de 2 a 4 de $\mathbf{M}$ (datos de Francia, Alemania y Reino Unido) o la submatriz con sus dos primeras columnas (años 2003 y 2004) son

$${}_{(2,3,4,)|}\mathbf{M} = \begin{bmatrix} 0.9 & 2.0 & 1.6 & 2.0 \\ -0.2 & 1.1 & 0.7 & 0.9 \\ 2.5 & 3.2 & 1.9 & 1.5 \end{bmatrix}; \quad \text{y} \quad \mathbf{M}_{|(1,2,)} = \begin{bmatrix} 2.9 & 3.1 \\ 0.9 & 2.0 \\ -0.2 & 1.1 \\ 2.5 & 3.2 \\ 1.8 & 3.1 \end{bmatrix}.$$

Podemos crear submatrices seleccionando simultáneamente filas y columnas; si $\boldsymbol{\beta}$ es otra lista de índices:

$${}_{\boldsymbol{\alpha}|}\big(\mathbf{A}_{|\boldsymbol{\beta}}\big) = \big({}_{\boldsymbol{\alpha}|}\mathbf{I}\big)\mathbf{A}\big(\mathbf{I}_{|\boldsymbol{\beta}}\big) = \big({}_{\boldsymbol{\alpha}|}\mathbf{A}\big)_{|\boldsymbol{\beta}},$$

que, aprovechando la asociatividad, podemos expresar de manera sencilla como: ${}_{\boldsymbol{\alpha}|}\mathbf{A}_{|\boldsymbol{\beta}}$.

Así, la submatriz que contiene los datos de Francia, Alemania y Reino Unido de los años 2003 y 2004 es

$${}_{(2,3,4,)|}\mathbf{M}_{|(1,2,)} = \begin{bmatrix} 0.9 & 2.0 \\ -0.2 & 1.1 \\ 2.5 & 3.2 \end{bmatrix}.$$

Nótese que $\big({}_{\boldsymbol{\alpha}|}\mathbf{A}\big)\big(\mathbf{B}_{|\boldsymbol{\beta}}\big) = \big({}_{\boldsymbol{\alpha}|}\mathbf{I}\big)\mathbf{A}\mathbf{B}\big(\mathbf{I}_{|\boldsymbol{\beta}}\big) = {}_{\boldsymbol{\alpha}|}\big(\mathbf{A}\mathbf{B}\big)_{|\boldsymbol{\beta}}$; en particular $\big(\mathbf{A}\mathbf{B}\big)_{|\boldsymbol{\beta}} = \mathbf{A}\big(\mathbf{B}_{|\boldsymbol{\beta}}\big)$ y ${}_{\boldsymbol{\alpha}|}\big(\mathbf{A}\mathbf{B}\big) = \big({}_{\boldsymbol{\alpha}|}\mathbf{A}\big)\mathbf{B}$; por lo que podemos escribir sencillamente: ${}_{\boldsymbol{\alpha}|}\mathbf{A}\mathbf{B}_{|\boldsymbol{\beta}}$. También se verifica que

$${}_{\boldsymbol{\alpha}|}\big(\mathbf{A}+\mathbf{B}\big)_{|\boldsymbol{\beta}} = {}_{\boldsymbol{\alpha}|}\mathbf{A}_{|\boldsymbol{\beta}} + {}_{\boldsymbol{\alpha}|}\mathbf{B}_{|\boldsymbol{\beta}} \qquad \text{y} \qquad \lambda\big({}_{\boldsymbol{\alpha}|}\mathbf{A}_{|\boldsymbol{\beta}}\big) = {}_{\boldsymbol{\alpha}|}\big(\lambda\mathbf{A}\big)_{|\boldsymbol{\beta}}. \tag{3.2}$$

**Definición 3.5.** *Llamamos* submatriz principal *de* $\underset{n\times n}{\mathbf{A}}$ *a cualquier submatriz de la forma* ${}_{\boldsymbol{\alpha}|}\mathbf{A}_{|\boldsymbol{\alpha}}$; *donde* $\boldsymbol{\alpha}$ *es una lista de* $k$ *índices tal que* $1 \le \alpha_1 < \alpha_2 < \cdots < \alpha_k \le n$. *Es decir, los índices son distintos, ordenados y comunes en la selección de filas y columnas. El caso particular en el que se seleccionan las* $k$ *primeras filas y columnas:* ${}_{(1:k)|}\mathbf{A}_{|(1:k)}$ *se denomina* submatriz principal superior.[3]

## 3.C. Producto matricial como suma de productos de submatrices

☞ En el Ejercicio 85 en la página 182 de la Sección 15.3 se emplea el *producto de matrices por bloques* para demostrar cómo es el determinante de una matriz diagonal por bloques (resultado que se empleará en el Corolario 17.2.4 en la página 203 de la lección sobre diagonalización por semejanza).

Este apéndice muestra cómo calcular el *producto de matrices por bloques* (que es una aplicación de la Ecuación 3.7 en la página 44 cuando las submatrices son el resultado de "cortar" las matrices en bloques).

Considere una lista $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_q,)$ de $q$ índices no repetidos, así como la submatriz ${}_{\boldsymbol{\alpha}|}\mathbf{I}$, de la matriz identidad $\mathbf{I}$ de orden mayor o igual al mayor de los índices en $\boldsymbol{\alpha}$. Si $\alpha_j$ es la $j$-*ésima* componente de la lista

[3] Aludiremos a las submatrices principales cuando abordemos el determinante de una matriz.

$\boldsymbol{\alpha}$, entonces la columna $\left({}_{\boldsymbol{\alpha}|}\mathbf{I}\right)_{|\alpha_j}$ resulta ser igual a $\mathbf{I}_{|j} \in \mathbb{R}^q$ (donde $q$ es la longitud de la lista $\boldsymbol{\alpha}$):

$$ {}_{\boldsymbol{\alpha}|}\mathbf{I}_{|k} = {}_{(\alpha_1,\ldots,\alpha_q,)|}\mathbf{I}_{|k} = {}_{(\alpha_1,\ldots,\alpha_q,)|}\left(\mathbf{I}_{|k}\right) = \begin{cases} \mathbf{I}_{|j} \in \mathbb{R}^q & \text{si } k = \alpha_j \\ & \in \mathbb{R}^q, \\ \mathbf{0} \in \mathbb{R}^q & \text{si } k \notin \boldsymbol{\alpha} \end{cases} \tag{3.3} $$

puesto que si de $\mathbf{I}_{|k}$ (cuya única componente no nula es la $k$-ésima) seleccionamos un subvector de $q$ componentes que coloca su $k$-ésima componente en la posición $j$, obtenemos $\mathbf{I}_{|j} \in \mathbb{R}^q$, y si seleccionamos un subvector que no toma su única componente no nula, obtenemos un subvector nulo: $\mathbf{0} \in \mathbb{R}^q$.

*Ejemplo* 4. Consideremos la lista de 3 índices $\boldsymbol{\alpha} = (4, 1, 2, )$ (cuyo mayor índice es el 4) y una matriz identidad de orden mayor o igual que 4; por ejemplo $\mathbf{I}$ de orden 4. Entonces

$$ {}_{\boldsymbol{\alpha}|}\mathbf{I} = {}_{(4,1,2,)|}\mathbf{I} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. $$

Fijándonos en las columnas de la anterior matriz, observamos que:

- como 4 es el *primer* componente de la lista $\boldsymbol{\alpha}$, entonces la cuarta columna $\left({}_{(4,1,2,)|}\mathbf{I}\right)_{|4}$ es igual a la *primera* columna de la identidad de orden 3 (pues 3 es la longitud de la lista):

$$ {}_{(4,1,2,)|}\mathbf{I}_{|4} = {}_{(4,1,2,)|}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. $$

- como 1 es la *segunda* componente de la lista $\boldsymbol{\alpha}$, entonces la primera columna $\left({}_{(4,1,2,)|}\mathbf{I}\right)_{|1}$ es igual a la *segunda* columna de la identidad:

$$ {}_{(4,1,2,)|}\mathbf{I}_{|1} = {}_{(4,1,2,)|}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. $$

- como 2 es la *tercera* componente de la lista $\boldsymbol{\alpha}$, entonces la segunda columna $\left({}_{(4,1,2,)|}\mathbf{I}\right)_{|2}$ es igual a la *tercera* columna de la identidad:

$$ {}_{(4,1,2,)|}\mathbf{I}_{|2} = {}_{(4,1,2,)|}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. $$

- y como 3 no está en la lista $\boldsymbol{\alpha}$, entonces la tercera columna $\left({}_{(4,1,2,)|}\mathbf{I}\right)_{|3}$ es nula

$$ {}_{(4,1,2,)|}\mathbf{I}_{|3} = {}_{(4,1,2,)|}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. $$

Módulo NAcAL para Python

```
alpha = (4,1,2,)                                  # índices
Sistema([ alpha|I(99)|j for j in (4,1,2,3,) ]) # las cuatro columnas del ejemplo
```

$$\left[\ \begin{pmatrix}1\\0\\0\end{pmatrix};\quad \begin{pmatrix}0\\1\\0\end{pmatrix};\quad \begin{pmatrix}0\\0\\1\end{pmatrix};\quad \begin{pmatrix}0\\0\\0\end{pmatrix};\ \right]$$

Y ahora considere las columnas del producto de matrices $(\mathbf{I}_{|\boldsymbol{\alpha}})(_{\boldsymbol{\alpha}|}\mathbf{I})$. Usando (3.3) se deduce que

$$\Big((\mathbf{I}_{|\boldsymbol{\alpha}})(_{\boldsymbol{\alpha}|}\mathbf{I})\Big)_{|k} = (\mathbf{I}_{|\boldsymbol{\alpha}})(_{\boldsymbol{\alpha}|}\mathbf{I}_{|k}) = \begin{cases}(\mathbf{I}_{|\boldsymbol{\alpha}})\mathbf{I}_{|j} = (\mathbf{I}_{|\boldsymbol{\alpha}})_{|j} = \mathbf{I}_{|k} & \text{si } k = \alpha_j\\ (\mathbf{I}_{|\boldsymbol{\alpha}})\mathbf{0} \ = \mathbf{0} & \text{si } k \notin \boldsymbol{\alpha}\end{cases}; \tag{3.4}$$

Así, para $\boldsymbol{\alpha} = (4,1,2,)$ e $\underset{4\times 4}{\mathbf{I}}$ tenemos que

$$(\mathbf{I}_{|(4,1,2,)})(_{(4,1,2,)|}\mathbf{I}) = \begin{bmatrix}0&1&0\\0&0&1\\0&0&0\\1&0&0\end{bmatrix}\begin{bmatrix}0&0&0&1\\1&0&0&0\\0&1&0&0\end{bmatrix} = \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&1\end{bmatrix},$$

de manera que solo hay unos en las posiciones 1, 2 y 4 de la diagonal.

Módulo NAcAL para Python

```
(I(4)|alpha) * (alpha|I(4)) # fíjese en la necesidad de los paréntesis
```

Ahora considere las matrices $\underset{m\times p}{\mathbf{A}}$ y $\underset{p\times n}{\mathbf{B}}$; y las listas de índices $\boldsymbol{\gamma}_1,\ldots,\boldsymbol{\gamma}_h$ tales que cada índice $j = 1:p$ está en una, y solo una, de las listas $\boldsymbol{\gamma}_i$. Por ejemplo, si $p = 5$, $\boldsymbol{\gamma}_1 = (3,1,)$, $\boldsymbol{\gamma}_2 = (5,)$ y $\boldsymbol{\gamma}_3 = (4,2,)$; entonces

$$\mathbf{A}_{|\boldsymbol{\gamma}_1\#\boldsymbol{\gamma}_2\#\boldsymbol{\gamma}_3} \ = \ \mathbf{A}_{|(3,1,5,4,2,)};$$

donde $\#$ es el operador que concatena. Es decir, $\boldsymbol{\gamma}_1\#\boldsymbol{\gamma}_2\#\boldsymbol{\gamma}_3 = (3,1,5,4,2,)$ es la concatenación de las listas $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$ y $\boldsymbol{\gamma}_3$. Entonces $\mathbf{A}_{|\boldsymbol{\gamma}_1\#\boldsymbol{\gamma}_2\#\boldsymbol{\gamma}_3}$ es un reordenamiento (o permutación) de las columnas, pues todos sus índices aparecen una, y solo una, vez en la lista $\boldsymbol{\gamma}_1\#\boldsymbol{\gamma}_2\#\boldsymbol{\gamma}_3 = (3,1,5,4,2,)$.

Si empleamos las listas $\boldsymbol{\gamma}_1,\ldots,\boldsymbol{\gamma}_h$ para reordenar tanto las columnas de $\mathbf{A}$ como las filas de $\mathbf{B}$, el producto de las matrices resultantes resulta ser igual al producto $\mathbf{AB}$:

$$(\mathbf{A}_{|\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h})(_{\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h|}\mathbf{B}) = \mathbf{A}(\mathbf{I}_{|\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h})(_{\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h|}\mathbf{I})\mathbf{B} = \mathbf{AIB} = \mathbf{AB},$$

pues, como todos los índices están incluidos por (3.4), sabemos que $(\mathbf{I}_{|\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h})(_{\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h|}\mathbf{I}) = \mathbf{I}$.

*Ejemplo* 5. Si usamos las listas $\boldsymbol{\gamma}_1 = (3,1,)$ y $\boldsymbol{\gamma}_2 = (4,2,)$ con $\underset{4\times 4}{\mathbf{I}}$, entonces

Módulo NAcAL para Python

```
g1=(3,1,); g2=(4,2,)
(I(4)|g1)*(g1|I(4)) + (I(4)|g2)*(g2|I(4))
```

$$(\mathbf{I}_{|(3,1,)})(_{(3,1,)|}\mathbf{I}) + (\mathbf{I}_{|(4,2,)})(_{(4,2,)|}\mathbf{I}) = \begin{bmatrix}1&0&0&0\\0&0&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix} + \begin{bmatrix}0&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&1\end{bmatrix} = \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}.$$

Consecuentemente, podemos calcular $\mathbf{AB}$ mediante la suma de $k$ productos de submatrices:

$$\sum_{k=1}^{h} (\mathbf{A}_{|\boldsymbol{\gamma}_k})(_{\boldsymbol{\gamma}_k|}\mathbf{B}) = \sum_{k=1}^{h} \Big(\mathbf{A}(\mathbf{I}_{|\boldsymbol{\gamma}_k})(_{\boldsymbol{\gamma}_k|}\mathbf{I})\mathbf{B}\Big) = \mathbf{A}\left(\sum_{k=1}^{h} (\mathbf{I}_{|\boldsymbol{\gamma}_k})(_{\boldsymbol{\gamma}_k|}\mathbf{I})\right)\mathbf{B} = \mathbf{AIB} = \mathbf{AB}, \tag{3.5}$$

donde, por (3.4), cada sumando de $\sum_{k=1}^{h} (\mathbf{I}_{|\boldsymbol{\gamma}_k})(_{\boldsymbol{\gamma}_k|}\mathbf{I})$ es una matriz cuadrada cuya $j$-ésima columna es $\mathbf{I}_{|j}$, cuando $j \in \boldsymbol{\gamma}_k$, o $\mathbf{0}$, en caso contrario. Y como cada índice está incluido en una y solo una de las listas $\boldsymbol{\gamma}_k$, la suma de dichas matrices es igual a la identidad.

*Ejemplo* 6. Para $\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -2 & 0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ -1 & 1 \end{bmatrix}$; y las listas $\boldsymbol{\gamma}_1 = (3,1,)$ y $\boldsymbol{\gamma}_2 = (2,)$; tenemos

$$\begin{aligned}(\mathbf{A}_{|\boldsymbol{\gamma}_1\#\boldsymbol{\gamma}_2})(_{\boldsymbol{\gamma}_1\#\boldsymbol{\gamma}_2|}\mathbf{B}) &= (\mathbf{A}_{|\boldsymbol{\gamma}_1})(_{\boldsymbol{\gamma}_1|}\mathbf{B}) + (\mathbf{A}_{|\boldsymbol{\gamma}_2})(_{\boldsymbol{\gamma}_2|}\mathbf{B}) = (\mathbf{A}_{|(3,1,)})(_{(3,1,)|}\mathbf{B}) + (\mathbf{A}_{|(2,)})(_{(2,)|}\mathbf{B}) \\ &= \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 2 \\ -2 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 4 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ -2 & -2 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 2 & 0 \end{bmatrix} = \mathbf{AB}.\end{aligned}$$

Módulo NAcAL para Python

```
A = Matrix([ [1,2,-1], [2,-2,0] ])
B = Matrix([ Vector([2,1,-1]), Vector([1,1,1])]
g1=(3,1,); g2=(2,)                          # Fíjese en la coma tras el 2
(A|g1)*(g1|B) + (A|g2)*(g2|B) )
```

Nota: las listas de un solo elemento se escriben en Python con una coma después del único elemento. Por ejemplo `(2,)` es la lista que solo contiene el 2. Así se distingue el *número* 2 entre paréntesis de la *lista* que contiene al 2. Por claridad, y puesto que la distinción es importante, hemos seguido el mismo criterio de notación:[4] así, $\mathbf{A}_{|2}$ es una columna de $\mathbf{A}$ (un vector) y $\mathbf{A}_{|(2,)}$ es una submatriz (una matriz columna).

### 3.C.1. Producto matricial expresado como una suma de matrices

Como caso particular, considere que cada lista $k$-ésima contiene únicamente el índice $k$, es decir, $\boldsymbol{\gamma}_k = (k,)$. Entonces, el producto $\mathbf{AB}$ queda descrito como una suma de matrices:

$$(\underset{m\times p}{\mathbf{A}})(\underset{p\times n}{\mathbf{B}}) = \sum_{k=1}^{p} (\mathbf{A}_{|(k,)})(_{(k,)|}\mathbf{B}), \tag{3.6}$$

donde cada sumando $(\underset{m\times 1}{\mathbf{A}_{|(k,)}})(\underset{1\times n}{_{(k,)|}\mathbf{B}})$ es una matriz[5] de orden $m$ por $n$. Ejemplo:

$$\begin{aligned}\mathbf{AB} &= \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} = (\mathbf{A}_{|(1,)})(_{(1,)|}\mathbf{B}) + (\mathbf{A}_{|(2,)})(_{(2,)|}\mathbf{B}) + (\mathbf{A}_{|(3,)})(_{(3,)|}\mathbf{B}) \\ &= \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\begin{bmatrix} -1 & 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}\begin{bmatrix} 0 & 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ -1 & 2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 3 & 0 \\ 1 & 0 \\ 2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 2 \\ 0 & 2 \\ 2 & 0 \end{bmatrix}.\end{aligned}$$

[4] véase la nota a pie 5 en la página 4

[5] Cuando veamos el concepto de rango, entenderá que esta expresión es una descripción del producto matricial como suma de matrices de rango uno como máximo.

### 3.C.2. Submatriz de un producto como suma de productos de submatrices

Considere una lista de listas $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_h$, donde cada índice $j = 1:p$ está incluido en una y solo en una de las listas $\boldsymbol{\gamma}_k$; por (3.5) sabemos que:

$$\mathbf{AB} \;=\; \sum_{k=1}^{h} (\mathbf{A}_{|\boldsymbol{\gamma}_k})(_{\boldsymbol{\gamma}_k|}\mathbf{B}).$$

Por tanto, seleccionando algunas filas de $\underset{m\times p}{\mathbf{A}}$ y algunas columnas de $\underset{p\times n}{\mathbf{B}}$, la submatriz $_{\boldsymbol{\alpha}|}(\mathbf{AB})_{|\boldsymbol{\beta}}$ resulta ser

$$_{\boldsymbol{\alpha}|}(\mathbf{AB})_{|\boldsymbol{\beta}} \;=\; \sum_{k=1}^{h} (_{\boldsymbol{\alpha}|}\mathbf{A}_{|\boldsymbol{\gamma}_k})(_{\boldsymbol{\gamma}_k|}\mathbf{B}_{|\boldsymbol{\beta}}). \tag{3.7}$$

**Matrices partidas en bloques**

☞ Y por fin podemos presentar un modo particular de multiplicar matrices (que era objetivo de este apéndice).

Fíjese que en (3.7) no se requiere que las sublistas $\boldsymbol{\gamma}_k$ estén formadas por índices consecutivos; pero este suele ser el caso más habitual en los textos. En tal caso, las matrices quedan "cortadas" en bloques. Fíjese en los dos siguientes ejemplos (en los que hemos marcado los bloques mediante líneas horizontales y verticales).

*Ejemplo* 7. Sean $\boldsymbol{\alpha} = \big((1,2,),\ (3,4,)\big)$, $\boldsymbol{\gamma} = \big((1,2,),\ (3,4,5,)\big)$; y las matrices

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{C} & \mathbf{E} \\ \hline \mathbf{D} & \mathbf{F} \end{array}\right] = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 2 & -1 & 4 & 2 & 1 \\ 3 & 1 & -1 & 7 & 5 \end{array}\right] \quad \text{y} \quad \mathbf{B} = \left[\begin{array}{c} \mathbf{G} \\ \hline \mathbf{H} \end{array}\right] = \left[\begin{array}{cc} 4 & -2 \\ 5 & 6 \\ \hline 7 & 3 \\ -1 & 0 \\ 1 & 6 \end{array}\right];$$

donde $\mathbf{C} = {}_{(1:2)|}\mathbf{A}_{|(1:2)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{E} = {}_{(1:2)|}\mathbf{A}_{|(3:5)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $\mathbf{D} = {}_{(3:4)|}\mathbf{A}_{|(1:2)} = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$ y $\mathbf{F} = {}_{(3:4)|}\mathbf{A}_{|(3:5)} = \begin{bmatrix} 4 & 2 & 1 \\ -1 & 7 & 5 \end{bmatrix}$. Y donde $\mathbf{G} = {}_{(1:2)|}\mathbf{B} = \begin{bmatrix} 4 & -2 \\ 5 & 6 \end{bmatrix}$ y $\mathbf{H} = {}_{(3:5)|}\mathbf{B} = \begin{bmatrix} 7 & 3 \\ -1 & 0 \\ 1 & 6 \end{bmatrix}$.

$$\text{Entonces,} \qquad \mathbf{AB} = \left[\begin{array}{c|c} \mathbf{C} & \mathbf{E} \\ \hline \mathbf{D} & \mathbf{F} \end{array}\right]\left[\begin{array}{c} \mathbf{G} \\ \hline \mathbf{H} \end{array}\right] = \left[\begin{array}{c} \mathbf{CG}+\mathbf{EH} \\ \hline \mathbf{DG}+\mathbf{FH} \end{array}\right] = \left[\begin{array}{cc} 4 & -2 \\ 5 & 6 \\ \hline 30 & 8 \\ 8 & 27 \end{array}\right].$$

*Ejemplo* 8. De manera similar, sean $\boldsymbol{\alpha} = \big((1,2,),\ (3,4,)\big)$, $\boldsymbol{\gamma} = \big((1,2,),\ (3,4,5,)\big)$, $\boldsymbol{\beta} = \big((1,2,),\ (3,)\big)$,

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{C} & \mathbf{E} \\ \hline \mathbf{D} & \mathbf{F} \end{array}\right] = \left[\begin{array}{cc|cc} 2 & -1 & 3 & 1 \\ 1 & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \quad \text{y} \quad \mathbf{B} = \left[\begin{array}{c|c} \mathbf{G} & \mathbf{J} \\ \hline \mathbf{H} & \mathbf{K} \end{array}\right] = \left[\begin{array}{cc|c} 1 & 2 & 0 \\ -1 & 0 & 0 \\ \hline 0 & 5 & 1 \\ 1 & -1 & 0 \end{array}\right]; \quad \text{entonces}$$

$$\mathbf{AB} = \left[\begin{array}{c|c} \mathbf{CG}+\mathbf{EH} & \mathbf{CJ}+\mathbf{EK} \\ \hline \mathbf{DG}+\mathbf{FH} & \mathbf{DH}+\mathbf{FK} \end{array}\right] = \left[\begin{array}{cc|c} 4 & 18 & 3 \\ 3 & 5 & 1 \\ \hline 0 & 5 & 1 \\ 1 & -1 & 0 \end{array}\right].$$

### 3.C.3. Matriz por vector como suma de submatrices por subvectores

**Definición 3.6.** *Sea* $_{\boldsymbol{\gamma}|}\boldsymbol{b}$ *el vector formado por las componentes de* $\boldsymbol{b}$ *indicadas en* $\boldsymbol{\gamma}$*, es decir,* $_{\boldsymbol{\gamma}|}\boldsymbol{b} = (_{\boldsymbol{\gamma}|}\mathbf{I})\boldsymbol{b}$.

Si $\mathbf{A}$ tiene $n$ columnas y $\boldsymbol{b} \in \mathbb{R}^n$; y si $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_h$ son listas de índices tales que cada índice $j = 1 : n$ está en una, y solo en una, de las listas $\boldsymbol{\gamma}_i$, entonces es inmediato ver que la combinación lineal $\mathbf{A}\boldsymbol{b}$ se puede expresar como suma de varias combinaciones lineales

$$\mathbf{A}\boldsymbol{b} = \sum_{k=1}^{h} (\mathbf{A}_{|\boldsymbol{\gamma}_k})(_{\boldsymbol{\gamma}_k|}\boldsymbol{b}). \tag{3.8}$$

Por ejemplo, si $n = 4$, $\boldsymbol{\gamma}_1 = (1, 3, )$, y $\boldsymbol{\gamma}_2 = (2, 4, )$. Entonces, reordenando los sumandos, tenemos

$$\mathbf{A}\boldsymbol{b} = \overbrace{(\mathbf{A}_{|1})(_{1|}\boldsymbol{b}) + (\mathbf{A}_{|3})(_{3|}\boldsymbol{b})}^{\text{combinación de columnas de índice impar}} + \overbrace{(\mathbf{A}_{|2})(_{2|}\boldsymbol{b}) + (\mathbf{A}_{|4})(_{4|}\boldsymbol{b})}^{\text{combinación de columnas de índice par}} = (\mathbf{A}_{|\boldsymbol{\gamma}_1})(_{\boldsymbol{\gamma}_1|}\boldsymbol{b}) + (\mathbf{A}_{|\boldsymbol{\gamma}_2})(_{\boldsymbol{\gamma}_2|}\boldsymbol{b}),$$

es decir, la suma de una combinación lineal de las columnas de índice impar, $(\mathbf{A}_{|\boldsymbol{\gamma}_1})(_{\boldsymbol{\gamma}_1|}\boldsymbol{b})$, más otra combinación de las columnas de índice par, $(\mathbf{A}_{|\boldsymbol{\gamma}_2})(_{\boldsymbol{\gamma}_2|}\boldsymbol{b})$.

## 3.D. Propiedades de la traza

Vamos a aprovechar la visión del producto de matrices como una suma de matrices para deducir algunas de las propiedades de la traza de una matriz.

**Definición 3.7.** *La* traza *de* $\underset{m\times n}{\mathbf{A}}$ *es la suma de los elementos de su diagonal principal, es decir:*

$$\operatorname{tr}(\mathbf{A}) = a_{11} + a_{22} + \cdots + a_{kk}, \quad \textit{donde } k = \min(m, n).$$

Ejercicio 26. Sabiendo que los operadores selectores son lineales y que $\underset{m\times p}{\mathbf{A}}\ \underset{p\times n}{\mathbf{B}} = \sum_{k=1}^{p} (\mathbf{A}_{|(k,)})(_{(k,)|}\mathbf{B})$, demuestre que

(a) la traza es un operador lineal,
(b) $\operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{A}^\intercal)$,
(c) $\operatorname{tr}(\mathbf{A}^\intercal\mathbf{A}) = \operatorname{tr}(\mathbf{A}\mathbf{A}^\intercal)$,
(d) Si $\mathbf{A}$ y $\mathbf{B}$ son de orden $m \times n$, entonces: $\operatorname{tr}(\mathbf{A}^\intercal\mathbf{B}) = \operatorname{tr}(\mathbf{B}\mathbf{A}^\intercal) = \operatorname{tr}(\mathbf{A}\mathbf{B}^\intercal) = \operatorname{tr}(\mathbf{B}^\intercal\mathbf{A})$.

Consecuentemente, si $\mathbf{A}$ es de orden $m \times n$ y $\mathbf{C}$ de orden $n \times m$, se tiene que

$$\operatorname{tr}(\mathbf{A}\mathbf{C}) = \operatorname{tr}(\mathbf{C}\mathbf{A}). \tag{3.9}$$

Basta renombrar $\mathbf{C}$ como $\mathbf{B}^\intercal$ y aplicar (d) para verlo: $\operatorname{tr}(\mathbf{A}\mathbf{C}) = \operatorname{tr}(\mathbf{A}\mathbf{B}^\intercal) = \operatorname{tr}(\mathbf{B}^\intercal\mathbf{A}) = \operatorname{tr}(\mathbf{C}\mathbf{A})$.

Por último, para $\underset{m\times n}{\mathbf{A}}$, $\underset{n\times p}{\mathbf{B}}$ y $\underset{p\times m}{\mathbf{C}}$ se verifica que $\operatorname{tr}(\mathbf{A}\mathbf{B}\mathbf{C}) = \operatorname{tr}(\mathbf{C}\mathbf{A}\mathbf{B}) = \operatorname{tr}(\mathbf{B}\mathbf{C}\mathbf{A})$, pues aplicando (3.9):

$$\operatorname{tr}((\mathbf{A}\mathbf{B})\mathbf{C}) = \operatorname{tr}(\mathbf{C}(\mathbf{A}\mathbf{B})) = \operatorname{tr}((\mathbf{C}\mathbf{A})\mathbf{B}) = \operatorname{tr}(\mathbf{B}(\mathbf{C}\mathbf{A})).$$

## 3.E. Elemento neutro multiplicativo universal

Llamaremos *elemento neutro multiplicativo universal por la izquierda* a cualquier matriz $\mathbf{F}$ de orden $m$ tal que:

$$\mathbf{F}\mathbf{X} = \mathbf{X} \qquad \text{para toda matriz } \underset{m\times n}{\mathbf{X}}\ .$$

Para ver que el único *elemento neutro multiplicativo universal por la izquierda* es la matriz identidad $\mathbf{I}$ de orden $m$, consideremos el caso particular en el que la matriz $\mathbf{X}$ es la matriz identidad $\mathbf{I}$ de orden $m$.

Por una parte, como $\mathbf{F}$ es un *elemento neutro multiplicativo universal por la izquierda*, tenemos que

$$\mathbf{FI} = \mathbf{I}.$$

Por otra parte, sabemos que dos matrices son iguales si y solo si lo son columna a columna. Por tanto,

$$\mathbf{I}_{|j} = \big(\mathbf{FI}\big)_{|j} = \mathbf{F}\big(\mathbf{I}_{|j}\big) = \mathbf{F}_{|j}.$$

Por lo tanto, $\mathbf{F} = \mathbf{I}$.

Y llamaremos *elemento neutro multiplicativo universal por la derecha* a cualquier matriz $\mathbf{C}$ de orden $n$ tal que

$$\mathbf{XC} = \mathbf{X} \qquad \text{para toda matriz } \underset{m\times n}{\mathbf{X}} \text{ ;}$$

Para ver que el único *elemento neutro multiplicativo universal por la derecha* es la matriz identidad $\mathbf{I}$ de orden $n$, consideremos el caso particular en el que la matriz $\mathbf{X}$ es la matriz identidad $\mathbf{I}$ de orden $n$.

Como $\mathbf{C}$ es un *elemento neutro multiplicativo universal por la derecha*, tenemos que

$$\mathbf{IC} = \mathbf{I}.$$

y como dos matrices son iguales si y solo si lo son fila a fila, tenemos que

$${}_{i|}\mathbf{I} = {}_{i|}\big(\mathbf{IC}\big) = \big({}_{i|}\mathbf{I}\big)\mathbf{C} = {}_{i|}\mathbf{C}.$$

Concluimos que, entre las matrices de orden $n$, el único *elemento neutro multiplicativo universal* (tanto por la derecha como por la izquierda) es la matriz identidad de orden $n$.

## 3.F. Potencias de una matriz cuadrada

Con $\mathbf{A}^m$ (donde $m$ es un número natural mayor que cero) denotamos el producto de $\mathbf{A}$ consigo misma $m$ veces:

$$\mathbf{A}^m = \underbrace{\mathbf{A}\cdot\cdots\cdot\mathbf{A}}_{m \text{ veces}}.$$

Fíjese que la notación satisface la *ley de exponentes*:

$$\mathbf{A}^m \cdot \mathbf{A}^n = \mathbf{A}^{m+n},$$

donde $m$ y $n$ son números naturales mayores que cero. Queremos extender esta notación a exponentes enteros.

### 3.F.1. Potencia cero de una matriz cuadrada

Extendamos la definición de potencias de una matriz para que siga siendo válida la ley de exponentes cuando $m = 0$. Aplicando ley de exponentes para $m = 0$, obtenemos

$$\mathbf{A}^0 \cdot \mathbf{A}^n = \mathbf{A}^{0+n} = \mathbf{A}^n \qquad \text{para todo } n \geq 0;$$

es decir, $\mathbf{A}^0$ tiene que actuar como elemento neutro multiplicativo sobre las potencias de $\mathbf{A}$.

Pero hemos visto que, en el álgebra de matrices, el único elemento neutro multiplicativo universal es la matriz identidad $\mathbf{I}$ de orden $n$; caracterizada por

$$\mathbf{IX} = \mathbf{XI} = \mathbf{X}$$

para toda matriz $\mathbf{X}$ de orden $n$. Consecuentemente, la única definición compatible para la potencia cero de una matriz cuadrada es

$$\mathbf{A}^0 = \mathbf{I}.$$

**Compatibilidad con exponentes negativos**

Para extender la definición de potencias de matrices de manera que también siga siendo válida la ley de exponentes cuando $m < 0$, debemos asumir que la matriz es invertible[6] Sea $\mathbf{A}$ invertible. Queremos que

$$\mathbf{A}^m \cdot \mathbf{A}^{-m} = \mathbf{A}^{m+m} = \mathbf{A}^0.$$

Pero por la definición de matriz inversa,

$$\mathbf{A}^m \cdot \mathbf{A}^{-m} = \mathbf{A}^m \cdot (\mathbf{A}^m)^{-1} = \mathbf{I}.$$

Es decir, la extensión a enteros negativos también requiere que $\mathbf{A}^0 = \mathbf{I}$.

[6]Este apartado hace uso del concepto de matriz inversa; que no se verá hasta la lección 5.

# Parte II

# Transformaciones elementales, métodos de eliminación y matriz inversa

# Transformaciones elementales y métodos de eliminación

Como veremos a lo largo de este curso, hay un algoritmo que permite resolver una gran cantidad de problemas: decidir si una matriz es invertible y, si lo es, encontrar su inversa; decidir si un sistema de ecuaciones lineales es resoluble y, si lo es, resolverlo; encontrar una base del complemento ortogonal de un subespacio de $\mathbb{R}^n$; pasar de las ecuaciones paramétricas a las cartesianas (y viceversa); calcular determinantes, etc. De hecho, ¡prácticamente todos los problemas que veremos en este curso se pueden resolver con un único algoritmo!

Este algoritmo se denomina *Método de Eliminación* y es sorprendentemente simple: *consiste en aplicar una secuencia de transformaciones sobre las columnas (o sobre las filas) de una matriz.* Cada transformación en la secuencia (cada paso en la secuencia) realiza uno de estos dos tipos de operaciones:

**Tipo I.** *Sumar* un múltiplo de un vector a *otro* vector. *Ejemplo:* $\big[\boldsymbol{a};\ \boldsymbol{b};\big] \to \big[\boldsymbol{a};\ (\boldsymbol{b}+\lambda\boldsymbol{a});\big]$.

**Tipo II.** *Multiplicar* un vector por un número *distinto de cero*. *Ejemplo:* $\big[\boldsymbol{a};\ \boldsymbol{b};\big] \to \big[\boldsymbol{a};\ (\lambda\boldsymbol{b});\big]$.

Ambas operaciones se denominan *transformaciones elementales.*[1]

## 4.1. Transformaciones y matrices elementales

Denotaremos con "$\boldsymbol{\tau}$" al operador correspondiente a la transformación aplicada. Lo situaremos como subíndice a la derecha de la matriz para indicar que opera sobre las columnas: $\mathbf{A}_{\boldsymbol{\tau}}$. Por debajo de "$\boldsymbol{\tau}$" [*entre corchetes*] indicaremos los detalles de la transformación: $\underset{[detalles]}{\boldsymbol{\tau}}$.

La matriz que resulta de aplicar una sola transformación elemental sobre $\mathbf{I}$ tiene un nombre especial:

**Definición 4.1.** Matriz elemental $\mathbf{I}_{\boldsymbol{\tau}}$ *es cualquier matriz obtenida tras aplicar una sola transformación elemental sobre las columnas de una matriz identidad.*

En esta sección veremos que $\mathbf{A}_{\boldsymbol{\tau}} = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}}\right)$. Es decir, que mediante un producto de matrices se puede aplicar la transformación elemental $\boldsymbol{\tau}$ a las columnas de $\mathbf{A}$. Antes debemos detenernos en las transformaciones (y matrices) elementales de tipo I y II.

### 4.1.1. Transformaciones elementales (y matrices elementales) de Tipo I

**Transf. elemental de Tipo I:** $\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}$ suma $\lambda$ veces el vector $\boldsymbol{i}$-ésimo al vector $\boldsymbol{j}$-ésimo (con $\boldsymbol{j} \neq \boldsymbol{i}$).
(*solo modifica el vector cuyo índice aparece en última posición en el corchete*).

Módulo NAcAL para Python

```
T( (8,2,3) )  # Suma ocho veces el segundo vector al tercero
```

$$\underset{[(8)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}$$

[1] En muchos manuales consideran el intercambio como una tercera transformación elemental, pero aquí no lo haremos.

**Una matriz elemental *Tipo I*** es la obtenida *tras aplicar una sola transformación elemental Tipo I* sobre las columnas de una matriz identidad:

$$\mathbf{I}_{\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}}$$

Por ejemplo, si sobre las columnas de la matriz identidad de orden 3 aplicamos la transformación $\boldsymbol{\tau}_{[(\lambda)\mathbf{2}+\mathbf{3}]}$ (*que suma "$\lambda$ veces" la segunda columna a la tercera*) obtenemos:

$$\mathbf{I}_{\boldsymbol{\tau}_{[(\lambda)\mathbf{2}+\mathbf{3}]}} = \begin{bmatrix} 1 & & \\ & 1 & \lambda \\ & & 1 \end{bmatrix};$$

que es como la matriz identidad de orden 3, salvo porque la segunda componente de su tercera columna ahora es $\lambda$ (en lugar de cero).

Fíjese que la notación $[(\lambda)\,\boldsymbol{i}+\boldsymbol{j}]$ indica que sumamos $\lambda$ veces la $\boldsymbol{i}$-ésima columna a la columna $\boldsymbol{j}$-ésima; pero también describe la matriz elemental correspondiente: en negrita aparecen los índices (fila $\boldsymbol{i}$-ésima, columna $\boldsymbol{j}$-ésima) de la componente de la matriz $\mathbf{I}$ cuyo valor ahora es $(\lambda)$.

Módulo NAcAL para Python

```
I(3) & T( (8,2,3) )  # Transf. de las columnas de la matriz identidad de orden 3
                     # (suma 8 veces la segunda columna a la tercera)
```

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 8 \\ 0 & 0 & 1 \end{bmatrix}$$

Aplicar $\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}$ sobre las columnas de $\mathbf{I}$ solo cambia la $j$-ésima columna. Las columnas de la nueva matriz son

$$\left(\mathbf{I}_{\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}}\right)_{|k} = \begin{cases} \text{cuando } k=j: & \left(\mathbf{I}_{\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}}\right)_{|j} = \lambda\mathbf{I}_{|i} + \mathbf{I}_{|j} & \text{(nueva columna } j\text{-ésima)} \\ \text{cuando } k\neq j: & \left(\mathbf{I}_{\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}}\right)_{|k} = \mathbf{I}_{|k} & \text{(el resto de columnas no cambian)} \end{cases} \tag{4.1}$$

donde $k = 1:n$. Veamos qué pasa si aplicamos una transformación elemental Tipo I sobre $\mathbf{A}$.

**Transformación de las columnas de A:** Si aplicamos $\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}$ a las columnas de $\underset{m\times n}{\mathbf{A}}$ obtenemos una matriz cuyas columnas son:

$$\left(\mathbf{A}_{\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}}\right)_{|k} = \left\{\begin{array}{lll} \lambda\mathbf{A}_{|i} + \mathbf{A}_{|j} & = \mathbf{A}\left(\lambda\mathbf{I}_{|i} + \mathbf{I}_{|j}\right) = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}}\right)_{|j} & \text{para } k=j \\ \mathbf{A}_{|k} & = \mathbf{A}\left(\mathbf{I}_{|k}\right) = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}}\right)_{|k} & \text{para } k\neq j \end{array}\right\} = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}}\right)_{|k},$$

donde $k = 1:n$ y donde hemos usado (4.1).

Por tanto, aplicar sobre las columnas de $\underset{m\times n}{\mathbf{A}}$ la transformación $\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}$ es equivalente a multiplicar $\mathbf{A}$ por la matriz elemental $\mathbf{I}_{\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}}$ de orden $n$ por la derecha,

$$\mathbf{A}_{\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}} = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}}\right).$$

Por ejemplo, si $\mathbf{A} = \begin{bmatrix} 0 & 0 & c \\ 0 & b & c \\ a & b & c \end{bmatrix}$; al sumar "$\lambda$ veces" la segunda columna a la tercera obtenemos:

$$\mathbf{A}_{\underset{[(\lambda)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}} = \begin{bmatrix} 0 & 0 & (\lambda 0 + c) \\ 0 & b & (\lambda b + c) \\ a & b & (\lambda b + c) \end{bmatrix} = \begin{bmatrix} 0 & 0 & c \\ 0 & b & c \\ a & b & c \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \lambda \\ & & 1 \end{bmatrix} = \mathbf{A}\left(\mathbf{I}_{\underset{[(\lambda)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}}\right).$$

### 4.1.2. Transformaciones elementales (y matrices elementales) de Tipo II

**Transf. elemental de Tipo II:** $\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}$ multiplica por $\alpha$ el $\boldsymbol{i}$-ésimo vector (con $\alpha \neq 0$).
(*solo modifica el vector cuyo índice aparece en última posición en el corchete*).

Módulo NAcAL para Python

```
T( (5, 1) )    # multiplica por 5 el primer vector
```

$$\underset{[(5)\mathbf{1}]}{\boldsymbol{\tau}}$$

**Una matriz elemental *Tipo II*** es la obtenida *tras aplicar una sola transformación elemental Tipo II* sobre las columnas de una matriz identidad:

$$\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}$$

Si sobre las columnas de la matriz identidad de orden 3 aplicamos la transformación $\underset{[(\alpha)\mathbf{2}]}{\boldsymbol{\tau}}$ obtenemos:

$$\mathbf{I}_{\underset{[(\alpha)\mathbf{2}]}{\boldsymbol{\tau}}} = \begin{bmatrix} 1 & & \\ & \alpha & \\ & & 1 \end{bmatrix};$$

que es como la matriz $\underset{3\times 3}{\mathbf{I}}$ salvo porque la segunda componente de la diagonal es $\alpha$.

Módulo NAcAL para Python

```
I(3) & T( (5,1) )  # Transf. de las columnas de la matriz identidad de orden 3
                   # (multiplica por 5 la primera columna)
```

$$\begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Aplicar $\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}$ sobre las columnas de $\mathbf{I}$ solo cambia la $i$-ésima columna. Las columnas de la nueva matriz son

$$\left(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\right)_{|k} = \begin{cases} \text{cuando } k = i: & \left(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\right)_{|i} = \alpha \mathbf{I}_{|i} \quad \text{(nueva columna } i\text{-ésima)} \\ \text{cuando } k \neq i: & \left(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\right)_{|k} = \mathbf{I}_{|k} \quad \text{(el resto de columnas no cambian)} \end{cases}, \text{ donde } k = 1:n.$$

Es decir, la matriz $\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}$ es como $\mathbf{I}$ salvo por la $\boldsymbol{i}$-ésima componente de su diagonal, que es $\alpha$.

**Transformación de las columnas de $\mathbf{A}$:** Si aplicamos $\underset{[(\alpha)i]}{\boldsymbol{\tau}}$ a las columnas de $\underset{m\times n}{\mathbf{A}}$ obtenemos una matriz cuyas columnas son:

$$\left(\mathbf{A}\,\underset{[(\alpha)i]}{\boldsymbol{\tau}}\right)_{|k} = \left\{\begin{array}{ll} \alpha\mathbf{A}_{|k} = \mathbf{A}(\alpha\mathbf{I}_{|k}) & \text{si } k=i \\ \\ \mathbf{A}_{|k} = \mathbf{A}(\mathbf{I}_{|k}) & \text{para el resto de columnas} \end{array}\right\} = \mathbf{A}\left(\mathbf{I}\,\underset{[(\alpha)i]}{\boldsymbol{\tau}}\right)_{|k}, \quad \text{donde } k=1:n.$$

De nuevo, aplicar una transformación elemental (ahora de *Tipo II*) sobre las columnas de $\underset{m\times n}{\mathbf{A}}$ es equivalente a multiplicar dicha matriz $\mathbf{A}$ por la correspondiente matriz elemental (de *Tipo II*) de orden $n$ por la derecha.

$$\mathbf{A}\,\underset{[(\alpha)i]}{\boldsymbol{\tau}} = \mathbf{A}\left(\mathbf{I}\,\underset{[(\alpha)i]}{\boldsymbol{\tau}}\right).$$

Por ejemplo, si $\mathbf{A} = \begin{bmatrix} 0 & 0 & c \\ 0 & b & c \\ a & b & c \end{bmatrix}$ entonces al multiplicar la segunda columna por $\alpha$ resulta que:

$$\mathbf{A}\,\underset{[(\alpha)2]}{\boldsymbol{\tau}} = \begin{bmatrix} 0 & 0 & c \\ 0 & \alpha b & c \\ a & \alpha b & c \end{bmatrix} = \begin{bmatrix} 0 & 0 & c \\ 0 & b & c \\ a & b & c \end{bmatrix}\begin{bmatrix} 1 & & \\ & \alpha & \\ & & 1 \end{bmatrix} = \mathbf{A}\left(\mathbf{I}\,\underset{[(\alpha)2]}{\boldsymbol{\tau}}\right).$$

### 4.1.3. Transformación de las columnas de una matriz

Como, al aplicar cualquier transformación elemental $\boldsymbol{\tau}$ sobre las columnas de $\mathbf{A}$, siempre se verifica que:

$$\boxed{\mathbf{A}_{\boldsymbol{\tau}} = \mathbf{A}(\mathbf{I}_{\boldsymbol{\tau}})},$$

disponemos de una nueva regla de reescritura asociativa:

$$\boxed{(\mathbf{AB})_{\boldsymbol{\tau}} \;=\; \mathbf{AB}(\mathbf{I}_{\boldsymbol{\tau}}) \;=\; \mathbf{A}(\mathbf{B}_{\boldsymbol{\tau}})}.$$

Por lo tanto, no son necesarios los paréntesis. Basta escribir: $\mathbf{AB}_{\boldsymbol{\tau}}$

☞ Recuerde que

- **multiplicar** una columna **por cero no es una transformación elemental**;
- **restar una columna a sí misma tampoco lo es**.

### 4.1.4. Sobre las matrices elementales y su notación

Las matrices elementales (de orden $n$) se obtienen al aplicar las correspondientes transformaciones elementales sobre la matriz identidad (de orden $n$). La notación empleada *"casi" describe por completo las correspondientes matrices elementales* (solo falta indicar el orden de la matriz identidad). Así,

- La matriz $\underset{[(-7)1+2]}{\mathbf{I}\,\boldsymbol{\tau}}$ (de orden $n$) es como la matriz identidad (de orden $n$) salvo por la componente de la fila 1 y columna 2, que es un $(-7)$. Por tanto

$$\mathbf{I}\,\underset{[(-7)1+2]}{\boldsymbol{\tau}} = \begin{bmatrix} 1 & -7 \\ 0 & 1 \end{bmatrix}, \quad \text{pero también} \quad \mathbf{I}\,\underset{[(-7)1+2]}{\boldsymbol{\tau}} = \begin{bmatrix} 1 & -7 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{ó} \quad \mathbf{I}\,\underset{[(-7)1+2]}{\boldsymbol{\tau}} = \begin{bmatrix} 1 & -7 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \ldots$$

- La matriz $\mathbf{I}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}}$ es como la matriz identidad (de orden $n$) salvo porque la primera componente de su diagonal es un (3). Por tanto

$$\mathbf{I}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}} = \begin{bmatrix}3\end{bmatrix}; \quad \text{pero también,} \quad \mathbf{I}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}} = \begin{bmatrix}3 & 0\\ 0 & 1\end{bmatrix}; \quad \text{ó} \quad \mathbf{I}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}} = \begin{bmatrix}3 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}, \quad \text{ó}\ldots$$

Fíjese que la matriz identidad es simultáneamente de **Tipo I** y de **Tipo II**: $\mathbf{I}_{\underset{[(0)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}} = \mathbf{I} = \mathbf{I}_{\underset{[(1)\boldsymbol{i}]}{\boldsymbol{\tau}}}$.

**Nota 3.** *Es fácil comprobar que* la transpuesta de una matriz elemental es otra matriz elemental.

- *Para las matrices elementales de Tipo I resulta evidente que* $\left(\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)^{\mathsf{T}} = \mathbf{I}_{\underset{[(\lambda)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}}$,
- *y las matrices elementales de Tipo II son simétricas.*

☞ Solo hay dos tipos de *transformaciones elementales*:

**de Tipo I.** Suman a un vector un múltiplo de *otro* vector: $\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}$

**de Tipo II.** Multiplican un vector por un número *distinto de cero*: $\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}$

Llamamos *matriz elemental* a la resultante de aplicar una sola transformación elemental $\boldsymbol{\tau}$ sobre las columnas de la matriz identidad, $\mathbf{I}_{\boldsymbol{\tau}}$. Solo hay dos tipos:

**de Tipo I.** $\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}$, donde $j \neq i$

**de Tipo II.** $\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}$, donde $\alpha \neq 0$.

Al aplicar una transformación elemental $\boldsymbol{\tau}$ sobre las columnas de $\mathbf{A}$, se verifica que:

$$\boxed{\mathbf{A}_{\boldsymbol{\tau}} = \mathbf{A}(\mathbf{I}_{\boldsymbol{\tau}}).}$$

Consecuentemente, tenemos una nueva propiedad asociativa:

$$\boxed{(\mathbf{AB})_{\boldsymbol{\tau}} = \mathbf{A}(\mathbf{B}_{\boldsymbol{\tau}}) = \mathbf{AB}_{\boldsymbol{\tau}}.}$$

*Además, la transpuesta de una matriz elemental también es elemental* (y del mismo *tipo*).

Ejercicio 27. Escriba las siguientes matrices elementales de orden 4.

(a) $\mathbf{I}_{\underset{[(-7)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}}$ (b) $\mathbf{I}_{\underset{[(1)\mathbf{1}+\mathbf{4}]}{\boldsymbol{\tau}}}$ (c) $\mathbf{I}_{\underset{[(3)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}}$ (d) $\mathbf{I}_{\underset{[(-10)\mathbf{3}]}{\boldsymbol{\tau}}}$

Ejercicio 28. ¿Qué operaciones realizan las matrices anteriores si multiplican por la derecha a $\mathbf{B}$?

Ejercicio 29. Sea la matriz $\mathbf{A} = \begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 1 & 1 & 1\\ 1 & 2 & 3 & 4\end{bmatrix}$, exprese las siguientes transformaciones elementales de $\mathbf{A}$ como productos de $\mathbf{A}$ por matrices elementales y escriba el resultado:

(a) Multiplicar la primera columna por 3. (b) Restar la primera columna de la segunda.

### 4.1.5. Secuencias de transformaciones elementales. Parte I

Considere la aplicación de una secuencia de $k$ transformaciones elementales sobre las columnas de $\mathbf{A}$:

$$\Big(\big(\dots(\mathbf{A}_{\boldsymbol{\tau}_1})_{\boldsymbol{\tau}_2}\dots\big)_{\boldsymbol{\tau}_{(k-1)}}\Big)_{\boldsymbol{\tau}_k}$$

Denotamos la aplicación de dicha secuencia con $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$. Es fácil ver que dicha aplicación es equivalente al producto de $\mathbf{A}$ por una secuencia de $k$ matrices elementales:

$$\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \;=\; \Big(\big(\dots(\mathbf{A}_{\boldsymbol{\tau}_1})_{\boldsymbol{\tau}_2}\dots\big)_{\boldsymbol{\tau}_{(k-1)}}\Big)_{\boldsymbol{\tau}_k} \;=\; \Big(\big(\dots(\mathbf{A}\mathbf{I}_{\boldsymbol{\tau}_1})\mathbf{I}_{\boldsymbol{\tau}_2}\dots\big)\mathbf{I}_{\boldsymbol{\tau}_{(k-1)}}\Big)\mathbf{I}_{\boldsymbol{\tau}_k} \;=\; \mathbf{A}\big(\mathbf{I}_{\boldsymbol{\tau}_1}\big)\cdots\big(\mathbf{I}_{\boldsymbol{\tau}_k}\big).$$

En particular, si $\mathbf{A}=\mathbf{I}$ entonces

$$\boxed{\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \big(\mathbf{I}_{\boldsymbol{\tau}_1}\big)\cdots\big(\mathbf{I}_{\boldsymbol{\tau}_k}\big)}, \tag{4.2}$$

es decir, la matriz $\big(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\big)$ es producto de matrices elementales; y, por tanto,

$$\boxed{\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathbf{A}\big(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\big)}.$$

EJERCICIO 30. Demuestre que la aplicación de una secuencia de transformaciones elementales de las columnas es un *operador lineal*.

☞ Puesto que la aplicación de una secuencia de transformaciones elementales es un producto de matrices:

$$\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \;=\; \mathbf{A}\big(\mathbf{I}_{\boldsymbol{\tau}_1}\big)\cdots\big(\mathbf{I}_{\boldsymbol{\tau}_k}\big) \;=\; \mathbf{A}\Big(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\Big)$$

la aplicación de una secuencia de transformaciones elementales de las columnas es un **operador lineal**.

**Notación** Denotaremos una secuencia de $k$ transformaciones de las *columnas* de $\mathbf{A}$ con el siguiente esquema:

$$\mathbf{A} \xrightarrow{\boldsymbol{\tau}_1} \mathbf{A}_{\boldsymbol{\tau}_1} \xrightarrow{\boldsymbol{\tau}_2} \mathbf{A}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2} \xrightarrow{\boldsymbol{\tau}_3} \cdots \xrightarrow{\boldsymbol{\tau}_k} \mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \quad \text{o agregando varios pasos} \quad \mathbf{A} \xrightarrow[\boldsymbol{\tau}_p]{\overset{\boldsymbol{\tau}_1}{\vdots}} \mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p} \xrightarrow[\boldsymbol{\tau}_k]{\overset{\boldsymbol{\tau}_{p+1}}{\vdots}} \mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

Por ejemplo, la sucesión de transformaciones $\mathbf{A}_{\substack{\boldsymbol{\tau}\\ [(3)\mathbf{2}]}\,\substack{\boldsymbol{\tau}\\ [(2)\mathbf{1}+\mathbf{2}]}}$ donde $\mathbf{A}=\begin{bmatrix}0&1&5\\7&0&4\end{bmatrix}$, se escribe como:

$$\begin{bmatrix}0&1&5\\7&0&4\end{bmatrix} \xrightarrow{\boldsymbol{\tau}_{[(3)\mathbf{2}]}} \begin{bmatrix}0&3&5\\7&0&4\end{bmatrix} \xrightarrow{\boldsymbol{\tau}_{[(2)\mathbf{1}+\mathbf{2}]}} \begin{bmatrix}0&3&5\\7&14&4\end{bmatrix}; \qquad \text{o bien como} \qquad \begin{bmatrix}0&1&5\\7&0&4\end{bmatrix} \xrightarrow[[(2)\mathbf{1}+\mathbf{2}]]{\boldsymbol{\tau}_{[(3)\mathbf{2}]}} \begin{bmatrix}0&3&5\\7&14&4\end{bmatrix}.$$

Alternativamente, podemos expresar una secuencia de transformaciones elementales mediante productos de matrices elementales. Por ejemplo, para la secuencia anterior:

$$\begin{bmatrix}0&1&5\\7&0&4\end{bmatrix}_{\substack{\boldsymbol{\tau}\\ [(3)\mathbf{2}]}\,\substack{\boldsymbol{\tau}\\ [(2)\mathbf{1}+\mathbf{2}]}} = \begin{bmatrix}0&1&5\\7&0&4\end{bmatrix}\Big(\mathbf{I}_{\substack{\boldsymbol{\tau}\\ [(3)\mathbf{2}]}}\Big)\Big(\mathbf{I}_{\substack{\boldsymbol{\tau}\\ [(2)\mathbf{1}+\mathbf{2}]}}\Big) = \begin{bmatrix}0&3&5\\7&14&4\end{bmatrix}.$$

Módulo NAcAL para Python

```
T( (2,3,4) ) & T( (-7,1) ) & T( (-1,3,1) )
```

$$\begin{matrix}\boldsymbol{\tau}\\ [(2)\mathbf{3}{+}\mathbf{4}]\\ [(-7)\mathbf{1}]\\ [(-1)\mathbf{3}{+}\mathbf{1}]\end{matrix}$$

### 4.1.6. Intercambios

Aunque solo hemos considerado como transformaciones elementales las de *Tipo I* y *II*, la mayoría de manuales de Álgebra Lineal consideran como tercer tipo el *intercambio*:

**Intercambio:** $\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}$ → intercambia los vectores $i$-ésimo y $j$-ésimo.

Módulo NAcAL para Python

```
T( {1, 3} )   # intercambia los vectores primero y tercero
```

$$\underset{[\mathbf{1}\rightleftharpoons\mathbf{3}]}{\boldsymbol{\tau}}$$

**Definición 4.2** (Matriz de intercambio)**.** *Cualquier matriz obtenida tras intercambiar las columnas $p$ y $q$ de la matriz identidad de orden $n$:* $\mathbf{I}_{\underset{[p\rightleftharpoons q]}{\boldsymbol{\tau}}}$

Puesto que en $\mathbf{I}_{\underset{[p\rightleftharpoons q]}{\boldsymbol{\tau}}}$ solo han cambiado las columnas $p$ y $q$ de $\mathbf{I}$, se deduce que $\mathbf{I}_{\underset{[p\rightleftharpoons q]}{\boldsymbol{\tau}}}$ es simétrica:

$$\left(\mathbf{I}_{\underset{[p\rightleftharpoons q]}{\boldsymbol{\tau}}}\right)_{|j}=\begin{cases}\mathbf{I}_{|q} & \text{si } j=p\\ \mathbf{I}_{|p} & \text{si } j=q\\ \mathbf{I}_{|j} & \text{si } j\neq p,q\end{cases}\quad\Longrightarrow\quad {}_{k|}\left(\mathbf{I}_{\underset{[p\rightleftharpoons q]}{\boldsymbol{\tau}}}\right)_{|j}=\begin{cases}1 & \text{si}\quad k=q\quad \text{y}\quad j=p\\ 1 & \text{si}\quad k=p\quad \text{y}\quad j=q\\ 1 & \text{si}\quad k=j\neq p,q\\ 0 & \text{en el resto de casos}\end{cases}={}_{j|}\left(\mathbf{I}_{\underset{[p\rightleftharpoons q]}{\boldsymbol{\tau}}}\right)_{|k},$$

es decir, transponiendo (intercambiando los índices $k$ y $j$) la matriz no cambia. Por tanto:

*toda matriz intercambio* $\mathbf{I}_{\underset{[p\rightleftharpoons q]}{\boldsymbol{\tau}}}$ *es simétrica.*

No consideramos el intercambio como una transformación elemental puesto que, en realidad, es una sucesión de transformaciones elementales de **Tipo I** más una transformación de **Tipo II** que multiplica una de las dos columnas intercambiadas por $-1$. El siguiente ejercicio le pide que lo compruebe.

EJERCICIO 31.

(a) Mediante una secuencia de transformaciones elementales Tipo I y II transforme $\mathbf{I}_{3\times 3}$ en la matriz $\mathbf{I}_{\underset{[\mathbf{1}\rightleftharpoons\mathbf{2}]}{\boldsymbol{\tau}}}$.

(b) Mediante el producto de las matrices elementales correspondientes a los pasos dados en el apartado anterior, obtenga la matriz intercambio $\mathbf{I}_{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}}$ de orden $n$.

Módulo NAcAL para Python

```
# Ejemplo de intercambio en una matriz identidad de orden 4
i = 2; j= 4                                  # indicamos quienes son j e i
I(4) & T( [(-1,j),(-1,j,i),(1,i,j),(-1,j,i)] )# aplicación de lista de transf. elem.
# I(4) & T( {i,j} )                          # es más sencillo aplicar intercambio
```

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

☞ $\mathbf{I}_{\underset{[i\rightleftharpoons j]}{\tau}}$ se puede factorizar en un producto de matrices elementales **Tipo I** y una matriz elemental **Tipo II**, (que multiplica por $-1$ una de las columnas intercambiadas) cuando $i \neq j$. Si $i = j$, entonces $\mathbf{I}_{\underset{[i\rightleftharpoons j]}{\tau}} = \mathbf{I}$.

**Intercambio de las columnas de A:**

$$\left(\mathbf{A}_{\underset{[i\rightleftharpoons j]}{\tau}}\right)_{|k} = \left.\begin{cases} \mathbf{A}_{|j} = \mathbf{A}\mathbf{I}_{|j} & \text{si } k = i \text{ (es decir, en la posición } i \text{ se coloca la columna } j) \\ \mathbf{A}_{|i} = \mathbf{A}\mathbf{I}_{|i} & \text{si } k = j \text{ (es decir, en la posición } j \text{ se coloca la columna } i) \\ \mathbf{A}_{|k} = \mathbf{A}\mathbf{I}_{|k} & \text{en el resto de casos se mantienen las columnas en su sitio} \end{cases} \right\} = \mathbf{A}\left(\mathbf{I}_{\underset{[i\rightleftharpoons j]}{\tau}}\right)_{|k}.$$

Si $\mathbf{A} = \begin{bmatrix} a & 0 & 0 \\ b & b & 0 \\ c & c & c \end{bmatrix}$, entonces $\mathbf{A}_{\underset{[1\rightleftharpoons 2]}{\tau}} = \mathbf{A}\left(\mathbf{I}_{\underset{[1\rightleftharpoons 2]}{\tau}}\right) = \begin{bmatrix} a & 0 & 0 \\ b & b & 0 \\ c & c & c \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ intercambia las columnas 1 y 2.

Módulo NAcAL para Python

```
a, b, c, = sympy.symbols('a b c')          # definimos variables simbólicas
A = Matrix([ [a,0,0], [a,b,0], [a,b,c] ]) # definimos Matrix fila a fila
A & T( {1,2} )                             # intercambio de las dos primeras columnas
```

$$\begin{bmatrix} 0 & a & 0 \\ b & a & 0 \\ b & a & c \end{bmatrix}$$

### 4.1.7. Permutaciones

Mediante una sucesión de intercambios se logra un reordenamiento o *permutación* de elementos de un sistema.

**Permutación:** $\underset{[\mathfrak{S}]}{\tau} \rightarrow$ reordena una lista mediante *una sucesión de intercambios.*

**Definición 4.3** (Matriz permutación $\mathbf{I}_{\underset{[\mathfrak{S}]}{\tau}}$)**.** *Cualquier matriz obtenida tras realizar un número arbitrario de intercambios entre las columnas de la matriz identidad.*

Módulo NAcAL para Python

```
I(4) & T( [{1,3}, {4,2}, {3,2}] )  # ejemplo de matriz permutación
```

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Por definición, una matriz permutación se puede factorizar como producto de matrices intercambio. Aunque

las matrices intercambio $\mathbf{I}_{\boldsymbol{\tau}_{[p\rightleftharpoons q]}}$ son simétricas, las matrices permutación $\mathbf{I}_{\boldsymbol{\tau}_{[\mathfrak{S}]}}$ *no* son simétricas en general:[2]

$$\left(\mathbf{I}_{\boldsymbol{\tau}_{[1\rightleftharpoons 3]}}\right)\left(\mathbf{I}_{\boldsymbol{\tau}_{[1\rightleftharpoons 2]}}\right) = \begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix} = \begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}.$$

☞ El intercambio, $\boldsymbol{\tau}_{[i\rightleftharpoons j]}$, es una secuencia de transformaciones *Tipo I* junto a una de *Tipo II* (*que multiplica por* $-1$ *uno de los vectores*). Las matrices intercambio, $\mathbf{I}_{\boldsymbol{\tau}_{[i\rightleftharpoons j]}}$, son simétricas y su cuadrado es la matriz identidad (es decir, "intercambiar dos veces los mismos vectores nos deja como al principio").

Una secuencia de intercambios da lugar a un reordenamiento de los vectores. Llamamos matriz permutación al resultado de aplicar una sucesión de intercambios sobre las columnas de $\mathbf{I}$. En general las matrices permutación $\mathbf{I}_{\boldsymbol{\tau}_{[\mathfrak{S}]}}$ no son simétricas.

## 4.2. Eliminación por columnas

**Definición 4.4.** *Llamamos* pivote *de una columna no nula, a su primer componente no nulo, y* posición de pivote *al índice de la fila en la que se encuentra el pivote.*

**Definición 4.5.** *Una matriz* $\mathbf{K}$ es pre-escalonada *si las componentes a la derecha de cada pivote son nulas.*

*Ejemplo* 9.

$$\mathbf{K} = \begin{bmatrix}0&0&0&0\\0&0&0&7\\0&0&0&6\\3&0&0&0\\6&0&1&0\end{bmatrix}$$

El pivote de la primera columna es 3, el de la tercera es 1, y el de la cuarta 7. La posición de pivote de la primera columna es 4 (cuarta fila), la de la tercera es 5 (quinta fila) y la de la última es 2 (segunda fila).

### 4.2.1. Método de eliminación (o eliminación de "izquierda a derecha")

El método de eliminación consiste en *pre-escalonar* una matriz mediante transformaciones elementales, es decir, en lograr una matriz en la que todas las componentes a la derecha de cada pivote son nulas:

$$\begin{bmatrix}0&1&1\\2&-2&2\\2&-8&1\end{bmatrix} \xrightarrow{\boldsymbol{\tau}_{[(-1)\mathbf{2}+\mathbf{3}]}} \begin{bmatrix}0&1&0\\2&-2&4\\2&-8&9\end{bmatrix} \xrightarrow[\boldsymbol{\tau}_{[(-2)\mathbf{1}+\mathbf{3}]}]{\boldsymbol{\tau}_{[(1)\mathbf{1}+\mathbf{2}]}} \begin{bmatrix}0&1&0\\2&0&0\\2&-6&5\end{bmatrix}$$

Toda matriz se puede *pre-escalonar* mediante transformaciones elementales. *¡Prácticamente el resto del curso descansa sobre este teorema!* Lo demostraremos por inducción: demostrando que si es cierto para cualquier matriz de $n$ columnas, también lo es para cualquier matriz de $n+1$ columnas.

**Teorema 4.2.1** (Eliminación)**.** *Para toda* $\mathbf{A}$ *existe una secuencia* $\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k$ *tal que* $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ *es* pre-escalonada.

*Demostración. Por inducción sobre el número de filas de* $\mathbf{A}$.

Paso de inducción: Supongamos que $\mathbf{A}$ tiene $m$ filas y que el resultado es cierto para matrices de $(m-1)$ filas. Es decir, existe una sucesión $\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k$ de transformaciones elementales de las columnas tal que

$$\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \left[\begin{array}{ccc}&\mathbf{K}&\\\hline b_{m1}&\cdots&b_{mn}\end{array}\right]; \qquad \text{y donde la submatriz } \mathbf{K} \text{ es } \textit{pre-escalonada.}$$

[2] El producto de dos matrices simétricas es simétrico si y solo si dichas matrices conmutan, es decir, si $\mathbf{AB} = \mathbf{BA}$.

Entonces $\mathbf{A}_{\tau_1 \cdots \tau_k}$ también es pre-escalonada si se da alguno de los siguientes *supuestos* en la última fila:

- *Supuesto 1.* La última fila es nula, es decir, $b_{mj} = 0$ para $j = 1 : n$.
- *Supuesto 2.* Son nulos los componentes de la última fila que están bajo las columnas nulas de $\mathbf{K}$ (como en el caso de la segunda columna del Ejemplo 9 en la página anterior).
- *Supuesto 3.* De los componentes de la última fila situados bajo las columnas nulas de $\mathbf{K}$, solo uno es distinto de cero, y las componentes a su derecha son nulas (como en la última fila del Ejemplo 9, donde 1 es el primer componente no nulo de aquellos con una columna de ceros por encima).

Si no se da ninguno de estos supuestos, entonces necesariamente se da el siguiente caso:

- como no se da el supuesto 1, la última fila es no nula.
- como no se da el supuesto 2, la última fila contiene algún pivote.
- como no se da el supuesto 3, a la derecha del primer pivote en la última fila hay coeficientes no nulos.

Si $b_{mr}$ es el primer pivote de los situados en la última fila (es decir, de aquellos componentes situados bajo las columnas nulas de $\mathbf{K}$, es el situado más a la izquierda), podemos *eliminar* los componentes a su derecha restando de cada columna $j$-ésima (con $j \geq r$) la columna $r$-ésima multiplicada por $\left(\frac{b_{mj}}{b_{mr}}\right)$. Así, aplicando la secuencia de transformaciones [3]

$$\left[\left(\tfrac{-b_{mj}}{b_{mr}}\right)\boldsymbol{r}+\boldsymbol{j}\right] \quad \text{para todo } j \geq r,$$

obtenemos una matriz que cumple el *Supuesto 3.*

Base de inducción: Se prueba igual que en el paso de inducción, pero sin $\mathbf{K}$, es decir, con una matriz $\mathbf{A}$ de una sola fila donde $\mathbf{A}_{|r}$ es la primera columna no nula. □

Llamaremos a este Método *"Eliminación de izquierda a derecha"* o sencillamente *"Eliminación"*.

Módulo NAcAL para Python

```
A = Matrix( [ [0,2,2], [1,-2,-8], [1,1,1] ] )
Elim(A,1)                                    # Eliminación por columnas
```

$$\begin{bmatrix} 0 & 2 & 2 \\ 1 & -2 & -8 \\ 1 & 1 & 1 \end{bmatrix} \xrightarrow[[(8)\mathbf{1}+\mathbf{3}]]{\overset{\boldsymbol{\tau}}{[(2)\mathbf{1}+\mathbf{2}]}} \begin{bmatrix} 0 & 2 & 2 \\ 1 & 0 & 0 \\ 1 & 3 & 9 \end{bmatrix} \xrightarrow{\boldsymbol{\tau}\,[(-1)\mathbf{2}+\mathbf{3}]} \begin{bmatrix} 0 & 2 & 0 \\ 1 & 0 & 0 \\ 1 & 3 & 6 \end{bmatrix}$$

El método de eliminación está implementado en NAcAL, pero es importante que usted sepa aplicarlo con lápiz y papel (o mejor aún, que sepa programarlo). [4]

### 4.2.2. Método de eliminación gaussiana

**Definición 4.6.** *Decimos que la matriz* $\mathbf{L}$ es escalonada *si toda columna que precede a una no nula* $\mathbf{L}_{|k}$ *es no nula y su posición de pivote es anterior a la posición de pivote de* $\mathbf{L}_{|k}$.

(Esta definición implica que a la derecha de cualquier pivote solo pueden haber ceros).

En una matriz escalonada, las columnas no nulas están a la izquierda y las nulas a la derecha (si las hubiere). Además, conforme nos movemos de izquierda a derecha, los pivotes aparecen cada vez más abajo, es decir, los pivotes describen una escalera descendente (de ahí el nombre de matriz "escalonada").

[3] que no modifican $\mathbf{K}$ puesto que por encima del pivote $b_{mr}$ todo son ceros
[4] Use el módulo solo para verificar los resultados que usted ha obtenido previamente usando lápiz y papel.

*Ejemplo* 10.

$$\mathbf{L} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 7 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 6 & 1 & 0 \end{bmatrix}$$

El pivote de la primera columna es 7, el de la segunda es 3 y el de la tercera 1. La posición de pivote de la primera columna es 2 (segunda fila), la de la segunda es 4 (cuarta fila) y la de la tercera es 5 (quinta fila).

Fíjese que aunque toda matriz *escalonada* es *pre-escalonada*, las matrices *pre-escalonadas* no suelen ser *escalonadas* (Ejemplo 9 en la página 59).

Toda matriz se puede escalonar mediante transformaciones elementales tal como afirma el siguiente

**Corolario 4.2.2** (Eliminación gaussiana)**.** *Para toda* $\mathbf{A}$ *existen* $\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k$ *tales que* $\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ *es escalonada.*

*Demostración.* Basta pre-escalonar $\mathbf{A}$ y reordenar sus columnas para que los pivotes describan una escalera descendente, dejando las columnas nulas a la derecha de las no nulas. □

Dicho procedimiento se denomina *Método de Eliminación gaussiana*.

Módulo NAcAL para Python

```
A = Matrix( [ [0,2,2], [1,-2,-8], [1,1,1] ] )
L = ElimG(A,1)                                  # Eliminación gaussiana por columnas
```

$$\begin{bmatrix} 0 & 2 & 2 \\ 1 & -2 & -8 \\ 1 & 1 & 1 \end{bmatrix} \xrightarrow[[(8)\mathbf{1}+\mathbf{3}]]{\overset{\boldsymbol{\tau}}{[(2)\mathbf{1}+\mathbf{2}]}} \begin{bmatrix} 0 & 2 & 2 \\ 1 & 0 & 0 \\ 1 & 3 & 9 \end{bmatrix} \xrightarrow{\overset{\boldsymbol{\tau}}{[(-1)\mathbf{2}+\mathbf{3}]}} \begin{bmatrix} 0 & 2 & 0 \\ 1 & 0 & 0 \\ 1 & 3 & 6 \end{bmatrix} \xrightarrow{\overset{\boldsymbol{\tau}}{[\mathbf{1} \rightleftharpoons \mathbf{2}]}} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 1 & 6 \end{bmatrix}$$

### 4.2.3. Método de eliminación Gauss-Jordan

La eliminación gaussiana se puede llevar más lejos, obteniendo una *forma escalonada reducida.*

**Definición 4.7.** *Decimos que una matriz* $\mathbf{R}$ *es* escalonada reducida*, si es escalonada, todos los pivotes son "unos" y los componentes situados a la derecha e izquierda de los pivotes son ceros.*

*Ejemplo* 11.

$$\mathbf{R} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \frac{6}{7} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

**Corolario 4.2.3** (Eliminación Gauss-Jordan)**.** *Para toda* $\mathbf{A}$ *existen* $\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k$ *tales que* $\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ *es escalonada reducida.*

*Demostración.* Basta escalonar la matriz, y con cada pivote eliminar el resto de componentes no nulos de su fila. Después se divide cada columna no nula por el valor de su pivote para lograr los "unos". □

Dicho procedimiento se denomina *Método de Eliminación Gauss-Jordan*.

Módulo NAcAL para Python

```
A = Matrix( [ [0,2,2], [1,-2,-8], [1,1,1] ] )
R = ElimGJ(A,1)                                 # Forma escalonada reducida
```

$$\begin{bmatrix} 0 & 2 & 2 \\ 1 & -2 & -8 \\ 1 & 1 & 1 \end{bmatrix} \xrightarrow[\substack{[(-1)\mathbf{2}+\mathbf{3}]\\ [(2)\mathbf{1}+\mathbf{2}]\\ [(6)\mathbf{1}+\mathbf{3}]}]{\boldsymbol{\tau}} \begin{bmatrix} 0 & 2 & 0 \\ 1 & 0 & 0 \\ 1 & 3 & 6 \end{bmatrix} \xrightarrow[[\mathbf{1}\rightleftharpoons\mathbf{2}]]{\boldsymbol{\tau}} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 1 & 6 \end{bmatrix} \xrightarrow[\substack{[(2)\mathbf{1}]\\ [(-1)\mathbf{3}+\mathbf{1}]\\ [(6)\mathbf{2}]\\ [(-1)\mathbf{3}+\mathbf{2}]}]{\boldsymbol{\tau}} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} \xrightarrow[\substack{[(\frac{1}{4})\mathbf{1}]\\ [(\frac{1}{6})\mathbf{2}]\\ [(\frac{1}{6})\mathbf{3}]}]{\boldsymbol{\tau}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

☞ Cualquier matriz se puede *pre-escalonar*, escalonar o *escalonar y reducir* mediante una secuencia de transformaciones elementales.

- Una matriz $\mathbf{K}$ es *pre-escalonada* si las componentes a la derecha de cada pivote son cero.
- Una matriz $\mathbf{L}$ es *escalonada* si está pre-escalonada y los pivotes describen una escalera descendente (cada pivote está más bajo que el pivote de la columna anterior) y las columnas nulas (si las hubiere) están todas a la derecha de la matriz.
- Una matriz $\mathbf{R}$ es *escalonada reducida* si es escalonada, cada pivote es un "1" y el resto de componentes de su fila son ceros (por ejemplo, la matriz identidad es escalonada reducida).

# *Apéndices a la lección*

## 4.A. Transformaciones elementales de las filas

☞ Casi todos los manuales de Álgebra Lineal aplican las transformaciones elementales principalmente sobre las filas. Aquí usaremos las transformaciones sobre las filas en la Página 80 para demostrar la Proposición 5.2.4, y serán fundamentales para diagonalizar matrices cuadradas tanto por semejanza como por congruencia (Parte VI del libro). Este apéndice trata las transformaciones sobre las filas de una matriz y su relación con las transformaciones sobre las columnas.

La primera observación es que siempre podemos aplicar una transformación elemental a las filas de $\mathbf{A}$ sencillamente transformando las columnas de $\mathbf{A}^\intercal$ y transponiendo de nuevo el resultado. Así concluimos que

$$\big((\mathbf{A}^\intercal)_{\boldsymbol{\tau}}\big)^\intercal = \big((\mathbf{A}^\intercal)\mathbf{I}_{\boldsymbol{\tau}}\big)^\intercal = \big((\mathbf{I}_{\boldsymbol{\tau}})^\intercal\big)\Big(\big(\mathbf{A}^\intercal\big)^\intercal\Big) = \big((\mathbf{I}_{\boldsymbol{\tau}})^\intercal\big)\mathbf{A},$$

donde $\big(\mathbf{I}_{\boldsymbol{\tau}}\big)^\intercal$ es una matriz elemental (por ser la transpuesta de una elemental —Nota 3 en la página 55).

Por tanto, al multiplicar por una matriz elemental *a la izquierda* de $\mathbf{A}$, transformamos las filas de $\mathbf{A}$. Ello sugiere usar la notación ${}_{\boldsymbol{\tau}}\mathbf{A}$ para la transformación $\boldsymbol{\tau}$ de las filas de $\mathbf{A}$;

$${}_{\boldsymbol{\tau}}\mathbf{A} = \big(\mathbf{I}_{\boldsymbol{\tau}}\big)^\intercal\mathbf{A}.$$

Consecuentemente, ${}_{\boldsymbol{\tau}}\mathbf{I}$ denota la transformación $\boldsymbol{\tau}$ de las filas de $\mathbf{I}$; es decir,[5]

$${}_{\boldsymbol{\tau}}\mathbf{I} = \big(\mathbf{I}_{\boldsymbol{\tau}}\big)^\intercal.$$

Así pues,

$${}_{\boldsymbol{\tau}}\mathbf{A} = \big({}_{\boldsymbol{\tau}}\mathbf{I}\big)\mathbf{A}.$$

Además, como el producto de matrices es asociativo y $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathbf{I}_{\boldsymbol{\tau}_1}\cdots\mathbf{I}_{\boldsymbol{\tau}_k}$, deducimos que

$$\big(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\big)^\intercal = \big(\mathbf{I}_{\boldsymbol{\tau}_1}\cdots\mathbf{I}_{\boldsymbol{\tau}_k}\big)^\intercal = \big(\mathbf{I}_{\boldsymbol{\tau}_k}\big)^\intercal\cdots\big(\mathbf{I}_{\boldsymbol{\tau}_1}\big)^\intercal = {}_{\boldsymbol{\tau}_k}\mathbf{I}\cdots{}_{\boldsymbol{\tau}_1}\mathbf{I} = {}_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{I}. \tag{4.3}$$

[5] ¡Fíjese que nuevamente la transposición supone cambiar de lado el subíndice!

Nótese que al transponer no solo cambiamos de lado los subíndices, sino que también invertimos el orden de la secuencia de transformaciones (de manera similar a cuando se transpone un producto de matrices).

Por todo lo anterior, denotaremos con la transposición a la operación de *invertir el orden de una secuencia de transformaciones elementales*:

$$(\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k)^\intercal = \boldsymbol{\tau}_k \cdots \boldsymbol{\tau}_1. \tag{4.4}$$

Módulo NAcAL para Python

```
Tr = T( [(66,1,2), {3,2}, (4,1)], rpr='h' )
Tr
```

$$\underset{[(66)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{[\mathbf{2}\rightleftharpoons\mathbf{3}]}{\boldsymbol{\tau}}\underset{[(4)\mathbf{1}]}{\boldsymbol{\tau}}$$

Módulo NAcAL para Python

```
~Tr                    # La transposición invierte el orden
```

$$\underset{[(4)\mathbf{1}]}{\boldsymbol{\tau}}\underset{[\mathbf{2}\rightleftharpoons\mathbf{3}]}{\boldsymbol{\tau}}\underset{[(66)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}$$

Módulo NAcAL para Python

```
I(3) & Tr
```

$$\begin{bmatrix} 4 & 0 & 66 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Módulo NAcAL para Python

```
(~Tr) & I(3)          # transpuesta de la matriz anterior
```

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 1 \\ 66 & 1 & 0 \end{bmatrix}$$

**Matrices simétricas.** Por (4.3) y (4.4) sabemos que $\left(\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}\right)^\intercal = {}_{\boldsymbol{\tau}_k \cdots \boldsymbol{\tau}_1}\mathbf{I} = {}_{(\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k)^\intercal}\mathbf{I}$. Consecuentemente, la matriz $\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ es simétrica si y solo si

$${}_{(\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k)^\intercal}\mathbf{I} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}. \tag{4.5}$$

Módulo NAcAL para Python

```
p = T([(9,2), (18,1), (-1,2,1), (fracc(1,9),1), (-1,1,2), (fracc(1,2),2)], rpr='h')
p
```

$$\underset{[(9)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(18)\mathbf{1}]}{\boldsymbol{\tau}}\underset{[(-1)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}\underset{\left[\left(\frac{1}{9}\right)\mathbf{1}\right]}{\boldsymbol{\tau}}\underset{[(-1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{\left[\left(\frac{1}{2}\right)\mathbf{2}\right]}{\boldsymbol{\tau}}$$

Módulo NAcAL para Python

```
(~p & I(2)) == (I(2) & p) # matriz simétrica si se verifica la igualdad
```

```
True
```

Módulo NAcAL para Python

```
I(2) & p # veamos que efectivamente la matriz es simétrica
```

$$\begin{bmatrix} 2 & -1 \\ -1 & 5 \end{bmatrix}$$

**Una secuencia de transformaciones elementales se comporta casi como un producto de matrices:** fíjese que aunque la secuencia $\tau_1 \cdots \tau_k$ no es un producto de matrices… ¡casi se puede interpretar como tal!… Tan solo falta que las transformaciones actúen sobre una matriz para saber el orden $n$ de las matrices elementales, y entonces sí actúan como un producto de matrices elementales. De tal manera que al igual que $\mathbf{A}\mathbf{B}^\intercal = (\mathbf{B}^\intercal)\,(\mathbf{A}^\intercal)$, también

$$\big(\mathbf{A}_{\tau_1\cdots\tau_k}\big)^\intercal \;=\; \big(\mathbf{A}(\mathbf{I}_{\tau_1\cdots\tau_k})\big)^\intercal \;=\; \big({}_{\tau_k\cdots\tau_1}\mathbf{I}\big)(\mathbf{A}^\intercal) \;=\; {}_{(\tau_k\cdots\tau_1)}(\mathbf{A}^\intercal) \;=\; {}_{(\tau_1\cdots\tau_k)^\intercal}(\mathbf{A}^\intercal).$$

Por (4.3) y (4.4) sabemos que

$${}_{\tau_k\cdots\tau_1}\mathbf{I} = {}_{(\tau_1\cdots\tau_k)^\intercal}\mathbf{I} = \big(\mathbf{I}_{\tau_1\cdots\tau_k}\big)^\intercal.$$

No obstante, en general tanto

$${}_{\tau_1\cdots\tau_k}\mathbf{I} \neq \mathbf{I}_{\tau_1\cdots\tau_k} \qquad \text{como} \qquad {}_{\tau_k\cdots\tau_1}\mathbf{I} \neq \mathbf{I}_{\tau_1\cdots\tau_k};$$

salvo cuando $\mathbf{I}_{\tau_1\cdots\tau_k}$ es simétrica, en cuyo caso ya hemos visto que ${}_{\tau_k\cdots\tau_1}\mathbf{I} = \mathbf{I}_{\tau_1\cdots\tau_k}$.
(fíjese en el orden de las transformaciones tanto en las igualdades como en las desigualdades).

Por último, nótese que incluso cuando operamos con las filas, la notación sigue siendo asociativa:

$$\boxed{\big({}_{\tau_1\cdots\tau_k}\mathbf{A}\big)\mathbf{B} = {}_{\tau_1\cdots\tau_k}(\mathbf{A}\mathbf{B})}.$$

Así que podemos escribir sencillamente ${}_{\tau_1\cdots\tau_k}\mathbf{A}\mathbf{B}$.

**Notación para una secuencia de transformaciones sobre las filas de una matriz.** De igual modo que una secuencia de transformaciones de las columnas se puede expresar de varias formas, también disponemos de representaciones alternativas para una secuencia de transformaciones de las *filas* de una matriz. Por ejemplo:

$${}_{\tau_{[(4)\mathbf{1}]}\tau_{[\mathbf{2}\rightleftharpoons\mathbf{3}]}\tau_{[(66)\mathbf{1}+\mathbf{2}]}}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = {}_{\substack{\tau_{[(4)\mathbf{1}]}\\ \tau_{[\mathbf{2}\rightleftharpoons\mathbf{3}]}\\ \tau_{[(66)\mathbf{1}+\mathbf{2}]}}}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 1 \\ 66 & 1 & 0 \end{bmatrix}.$$

es decir, ${}_{\tau_{[(4)\mathbf{1}]}\tau_{[\mathbf{2}\rightleftharpoons\mathbf{3}]}\tau_{[(66)\mathbf{1}+\mathbf{2}]}}\mathbf{A} = \big({}_{\tau_{[(4)\mathbf{1}]}}\mathbf{I}\big)\big({}_{\tau_{[\mathbf{2}\rightleftharpoons\mathbf{3}]}}\mathbf{I}\big)\big({}_{\tau_{[(66)\mathbf{1}+\mathbf{2}]}}\mathbf{I}\big)\mathbf{A} = {}_{\substack{\tau_{[(4)\mathbf{1}]}\\ \tau_{[\mathbf{2}\rightleftharpoons\mathbf{3}]}\\ \tau_{[(66)\mathbf{1}+\mathbf{2}]}}}\mathbf{A}.$

Al operar sobre las filas, las primeras transformaciones en actuar son las que quedan más a la derecha de ${}_{\tau_{[(4)\mathbf{1}]}\tau_{[\mathbf{2}\rightleftharpoons\mathbf{3}]}\tau_{[(66)\mathbf{1}+\mathbf{2}]}}$ o más abajo de ${}_{\substack{\tau_{[(4)\mathbf{1}]}\\ \tau_{[\mathbf{2}\rightleftharpoons\mathbf{3}]}\\ \tau_{[(66)\mathbf{1}+\mathbf{2}]}}}$.

Además, para indicar que una secuencia de transformaciones actúa sobre las *filas* de una matriz, escribiremos las abreviaturas por debajo de las fechas. Por tanto, los pasos en el esquema de la transformación ${}_{\tau_{[(4)\mathbf{1}]}\tau_{[\mathbf{2}\rightleftharpoons\mathbf{3}]}\tau_{[(66)\mathbf{1}+\mathbf{2}]}}\mathbf{I}$ son los siguientes:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[\tau_{[(66)\mathbf{1}+\mathbf{2}]}]{} \begin{bmatrix} 1 & 0 & 0 \\ 66 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[\tau_{[\mathbf{2}\rightleftharpoons\mathbf{3}]}]{} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 66 & 1 & 0 \end{bmatrix} \xrightarrow[\tau_{[(4)\mathbf{1}]}]{} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 1 \\ 66 & 1 & 0 \end{bmatrix}.$$

### 4.A.1. Matriz escalonada por filas

**Definición 4.8.** *Una matriz está* escalonada por filas *si su transpuesta está escalonada por columnas.*

Módulo NAcAL para Python

```
A = Matrix([ [0,2,2], [1,-2,-8], [3,6,-6] ])
U = ElimGF(A,1)                # Escalonamiento operando con las filas
```

$$\begin{bmatrix} 0 & 2 & 2 \\ 1 & -2 & -8 \\ 3 & 6 & -6 \end{bmatrix} \xrightarrow[\substack{[(-3)\mathbf{1}+\mathbf{3}]\\[(1)\mathbf{1}+\mathbf{2}]}]{\boldsymbol{\tau}} \begin{bmatrix} 0 & 2 & 2 \\ 1 & 0 & -6 \\ 3 & 0 & -12 \end{bmatrix} \xrightarrow[[(-3)\mathbf{2}+\mathbf{3}]]{\boldsymbol{\tau}} \begin{bmatrix} 0 & 2 & 2 \\ 1 & 0 & -6 \\ 0 & 0 & 6 \end{bmatrix} \xrightarrow[[\mathbf{1}\rightleftharpoons\mathbf{2}]]{\boldsymbol{\tau}} \begin{bmatrix} 1 & 0 & -6 \\ 0 & 2 & 2 \\ 0 & 0 & 6 \end{bmatrix}$$

(recuerde que en el primer paso primero actúa $\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}$ y luego $\underset{[(-3)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}$)

## 4.B. Transformación elemental "*espejo*" de otra transformación

☞ Usaremos las transformaciones *"espejo"* en la Lección 17 cuando veamos *matrices semejantes*. Si lo prefiere, puede posponer la lectura de esta sección hasta ese momento.

Resulta que podemos obtener una misma matriz elemental operando sobre las filas o sobre las columnas de la matriz identidad. Si la matriz elemental es de *Tipo I*, por ejemplo $\mathbf{I}_{\underset{[(\alpha)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}}$, es fácil verificar que

$$\underset{[(\alpha)\mathbf{3}+\mathbf{2}]}{\boldsymbol{\tau}}\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \alpha \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}\underset{[(\alpha)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}.$$

En general, para una misma matriz elemental de *Tipo I* tenemos que

$$\underset{[(\alpha)\boldsymbol{k}+\boldsymbol{j}]}{\boldsymbol{\tau}}\mathbf{I} = \mathbf{I}_{\underset{[(\alpha)\boldsymbol{j}+\boldsymbol{k}]}{\boldsymbol{\tau}}}.$$

Fíjese en el *intercambio entre los índices* $\boldsymbol{j}$ y $\boldsymbol{k}$ en función de si se transforman las filas o las columnas de la matriz identidad. Es decir, la transformación elemental necesaria para crear una matriz elemental de *Tipo I* es distinta en función de si actuamos sobre las filas o sobre las columnas de $\mathbf{I}$.

A falta de mejor nombre, llamaremos a cada una de estas transformaciones el "*espejo*" de la otra:

**Definición 4.9.** *Llamamos* espejo *de* $\boldsymbol{\tau}$ *(que denotamos con "*$\mathrm{esp}(\boldsymbol{\tau})$*") a aquella transformación elemental que actuando por el otro lado de la matriz identidad arroja el mismo resultado; es decir*

$$_{\mathrm{esp}(\boldsymbol{\tau})}\mathbf{I} = \mathbf{I}_{\boldsymbol{\tau}} \qquad \text{ó} \qquad {}_{\boldsymbol{\tau}}\mathbf{I} = \mathbf{I}_{\mathrm{esp}(\boldsymbol{\tau})}.$$

El *espejo* de una transformación elemental *Tipo I* $\underset{[(\alpha)\boldsymbol{j}+\boldsymbol{k}]}{\boldsymbol{\tau}}$ es la transformación que resulta de intercambiar los índices $j$ y $k$, es decir,

$$\mathrm{esp}\left(\underset{[(\alpha)\boldsymbol{j}+\boldsymbol{k}]}{\boldsymbol{\tau}}\right) = \underset{[(\alpha)\boldsymbol{k}+\boldsymbol{j}]}{\boldsymbol{\tau}}.$$

En cuanto a las transformaciones elementales *Tipo II* y los *intercambios...* ¡son su propio "*espejo*"! pues tanto las matrices elementales *Tipo II* como las matrices intercambio son simétricas ($\mathbf{I}_{\boldsymbol{\tau}} = {}_{\boldsymbol{\tau}}\mathbf{I} = {}_{\mathrm{esp}(\boldsymbol{\tau})}\mathbf{I}$). Por ejemplo

$$\underset{[(\alpha)\mathbf{2}]}{\boldsymbol{\tau}}\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & \alpha & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}\underset{[(\alpha)\mathbf{2}]}{\boldsymbol{\tau}} \quad \text{y} \quad \underset{[\mathbf{2}\rightleftharpoons\mathbf{3}]}{\boldsymbol{\tau}}\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}\underset{[\mathbf{2}\rightleftharpoons\mathbf{3}]}{\boldsymbol{\tau}}.$$

Además, el *espejo* de una sucesión de transformaciones elementales es la sucesión de transformaciones espejo:

$$\mathrm{esp}(\boldsymbol{\tau}_1)\ldots\mathrm{esp}(\boldsymbol{\tau}_k) \;=\; \mathrm{esp}\big(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k\big);$$

es decir

$$\mathrm{esp}(\boldsymbol{\tau}_1)\ldots\mathrm{esp}(\boldsymbol{\tau}_k)\mathbf{I} \;=\; \mathrm{esp}(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)\mathbf{I} \;=\; \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

Módulo NAcAL para Python

```
Tr = T((3,2,1)) & T((-2,3,2)) & T((7,1,3)) & T({2,3}) & T((20,1)) # Suc. transf.
A  = I(3) & Tr            # Transformación de las columnas de la Identidad
B  = Tr.espejo() & I(3)    # Transformación espejo sobre las filas de la Identidad
A == B                     # Verificación de que A y B son iguales
```

```
True
```

**Matrices simétricas.** Por la Ecuación 4.5 en la página 63, la matriz $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ es simétrica si y solo si

$$\mathrm{esp}(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)\mathbf{I} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = (\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{\intercal}\mathbf{I}.$$

Módulo NAcAL para Python

```
p = T([(9,2), (18,1), (-1,2,1), (fracc(1,9),1), (-1,1,2), (fracc(1,2),2)], rpr='h')
p
```

$$\underset{[(9)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(18)\mathbf{1}]}{\boldsymbol{\tau}}\underset{[(-1)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}\underset{\left[\left(\frac{1}{9}\right)\mathbf{1}\right]}{\boldsymbol{\tau}}\underset{[(-1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{\left[\left(\frac{1}{2}\right)\mathbf{2}\right]}{\boldsymbol{\tau}}$$

Módulo NAcAL para Python

```
(p.espejo() & I(2)) == (~p & I(2)) # matriz simétrica si se verifica la igualdad
```

```
True
```

Módulo NAcAL para Python

```
p.espejo() & I(2) # veamos que efectivamente la matriz es simétrica
```

$$\begin{bmatrix} 2 & -1 \\ -1 & 5 \end{bmatrix}$$

Aun cuando $\mathrm{esp}(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)\mathbf{I} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = (\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{\intercal}\mathbf{I}$ (por ser $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ simétrica), las tres sucesiones de transformaciones son en general distintas. En el ejemplo de más arriba: la sucesión $\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k$, la sucesión $\mathrm{esp}(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)$ y la sucesión $(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{\intercal}$ son respectivamente

$$\begin{matrix}\boldsymbol{\tau}\\ [(9)\mathbf{2}]\\ [(18)\mathbf{1}]\\ [(-1)\mathbf{2}+\mathbf{1}]\\ \left[\left(\frac{1}{9}\right)\mathbf{1}\right]\\ [(-1)\mathbf{1}+\mathbf{2}]\\ \left[\left(\frac{1}{2}\right)\mathbf{2}\right]\end{matrix}\quad,\quad\begin{matrix}\boldsymbol{\tau}\\ [(9)\mathbf{2}]\\ [(18)\mathbf{1}]\\ [(-1)\mathbf{1}+\mathbf{2}]\\ \left[\left(\frac{1}{9}\right)\mathbf{1}\right]\\ [(-1)\mathbf{2}+\mathbf{1}]\\ \left[\left(\frac{1}{2}\right)\mathbf{2}\right]\end{matrix}\quad\text{y}\quad\begin{matrix}\boldsymbol{\tau}\\ \left[\left(\frac{1}{2}\right)\mathbf{2}\right]\\ [(-1)\mathbf{1}+\mathbf{2}]\\ \left[\left(\frac{1}{9}\right)\mathbf{1}\right]\\ [(-1)\mathbf{2}+\mathbf{1}]\\ [(18)\mathbf{1}]\\ [(9)\mathbf{2}]\end{matrix}\quad.$$

(si se fija en los índices de las transformaciones elementales de Tipo I de las dos primeras sucesiones, constatará que las tres sucesiones son distintas).

***Espejo* de la inversa de una transformación elemental.** Cuando una transformación elemental deshace la operación realizada por otra, se dice de ambas que *"una es la inversa de la otra"* [6].

Por ejemplo, si sumamos el triple del vector $\boldsymbol{x}$ al vector $\boldsymbol{y}$, podemos deshacer (o invertir) la operación restando el triple de $\boldsymbol{x}$ a $\boldsymbol{y}$. Así, la transformación inversa de $\underset{[(\alpha)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}$ es $\underset{[(-\alpha)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}$.

De manera similar, si multiplicamos por 2 un vector, podemos deshacer (o invertir) la operación dividiendo dicho vector por 2. Así tenemos que la transformación inversa de $\underset{[(\alpha)\boldsymbol{j}]}{\boldsymbol{\tau}}$ es $\underset{\left[\left(\frac{1}{\alpha}\right)\boldsymbol{j}\right]}{\boldsymbol{\tau}}$.

Consecuentemente, empleando las transformaciones espejo concluimos que

$$\underset{[(-\alpha)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}\mathbf{I}\underset{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}} = \mathbf{I} \qquad \text{y que} \qquad \underset{\left[\left(\frac{1}{\alpha}\right)\boldsymbol{i}\right]}{\boldsymbol{\tau}}\mathbf{I}\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}, = \mathbf{I};$$

así, si denotamos con $\boldsymbol{\tau}^{-1}$ la *inversa* de $\boldsymbol{\tau}$, tenemos que

$$\operatorname{esp}(\boldsymbol{\tau})^{-1}\mathbf{I}_{\boldsymbol{\tau}} = \operatorname{esp}(\boldsymbol{\tau}^{-1})\mathbf{I}_{\boldsymbol{\tau}} = \mathbf{I};$$

y en general, para una secuencia de transformaciones elementales[7]

$$\operatorname{esp}(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{-1}\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathbf{I}.$$

Módulo NAcAL para Python

```
Tr = T((3,2,1)) & T((-2,3,2)) & T((7,1,3)) & T({2,3}) & T((20,1))
(Tr**-1).espejo() & I(3) & Tr
```

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

☞ Veremos que la Lección 17 se basa en aplicar, a matrices *cuadradas*, transformaciones de la forma

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)^{-1}\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \operatorname{esp}(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{-1}\mathbf{I}\mathbf{A}\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \operatorname{esp}(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{-1}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k},$$

donde $\mathbf{S} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$.

La Lección 19 se basa en aplicar, a matrices *simétricas*, transformaciones de la forma

$$\mathbf{B}^{\intercal}\mathbf{A}\mathbf{B} = \left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)^{\intercal}\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = (\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{\intercal}\mathbf{I}\mathbf{A}\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = (\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{\intercal}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k},$$

donde $\mathbf{B} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$.

La Lección 18 es un caso especial en el que se aplican, a matrices *simétricas*, transformaciones de la forma

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \operatorname{esp}(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{-1}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = (\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{\intercal}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$$

donde $\mathbf{Q} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ es tal que $\mathbf{Q}^{-1} = \mathbf{Q}^{\intercal}$, es decir, $\mathbf{Q}^{-1} = \mathbf{I}_{(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{-1}} = \operatorname{esp}(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{-1}\mathbf{I} = (\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{\intercal}\mathbf{I} = \mathbf{Q}^{\intercal}$; por lo que es un caso particular de lo tratado tanto en la Lección 17 como en la Lección 19.

[6] La Lección 5 trata sobre este tema.

[7] véase la Sección 5.A en la página 82 la justificación para la notación: $\operatorname{esp}(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{-1}$.

# Inversa de una matriz. Rango de una matriz

**Matrices invertibles por un lado.** Decimos que $\mathbf{A}$ es *invertible por la derecha* si existe otra matriz $\mathbf{B}$ tal que $\mathbf{AB} = \mathbf{I}$. Y decimos que $\mathbf{A}$ es *invertible por la izquierda* si existe otra matriz $\mathbf{B}$ tal que $\mathbf{BA} = \mathbf{I}$. Esta lección se centra en las matrices *cuadradas* e invertibles por ambos lados a la vez...

## 5.1. Matrices invertibles

**Definición 5.1** (Matriz invertible)**.** *Se dice que* $\mathbf{A}$ *es* invertible (o tiene inversa) *si existe otra matriz* $\mathbf{B}$ *del mismo orden tal que*

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}.$$

(nótese que $\mathbf{AB} = \mathbf{BA}$ implica que $\mathbf{A}$ y $\mathbf{B}$ son necesariamente cuadradas).

Si $\mathbf{A}$ es invertible, su inversa es única:

**Proposición 5.1.1** (La inversa es única)**.** *Si* $\mathbf{A}$, $\mathbf{B}$ *y* $\mathbf{B'}$ *verifican que*

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I} \qquad \text{y que} \qquad \mathbf{AB'} = \mathbf{B'A} = \mathbf{I}$$

*entonces* $\mathbf{B}$ *y* $\mathbf{B'}$ *son iguales.*

*Demostración.* En efecto: $\mathbf{B} = \mathbf{BI} = \mathbf{B}\underbrace{(\mathbf{AB'})}_{\mathbf{I}=\mathbf{AB'}} = \underbrace{(\mathbf{BA})}_{\mathbf{BA}=\mathbf{I}}\mathbf{B'} = \mathbf{IB'} = \mathbf{B'}$. □

La unicidad nos permite dar una definición de *"la" matriz inversa de* $\mathbf{A}$:

**Definición 5.2** (Matriz inversa)**.** *Si* $\mathbf{A}$ *es invertible, denotaremos con* $\mathbf{A}^{-1}$ *a la única matriz tal que*

$$\mathbf{A}(\mathbf{A}^{-1}) = (\mathbf{A}^{-1})\mathbf{A} = \mathbf{I}.$$

*Además, diremos que* $\mathbf{A}^{-1}$ *es la matriz inversa de* $\mathbf{A}$.

**Definición 5.3** (Matriz singular)**.** *Si una matriz cuadrada* no tiene inversa *se dice que es* singular.

A continuación, veremos varias propiedades de las matrices inversas. En particular, que el *producto de matrices invertibles* es invertible; que la transpuesta de una *matriz invertible* también es invertible; y que si una matriz tiene alguna fila o columna nula, entonces no es invertible.

Ejercicio 32. Demuestre las siguientes proposiciones:

(a) **Proposición 5.1.2.** *Si* $\mathbf{A}$ *y* $\mathbf{B}$ *(de orden n) tienen inversa, entonces:* $\boxed{(\mathbf{AB})^{-1} = (\mathbf{B}^{-1})(\mathbf{A}^{-1}).}$

(b) **Proposición 5.1.3.** *Si* $\mathbf{A}_k, \ldots, \mathbf{A}_n$ *tienen inversa, entonces* $\boxed{(\mathbf{A}_1 \cdots \mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1} \cdots \mathbf{A}_1^{-1}.}$

*Pista.* $(\mathbf{A}_1 \cdots \mathbf{A}_k)^{-1} = (\mathbf{A}_1(\mathbf{A}_2 \cdots \mathbf{A}_k))^{-1} = (\mathbf{A}_2 \cdots \mathbf{A}_k)^{-1}(\mathbf{A}_1^{-1})$.

(c) **Proposición 5.1.4.** *Sea* $\mathbf{A}$ *de orden* $n$ *y* $\mathbf{B}$ *invertible de orden* $n$. *Entonces* $\mathbf{AB}$ *es invertible si y solo si* $\mathbf{A}$ *es invertible.*

(d) **Proposición 5.1.5.** *Sea* $\mathbf{A}$ *invertible de orden* $n$ *y* $\mathbf{B}$ *de orden* $n$. *Entonces* $\mathbf{AB}$ *es invertible si y solo si* $\mathbf{B}$ *es invertible.*

(e) **Proposición 5.1.6.** *Si* $\mathbf{A}$ *de orden* $n$ *tiene inversa, entonces la inversa de* $\mathbf{A}^\intercal$ *es:* $\boxed{\left(\mathbf{A}^\intercal\right)^{-1} = (\mathbf{A}^{-1})^\intercal.}$

*Pista.* Recuerde que $\mathbf{B}^\intercal\mathbf{A}^\intercal = \left(\mathbf{AB}\right)^\intercal$ y aplíquelo a $\mathbf{A}^\intercal(\mathbf{A}^{-1})^\intercal$ y a $(\mathbf{A}^{-1})^\intercal\mathbf{A}^\intercal$.

(f) **Proposición 5.1.7.** *Si alguna fila o columna de* $\mathbf{A}$ *es nula, entonces la matriz es* singular.

☞ $\mathbf{A}$ es invertible si existe $\mathbf{B}$ tal que $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$; su inversa $\mathbf{B}$ es única y se denota con $\mathbf{A}^{-1}$.

- El producto de matrices invertibles es invertible: $\left(\mathbf{AB}\right)^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
- La transpuesta de una matriz invertible es invertible: $\left(\mathbf{A}^\intercal\right)^{-1} = (\mathbf{A}^{-1})^\intercal$.
- Si alguna fila o columna de $\mathbf{A}$ es nula, la matriz es *singular* (no tiene inversa).

### 5.1.1. Inversa de las matrices (y las transformaciones) elementales.

☞ ¡No todas las matrices cuadradas son invertibles!

¿Qué hace que una matriz cuadrada sea invertible?... Todo tiene que ver con las matrices elementales. Veamos primero que las transformaciones elementales (y las matrices elementales) son invertibles.

Las transformaciones elementales son reversibles. Por ejemplo, sumar el doble de la primera columna a la tercera, y luego restar el doble de la primera columna de la tercera, nos deja como al principio:

$$\mathbf{A}_{\overset{\tau}{[(2)1+3]}\overset{\tau}{[(-2)1+3]}} = \mathbf{A}\left(\mathbf{I}_{\overset{\tau}{[(2)1+3]}}\right)\left(\mathbf{I}_{\overset{\tau}{[(-2)1+3]}}\right) = \mathbf{A}\begin{bmatrix}1&0&2\\0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&-2\\0&1&0\\0&0&1\end{bmatrix} = \mathbf{A}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} = \mathbf{A}.$$

Algo similar ocurre cuando multiplicamos y dividimos una columna por $\alpha \neq 0$:

$$\mathbf{A}_{\overset{\tau}{[(5)2]}\overset{\tau}{\left[\left(\frac{1}{5}\right)2\right]}} = \mathbf{A}\left(\mathbf{I}_{\overset{\tau}{[(5)2]}}\right)\left(\mathbf{I}_{\overset{\tau}{\left[\left(\frac{1}{5}\right)2\right]}}\right) = \mathbf{A}\begin{bmatrix}1&0&0\\0&5&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&\frac{1}{5}&0\\0&0&1\end{bmatrix} = \mathbf{A}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} = \mathbf{A}$$

En general,

$$\mathbf{A}_{\overset{\tau}{[(\lambda)i+j]}\overset{\tau}{[(-\lambda)i+j]}} = \mathbf{A}_{\overset{\tau}{[(-\lambda)i+j]}\overset{\tau}{[(\lambda)i+j]}} = \mathbf{A} \qquad \text{y} \qquad \mathbf{A}_{\overset{\tau}{[(\alpha)i]}\overset{\tau}{\left[\left(\frac{1}{\alpha}\right)i\right]}} = \mathbf{A}_{\overset{\tau}{\left[\left(\frac{1}{\alpha}\right)i\right]}\overset{\tau}{[(\alpha)i]}} = \mathbf{A},$$

es decir, para toda transformación elemental $\boldsymbol{\tau}$ existe otra, que denotamos $\boldsymbol{\tau}^{-1}$, tal que

$$\mathbf{A}_{(\boldsymbol{\tau}^{-1})\boldsymbol{\tau}} = \mathbf{A}_{\boldsymbol{\tau}(\boldsymbol{\tau}^{-1})} = \mathbf{A}. \tag{5.1}$$

Recordando que $\mathbf{A}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2} = \left(\mathbf{A}_{\boldsymbol{\tau}_1}\right)_{\boldsymbol{\tau}_2} = \left(\mathbf{A}_{\boldsymbol{\tau}_1}\right)\mathbf{I}_{\boldsymbol{\tau}_2}$ podemos expresar (5.1) como

$$\left(\mathbf{A}_{\boldsymbol{\tau}^{-1}}\right)\mathbf{I}_{\boldsymbol{\tau}} = \left(\mathbf{A}_{\boldsymbol{\tau}}\right)\mathbf{I}_{\boldsymbol{\tau}^{-1}} = \mathbf{A};$$

si en particular $\mathbf{A} = \mathbf{I}$ tenemos que

$$\left(\mathbf{I}_{\boldsymbol{\tau}^{-1}}\right)\mathbf{I}_{\boldsymbol{\tau}} = \left(\mathbf{I}_{\boldsymbol{\tau}}\right)\mathbf{I}_{\boldsymbol{\tau}^{-1}} = \mathbf{I} \qquad \Longrightarrow \qquad \left(\mathbf{I}_{\boldsymbol{\tau}^{-1}}\right) = \left(\mathbf{I}_{\boldsymbol{\tau}}\right)^{-1}.$$

Así pues, llegamos al siguiente

**Teorema 5.1.8.** *Para toda matriz elemental* $\mathbf{I}_{\boldsymbol{\tau}}$*, existe otra matriz elemental* $\mathbf{I}_{\boldsymbol{\tau}^{-1}}$*, tal que*

$$\left(\mathbf{I}_{\boldsymbol{\tau}}\right)\left(\mathbf{I}_{\boldsymbol{\tau}^{-1}}\right) = \left(\mathbf{I}_{\boldsymbol{\tau}^{-1}}\right)\left(\mathbf{I}_{\boldsymbol{\tau}}\right) = \mathbf{I}.$$

En particular, si $i \neq j$ y $\alpha \neq 0$:

$$\boxed{\left(\mathbf{I}_{\underset{[(-\lambda)i+j]}{\boldsymbol{\tau}}}\right)\left(\mathbf{I}_{\underset{[(\lambda)i+j]}{\boldsymbol{\tau}}}\right) = \left(\mathbf{I}_{\underset{[(\lambda)i+j]}{\boldsymbol{\tau}}}\right)\left(\mathbf{I}_{\underset{[(-\lambda)i+j]}{\boldsymbol{\tau}}}\right) = \mathbf{I}} \qquad \text{y} \qquad \boxed{\left(\mathbf{I}_{\underset{\left[\left(\frac{1}{\alpha}\right)i\right]}{\boldsymbol{\tau}}}\right)\left(\mathbf{I}_{\underset{[(\alpha)i]}{\boldsymbol{\tau}}}\right) = \left(\mathbf{I}_{\underset{[(\alpha)i]}{\boldsymbol{\tau}}}\right)\left(\mathbf{I}_{\underset{\left[\left(\frac{1}{\alpha}\right)i\right]}{\boldsymbol{\tau}}}\right) = \mathbf{I}}.$$

Como toda transformación elemental es invertible, también lo es cualquier sucesión de transformaciones elementales (por ejemplo, un intercambio). Es decir, se verifica la siguiente

**Proposición 5.1.9.** *Si* $\mathbf{A}$ *es producto de matrices elementales, entonces existe* $\mathbf{B}$ *(también producto de matrices elementales) tal que* $\mathbf{AB} = \mathbf{I} = \mathbf{BA}$ *(es decir, si* $\mathbf{A} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$*, entonces es invertible).*

Pero hay que tener en cuenta que la última transformación realizada ha de ser la primera en ser invertida (como cuando uno se pone los calcetines y luego los zapatos... para descalzarse se empieza por los zapatos y se finaliza por los calcetines); así, si se realizan una serie de transformaciones elementales en determinado orden, se deshacen en el orden inverso —véase la Proposición 5.1.3.

☞ Cualquier matriz $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ es invertible por ser producto de matrices elementales, en particular,

$$\left(\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}\right)\left(\mathbf{I}_{\boldsymbol{\tau}_k^{-1} \cdots \boldsymbol{\tau}_1^{-1}}\right) = \left(\mathbf{I}_{\boldsymbol{\tau}_1}\right) \cdots \left(\mathbf{I}_{\boldsymbol{\tau}_k}\right) \left(\mathbf{I}_{\boldsymbol{\tau}_k^{-1}}\right) \cdots \left(\mathbf{I}_{\boldsymbol{\tau}_1^{-1}}\right) = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k\, \boldsymbol{\tau}_k^{-1} \cdots \boldsymbol{\tau}_1^{-1}} = \mathbf{I};$$

por lo que podemos denotar $\boldsymbol{\tau}_k^{-1} \cdots \boldsymbol{\tau}_1^{-1}$ como $\left(\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k\right)^{-1}$. Así pues,

$$\mathbf{E}^{-1} = \mathbf{I}_{\boldsymbol{\tau}_k^{-1} \cdots \boldsymbol{\tau}_1^{-1}} = \mathbf{I}_{(\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k)^{-1}}.$$

Módulo NAcAL para Python

```
Tr = T([ (3,2), (5,2,3), (10,1,2) ], rpr='h'); Tr # Secuencia de transformaciones
                                                  # el parámetro rpr='h' es para mostrar
                                                  # la secuencia en rpr='h'orizontal
```

$$\underset{[(3)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(5)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}\underset{[(10)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}$$

```
Tr**(-1)                                          # Inversa de la secuencia de transf.
```

$$\underset{[(-10)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-5)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}\underset{\left[\left(\frac{1}{3}\right)\mathbf{2}\right]}{\boldsymbol{\tau}}$$

```
A    = I(3) & Tr                                 # A: producto de elementales
B    = I(3) & Tr**(-1)                           # B es la inversa de A
A*B                                               # Comprobación de que son inversas
```

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

### 5.1.2. Matrices preescalonadas con inversa.

☞ Como primer paso para encontrar un algoritmo que calcule la inversa, demostraremos que una matriz $\mathbf{K}$, de orden $n$, preescalonada y *sin columnas nulas*, se puede transformar en la matriz identidad $\mathbf{I}$ mediante una secuencia de transformaciones elementales

$$\mathbf{K}_{\tau_1\cdots\tau_k} \;=\; \mathbf{K}\big(\mathbf{I}_{\tau_1\cdots\tau_k}\big) \;=\; \mathbf{I} \qquad \text{(Teorema 5.1.10).}$$

En consecuencia, $\mathbf{K}^{-1}$ es el producto de las correspondientes matrices elementales; $\boxed{\mathbf{K}^{-1} = \mathbf{I}_{\tau_1\cdots\tau_k}}$.

**Teorema 5.1.10.** *Si una matriz $\mathbf{K}$, de orden $n$, es preescalonada y sin columnas nulas, existen $k$ transformaciones elementales tales que $\mathbf{K}_{\tau_1\cdots\tau_k} = \mathbf{I}$.*

*Demostración.* Toda matriz preescalonada sin columnas nulas tiene pivotes en todas sus columnas; si además es *cuadrada*, entonces tiene pivotes en todas sus *filas* (pues cada pivote ocupa una fila y columna distintas).

Por tanto, basta aplicar las $k$ transformaciones $\tau_1,\dots\tau_k$ de la eliminación Gauss-Jordan (Corolario 4.2.3 en la página 61) que transforman $\mathbf{K}$ en su forma escalonada reducida $\mathbf{K}_{\tau_1\cdots\tau_k}$; que en este caso resulta ser la matriz identidad de orden $n$ (pues $\mathbf{K}$, de orden $n$, tiene $n$ pivotes). □

Módulo NAcAL para Python

```
K = Matrix( [[0,5,0], [2,0,0], [6,9,3]] )
R = ElimGJ(K,1)                          # Forma Escalonada Reducida
```

$$\begin{bmatrix} 0 & 5 & 0 \\ 2 & 0 & 0 \\ 6 & 9 & 3 \end{bmatrix} \xrightarrow[[(-3)\mathbf{3}+\mathbf{2}]]{\overset{\tau}{[(-2)\mathbf{3}+\mathbf{1}]}} \begin{bmatrix} 0 & 5 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix} \xrightarrow[[(\frac{1}{3})\mathbf{3}]]{\overset{\tau}{[(\frac{1}{2})\mathbf{1}]}\\ [(\frac{1}{5})\mathbf{2}]} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{[\mathbf{1}\overset{\tau}{\rightleftharpoons}\mathbf{2}]} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

**Corolario 5.1.11.** *Si $\mathbf{K}$ es preescalonada y cuadrada, las siguientes propiedades son equivalentes:*

1. *$\mathbf{K}$ no tiene columnas nulas.*
2. *$\mathbf{K}$ es un producto de matrices elementales.*
3. *$\mathbf{K}$ tiene inversa.*

*Demostración.*

*1. ⇒ 2.* Si $\mathbf{K}$ no tiene columnas nulas, existen transformaciones elementales $\tau_1\cdots\tau_k$ tales que $\mathbf{K}_{\tau_1\cdots\tau_k} = \mathbf{I}$; así que

$$\begin{aligned} \mathbf{I} &= \mathbf{K}_{\tau_1\cdots\tau_k} = \mathbf{K}\big(\mathbf{I}_{\tau_1\cdots\tau_k}\big), \\ \mathbf{I}\big(\mathbf{I}_{\tau_1\cdots\tau_k}\big)^{-1} &= \mathbf{K}\big(\mathbf{I}_{\tau_1\cdots\tau_k}\big)\big(\mathbf{I}_{\tau_1\cdots\tau_k}\big)^{-1}; \quad \text{multiplicando ambos lados por } \big(\mathbf{I}_{\tau_1\cdots\tau_k}\big)^{-1}, \\ \big(\mathbf{I}_{\tau_1\cdots\tau_k}\big)^{-1} &= \mathbf{K}, \end{aligned}$$

donde $\big(\mathbf{I}_{\tau_1\cdots\tau_k}\big)^{-1} = \mathbf{I}_{(\tau_1\cdots\tau_k)^{-1}} = \mathbf{I}_{\tau_k^{-1}\cdots\tau_1^{-1}}$ es un producto de matrices elementales.

*2. ⇒ 3.* ¡Ya se sabe! Proposición 5.1.9 en la página anterior.

*3. ⇒ 1.* ¡Ya se sabe!... si tuviera columnas nulas sería singular por la Proposición 5.1.7 en la página 70. □

Módulo NAcAL para Python

```
(R.TrC)**-1 # inversa de la secuencia de transformaciones de las
            # columnas en el proceso de eliminación K --> R
```

$$\overset{\tau}{[1\rightleftharpoons 2]}\overset{\tau}{[(3)3]}\overset{\tau}{[(5)2]}\overset{\tau}{[(2)1]}\overset{\tau}{[(3)3+2]}\overset{\tau}{[(2)3+1]}$$

(compare con la salida del bloque de código anterior y observe cómo la primera transformación inversa deshace la última del bloque anterior, la segunda deshace la penúltima, etc.)

Módulo NAcAL para Python

```
I(K.n) & ( R.TrC )**-1   # al aplicar la inversa de las
                         # transformaciones sobre I(3) recuperamos K
```

$$\begin{bmatrix} 0 & 5 & 0 \\ 2 & 0 & 0 \\ 6 & 9 & 3 \end{bmatrix}$$

(`K.n` es el número de columnas de $\mathbf{K}$, así que `I(K.n)` es la matriz identidad de orden 3)

### 5.1.3. Matrices invertibles

☞ El siguiente corolario caracteriza las matrices con inversa. Además, nos dice cómo verificar si $\mathbf{A}$ es invertible:

- $\mathbf{A}$ es invertible si y solo si tiene alguna forma pre-escalonada sin columnas nulas.

**Corolario 5.1.12** (Caracterización de las matrices invertibles)**.** *Dada* $\mathbf{A}$ *de orden* $n$*, las siguientes propiedades son equivalentes:*

1. *Al (pre)escalonar* $\mathbf{A}$ *se obtiene una matriz sin columnas nulas.*
2. $\mathbf{A}$ *es producto de matrices elementales.*
3. $\mathbf{A}$ *tiene inversa.*

*Demostración.*

1. $\Rightarrow$ 2. Por el Teorema 4.2.1 sabemos que existen $\tau_1 \cdots \tau_k$ tales que $\mathbf{A}_{\tau_1 \cdots \tau_k} = \mathbf{K}$ es pre-escalonada; y si $\mathbf{K}$ no tiene columnas nulas, entonces es producto de $p$ matrices elementales, $\mathbf{K} = \mathbf{I}_{\tau_1' \cdots \tau_p'}$:

$$\begin{aligned} \mathbf{A}\left(\mathbf{I}_{\tau_1 \cdots \tau_k}\right) &= \left(\mathbf{I}_{\tau_1' \cdots \tau_p'}\right) \\ \mathbf{A}\left(\mathbf{I}_{\tau_1 \cdots \tau_k}\right)\left(\mathbf{I}_{\tau_1 \cdots \tau_k}\right)^{-1} &= \left(\mathbf{I}_{\tau_1' \cdots \tau_p'}\right)\left(\mathbf{I}_{\tau_1 \cdots \tau_k}\right)^{-1} \\ \mathbf{A} &= \left(\mathbf{I}_{\tau_1' \cdots \tau_p'}\right)\left(\mathbf{I}_{(\tau_1 \cdots \tau_k)^{-1}}\right) = \mathbf{I}_{\tau_1' \cdots \tau_p' \tau_k^{-1} \cdots \tau_1^{-1}}, \end{aligned}$$

2. $\Rightarrow$ 3. ¡Ya se sabe! Proposición 5.1.9 en la página 71

3. $\Rightarrow$ 1. Sean $\tau_1 \cdots \tau_k$ tales que $\mathbf{A}\left(\mathbf{I}_{\tau_1 \cdots \tau_k}\right) = \mathbf{K}$ es pre-escalonada. Entonces, puesto que $\mathbf{I}_{\tau_1 \cdots \tau_k}$ es invertible, $\mathbf{A}$ es invertible si y solo si $\mathbf{K}$ es invertible (Proposición 5.1.4 en la página 70).

□

Por tanto, para saber si una matriz cuadrada es singular o si es invertible, basta con aplicar la eliminación y verificar si la forma pre-escalonada encontrada tiene columnas nulas o no.

Módulo NAcAL para Python

```
A  = Matrix( [ [2,-1,0], [-1,2,-1], [0,-1,2] ] )
K  = Elim(A,1)              # una forma pre-escalonada de A
```

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \xrightarrow[[(1)\mathbf{1}+\mathbf{2}]]{\overset{\boldsymbol{\tau}}{[(2)\mathbf{2}]}} \begin{bmatrix} 2 & 0 & 0 \\ -1 & 3 & -1 \\ 0 & -2 & 2 \end{bmatrix} \xrightarrow[[(1)\mathbf{2}+\mathbf{3}]]{\overset{\boldsymbol{\tau}}{[(3)\mathbf{3}]}} \begin{bmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 0 & -2 & 4 \end{bmatrix}$$

(Como **A** es cuadrada y la forma pre-escalonada **K** no tiene columnas nulas, **A** es invertible.)

Podemos ver las transformaciones que hemos aplicado para llegar a la forma pre-escalonada **K** con:

Módulo NAcAL para Python

```
Tr = K.TrC                  # Transformaciones aplicadas a las columnas
T(Tr, rpr='h')              # Las representaremos en rpr='h'orizontal
```

$$\overset{\boldsymbol{\tau}}{[(2)\mathbf{2}]}\overset{\boldsymbol{\tau}}{[(1)\mathbf{1}+\mathbf{2}]}\overset{\boldsymbol{\tau}}{[(3)\mathbf{3}]}\overset{\boldsymbol{\tau}}{[(1)\mathbf{2}+\mathbf{3}]}$$

Aplicado la inversa de dichas transformaciones sobre **K** recuperamos la matriz original **A**:

Módulo NAcAL para Python

```
Matrix(K) & ( Tr**-1 )      # es A
```

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

Es decir, $\mathbf{A} = \left(\mathbf{I}_{\boldsymbol{\tau}_1'\cdots\boldsymbol{\tau}_p'}\right)_{\boldsymbol{\tau}_k^{-1}\cdots\boldsymbol{\tau}_1^{-1}}$ donde $\mathbf{I}_{\boldsymbol{\tau}_1'\cdots\boldsymbol{\tau}_p'} = \mathbf{K}$ es pre-escalonada, como en la demostración anterior:

$$\mathbf{K}_{\overset{\boldsymbol{\tau}}{[(-1)\mathbf{2}+\mathbf{3}]}\left[\left(\frac{1}{3}\right)\mathbf{3}\right]\overset{\boldsymbol{\tau}}{[(-1)\mathbf{1}+\mathbf{2}]}\left[\left(\frac{1}{2}\right)\mathbf{2}\right]} = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 0 & -2 & 4 \end{bmatrix}_{\overset{\boldsymbol{\tau}}{[(-1)\mathbf{2}+\mathbf{3}]}\left[\left(\frac{1}{3}\right)\mathbf{3}\right]\overset{\boldsymbol{\tau}}{[(-1)\mathbf{1}+\mathbf{2}]}\left[\left(\frac{1}{2}\right)\mathbf{2}\right]} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = \mathbf{A}.$$

☞ La siguiente proposición nos ofrece una nueva característica de las matrices NO invertibles:

- Si **A** tiene columnas que son combinación lineal del resto, entonces es singular.

Ejercicio 33. Demuestre la siguiente proposición.

**Proposición 5.1.13.** *Sea* **A** *de orden n; si alguna de sus columnas es combinación lineal del resto (o si alguna de sus filas es combinación lineal del resto), entonces* **A** *es* singular.

☞ Y ahora veamos que si **A** y **B** son de orden $n$ y $\mathbf{AB} = \mathbf{I}$, entonces **A** y **B** son una la inversa de la otra.

Este resultado es de gran importancia, pues nos dice que para comprobar que una matriz *cuadrada* es invertible, basta ver si existe inversa por uno cualquiera de sus lados.

Como nos apoyaremos en la eliminación para su demostración, no hemos podido ver antes este resultado.

**Proposición 5.1.14.** *Para* $\mathbf{A}$ *y* $\mathbf{B}$ *de orden* $n$ *se verifican las siguientes propiedades:*

- *Si* $\mathbf{B}$ *es singular, entonces* $\mathbf{AB}$ *es singular.*
- *Si* $\mathbf{A}$ *es singular, entonces* $\mathbf{AB}$ *es singular.*
- *Si* $\mathbf{AB} = \mathbf{I}$*, entonces tanto* $\mathbf{A}$ *como* $\mathbf{B}$ *son invertibles y* $\mathbf{B} = \mathbf{A}^{-1}$.

*Demostración.* Vayamos punto por punto:

- Si $\mathbf{B}$ es singular, la última columna de cualquiera de sus formas escalonadas es nula. Consecuentemente, si $\mathbf{E}$ es invertible y tal que $\mathbf{BE}$ es escalonada, entonces la última columna de $\mathbf{ABE}$ es nula. Y, por tanto, $\mathbf{ABE}$ es singular, pero como $\mathbf{E}$ es invertible, necesariamente $\mathbf{AB}$ es singular.
- Como $\mathbf{A}$ es singular, su transpuesta también. Por tanto, $\mathbf{B}^\intercal\mathbf{A}^\intercal$ es singular (punto anterior). Pero como $\mathbf{B}^\intercal\mathbf{A}^\intercal = (\mathbf{AB})^\intercal$ es singular, entonces su transpuesta $\mathbf{AB}$ también.
- Como $\mathbf{I}$ es invertible, por el punto anterior, $\mathbf{A}$ tiene que ser invertible, es decir, existe $\mathbf{A}^{-1}$. Multiplicando ambos lados por $\mathbf{A}^{-1}$ obtenemos el resultado deseado: $\mathbf{A}^{-1}(\mathbf{AB}) = \mathbf{A}^{-1}\mathbf{I} \quad \Rightarrow \quad \mathbf{B} = \mathbf{A}^{-1}$.

□

☞ ¡Ya tenemos las piezas para diseñar un algoritmo que encuentre la inversa de una matriz!
(lo hace buscando la inversa por el lado derecho mediante eliminación por columnas).

La eliminación permite encontrar una forma pre-escalonada de toda matriz (Teorema 4.2.1 en la página 59)

$$\mathbf{A}_{\tau_1\cdots\tau_p} = \mathbf{K},$$

y si $\mathbf{K}$ no tiene columnas nulas, aplicando Gauss-Jordan, se puede continuar con las transformaciones elementales hasta reducir $\mathbf{K}$ hasta $\mathbf{I}$ (Teorema 5.1.10 en la página 72)

$$\mathbf{K}_{\tau_{(p+1)}\cdots\tau_k} = \mathbf{I}.$$

Por tanto, la sucesión de $k$ transformaciones elementales $\tau_1\cdots\tau_p, \tau_{(p+1)}\cdots\tau_k$ transforma $\mathbf{A}$ en $\mathbf{I}$

$$\mathbf{A}_{\tau_1\cdots\tau_k} = \mathbf{A}(\mathbf{I}_{\tau_1\cdots\tau_k}) = \mathbf{I};$$

es decir, si $\mathbf{K}$ no tiene columnas nulas, entonces $\boxed{\mathbf{A}^{-1} = \mathbf{I}_{\tau_1\cdots\tau_k}}$; y si las tiene, entonces $\mathbf{A}^{-1}$ no existe.

Módulo NAcAL para Python

```
a, b, c, d = sympy.symbols('a b c d') # definimos las variables simbólicas
A = Matrix([[a,b], [c,d]])            # definimos la matriz
R = ElimGJ(A,1)                       # Eliminación hasta obtener la Identidad
Tr = R.TrC                            # Transformaciones de las Columnas
```

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \xrightarrow{[(-\frac{b}{a})\mathbf{1}+\mathbf{2}]^{\boldsymbol{\tau}}} \begin{bmatrix} a & 0 \\ c & \frac{ad-bc}{a} \end{bmatrix} \xrightarrow{[(-\frac{ac}{ad-bc})\mathbf{2}+\mathbf{1}]^{\boldsymbol{\tau}}} \begin{bmatrix} a & 0 \\ 0 & \frac{ad-bc}{a} \end{bmatrix} \xrightarrow{[(\frac{1}{a})\mathbf{1}]^{\boldsymbol{\tau}}\,[(\frac{a}{ad-bc})\mathbf{2}]} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(asumiendo que $a \neq 0$ y que $ad - bc \neq 0$)

Ahora ya sabemos que aplicando las transformaciones sobre las columnas de $\mathbf{I}$ obtenemos la inversa

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{[(-\frac{b}{a})\boldsymbol{\tau}_{1+2}][(-\frac{ac}{ad-bc})\boldsymbol{\tau}_{2+1}][(\frac{1}{a})\boldsymbol{\tau}_{1}][(\frac{a}{ad-bc})\boldsymbol{\tau}_{2}]} = \begin{bmatrix} \frac{d}{ad-bc} & -\frac{b}{ad-bc} \\ -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}$$

Módulo NAcAL para Python

```
A * (I(2) & Tr)        # comprobación de que I(2) & Tr es la inversa de A
```

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

☞ ...Pero aplicar $k$ transformaciones elementales $\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k$ a las columnas de $\mathbf{A}$ hasta tener $\mathbf{I}$, y después aplicar dichas transformaciones a las columnas de la matriz identidad, $\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$, puede ser bastante pesado.

¿Hay alguna forma de realizar ambas operaciones a la vez: la transformación $\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} = \mathbf{I}$ y la transformación $\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} = \mathbf{A}^{-1}$?

¡Sí! y es sencillísimo... basta "alargar" las columnas de $\mathbf{A}$ con $\mathbf{I}$ y entonces aplicar $\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k$ a $\begin{bmatrix} \mathbf{A} \\ \mathbf{I} \end{bmatrix}$.

### 5.1.4. Secuencias de transformaciones elementales. Parte II

Al apilar $\mathbf{A}$ y $\mathbf{B}$ (ambas con el mismo número de columnas), es decir, alargando las columnas de $\mathbf{A}$ al colocar las de $\mathbf{B}$ por debajo, obtenemos $\begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}$; podemos aplicar una transformación elemental $\boldsymbol{\tau}$ a las columnas de esta matriz "alargada", y el resultado es equivalente a transformar tanto $\mathbf{A}$ como $\mathbf{B}$ y apilar luego las matrices resultantes:

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix} \xrightarrow{\boldsymbol{\tau}} \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}_{\boldsymbol{\tau}} = \begin{bmatrix} \mathbf{A}_{\boldsymbol{\tau}} \\ \mathbf{B}_{\boldsymbol{\tau}} \end{bmatrix},$$

de manera que la columna $j$-ésima $\begin{bmatrix} \mathbf{A}_{\boldsymbol{\tau}} \\ \mathbf{B}_{\boldsymbol{\tau}} \end{bmatrix}_{|j}$ es la concatenación de las columnas $\left(\mathbf{A}_{\boldsymbol{\tau}}\right)_{|j}$ y $\left(\mathbf{B}_{\boldsymbol{\tau}}\right)_{|j}$.

Lo mismo sucede si aplicamos una secuencia de $k$ transformaciones elementales:

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} = \begin{bmatrix} \mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} \\ \mathbf{B}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} \end{bmatrix}.$$

Por tanto, si antes de realizar las transformaciones "pegamos" por debajo de $\mathbf{A} = \begin{bmatrix} 0 & 1 & 5 \\ 7 & 0 & 4 \end{bmatrix}$ la matriz $\mathbf{I}$ de orden 3 y aplicamos algunas transformaciones:

$$\begin{bmatrix} \mathbf{A} \\ \mathbf{I} \end{bmatrix} = \left[\begin{array}{ccc} 0 & 1 & 5 \\ 7 & 0 & 4 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow{[(7)\boldsymbol{\tau}_{3}][(-4)\boldsymbol{\tau}_{1+3}]} \left[\begin{array}{ccc} 0 & 1 & 35 \\ 7 & 0 & 0 \\ \hline 1 & 0 & -4 \\ 0 & 1 & 0 \\ 0 & 0 & 7 \end{array}\right] \xrightarrow{[(-35)\boldsymbol{\tau}_{2+3}]} \left[\begin{array}{ccc} 0 & 1 & 0 \\ 7 & 0 & 0 \\ \hline 1 & 0 & -4 \\ 0 & 1 & -35 \\ 0 & 0 & 7 \end{array}\right] = \begin{bmatrix} \mathbf{A}_{[(7)\boldsymbol{\tau}_{3}][(-4)\boldsymbol{\tau}_{1+3}][(-35)\boldsymbol{\tau}_{2+3}]} \\ \mathbf{I}_{[(7)\boldsymbol{\tau}_{3}][(-4)\boldsymbol{\tau}_{1+3}][(-35)\boldsymbol{\tau}_{2+3}]} \end{bmatrix},$$

entonces, sin necesidad de multiplicar, sabemos por una parte que

$$\mathbf{I}_{\overset{\tau}{[(7)\mathbf{3}]}[(-4)\overset{\tau}{\mathbf{1}}+\mathbf{3}]}\overset{\tau}{[(-35)\mathbf{2}+\mathbf{3}]} = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & -35 \\ 0 & 0 & 7 \end{bmatrix}: \qquad \text{y, por la otra, que} \qquad \mathbf{A}_{\overset{\tau}{[(7)\mathbf{3}]}[(-4)\overset{\tau}{\mathbf{1}}+\mathbf{3}]}\overset{\tau}{[(-35)\mathbf{2}+\mathbf{3}]} = \begin{bmatrix} 0 & 1 & 0 \\ 7 & 0 & 0 \end{bmatrix}.$$

Así pues, es fácil encontrar la inversa de $\mathbf{A}$: basta con calcular la forma escalonada reducida de la matriz que resulta de apilar $\mathbf{A}$ con la matriz identidad que tiene el mismo número de columnas:

*Ejemplo* 12.

$$\left[\begin{array}{c}\mathbf{A}\\\hline \mathbf{I}\end{array}\right] = \left[\begin{array}{cc} 5 & 0 \\ 2 & -2 \\ \hline 1 & 0 \\ 0 & 1 \end{array}\right] \xrightarrow{\overset{\tau}{[(1)\mathbf{2}+\mathbf{1}]}} \left[\begin{array}{cc} 5 & 0 \\ 0 & -2 \\ \hline 1 & 0 \\ 1 & 1 \end{array}\right] \xrightarrow[\overset{\tau}{[(-\frac{1}{2})\mathbf{2}]}]{\overset{\tau}{[(\frac{1}{5})\mathbf{1}]}} \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 1/5 & 0 \\ 1/5 & -1/2 \end{array}\right] = \left[\begin{array}{c}\mathbf{A}_{\tau_1\tau_2\tau_3}\\\hline \mathbf{I}_{\tau_1\tau_2\tau_3}\end{array}\right].$$

Puesto que $\mathbf{A}_{\tau_1\tau_2\tau_3} = \mathbf{I}$, entonces $\mathbf{I}_{\tau_1\tau_2\tau_3} = \begin{bmatrix} 1/5 & 0 \\ 1/5 & -1/2 \end{bmatrix} = \mathbf{A}^{-1}$.

☞

**Método para encontrar la inversa de una matriz (cuadrada)**

1. Con transformaciones elementales de las columnas, encuentre una forma escalonada reducida de $\left[\begin{array}{c}\mathbf{A}\\\hline \mathbf{I}\end{array}\right]$.
2. Si en la matriz obtenida $\left[\begin{array}{c}\mathbf{R}\\\hline \mathbf{I}_{\tau_1\cdots\tau_k}\end{array}\right]$ la submatriz $\mathbf{R}$ tiene columnas nulas, entonces $\mathbf{A}$ no tiene inversa.
3. En caso contrario, $\mathbf{I}_{\tau_1\cdots\tau_k}$ es la inversa de $\mathbf{A}$.

Módulo NAcAL para Python

```
A = Matrix([ [1, -1, -3], [0, 1, 0], [0, -1, 1] ])
InvMat(A,1)          # Calcula la inversa por eliminación Gauss-Jordan
```

$$\left[\begin{array}{ccc} 1 & -1 & -3 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow[\overset{\tau}{[(3)\mathbf{1}+\mathbf{3}]}]{\overset{\tau}{[(1)\mathbf{1}+\mathbf{2}]}} \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \\ \hline 1 & 1 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow{\overset{\tau}{[(1)\mathbf{3}+\mathbf{2}]}} \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 1 & 4 & 3 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right]$$

### 5.1.5. Inversa de una matriz triangular

☞ Veamos primero que una matriz triangular no puede tener ceros en la diagonal principal y ser invertible.

Por ejemplo, la siguiente matriz no es invertible, pues es imposible encontrar un vector $(x, y, z)$ tal que

$$\begin{bmatrix} 1 & a & b \\ & 0 & c \\ & & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \mathbf{I}_{|2}.$$

Por una parte, $z$ es necesariamente cero, y por otra, con la primera y segunda columnas no podemos obtener la segunda componente del vector del lado derecho... Ahora vayamos con la demostración general.

**Proposición 5.1.15.** *Una matriz triangular invertible no tiene ceros en la diagonal principal.*

*Demostración.* Si $\mathbf{A}$ es triangular superior (si *las componentes por debajo de la diagonal son nulas*). ¿Puede haber un cero en $a_{11}$? (en la primera posición de la diagonal). Puesto que las componentes por debajo de la diagonal son cero, si además $a_{11}$ también fuera cero, entonces la primera columna sería nula, lo cual es imposible dado que $\mathbf{A}$ es invertible. **Por tanto,** $a_{11}$ **no puede ser cero.**

¿Puede haber un *primer* cero en la posición $j$-ésima de la diagonal? Por ser el *primer* cero de la diagonal, las anteriores componentes de la diagonal serían distintas de cero. Así que, aplicando eliminación, podemos anular todas las componentes por encima de $a_{jj}$; pero como todas las componentes por debajo también son cero (por ser la matriz triangular superior), anularíamos la columna; algo imposible, pues $\mathbf{A}$ es invertible. **Así que** $a_{jj}$ **tampoco puede ser cero.** (La demostración es semejante para una triangular inferior) □

Si $\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k$ es una sucesión de transformaciones elementales de *Tipo I* de un proceso de eliminación de "izquierda a derecha", la matriz $\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ es triangular superior unitaria, ya que solo cambian las componentes a la derecha de la diagonal, puesto que ninguna columna es multiplicada por un escalar. Así tenemos el siguiente

**Teorema 5.1.16.** *La aplicación del proceso de eliminación de "izquierda a derecha" sobre una matriz identidad da lugar a una matriz triangular superior unitaria e invertible.*

Ahora considere una matriz $\mathbf{U}$ triangular superior unitaria. Para calcular su inversa, basta aplicar la eliminación de "izquierda a derecha" mediante una sucesión de transformaciones $\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k$ que eliminen todas las componentes a la derecha de la diagonal:

$$\left[\begin{array}{c}\mathbf{U}\\\hline \mathbf{I}\end{array}\right] \xrightarrow{\overbrace{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}^{\text{(de "izquierda a derecha")}}} \left[\begin{array}{c}\mathbf{I}\\\hline \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}\end{array}\right] \quad \text{por tanto} \quad \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} = \mathbf{U}^{-1} \text{ es triangular superior unitaria.}$$

Así pues, tenemos el siguiente

**Corolario 5.1.17.** *Si* $\mathbf{U}$ *es cuadrada y triangular superior unitaria,* $\mathbf{U}^{-1}$ *también es triangular superior unitaria.*

Por ejemplo:
$$\left[\begin{array}{ccc}1&b&c\\0&1&e\\0&0&1\\\hline 1&0&0\\0&1&0\\0&0&1\end{array}\right] \xrightarrow[\boldsymbol{[(-c)1+3]}]{\boldsymbol{\tau}\ \boldsymbol{[(-b)1+2]}} \left[\begin{array}{ccc}1&0&0\\0&1&e\\0&0&1\\\hline 1&-b&-c\\0&1&0\\0&0&1\end{array}\right] \xrightarrow{\boldsymbol{\tau}\ \boldsymbol{[(-e)2+3]}} \left[\begin{array}{ccc}1&0&0\\0&1&0\\0&0&1\\\hline 1&-b&be-c\\0&1&-e\\0&0&1\end{array}\right].$$

Si $\mathbf{U}$ no es unitaria, aunque sí invertible, tras escalonar $\mathbf{U}$ por eliminación, será necesario dividir cada columna por el valor de su correspondiente pivote (última operación en la eliminación Gauss-Jordan). Consecuentemente, la inversa de $\mathbf{U}$ es triangular superior (por el proceso de eliminación), pero no es unitaria (pues el último paso de la eliminación Gauss-Jordan modifica la diagonal). Así hemos demostrado la siguiente

**Proposición 5.1.18.** *Si* $\mathbf{U}$ *es triangular superior e invertible,* $\mathbf{U}^{-1}$ *también es triangular superior.*

Por ejemplo:

$$\left[\begin{array}{ccc}a&b&c\\0&d&e\\0&0&f\\\hline 1&0&0\\0&1&0\\0&0&1\end{array}\right] \xrightarrow[\boldsymbol{[(-\frac{c}{a})1+3]}]{\boldsymbol{\tau}\ \boldsymbol{[(-\frac{b}{a})1+2]}} \left[\begin{array}{ccc}a&0&0\\0&d&e\\0&0&f\\\hline 1&-\frac{b}{a}&-\frac{c}{a}\\0&1&0\\0&0&1\end{array}\right] \xrightarrow{\boldsymbol{\tau}\ \boldsymbol{[(-\frac{e}{d})2+3]}} \left[\begin{array}{ccc}a&0&0\\0&d&0\\0&0&f\\\hline 1&-\frac{b}{a}&\frac{be-cd}{ad}\\0&1&-\frac{e}{d}\\0&0&1\end{array}\right] \xrightarrow[\boldsymbol{[(\frac{1}{f})3]}]{\boldsymbol{\tau}\ \boldsymbol{[(\frac{1}{d})2]}\ \boldsymbol{[(\frac{1}{a})1]}} \left[\begin{array}{ccc}1&0&0\\0&1&0\\0&0&1\\\hline \frac{1}{a}&-\frac{b}{ad}&\frac{be-cd}{adf}\\0&\frac{1}{d}&-\frac{e}{df}\\0&0&\frac{1}{f}\end{array}\right]$$

Para terminar esta sección, como la inversa de la transpuesta es la transpuesta de la inversa (Proposición 5.1.6 en la página 70), llegamos a los siguientes corolarios:

**Corolario 5.1.19.** *Si* $\mathbf{L}$ *es cuadrada y triangular inferior unitaria,* $\mathbf{L}^{-1}$ *también es triangular inferior unitaria.*

**Corolario 5.1.20.** *Si* $\mathbf{L}$ *es triangular inferior e invertible,* $\mathbf{L}^{-1}$ *también es triangular inferior.*

## 5.2. Rango de una matriz

☞ Vimos (Corolario 5.1.12) que si una matriz es invertible, ninguna de sus formas preescalonadas tiene columnas nulas. Una generalización de este resultado da lugar a la definición de rango que utilizaremos aquí.

Considere una matriz $\mathbf{A}$ tal que $\mathbf{A} = \mathbf{KB}$, donde $\mathbf{K}$ es una matriz preescalonada. Puesto que las columnas de $\mathbf{A}$ son combinaciones lineales de las de $\mathbf{K}$, resulta que la posición de pivote de cada una de las columnas no nulas de $\mathbf{A}$ coincide con la posición de pivote de *alguna* (solo una) de las columnas de $\mathbf{K}$.

El motivo es que la posición de pivote de la columna no nula $\mathbf{A}_{|j} = \mathbf{KB}_{|j}$ es, de entre aquellas columnas de $\mathbf{K}$ multiplicadas por coeficientes no nulos de $\mathbf{B}_{|j}$, la posición del pivote que se encuentre más "arriba", es decir, la de menor posición (recuerde que, como $\mathbf{K}$ está preescalonada, *cada pivote en* $\mathbf{K}$ *ocupa una fila distinta*).

Para verlo, nótese que, si $\mathbf{K}$ tiene $n$ columnas, entonces, para toda lista $\boldsymbol{\gamma}$ de índices entre 1 y $n$, no repetidos y ordenados de menor a mayor, la submatriz $\mathbf{K}_{|\boldsymbol{\gamma}}$ es preescalonada (pues como mucho habrán desaparecido algunas columnas, por lo que sigue sin haber elementos distintos de cero a la derecha de cada pivote). En particular, sea $\boldsymbol{\gamma}$ la lista formada por los índices $h$ tales que $\mathbf{K}_{|h}$ y ${}_{h|}\mathbf{B}_{|j}$ son distintos de cero, es decir $\boldsymbol{\gamma} = \big(h \mid \mathbf{K}_{|h} \neq \mathbf{0} \text{ y } {}_{h|}\mathbf{B}_{|j} \neq 0\big)$; y sea $\boldsymbol{\nu}$ la lista de índices de 1 a $n$ que no están en $\boldsymbol{\gamma}$. Entonces, por la Ecuación 3.8 sabemos que

$$\mathbf{A}_{|j} = \mathbf{KB}_{|j} = \big(\mathbf{K}_{|\boldsymbol{\gamma}}\big)\big({}_{\boldsymbol{\gamma}|}\mathbf{B}_{|j}\big) + \underbrace{\big(\mathbf{K}_{|\boldsymbol{\nu}}\big)\big({}_{\boldsymbol{\nu}|}\mathbf{B}_{|j}\big)}_{\mathbf{0}} = \big(\mathbf{K}_{|\boldsymbol{\gamma}}\big)\big({}_{\boldsymbol{\gamma}|}\mathbf{B}\big)_{|j}.$$

Sea $i$ la *menor posición* de los pivotes en $\mathbf{K}_{|\boldsymbol{\gamma}}$, entonces (por ser preescalonada) el pivote con dicha posición es el único elemento distinto de cero en la fila ${}_{i|}\big(\mathbf{K}_{|\boldsymbol{\gamma}}\big)$. Así, como el subvector $\big({}_{\boldsymbol{\gamma}|}\mathbf{B}\big)_{|j}$ no tiene elementos nulos, necesariamente ${}_{i|}\mathbf{A}_{|j} = {}_{i|}\big(\mathbf{K}_{|\boldsymbol{\gamma}}\big)\big({}_{\boldsymbol{\gamma}|}\mathbf{B}\big)_{|j} \neq 0$ es el pivote de la columna $\mathbf{A}_{|j}$.

*Resumiendo*: sabemos que si $\mathbf{A} = \mathbf{KB}$ (con $\mathbf{K}$ preescalonada), la posición de pivote de cada una de las columnas no nulas de $\mathbf{A}$ coincide con la posición de pivote de *alguna* (solo una) de las columnas de $\mathbf{K}$. Por tanto, *el conjunto de posiciones de pivote en* $\mathbf{K}$ *contiene al conjunto de posiciones de pivote en* $\mathbf{A}$ (pues todas las posiciones de pivote en $\mathbf{A}$ están en $\mathbf{K}$).

*Demostremos ahora que las formas preescalonadas de una matriz* $\mathbf{A}$ *tienen el mismo número de pivotes:*

Supongamos que $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ y $\mathbf{E}' = \mathbf{I}_{\boldsymbol{\tau}'_1 \cdots \boldsymbol{\tau}'_p}$ son matrices tales que $\mathbf{AE} = \mathbf{K}$ y $\mathbf{AE}' = \mathbf{K}'$ son preescalonadas. Entonces las posiciones de los pivotes de $\mathbf{K}$ y $\mathbf{K}'$ son coincidentes. Veámoslo.

Como $\mathbf{A} = \mathbf{KE}^{-1}$, tenemos que

$$\mathbf{K}' = \mathbf{AE}' = \mathbf{KE}^{-1}\mathbf{E}';$$

por tanto, las columnas de $\mathbf{K}' = \mathbf{K}\big(\mathbf{E}^{-1}\mathbf{E}'\big)$ son combinación lineal de las de la matriz preescalonada $\mathbf{K}$; así que *el conjunto de posiciones de pivote en* $\mathbf{K}$ *contiene al de* $\mathbf{K}'$. Pero también $\mathbf{K} = \mathbf{K}'\big(\mathbf{E}'^{-1}\mathbf{E}\big)$, por tanto, las columnas de $\mathbf{K}$ son combinación lineal de las de la matriz preescalonada $\mathbf{K}'$; así que *el conjunto de posiciones de pivote en* $\mathbf{K}'$ *contiene al de* $\mathbf{K}$. Consecuentemente, ambos conjuntos son iguales. Y como ambas matrices

están preescalonadas, es decir, como en cada matriz cada pivote ocupa una posición distinta, necesariamente ambas matrices tienen el mismo número de pivotes.

Este resultado nos permite dar la siguiente definición:

**Definición 5.4** (Rango de una matriz)**.** *El rango de* $\mathbf{A}$ *es el número de columnas no nulas de cualquiera de sus formas preescalonadas. Denotamos el rango de* $\mathbf{A}$ *con* $\operatorname{rg}(\mathbf{A})$.

Es decir, el rango de $\mathbf{A}$ es el número de pivotes de cualquiera de sus formas preescalonadas.

### 5.2.1. Algunas propiedades del rango de una matriz

**Proposición 5.2.1.** *Si* $\mathbf{E}$ *es invertible, entonces* $\operatorname{rg}(\mathbf{A}) = \operatorname{rg}(\mathbf{AE})$.

*Demostración.* Con ambas matrices se pueden obtener idéntica forma preescalonada $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$. Veámoslo: sea $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}'_1\cdots\boldsymbol{\tau}'_p}$, entonces $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \left(\mathbf{AEE}^{-1}\right)_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \left(\mathbf{AE}\right)_{(\boldsymbol{\tau}'_1\cdots\boldsymbol{\tau}'_p)^{-1}\,\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$. □

**Proposición 5.2.2.** *El rango de* $\mathbf{AB}$ *es menor o igual que el de* $\mathbf{B}$.

*Demostración.* Sea $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ tal que $\mathbf{BE} = \mathbf{L}$ es escalonada. Entonces

- Aplicando la proposición anterior tenemos que en particular: $\operatorname{rg}(\mathbf{AB}) = \operatorname{rg}((\mathbf{AB})\mathbf{E})$.
- Por otra parte tenemos que $(\mathbf{AB})\mathbf{E} = \mathbf{A}(\mathbf{BE}) = \mathbf{AL}$, y si $k$ es el número de columnas nulas de la matriz escalonada $\mathbf{L}$, las últimas $k$ columnas de $\mathbf{AL}$ son necesariamente nulas.
- Si realizamos transformaciones elementales sobre las columnas de $\mathbf{AL}$ de manera que no modificamos las $k$ últimas, obtendremos una forma escalonada con al menos $k$ columnas nulas. Por tanto, el rango de $\mathbf{AL}$ es menor o igual que el rango de $\mathbf{L}$, por lo que finalmente concluimos que

$$\operatorname{rg}(\mathbf{AB}) = \operatorname{rg}(\mathbf{ABE}) = \operatorname{rg}(\mathbf{AL}) \leq \operatorname{rg}(\mathbf{L}) = \operatorname{rg}(\mathbf{B}).$$ □

**Proposición 5.2.3.** *Si* $\mathbf{E}$ *es invertible, entonces* $\operatorname{rg}(\mathbf{EA}) = \operatorname{rg}(\mathbf{A})$.

*Demostración.* La proposición anterior nos indica que $\operatorname{rg}(\mathbf{EA}) \leq \operatorname{rg}(\mathbf{A})$, pero también nos indica que $\operatorname{rg}(\mathbf{E}^{-1}(\mathbf{EA})) \leq \operatorname{rg}(\mathbf{EA})$; por tanto $\operatorname{rg}(\mathbf{A}) \leq \operatorname{rg}(\mathbf{EA})$. Por tanto, $\operatorname{rg}(\mathbf{EA}) = \operatorname{rg}(\mathbf{A})$. □

Aplicando eliminación *gaussiana* por filas[1] a una matriz escalonada reducida $\mathbf{R}$ (anulando las componentes por debajo de los pivotes y dejando las filas nulas por debajo de las no nulas), obtenemos una matriz $\mathbf{S}$ escalonada reducida cuyos pivotes están en la diagonal principal. Así, tanto $\mathbf{S}$ como su transpuesta son diagonales (con idéntica diagonal); consecuentemente $\operatorname{rg}(\mathbf{S}) = \operatorname{rg}(\mathbf{S}^\intercal)$:

Módulo NAcAL para Python

```
R = Matrix( [ [1,0,0,0], [-1,0,0,0], [0,1,0,0], [0,0,1,0], [0,0,2,0] ] )
S = ElimGF(R,1)              # Eliminación Gaussiana por filas
~S                           # transpuesta de S
```

[1] Véase el Apéndice 4.A en la página 62

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix} \xrightarrow[[(1)\mathbf{1}+\mathbf{2}]]{\boldsymbol{\tau}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix} \xrightarrow[[(-2)\mathbf{4}+\mathbf{5}]]{\boldsymbol{\tau}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[\substack{[\mathbf{3}\rightleftharpoons\mathbf{4}] \\ [\mathbf{2}\rightleftharpoons\mathbf{3}]}]{\boldsymbol{\tau}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathbf{S}$$

$$\mathbf{S}^\intercal = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Aprovechemos este hecho para demostrar las siguientes dos proposiciones.

**Proposición 5.2.4.** *Si* $\mathbf{R}$ *es escalonada reducida, entonces* $\operatorname{rg}(\mathbf{R}) = \operatorname{rg}(\mathbf{R}^\intercal)$.

*Demostración.* Puesto que podemos encontrar una matriz invertible $\mathbf{E} = {}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}\mathbf{I}$ tal que $\mathbf{ER} = \mathbf{S}$ es diagonal. Así, por las proposiciones 5.2.3 y 5.2.1, y puesto que $\mathbf{E}^\intercal$ también es invertible:

$$\operatorname{rg}(\mathbf{R}) = \operatorname{rg}(\mathbf{ER}) = \operatorname{rg}(\mathbf{S}) = \operatorname{rg}(\mathbf{S}^\intercal) = \operatorname{rg}(\mathbf{R}^\intercal\mathbf{E}^\intercal) = \operatorname{rg}(\mathbf{R}^\intercal). \qquad \square$$

**Proposición 5.2.5.** $\operatorname{rg}(\mathbf{A}) = \operatorname{rg}(\mathbf{A}^\intercal)$.

*Demostración.* Sea $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ tal que $\mathbf{AE} = \mathbf{R}$ es escalonada reducida, entonces (aplicando la Proposición 5.2.1 en la última igualdad) tenemos: $\operatorname{rg}(\mathbf{A}) = \operatorname{rg}(\mathbf{AE}) = \operatorname{rg}(\mathbf{R}) = \operatorname{rg}(\mathbf{R}^\intercal) = \operatorname{rg}(\mathbf{E}^\intercal\mathbf{A}^\intercal) = \operatorname{rg}(\mathbf{A}^\intercal)$. $\square$

**Proposición 5.2.6.** *El rango de* $\mathbf{AB}$ *es menor o igual que el de* $\mathbf{A}$.

*Demostración.* $\operatorname{rg}(\mathbf{AB}) = \operatorname{rg}(\mathbf{B}^\intercal\mathbf{A}^\intercal) \leq \operatorname{rg}(\mathbf{A}^\intercal) = \operatorname{rg}(\mathbf{A})$. $\square$

Módulo NAcAL para Python

```
Elim(A).rango    # rango es un atributo de la clase Elim
A.rango()        # rango() es un método de la clase Matrix
```

**Rango completo por columnas (por filas).** Una matriz cuyas formas preescalonadas tienen pivote en todas las columnas se dice que es *de rango completo por columnas*; y si tienen pivote en todas las filas se dice que es *de rango completo por filas*.

**Rango completo.** Una matriz que es de rango completo tanto por filas como por columnas (y que por tanto es cuadrada) se dice que es *de rango completo*.

Así pues, por el Corolario 5.1.11, *una matriz de rango completo es invertible*, pues es cuadrada y sin columnas nulas (por tener todas pivote).

## *Apéndices a la lección*

☞ Vamos a aclarar algo sobre la notación empleada en la Sección 4.B en la página 65. Recuerde que puede posponer el estudio de las transformaciones elementales *"espejo"* hasta que llegue a la Lección 17.

## 5.A. *"Espejo"* de la inversa de una secuencia de transformaciones elementales

Fíjese que ${}_{\mathrm{esp}(\boldsymbol{\tau}^{-1})}\mathbf{I} = \mathbf{I}_{\boldsymbol{\tau}^{-1}} = \big(\mathbf{I}_{\boldsymbol{\tau}}\big)^{-1} = \big({}_{\mathrm{esp}(\boldsymbol{\tau})}\mathbf{I}\big)^{-1} = {}_{(\mathrm{esp}(\boldsymbol{\tau}))^{-1}}\mathbf{I}$; es decir, que $\mathrm{esp}(\boldsymbol{\tau}^{-1}) = (\mathrm{esp}(\boldsymbol{\tau}))^{-1}$.

Como $\mathrm{esp}\big(\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k\big) = \mathrm{esp}(\boldsymbol{\tau}_1) \ldots \mathrm{esp}(\boldsymbol{\tau}_k)$, y como $\big(\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k\big)^{-1} = \boldsymbol{\tau}_k^{-1} \ldots \boldsymbol{\tau}_1^{-1}$; tenemos que

$\mathrm{esp}\big((\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k)^{-1}\big) = \mathrm{esp}\big(\boldsymbol{\tau}_k^{-1} \ldots \boldsymbol{\tau}_1^{-1}\big) = \mathrm{esp}(\boldsymbol{\tau}_k^{-1}) \ldots \mathrm{esp}(\boldsymbol{\tau}_1^{-1}) = \big(\mathrm{esp}(\boldsymbol{\tau}_k)\big)^{-1} \ldots \big(\mathrm{esp}(\boldsymbol{\tau}_1)\big)^{-1} = \big(\mathrm{esp}(\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k)\big)^{-1}$;

así que podemos escribir sencillamente: $\boxed{\mathrm{esp}\,(\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k)^{-1}}$.

Módulo NAcAL para Python

```
Tr = T((3,2,1)) & T((-2,3,2)) & T((7,1,3)) & T({2,3}) & T((20,1)) # Suc. transf.
(Tr**-1).espejo() == (Tr.espejo()**-1)                              # ¿son iguales?
```

```
True
```

# Parte III

# Subespacios y resolución de sistemas de ecuaciones lineales

# Espacios vectoriales y funciones lineales

☞ El propósito de esta lección es introducir terminología abstracta que utilizaremos a partir de ahora. De paso, recapitularemos lo visto hasta aquí empleando la nueva terminología.

Comenzaremos con los espacios vectoriales y las funciones lineales. Luego veremos formas de definir espacios más pequeños que uno dado (los subespacios) y espacios más grandes a partir de otros subespacios: mediante productos cartesianos y con funciones cuya imagen está contenida en un espacio vectorial.

Casi cualquier espacio vectorial con el que se encuentre en otras disciplinas será fruto de un encadenamiento de espacios vectoriales de funciones y/o productos cartesianos, o subespacios de estos.

## 6.1. Espacios vectoriales

Desde la escuela usted sabe que, con la suma y el producto habituales entre *números reales*, se verifican las siguientes propiedades:

1. $a + b = b + a$.
2. $a + (b + c) = (a + b) + c$.
3. $a + 0 = a$.
4. $a + (-a) = 0$.
5. $a(b + c) = ab + ac$.
6. $(a + b)c = ac + bc$.
7. $a(bc) = (ab)c$.
8. $1a = a$.

Además, en la primera lección vimos que, al definir la suma y el producto por escalares para $\mathbb{R}^n$ como

$$(\boldsymbol{a} + \boldsymbol{b})_{|i} = \boldsymbol{a}_{|i} + \boldsymbol{b}_{|i} \qquad \text{y} \qquad (\lambda \boldsymbol{b})_{|i} = \lambda(\boldsymbol{b}_{|i}),$$

se verifica que (Proposición 1.2.1 en la página 9):

1. $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{b} + \boldsymbol{a}$
2. $\boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}) = (\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c}$
3. $\boldsymbol{a} + \boldsymbol{0} = \boldsymbol{a}$
4. $\boldsymbol{a} + (-\boldsymbol{a}) = \boldsymbol{0}$
5. $\lambda(\boldsymbol{a} + \boldsymbol{b}) = \lambda\boldsymbol{a} + \lambda\boldsymbol{b}$
6. $(\lambda + \varphi)\boldsymbol{a} = \lambda\boldsymbol{a} + \varphi\boldsymbol{a}$
7. $\lambda(\varphi\boldsymbol{a}) = (\lambda\varphi)\boldsymbol{a}$
8. $1\boldsymbol{a} = \boldsymbol{a}$

Asimismo, si definimos la suma y el producto por escalares para $\mathbb{R}^{m\times n}$ (*matrices de orden $m$ por $n$*) como

$$(\mathbf{A} + \mathbf{B})_{|j} = \mathbf{A}_{|j} + \mathbf{B}_{|j} \qquad \text{y} \qquad (\lambda\mathbf{A})_{|j} = \lambda(\mathbf{A}_{|j}),$$

también se verifica que (Proposición 1.5.1 en la página 22):

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
2. $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$
3. $\mathbf{A} + \mathbf{0} = \mathbf{A}$
4. $\mathbf{A} + (-\mathbf{A}) = \mathbf{0}$
5. $\lambda(\mathbf{A} + \mathbf{B}) = \lambda\mathbf{A} + \lambda\mathbf{B}$
6. $(\lambda + \varphi)\mathbf{A} = \lambda\mathbf{A} + \varphi\mathbf{A}$
7. $\lambda(\varphi\mathbf{A}) = (\lambda\varphi)\mathbf{A}$
8. $1\mathbf{A} = \mathbf{A}$

A primera vista, es evidente que subyace una estructura común a $\mathbb{R}$, a $\mathbb{R}^n$ y al conjunto de matrices $\mathbb{R}^{m\times n}$ (con sus respectivas operaciones de suma y producto por escalares). Dicha estructura se denomina *espacio vectorial*.

**Definición 6.1.** *Un* espacio vectorial *es un conjunto $\mathcal{V}$ de objetos*[1] *junto con dos operaciones que denominamos:* suma *y* producto por escalares. *En cuanto a los elementos de $\mathcal{V}$:*

- *Se denominan* vectores *y los denotaremos genéricamente con letras minúsculas en cursiva y con una flecha por encima: $\vec{x}$, $\vec{y}$, $\vec{z}$,...*[2]

*En cuanto a las operaciones:*

- *La* suma *asocia cualquier par de elementos de $\mathcal{V}$ con otro elemento de $\mathcal{V}$.*

  *Dicho de otro modo, la suma de dos vectores, $\vec{x}$ e $\vec{y}$, de $\mathcal{V}$ es otro vector de $\mathcal{V}$ que denotamos como $\vec{x}+\vec{y}$. Esto se representa esquemáticamente con*[3]

$$\begin{aligned} _ + _ : \mathcal{V}\times\mathcal{V} &\longrightarrow \mathcal{V} \\ (\vec{x},\vec{y}) &\longmapsto \vec{x}+\vec{y} \end{aligned}$$

  (Puesto que al sumar vectores de $\mathcal{V}$ se obtienen elementos de $\mathcal{V}$, se dice que $\mathcal{V}$ *es cerrado para la suma*).

  *La operación* suma *debe verificar las siguientes cuatro propiedades:*

  1. $\vec{x}+\vec{y}=\vec{y}+\vec{x}$
  2. $\vec{x}+(\vec{y}+\vec{z})=(\vec{x}+\vec{y})+\vec{z}$
  3. *Existe un vector,*[4]*, que denotamos con* $\vec{0}$ (vector nulo) *tal que* $\vec{x}+\vec{0}=\vec{x}$, *para todo* $\vec{x}\in\mathcal{V}$.
  4. *Para cada* $\vec{x}\in\mathcal{V}$ *existe un vector, que denotamos con* $-\vec{x}$ (vector opuesto) *tal que* $\vec{x}+(-\vec{x})=\vec{0}$.

- *El* producto por escalares *asocia cualquier par, formado por un escalar y un vector de $\mathcal{V}$, con un vector de $\mathcal{V}$.*

  *Dicho de otro modo, el producto de $a\in\mathbb{R}$ por $\vec{x}\in\mathcal{V}$ es otro vector de $\mathcal{V}$ que denotamos como $a\cdot\vec{x}$*

$$\begin{aligned} _ \cdot _ : \mathbb{R}\times\mathcal{V} &\longrightarrow \mathcal{V} \\ (a,\vec{x}) &\longmapsto a\cdot\vec{x} \end{aligned}$$

  (normalmente omitimos el "punto" y escribimos sencillamente $a\vec{x}$, que llamaremos *múltiplo* de $\vec{x}$) .

  *La operación* producto por un escalar *debe verificar las siguientes cuatro propiedades:*

  5. $a(\vec{x}+\vec{y})=a\vec{x}+a\vec{y}$
  6. $(a+b)\vec{x}=a\vec{x}+b\vec{x}$
  7. $(ab)\vec{x}=a(b\vec{x})$
  8. $1\vec{x}=\vec{x}$

Por tanto, y recordando lo visto en la Lección 1, sabemos que el conjunto de números reales $\mathbb{R}$, el conjunto de listas ordenadas $\mathbb{R}^n$ y el conjunto de matrices $\mathbb{R}^{m\times n}$ (junto con las operaciones de suma y producto por escalares definidas respectivamente en cada conjunto) son espacios vectoriales.

Ejercicio 34. Demuestre que
(a) Si $\vec{z}+\vec{x}=\vec{x}$ para cualquier $\vec{x}$, entonces necesariamente $\vec{z}=\vec{0}$.

[1] Objetos del mismo tipo: o números, o listas de números, o matrices, o funciones, o polinomios, o variables aleatorias, etc.
[2] Por tanto, según qué tipo de objetos contenga $\mathcal{V}$, resultará que $\vec{x}$ será un número, o una lista de números, etc.
[3] Véase la Sección 6.A.1 en la página 96 del apéndice a esta lección.
[4] Y por tanto $\mathcal{V}\neq\emptyset$

(b) Si $\vec{z} + \vec{x} = \vec{0}$, entonces $\vec{z} = -\vec{x}$.

Por tanto, en todo espacio vectorial los vectores $\vec{0}$ y, para cada $\vec{x}$, el vector opuesto son únicos.

**Notación.** Para denotar espacios vectoriales genéricos, usaremos caracteres como: $\mathcal{A}$, $\mathcal{B}$, ..., $\mathcal{Y}$, $\mathcal{Z}$... y para denotar sus vectores (sus elementos), usaremos $\vec{a}$, $\vec{w}$, $\vec{0}$, etc.

No obstante, en el caso particular de $\mathbb{R}^n$ denotamos sus vectores con $\boldsymbol{a}$, $\boldsymbol{w}$, $\mathbf{0}$, etc., y en el caso de $\mathbb{R}^{m\times n}$ los denotamos con $\mathbf{A}$, $\mathbf{W}$, $\mathbf{0}$, etc.

☞ Fíjese que la definición de espacio vectorial es abstracta, ya que nada dice sobre la naturaleza de los objetos que constituyen el conjunto $\mathcal{V}$, ni tampoco sobre la definición de las operaciones *suma de vectores* y *producto de un escalar por un vector*; solo se enumeran las propiedades que deben verificar dichas operaciones.

Consecuentemente, para definir un espacio vectorial concreto necesitaremos indicar qué elementos constituyen el conjunto y definir las operaciones (suma y producto por escalares) de manera que verifiquen las propiedades indicadas. Implícitamente así lo hicimos en la Lección 1.

A modo de ejemplo, en el siguiente ejercicio se define un espacio vectorial sobre los números reales positivos, definiendo las dos operaciones de una manera completamente distinta de la habitual.

EJERCICIO 35. Considere el conjunto de números reales *positivos*, $\mathbb{R}^+$, junto con las siguientes definiciones: suma de $\vec{x}$ e $\vec{y}$ es $xy$; y el producto de un escalar $c$ por el vector $\vec{x}$ es $x^c$, es decir,

$$\begin{array}{rcl} _ + _ : \mathbb{R}^+ \times \mathbb{R}^+ & \longrightarrow & \mathbb{R}^+ \\ (\vec{x}, \vec{y}) & \longmapsto & xy \end{array} \qquad \text{y} \qquad \begin{array}{rcl} _ \cdot _ : \mathbb{R} \times \mathbb{R}^+ & \longrightarrow & \mathbb{R}^+ \\ (c, \vec{x}) & \longmapsto & x^c \end{array}$$

Demuestre que $\mathbb{R}^+$ junto con las operaciones indicadas es un espacio vectorial. ¿Quién es $\vec{0}$ en este caso?

### 6.1.1. Algunas propiedades

Hay algunas propiedades muy elementales que se cumplen en cualquier espacio vectorial $\mathcal{V}$ y que se deducen de las ocho propiedades descritas anteriormente: para el vector nulo $\vec{0} \in \mathcal{V}$ y cualquier escalar $a$:

$$a\,\vec{0} = a(\vec{0} + \vec{0}) = a\,\vec{0} + a\,\vec{0};$$

y restando $-(a\,\vec{0})$ a ambos lados deducimos que $\vec{0} = a\,\vec{0}$. Y para el escalar 0 y cualquier vector $\vec{x} \in \mathcal{V}$:

$$0\vec{x} = (0+0)\vec{x} = 0\vec{x} + 0\vec{x};$$

y restando $-(0\vec{x})$ a ambos lados deducimos que $\vec{0} = 0\vec{x}$. Además, $(-1)\vec{x} = -\vec{x}$, puesto que

$$\vec{x} + (-1)\vec{x} = 1\vec{x} + (-1)\vec{x} = (1-1)\vec{x} = 0\vec{x} = \vec{0}.$$

Por último, si $a\vec{v} = \vec{0}$, hay dos casos posibles. O bien $a = 0$; o si es distinto de cero, multiplicando ambos lados por el inverso del escalar, $a^{-1}$, tenemos $a^{-1}(a\vec{v}) = a^{-1}\,\vec{0}$. Desarrollando el lado izquierdo llegamos a $a^{-1}(a\vec{v}) = (a^{-1}a)\vec{v} = 1\vec{v} = \vec{v}$; y el lado derecho es $\vec{0}$, puesto que $\vec{0}$ por cualquier escalar es $\vec{0}$. Así llegamos a que si $a \neq 0$ entonces $\vec{v} = \vec{0}$. Por tanto, acabamos de demostrar el siguiente

**Teorema 6.1.1.** *Si $\mathcal{V}$ es un espacio vectorial*

1. $\vec{0} = a\,\vec{0}$ *para* $\vec{0} \in \mathcal{V}$ *y* $a \in \mathbb{R}$.
2. $\vec{0} = 0\vec{x}$ *para* $\vec{x} \in \mathcal{V}$ *y* $0 \in \mathbb{R}$.
3. $(-1)\vec{x} = -\vec{x}$ *para todo* $\vec{x} \in \mathcal{V}$.
4. *Si* $a\vec{v} = \vec{0}$ *o bien* $a = 0$ *o bien* $\vec{v} = \vec{0}$.

Particularizando a los espacios $\mathbb{R}^n$ y $\mathbb{R}^{m\times n}$ tenemos que:

En $\mathbb{R}^n$

1. $\mathbf{0} = a\mathbf{0}$ para $\mathbf{0} \in \mathbb{R}^n$ y $a \in \mathbb{R}$.
2. $\mathbf{0} = 0\boldsymbol{x}$ para $\boldsymbol{x} \in \mathbb{R}^n$ y $0 \in \mathbb{R}$.
3. $(-1)\boldsymbol{x} = -\boldsymbol{x}$ para todo $\boldsymbol{x} \in \mathbb{R}^n$.
4. Si $a\boldsymbol{x} = \mathbf{0}$ o bien $a = 0$ o bien $\boldsymbol{x} = \mathbf{0}$.

En $\mathbb{R}^{m\times n}$

1. $\mathbf{0} = a\mathbf{0}$ para $\mathbf{0} \in \mathbb{R}^{m\times n}$ y $a \in \mathbb{R}$.
2. $\mathbf{0} = 0\mathbf{A}$ para $\mathbf{A} \in \mathbb{R}^{m\times n}$ y $0 \in \mathbb{R}$.
3. $(-1)\mathbf{A} = -\mathbf{A}$ para todo $\mathbf{A} \in \mathbb{R}^{m\times n}$.
4. Si $a\mathbf{A} = \mathbf{0}$ o bien $a = 0$ o bien $\mathbf{A} = \mathbf{0}$.

☞ Hemos visto que el conjunto de matrices $m$ por $n$ ($\mathbb{R}^{m\times n}$), las listas ordenadas de $n$ números ($\mathbb{R}^n$) y el conjunto de números reales $\mathbb{R}$, comparten una estructura común: la estructura de *espacio vectorial*.

Dicha estructura consiste en un conjunto de objetos junto con las operaciones de *suma* y *producto por escalares* que verifican las mismas propiedades que ya vimos en $\mathbb{R}^n$ para la suma y producto por escalares en la Lección 1. Los elementos de un espacio vectorial $\mathcal{V}$ se denominan *vectores* de $\mathcal{V}$.

En la Definición 2.2 de la Página 28 definimos las *combinaciones lineales* de vectores de $\mathbb{R}^n$. Ahora daremos una definición de combinación lineal para vectores de un espacio vectorial genérico $\mathcal{V}$:

**Definición 6.2** (Combinación lineal)**.** *Sea* $\left[\vec{v}_1;\ \vec{v}_2;\ \ldots\ \vec{v}_n;\right]$ *un sistema de* $n$ *vectores de* $\mathcal{V}$*. Llamamos combinación lineal a cualquier suma de múltiplos de dichos vectores:*

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n$$

*donde los "*$a_i$*" son los* coeficientes *de la combinación lineal.*

☞ En este curso trabajaremos principalmente con el espacio vectorial $\mathbb{R}^n$, pero debe saber que existen otros conjuntos que (con una adecuada definición de suma y producto por escalares) poseen estructura de espacio vectorial. Usted se topará con ellos en otras asignaturas (véase el final de la Sección 6.3.3).

## 6.2. Funciones lineales

☞ El segundo concepto abstracto de la lección son las funciones lineales...

A lo largo de estas lecciones hemos indicado en numerosas ocasiones que ciertos operadores eran *lineales*.[5] En cada una de esas ocasiones lo que hemos visto es un ejemplo de función lineal.[6]

**Definición 6.3** (Función lineal)**.** *Sean* $\mathcal{D}$ *y* $\mathcal{V}$ *dos espacios vectoriales. Decimos que la función* $f\colon \mathcal{D} \to \mathcal{V}$ *es* lineal *si satisface las siguientes propiedades:*

1. *Para todo* $\vec{x}, \vec{y} \in \mathcal{D}$, $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$.
2. *Para todo* $\vec{x} \in \mathcal{D}$ *y para todo* $\alpha \in \mathbb{R}$, $f(\alpha\vec{x}) = \alpha f(\vec{x})$.

☞ Repasemos las operaciones empleadas hasta el momento y veremos que casi todas son funciones lineales...

### 6.2.1. Ejemplos de funciones lineales que ya hemos usado

*Ejemplo* 13. Las definiciones de suma y producto por escalares dadas en la Sección 1.2 en la página 6 hacen que la función $f(\boldsymbol{x}) = \boldsymbol{x}_{|i}$ que selecciona la componente $i$-ésima de un vector de $\mathbb{R}^n$ sea una función lineal.

[5] Consideraremos como sinónimos los términos: función lineal, aplicación lineal, transformación lineal u operador lineal.
[6] Se supone que usted conoce las funciones, pero dispone del Apéndice 6.A en la página 95 donde se expone todo lo necesario.

Podemos describir la función "*selector de una componente*", $f(\boldsymbol{x}) = \boldsymbol{x}_{|i}$, con el siguiente esquema:

$$\begin{aligned} -_{|i} : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\longmapsto x_i \end{aligned}$$

Sabemos que dicha función es lineal puesto que

$$\begin{aligned} (\boldsymbol{x}+\boldsymbol{y})_{|i} &= \boldsymbol{x}_{|i} + \boldsymbol{y}_{|i} \\ (\lambda\boldsymbol{x})_{|i} &= \lambda(\boldsymbol{x}_{|i}). \end{aligned}$$

EJERCICIO 36. Describa con un esquema (como en el ejemplo anterior) las siguientes funciones lineales, e indique (como en el ejemplo anterior) por qué son funciones lineales:

(a) La selección de la columna $j$-ésima de una matriz: $\mathbf{A} \to \mathbf{A}_{|j}$ (Página 20)
(b) La selección de la componente de la fila $i$-ésima y columna $j$-ésima de una matriz: $\mathbf{A} \to {}_{i|}\mathbf{A}_{|j}$ (Página 21)
(c) La selección de la fila $i$-ésima de una matriz: $\mathbf{A} \to {}_{i|}\mathbf{A}$ (Página 23)
(d) La transposición de una matriz: $\mathbf{A} \to \mathbf{A}^\intercal$ (Página 23)
(e) El producto de una matriz por el vector $\boldsymbol{b}$ por su derecha: $\mathbf{A} \to \mathbf{A}\boldsymbol{b}$ (Página 31)
(f) El producto de un vector por una matriz $\mathbf{A}$ a su izquierda: $\boldsymbol{b} \to \boldsymbol{b}\mathbf{A}$ (Página 28)
(g) El producto de una matriz por el vector $\boldsymbol{a}$ por su izquierda: $\mathbf{B} \to \boldsymbol{a}\mathbf{B}$ (Página 31)
(h) El producto de un vector por una matriz $\mathbf{B}$ a su derecha: $\boldsymbol{a} \to \boldsymbol{a}\mathbf{B}$ (Página 33)
(i) El producto de una matriz por una matriz $\mathbf{B}$ a su derecha: $\mathbf{A} \to \mathbf{A}\mathbf{B}$ (Página 37)
(j) El producto de una matriz por una matriz $\mathbf{A}$ a su izquierda: $\mathbf{B} \to \mathbf{A}\mathbf{B}$ (Página 37)
(k) Una transformación elemental $\boldsymbol{\tau}$ de las columnas de una matriz: $\mathbf{A} \to \mathbf{A}_{\boldsymbol{\tau}}$ (Página 54)
(l) Una transformación elemental $\boldsymbol{\tau}$ de las filas de una matriz: $\mathbf{A} \to {}_{\boldsymbol{\tau}}\mathbf{A}$ (Página 62)

☞ ¡Un título alternativo para las cuatro primeras lecciones podría ser... *"Ejemplos de funciones lineales"*!

### 6.2.2. Algunas propiedades de las funciones lineales

A continuación, vamos a ver algunos resultados básicos de las funciones lineales. Empecemos con la *composición* y la *inversa* de funciones lineales:

EJERCICIO 37. Demuestre las siguientes proposiciones (ambas demostraciones son muy similares):

(a) **Proposición 6.2.1.** *La composición $g \circ f$ de dos funciones lineales $f : \mathcal{D} \to \mathcal{V}$ y $g : \mathcal{V} \to \mathcal{W}$ es una función lineal.*
(b) **Proposición 6.2.2.** *Si $f : \mathcal{D} \to \mathcal{V}$ es lineal e invertible, su inversa $f^{-1} : \mathcal{V} \to \mathcal{D}$ también es lineal.*

¡En las primeras lecciones nos hemos topado con composiciones de funciones lineales! Por ejemplo:

- La selección de la componente de la fila $i$-ésima y columna $j$-ésima de una matriz $f(\mathbf{A}) = {}_{i|}\mathbf{A}_{|j}$:

  $$ {}_{i|}-_{|j} : \mathbb{R}^{m\times n} \to \mathbb{R}$$

  es la composición de la función que selecciona la $i$-ésima fila de la matriz, con la función que selecciona la componente $j$-ésima de la fila.

  $$ {}_{i|}-_{|j} : \mathbb{R}^{m\times n} \xrightarrow{\ i|-\ } \mathbb{R}^n \xrightarrow{\ -|j\ } \mathbb{R}$$

  (de igual modo, también es la composición de la función que selecciona la $j$-ésima columna de la matriz, con la función que selecciona la componente $i$-ésima de dicha columna.)

- El producto de dos matrices por un vector $\mathbf{AB}\boldsymbol{x}$ es un ejemplo de composición de funciones lineales; pues si $f_{\mathbf{A}}(\boldsymbol{x}) = \mathbf{A}\boldsymbol{x}$ y $f_{\mathbf{B}}(\boldsymbol{x}) = \mathbf{B}\boldsymbol{x}$; entonces

$$\big[f_{\mathbf{A}} \circ f_{\mathbf{B}}\big](\boldsymbol{x}) \;=\; f_{\mathbf{A}}\big(f_{\mathbf{B}}(\boldsymbol{x})\big) \;=\; f_{\mathbf{A}}\big(\mathbf{B}\boldsymbol{x}\big) \;=\; \mathbf{AB}\boldsymbol{x} \;=\; f_{\mathbf{AB}}(\boldsymbol{x}).$$

  Fíjese que, aunque en el ejemplo anterior el orden en el que aparecen compuestas las funciones coincide con el orden en el que aparecen las matrices en el producto, ocurre lo contrario cuando el vector aparece a la izquierda del producto: si $g_{\mathbf{A}}(\boldsymbol{x}) = \boldsymbol{x}\mathbf{A}$ y $g_{\mathbf{B}}(\boldsymbol{x}) = \boldsymbol{x}\mathbf{B}$, entonces

$$\big[g_{\mathbf{A}} \circ g_{\mathbf{B}}\big](\boldsymbol{x}) \;=\; g_{\mathbf{A}}\big(g_{\mathbf{B}}(\boldsymbol{x})\big) \;=\; g_{\mathbf{A}}\big(\boldsymbol{x}\mathbf{B}\big) \;=\; \boldsymbol{x}\mathbf{BA} \;=\; g_{\mathbf{BA}}(\boldsymbol{x}).$$

  (Otros ejemplos son el producto de varias matrices: tanto $\mathbf{ABX}$ como $\mathbf{XBA}$.)

- Una sucesión de $k$ transformaciones elementales de las columnas $\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ es una composición de $k$ funciones lineales; y lo mismo se aplica a una sucesión de transformaciones de las filas … o también a una sucesión de transformaciones elementales tanto de las filas como de las columnas (puesto que las transformaciones elementales sobre $\mathbf{A}$ se pueden representar como un producto de varias matrices).

¡También nos hemos topado con ejemplos de funciones lineales invertibles!:

- Las transformaciones elementales.
- Cuando $\mathbf{A}$ es invertible, la función $f_{\mathbf{A}}$ es invertible; en particular, $\big(f_{\mathbf{A}}\big)^{-1} = f_{\mathbf{A}^{-1}}$.

  *Explicación*: por un lado $f_{\mathbf{A}}\big(f_{\mathbf{A}^{-1}}(\boldsymbol{x})\big) = \mathbf{A}\mathbf{A}^{-1}\boldsymbol{x} = \boldsymbol{x}$ y, por otro lado, $f_{\mathbf{A}^{-1}}\big(f_{\mathbf{A}}(\boldsymbol{y})\big) = \mathbf{A}^{-1}\mathbf{A}\boldsymbol{y} = \boldsymbol{y}$. (Proposición 6.A.1 en la página 97 del apéndice de la lección).

## 6.3. Creación de nuevos espacios vectoriales

☞ Comencemos con la creación de espacios vectoriales dentro de espacios vectoriales.

### 6.3.1. Subespacios

Algunos subconjuntos de un espacio vectorial $\mathcal{V}$ (con las mismas operaciones definidas en $\mathcal{V}$) son espacios vectoriales por sí mismos. Dichos subconjuntos se denominan *subespacios*.

**Definición 6.4.** *Un subconjunto $\mathcal{W}$ no vacío de un espacio vectorial $\mathcal{V}$ es un* subespacio *de $\mathcal{V}$ si es cerrado para la suma y el producto por escalares. Es decir, si para todo $\vec{x}$ e $\vec{y}$ en $\mathcal{W}$ y todo escalar $a$,*

1. *$\vec{x} + \vec{y}$ está en el subespacio $\mathcal{W}$*
2. *$a\vec{x}$ está en el subespacio $\mathcal{W}$*

Así, para demostrar que $\mathcal{W}$ es un subespacio de $\mathcal{V}$ basta comprobar que es cerrado tanto para la suma como para el producto por escalares. Alternativamente, también basta demostrar que *es cerrado para las combinaciones lineales*, pues si para todo $\vec{x}$ e $\vec{y}$ en $\mathcal{W}$ y para todo $a, b \in \mathbb{R}$

$$a\vec{x} + b\vec{y} \quad \text{está en el subespacio } \mathcal{W}$$

entonces con $a = b = 1$ satisfacemos la primera condición de la definición y con $b = 0$ la segunda. Por tanto, decir que *el conjunto es cerrado para las combinaciones lineales* es equivalente a decir que el conjunto es cerrado para la suma y el producto por escalares.

Ejercicio 38. ¿Es el conjunto de combinaciones lineales de $n$ vectores un subespacio?

Fíjese que si $W$ es un subconjunto de un espacio vectorial $\mathcal{V}$; sabremos inmediatamente que $W$ no es un subespacio si $\vec{0}$ no pertenece a $W$, o si pertenece $\vec{x}$ pero no $-\vec{x}$.

☞ Un subespacio es un espacio vectorial que está contenido dentro de otro espacio vectorial.

Para comprobar que $\mathcal{W}$ es subespacio basta verificar si es cerrado para la suma y el producto por escalares (es decir, que es cerrado para las combinaciones lineales).

**Ejemplos de subespacios**

- Los vectores de $\mathbb{R}^3$ cuya primera componente es cero.

  Si sumamos dos vectores con primera componente nula, obtenemos otro vector con la primera componente nula; y lo mismo ocurre con cualquier múltiplo de un vector con primera componente nula.

- Las funciones reales ($\mathbb{R} \to \mathbb{R}$) que además son continuas (véase la Sección 6.3.3 en la página siguiente).

  Si sumamos dos funciones reales y continuas, obtenemos otra función real y continua; y lo mismo ocurre si multiplicamos una función real y continua por un escalar.

- El conjunto de polinomios ($\mathbb{R}[x]$) de grado menor o igual a $n$.
- Variables aleatorias con esperanza cero: su suma tiene esperanza cero y cualquier múltiplo también.
- Variables aleatorias independientes y con distribución normal (gaussiana).

☞ Nótese que como todo conjunto está contenido en sí mismo, todo espacio vectorial es un *subespacio* de sí mismo. Por tanto, para cualquier espacio vectorial $\mathcal{V}$, son subespacios tanto $\mathcal{V}$ como el subconjunto $\{\vec{0}\}$ que solo contiene el vector nulo (pues obviamente es cerrado para la suma y el producto por escalares).

**Intersección de subespacios**

La intersección de subespacios es un subespacio:

**Teorema 6.3.1.** *Sean $\mathcal{V}_1$ y $\mathcal{V}_2$ subespacios de $\mathcal{S}$, entonces la intersección de $\mathcal{W} = \mathcal{V}_1 \cap \mathcal{V}_2$ también lo es.*

*Demostración.* Sabemos que la intersección es no vacía pues el $\vec{0}$ está en ambos subespacios. Así que basta probar que la intersección es cerrada para las combinaciones lineales de dos vectores. Para ello tomamos $a$ y $b$ de $\mathbb{R}$ y $\vec{x}$ e $\vec{y}$ de $\mathcal{W}$. Como $\mathcal{W}$ es la intersección, entonces $\vec{x}$ e $\vec{y}$ pertenecen tanto a $\mathcal{V}_1$ como a $\mathcal{V}_2$.

- Por ser $\mathcal{V}_1$ un subespacio (i.e., cerrado para las comb. lineales) sabemos que $a\vec{x} + b\vec{y} \in \mathcal{V}_1$.
- Por ser $\mathcal{V}_2$ un subespacio (i.e., cerrado para las comb. lineales) sabemos que $a\vec{x} + b\vec{y} \in \mathcal{V}_2$.

Por tanto, si $\vec{x}$ y $\vec{y}$ pertenecen a $\mathcal{W}$, entonces $a\vec{x} + b\vec{y} \in \mathcal{V}_1 \cap \mathcal{V}_2 \;=\; \mathcal{W}$. □

Sin embargo, la unión de subespacios no suele ser cerrada para la suma (piense en la unión del conjunto de *vectores de $\mathbb{R}^3$ cuya primera componente es cero* y el conjunto de *vectores de $\mathbb{R}^3$ cuya última componente es cero*; si sumamos un vector del primer grupo con uno del segundo, generalmente obtendremos un vector en el que no es cero ni la primera, ni la última componente (al sumar vectores nos "salimos" del conjunto).

### 6.3.2. Espacios vectoriales de productos cartesianos

☞ Y ahora una forma de construir espacios vectoriales "más grandes" a partir de otros espacios vectoriales.

El *Producto Cartesiano* de dos conjuntos $A$ y $B$, que denotamos con $A \times B$, es el conjunto de pares ordenados $(a, b)$ donde $a \in A$ y $b \in B$:

$$A \times B = \{\, (a, b) \mid a \in A \;\text{ y }\; b \in B \,\}.$$

Si $\mathcal{A}$ y $\mathcal{B}$ son espacios vectoriales, definimos la suma de pares ordenados del conjunto $\mathcal{A} \times \mathcal{B}$ como el par ordenado formado por la suma de las primeras componentes y la suma de las segundas:

$$(\vec{a}, \vec{b}) + (\vec{c}, \vec{d}) = (\vec{a} + \vec{c},\ \vec{b} + \vec{d}); \tag{6.1}$$

donde es importante hacer notar que el símbolo "+" de la izquierda representa la suma de pares ordenados; pero que los símbolos "+" de la derecha representan las sumas de los vectores de los respectivos espacios vectoriales. Por ejemplo, si consideramos pares del producto cartesiano $\mathbb{R}^n \times \mathbb{R}^{p\times q}$, entonces

$$(\boldsymbol{x}, \mathbf{A}) + (\boldsymbol{y}, \mathbf{B}) = (\boldsymbol{x} + \boldsymbol{y},\ \mathbf{A} + \mathbf{B})$$

evidentemente los tres tipos de sumas, lo son de objetos muy distintos: pares, vectores de $\mathbb{R}^n$ y matrices.

Y definimos el producto de un escalar $\lambda$ por un par ordenado como el par ordenado que resulta de multiplicar cada componente por el escalar $\lambda$:

$$\lambda(\vec{a}, \vec{b}) = (\lambda\vec{a}, \lambda\vec{b}); \tag{6.2}$$

haciendo la misma advertencia que con las sumas en la definición anterior; piense en el siguiente ejemplo:

$$\lambda(\boldsymbol{x}, \mathbf{A}) = (\lambda\boldsymbol{x},\ \lambda\mathbf{A}).$$

Ejercicio 39. Demuestre la siguiente

**Proposición 6.3.2.** *El producto cartesiano $\mathcal{A} \times \mathcal{X}$ de los espacios vectoriales $\mathcal{A}$ y $\mathcal{X}$, junto con la suma de pares ordenados (6.1) y el producto de un par ordenado por un escalar (6.2) es un espacio vectorial.*

La generalización a $k$ conjuntos $A, \ldots, M$ del *producto cartesiano* (que denotamos con $A \times \cdots \times M$) es el conjunto de $k$-tuplas ordenadas $(a, b, \ldots, m)$ donde $a \in A$, $b \in B$, $\ldots$ y $m \in M$:

$$A \times B \times \cdots \times M = \{\, (a, b, \ldots, m) \mid a \in A,\ b \in B, \ldots \text{ y } m \in M \,\}.$$

Pero alternativamente podemos considerar esta generalización como el producto cartesiano del primer conjunto con el producto cartesiano de los $n - 1$ restantes, es decir:

$$A \times B \times \cdots \times M = A \times (B \times \cdots \times M).$$

*Ejemplo* 14.

$$A \times B \times C \times D = A \times \big(B \times (C \times D)\big).$$

Empleando dicha visión alternativa podemos aprovechar la Proposición 6.3.2 para saber que el producto cartesiano $\mathcal{A} \times \mathcal{B} \times \cdots \mathcal{X}$ de los espacios vectoriales $\mathcal{A}, \mathcal{B}, \ldots, \mathcal{X}$, junto con la suma de pares ordenados (6.1) y el producto de un par ordenado por un escalar (6.2) es un espacio vectorial.

### 6.3.3. Espacios vectoriales de funciones

☞ La *suma de funciones* y el *producto de un escalar por una función* son potentes herramientas para crear nuevos espacios vectoriales. Algunos de ellos aparecerán en otras asignaturas.

Definamos la suma y el producto por escalares para funciones con el mismo dominio $X$ y cuyas imágenes están contenidas en el mismo espacio vectorial $\mathcal{V}$:

**Definición 6.5.** *Se define la suma de dos funciones $f\colon X \to \mathcal{V}$ y $g\colon X \to \mathcal{V}$ como la función*

$$\begin{aligned}[f+g]\colon X &\longrightarrow \mathcal{V} \\ \vec{x} &\longmapsto f(\vec{x}) + g(\vec{x})\end{aligned}. \tag{6.3}$$

**Definición 6.6.** *Se define el producto de un escalar $\alpha$ por una función $f\colon X \to \mathcal{V}$ como la función*

$$\begin{aligned}[\alpha \cdot f]\colon X &\longrightarrow \mathcal{V} \\ \vec{x} &\longmapsto \alpha\ f(\vec{x})\end{aligned}. \tag{6.4}$$

Con $\mathcal{V}^D$ denotaremos el conjunto de funciones cuyo dominio es $D$ y cuya imagen está contenida en $\mathcal{V}$:

$$\mathcal{V}^D = \{f \mid f\colon D \to \mathcal{V}\}\,. \tag{6.5}$$

Ejercicio 40. Demuestre la siguiente

**Proposición 6.3.3.** *El conjunto $\mathcal{V}^D$ de funciones cuyo dominio es $D$ y cuya imagen está contenida en $\mathcal{V}$, junto con la suma (6.3) y el producto por escalares (6.4), es un espacio vectorial.*

**Ejemplos de este tipo de espacios que probablemente encontrará en futuras asignaturas**

... aunque quizá no le indiquen explícitamente que está trabajando con espacios vectoriales:

- $\mathbb{R}^{\mathbb{R}}$: funciones reales de variable real: $f\colon \mathbb{R} \to \mathbb{R}$.
- $\mathbb{R}^{(\mathbb{R}^n)}$: funciones reales de varias variables reales: $f\colon \mathbb{R}^n \to \mathbb{R}$.
- $\mathbb{R}^{\mathbb{N}}$: sucesiones infinitas de números reales (con índice sobre los números naturales): $\mathbf{a}\colon \mathbb{N} \to \mathbb{R}$.
- $\mathbb{R}^{\mathbb{Z}}$: sucesiones infinitas de números reales (con índice sobre los números enteros): $\mathbf{a}\colon \mathbb{Z} \to \mathbb{R}$.
- $\mathbb{R}^{\Omega}$: variables aleatorias (donde $\Omega$ es el conjunto de sucesos elementales): $X\colon \Omega \to \mathbb{R}$.
- $\left(\mathbb{R}^{\Omega}\right)^{\mathbb{Z}}$: procesos estocásticos: $\mathbf{Y}\colon \mathbb{Z} \to \mathbb{R}^{\boldsymbol{\Omega}}$.
- $\mathbb{R}^N$: series temporales con un tamaño muestral $N$.
- Etcétera.

### 6.3.4. Subespacio de funciones lineales

Ejercicio 41. Demuestre las siguientes proposiciones (ambas demostraciones son muy similares):

(a) **Proposición 6.3.4.** *La suma de dos funciones lineales es una función lineal.*
(b) **Proposición 6.3.5.** *El producto de un escalar por una función lineal es una función lineal.*

Conforme con lo anterior, el conjunto de funciones lineales que van del espacio vectorial $\mathcal{D}$ al espacio vectorial $\mathcal{V}$ es un subespacio del espacio vectorial $\mathcal{V}^{\mathcal{D}}$ (el conjunto de funciones cuyo dominio es $\mathcal{D}$ y cuya imagen está contenida en el espacio vectorial $\mathcal{V}$). Cerramos la lección enunciando este resultado en el siguiente

**Corolario 6.3.6.** *Dados $\mathcal{D}$ y $\mathcal{V}$, el conjunto de funciones lineales $f\colon \mathcal{D} \to \mathcal{V}$ es un subespacio de $\mathcal{V}^{\mathcal{D}}$.*

(hemos citado algunos ejemplos en la Sección 6.3.1)

## 6.4. Generalizaciones o visiones alternativas (★)

☞ Hay una estrecha relación entre subespacios y funciones lineales que no se suele contar.

### 6.4.1. Subespacios y funciones lineales

**Proposición 6.4.1.** *Sea $f \subset \mathcal{V} \times \mathcal{W}$ donde $\mathcal{V}$ y $\mathcal{W}$ son subespacios. Las siguientes propiedades son equivalentes:*

1. *$f\colon \mathcal{V} \to \mathcal{W}$ es una aplicación lineal.*
2. *$f$ es un subespacio vectorial de $\mathcal{V} \times \mathcal{W}$ suplementario[7] de $\{\vec{0}\} \times \mathcal{W}$.*

*Demostración.* 1. $\Rightarrow$ 2. Veamos primero que $f$ es un subespacio vectorial de $\mathcal{V} \times \mathcal{W}$.

[7] es decir, $\mathcal{V} \times \mathcal{W} = f \oplus (\{\vec{0}\} \times \mathcal{W})$. (véase la Definición 10.2 en la página 133)

1. Dado que $f$ es una función cuyo dominio es $\mathcal{V}$ y cuya imagen está contenida en $\mathcal{W}$, cualquier elemento de $f$ es un par ordenado cuya primera componente pertenece a $\mathcal{V}$ y cuya segunda componente pertenece a $\mathcal{W}$. Por tanto, $f \subset \mathcal{V} \times \mathcal{W}$.

2. Como $\mathcal{V} \neq \emptyset$, por ser $\mathcal{V}$ un subespacio, $\text{dom}(f) = \mathcal{V} \neq \emptyset$ y, en consecuencia, $f \neq \emptyset$.

3. Si $(\vec{v}_1, \vec{w}_1)$ y $(\vec{v}_2, \vec{w}_2)$ pertenecen a $f$ y $\alpha, \beta \in \mathbb{R}$, entonces, como $f(\vec{v}_1) = \vec{w}_1$ y $f(\vec{v}_2) = \vec{w}_2$ y $f$ es lineal, se tendrá que

$$f(\alpha\vec{v}_1 + \beta\vec{v}_2) = \alpha f(\vec{v}_1) + \beta f(\vec{v}_2) = \alpha\vec{w}_1 + \beta\vec{w}_2.$$

   Por lo tanto, $(\alpha\vec{v}_1 + \beta\vec{v}_2,\ \alpha\vec{w}_1 + \beta\vec{w}_2) \in f$ y, así,

$$\alpha(\vec{v}_1, \vec{w}_1) + \beta(\vec{v}_2, \vec{w}_2) = (\alpha\vec{v}_1 + \beta\vec{v}_2, \alpha\vec{w}_1 + \beta\vec{w}_2) \in f.$$

Veamos en segundo lugar que $f$ y $\{\vec{0}\} \times \mathcal{W}$ son subespacios suplementarios:

1. Si $(\vec{v}, \vec{w}) \in f \cap (\{\vec{0}\} \times \mathcal{W})$, tenemos que $\vec{w} = f(\vec{v})$ y que $\vec{v} = \vec{0}$. Luego, por ser $f$ lineal, necesariamente $\vec{w} = \vec{0}$. Por tanto,

$$f \cap (\{\vec{0}\} \times \mathcal{W}) = (\vec{0}, \vec{0}).$$

2. Dado $(\vec{v}, \vec{w}) \in \mathcal{V} \times \mathcal{W}$, como $\mathcal{V} = \text{dom}(f)$, tenemos que

$$(\vec{v}, \vec{w}) = \big(\vec{v}, f(\vec{v})\big) + \big(\vec{0}, \vec{w} - f(\vec{v})\big) \in f + (\{\vec{0}\} \times \mathcal{W}).$$

   Por tanto, $\mathcal{V} \times \mathcal{W} = f \oplus (\{\vec{0}\} \times \mathcal{W})$.

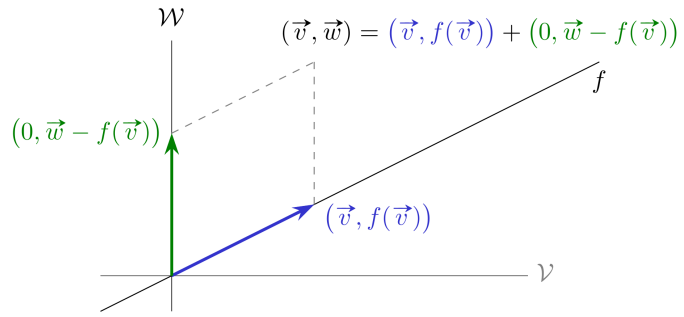



Figura 6.1: Representación esquemática de que $\mathcal{V} \times \mathcal{W} = f \oplus (\{\vec{0}\} \times \mathcal{W})$.

1. $\Leftarrow$ 2. Verificaremos primero que $f$ es una función cuyo dominio es $\mathcal{V}$ y cuya imagen está contenida en $\mathcal{W}$.

1. Como $f \subset \mathcal{V} \times \mathcal{W}$, los elementos de $f$ son pares ordenados con primeras componentes en $\mathcal{V}$ y segundas en $\mathcal{W}$.

2. Si $(\vec{v}, \vec{w}_1)$ y $(\vec{v}, \vec{w}_2)$ pertenecen a $f$, por ser $f$ un subespacio vectorial de $\mathcal{V} \times \mathcal{W}$, tendremos que su diferencia $(\vec{0}, \vec{w}_1 - \vec{w}_2) \in f$. Como esta diferencia pertenece al subespacio $\{\vec{0}\} \times \mathcal{W}$, que es suplementario de $f$, necesariamente[8] $\vec{w}_1 - \vec{w}_2 = \vec{0}$. Por tanto, $\vec{w}_1 = \vec{w}_2$, con lo que es imposible encontrar dos pares distintos en $f$ cuyas primeras componentes sean coincidentes.

[8] $f \cap (\{\vec{0}\} \times \mathcal{W}) = (\vec{0}, \vec{0})$

3. En este punto, ya sabemos que $f$ es una función cuyo dominio está contenido en $\mathcal{V}$ y cuya imagen está contenida en $\mathcal{W}$. Comprobemos que $\mathcal{V} = \mathrm{dom}(f)$.

   Si $\vec{v} \in \mathcal{V}$, como $\{\vec{0}\} \times \mathcal{W}$ y $f$ son suplementarios, el vector $(\vec{v}, \vec{0})$ se descompone en suma de un vector $(\vec{x}, \vec{y})$ de $f$ y otro $(\vec{0}, \vec{z})$ de $\{\vec{0}\} \times \mathcal{W}$.

   $$(\vec{v}, \vec{0}) = (\vec{x}, \vec{y}) + (\vec{0}, \vec{z}).$$

   Como en esta descomposición necesariamente $\vec{x} = \vec{v}$, concluimos que $\vec{v} \in \mathrm{dom}(f)$.

Solo falta comprobar que $f : \mathcal{V} \to \mathcal{W}$ es lineal.

Si $\vec{v}_1, \vec{v}_2 \in \mathcal{V}$ y $\alpha_1, \alpha_2 \in \mathbb{R}$, como $\big(\vec{v}_1, f(\vec{v}_1)\big)$ y $\big(\vec{v}_2, f(\vec{v}_2)\big)$ pertenecen al subespacio $f$, se tendrá que

$$\big(\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2, \alpha_1 f(\vec{v}_1) + \alpha_2 f(\vec{v}_2)\big) = \alpha_1\big(\vec{v}_1, f(\vec{v}_1)\big) + \alpha_2\big(\vec{v}_2, f(\vec{v}_2)\big) \in f.$$

Luego, $f(\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2) = \alpha_1 f(\vec{v}_1) + \alpha_2 f(\vec{v}_2)$.

□

## *Apéndices a la lección*

☞ Se suele asumir que los estudiantes conocen las funciones, pero creo oportuno añadir el siguiente apéndice.

## 6.A. Funciones

**Definición 6.7.** *Una función es un conjunto de pares ordenados en el que no existen dos pares distintos que tengan sus primeras componentes iguales.*

Dada una función $f$, se llama *dominio* de $f$ al conjunto de primeras componentes de los pares de $f$.

$$\mathrm{dom}(f) = \{x \mid \text{existe } y \text{ tal que } (x, y) \in f\}.$$

Dada una función $f$, se llama *imagen* de $f$ al conjunto de segundas componentes de $f$.

$$\mathrm{imag}(f) = \{y \mid \text{existe } x \text{ tal que } (x, y) \in f\}.$$

**Ejemplos.**

*Ejemplo* 15. El siguiente conjunto (en el que las segundas componentes son el cuadrado de las primeras) es una función

$$f = \{(x, x) \mid x \in \mathbb{Z}\} = \{(0,0),\ (1,1),\ (-1,1),\ (2,4),\ (-2,4),\ (3,9),\ (-3,9),\ \dots\}$$

donde $\mathbb{Z}$ es el conjunto de números enteros.

Sin embargo, este otro conjunto (en el que las primeras componentes son el cuadrado de las segundas) NO es una función:

$$\big\{(x^2, x) \mid x \in \mathbb{Z}\big\} = \{(0,0),\ (1,1),\ (1,-1),\ (4,2),\ (4,-2),\ (9,3),\ (9,-3),\ \dots\}$$

pues hay pares distintos que tienen la misma primera componente; por ejemplo $(1,1)$ y $(1,-1)$.

*Ejemplo* 16 (Función nula). La *función nula* es aquella cuyas segundas componentes son todas nulas, es decir

$$n = \{(x, 0) \mid x \in \mathrm{dom}(n)\} = \{(x,0),\ (y,0),\ (z,0), \dots\};$$

(donde con el símbolo "0" denotamos el elemento *nulo* del conjunto que contiene la *imagen* de la función —véase un ejemplo de función nula más abajo).

*Ejemplo* 17 (Función identidad). La *función identidad* es aquella cuyas segundas componentes son iguales a las primeras, es decir

$$id = \{(x, x) \mid x \in \mathrm{dom}(id)\} = \{(x,x),\ (y,y),\ (z,z), \dots\}.$$

### 6.A.1. Notación

Son habituales distintas formas de notación relacionadas con las funciones. Por ejemplo, la expresión

$$f\colon X \to Y \qquad \text{o equivalentemente} \qquad X \xrightarrow{f} Y$$

es una forma abreviada de expresar lo siguiente:

1. $f$ es una función
2. $\mathrm{dom}(f) = X$
3. $\mathrm{imag}(f) \subset Y$

Cuando se usa esta notación, el conjunto $Y$ que contiene la imagen de la función se denomina *codominio*. Por otra parte, el esquema

$$\begin{aligned} f\colon X &\longrightarrow Y \\ x &\longmapsto \textit{expresión de } x \end{aligned}$$

es una forma abreviada de expresar el siguiente conjunto de pares:

$$f = \{(x, \textit{expresión de } x) \mid x \in X\}\,.$$

De manera que

- el Ejemplo 15 en la página anterior se puede expresar como $\begin{aligned} f\colon \mathbb{Z} &\longrightarrow \mathbb{N} \\ x &\longmapsto x^2 \end{aligned}$.
- la función *nula* que asigna a todo número real el vector nulo de $\mathbb{R}^2$ se puede expresar como

$$\begin{aligned} n\colon \mathbb{R} &\longrightarrow \mathbb{R}^2, \\ x &\longmapsto \mathbf{0} \end{aligned} \quad \text{donde } \mathbf{0} = (0, 0,) \text{ es el elemento } \textit{nulo} \text{ en } \mathbb{R}^2.$$

- y podemos expresar la función *identidad* en $\mathbb{R}^3$ como $\begin{aligned} id\colon \mathbb{R}^3 &\longrightarrow \mathbb{R}^3. \\ \boldsymbol{x} &\longmapsto \boldsymbol{x} \end{aligned}$

**Uso de la notación funcional.** Posiblemente esté más familiarizado con la siguiente notación funcional

$$f(x) = y$$

que equivale a escribir $(x, y) \in f$. Así, el Ejemplo 15 en la página anterior se puede expresar como:

$$f(x) = x^2, \qquad \text{donde } x \in \mathbb{Z};$$

pues indica que cada par $(x, x^2)$ pertenece a $f$. Seguramente alguna vez haya tenido que representar gráficamente los pares $(x, x^2)$ de esta función, y sabrá que describen una parábola (Figura 6.2). La representación de los pares se denomina *grafo* de la función.

De igual modo, expresamos la función *nula* que asigna a todo número real el vector nulo de $\mathbb{R}^2$:

$$f(x) = \mathbf{0}, \qquad \text{donde } x \in \mathbb{R} \quad \text{y} \quad \mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Y podemos expresar la función *identidad* en $\mathbb{R}^3$ como:

$$f(\boldsymbol{x}) = \boldsymbol{x}, \qquad \text{donde } \boldsymbol{x} \in \mathbb{R}^3.$$

Aprovechando la *notación funcional* podemos expresar la *imagen de $f$* del siguiente modo:

$$\mathrm{imag}(f) = \{y \mid \text{existe } x \in \mathrm{dom}(f) \text{ tal que } f(x) = y\}\,. \tag{6.6}$$

A veces también se usa la notación $f(X)$ para indicar *la imagen de $f\colon X \to Y$* (donde $X$ es el dominio):

$$f(X) = \{y \mid \text{existe } x \in X \text{ tal que } f(x) = y\}\,.$$

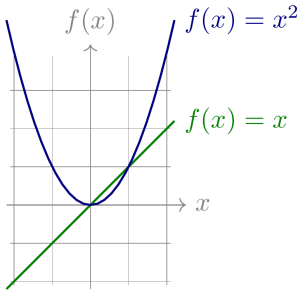



Figura 6.2: Representación de los conjuntos de pares $(x, x)$ en verde, $(x, x^2)$ en azul.

### 6.A.2. Invertibilidad

Decimos que una función es invertible si al invertir el orden de los pares se obtiene una función. En tal caso, dicha función se denota $f^{-1}$ y se llama *función inversa* de $f$.

Fíjese que de los tres ejemplos anteriores, únicamente la función identidad es invertible. En los otros dos, al invertir el orden de los pares se obtienen pares distintos pero que comparten la misma primera componente (si invertimos el orden de los pares de la función nula, todos tendrán como primera componente el vector $\mathbf{0}$; y ya vimos qué pasaba al invertir los pares ordenados del Ejemplo 15 en la página 95).

Puesto que intercambiando dos veces el orden de los pares recuperamos los pares originales, tenemos que

$$\big(f^{-1}\big)^{-1} = f.$$

También es evidente que $\text{dom}(f^{-1}) = \text{imag}(f)$, y que $\text{imag}(f^{-1}) = \text{dom}(f)$.

**Proposición 6.A.1.** *Si $f$ y $g$ son dos funciones tales que*

$$g(f(x)) = x \ \text{ para todo } \ x \in \text{dom}(f) \qquad y \qquad f(g(y)) = y \ \text{ para todo } \ y \in \text{dom}(g)$$

*entonces $f$ es invertible y su inversa es $g$.*

*Demostración.* Puesto que $f$ y $g$ son funciones, basta comprobar que $(x, y) \in f$ si y solo si $(y, x) \in g$.
Por una parte: $(x, y) \in f \ \Rightarrow\ y = f(x) \ \Rightarrow\ g(y) = g(f(x)) = x \ \Rightarrow\ (y, x) \in g$.
Y por otra: $(y, x) \in g \ \Rightarrow\ g(y) = x \ \Rightarrow\ f(g(y)) = f(x) = y \ \Rightarrow\ (x, y) \in f$. □

### 6.A.3. Composición de funciones

Si tenemos dos funciones $f$ y $g$ tales que el dominio de $g$ contiene a la imagen de $f$, llamamos *composición* de $f$ y $g$ al conjunto de pares

$$g \circ f = \big\{(x, z) \;\big|\; \text{ existe } \ z \ \text{ de modo que } \ (x, y) \in f \ \text{ y } \ (y, z) \in g\big\}.$$

La composición de funciones se escribe usando la notación funcional como

$$[g \circ f](x) \ = \ g\big(f(x)\big).$$

pues $z = [g \circ f](x)$ siempre que $y = f(x)$ y $z = g(y)$.

En consecuencia, resulta evidente que si $f$ es invertible entonces

$$[f^{-1} \circ f](x) = f^{-1}\big(f(x)\big) = x, \qquad \text{para todo } \ x \in \text{dom}(f).$$

Y es igualmente evidente que

$$[f \circ f^{-1}](y) = f\big(f^{-1}(y)\big) = y, \qquad \text{para todo } \ y \in \text{dom}(f^{-1}).$$

### 6.A.4. Restricción de una función

En ocasiones, solo queremos considerar aquellos pares de la función cuya primera componente cumple una condición. Dicho subconjunto de pares se denomina restricción de la función. Nótese que, como el conjunto inicial de pares es una función, cualquier subconjunto de dichos pares también será una función.

**Definición 6.8** (Restricción de una función a un subconjunto)**.** *Sea $f\colon X \to Y$ una función y sea $A$ un subconjunto de su dominio. El subconjunto de pares*

$$f_{|A} = \{(x, y) \mid (x, y) \in f \text{ con } x \in A \subset X\}$$

*se denomina la restricción de $f$ al subconjunto $A$. Por tanto, es una nueva función cuyo dominio es $A$, es decir $f_{|A}\colon A \to Y$.*

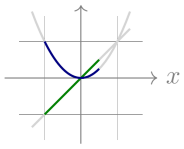



Figura 6.3: Representación de los conjuntos de pares $(x, x)$ en verde, $(x, x^2)$ en azul restringidos al subconjunto $A = \{x \mid x \in (-1, 0.5)\}$. Compárese con la Figura 6.2

### 6.A.5. Funciones inyectivas, sobreyectivas y biyectivas

Dada una función $f\colon X \to Y$

- La función es *inyectiva* (o *uno a uno*) si cada elemento de la imagen es imagen de uno solo de los elementos del dominio:

  $$\text{para todo } x, x' \in X; \quad f(x) = f(x') \implies x = x'.$$

  Es decir, si dos pares tienen segundas componentes distintas, entonces también son distintas las primeras: si $(x, y)$, $(x', y') \in f$; entonces $y \neq y' \Rightarrow x \neq x'$.

- La función es *sobreyectiva* si la imagen y el codominio son iguales $\text{imag}(f) = Y$; es decir, si todo elemento de $Y$ es imagen de algún elemento de $X$:

  $$\text{para todo } y \in Y, \text{ existe } x \in X \text{ tal que } y = f(x).$$

Una función es *biyectiva* si simultáneamente es *inyectiva* y *sobreyectiva*.

### 6.A.6. Los sistemas (listas ordenadas) son funciones

☞ Iniciamos la Lección 1 haciendo uso del concepto "intuitivo" de *lista ordenada*. Para poder hablar con propiedad de conceptos tales como *el conjunto de vectores de* $\mathbb{R}^3$ o *el conjunto de matrices de orden n* necesitamos contar con una definición formal de *sistema* (lista ordenada).

Un *sistema* (o lista ordenada) de $n$ elementos *es una función cuyo dominio es el conjunto de números naturales menores o iguales a n y mayores que cero*, es decir, cuyo dominio es el conjunto de índices $\{i \mid 1 \leq i \leq n\}$.

Así, el sistema $[\boldsymbol{a};\ 5;\ \mathbf{B};]$ es la función con dominio $\{1, 2, 3\}$ consistente en el conjunto de pares:

$$\Big\{(1, \boldsymbol{a}),\ (2, 5),\ (3, \mathbf{B})\Big\};$$

y el vector $(0,\, -2,\, \pi,\,)$ es la función (el conjunto de pares):

$$\Big\{(1,0),\ (2,-2),\ (3,\pi)\Big\}.$$

Visto así, $\mathbb{R}^3$ es el conjunto de funciones de la forma $\boldsymbol{v}\colon \{1,2,3\} \to \mathbb{R}$, es decir, el conjunto de funciones cuyo dominio es $\{1,2,3\}$ y cuya imagen está contenida en $\mathbb{R}$. De hecho, el símbolo $\mathbb{R}^3$ es (siguiendo la notación de la ecuación 6.5 en la página 93) una forma abreviada de escribir $\mathbb{R}^{\{1,2,3\}}$, donde

$$\mathbb{R}^{\{1,2,3\}} = \Big\{\boldsymbol{v} \;\Big|\; \boldsymbol{v}\colon \{1,2,3\} \to \mathbb{R}\Big\}.$$

**Listas vacías.** Una lista vacía (un sistema vacío) es cualquier función $[\ ]\colon \emptyset \to X$ cuyo dominio es el conjunto vacío $\emptyset$; consecuentemente, dicha función es *el conjunto vacío*, pues no contiene ningún par:[9]

$$[\ ] \equiv \emptyset. \tag{6.7}$$

[9] esta equivalencia matemática no se da en el lenguaje de programación Python, puesto que la implementación informática de conjuntos y listas es diferente, así que en Python un objeto lista (`list`) y un objeto conjunto (`set`) nunca pueden ser lo mismo.

# Resolviendo Ax = 0

## 7.1. Sistemas de ecuaciones lineales

☞ La Parte III del curso trata sobre sistemas de ecuaciones lineales. Como veremos enseguida, los sistemas de ecuaciones lineales están íntimamente relacionados con los subespacios de $\mathbb{R}^n$.

**Una advertencia.** Quizá conozca algún método de resolución de sistemas. Mi experiencia es que generalmente los alumnos saben ejecutar unos pasos a ciegas para obtener una solución, pero sin entender bien el método... y con cierta frecuencia no son capaces de ejecutar el método correctamente.

El método que veremos es deliberadamente distinto del que se suele enseñar. Con ello pretendo, entre otras cosas, que el estudiante no aplique ciegamente una batería de recetas (mal aprendidas en muchos casos). Así que, *por el momento olvide lo que sabe y trate de entender el método que expondré aquí.*

Afortunadamente, el método es muy sencillo y aplicable a otros problemas que veremos más adelante... y para mayor fortuna, ¡usted ya lo ha empleado en la Lección 5 para invertir matrices!

Llamamos *ecuación lineal* a aquella que se puede escribir de la siguiente forma

$$a_1x_1 + \cdots + a_nx_n = b,$$

donde $b$ y los *coeficientes* $a_1, \dots a_n$ denotan números fijos y $x_1, \dots x_n$ son las variables, es decir, etiquetas para ser reemplazadas por números.
Llamamos *solución* a los valores que, reemplazando a las variables, hacen cierta la igualdad.

**Sistemas de ecuaciones lineales** Se llama sistema de ecuaciones lineales a una colección de ecuaciones en la que cada variable es reemplazada por un mismo valor en todas las ecuaciones donde aparece. Por ejemplo,

$$\begin{cases} x + y - 2z = 1 \\ x - 2y + z = 2 \\ -2x + y + z = 3 \end{cases},$$

es un sistema de tres ecuaciones con tres variables (o incógnitas) $x$, $y$ y $z$. La llave indica que cada variable debe ser reemplazada por el mismo valor en cada una de las tres ecuaciones; es decir, si asignamos el valor $x = 5$, lo hacemos para las tres ecuaciones a la vez. No hay un orden ni en la disposición de las ecuaciones ni en las sumas dentro de cada ecuación; de manera que el sistema anterior también lo podemos escribir como

$$\begin{cases} x + z - 2y = 2 \\ y + x - 2z = 1 \\ z + y - 2x = 3 \end{cases},$$

pero esta libertad en la notación tradicional no ayuda a trabajar con ellos.

### 7.1.1. Notación matricial de los sistemas de ecuaciones lineales

Aquí emplearemos la notación matricial en lugar de la tradicional. Ello nos permitirá aprovechar fácilmente toda la potencia de los conceptos de espacio vectorial.

El primer paso para usar la notación matricial requiere tomar una decisión inicial y arbitraria. Hemos de establecer quién es la primera variable, quién la segunda y quién la tercera; y también decidir qué ecuación será la primera, cuál la segunda y cuál la tercera. La decisión que tomemos no es importante, pero una vez tomada, estamos obligados a mantenerla.

Volviendo al ejemplo anterior, podemos decidir que el orden de las variables es, primero $x$, luego $y$ y por último $z$; así definimos el *vector de incógnitas*: $\boldsymbol{x} = (x, y, z, )$. También podemos establecer que las ecuaciones serán ordenadas tal como aparecen dentro de la llave en la primera versión del ejemplo. Siguiendo este criterio, generamos una *matriz de coeficientes* cuya primera fila contiene los coeficientes de la primera ecuación (y en el orden que hayamos establecido para las variables), cuya segunda fila contiene los coeficientes de la segunda ecuación (respetando el mismo orden de las variables), etc.

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 2 \end{bmatrix}.$$

El orden de las ecuaciones también determina el orden de las componentes del vector que contiene los números que aparecen a la derecha de cada ecuación. En el ejemplo anterior será $\boldsymbol{b} = (1, 2, 3, )$. A este vector lo denominamos *vector del lado derecho*.

Así, con la siguiente ecuación matricial:

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b},$$

denotamos de manera compacta un sistema de ecuaciones, y donde la variable $\boldsymbol{x}$ es una etiqueta para ser reemplazada por un vector. Así, con

$$\begin{bmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 2 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{denotamos} \quad \begin{cases} x + y - 2z = 1 \\ x - 2y + z = 2 \\ -2x + y + z = 3 \end{cases}.$$

☞ La ecuación matricial $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ define un *sistema de ecuaciones lineales*.

Decimos que el vector $\boldsymbol{z}$ es *una solución* del sistema de ecuaciones $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ si reemplazar $\boldsymbol{x}$ por $\boldsymbol{z}$ hace cierta la igualdad; y llamamos "*resolver*" un sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ a encontrar el conjunto de sus soluciones.

## 7.2. Sistemas de ecuaciones lineales homogéneos

Cuando el vector del lado derecho es un vector nulo, el sistema de ecuaciones se dice que es *homogéneo*. En esta Lección 7 solo nos ocuparemos de los sistemas de ecuaciones que son homogéneos:

$$\mathbf{A}\boldsymbol{x} = \mathbf{0}.$$

Fíjese que el lado izquierdo de un sistema de ecuaciones está formado por el producto $\mathbf{A}\boldsymbol{x}$, es decir, *el lado izquierdo es una combinación lineal* de las columnas de la matriz de coeficientes. Por tanto, el sistema de más arriba nos está preguntando: ¿qué combinaciones lineales de las columnas de $\mathbf{A}$ son vectores nulos? Evidentemente, si reemplazamos el vector de incógnitas $\boldsymbol{x}$ por $\mathbf{0}$, tenemos que

$$\mathbf{A}\mathbf{0} = \mathbf{0}.$$

Pero, aparte de la solución trivial $\boldsymbol{x} = \mathbf{0}$, ¿existen otros vectores $\boldsymbol{x} \neq \mathbf{0}$ para los que la combinación lineal $\mathbf{A}\boldsymbol{x}$ es el vector nulo? Veamos un par de ejemplos. Imagine que la matriz de coeficientes es la matriz identidad de orden 3, entonces tenemos el sistema $\mathbf{I}\boldsymbol{x} = \mathbf{0}$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Puesto que el lado izquierdo del sistema es

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

y tiene que ser igual al vector del lado derecho (que es $\mathbf{0}$); no hay más solución que aquella en la que todas las componentes de $\boldsymbol{x}$ son nulas (es decir, la *solución trivial* $\boldsymbol{x} = 0$).

Sin embargo, es posible encontrar soluciones no triviales para el siguiente sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$:

$$\begin{bmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix};$$

por ejemplo, $\boldsymbol{x} = (2, 2, 2, )$. Piense un poco y seguro que es capaz de encontrar muchísimas más.

## 7.3. Espacio nulo de una matriz $\mathcal{N}(\mathbf{A})$

Los ejemplos anteriores muestran que el conjunto de soluciones de algunos sistemas de ecuaciones lineales homogéneos *únicamente* contiene el vector nulo $\mathbf{0}$ (la *solución trivial*), pero que, para otros sistemas, el conjunto de soluciones *también contiene* vectores distintos de $\mathbf{0}$.

Vamos a dar un nombre corto al "*conjunto de soluciones de un sistema homogéneo* $\mathbf{A}\boldsymbol{x} = \mathbf{0}$":

**Definición 7.1.** *Sea* $\mathbf{A}$ *de orden* $m \times n$. *Denominamos* espacio nulo *de* $\mathbf{A}$ (que denotamos con $\mathcal{N}(\mathbf{A})$) *al subconjunto de vectores de* $\mathbb{R}^n$ *que son solución del sistema* $\mathbf{A}\boldsymbol{x} = \mathbf{0}$, *es decir* [1]

$$\mathcal{N}(\mathbf{A}) = \{\boldsymbol{x} \in \mathbb{R}^n | \, \mathbf{A}\boldsymbol{x} = \mathbf{0}\}.$$

En el primer ejemplo de la página anterior, vimos que el *espacio nulo de* $\mathbf{I}$ solo contiene el vector $\mathbf{0}$, es decir, $\mathcal{N}(\mathbf{I}) = \{\mathbf{0}\}$. En el segundo ejemplo, $\mathcal{N}(\mathbf{A})$ contiene infinitos vectores (además de $\mathbf{0}$).

Ahora bien, en todos los casos, el *espacio nulo* de una matriz de $n$ columnas es un subespacio de $\mathbb{R}^n$.

Ejercicio 42. Demuestre la siguiente

**Proposición 7.3.1.** *Para cualquier* $\mathbf{A}$ *de orden* $m \times n$, *el espacio nulo* $\mathcal{N}(\mathbf{A})$ *es un subespacio de* $\mathbb{R}^n$.

*Pista.* Basta comprobar que el conjunto es cerrado para las combinaciones lineales.

Es importante subrayar que, aunque el conjunto de soluciones de cualquier sistema homogéneo es un subespacio, *el conjunto de soluciones de un sistema NO homogéneo* $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ (con $\boldsymbol{b} \neq \mathbf{0}$) *nunca es un subespacio*, pues el vector nulo $\mathbf{0}$ no pertenece al conjunto de soluciones.

[1] Esta forma de caracterizar los elementos de un conjunto como soluciones de un sistema de ecuaciones se denomina *ecuación cartesiana*.

## 7.4. Resolución de un Sistema Homogéneo por eliminación

Dado que $\mathcal{N}\left(\mathbf{A}\right)$ es un subespacio, sabemos que si conocemos algunas soluciones del sistema $\mathbf{A}\boldsymbol{x} = \mathbf{0}$, también son solución sus combinaciones lineales. Pero, ¿cómo encontrar todas las soluciones? Es decir, ¿cómo calcular $\mathcal{N}\left(\mathbf{A}\right)$?

**Revisitando el método de eliminación.** Vimos en la Sección 5.1.4 que al aplicar una secuencia de transformaciones elementales $\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k$ a las columnas de la matriz resultante de apilar $\mathbf{A}$ con $\mathbf{I}$ (con idéntico número de columnas), obtenemos:

$$\left[\frac{\mathbf{A}}{\mathbf{I}}\right]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \left[\frac{\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}{\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}\right] = \left[\frac{\mathbf{A}(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k})}{\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}\right] = \left[\frac{\mathbf{AE}}{\mathbf{E}}\right], \qquad \text{donde} \quad \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

Ahora fijémonos en la columna $j$-ésima de la matriz resultante

$$\left[\frac{\mathbf{AE}}{\mathbf{E}}\right]_{|j} = \left(\frac{\mathbf{A}(\mathbf{E}_{|j})}{\mathbf{E}_{|j}}\right) \qquad \text{(que es la concatenación } \left(\mathbf{A}(\mathbf{E}_{|j})\right) \# \mathbf{E}_{|j}\text{);}$$

es decir, para cada columna, el subvector de la parte inferior, $\mathbf{E}_{|j}$, contiene los coeficientes ("la receta") de la combinación lineal de las columnas de $\mathbf{A}$ que utilizamos para calcular ("cocinar") la parte superior de la columna, $\mathbf{AE}_{|j}$.

*Ejemplo* 18. Apliquemos la eliminación (de izquierda a derecha) tras apilar la matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}$$

con la matriz identidad de orden 4. Fíjese en cómo la parte inferior de cada columna, de cada matriz en cada uno de los pasos de la eliminación, indica cómo calcular la parte superior de esa misma columna:

$$\left[\frac{\mathbf{A}}{\mathbf{I}}\right] = \left[\begin{array}{cccc} 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{[(-2)\mathbf{1}+\mathbf{2}]\\[(-1)\mathbf{1}+\mathbf{4}]}]{\boldsymbol{\tau}} \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 1 & -2 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[[(-1)\mathbf{2}+\mathbf{3}]]{\boldsymbol{\tau}} \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 1 & -2 & 2 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] = \left[\frac{\mathbf{A}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2\boldsymbol{\tau}_3}}{\mathbf{I}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2\boldsymbol{\tau}_3}}\right]$$

Por ejemplo, la 1ª columna de la 1ª matriz, la 2ª columna de la 2ª matriz, y la 3ª columna de la 3ª matriz verifican respectivamente que:

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \mathbf{A}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}; \qquad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \mathbf{A}\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}; \quad \text{ó} \quad \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \mathbf{A}\begin{pmatrix} 2 \\ -1 \\ 1 \\ 0 \end{pmatrix}.$$

Y no puede ser de otra manera, ya que la relación $\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)_{|j} = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)_{|j}$ se cumple en todas y cada una de las columnas, de todas y cada una de las matrices en todos y cada uno de los pasos de eliminación.

☞ En todo momento a lo largo de la secuencia de transformaciones

$$\begin{bmatrix}\mathbf{A}\\ \hline \mathbf{I}\end{bmatrix} \to \begin{bmatrix}\mathbf{A}\\ \hline \mathbf{I}\end{bmatrix}_{\boldsymbol{\tau}_1} \to \ldots \to \begin{bmatrix}\mathbf{A}\\ \hline \mathbf{I}\end{bmatrix}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \begin{bmatrix}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\end{bmatrix}.$$

La parte inferior de cada columna indica qué combinación lineal de las columnas de $\mathbf{A}$ hemos utilizado para calcular la parte superior de la columna: $\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_h}\right)_{|j} = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_h}\right)_{|j}$; $h = 1:k$.

Por el Teorema 4.2.2 sabemos que para cualquier $\mathbf{A}$, existen $k$ transformaciones elementales tales que $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathbf{K}$ es pre-escalonada. Por tanto

$$\begin{bmatrix}\mathbf{A}\\ \hline \mathbf{I}\end{bmatrix}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \begin{bmatrix}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\end{bmatrix} = \begin{bmatrix}\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)\\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\end{bmatrix} = \begin{bmatrix}\mathbf{AE}\\ \hline \mathbf{E}\end{bmatrix} = \begin{bmatrix}\mathbf{K}\\ \hline \mathbf{E}\end{bmatrix}; \quad \text{donde} \quad \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

Solo caben dos posibilidades: que $\mathbf{K}$ tenga columnas nulas o que no tenga:

**Cuando $\mathbf{K}$ tiene columnas nulas:** si la columna $\mathbf{K}_{|j}$ es nula, sabemos que

$$\mathbf{K}_{|j} = \mathbf{A}\left(\mathbf{E}_{|j}\right) = \mathbf{0} \qquad \text{y, por lo tanto,} \qquad \mathbf{E}_{|j} \in \mathcal{N}\left(\mathbf{A}\right).$$

Es decir, si $\mathbf{K}$ tiene columnas nulas, los vectores que aparecen por debajo de las columnas nulas son *soluciones* al sistema $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ (así como todas las combinaciones lineales de dichos vectores).

Por lo tanto, buscar soluciones no nulas del sistema $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ es tan sencillo como alargar $\mathbf{A}$ colocando la matriz identidad por debajo, y aplicar transformaciones elementales a las columnas de la matriz por bloques hasta pre-escalonar el bloque superior. Si la forma pre-escalonada $\mathbf{K}$ tiene columnas nulas, los vectores que hay debajo son soluciones no nulas de $\mathbf{A}\boldsymbol{x} = \mathbf{0}$. Siguiendo la nomenclatura de G. Strang, llamaremos *soluciones especiales* a estas soluciones encontradas por eliminación.

Volviendo al Ejemplo 18 en la página anterior, sabemos que los vectores

$$\begin{pmatrix}2\\-1\\1\\0\end{pmatrix} \qquad \text{y} \qquad \begin{pmatrix}-1\\0\\0\\1\end{pmatrix}$$

son soluciones *especiales* del sistema $\mathbf{A}\boldsymbol{x} = \mathbf{0}$. También sabemos que es solución cualquier combinación lineal de ambos vectores. Por ejemplo, si sumamos ambos vectores, obtenemos una nueva solución

$$\begin{bmatrix}1&2&0&1\\0&1&1&0\\1&2&0&1\end{bmatrix}\left(\begin{pmatrix}2\\-1\\1\\0\end{pmatrix}+\begin{pmatrix}-1\\0\\0\\1\end{pmatrix}\right) = \begin{bmatrix}1&2&0&1\\0&1&1&0\\1&2&0&1\end{bmatrix}\begin{pmatrix}1\\-1\\1\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

**Cuando $\mathbf{K}$ no tiene columnas nulas:** Cuando $\mathbf{K}$ es de rango completo por columnas (i.e., no tiene columnas nulas), podemos continuar con la eliminación Gauss-Jordan hasta alcanzar una forma escalonada reducida $\mathbf{AE} = \mathbf{R}$. Como cada columna tiene un pivote y cada pivote es el único elemento no nulo de su fila, es evidente que $\mathbf{R}\boldsymbol{x} = \mathbf{0}$ si y solo si $\boldsymbol{x} = \mathbf{0}$. Así, tenemos que

$$\mathbf{A}\boldsymbol{x} = \mathbf{0} \quad \Rightarrow \quad \mathbf{AE}\mathbf{E}^{-1}\boldsymbol{x} = \mathbf{0} \quad \Rightarrow \quad \mathbf{R}\left(\mathbf{E}^{-1}\boldsymbol{x}\right) = \mathbf{0} \quad \Rightarrow \quad \mathbf{E}^{-1}\boldsymbol{x} = \mathbf{0} \quad \Rightarrow \quad \mathbf{E}\mathbf{E}^{-1}\boldsymbol{x} = \mathbf{E}\mathbf{0} \quad \Rightarrow \quad \boldsymbol{x} = \mathbf{0}.$$

Así que la única solución es $\boldsymbol{x} = \mathbf{0}$ (la solución trivial).

Y ahora puede surgir una duda: ¿existen otras soluciones que no sean combinaciones lineales de las *soluciones especiales* que hemos encontrado mediante la eliminación?

La Proposición 7.4.3 demuestra que no; es decir, que *todas las soluciones son combinación lineal de las soluciones especiales*.

Pero antes necesitamos demostrar un resultado previo:

Ejercicio 43. Demuestre el siguiente

**Lema 7.4.1.** *Si* $\mathbf{K}$ *es preescalonada, entonces* $\mathbf{K}\boldsymbol{x} = \mathbf{0}$ *si y solo si son nulos los coeficientes* $x_j$ *correspondientes a las columnas no nulas de* $\mathbf{K}$.

Veamos ahora el resultado fundamental de esta lección...

**Proposición 7.4.2.** *Si* $\mathbf{A}$ *tiene soluciones especiales, todo vector de* $\mathcal{N}(\mathbf{A})$ *es combinación lineal de ellas.*

*Demostración.* Sea $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ tal que $\mathbf{AE} = \mathbf{K}$ es preescalonada (con $n$ columnas). Puesto que $\mathbf{E}$ es invertible, para cada $\boldsymbol{x} \in \mathcal{N}(\mathbf{A})$ existe un *único* vector $\boldsymbol{y}$ tal que

$$\mathbf{E}\boldsymbol{y} = \boldsymbol{x};$$

concretamente el vector $\boldsymbol{y} = (\mathbf{E}^{-1})\boldsymbol{x}$. Por tanto, como $\mathbf{A}\boldsymbol{x} = \mathbf{0}$, entonces $\mathbf{AE}\boldsymbol{y} = \mathbf{K}\boldsymbol{y} = \mathbf{0}$.

Sea $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_r,)$ la lista de índices de las $r$ columnas no nulas de $\mathbf{K}$ (de sus $r$ columnas con pivote). Por el Lema 7.4.1 sabemos que la condición necesaria y suficiente para que $\mathbf{K}\boldsymbol{y} = \mathbf{0}$ es que las componentes del vector $\boldsymbol{y}$ que multiplican a las columnas con pivote sean nulas, es decir[2], ${}_{\boldsymbol{\rho}|}\boldsymbol{y} = \mathbf{0}$.

Si $\boldsymbol{\nu} \in \mathbb{R}^{n-r}$ es la lista de índices de las columnas nulas de $\mathbf{K}$, entonces cada índice $j = 1 : n$ está en una y solo una de las dos listas: $\boldsymbol{\rho}$ ó $\boldsymbol{\nu}$. Así, usando la Ecuación 3.8 en la página 45 tenemos que para cualquier $\boldsymbol{x}$ de $\mathcal{N}(\mathbf{A})$

$$\boldsymbol{x} = \mathbf{E}\boldsymbol{y} = (\mathbf{E}_{|\boldsymbol{\rho}})({}_{\boldsymbol{\rho}|}\boldsymbol{y}) + (\mathbf{E}_{|\boldsymbol{\nu}})({}_{\boldsymbol{\nu}|}\boldsymbol{y}) = (\mathbf{E}_{|\boldsymbol{\nu}})({}_{\boldsymbol{\nu}|}\boldsymbol{y})$$

por ser nulo el vector ${}_{\boldsymbol{\rho}|}\boldsymbol{y}$. Como $\mathbf{E}_{|\boldsymbol{\nu}}$ es una matriz (la submatriz formada por las soluciones especiales) y como ${}_{\boldsymbol{\nu}|}\boldsymbol{y}$ es un vector, todo vector $\boldsymbol{x}$ de $\mathcal{N}(\mathbf{A})$ es combinación lineal de las soluciones especiales. □

Por tanto, tenemos el siguiente

**Corolario 7.4.3.** *O todas las soluciones de* $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ *son combinaciones lineales de las soluciones especiales, o bien, la forma preescalonada de* $\mathbf{A}$ *no tiene columnas nulas y la única solución es* $\mathbf{0}$*, es decir:*

$$\acute{O} \quad \mathcal{N}(\mathbf{A}) = \{\textit{combinaciones lineales de las soluciones especiales}\} \quad \acute{o} \quad \mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}.$$

Ahora ya podemos terminar el Ejemplo 18 en la página 104 concluyendo que el espacio nulo de $\mathbf{A}$ (es decir, el conjunto de soluciones del sistema $\mathbf{A}\boldsymbol{x} = \mathbf{0}$) es:

$$\mathcal{N}(\mathbf{A}) = \left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^2,\ \boldsymbol{v} = \begin{bmatrix} 2 & -1 \\ -1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{p} \right\}.$$

[2] Repase la Sección 3.C.3 en la página 45

Es decir, $\mathcal{N}\left(\mathbf{A}\right)$ *es el subconjunto de vectores* $\boldsymbol{v}$ *de* $\mathbb{R}^4$ *que verifican que, para cada* $\boldsymbol{v}$*, existen dos números reales* $\boldsymbol{p}_{|1}$ *y* $\boldsymbol{p}_{|2}$ *tales que* $\boldsymbol{p}_{|1}$ *veces la primera solución especial más* $\boldsymbol{p}_{|2}$ *veces la segunda es igual a* $\boldsymbol{v}$.[3]

**Nomenclatura usada por G. Strang** Siguiendo la nomenclatura de G. Strang, a las columnas de $\mathbf{A}$ que acaban siendo nulas en la forma preescalonada $\mathbf{K}$ se les denominará *columnas libres* y el resto (i.e. las que mantienen un pivote) se llamarán *columnas pivote*.[4]

Las variables (o incógnitas) correspondientes a columnas pivote se denominarán *variables pivote* (o variables endógenas), el resto se denominarán *variables libres* (o variables exógenas); y las columnas que quedan debajo de las columnas nulas de $\mathbf{K}$ se denominarán *soluciones especiales* (realmente no son especiales, pero así es como las llama G. Strang).

☞

**Algoritmo para resolver $\mathbf{A}\boldsymbol{x}=\mathbf{0}$**

1. Aplicamos la eliminación: $\left[\begin{array}{c}\mathbf{A}\\\hline\mathbf{I}\end{array}\right] \xrightarrow{\tau_1,\ldots,\tau_k} \left[\begin{array}{c}\mathbf{K}\\\hline\mathbf{E}\end{array}\right]$, donde $\mathbf{K}=\mathbf{A}_{\tau_1\cdots\tau_k}$ y $\mathbf{E}=\mathbf{I}_{\tau_1\cdots\tau_k}$.
2. Si hay *soluciones especiales* (columnas de $\mathbf{E}$ bajo las columnas nulas de $\mathbf{K}$)
   - Solución completa: $\mathcal{N}\left(\mathbf{A}\right)=\{\text{combinaciones lineales de las } \textit{soluciones especiales}\}$
3. Si no hay *soluciones especiales* (si $\mathbf{K}$ no tiene columnas nulas)
   - Solución completa: $\mathcal{N}\left(\mathbf{A}\right)=\{\mathbf{0}\}$

Módulo NAcAL para Python

```
A = Matrix( [ [1,2,0,1],[0,1,1,0],[1,2,0,1] ] )
H = Homogenea(A, 1)                                  # Resuelve la ecuación homogenea
```

$$\left[\begin{array}{cccc}1&2&0&1\\0&1&1&0\\1&2&0&1\\\hline1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right] \xrightarrow[[(-1)\mathbf{1}+\mathbf{4}]]{\boldsymbol{\tau}[(-2)\mathbf{1}+\mathbf{2}]} \left[\begin{array}{cccc}1&0&0&0\\0&1&1&0\\1&0&0&0\\\hline1&-2&0&-1\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right] \xrightarrow{\boldsymbol{\tau}[(-1)\mathbf{2}+\mathbf{3}]} \left[\begin{array}{cccc}1&0&0&0\\0&1&0&0\\1&0&0&0\\\hline1&-2&2&-1\\0&1&-1&0\\0&0&1&0\\0&0&0&1\end{array}\right]$$

[3] Esta forma de caracterizar los elementos $\boldsymbol{v}$ de un subconjunto como aquellos para los que es posible encontrar parámetros $\boldsymbol{p}$ que permitan expresar cada $\boldsymbol{v}$ como una combinación lineal se denomina *ecuación paramétrica*.

Fíjese que con las *ecuaciones paramétricas* es muy fácil obtener elementos del conjunto; basta elegir dos valores $p_1$ y $p_2$ cualesquiera para obtener una solución. Por el contrario, las *ecuaciones cartesianas* (Definición 7.1) nos sirven para verificar si un vector $\boldsymbol{x}$ pertenece o no al conjunto; basta verificar que $\mathbf{A}\boldsymbol{x}$ es cero. ¡Cada tipo de ecuaciones tiene una utilidad distinta, unas sirven para generar ejemplos y las otras para verificar la pertenencia al conjunto!

[4] Nótese que la clasificación indicada por esta nomenclatura depende de la secuencia de transformaciones $\tau_1\cdots\tau_k$ empleada, pues, en general, una columna acabará siendo nula o no dependiendo de dicha secuencia.

# Resolviendo Ax = b

En esta lección veremos que el método de eliminación también nos permite deducir si es resoluble $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ y, cuando lo es, encontrar el conjunto de vectores $\boldsymbol{x}$ tales que $\mathbf{A}\boldsymbol{x}$ es igual a $\boldsymbol{b}$.

## 8.1. Eliminación sobre la matriz ampliada

Considere el *sistema NO homogéneo* $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ (es decir, con $\boldsymbol{b} \neq \mathbf{0}$). Podemos reescribir el sistema para que tenga una apariencia de *sistema homogéneo*:

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b} \quad \Longleftrightarrow \quad \mathbf{A}\boldsymbol{x} - \boldsymbol{b} = \mathbf{0} \quad \Longleftrightarrow \quad \left[\begin{array}{c|c} \mathbf{A} & -\boldsymbol{b} \end{array}\right] \begin{pmatrix} \boldsymbol{x} \\ \hline 1 \end{pmatrix} = \mathbf{0}. \tag{8.1}$$

El nuevo vector de incógnitas, $(\boldsymbol{x}|1,) = (x_1, \ldots, x_n, 1,)$. es la concatenación del vector $\boldsymbol{x}$ con el vector $(1,)$. Por tanto, el nuevo vector de incógnitas tiene como restricción que *su última componente debe ser un 1.* La matriz $\left[\mathbf{A}|-\boldsymbol{b}\right]$, que es la matriz que obtenemos al *añadir* el vector opuesto de $\boldsymbol{b}$ a la lista de columnas de $\mathbf{A}$, se denomina *matriz de coeficientes "ampliada"*. Dibujamos una línea vertical para recordar que la parte izquierda de la matriz corresponde a la *matriz de coeficientes* del sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ y que la columna de la derecha es el **opuesto** del *vector del lado derecho...* **¡fíjese en el cambio de signo!**

El método de resolución que vamos a ver intenta resolver el "pseudo-*sistema homogéneo*" (8.1) buscando soluciones que tengan un 1 en su última componente (es decir, trataremos de *encontrar algún vector de* $\mathcal{N}\left(\left[\mathbf{A}|-\boldsymbol{b}\right]\right)$ *que tenga un 1 en la última componente*).

☞ Las implicaciones "si y solo si" de más arriba indican que encontrar soluciones de $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ es equivalente a encontrar vectores del espacio nulo de la matriz ampliada *que tengan un 1 como última componente*.

La mecánica consiste en aplicar el método de eliminación como en la lección anterior, pero usando ahora la *matriz de coeficientes "ampliada"*. Si logramos anular la última columna de la matriz de coeficientes *"ampliada"*, entonces el conjunto de soluciones del sistema no es vacío. ¡Si además logramos mantener el 1 de la última componente de la solución especial correspondiente a la última columna, entonces tendremos a la vista una solución de $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$!

Para resolver por eliminación el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ (de $n$ incógnitas), "apilamos" la matriz ampliada $\left[\mathbf{A}|-\boldsymbol{b}\right]$, de $(n+1)$ columnas, con una matriz identidad de orden $(n+1)$ por debajo:

$$\left[\begin{array}{c|c} \underset{m\times n}{\mathbf{A}} & -\boldsymbol{b} \\ \hline \underset{n\times n}{\mathbf{I}} & \\ \hline & 1 \end{array}\right] \qquad \text{donde} \qquad \underset{(n+1)\times(n+1)}{\mathbf{I}} = \left[\begin{array}{c|c} \underset{n\times n}{\mathbf{I}} & \\ \hline & 1 \end{array}\right];$$

y distinguimos los bloques correspondientes a las distintas partes con líneas horizontales y verticales. Después pre-escalonamos la matriz como hicimos en la lección anterior con los sistemas homogéneos.

Probemos con el siguiente ejemplo de sistema de ecuaciones "*sin solución*":

*Ejemplo* 19. Considere el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{d}$ con $\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix}$ y $\boldsymbol{d} = \begin{pmatrix} 5 \\ -1 \\ 1 \\ -1 \end{pmatrix}$.

$$\left[\begin{array}{c|c} \mathbf{A} & -\boldsymbol{d} \\ \hline \mathbf{I} & \\ \hline & 1 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 1 & 2 & -5 \\ 2 & 2 & 4 & 1 \\ 3 & 1 & 4 & -1 \\ 4 & 1 & 5 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{[(-1)\mathbf{1}+\mathbf{2}] \\ [(-2)\mathbf{1}+\mathbf{3}] \\ [(5)\mathbf{1}+\mathbf{4}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 11 \\ 3 & -2 & -2 & 14 \\ 4 & -3 & -3 & 21 \\ \hline 1 & -1 & -2 & 5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{[(-1)\mathbf{2}+\mathbf{3}] \\ [(7)\mathbf{2}+\mathbf{4}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 11 \\ 3 & -2 & 0 & 0 \\ 4 & -3 & 0 & 0 \\ \hline 1 & -1 & -1 & -2 \\ 0 & 1 & -1 & 7 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

Como no encontramos un vector de $\mathcal{N}\left([\mathbf{A}|-\boldsymbol{d}]\right)$ cuya última componente sea un 1, el conjunto de soluciones de $\mathbf{A}\boldsymbol{x} = \boldsymbol{d}$ es vacío (entonces se dice que el sistema es incompatible, o que *NO tiene solución*).

*Ejemplo* 20. Ahora considere el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ con $\mathbf{A} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix}$ y $\boldsymbol{b} = \begin{pmatrix} 5 \\ 10 \\ 1 \\ -1 \end{pmatrix}$. Aplicando el método de eliminación sobre la matriz ampliada tenemos

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & -5 \\ 2 & 2 & 4 & -10 \\ 3 & 1 & 4 & -1 \\ 4 & 1 & 5 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{[(-1)\mathbf{1}+\mathbf{2}] \\ [(-2)\mathbf{1}+\mathbf{3}] \\ [(5)\mathbf{1}+\mathbf{4}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 3 & -2 & -2 & 14 \\ 4 & -3 & -3 & 21 \\ \hline 1 & -1 & -2 & 5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{[(-1)\mathbf{2}+\mathbf{3}] \\ [(7)\mathbf{2}+\mathbf{4}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ 4 & -3 & 0 & 0 \\ \hline 1 & -1 & -1 & -2 \\ 0 & 1 & -1 & 7 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

Hemos encontrado dos soluciones especiales (los vectores por debajo de las dos columnas nulas), de las cuales la segunda tiene un 1 como última componente. Así pues, por las implicaciones "si y solo si" de la Ecuación 8.1 sabemos que el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ es resoluble y que $\boldsymbol{x} = \begin{pmatrix} -2, & 7, & 0 \end{pmatrix}$ es *una* solución, pues

$$[\mathbf{A}|-\boldsymbol{b}] \begin{pmatrix} \boldsymbol{x} \\ \hline 1 \end{pmatrix} = \mathbf{0} \iff \mathbf{A}\boldsymbol{x} - \boldsymbol{b} = \mathbf{0} \iff \mathbf{A}\boldsymbol{x} = \boldsymbol{b}.$$

**¿Una solución o infinitas soluciones?**

Fíjese en la última matriz obtenida en el proceso de eliminación del último ejemplo. Sumando cualquier múltiplo de la tercera columna a la cuarta, obtenemos otro vector de $\mathcal{N}\left([\mathbf{A}|-\boldsymbol{b}]\right)$ con un 1 como última componente. Es decir, el conjunto de vectores de $\mathcal{N}\left([\mathbf{A}|-\boldsymbol{b}]\right)$ con un 1 en la última componente es

$$\left\{ \boldsymbol{y} \in \mathbb{R}^4 \;\middle|\; \text{existe } p \in \mathbb{R} \text{ tal que } \boldsymbol{y} = \begin{pmatrix} -2 \\ 7 \\ 0 \\ \hline 1 \end{pmatrix} + p \begin{pmatrix} -1 \\ -1 \\ 1 \\ \hline 0 \end{pmatrix} \right\}$$

Consecuentemente (Ecuación 8.1), el conjunto de soluciones del sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ es

$$\left\{ \boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \text{existe } p \in \mathbb{R} \text{ tal que } \boldsymbol{v} = \begin{pmatrix} -2 \\ 7 \\ 0 \end{pmatrix} + p \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\}$$

Las ecuaciones paramétricas permiten obtener ejemplos de soluciones; basta asignar valores arbitrarios a los parámetros (en este caso a $p$). Obtenga alguna otra solución $\boldsymbol{v}$ y compruebe que $\mathbf{A}\boldsymbol{v} = \boldsymbol{b}$.

☞ Una vez visto el funcionamiento en un ejemplo, veamos el caso general del sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ con solución.

Considere un sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ que tiene solución y donde $r$ es el rango de la matriz de coeficientes $\mathbf{A}$ de $n$ columnas. Sean $\big[\mathbf{A}|-\boldsymbol{b}\big]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ una forma pre-escalonada y $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$.

Sabemos que el espacio nulo $\mathcal{N}\left(\big[\mathbf{A}|-\boldsymbol{b}\big]\right)$ contiene todas las combinaciones lineales de las soluciones especiales. Como en la demostración de la Proposición 7.4.2 en la página 106, sea $\boldsymbol{\nu} = (\nu_1, ..., \nu_{n-r},\ n+1,)$ la lista de índices correspondientes a las columnas nulas de la forma pre-escalonada (sabemos que el último índice es $n+1$ puesto que $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ tiene solución).

Por conveniencia, expresaré $\boldsymbol{\nu}$ como concatenación de dos listas: $\boldsymbol{\nu} = \boldsymbol{\alpha}\#(n+1,)$; donde $\boldsymbol{\alpha} = (\nu_1, ..., \nu_{n-r},)$ es la lista con los $n-r$ primeros índices, y la segunda lista, $(n+1,)$, tan solo contiene el último índice. Entonces:

$$\mathcal{N}\left(\big[\mathbf{A}|-\boldsymbol{b}\big]\right) = \left\{\boldsymbol{y} \in \mathbb{R}^{(n+1)} \,\middle|\, \text{existen } \boldsymbol{x} \in \mathbb{R}^{n-r} \text{ y } z \in \mathbb{R} \text{ tales que } \boldsymbol{y} = \big(\mathbf{E}_{|\boldsymbol{\alpha}}\big)\boldsymbol{x} + (\mathbf{E}_{|n+1})z\right\};$$

y como todos los vectores de $\mathcal{N}\left(\big[\mathbf{A}|-\boldsymbol{b}\big]\right)$ tienen un 1 en su última componente y la única solución especial con un 1 en la última componente[1] es $\mathbf{E}_{|n+1}$, necesariamente $z = 1$. En consecuencia,

$$\mathcal{N}\left(\big[\mathbf{A}|-\boldsymbol{b}\big]\right) = \left\{\boldsymbol{v} \in \mathbb{R}^{(n+1)} \,\middle|\, \text{existe } \boldsymbol{x} \in \mathbb{R}^{n-r} \text{ tal que } \boldsymbol{v} = \big(\mathbf{E}_{|\boldsymbol{\alpha}}\big)\boldsymbol{x} + \mathbf{E}_{|n+1}\right\}. \tag{8.2}$$

De esto se deduce que *cuando* $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ *es resoluble:*

1. El conjunto de soluciones del sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ es el resultado de quitar la última componente de las soluciones especiales de la Ecuación (8.2).
2. Si la forma pre-escalonada de $\mathbf{A}$ tiene columnas nulas (si $\operatorname{rg}\big(\mathbf{A}\big) < n$), entonces el sistema tiene infinitas soluciones. Y si, por el contrario, $\mathcal{N}\big(\mathbf{A}\big) = \{\mathbf{0}\}$, entonces existe un único vector solución... (el vector cuyas $n$ componentes son las $n$ primeras componentes de la solución especial $\mathbf{E}_{|n+1}$).

## Algoritmo de resolución de un sistema no homogéneo

Aplicando el método de eliminación de izquierda a derecha para pre-escalonar, logramos:

$$\left[\begin{array}{c|c} \mathbf{A}_{m\times n} & -\boldsymbol{b} \\ \hline \mathbf{I}_{n\times n} & \\ \hline & 1 \end{array}\right]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_p} = \left[\begin{array}{c|c} \mathbf{K} & \boldsymbol{c} \\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \boldsymbol{w} \\ \hline & d \end{array}\right], \tag{8.3}$$

donde $\big[\mathbf{K}|\boldsymbol{c}\big]$ es una matriz pre-escalonada de la matriz de coeficientes ampliada $\big[\mathbf{A}|-\boldsymbol{b}\big]$; y donde la única componente distinta de cero de la última fila es $d$ (en la última columna).

Si se siguen los pasos de la demostración del Teorema 4.2.1 en la página 59, entonces $d$ siempre será igual a 1, pues no se habrán empleado transformaciones de *Tipo II*. Pero si usted opera como lo hace el módulo NAcAL (que multiplica algunas columnas por números no nulos para evitar operar con fracciones en la medida de lo posible), entonces $d$ podrá tomar cualquier valor no nulo.

[1] recuerde que la última componente del resto de vectores $\mathbf{E}_{|j}$ con $j \neq (n+1)$ es un 0

Tras la eliminación en (8.3), solo caben dos posibilidades:

1. Si $\boldsymbol{c} = \mathbf{0}$, entonces el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ tiene solución. Si $d \neq 1$, basta dividir la última columna por $d$ para obtener una solución especial cuya última componente es un 1.

$$\left[\begin{array}{c|c} \underset{m\times n}{\mathbf{K}} & \mathbf{0} \\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \boldsymbol{w} \\ \hline & d \end{array}\right] \xrightarrow{\boldsymbol{\tau}\left[\left(\frac{1}{d}\right)(n+1)\right]} \left[\begin{array}{c|c} \underset{m\times n}{\mathbf{K}} & \mathbf{0} \\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \boldsymbol{s} \\ \hline & 1 \end{array}\right],$$

   donde $\boldsymbol{s}$ es tal que $\mathbf{A}\boldsymbol{s} = \boldsymbol{b}$. Al vector $\boldsymbol{s}$ se le suele llamar *"solución particular"* (aunque no tiene nada de particular, tan solo que lo hemos encontrado aplicando la eliminación).

2. Si $\boldsymbol{c} \neq \mathbf{0}$, entonces el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ NO tiene solución, pues $\boldsymbol{b} \notin \mathcal{C}\left(\mathbf{A}\right)$.

Por tanto, *el sistema tiene solución si y solo si logramos transformar $-\boldsymbol{b}$ en el vector $\mathbf{0}$*.

☞ Mediante la eliminación, podemos saber si $\underset{m\times n}{\mathbf{A}}\ \boldsymbol{x} = \boldsymbol{b}$ es resoluble (y encontrar *una* solución cuando lo es).

**Algoritmo para resolver $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$**

Pre-escalonamos (*y luego dividimos la última columna por su última componente, $d$*)[a]:

$$\left[\begin{array}{c|c} \mathbf{A} & -\boldsymbol{b} \\ \hline \mathbf{I} & \\ \hline & 1 \end{array}\right] \xrightarrow{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_p} \left[\begin{array}{c|c} \mathbf{K} & \boldsymbol{f} \\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \boldsymbol{w} \\ \hline & d \end{array}\right] \xrightarrow{\boldsymbol{\tau}\left[\left(\frac{1}{d}\right)(n+1)\right]} \left[\begin{array}{c|c} \mathbf{K} & \boldsymbol{c} \\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \boldsymbol{s} \\ \hline & 1 \end{array}\right]$$

donde $\mathbf{K} = \mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ está pre-escalonada; y por tanto: $\mathbf{K}_{|j} = \mathbf{0} \implies \left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)_{|j} \in \mathcal{N}\left(\mathbf{A}\right)$.

Solo caben dos posibilidades:

- Si $\boldsymbol{c} \neq \mathbf{0}$, el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ NO tiene solución.
- Si $\boldsymbol{c} = \mathbf{0}$, entonces $\mathbf{A}\boldsymbol{s} = \boldsymbol{b}$, y el conjunto de todas las soluciones es

$$\left\{\boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \text{existe } \boldsymbol{v} \in \mathcal{N}\left(\mathbf{A}\right) \text{ tal que } \boldsymbol{v} = \boldsymbol{s} + \boldsymbol{v}\right\}.$$

Por tanto, cuando $\mathbf{K}$ no tiene columnas nulas ($\mathcal{N}\left(\mathbf{A}\right) = \{\mathbf{0}\}$), la única solución es $\boldsymbol{s}$.

[a] Nótese que $d$ siempre es igual a 1 si se aplica el algoritmo de la demostración del Teorema 4.2.1 en la página 59, pero si también se emplean transformaciones *Tipo II*, entonces $d$ podría tomar cualquier valor no nulo.

Decimos que un sistema de ecuaciones es *incompatible* si el conjunto de soluciones es *vacío*; y *compatible* en caso contrario. A su vez, los sistemas compatibles se dividen en *determinados* cuando el vector solución es único, e *indeterminado* cuando el conjunto de soluciones contiene más de un vector.

*"Resolver un sistema de ecuaciones"* consiste en indicar que el conjunto de soluciones es vacío cuando el sistema es *incompatible*; indicar la única solución cuando el sistema es *compatible determinado*; y *expresar el conjunto de soluciones mediante ecuaciones paramétricas* cuando el sistema es *compatible indeterminado*.

Módulo NAcAL para Python

```
A = Matrix( [ [7,-2,0], [6,4,2] ] ); b = Vector([4,12])
SEL( A, b, 1 )        # resuelve el Sistema de Ecuaciones Lineales
```

$$\left[\begin{array}{ccc|c} 7 & -2 & 0 & -4 \\ 6 & 4 & 2 & -12 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{[(7)\mathbf{2}]\\ [(2)\mathbf{1}+\mathbf{2}]\\ [(7)\mathbf{4}]\\ [(4)\mathbf{1}+\mathbf{4}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccc|c} 7 & 0 & 0 & 0 \\ 6 & 40 & 2 & -60 \\ \hline 1 & 2 & 0 & 4 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 7 \end{array}\right] \xrightarrow[\substack{[(20)\mathbf{3}]\\ [(-1)\mathbf{2}+\mathbf{3}]\\ [(2)\mathbf{4}]\\ [(3)\mathbf{2}+\mathbf{4}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccc|c} 7 & 0 & 0 & 0 \\ 6 & 40 & 0 & 0 \\ \hline 1 & 2 & -2 & 14 \\ 0 & 7 & -7 & 21 \\ 0 & 0 & 20 & 0 \\ \hline 0 & 0 & 0 & 14 \end{array}\right] \xrightarrow[[(\frac{1}{14})\mathbf{4}]]{\boldsymbol{\tau}} \left[\begin{array}{ccc|c} 7 & 0 & 0 & 0 \\ 6 & 40 & 0 & 0 \\ \hline 1 & 2 & -2 & 1 \\ 0 & 7 & -7 & \frac{3}{2} \\ 0 & 0 & 20 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

$$\left\{ \boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^1,\ \boldsymbol{v} = \begin{pmatrix} 1 \\ \frac{3}{2} \\ 0 \end{pmatrix} + \begin{bmatrix} -2 \\ -7 \\ 20 \end{bmatrix} \boldsymbol{p} \right\}.$$

(Como multiplicamos la cuarta columna por 7 y por 2, al final debemos dividir la cuarta columna por 14.)

### 8.1.1. El Teorema de Rouché-Frobenius

☞ Del procedimiento anterior se deduce que los rangos de la matriz de coeficientes $\mathbf{A}$ y de la matriz ampliada $\big[\mathbf{A}|-\boldsymbol{b}\big]$ clasifican los posibles casos en cuanto al número de soluciones del sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$.

Sabemos que tras aplicar el método de eliminación sobre las columnas de la matriz de coeficientes ampliada

$$\big[\mathbf{A}|-\boldsymbol{b}\big] \xrightarrow{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} \big[\mathbf{K}\,|\,\boldsymbol{c}\big]$$

solo se pueden presentar dos casos: que la última columna sea nula ($\boldsymbol{c} = \mathbf{0}$), o que no lo sea ($\boldsymbol{c} \neq \mathbf{0}$):

1. Si $\boldsymbol{c} \neq \mathbf{0}$, entonces la matriz ampliada tiene una columna pivote adicional, es decir,

$$\operatorname{rg}\big(\big[\mathbf{A}|-\boldsymbol{b}\big]\big) > \operatorname{rg}\big(\mathbf{A}\big) \qquad \Longleftrightarrow \qquad \text{el sistema NO tiene solución.}$$

2. Si $\boldsymbol{c} = \mathbf{0}$, entonces

$$\operatorname{rg}\big(\big[\mathbf{A}|-\boldsymbol{b}\big]\big) = \operatorname{rg}\big(\mathbf{A}\big) \qquad \Longleftrightarrow \qquad \text{el sistema es resoluble.}$$

   Si además $\mathbf{A}$ es de rango completo por columnas, entonces la forma pre-escalonada $\mathbf{K}$ no tiene columnas nulas (es decir, $\mathcal{N}\big(\mathbf{A}\big) = \{\mathbf{0}\}$), por lo que el sistema tiene una única solución.

   Sin embargo, si el rango es menor que $n$, entonces el sistema tiene infinitas soluciones, ya que a una solución $\boldsymbol{x}$ se le puede sumar cualquier vector de $\mathcal{N}\big(\mathbf{A}\big)$ para obtener otra solución.

El resultado anterior se resume en el siguiente

**Teorema 8.1.1** (Rouché-Frobenius)**.** *Un sistema de ecuaciones lineales $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ con $n$ incógnitas tiene solución si y solo si el rango de la matriz de coeficientes $\mathbf{A}$ es igual al rango de la matriz de coeficientes ampliada $[\mathbf{A}|-\boldsymbol{b}]$. En particular:*

1. *Si $\operatorname{rg}\big(\mathbf{A}\big) = n$, la solución es única.*
2. *En caso contrario, hay infinitas soluciones.*

Si además $\mathbf{A}$ es de rango completo por filas (si su rango es igual a su número de filas $m$), el sistema tiene solución para cualquier vector $\boldsymbol{b} \in \mathbb{R}^m$; pues al pre-escalonar la matriz ampliada se anulan todas las componentes de $-\boldsymbol{b}$, por estar a la derecha de $\mathbf{A}$ y haber un pivote en cada fila de $\mathbf{A}$).

Así, para el caso de matrices cuadradas, tenemos el siguiente

**Corolario 8.1.2.** *Para toda matriz cuadrada* $\mathbf{A}$ *de orden* $n$*, las siguientes afirmaciones son equivalentes:*

1. *El rango de* $\mathbf{A}$ *es* $n$ *(o* $\mathbf{A}$ *es de rango completo).*
2. $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ *si y solo si* $\boldsymbol{x} = \mathbf{0}$.
3. $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ *tiene solución única.*
4. $\mathbf{A}$ *no es singular.*
5. $\mathbf{A}$ *es invertible.*
6. $\mathbf{A}$ *es producto de matrices elementales:* $\mathbf{A} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$.
7. $\boldsymbol{x} = \mathbf{A}^{-1}\boldsymbol{b}$ *es la única solución a* $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$.

## 8.2. Espacio columna de una matriz

Considere el conjunto de todas las combinaciones lineales de las columnas de $\mathbf{A}$ de orden $m$ por $n$

$$\left\{ \boldsymbol{b} \in \mathbb{R}^m \;\middle|\; \text{existen } x_1, \ldots, x_n \in \mathbb{R} \text{ tales que } \boldsymbol{b} = x_1 \mathbf{A}_{|1} + \cdots + x_n \mathbf{A}_{|n} \right\}$$

o bien, usando la notación matricial,

$$\{ \boldsymbol{b} \in \mathbb{R}^n \mid \text{existe } \boldsymbol{x} \in \mathbb{R}^n \text{ tal que } \boldsymbol{b} = \mathbf{A}\boldsymbol{x} \}.$$

Dado que este subconjunto de $\mathbb{R}^n$ *contiene todas* las combinaciones lineales de las columnas, *es cerrado* para las combinaciones lineales. Por tanto, este conjunto es un *subespacio* de $\mathbb{R}^n$. Lo denominamos *espacio columna* de $\mathbf{A}$, y lo denotamos por $\mathcal{C}(\mathbf{A})$. Por ejemplo

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}; \qquad \Rightarrow \qquad \mathcal{C}(\mathbf{A}) = \left\{ \boldsymbol{b} \in \mathbb{R}^n \;\middle|\; \text{existen } x, y, z \in \mathbb{R} \text{ tales que } \boldsymbol{b} = x \begin{pmatrix} 1 \\ 1 \end{pmatrix} + y \begin{pmatrix} 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

Evidentemente

$$\boxed{\mathbf{A}\boldsymbol{x} = \boldsymbol{b} \quad \text{tiene solución si y solo si} \quad \boldsymbol{b} \in \mathcal{C}(\mathbf{A}).}$$

**El espacio columna y la eliminación**

El subespacio $\mathcal{C}(\mathbf{A})$ es cerrado para la suma y el producto por escalares, pero precisamente ese es el tipo de operaciones que realizan las transformaciones elementales de las columnas. Por tanto, al aplicar transformaciones elementales sobre las columnas de $\mathbf{A}$ obtenemos nuevas matrices cuyas columnas pertenecen a $\mathcal{C}(\mathbf{A})$. Así, el espacio columna de $\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ está contenido en el espacio columna de $\mathbf{A}$:

$$\mathcal{C}(\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}) \subset \mathcal{C}(\mathbf{A});$$

pero como las transformaciones elementales son "reversibles", mediante una sucesión de transformaciones elementales se puede retornar desde $\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ a $\mathbf{A}$, pues $\mathbf{A} = \left(\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}\right)_{\boldsymbol{\tau}_k^{-1} \cdots \boldsymbol{\tau}_1^{-1}}$. Por tanto, también ocurre que

$$\mathcal{C}(\mathbf{A}) \subset \mathcal{C}(\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}).$$

Por tanto

$$\boxed{\mathcal{C}(\mathbf{A}) = \mathcal{C}(\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k})}.$$

¡Al aplicar la eliminación, todas las matrices que aparecen en el proceso tienen el mismo espacio columna!

Así, concluimos que

$$\boxed{\text{Si} \quad \mathbf{B} = \mathbf{A}\mathbf{E} \quad \text{y} \quad \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}, \quad \text{entonces} \quad \mathcal{C}(\mathbf{A}) = \mathcal{C}(\mathbf{B}).}$$

## 8.3. Generalizaciones o visiones alternativas (★)

**Resolviendo varios sistemas a la vez** El método de resolución que hemos visto permite (con una mínima variación) resolver simultáneamente sistemas que compartan la misma matriz de coeficientes. Por ejemplo, podemos tratar de resolver simultáneamente los ejemplos 19 y 20:

$$\left[\begin{array}{c|cc} \mathbf{A} & -\boldsymbol{b} & -\boldsymbol{d} \\ \mathbf{I} & & \\ \hline & 1 & 0 \\ & 0 & 1 \end{array}\right] = \left[\begin{array}{ccc||c|c} 1&1&2&-5&-5\\ 2&2&4&-10&1\\ 3&1&4&-1&-1\\ 4&1&5&1&1\\ \hline 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&1&0&0\\ \hline 0&0&0&1&0\\ \hline 0&0&0&0&1 \end{array}\right] \to \left[\begin{array}{ccc||c|c} 1&0&0&0&0\\ 2&0&0&0&11\\ 3&-2&-2&14&14\\ 4&-3&-3&21&21\\ \hline 1&-1&-2&5&5\\ 0&1&0&0&0\\ 0&0&1&0&0\\ \hline 0&0&0&1&0\\ \hline 0&0&0&0&1 \end{array}\right] \to \left[\begin{array}{ccc||c|c} 1&0&0&0&0\\ 2&0&0&0&11\\ 3&-2&0&0&0\\ 4&-3&0&0&0\\ \hline 1&-1&-1&-2&-2\\ 0&1&-1&7&7\\ 0&0&1&0&0\\ \hline 0&0&0&1&0\\ \hline 0&0&0&0&1 \end{array}\right]$$

Es decir, $\mathbf{A}\boldsymbol{x} = \boldsymbol{d}$ NO tiene solución, pero $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ sí es resoluble, y el conjunto de todas las soluciones es

$$\left\{ \boldsymbol{x} \in \mathbb{R}^3 \;\middle|\; \text{existe } a \in \mathbb{R} \text{ tal que } \boldsymbol{x} = \begin{pmatrix} -2 \\ 7 \\ 0 \end{pmatrix} + a \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\}.$$

(Tenga en cuenta que aquí no puede usar la penúltima columna para eliminar componentes de la última).

# Independencia, base y dimensión

## 9.1. Combinaciones lineales de vectores de $\mathcal{V}$

[Recuerde las definiciones de *Espacio Vectorial* (Pag. 86) y de *Subespacio* (Pag. 90)]

En esta lección abordaremos los *sistemas* de vectores en un subespacio genérico $\mathcal{V}$:

$$\mathsf{Z} = \left[\vec{z}_1;\ \vec{z}_2; \ldots \vec{z}_n;\right], \quad \text{con} \quad \vec{z}_i \in \mathcal{V};$$

denotaremos tales sistemas utilizando caracteres del tipo: A, B,..., Y, Z. También utilizaremos el operador selector para indicar el $j$-ésimo vector de un sistema $\mathsf{Z} = \left[\vec{z}_1; \ldots \vec{z}_n;\right]$:

$$\mathsf{Z}_{|j} = \vec{z}_j; \quad \text{por lo tanto} \quad \mathsf{Z} = \left[\mathsf{Z}_{|1};\ \mathsf{Z}_{|2}; \ldots \mathsf{Z}_{|n};\right].$$

Es importante notar que las matrices son un caso particular; ya que, una matriz $\mathbf{A}$ de orden $m$ por $n$ representa un sistema de $n$ vectores en $\mathbb{R}^m$:

$$\mathbf{A} = \left[\mathbf{A}_{|1};\ \mathbf{A}_{|2}; \ldots \mathbf{A}_{|n};\right], \quad \text{con} \quad \mathbf{A}_{|j} \in \mathbb{R}^m.$$

Módulo NAcAL para Python

```
A = Matrix( [ [1,0,0], [-1,0,1] ] )
A.sis()              # Devuelve el Sistema correspondiente a las columnas de A
```

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} = \left[ \begin{pmatrix} 1 \\ -1 \end{pmatrix}; \quad \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \right].$$

### 9.1.1. Extensión de la notación matricial a los espacios vectoriales

Ampliaremos la notación matricial a subespacios genéricos[1] comenzando con el...

**Producto de un sistema de $n$ vectores de $\mathcal{V}$ por un vector de $\mathbb{R}^n$**

Definimos el producto $\mathsf{Z}\boldsymbol{a}$ de un *sistema de $n$ vectores de $\mathcal{V}$ por un vector de $\mathbb{R}^n$* como la combinación lineal de los vectores de $\mathsf{Z}$, cuyos coeficientes son las componentes de $\boldsymbol{a}$:

$$\mathsf{Z}\boldsymbol{a} = (\mathsf{Z}_{|1})a_1 + \cdots + (\mathsf{Z}_{|n})a_n.$$

Por tanto, $\mathsf{Z}\boldsymbol{a}$ es un vector de $\mathcal{V}$. En el caso particular del producto $\mathbf{A}\boldsymbol{x}$ se obtiene un vector de $\mathbb{R}^m$, mientras que en el caso general de $\mathsf{Z}\boldsymbol{a}$, dependiendo de la naturaleza de $\mathcal{V}$, el vector resultante podría ser una función, una matriz, un polinomio, una variable aleatoria, etc.... todo depende de la naturaleza de $\mathcal{V}$.

[1] subespacios genéricos pero *de dimensión finita*.

*Ejemplo* 21. Consideremos un ejemplo de un sistema de tres matrices multiplicado por un vector de $\mathbb{R}^3$:

Módulo NAcAL para Python

```
A = Matrix([ [ 1, 1, 1], [1, 0, 0] ])
B = Matrix([ [ 4, 4, 4], [0, 4, 0] ])
C = Matrix([ [ 2, 0, 1], [0, 0, 1] ])
Z = Sistema( [ A, B, C] )           # Sistema de Matrices
a = Vector ( [ 1, 0, 1] )
Z*a                                 # Sistema por Vector (comb. lineal de matrices)
```

$$\mathsf{Z}\boldsymbol{a} = \left[ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 4 & 4 & 4 \\ 0 & 4 & 0 \end{bmatrix}; \begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}; \right] \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix}.$$

El ejercicio siguiente demuestra que el producto $\mathsf{Z}\boldsymbol{a}$ cumple con las *propiedades de linealidad*: la primera respecto a la suma, $\mathsf{Z}(\boldsymbol{b}+\boldsymbol{c}) = \mathsf{Z}\boldsymbol{b}+\mathsf{Z}\boldsymbol{c}$, y la segunda respecto al producto por un escalar, $\mathsf{Z}(\lambda\boldsymbol{b}) = \lambda(\mathsf{Z}\boldsymbol{b})$.

Ejercicio 44. Demuestre las siguientes proposiciones.

(a) **Proposición 9.1.1.** *Sea* $\mathsf{Z}$ *un sistema de* $n$ *vectores de* $\mathcal{V}$*, y sean* $\boldsymbol{b}$ *y* $\boldsymbol{c}$ *vectores de* $\mathbb{R}^n$*; entonces*

$$\mathsf{Z}(\boldsymbol{b}+\boldsymbol{c}) = \mathsf{Z}\boldsymbol{b}+\mathsf{Z}\boldsymbol{c}.$$

(b) **Proposición 9.1.2.** *Sea* $\mathsf{Z}$*, un sistema de* $n$ *vectores de* $\mathcal{V}$*, el escalar* $\lambda$ *y* $\boldsymbol{b}$ *un vector de* $\mathbb{R}^n$*; entonces*

$$\mathsf{Z}(\lambda\boldsymbol{b}) = \lambda(\mathsf{Z}\boldsymbol{b}).$$

Podemos extender aún más la notación matricial.

**Sistema de $n$ vectores de $\mathcal{V}$ por una matriz de $\mathbb{R}^{n\times p}$**

Si $\mathsf{Z}$ es un sistema de $n$ vectores y $\mathbf{A}$ es una matriz de $n$ filas y $p$ columnas, denotamos por $\mathsf{Z}\mathbf{A}$ al sistema de vectores:

$$\mathsf{Z}\mathbf{A} = \left[\mathsf{Z}(\mathbf{A}_{|1}); \ldots \mathsf{Z}(\mathbf{A}_{|p});\right];$$

es decir, el vector $j$-ésimo del sistema de $\mathsf{Z}\mathbf{A}$ es

$$(\mathsf{Z}\mathbf{A})_{|j} = \mathsf{Z}(\mathbf{A}_{|j});$$

(compare esta expresión con la Definición 3.1). Por lo tanto, podemos escribir sin ambigüedad el vector $\mathsf{Z}\mathbf{A}_{|j}$.

Continuando con el ejemplo anterior:

Módulo NAcAL para Python

```
D = Matrix([ Vector([0,2,0]), Vector([1,0,1]) ])
Z*D                                 # Sistema por Matriz (Sistema de Matrices)
```

$$\mathsf{Z}\mathbf{D} = \left[ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}; \begin{bmatrix} 4 & 4 & 4 \\ 0 & 4 & 0 \end{bmatrix}; \begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}; \right] \begin{bmatrix} 0 & 1 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} = \left[ \begin{bmatrix} 8 & 8 & 8 \\ 0 & 8 & 0 \end{bmatrix}; \begin{bmatrix} 3 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix}; \right].$$

**Definición 9.1.** *Sea* $\mathsf{Z}_{|\boldsymbol{\gamma}}$ *el sistema formado por los vectores de* $\mathsf{Z}$ *indicados en* $\boldsymbol{\gamma}$*, es decir,* $\mathsf{Z}_{|\boldsymbol{\gamma}} = \mathsf{Z}(\mathbf{I}_{|\boldsymbol{\gamma}})$.

Veamos que también se cumple la asociatividad para el producto: $\mathsf{Z}(\mathbf{B}\boldsymbol{c}) = (\mathsf{Z}\mathbf{B})\boldsymbol{c}$ y $\mathsf{Z}\big(\mathbf{B}\mathbf{C}\big) = \big(\mathsf{Z}\mathbf{B}\big)\mathbf{C}$ :

EJERCICIO 45. Demuestre las siguientes proposiciones.

(a) **Proposición 9.1.3.** *Sea* $\mathsf{Z}$ *un sistema de* $m$ *vectores de* $\mathcal{V}$*, la matriz* $\underset{m\times n}{\mathbf{B}}$ *, y el vector* $\boldsymbol{c}$ *de* $\mathbb{R}^n$*; entonces:*

$$\mathsf{Z}(\mathbf{B}\boldsymbol{c}) = (\mathsf{Z}\mathbf{B})\boldsymbol{c}.$$

(b) **Proposición 9.1.4.** *Sea* $\mathsf{Z}$ *un sistema de* $p$ *vectores de* $\mathcal{V}$*, y* $\underset{p\times q}{\mathbf{B}}$ *y* $\underset{q\times n}{\mathbf{C}}$ *, entonces* $\mathsf{Z}\big(\mathbf{B}\mathbf{C}\big) = \big(\mathsf{Z}\mathbf{B}\big)\mathbf{C}$.

De esta manera, podemos escribir sin ambigüedad el vector $\mathsf{Z}\mathbf{B}\boldsymbol{c}$ y el sistema $\mathsf{Z}\mathbf{B}\mathbf{C}$.

☞ Al extender la notación matricial para expresar combinaciones lineales en espacios vectoriales genéricos, hemos encontrado resultados sobre combinaciones lineales de sistemas de vectores de $\mathcal{V}$ que serán de utilidad.

- $\mathsf{Z}\big(\boldsymbol{b}+\boldsymbol{c}\big) = \mathsf{Z}\boldsymbol{b} + \mathsf{Z}\boldsymbol{c}$
- $\mathsf{Z}\big(\lambda\boldsymbol{b}\big) = \lambda\big(\mathsf{Z}\boldsymbol{b}\big)$
- $\mathsf{Z}(\mathbf{B}\boldsymbol{c}) = (\mathsf{Z}\mathbf{B})\boldsymbol{c}$
- $\mathsf{Z}(\mathbf{B}\mathbf{C}) = (\mathsf{Z}\mathbf{B})\mathbf{C}$

## 9.2. Sistemas generadores

En la Sección 8.2 definimos el espacio columna de la matriz $\mathbf{A}$ de orden $m$ por $n$ como el conjunto de todas las combinaciones lineales de sus columnas:

$$\mathcal{C}\left(\mathbf{A}\right) = \left\{\boldsymbol{y}\in\mathbb{R}^m \;\middle|\; \text{existen } x_1,\ldots,x_n\in\mathbb{R} \text{ tales que } \boldsymbol{y} = \left(\mathbf{A}_{|1}\right)x_1 + \cdots + \left(\mathbf{A}_{|n}\right)x_n\right\}.$$

o usando notación matricial

$$\mathcal{C}\left(\mathbf{A}\right) = \{\boldsymbol{y}\in\mathbb{R}^n \mid \text{existe } \boldsymbol{x}\in\mathbb{R}^n \text{ tal que } \boldsymbol{y} = \mathbf{A}\boldsymbol{x}\}.$$

Ahora vamos a generalizar esta idea a cualquier espacio vectorial $\mathcal{V}$.

**Definición 9.2** ($\mathcal{L}(\mathsf{Z})$ : subespacio engendrado por un sistema $\mathsf{Z}$)**.** *Dado un sistema* $\mathsf{Z} = \left[\vec{z}_1;\ldots\vec{z}_n;\right]$ *de* $n$ *vectores del espacio vectorial* $\mathcal{V}$*, el* subespacio engendrado por dicho sistema *es*

$$\mathcal{L}(\mathsf{Z}) = \left\{\vec{v}\in\mathcal{V} \;\middle|\; \textit{existen } a_1,\ldots,a_n\in\mathbb{R} \textit{ tales que } \vec{v} = \vec{z}_1a_1 + \cdots + \vec{z}_na_n\right\},$$

*es decir* $\mathcal{L}\big(\mathsf{Z}\big)$ es el conjunto de combinaciones lineales de $\mathsf{Z}_{|j}$ *con* $j = 1:n$*; o usando notación matricial*

$$\mathcal{L}(\mathsf{Z}) = \left\{\vec{v}\in\mathcal{V} \;\middle|\; \textit{existe } \boldsymbol{a}\in\mathbb{R}^n \textit{ tal que } \vec{v} = \mathsf{Z}\boldsymbol{a}\right\}. \tag{9.1}$$

**Definición 9.3** (Sistema generador de $\mathcal{V}$)**.** *Decimos que* $\mathsf{Z}$ *es un sistema generador de* $\mathcal{V}$ *si* $\mathcal{V} = \mathcal{L}(\mathsf{Z})$.

*Ejemplo* 22.

- Las columnas de $\mathbf{A}$ son un *sistema generador* de $\mathcal{C}\left(\mathbf{A}\right)$; es decir $\mathcal{L}\big(\mathbf{A}\big) = \mathcal{C}\left(\mathbf{A}\right)$.

Módulo NAcAL para Python

```
A = Matrix([ Vector([1,0,0]), Vector([-1,0,1]) ])
SubEspacio( A.sis() ) # SubEspacio generado por el Sistema de columnas de A
```

$$\left\{\boldsymbol{v}\in\mathbb{R}^3 \;\middle|\; \exists\boldsymbol{p}\in\mathbb{R}^2,\ \boldsymbol{v} = \begin{bmatrix}1 & -1\\ 0 & 0\\ 0 & 1\end{bmatrix}\boldsymbol{p}\right\} = \left\{\boldsymbol{v}\in\mathbb{R}^3 \;\middle|\; \begin{bmatrix}0 & 1 & 0\end{bmatrix}\boldsymbol{v} = \big(0,\big)\right\}$$

(*Ecuaciones paramétricas a la izquierda y cartesianas a la derecha*)

- Las soluciones especiales de $\mathbf{A}\boldsymbol{x}=\mathbf{0}$ son un *sistema generador* de $\mathcal{N}\left(\mathbf{A}\right)$.

Módulo NAcAL para Python

```
C = Matrix([ [1,0,0], [1,1,1] ])
Homogenea(C,1).sgen    # Sistema generador del espacio nulo de C
```

$$\left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 1 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow{\boldsymbol{\tau}_{[(-1)\mathbf{2}+\mathbf{3}]}} \left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \hline 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{array}\right]$$

$$\left[\ \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix};\ \right]$$

Módulo NAcAL para Python

```
Homogenea(C).enulo    # Espacio nulo de C
```

$$\left\{\boldsymbol{v}\in\mathbb{R}^3 \;\middle|\; \exists \boldsymbol{p}\in\mathbb{R}^1,\ \boldsymbol{v}=\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\boldsymbol{p}\right\} = \left\{\boldsymbol{v}\in\mathbb{R}^3 \;\middle|\; \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}\boldsymbol{v}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right\}$$

☞ $\mathcal{L}(\mathsf{Z})$ es el conjunto de combinaciones lineales de los vectores de $\mathsf{Z}$; y puesto que

- la suma de dos combinaciones lineales $\mathsf{Z}\boldsymbol{b}+\mathsf{Z}\boldsymbol{c}$ es la combinación lineal $\mathsf{Z}\left(\boldsymbol{b}+\boldsymbol{c}\right)$
- el producto de una combinación lineal por un escalar $\left(\mathsf{Z}\boldsymbol{b}\right)\lambda$ es la combinación lineal $\mathsf{Z}\left(\boldsymbol{b}\lambda\right)$

$\mathcal{L}(\mathsf{Z})$ *es un subespacio*, pues es cerrado para la suma y el producto por escalares.

Además, puesto que $\mathcal{L}(\mathsf{Y})$ contiene *únicamente* las combinaciones lineales de $\mathsf{Y}=\left[\vec{y}_1;\ldots\vec{y}_n;\right]$, el subespacio $\mathcal{L}(\mathsf{Y})$ *está contenido en cualquier otro subespacio que también contenga los vectores* $\vec{y}_1,\ldots,\vec{y}_n$; es decir, $\mathcal{L}(\mathsf{Y})$ es el subespacio más pequeño que contiene los vectores $\vec{y}_1,\ldots,\vec{y}_n$ (véase la Sección 9.6).

Un sistema de vectores $\mathsf{Z}$ es un sistema generador de $\mathcal{V}$ si: $\mathcal{V}=\mathcal{L}(\mathsf{Z})$

Ahora ya podemos establecer un criterio para saber si dos sistemas $\mathsf{Y}$ y $\mathsf{Z}$ generan el mismo subespacio:

Sea $\mathsf{Y}$ un sistema de $n$ vectores; puesto que $\mathcal{L}(\mathsf{Y})\subset\mathcal{L}(\mathsf{Z})$ si y solo si $\mathsf{Y}_{|i}\in\mathcal{L}(\mathsf{Z}),\ i=1:n,$ y puesto que dos conjuntos $A$ y $B$ son iguales si y solo si $A\subset B$ y $B\subset A,$ llegamos al siguiente

**Corolario 9.2.1.** *Sean* $\mathsf{Y}$ *con* $n$ *vectores y* $\mathsf{Z}$ *con* $k$ *vectores; entonces* $\mathcal{L}(\mathsf{Y})=\mathcal{L}(\mathsf{Z})$ *si y sólo si*

$$\mathsf{Y}_{|i}\in\mathcal{L}(\mathsf{Z}) \text{ para todo } i=1:n \qquad y \qquad \mathsf{Z}_{|j}\in\mathcal{L}(\mathsf{Y}) \text{ para todo } j=1:k.$$

**Definición 9.4.** *Diremos que dos sistemas generadores son* equivalentes *si generan el mismo subespacio.*

Módulo NAcAL para Python

```
A = Matrix( [[2,-1,0],[-1,2,-1]] )
B = ElimGJ(Matrix(A))
SubEspacio( A.sis() ) == SubEspacio( B.sis() ) #¿Son equivalentes?
```

```
True
```

## 9.3. Transformaciones elementales sobre sistemas de vectores

Al igual que con las matrices, podemos aplicar transformaciones elementales sobre los sistemas de vectores:

- *(Transformación* Tipo I *sobre el sistema de vectores)* Si $i \neq j$ entonces

$$\mathsf{Z} = \left[\vec{z}_1; \dots \vec{z}_i; \dots \vec{z}_j; \dots \vec{z}_n;\right] \xrightarrow{\overset{\tau}{[(\alpha)i+j]}} \left[\vec{z}_1; \dots \vec{z}_i; \dots (\alpha\vec{z}_i + \vec{z}_j); \dots \vec{z}_n;\right] = \mathsf{Z}_{\overset{\tau}{[(\alpha)i+j]}} = \mathsf{Z}\left(\mathsf{I}_{\overset{\tau}{[(\alpha)i+j]}}\right)$$

- *(Transformación* Tipo II *sobre el sistema de vectores)* Si $\alpha \neq 0$ entonces

$$\mathsf{Z} = \left[\vec{z}_1; \dots \vec{z}_i; \dots \vec{z}_n;\right] \xrightarrow{\overset{\tau}{[(\alpha)i]}} \left[\vec{z}_1; \dots (\alpha\vec{z}_i); \dots \vec{z}_n;\right] = \mathsf{Z}_{\overset{\tau}{[(\alpha)i]}} = \mathsf{Z}\left(\mathsf{I}_{\overset{\tau}{[(\alpha)i]}}\right)$$

**Sistemas equivalentes**

Así, si mediante una sucesión de transformaciones elementales $\tau_1 \cdots \tau_k$ transformamos $\mathsf{Z}$ en un nuevo sistema $\mathsf{Y} = \mathsf{Z}_{\tau_1 \cdots \tau_k}$, puesto que por una parte, $\mathsf{Y}_{|j} = \mathsf{Z}\mathbf{E}_{|j}$, donde $\mathbf{E} = \mathsf{I}_{\tau_1 \cdots \tau_k}$; y por otra $\mathsf{Z}_{|j} = \mathsf{Y}\left(\mathbf{E}^{-1}\right)_{|j}$; por el Corolario 9.2.1 concluimos que ambos sistemas son equivalentes:

$$\mathcal{L}(\mathsf{Z}) = \mathcal{L}(\mathsf{Z}\mathbf{E}) = \mathcal{L}(\mathsf{Y}), \quad \text{donde } \mathbf{E} = \mathsf{I}_{\tau_1 \cdots \tau_k}. \tag{9.2}$$

Consideremos ahora una transformación de naturaleza distinta: *quitar vectores nulos de un sistema*:

- *(Quitando vectores nulos)* $\left[\vec{z}_1;\ \vec{0}; \dots \vec{z}_i;\ \vec{0}; \dots \vec{z}_n;\ \vec{0};\right] \xrightarrow{\text{Quitando vectores nulos}} \left[\vec{z}_1; \dots \vec{z}_i; \dots \vec{z}_n;\right]$.

Por el Corolario 9.2.1

$$\mathcal{L}\Big(\left[\vec{z}_1;\ \vec{0}; \dots \vec{z}_i;\ \vec{0}; \dots \vec{z}_n;\ \vec{0};\right]\Big) = \mathcal{L}\Big(\left[\vec{z}_1; \dots \vec{z}_i; \dots \vec{z}_n;\right]\Big).$$

Es más, podemos generalizar el resultado con la siguiente

**Proposición 9.3.1.** *Si* $\mathsf{Z} = \left[\vec{z}_1; \dots \vec{z}_n;\right]$ *es un sistema de vectores entonces*

$$\mathcal{L}\Big(\left[\vec{z}_1; \dots \vec{z}_{i-1};\ \vec{z}_i;\ \vec{z}_{i+1}, \dots \vec{z}_n;\right]\Big) = \mathcal{L}\Big(\left[\vec{z}_1; \dots \vec{z}_{i-1};\ \vec{z}_{i+1}; \dots \vec{z}_{n-1};\right]\Big).$$

*si, y solo si,* $\vec{z}_i$ *es combinación lineal del resto de vectores de* $\mathsf{Z}$.

*Demostración.* Basta aplicar el Corolario 9.2.1. □

☞ Como consecuencia, y volviendo al espacio generado por las columnas de una matriz $\mathbf{A}$:

Si $\mathbf{K}$ es una forma pre-escalonada de $\mathbf{A}$ y si $\boldsymbol{\rho}$ es la lista de índices de columnas no nulas de $\mathbf{K}$, entonces

$$\mathcal{C}(\mathbf{A}) = \mathcal{C}(\mathbf{K}) = \mathcal{C}(\mathbf{K}_{|\boldsymbol{\rho}}),$$

donde $\mathbf{K}_{|\boldsymbol{\rho}}$ es la submatriz correspondiente a las columnas no nulas de $\mathbf{K}$ (véase Apéndice 3.B en la página 39).

Por tanto, las columnas no nulas de cualquier forma pre-escalonada de $\mathbf{A}$ generan $\mathcal{C}(\mathbf{A})$.

*Observación.* Como $\vec{0} \in \mathcal{V}$, y dado que al quitar los vectores nulos de un sistema generador se engendra el mismo espacio, tenemos que $\mathcal{L}\left(\left[\vec{0};\ \vec{0};\ \vec{0};\right]\right) = \mathcal{L}\left(\left[\vec{0};\ \vec{0};\right]\right) = \mathcal{L}\left(\left[\vec{0};\right]\right) = \mathcal{L}\left(\left[\ \right]\right) = \mathcal{L}(\emptyset) = \{\vec{0}\}$; es decir, un sistema (o lista) vacío es el sistema generador del subespacio $\{\vec{0}\}$ (véase la Sección 9.6).

## 9.4. Sistemas linealmente dependientes *vs* independientes

EJERCICIO 46. Demuestre la siguiente...

**Proposición 9.4.1.** *Dado un sistema* $\mathsf{Z} = \left[\vec{z}_1; \dots \vec{z}_n;\right]$ *de* $n$ *vectores de* $\mathcal{V}$*, existe* $i \in \{1, \dots n\}$ *tal que* $z_i$ *es combinación lineal del resto de vectores de* $\mathsf{Z}$ *si, y solo si, existe* $\boldsymbol{a} \neq \mathbf{0}$ *en* $\mathbb{R}^n$ *tal que* $\mathsf{Z}\boldsymbol{a} = \vec{0}$.

Esto da lugar a la definición de *sistema linealmente dependiente* y de *sistema linealmente independiente*:

**Definición 9.5.** *Diremos que el sistema* $\mathsf{Z}$ *de* $n$ *vectores de* $\mathcal{V}$

- *es* linealmente dependiente *si existe* $\boldsymbol{a} \in \mathbb{R}^n$ *no nulo tal que,* $\mathsf{Z}\boldsymbol{a} = \vec{0}$.
- *es* linealmente independiente *si ocurre lo contrario,*[2] *por tanto,* $\mathsf{Z}\boldsymbol{a} = \vec{0}$ *si y solo si* $\boldsymbol{a} = \mathbf{0}$.

(véase también la Sección 9.6.)

**Definición 9.6.** *Sea* $\mathsf{Z}$ *de* $n$ *vectores de* $\mathcal{V}$*. Denominamos espacio nulo de* $\mathsf{Z}$ *al subconjunto* $\mathcal{N}\left(\mathsf{Z}\right)$ *de vectores* $\boldsymbol{x}$ *de* $\mathbb{R}^n$ *tales que* $\mathsf{Z}\boldsymbol{x} = \vec{0}$*, es decir*

$$\mathcal{N}\left(\mathsf{Z}\right) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \mathsf{Z}\boldsymbol{x} = \vec{0}\}.$$

Las anteriores definiciones implican que $\mathsf{Z}$ es linealmente independiente si y solo si $\mathcal{N}\left(\mathsf{Z}\right) = \{\mathbf{0}\}$.

**Ejemplos de sistemas linealmente independientes**

- Las columnas de una matriz $\mathbf{A}$ *invertible* constituyen un sistema linealmente independiente, puesto que

$$\mathbf{A}\boldsymbol{x} = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{A}^{-1}\mathbf{A}\boldsymbol{x} = \mathbf{A}^{-1}\mathbf{0} \quad \Leftrightarrow \quad \boldsymbol{x} = \mathbf{0}.$$

- Si, de un sistema $\mathsf{Z}$ de vectores linealmente independientes, se quita uno cualquiera de sus vectores, el sistema resultante sigue siendo linealmente independiente.
- Por tanto, cualquier sistema formado por una selección de columnas de una matriz invertible (sin repetición) $\mathbf{E}_{|\gamma}$ es un sistema linealmente independiente.
- Como las soluciones especiales son una selección de columnas de una matriz invertible $\mathbf{E}$, *las soluciones especiales constituyen un sistema linealmente independiente*.

Ahora que hemos definido los sistemas linealmente independientes, ya podemos enunciar la siguiente proposición sobre matrices de rango completo por columnas (son equivalencias vistas anteriormente):

**Proposición 9.4.2.** *Para* $\mathbf{A} = \left[\boldsymbol{v}_1; \dots \boldsymbol{v}_n;\right]$ *de* $n$ *vectores de* $\mathbb{R}^m$*, las siguientes propiedades son equivalentes:*

- *El rango de* $\mathbf{A} = \left[\boldsymbol{v}_1; \dots \boldsymbol{v}_n;\right]$ *es* $n$.
- $\mathcal{N}\left(\mathbf{A}\right) = \{\mathbf{0}\}$.
- *La combinación lineal* $\mathbf{A}\boldsymbol{x}$ *es* $\mathbf{0}$ *si y solo si* $\boldsymbol{x} = \mathbf{0}$.
- *Las columnas de* $\mathbf{A}$ *son linealmente independientes.*

## 9.5. Bases y dimensión

Hemos visto que en un sistema de vectores linealmente dependiente es posible retirar algún vector sin reducir el espacio generado. Pero si el sistema es linealmente independiente, retirar cualquiera de los vectores reduce el espacio engendrado por ellos. Un sistema que tenga el tamaño justo para generar el subespacio $\mathcal{V}$ sin que sobre (sin que se pueda retirar) ningún vector tiene un nombre especial:

**Definición 9.7.** *Diremos que el sistema* $\mathsf{B}$ *es* una base *de un subespacio* $\mathcal{V}$ *si simultáneamente*

1. $\mathsf{B}$ *genera el subespacio, es decir, si* $\mathcal{L}(\mathsf{B}) = \mathcal{V}$.
2. $\mathsf{B}$ *es linealmente independiente, es decir,* $\mathsf{B}\boldsymbol{x} = \vec{0}$ *si y solo si* $\boldsymbol{x} = \mathbf{0}$.

[2] Para ver por qué un sistema vacío, $\mathsf{Z} = [\ ]$, es linealmente independiente; véase la Sección 9.6.

Es decir, una base $\mathsf{B}$ de $\mathcal{V}$ es un sistema de vectores suficientemente reducido para ser linealmente independiente pero suficientemente grande como para generar el subespacio $\mathcal{V}$.

*Ejemplo* 23. Las soluciones *especiales* encontradas al resolver $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ constituyen una base de $\mathcal{N}\left(\mathbf{A}\right)$.

*Justificación:* Por la Proposición 7.4.2 en la página 106 sabemos que las soluciones especiales son un sistema generador de $\mathcal{N}\left(\mathbf{A}\right)$ y, puesto que son columnas de una matriz invertible, son linealmente independientes.

Un importante resultado es que todas las bases de un subespacio $\mathcal{V}$ tienen el mismo número de vectores. Usaremos la siguiente proposición para demostrarlo:

**Proposición 9.5.1.** *Si* $\mathsf{X}$ *y* $\mathsf{Z}$ *son dos sistemas con* $p$ *y* $q$ *vectores de* $\mathcal{V}$*, respectivamente, donde* $p > q$*; tales que* $\mathcal{L}\left(\mathsf{X}\right) \subset \mathcal{L}\left(\mathsf{Z}\right)$*, entonces* $\mathsf{X}$ *es linealmente dependiente.*

*Demostración.* Como cada $\mathsf{X}_{|i}$ es combinación lineal de los vectores del sistema $\mathsf{Z}$, existe $\underset{p\times q}{\mathbf{M}}$ tal que

$$\mathsf{X}_{|j} = \mathsf{Z}\mathbf{M}_{|j} \text{ para } j = 1 : p; \text{ es decir, tal que } \mathsf{X} = \mathsf{Z}\mathbf{M}.$$

Como $\mathbf{M}$ tiene más columnas que filas, existe $\boldsymbol{a} \neq \mathbf{0}$ tal que $\mathbf{M}\boldsymbol{a} = \mathbf{0}$; y por tanto $\mathsf{X}\boldsymbol{a} = \mathsf{Z}\mathbf{M}\boldsymbol{a} = \mathsf{Z}\mathbf{0} = \vec{0}$. □

Así pues, *dos bases de un mismo subespacio tienen el mismo número de vectores*; pues si una base tuviera más que la otra, por la Proposición 9.5.1, sería un sistema dependiente, y por tanto ya no sería una base.

**Definición 9.8.** *Decimos que un subespacio* $\mathcal{V}$ *es de dimensión finita cuando tiene una base con un número finito de vectores. En tal caso, llamamos dimensión de* $\mathcal{V}$ *al número de vectores de cualquiera de sus bases.*

En la sección anterior vimos un procedimiento para generar una base quitando vectores. Veamos otro procedimiento para completar una base añadiendo vectores.

**Proposición 9.5.2.** *Si* $\left[\vec{z}_1; \ldots \vec{z}_n;\right]$ *es un sistema linealmente independiente y* $\vec{z}_{n+1} \notin \mathcal{L}\Big(\left[\vec{z}_1; \ldots \vec{z}_n;\right]\Big)$ *entonces* $\left[\vec{z}_1; \ldots \vec{z}_n;\ \vec{z}_{n+1};\right]$ *es linealmente independiente.*

*Demostración.* Supongamos que $a_1\vec{z}_1 + \cdots + a_n\vec{z}_n + a_{n+1}\vec{z}_{n+1} = \vec{0}$. Como $\vec{z}_{n+1} \notin \mathcal{L}\Big(\left[\vec{z}_1; \ldots \vec{z}_n;\right]\Big)$, necesariamente $a_{n+1} = 0$. Y como $\left[\vec{z}_1; \ldots \vec{z}_n;\right]$ es linealmente independiente, también $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n = 0$. □

**Corolario 9.5.3.** *Cualquier subespacio* $\mathcal{W}$*, de un espacio* $\mathcal{V}$ *de dimensión finita, tiene dimensión finita menor o igual a la dimensión de* $\mathcal{V}$*.*

*Demostración.* Sea $\mathcal{W}$ un subespacio de $\mathcal{V}$. Sabemos que existen sistemas de vectores de $\mathcal{W}$ linealmente independientes. Por ejemplo, el vacío. Por otra parte, cualquier sistema de vectores linealmente independientes tiene como mucho tantos vectores como la dimensión del espacio. De todos los posibles sistemas linealmente independientes de $\mathcal{W}$, tomemos un sistema $\mathsf{M}$ con el mayor número posible de vectores. Entonces $\mathcal{L}\left(\mathsf{M}\right) \subset \mathcal{W}$. Veamos que si suponemos que existe $\vec{v} \in \mathcal{W}$ tal que $\vec{v} \notin \mathcal{L}\left(\mathsf{M}\right)$, llegamos a una contradicción. Basta añadir $\vec{v}$ al sistema $\mathsf{M}$ para obtener un sistema linealmente independiente formado por vectores de $\mathcal{W}$. Pero esto contradice que $\mathsf{M}$ sea un sistema linealmente independiente con el mayor número posible de vectores. □

**Corolario 9.5.4.** *Si* $\mathcal{W}$ *es un subespacio de* $\mathcal{V}$ *de la misma dimensión de* $\mathcal{V}$*, entonces* $\mathcal{W}$ *y* $\mathcal{V}$ *son iguales.*

*Demostración.* Tomemos una base $\mathsf{B}$ de $\mathcal{W}$. Entonces sabemos que $\mathcal{L}(\mathsf{B}) = \mathcal{W}$. Para demostrar que $\mathcal{L}(\mathsf{B}) = \mathcal{V}$, supongamos lo contrario. Supongamos que existe $\vec{v} \in \mathcal{V}$ tal que $\vec{v} \notin \mathcal{L}(\mathsf{B})$. Entonces podríamos incluir $\vec{v}$ en $\mathsf{B}$ y obtendríamos un sistema linealmente independiente con más elementos que la dimensión de $\mathcal{V}$. □

**Corolario 9.5.5.** *Cualquier sistema* $\mathsf{B}$ *de vectores de* $\mathcal{W}$ *linealmente independiente cuyo número de vectores coincida con la dimensión de* $\mathcal{W}$ *es necesariamente una base.*

### 9.5.1. Eliminación "de izquierda a derecha" y sistemas "acoplados" de vectores

☞ Vamos a diseñar un algoritmo para encontrar una base de $\mathcal{L}(\mathsf{Z})$. La mecánica es sencilla y se basa en la Ecuación 9.2, es decir, $\mathcal{L}(\mathsf{Z}) = \mathcal{L}(\mathsf{Z}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k})$

- consiste en aplicar la eliminación de "izquierda a derecha" sobre $\mathsf{Z}$, y observar qué vectores de $\mathsf{Z}$ no se anulan. Aquellos vectores de $\mathsf{Z}$ que no se anulen constituyen una base de $\mathcal{L}(\mathsf{Z})$.

Aunque el procedimiento es sencillo, demostrar por qué funciona no es inmediato. Aprovecharemos para introducir notación y nuevas operaciones; a cambio lograremos una comprensión más profunda de las propiedades de los sistemas obtenidos en el proceso de eliminación ("de izquierda a derecha").

**Notación.** Denotaremos con $\mathsf{Y}_{|(1:s)}$ al subsistema de $\mathsf{Y} = [\vec{y}_1; \dots \vec{y}_n;]$ formado por sus primeros $s$ vectores

$$\mathsf{Y}_{|(1:s)} = \left[\vec{y}_1; \dots \vec{y}_s;\right] = \left[\vec{y}_j \mid j < s\right] \quad \text{con} \quad s \leq n,$$

y asumiremos que $\mathcal{L}(\mathsf{Z}_{|(1:0)}) = \mathcal{L}([\,]) = \{\vec{0}\}$; (véase la Sección 9.6 en la página 128 y (a pie de página) la Nota 6 en la página 4).

**Definición 9.9** (Sistemas acoplados)**.** *Los sistemas* $\mathsf{Y}$ *y* $\mathsf{Z}$*, de* $n$ *vectores cada uno, están* acoplados *si*

$$\mathcal{L}\left([\vec{y}_1; \dots \vec{y}_i;]\right) = \mathcal{L}\left([\vec{z}_1; \dots \vec{z}_i;]\right) \quad \textit{para } i = 1:n,$$

*es decir, si*

$$\mathcal{L}\left(\mathsf{Y}_{|(1:i)}\right) = \mathcal{L}\left(\mathsf{Z}_{|(1:i)}\right) \quad \textit{para } i = 1:n.$$

**Proposición 9.5.6** (Eliminación "de izquierda a derecha" en $\mathsf{Z}$)**.** *Si transformamos* $\mathsf{Z} = [\vec{z}_1; \dots \vec{z}_n;]$ *mediante una secuencia* $\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k$ *de transformaciones elementales* "de izquierda a derecha"

$$\mathsf{Z}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} = \mathsf{Z}\mathbf{E} = \mathsf{Y}; \quad \textit{donde } \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}, \textit{ es triangular superior unitaria,}^{3} \textit{es decir,}$$

$$\mathsf{Y} = [\vec{y}_1; \dots \vec{y}_n;] = [\vec{z}_1; \dots \vec{z}_n;] \begin{bmatrix} 1 & e_{12} & e_{13} & \cdots & e_{1n} \\ & 1 & e_{23} & \cdots & e_{2n} \\ & & 1 & \cdots & e_{3n} \\ & & & \ddots & \vdots \\ & & & & 1 \end{bmatrix} = \mathsf{Z}\mathbf{E},$$

*entonces* $\mathsf{Y}$ *y* $\mathsf{Z}$ *están acoplados, es decir,* $\mathcal{L}\left(\mathsf{Y}_{|(1:i)}\right) = \mathcal{L}\left(\mathsf{Z}_{|(1:i)}\right)$ *para* $i = 1:n$.

*Demostración.* Por una parte, como $\mathbf{E}$ es triangular superior unitaria, cada vector $\mathsf{Y}_{|j} = \mathsf{Z}\mathbf{E}_{|j}$ es el vector $\mathsf{Z}_{|j}$ más una combinación lineal de los vectores $\mathsf{Z}_{|k}$ con $k < j$, y por tanto $\mathcal{L}(\mathsf{Y}_{|(1:j)}) \subset \mathcal{L}(\mathsf{Z}_{|(1:j)})$ para $j = 1:n$.

Por otra parte, como $\mathsf{Z} = \mathsf{Y}\mathbf{E}^{-1}$, y como $\mathbf{E}^{-1}$ también es triangular superior unitaria (Corolario 5.1.17), usando el mismo argumento del párrafo anterior concluimos que $\mathcal{L}(\mathsf{Z}_{|(1:j)}) \subset \mathcal{L}(\mathsf{Y}_{|(1:j)})$ para $j = 1:n$. □

**Notación.** Sean $\mathsf{Y} = [\vec{y}_1; \dots \vec{y}_n;]$ y $\mathsf{Z} = [\vec{z}_1; \dots \vec{z}_k;]$ dos sistemas de vectores de $\mathcal{V}$. Denotaremos con $\mathsf{Y} +\!\!+ \mathsf{Z}$ al sistema de $n + k$ vectores resultante de *concatenar* $\mathsf{Y}$ (de $n$ vectores) y $\mathsf{Z}$ (de $k$ vectores)

$$\mathsf{Y} +\!\!+ \mathsf{Z} = [\vec{y}_1; \dots \vec{y}_n;\ \vec{z}_1; \dots \vec{z}_k;].$$

[3] por el Teorema 5.1.16 en la página 78

**Lema 9.5.7.** *Sean* $\mathsf{Y} = [\vec{y}_1; \dots \vec{y}_n;]$ *y* $\mathsf{Z} = [\vec{z}_1; \dots \vec{z}_k;]$ *dos sistemas de vectores de* $\mathcal{V}$ *tales que* $\mathcal{L}(\mathsf{Y}) = \mathcal{L}(\mathsf{Z})$. *Entonces, para cualquier sistema* $\mathsf{X} = [\vec{x}_1; \dots \vec{x}_r;]$ *de vectores de* $\mathcal{V}$, *se verifica que*

$$\mathcal{L}\Big(\mathsf{Y} +\!\!+ \mathsf{X}\Big) = \mathcal{L}\Big(\mathsf{Z} +\!\!+ \mathsf{X}\Big).$$

*Demostración.* Hay que demostrar que $\{\vec{z}_1, \dots, \vec{z}_k,\ \vec{x}_1, \dots, \vec{x}_r\} \subset \mathcal{L}(\mathsf{Y} +\!\!+ \mathsf{X})$ y que $\{\vec{y}_1, \dots, \vec{y}_n,\ \vec{x}_1, \dots, \vec{x}_r\} \subset \mathcal{L}(\mathsf{Z} +\!\!+ \mathsf{X})$. Como ambos contenidos se prueban de la misma manera, solo comprobaremos el primero.

Por un lado tenemos que $\{\vec{z}_1, \dots, \vec{z}_k\} \subset \mathcal{L}(\mathsf{Z}) = \mathcal{L}(\mathsf{Y}) \subset \mathcal{L}(\mathsf{Y} +\!\!+ \mathsf{X})$. Por otro que $\{\vec{x}_1, \dots, \vec{x}_r\} \subset \mathcal{L}(\mathsf{Y} +\!\!+ \mathsf{X})$. Luego $\{\vec{z}_1, \dots, \vec{z}_k\} \cup \{\vec{x}_1, \dots, \vec{x}_r\} \subset \mathcal{L}(\mathsf{Y} +\!\!+ \mathsf{X})$. □

**Corolario 9.5.8.** *Si* $\mathsf{Y} = [\vec{y}_1; \dots \vec{y}_n;]$ *y* $\mathsf{Z} = [\vec{z}_1; \dots \vec{z}_n;]$ *son dos sistemas acoplados y* $\vec{z}_k$ *es combinación de los vectores que le anteceden en la lista* $\mathsf{Z}$

$$\vec{z}_k \in \mathcal{L}\Big(\mathsf{Z}_{|(1:k-1)}\Big) = \mathcal{L}\Big([\vec{z}_j \mid j < k]\Big), \quad \textit{entonces} \tag{9.3}$$

*los sistemas resultantes tras quitar el* $k$*-ésimo vector,* $[\vec{z}_j \mid j \neq k]$ *y* $[\vec{y}_j \mid j \neq k]$, *también están acoplados.*

*Demostración.* El resultado es trivial si $k = n$. Así que supongamos que $k < n$. Puesto que $\mathsf{Z}$ y $\mathsf{Y}$ están acoplados, $\mathcal{L}(\mathsf{Y}_{|(1:i)}) = \mathcal{L}(\mathsf{Z}_{|(1:i)})$ para todo $i < k$, y como consecuencia de (9.3) tendremos que

$$\mathcal{L}\Big(\mathsf{Y}_{|(1:k-1)}\Big) = \mathcal{L}\Big(\mathsf{Z}_{|(1:k-1)}\Big) \overset{(9.3)}{=} \mathcal{L}\Big(\mathsf{Z}_{|(1:k)}\Big) = \mathcal{L}\Big(\mathsf{Y}_{|(1:k)}\Big)$$

luego, aplicando el Lema 9.5.7 (en $\overset{*}{=}$) tendremos que para todo $k < i$

$$\begin{aligned}
\mathcal{L}\Big([\vec{y}_j \mid j \neq k]\Big) = \mathcal{L}\Big(\mathsf{Y}_{|(1:k-1)} +\!\!+ \mathsf{Y}_{|(k+1:i)}\Big) &\overset{*}{=} \mathcal{L}\Big(\mathsf{Y}_{|(1:k)} +\!\!+ \mathsf{Y}_{|(k+1:i)}\Big) = \mathcal{L}\Big(\mathsf{Y}_{|(1:i)}\Big) = \mathcal{L}\Big(\mathsf{Z}_{|(1:i)}\Big) \\
&= \mathcal{L}\Big(\mathsf{Z}_{|(1:k)} +\!\!+ \mathsf{Z}_{|(k+1:i)}\Big) \\
&\overset{*}{=} \mathcal{L}\Big(\mathsf{Z}_{|(1:k-1)} +\!\!+ \mathsf{Z}_{|(k+1:i)}\Big) = \mathcal{L}\Big([\vec{z}_j \mid j \neq k]\Big).
\end{aligned}$$

□

**Corolario 9.5.9.** *Si* $\mathsf{Y}$ *y* $\mathsf{Z}$ *son dos sistemas acoplados de* $n$ *vectores y* $\vec{y}_s = \vec{0}$ *entonces los sistemas* $[\vec{y}_j \mid j \neq s]$ *y* $[\vec{z}_j \mid j \neq s]$ *también están acoplados.*

**Definición 9.10.** *Si* $[\vec{y}_1; \dots \vec{z}_y;]$ *y* $[\vec{z}_1; \dots \vec{z}_n;]$ *son dos sistemas acoplados, diremos que el par* $(\vec{y}_s, \vec{z}_s)$ *es superfluo si* $\vec{y}_s = \vec{0}$. *En tal caso diremos que* $[\vec{y}_j \mid j \neq s]$ *y* $[\vec{z}_j \mid j \neq s]$ *son los sistemas acoplados resultantes de quitar el par superfluo* $(\vec{y}_s, \vec{z}_s)$.

**Corolario 9.5.10.** *Si* $[\vec{y}_1; \dots \vec{y}_n;]$ *y* $[\vec{z}_1; \dots \vec{z}_n;]$ *son sistemas acoplados, también lo son* $[\vec{y}_j \mid \vec{y}_j \neq \vec{0}]$ *y* $[\vec{z}_j \mid \vec{y}_j \neq \vec{0}]$.

*Demostración.* Basta con quitar los pares superfluos correspondientes a los ceros en el sistema $[\vec{y}_1; \dots \vec{y}_n;]$. □

☞ El anterior corolario nos muestra una forma de encontrar una base de $\mathcal{L}(\mathsf{Z})$ entre los vectores de $\mathsf{Z}$: Basta seleccionar el subsistema con los vectores de $\mathsf{Z}_{|j}$ que no se anulen tras aplicar el método de eliminación $\mathsf{Z}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$, es decir

$$\Big[\mathsf{Z}_{|j} \mid \mathsf{Z}\mathbf{E}_{|j} \neq \vec{0}\Big], \quad \text{donde} \quad \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}.$$

Así, tenemos otro ejemplo de base de un subespacio:

*Ejemplo* 24. Las columnas no nulas de una forma escalonada de $\mathbf{A}$ constituyen una base de $\mathcal{C}\left(\mathbf{A}\right)$.

*Justificación:* Sea $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ una forma escalonada de $\mathbf{A}$, por (9.2) sabemos que $\mathcal{L}\left(\mathbf{A}\right)=\mathcal{L}\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)$; así, al seleccionar las columnas no nulas de $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ creamos un sistema linealmente independiente.

**Encontrando dos bases de $\mathcal{C}\left(\mathbf{A}\right)$ y una base de $\mathcal{N}\left(\mathbf{A}\right)$**

Concluimos que aplicando la eliminación sobre la matriz resultante de concatenar las $n$ columnas de $\mathbf{A}$ con las columnas de la matriz identidad de orden $n$, encontramos dos bases de $\mathcal{C}\left(\mathbf{A}\right)$ y una base de $\mathcal{N}\left(\mathbf{A}\right)$:

$$\left[\begin{array}{c}\mathbf{A}\\\hline\mathbf{I}\end{array}\right]=\left[\begin{array}{cccc}0&0&1&2\\1&2&1&3\\2&4&3&8\\\hline1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right]\xrightarrow[\substack{[(-3)\mathbf{1}+\mathbf{4}]\\[(-1)\mathbf{1}+\mathbf{3}]\\[(-2)\mathbf{1}+\mathbf{2}]}]{\boldsymbol{\tau}}\left[\begin{array}{cccc}0&0&1&2\\1&0&0&0\\2&0&1&2\\\hline1&-2&-1&-3\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right]\xrightarrow[[(-2)\mathbf{3}+\mathbf{4}]]{\boldsymbol{\tau}}\left[\begin{array}{cccc}0&0&1&0\\1&0&0&0\\2&0&1&0\\\hline1&-2&-1&-1\\0&1&0&0\\0&0&1&-2\\0&0&0&1\end{array}\right]=\left[\begin{array}{c}\mathbf{K}\\\hline\mathbf{E}\end{array}\right];$$

$\mathcal{C}\left(\mathbf{A}\right)=\mathcal{L}\Big(\left[\mathbf{K}_{|1};\ \mathbf{K}_{|3}\right]\Big)$ es de dimensión 2 por ser $\mathbf{K}_{|1}$ y $\mathbf{K}_{|3}$ linealmente independientes; y por el Corolario 9.5.10 sabemos que $\mathcal{C}\left(\mathbf{A}\right)=\mathcal{L}\Big(\left[\mathbf{A}_{|1};\ \mathbf{A}_{|3}\right]\Big)$ y, como $\left[\mathbf{A}_{|1};\ \mathbf{A}_{|3}\right]$ tiene dos vectores, es una base de $\mathcal{C}\left(\mathbf{A}\right)$. Por otra parte, ya vimos (Página 123) que las soluciones especiales $\mathbf{E}_{|2}$ y $\mathbf{E}_{|4}$ son una base de $\mathcal{N}\left(\mathbf{A}\right)$.

Es evidente que, con este procedimiento, el número de columnas que se anulan (la dimensión de $\mathcal{N}\left(\mathbf{A}\right)$) más el número de columnas que no se anulan (la dimensión de $\mathcal{C}\left(\mathbf{A}\right)$) es el número total de columnas de $\mathbf{A}$. Por tanto, podemos enunciar el siguiente

**Corolario 9.5.11.** $\dim\mathcal{C}\left(\mathbf{A}\right)+\dim\mathcal{N}\left(\mathbf{A}\right)$ *es igual al número de columnas de* $\mathbf{A}$.

Módulo NAcAL para Python

```
A = Matrix([ [0,0,1,2],[1,2,1,3],[2,4,3,8]  ]);
SubEspacio( A.sis() ).base   # una base del SubEspacio
```

$$\left[\ \begin{pmatrix}0\\1\\2\end{pmatrix};\quad\begin{pmatrix}1\\1\\3\end{pmatrix};\ \right]$$

Módulo NAcAL para Python

```
SubEspacio( A.sis() ).dim    # dimensión del SubEspacio
```

```
2
```

Módulo NAcAL para Python

```
Homogenea(A).enulo.dim    # dimensión del Espacio Nulo
```

```
2
```

Pero nótese que, en general, ambos son subespacios de espacios distintos: $\mathcal{C}\left(\mathbf{A}\right)\subset\mathbb{R}^m$ y $\mathcal{N}\left(\mathbf{A}\right)\subset\mathbb{R}^n$.

Pero el anterior resultado no se limita a sistemas de vectores de $\mathbb{R}^m$. De manera análoga, si mediante eliminación anulamos todo vector $\mathsf{Z}_{|j}$ que sea combinación lineal de los vectores de $\mathsf{Z}$ que estén a su izquierda; es decir, si $\mathsf{Y} = \mathsf{Z}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$, donde $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ y

$$\left[\begin{array}{c}\mathsf{Z}\\ \hline \mathbf{I}\end{array}\right] \xrightarrow{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} \left[\begin{array}{c}\mathsf{Y}\\ \hline \mathbf{E}\end{array}\right] = \left[\begin{array}{ccccccc} \vec{y}_1; & \vec{0}; & \vec{y}_3; & \vec{0}; & \cdots & \vec{y}_n; \\ \hline 1 & e_{12} & e_{13} & e_{14} & \cdots & e_{1n} \\ & 1 & e_{23} & e_{24} & \cdots & e_{2n} \\ & & 1 & e_{34} & \cdots & e_{3n} \\ & & & 1 & \cdots & e_{3n} \\ & & & & \ddots & \vdots \\ & & & & & 1 \end{array}\right]; \tag{9.4}$$

el número total de vectores no anulados más el número total de soluciones especiales es igual al número de vectores del sistema $\mathsf{Z}$. Por tanto, también es cierto el siguiente

**Corolario 9.5.12.** $\dim \mathcal{L}(\mathsf{Z}) + \dim \mathcal{N}(\mathsf{Z})$ *es igual al número de vectores de* $\mathsf{Z}$.

Pero nótese que, en general, ambos son subespacios de espacios distintos: $\mathcal{L}(\mathsf{Z}) \subset \mathcal{V}$ y $\mathcal{N}(\mathbf{A}) \subset \mathbb{R}^n$.

☞ Si tenemos un sistema de vectores y aplicamos el método de eliminación (de *"izquierda a derecha"*), los sucesivos sistemas que van apareciendo tienen la propiedad de que los subespacios generados por los $j$ primeros vectores de los distintos sistemas generan el mismo espacio. Es decir, para $\mathsf{Z}$, y la sucesión de transformaciones elementales $\boldsymbol{\tau}_k, \ldots, \boldsymbol{\tau}_n$ tal que $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ es triangular superior unitaria:

$$\text{si} \quad \mathsf{Y} = \mathsf{Z}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} = \mathsf{Z}\mathbf{E}, \quad \text{entonces } \mathsf{Y} \text{ y } \mathsf{Z} \text{ están acoplados.}$$

Es decir,

$$\mathcal{L}\Big([\mathsf{Z}_{|1}; \ldots \mathsf{Z}_{|j};]\Big) \;=\; \mathcal{L}\Big([\mathsf{Y}_{|1}; \ldots \mathsf{Y}_{|j};]\Big) \quad \text{para } j = 1 : n.$$

Si algunos vectores de $\mathsf{Y} = \mathsf{Z}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ acaban anulándose (tras la eliminación de *"izquierda a derecha"*), los correspondientes vectores de $\mathsf{Z}$ se pueden quitar y el sistema resultante continúa generando el mismo subespacio (Corolario 9.5.10).

$$\mathcal{L}(\mathsf{Z}) = \mathcal{L}(\mathsf{Z}_{|\boldsymbol{\rho}}) \quad \text{donde} \quad \boldsymbol{\rho} = \big(i \mid (\mathsf{Z}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k})_{|i} \neq \vec{0}\big).$$

Hemos usado esta idea para encontrar una base de $\mathcal{C}(\mathbf{A})$ entre las columnas de $\mathbf{A}$ pre-escalonando la matriz mediante eliminación (de *"izquierda a derecha"*).

Recordando la Proposición 5.2.5, completemos el Corolario 8.1.2 con cuatro nuevas afirmaciones:

**Corolario 9.5.13.** *Para toda matriz cuadrada* $\mathbf{A}$ *de orden* $n$*, las siguientes afirmaciones son equivalentes:*

1. *El rango de* $\mathbf{A}$ *es* $n$ *(ó* $\mathbf{A}$ *es de rango completo).*
2. $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ *si y solo si* $\boldsymbol{x} = \mathbf{0}$.
3. $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ *tiene solución única.*
4. $\mathbf{A}$ *no es singular.*
5. $\mathbf{A}$ *es invertible.*
6. $\mathbf{A} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$. *(prod. de mat. elementales)*
7. *Las columnas de* $\mathbf{A}$ *son linealmente indep.*
8. *Las filas de* $\mathbf{A}$ *son linealmente independientes.*
9. *Las columnas de* $\mathbf{A}$ *forman una base de* $\mathbb{R}^n$.
10. *Las filas de* $\mathbf{A}$ *forman una base de* $\mathbb{R}^n$.

## 9.6. Generalizaciones o visiones alternativas (★)

**Conjunto generador de un subespacio**

En ocasiones es necesario definir el subespacio engendrado por un *conjunto* $Z$ de vectores en lugar del espacio engendrado por un *sistema* $\mathsf{Z}$ de vectores.[4]

**Definición 9.11.** *Dado un conjunto $Z$ de vectores de $\mathcal{V}$, el* subespacio $\mathcal{L}(Z)$ de $\mathcal{V}$ engendrado por $Z$ *es*

$$\mathcal{L}(Z) = \left\{ \vec{v} \in \mathcal{V} \;\middle|\; \text{existen } a_1, \dots, a_n \in \mathbb{R} \text{ y } \vec{z}_1, \dots, \vec{z}_n \in \mathcal{V} \text{ tales que } \vec{v} = a_1 \vec{z}_1 + \cdots + a_n \vec{z}_n \right\}.$$

**Envolvente lineal $\mathcal{L}(\mathsf{Z})$ de un conjunto $Z$.** En un recuadro de la Sección 9.2 se dio una segunda caracterización para $\mathcal{L}(\mathsf{Z})$ que, según el caso, resulta más operativa; y que enunciamos de nuevo en relación al *conjunto* generador $Z$:

*Dado un subconjunto $Z$ de vectores de $\mathcal{V}$, el* subespacio $\mathcal{L}(Z)$ de $\mathcal{V}$ engendrado por $Z$ *es el subespacio más pequeño de $\mathcal{V}$ que contiene a $Z$.*

$$\mathcal{L}(Z) = \bigcap \{ \mathcal{H} \mid \mathcal{H} \text{ es un subespacio vectorial tal que } Z \subset \mathcal{H} \;\}.$$

Por este motivo $\mathcal{L}(Z)$ también se denomina *envolvente lineal de $Z$*.

**Sistema (o conjunto) vacío como sistema generador.** Con esta caracterización se hace evidente que el subespacio de $\mathcal{V}$ engendrado por el conjunto vacío es el subespacio $\{\vec{0}\}$ de $\mathcal{V}$ que únicamente contiene el vector nulo $\vec{0} \in \mathcal{V}$; pues es el menor subespacio de $\mathcal{V}$ que contiene el conjunto vacío:

$$\mathcal{L}(\emptyset) = \mathcal{L}([\,]) = \{\vec{0}\};$$

recuerde que un sistema vacío es equivalente al conjunto vacío (Ecuación 6.7 en la página 99) y que todo subespacio contiene al vector nulo.

*Observación.* Esto nos indica que la notación $\mathcal{L}([\,])$ es incompleta por no indicar de quién es el subespacio que estamos generando con $[\,]$. Lo correcto sería escribir algo como[5]

$$\mathcal{L}_{\mathcal{V}}([\,]) = \{\vec{0}\} \subset \mathcal{V} \qquad \text{ó} \qquad \mathcal{L}_{\mathbb{R}^n}([\,]) = \{\mathbf{0}\} \subset \mathbb{R}^n;$$

pero arrastrar una notación de la forma $\mathcal{L}_{\mathcal{V}}(X)$, donde se indica explícitamente el espacio vectorial $\mathcal{V}$, sería pesado y normalmente innecesario, pues solo pueden surgir dudas cuando el sistema generador es vacío.

**Dependencia e independencia lineal** Con la última caracterización de subespacio engendrado por un conjunto también podemos caracterizar los conjuntos de vectores linealmente dependientes e independientes.

*Diremos que un subconjunto $Z$ de vectores de $\mathcal{V}$, es* linealmente dependiente *si es posible quitar algún vector de $Z$ y que el conjunto resultante $Z^*$ siga engendrando el mismo subespacio, es decir, que $\mathcal{L}(Z) = \mathcal{L}(Z^*)$.*

Un conjunto de vectores que no es linealmente dependiente se dice que es *linealmente independiente*. Por tanto, un sistema vacío, $\mathsf{Z} = [\,]$, es linealmente independiente. Consecuentemente, $[\,]$ es una base del subespacio $\{\vec{0}\}$, de lo cual se deduce que $\dim(\{\vec{0}\}) = 0$.

[4] Lo necesitaremos en la Sección 11.2; y también es necesario cuando se trabaja con espacios vectoriales de dimensión infinita.

[5] De hecho, lo correcto sería escribir $\mathcal{L}_{(\mathcal{V},+,\cdot)}([\,])$; pues para denotar un espacio vectorial es necesario indicar no solo el conjunto "$\mathcal{V}$", sino las operaciones suma "$+$" y producto por escalares "$\cdot$" (véase el Ejercicio 35 en la página 87).

**Sistemas de vectores y funciones lineales**

Las proposiciones 9.1.1 y 9.1.2 implican que $\mathsf{Z}\boldsymbol{x}$ es una función lineal; lo que sugiere el uso de la siguiente:

**Notación.** Sea $\mathsf{Z}$ un sistema de $n$ vectores de $\mathcal{V}$, denotaremos con $f_{\mathsf{Z}}$ a la función lineal:

$$\begin{aligned} f_{\mathsf{Z}} : \mathbb{R}^n &\longrightarrow \mathcal{V} \\ \boldsymbol{x} &\longmapsto \mathsf{Z}\boldsymbol{x} \end{aligned} .$$

**Proposición 9.6.1.** *Para cualquier función* lineal $f : \mathbb{R}^n \to \mathcal{V}$*, existe* $\mathsf{Z}$ *de* $n$ *vectores de* $\mathcal{V}$ *tal que* $f(\boldsymbol{x}) = \mathsf{Z}\boldsymbol{x}$ *para todo* $\boldsymbol{x} \in \mathbb{R}^n$*.*

*Demostración.*

$$\begin{aligned} f(\boldsymbol{x}) =& f(\mathbf{I}\boldsymbol{x}) \\ =& f\big(x_1\mathbf{I}_{|1} + \cdots + x_n\mathbf{I}_{|n}\big) && \text{por la definición de matriz por vector de } \mathbb{R}^n \\ =& x_1 f(\mathbf{I}_{|1}) + \cdots + x_n f(\mathbf{I}_{|n}) && \text{por ser } f \text{ una función lineal} \\ =& \big[f(\mathbf{I}_{|1});\ \ldots f(\mathbf{I}_{|n})\big]\boldsymbol{x} && \text{por la definición de sistema por vector de } \mathbb{R}^n \\ =& \mathsf{Z}\boldsymbol{x} && \text{donde } \mathsf{Z} = \big[f(\mathbf{I}_{|1});\ \ldots f(\mathbf{I}_{|n});\big]. \end{aligned}$$

□

Nótese que tanto $\mathsf{Z}\mathbf{B}\boldsymbol{c}$ como $\mathsf{Z}\mathbf{B}\mathbf{C}$ son composiciones de funciones lineales, ya que, por una parte: si $\mathsf{Z}$ es un sistema de $m$ vectores de $\mathcal{V}$, si $\mathbf{B}$ es de orden $m$ por $n$, y si $\boldsymbol{c} \in \mathbb{R}^n$; entonces

$$[f_{\mathsf{Z}} \circ f_{\mathbf{B}}](\boldsymbol{c}) \;=\; f_{\mathsf{Z}}\big(f_{\mathbf{B}}(\boldsymbol{c})\big) \;=\; f_{\mathsf{Z}}\big(\mathbf{B}\boldsymbol{c}\big) \;=\; \mathsf{Z}\mathbf{B}\boldsymbol{c} \;=\; f_{\mathsf{Z}\mathbf{B}}(\boldsymbol{c}),$$

es decir,

$$f_{\mathsf{Z}\mathbf{B}}(\boldsymbol{c}) \colon \mathbb{R}^n \xrightarrow{f_{\mathbf{B}}} \mathbb{R}^m \xrightarrow{f_{\mathsf{Z}}} \mathcal{V}.$$

Y, por otra parte, si $\underset{m\times q}{\mathbf{B}}$ y $\underset{q\times n}{\mathbf{C}}$,

$$[f_{\mathsf{Z}} \circ f_{\mathbf{B}}](\mathbf{C}) \;=\; f_{\mathsf{Z}}\big(f_{\mathbf{B}}(\mathbf{C})\big) \;=\; f_{\mathsf{Z}}\big(\mathbf{B}\mathbf{C}\big) \;=\; \mathsf{Z}\mathbf{B}\mathbf{C} \;=\; f_{\mathsf{Z}\mathbf{B}}(\mathbf{C});$$

es decir,

$$f_{\mathsf{Z}\mathbf{B}}(\mathbf{C}) \colon \mathbb{R}^{q\times n} \xrightarrow{f_{\mathbf{B}}} \mathbb{R}^{m\times n} \xrightarrow{f_{\mathsf{Z}}} \mathcal{V}^n,$$

donde $\mathcal{V}^n$ es el conjunto de sistemas de $n$ vectores de $\mathcal{V}$.

Ejercicio 47. Compruebe que el espacio engendrado por un sistema $\mathsf{Z}$ es la imagen de $f_{\mathsf{Z}}$.

Ejercicio 48. Considere $\underset{m\times n}{\mathbf{A}}$, y compruebe que $\mathcal{C}\left(\mathbf{A}\right) = \mathit{imagen}(f_{\mathbf{A}})$, donde $\begin{aligned} f_{\mathbf{A}} : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ \boldsymbol{x} &\longmapsto \mathbf{A}\boldsymbol{x} \end{aligned}$.

**Definición 9.12.** *Llamamos* núcleo *(o* espacio nulo*) de la función lineal* $f : \mathcal{V} \to \mathcal{W}$*, al siguiente conjunto*

$$\mathcal{N}\left(f\right) = \left\{\vec{v} \in \mathcal{V} \mid f(\vec{v}) = \vec{0}\right\}.$$

Ejercicio 49. Demuestre la siguiente

**Proposición 9.6.2.** *El núcleo de una función lineal es un subespacio.*

Fíjese que la Proposición 7.3.1 es un caso particular en el que la función es $f_{\mathsf{A}}:\mathbb{R}^n \longrightarrow \mathbb{R}^m$. $\boldsymbol{x} \longmapsto \mathsf{A}\boldsymbol{x}$

Ejercicio 50. Demuestre la siguiente

**Proposición 9.6.3.** *La imagen de una función lineal es un subespacio.*

Ejercicio 51. Demuestre la siguiente

**Proposición 9.6.4.** *Sean las funciones $f:\mathcal{V}\to\mathcal{W}$ y $g:\mathcal{W}\to\mathcal{Y}$, entonces la imagen de $g\circ f$ está contenida en la imagen de $g$.*

Ejercicio 52. Demuestre el siguiente

**Corolario 9.6.5.** *Sean las funciones $f:\mathcal{V}\to\mathcal{W}$ y $g:\mathcal{W}\to\mathcal{Y}$, y donde $f$ es invertible, entonces $imagen(g\circ f) = imagen(g)$.*

**Sistemas linealmente independientes y coordenadas**

**Proposición 9.6.6.** *Un sistema $\mathsf{Z}$ de $n$ vectores es independiente si y solo si $f_{\mathsf{Z}}$ es invertible.*

*Demostración.* Si $\mathsf{Z}$ es independiente, tenemos que comprobar que el conjunto de pares $(\mathsf{Z}\boldsymbol{x},\boldsymbol{x})$ tales que $\boldsymbol{x}\in\mathbb{R}^n$ es función. Supongamos que $\mathsf{Z}\boldsymbol{x}=\mathsf{Z}\boldsymbol{y}$, entonces $\mathsf{Z}(\boldsymbol{x}-\boldsymbol{y})=\mathbf{0}$. Como $\mathsf{Z}$ es independiente, necesariamente $(\boldsymbol{x}-\boldsymbol{y})=\mathbf{0}$; es decir $\boldsymbol{x}=\boldsymbol{y}$.

Supongamos que el conjunto de pares $(\mathsf{Z}\boldsymbol{x},\boldsymbol{x})$ tales que $\boldsymbol{x}\in\mathbb{R}^n$ es función, es decir, $(f_{\mathsf{Z}})^{-1}$; y que $\mathsf{Z}\boldsymbol{x}=\mathbf{0}$. Como también $\mathsf{Z}\mathbf{0}=\mathbf{0}$, los pares $(\mathsf{Z}\boldsymbol{x},\boldsymbol{x})$ y $(\mathsf{Z}\mathbf{0},\mathbf{0})$ pertenecen a $(f_{\mathsf{Z}})^{-1}$. Como las primeras componentes son nulas, necesariamente $\boldsymbol{x}$ tiene que ser $\mathbf{0}$. □

Cuando $\mathsf{Z}$ es independiente, la función inversa $(f_{\mathsf{Z}})^{-1}:\mathcal{L}(\mathsf{Z})\to\mathbb{R}^n$ se conoce como *función de coordenadas respecto a* $\mathsf{Z}$; y la denotaremos del siguiente modo:

$$\vec{v}_{/_{\mathsf{Z}}} = (f_{\mathsf{Z}})^{-1}(\vec{v}) \qquad \Rightarrow \qquad {}_{-/_{\mathsf{Z}}}:\mathcal{V}\to\mathbb{R}^n;$$

por tanto, $\vec{v}_{/_{\mathsf{Z}}}$ (las coordenadas de $\vec{v}$ respecto de $\mathsf{Z}$) es *el vector de $\mathbb{R}^n$ con los coeficientes de la única combinación lineal* de los vectores del sistema $\mathsf{Z}$ que es igual a $\vec{v}$. Es decir, $\boldsymbol{x}=\vec{v}_{/_{\mathsf{Z}}}$ es el *único vector de $\mathbb{R}^n$* tal que

$$\mathsf{Z}\boldsymbol{x}=\vec{v}; \qquad \text{pues} \qquad \mathsf{Z}\big(\vec{v}_{/_{\mathsf{Z}}}\big)=\vec{v}.$$

Como la función *función de coordenadas respecto a* $\mathsf{Z}$ es lineal, sabemos que

$$\big(\vec{x}+\vec{y}\big)_{/_{\mathsf{Z}}} = \vec{x}_{/_{\mathsf{Z}}} + \vec{y}_{/_{\mathsf{Z}}} \qquad \text{y} \qquad \big(a\vec{x}\big)_{/_{\mathsf{Z}}} = a\big(\vec{x}_{/_{\mathsf{Z}}}\big).$$

Ejercicio 53. Demuestre la siguiente

**Proposición 9.6.7.** *Si $\mathsf{B}$ es una base de $\mathcal{V}$, entonces para cualquier función lineal $f:\mathcal{V}\to\mathcal{W}$ existe un sistema $\mathsf{Z}$ de vectores de $\mathcal{W}$ tal que $f(\vec{v})=\mathsf{Z}\big(\vec{v}_{/_{\mathsf{B}}}\big)$.*

## *Apéndices a la lección*

# Los cuatro subespacios fundamentales de una matriz **A**

## 10.1. Suma de subespacios

Sean $\mathcal{A}$ y $\mathcal{B}$ subespacios de $\mathcal{V}$. Su *suma*, que escribimos como $\mathcal{A}+\mathcal{B}$, es el conjunto de vectores de $\mathcal{V}$ que se pueden escribir como $\vec{a}+\vec{b}$ con $\vec{a}\in\mathcal{A}$ y $\vec{b}\in\mathcal{B}$. Veamos que la suma $\mathcal{A}+\mathcal{B}$ es un subespacio:

**Proposición 10.1.1.** *Si $\mathcal{A}$ y $\mathcal{B}$ son subespacios de $\mathcal{V}$, entonces la suma $\mathcal{A}+\mathcal{B}$ también es un subespacio.*

*Demostración.* $\mathcal{A}+\mathcal{B}$ no es vacío por ser no vacío ni $\mathcal{A}$ ni $\mathcal{B}$ (pues ambos contienen $\vec{0}$). Sean $\vec{a}_1,\vec{a}_2\in\mathcal{A}$ y $\vec{b}_1,\vec{b}_2\in\mathcal{B}$. Si $\vec{p}=\vec{a}_1+\vec{b}_1$ y $\vec{q}=\vec{a}_2+\vec{b}_2$, entonces $\vec{p},\vec{q}\in\mathcal{A}+\mathcal{B}$. Veamos que cualquier combinación lineal también pertenece a $\mathcal{A}+\mathcal{B}$. Sean $x,y\in\mathbb{R}$, entonces

$$x(\vec{p})+y(\vec{q})=x\big(\vec{a}_1+\vec{b}_1\big)+y\big(\vec{a}_2+\vec{b}_2\big)=\big(x(\vec{a}_1)+y(\vec{a}_2)\big)+\big(x(\vec{b}_1)+y(\vec{b}_2)\big)\in\mathcal{A}+\mathcal{B},$$

pues $\big(x(\vec{a}_1)+y(\vec{a}_2)\big)\in\mathcal{A}$ y $\big(x(\vec{b}_1)+y(\vec{b}_2)\big)\in\mathcal{B}$. □

De la demostración anterior se deduce que si $\mathcal{A}=\mathcal{L}\big(A\big)$ y $\mathcal{B}=\mathcal{L}\big(B\big)$,

$$\mathcal{L}\big(A\big)+\mathcal{L}\big(B\big)=\mathcal{L}\big(\mathcal{L}\,(A)\cup\mathcal{L}\,(B)\,\big)=\mathcal{L}\big(A\cup B\big). \tag{10.1}$$

Módulo NAcAL para Python

```
A = Sistema([Vector([1, 0, 1, 0]), Vector([0, -1, 0, -1])])
B = Sistema([Vector([1, 1, 1, 1]), Vector([1,  0, 0,  0])])
SubEspacio(A) + SubEspacio(B) # SubEspacio suma de los SubEspacios L(A) y L(B)
# que es igual al espacio generado por la union:  SubEspacio(A.concatena(B))
```

$$\left\{\boldsymbol{v}\in\mathbb{R}^4\;\middle|\;\exists\boldsymbol{p}\in\mathbb{R}^3,\ \boldsymbol{v}=\begin{bmatrix}1&0&1\\0&-1&0\\1&0&0\\0&-1&0\end{bmatrix}\boldsymbol{p}\right\}=\left\{\boldsymbol{v}\in\mathbb{R}^4\;\middle|\;\begin{bmatrix}0&-1&0&1\end{bmatrix}\boldsymbol{v}=\big(0,\big)\right\}$$

**Proposición 10.1.2.** *Sean $\mathsf{A}=\big[\vec{a}_1;\dots\,\vec{a}_p;\big]$ y $\mathsf{B}=\big[\vec{b}_1;\dots\,\vec{b}_q;\big]$ dos sistemas de vectores de $\mathcal{V}$. Y sea $\mathbf{X}$ un sistema de $s$ vectores de $\mathbb{R}^p$ e $\mathbf{Y}$ un sistema de $s$ vectores de $\mathbb{R}^q$; tales que*

$$\big[\mathsf{A}\,\#\,\mathsf{B}\big]\begin{bmatrix}\mathbf{X}\\\mathbf{Y}\end{bmatrix}=\big[\vec{0};\dots\,\vec{0};\big],\quad \textit{es decir}\quad \mathcal{N}\left(\mathsf{A}\,\#\,\mathsf{B}\right)=\mathcal{C}\left(\begin{bmatrix}\mathbf{X}\\\mathbf{Y}\end{bmatrix}\right),$$

*donde $\mathsf{A}\,\#\,\mathsf{B}$ es el sistema de $p+q$ vectores resultante de concatenar $\mathsf{A}$ y $\mathsf{B}$. Entonces $\mathcal{L}\big(\mathsf{A}\mathbf{X}\big)=\mathcal{L}\big(\mathsf{A}\big)\cap\mathcal{L}\big(\mathsf{B}\big)$.*

*Si además la matriz "apilada" $\begin{bmatrix}\mathbf{X}\\\mathbf{Y}\end{bmatrix}$ es de rango completo por columnas y si $\mathsf{A}$ y $\mathsf{B}$ son linealmente independientes, también lo es el sistema $\mathsf{A}\mathbf{X}$ de vectores de $\mathcal{V}$.*

*Demostración.* Para ver que $\mathcal{L}\big(\mathsf{A}\mathbf{X}\big) = \mathcal{L}\big(\mathsf{A}\big)\cap\mathcal{L}\big(\mathsf{B}\big)$ tenemos que demostrar la pertenencia en ambos sentidos:

1. *Demostremos que* $\mathsf{A}\mathbf{X}_{|j} \in \mathcal{L}\big(\mathsf{A}\big) \cap \mathcal{L}\big(\mathsf{B}\big)$.

   $$\big[\mathsf{A} \# \mathsf{B}\big]\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}_{|j} = \vec{0}; \quad \Rightarrow \quad \mathsf{A}\mathbf{X}_{|j} + \mathsf{B}\mathbf{Y}_{|j} = \vec{0}; \quad \Rightarrow \quad \mathsf{A}\mathbf{X}_{|j} = -\mathsf{B}\mathbf{Y}_{|j}; \quad \Rightarrow \quad \mathsf{A}\mathbf{X}_{|j} \in \mathcal{L}\big(\mathsf{A}\big) \cap \mathcal{L}\big(\mathsf{B}\big),$$

   pues $\mathsf{A}\mathbf{X}_{|j}$ también es una combinación lineal de los vectores de $\mathsf{B}$.

2. *Demostremos que si* $\vec{v} \in \mathcal{L}\big(\mathsf{A}\big) \cap \mathcal{L}\big(\mathsf{B}\big)$ *entonces* $\vec{v} \in \mathcal{L}\big(\mathsf{A}\mathbf{X}\big)$, *es decir,* $\vec{v} = (\mathsf{A}\mathbf{X})\boldsymbol{z}$ *para algún* $\boldsymbol{z} \in \mathbb{R}^s$.
   Sea $\vec{v} \in \mathcal{L}\big(\mathsf{A}\big) \cap \mathcal{L}\big(\mathsf{B}\big)$, entonces existen $\boldsymbol{x}$ e $\boldsymbol{y}$ tales que $\vec{v} = \mathsf{A}\boldsymbol{x}$ y $\vec{v} = -\mathsf{B}\boldsymbol{y}$. Por tanto $\mathsf{A}\boldsymbol{x} + \mathsf{B}\boldsymbol{y} = \vec{0}$; es decir $\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix} \in \mathcal{N}\left(\mathsf{A} \# \mathsf{B}\right) = \mathcal{C}\left(\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}\right)$. Así pues, existe $\boldsymbol{z} \in \mathbb{R}^s$ tal que $\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix} = \begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}\boldsymbol{z}$. Consecuentemente $\boldsymbol{x} = \mathbf{X}\boldsymbol{z}$ y por tanto $\vec{v} = \mathsf{A}\boldsymbol{x} = \mathsf{A}\mathbf{X}\boldsymbol{z}$.

*Demostremos que si además la matriz "apilada"* $\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}$ *es de rango completo por columnas, y si* $\mathsf{A}$ *y* $\mathsf{B}$ *son linealmente independientes, también es linealmente independiente el sistema* $\mathsf{A}\mathbf{X}$ *de vectores de* $\mathcal{V}$.

Supongamos que $(\mathsf{A}\mathbf{X})\boldsymbol{z} = \vec{0}$. Entonces $\mathsf{A}(\mathbf{X}\boldsymbol{z}) = \vec{0}$, y como $\mathsf{A}$ es linealmente independiente, $\mathbf{X}\boldsymbol{z} = \vec{0}$. Por otra parte, como $\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}\boldsymbol{z} \in \mathcal{N}\left(\mathsf{A} \# \mathsf{B}\right)$, tendremos que $\vec{0} = \mathsf{A}(\mathbf{X}\boldsymbol{z}) + \mathsf{B}(\mathbf{Y}\boldsymbol{z}) = \mathsf{B}(\mathbf{Y}\boldsymbol{z})$; y al ser $\mathsf{B}$ linealmente independiente $\mathbf{Y}\boldsymbol{z} = \vec{0}$. Pero entonces tenemos que $\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}\boldsymbol{z} = \begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}$; y como $\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}$ es de rango completo por columnas, necesariamente $\boldsymbol{z} = 0$. □

Consecuentemente, si $\mathsf{A}$ y $\mathsf{B}$ son linealmente independientes y si las columnas de $\begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}$ son una base de $\mathcal{N}\left(\mathsf{A} \# \mathsf{B}\right)$, entonces $\mathsf{A}\mathbf{X} = \big[\mathsf{A}\mathbf{X}_{|1};\ \ldots \mathsf{A}\mathbf{X}_{|s};\big]$ es una base de $\mathcal{L}\big(\mathsf{A}\big) \cap \mathcal{L}\big(\mathsf{B}\big)$.

**Corolario 10.1.3.** *Si* $\mathcal{A}$ *y* $\mathcal{B}$ *son subespacios de* $\mathcal{V}$ *entonces* $\dim(\mathcal{A}+\mathcal{B}) = \dim(\mathcal{A}) + \dim(\mathcal{B}) - \dim(\mathcal{A}\cap\mathcal{B})$.

*Demostración.* (*Esta demostración muestra un método para encontrar una base de* $\mathcal{A}\cap\mathcal{B}$)
Por lo anterior, sabemos que si tomamos una base $\mathsf{A}$ de $\mathcal{A}$ y una base $\mathsf{B}$ de $\mathcal{B}$, las dimensiones de $\mathcal{L}\big(\mathsf{A}\big)\cap\mathcal{L}\big(\mathsf{B}\big)$ y de $\mathcal{N}\left(\mathsf{A} \# \mathsf{B}\right)$ son iguales; i.e., coinciden con el número de "soluciones especiales" del sistema $[\mathsf{A}\#\mathsf{B}]\boldsymbol{z} = \vec{0}$.

Sean $\mathsf{A} = \big[\vec{a}_1;\ldots;\vec{a}_p;\big]$ y $\mathsf{B} = \big[\vec{b}_1;\ldots;\vec{b}_q;\big]$ bases de $\mathcal{A}$ y $\mathcal{B}$, respectivamente. Consideremos el sistema $\mathsf{A} \# \mathsf{B} = \big[\vec{a}_1;\ldots;\vec{a}_p;\vec{b}_1;\ldots;\vec{b}_q;\big]$ resultante de concatenar $\mathsf{A}$ y $\mathsf{B}$. Evidentemente $\mathcal{L}\big(\mathsf{A} \# \mathsf{B}\big) = \mathcal{A} + \mathcal{B}$. Aplicando la eliminación *"de izquierda a derecha"* sobre el sistema $\mathsf{A} \# \mathsf{B}$ (como en la Proposición 9.5.6 o la Ecuación 9.4 en la página 127) observaremos el siguiente esquema:

$$\left[\frac{\big[\mathsf{A} \# \mathsf{B}\big]}{\mathbf{I}}\right]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \left[\frac{\big[\mathsf{A} \# \mathsf{B}\big]\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}{\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}\right] = \left[\frac{\big[\mathsf{A} \# \mathsf{B}\big]\mathbf{E}}{\mathbf{E}}\right]; \quad \text{donde} \quad \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

Con la eliminación sobre $\mathsf{A}\#\mathsf{B}$ habremos anulado todos los vectores que eran combinación lineal de los que les anteceden. Por tanto, los vectores no nulos de $\big[\mathsf{A} \# \mathsf{B}\big]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ constituyen una base de $\mathcal{A}+\mathcal{B}$ (Corolario 9.5.10). Así, si tras la eliminación se anulan $s$ vectores de $\mathsf{A} \# \mathsf{B}$, entonces $\boxed{\dim(\mathcal{A}+\mathcal{B}) = p+q-s}$.

Como complemento podemos construir una base de $\mathcal{A}\cap\mathcal{B}$. Sea $\mathbf{X} =_{(1:p)|}\mathbf{E}$ la submatriz formada por las $p$ primeras filas de $\mathbf{E}$, e $\mathbf{Y}$ la submatriz formada por las $q$ últimas. Entonces, si $\boldsymbol{\nu}$ es la lista con los índices de los $s$ vectores de $\mathsf{A} \# \mathsf{B}$ que se han anulado en el proceso de eliminación, tenemos que

$$\Big(\big[\mathsf{A} \# \mathsf{B}\big]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\Big)_{|\boldsymbol{\nu}} = \big[\mathsf{A} \# \mathsf{B}\big]\mathbf{E}_{|\boldsymbol{\nu}} = [\vec{0};\ldots\vec{0};], \quad \text{con} \quad \mathbf{E}_{|\boldsymbol{\nu}} = \begin{bmatrix}\mathbf{X}\\ \mathbf{Y}\end{bmatrix}_{|\boldsymbol{\nu}} \Rightarrow\ \mathsf{A}\mathbf{X}_{|\boldsymbol{\nu}} + \mathsf{B}\mathbf{Y}_{|\boldsymbol{\nu}} = [\vec{0};\ldots\vec{0};].$$

Así, por la Proposición 10.1.2 sabemos que $\boxed{\dim \mathcal{A} \cap \mathcal{B} = s}$ y que $\mathsf{A}\mathbf{X}_{|\nu}$ es una base de $\mathcal{L}(\mathsf{A}) \cap \mathcal{L}(\mathsf{B})$. □

EJERCICIO 54. Encuentre con NAcAL una base de la intersección de los subespacios generados por:

```
A = Sistema([Vector([1, 0, 1, 0]), Vector([0, -1, 0, -1])])
B = Sistema([Vector([1, 1, 1, 1]), Vector([1,  0, 0,  0])])
```

Y ahora verifiquemos que se cumple que $\dim(\mathcal{A}+\mathcal{B}) = \dim(\mathcal{A}) + \dim(\mathcal{B}) - \dim(\mathcal{A}\cap\mathcal{B})$:

Módulo NAcAL para Python

```
    DimSuma     = (SubEspacio(A) + SubEspacio(B)).dim # dimensión Suma A + B
    DimA        = SubEspacio(A).dim                   # dimensión A
    DimB        = SubEspacio(B).dim                   # dimensión B
    DimIntersec = (SubEspacio(A) & SubEspacio(B)).dim # dimensión Intersecc
    print( DimSuma, '=', DimA, '+', DimB, "-", DimIntersec)
```

```
3 = 2 + 2 - 1
```

### 10.1.1. Suma directa y subespacios suplementarios

**Definición 10.1** (Suma directa)**.** *Si $\mathcal{A}$ y $\mathcal{B}$ son subespacios de $\mathcal{V}$, y su intersección es $\mathcal{A}\cap\mathcal{B} = \vec{0}$; entonces la suma de los subespacios $\mathcal{A}$ y $\mathcal{B}$ se denomina* suma directa *y se denota con $\mathcal{A}\oplus\mathcal{B}$.*

**Definición 10.2** (Espacios suplementarios)**.** *Si $\mathcal{A}$ y $\mathcal{B}$ son subespacios de $\mathcal{V}$, y su suma directa es todo el espacio, $\mathcal{A}\oplus\mathcal{B} = \mathcal{V}$, entonces decimos que $\mathcal{A}$ y $\mathcal{B}$ son* subespacios suplementarios.

*Observación.* Existen muchos espacios suplementarios a uno dado. Observando la Figura 6.1 en la página 94 es fácil darse cuenta de que cualquier función lineal $g\colon \mathcal{V}\to\mathcal{W}$ es un espacio suplementario de $\{\vec{0}\}\times\mathcal{W}$.

Veamos que si $\mathcal{V} = \mathcal{A}\oplus\mathcal{B}$, cualquier vector $\vec{v}\in\mathcal{V}$ se puede descomponer *de manera única* en la suma de dos vectores, uno que pertenece a $\mathcal{A}$ y otro que pertenece a $\mathcal{B}$.

**Proposición 10.1.4.** *Si $\vec{v} = \vec{a} + \vec{b}$ con $\vec{a}\in\mathcal{A}$, $\vec{b}\in\mathcal{B}$ y $\mathcal{A}\cap\mathcal{B} = \{\vec{0}\}$, entonces $\vec{a}$ y $\vec{b}$ son únicos.*

*Demostración.* Imagine que $\vec{v} = \vec{a} + \vec{b}$ y $\vec{v} = \vec{a}_* + \vec{b}_*$, con $\vec{a}, \vec{a}_* \in \mathcal{A}$ y $\vec{b}, \vec{b}_* \in \mathcal{B}$. Entonces

$$\vec{v} - \vec{v} = (\vec{a} + \vec{b}) - (\vec{a}_* + \vec{b}_*) = \underbrace{(\vec{a} - \vec{a}_*)}_{\in\mathcal{A}} - \underbrace{(\vec{b} + \vec{b}_*)}_{\in\mathcal{B}} = \vec{0} \qquad \Longrightarrow \qquad (\vec{a} - \vec{a}_*) = -(\vec{b} + \vec{b}_*) \in \mathcal{B}.$$

Es decir, $(\vec{a} - \vec{a}_*) \in \mathcal{A}\cap\mathcal{B} = \{\vec{0}\}$, y por tanto $(\vec{a} - \vec{a}_*) = \vec{0} = (\vec{b} - \vec{b}_*)$, es decir, $\vec{a} = \vec{a}_*$ y $\vec{b} = \vec{b}_*$. □

EJERCICIO 55. Sea $\mathcal{V} = \mathcal{A}\oplus\mathcal{B}$ y considere $f\colon \mathcal{V} \longrightarrow \mathcal{A}\times\mathcal{B}$, tal que $\vec{v} = \vec{a} + \vec{b}$ con $\vec{a}\in\mathcal{A}$ y $\vec{b}\in\mathcal{B}$.
$\vec{v} \longmapsto (\vec{a}, \vec{b})$

(a) Demuestre que $f$ es lineal.

(b) Considere ahora una función parecida a la anterior, salvo que ahora a cada vector $\vec{v}$ le asocia únicamente la primera de las componentes de los anteriores pares cartesianos; es decir, a cada $\vec{v}$ le asocia la componente $\vec{a}$ correspondiente al subespacio $\mathcal{A}$; por tanto $f_{\mathcal{A}}\colon \mathcal{V} \longrightarrow \mathcal{A}$ de tal manera que $\vec{v} - \vec{a} \in \mathcal{B}$.
$\vec{v} \longmapsto \vec{a}$

Demuestre que $f_{\mathcal{A}}$ es lineal.[1]

(c) (Opcional) Demuestre que $\mathcal{N}(f_{\mathcal{A}}) = \mathcal{B}$.

[1] Esta función $f_{\mathcal{A}}$ es la proyección sobre el subespacio $\mathcal{A}$ de $\mathcal{V}$ paralela a $\mathcal{B}$. En la Lección 13 trataremos del caso particular en el que $\mathcal{A}$ y $\mathcal{B}$ son perpendiculares.

## 10.2. Los cuatro subespacios fundamentales de una matriz $\mathbf{A}$

☞ En esta sección se verá que toda matriz $\mathbf{A}$, de orden $m$ por $n$, define una descomposición de $\mathbb{R}^m$ y de $\mathbb{R}^n$ en subespacios suplementarios. En particular:

- $\mathbb{R}^n$ se descompone en el *espacio nulo* $\mathcal{N}\left(\mathbf{A}\right)$ y en el *espacio fila* $\mathcal{C}\left(\mathbf{A}^\intercal\right)$:

$$\boxed{\mathbb{R}^n = \mathcal{N}\left(\mathbf{A}\right) \oplus \mathcal{C}\left(\mathbf{A}^\intercal\right)}.$$

- $\mathbb{R}^m$ se descompone en el *espacio columna* $\mathcal{C}\left(\mathbf{A}\right)$ y en el *espacio nulo por la izquierda* $\mathcal{N}\left(\mathbf{A}^\intercal\right)$:

$$\boxed{\mathbb{R}^m = \mathcal{C}\left(\mathbf{A}\right) \oplus \mathcal{N}\left(\mathbf{A}^\intercal\right)}.$$

El profesor Strang (2007) se refiere a estos cuatro subespacios: $\mathcal{N}\left(\mathbf{A}\right)$, $\mathcal{C}\left(\mathbf{A}\right)$, $\mathcal{N}\left(\mathbf{A}^\intercal\right)$ y $\mathcal{C}\left(\mathbf{A}^\intercal\right)$, con el nombre de *"los cuatro subespacios fundamentales de la matriz* $\mathbf{A}$*"*.

La notación indica que los nuevos espacios $\mathcal{C}\left(\mathbf{A}^\intercal\right)$ y $\mathcal{N}\left(\mathbf{A}^\intercal\right)$ son el espacio columna y el nulo de $\mathbf{A}^\intercal$.

### 10.2.1. El espacio fila

El espacio fila de $\underset{m\times n}{\mathbf{A}}$ es el conjunto de las combinaciones lineales de las filas:

$$\mathcal{L}\big(\text{filas de } \mathbf{A}\big) \;=\; \mathcal{L}\big(\text{columnas de } \mathbf{A}^\intercal\big) \;=\; \mathcal{C}\left(\mathbf{A}^\intercal\right),$$

que es un subespacio de $\mathbb{R}^n$, dado que cada fila tiene $n$ componentes (cada uno situado en una columna diferente).

Como las filas de $\mathbf{A}$ son las columnas de $\mathbf{A}^\intercal$, tenemos que

$$\begin{aligned}\mathcal{C}\left(\mathbf{A}^\intercal\right) =& \{\boldsymbol{x} \mid \text{existe } \boldsymbol{y} \in \mathbb{R}^m : \boldsymbol{x} = (\mathbf{A}^\intercal)\boldsymbol{y}\}\\ =& \{\boldsymbol{x} \mid \text{existe } \boldsymbol{y} \in \mathbb{R}^m : \boldsymbol{x} = \boldsymbol{y}\mathbf{A}\} \qquad\qquad \text{ya que } (\mathbf{A}^\intercal)\boldsymbol{y} = \boldsymbol{y}\mathbf{A};\end{aligned}$$

es decir, *el espacio fila* $\mathcal{C}\left(\mathbf{A}^\intercal\right)$ está formado por los vectores que se pueden describir como combinaciones lineales de las filas de $\mathbf{A}$. Como el rango de una matriz no cambia al transponerla (Proposición 5.2.5), se cumple que

$$\dim \mathcal{C}\left(\mathbf{A}\right) \;=\; \operatorname{rg}\left(\mathbf{A}\right) \;=\; \operatorname{rg}\left(\mathbf{A}^\intercal\right) \;=\; \dim \mathcal{C}\left(\mathbf{A}^\intercal\right).$$

Es decir, para una matriz $\mathbf{A}$ de rango $r$, tenemos que $\dim \mathcal{C}\left(\mathbf{A}^\intercal\right) = r$.

Observe que si $\mathbf{A}$ es de orden $m\times n$, dado que $\dim \mathcal{N}\left(\mathbf{A}\right) = n-r$, se tiene que $\dim \mathcal{C}\left(\mathbf{A}^\intercal\right)+\dim \mathcal{N}\left(\mathbf{A}\right) = n$, que es igual a la dimensión de todo el espacio $\mathbb{R}^n$. Así, si probamos que $\mathcal{N}\left(\mathbf{A}\right) \cap \mathcal{C}\left(\mathbf{A}^\intercal\right) = \{\mathbf{0}\}$ (es decir, que la dimensión de la intersección es cero), por la Proposición 10.1.4 habremos demostrado que $\mathbb{R}^n = \mathcal{C}\left(\mathbf{A}^\intercal\right) \oplus \mathcal{N}\left(\mathbf{A}\right)$:

**Proposición 10.2.1.** *Para toda* $\mathbf{A}$ *se verifica que* $\mathcal{N}\left(\mathbf{A}\right) \cap \mathcal{C}\left(\mathbf{A}^\intercal\right) = \{\mathbf{0}\}$.

*Demostración.* Para todo vector $\boldsymbol{y} \in \mathcal{C}\left(\mathbf{A}^\intercal\right)$ (es decir, de la forma $\boldsymbol{y} = \boldsymbol{z}\mathbf{A}$) y para todo vector $\boldsymbol{x} \in \mathcal{N}\left(\mathbf{A}\right)$ (es decir, tal que $\mathbf{A}\boldsymbol{x} = \mathbf{0}$) el producto punto $\boldsymbol{y}\cdot\boldsymbol{x}$ es nulo:

$$\boldsymbol{y}\cdot\boldsymbol{x} = \boldsymbol{z}\mathbf{A}\boldsymbol{y} = \boldsymbol{z}\cdot\mathbf{0} = 0.$$

Por tanto, si $\boldsymbol{v}$ pertenece simultáneamente a los dos subespacios, entonces

$$\boldsymbol{v}\cdot\boldsymbol{v} = \sum v_i^2 = 0 \;\Rightarrow\; v_i = 0;$$

es decir, si $\boldsymbol{v} \in \mathcal{N}\left(\mathbf{A}\right) \cap \mathcal{C}\left(\mathbf{A}^\intercal\right)$, entonces necesariamente $\boldsymbol{v} = \mathbf{0}$. □

Puesto que $\mathcal{N}\left(\mathbf{A}\right) \cap \mathcal{C}\left(\mathbf{A}^{\intercal}\right) = \{\mathbf{0}\}$, por la Proposición 10.1.4 concluimos que $\boxed{\mathbb{R}^n = \mathcal{C}\left(\mathbf{A}^{\intercal}\right) \oplus \mathcal{N}\left(\mathbf{A}\right)}$.

### 10.2.2. El espacio nulo por la izquierda

Observe que $\mathcal{N}\left(\mathbf{A}^{\intercal}\right)$ es el conjunto de soluciones del sistema $\left(\mathbf{A}^{\intercal}\right)\boldsymbol{y} = \mathbf{0}$, que nos pregunta: ¿qué combinaciones lineales de las columnas de $\mathbf{A}^{\intercal}$ son iguales a cero?, es decir, ¿qué combinaciones lineales de las *filas* de $\mathbf{A}$ son iguales a cero? Por tanto, $\mathcal{N}\left(\mathbf{A}^{\intercal}\right)$ son las soluciones del siguiente *sistema de ecuaciones homogéneo*

$$\boldsymbol{y}\mathbf{A} = \mathbf{0},$$

lo que justifica el nombre de *espacio nulo por la izquierda* (pues el vector de incógnitas $\boldsymbol{y}$ multiplica por la izquierda). Así, definimos el *espacio nulo por la izquierda* de $\mathbf{A}$ (de orden $m$ por $n$) como

$$\mathcal{N}\left(\mathbf{A}^{\intercal}\right) \;=\; \{\boldsymbol{y} \mid \boldsymbol{y}\mathbf{A} = \mathbf{0}\}.$$

Si llamamos $\mathbf{B}$ a $\mathbf{A}^{\intercal}$, aplicando el resultado de la sección anterior tenemos que $\mathbb{R}^m = \mathcal{C}\left(\mathbf{B}^{\intercal}\right) \oplus \mathcal{N}\left(\mathbf{B}\right)$.

Es decir, $\boxed{\mathbb{R}^m = \mathcal{C}\left(\mathbf{A}\right) \oplus \mathcal{N}\left(\mathbf{A}^{\intercal}\right)}$.

### 10.2.3. El gran esquema

Así, dada $\underset{m\times n}{\mathbf{A}}$ de rango $r$, los espacios $\mathbb{R}^m$ y $\mathbb{R}^n$ quedan descompuestos según el siguiente esquema:

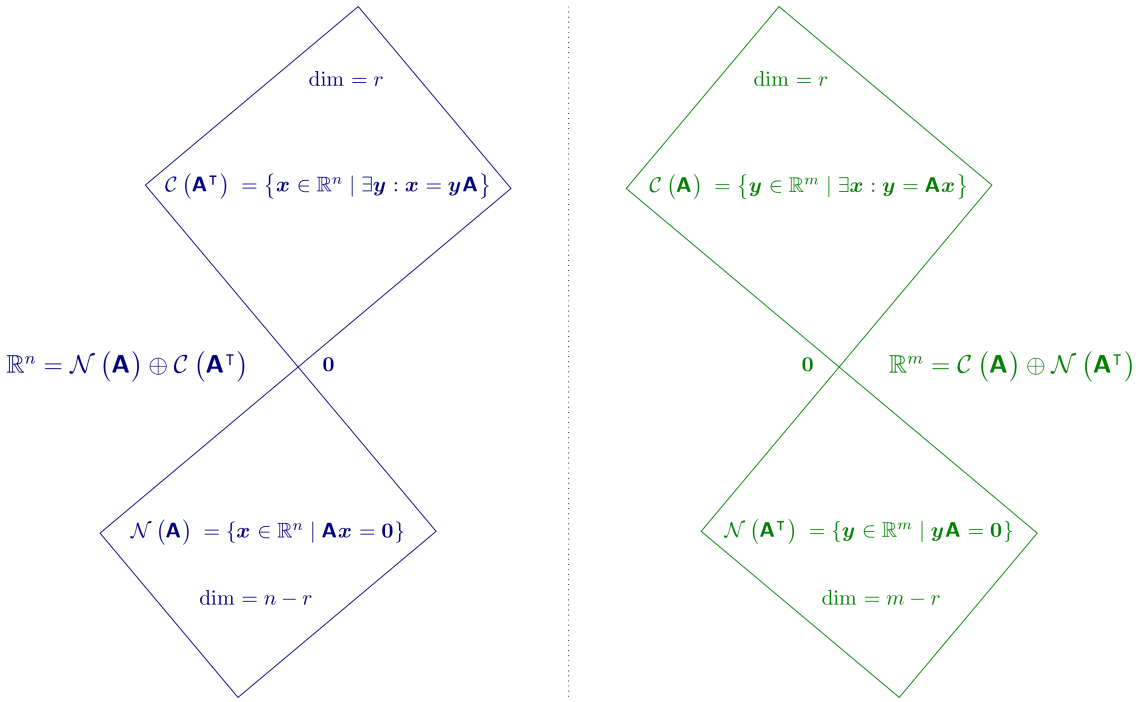



Figura 10.1: Esquema de descomposición de los cuatro subespacios fundamentales de una matriz (parte 1)

*Ejemplo* 25. Para $\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & -1 & -1 & 0 \end{bmatrix}$ de rango 1, comprobemos el Corolario 10.1.3, es decir, que

$$3 = \dim \mathbb{R}^m = \dim \mathcal{N}\left(\mathbf{A}^\intercal\right) + \dim \mathcal{C}\left(\mathbf{A}\right) - \dim\left(\mathcal{N}\left(\mathbf{A}^\intercal\right) \cap \mathcal{C}\left(\mathbf{A}\right)\right) = 2 + 1 - 0$$
$$4 = \dim \mathbb{R}^n = \dim \mathcal{N}\left(\mathbf{A}\right) + \dim \mathcal{C}\left(\mathbf{A}^\intercal\right) - \dim\left(\mathcal{N}\left(\mathbf{A}\right) \cap \mathcal{C}\left(\mathbf{A}^\intercal\right)\right) = 3 + 1 - 0$$

Módulo NAcAL para Python

```
A = Matrix([ [0,1,1,0], [0,2,2,0],[0,-1,-1,0] ])
print( A.rango() )              # Rango de A
```

```
1
```

Módulo NAcAL para Python

```
ColA = SubEspacio(A.sis()); # Espacio Columna de A (en R3)
print(ColA.dim);  ColA
```

```
1
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^1,\ \boldsymbol{v} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \boldsymbol{v} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$

Módulo NAcAL para Python

```
NulAT = SubEspacio( ~A );    # Espacio Nulo Izda de A  (en R3)
print(NulAT.dim);  NulAT
```

```
2
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^2,\ \boldsymbol{v} = \begin{bmatrix} -2 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \begin{bmatrix} -1 & -2 & 1 \end{bmatrix} \boldsymbol{v} = \begin{pmatrix} 0, \end{pmatrix} \right\}$$

Módulo NAcAL para Python

```
FilA = SubEspacio( (~A).sis() ); # Espacio Fila de A (en R4)
print(FilA.dim);  FilA
```

```
1
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^1,\ \boldsymbol{v} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

Módulo NAcAL para Python

```
NulA = SubEspacio( A );      # Espacio Nulo de A (en R4)
print(NulA.dim);  NulA
```

```
3
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^3,\ \boldsymbol{v} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix} \boldsymbol{v} = \begin{pmatrix} 0, \end{pmatrix} \right\}$$

Módulo NAcAL para Python

```
print( (ColA + NulAT).dim ) # Dim. del SubEspacio suma (suma es todo R3)
print( (ColA & NulAT).dim ) # Dim. de la intersección (que solo contiene el cero)
```

```
3
0
```

(son espacios suplementarios)

Módulo NAcAL para Python

```
print( (FilA + NulA).dim ) # Dim. del SubEspacio suma (suma es todo R4)
print( (FilA & NulA).dim ) # Dim. de la intersección (que solo contiene el cero)
```

```
4
0
```

(son espacios suplementarios)

## 10.3. Encontrando bases para los espacios $\mathcal{C}\left(\mathbf{A}^\intercal\right)$ y $\mathcal{N}\left(\mathbf{A}^\intercal\right)$

Sabemos que aplicando la eliminación *sobre las columnas* de $\mathbf{A}$ podemos encontrar bases para los espacios *columna* $\mathcal{C}\left(\mathbf{A}\right)$ y *nulo* $\mathcal{N}\left(\mathbf{A}\right)$. Por tanto, sabemos que aplicando la eliminación *sobre las columnas* de $\mathbf{A}^\intercal$ encontraremos bases para los nuevos subespacios $\mathcal{C}\left(\mathbf{A}^\intercal\right)$ y $\mathcal{N}\left(\mathbf{A}^\intercal\right)$. ¡Pero seguir esta estrategia supondría que, para encontrar bases de los cuatro espacios, necesitamos aplicar el método de eliminación dos veces! Una para $\mathbf{A}$ y otra para $\mathbf{A}^\intercal$.

¿Es posible encontrar bases para cada uno de los cuatro subespacios aplicando únicamente la eliminación sobre las *columnas* de $\mathbf{A}$?... ¡Afortunadamente, sí!

La siguiente proposición nos proporciona la pista necesaria para hacerlo.

**Proposición 10.3.1.** *Sean* $\mathbf{A}$ *de orden* $m$ *por* $n$ *y* $\mathbf{E}$ *invertible de orden* $n$*, entonces* $\mathcal{N}\left(\mathbf{A}^\intercal\right) = \mathcal{N}\left((\mathbf{AE})^\intercal\right)$*, es decir, dado* $\boldsymbol{y} \in \mathbb{R}^m$ *se verifica que*

$$\boldsymbol{y}\mathbf{A} = \mathbf{0} \quad \Longleftrightarrow \quad \boldsymbol{y}\mathbf{AE} = \mathbf{0}.$$

*Demostración.* Veamos las implicaciones en uno y otro sentido.

- Si $\boldsymbol{y} \in \mathcal{N}\left(\mathbf{A}^\intercal\right) \implies \boldsymbol{y}\mathbf{A} = \mathbf{0} \implies \boldsymbol{y}\mathbf{AE} = \mathbf{0E} \implies \boldsymbol{y}\big(\mathbf{AE}\big) = \mathbf{0} \implies \boldsymbol{y} \in \mathcal{N}\left((\mathbf{AE})^\intercal\right)$
- Si $\boldsymbol{y} \in \mathcal{N}\left((\mathbf{AE})^\intercal\right) \implies \boldsymbol{y}\big(\mathbf{AE}\big) = \mathbf{0} \implies \boldsymbol{y}\mathbf{AE}\mathbf{E}^{-1} = \mathbf{0}\mathbf{E}^{-1} \implies \boldsymbol{y}\mathbf{A} = \mathbf{0} \implies \boldsymbol{y} \in \mathcal{N}\left(\mathbf{A}^\intercal\right)$.

□

Consecuentemente, si $\mathbf{A}$ es de orden $m$ por $n$ y $\mathbf{E}$ es invertible y de orden $n$, entonces

- un sistema de vectores formado por varias filas $\left[{}_{i_1|}\mathbf{A};\ \ldots\ {}_{i_r|}\mathbf{A};\right]$ es linealmente independiente si y sólo si el sistema de vectores $\left[{}_{i_1|}\mathbf{AE};\ \ldots\ {}_{i_r|}\mathbf{AE};\right]$ es linealmente independiente.
- la dimensión del espacio fila de $\mathbf{A}$ es igual a la dimensión del espacio fila de $\mathbf{AE}$.

Así, podemos aplicar transformaciones elementales a las columnas de una matriz para obtener otra más simple donde verificar qué filas son combinación lineal de las filas que la anteceden y cuáles forman un sistema linealmente independiente.

El siguiente lema indica que las filas con pivote de cualquier forma pre-escalonada son linealmente independientes.

Ejercicio 56. Demuestre que

**Lema 10.3.2.** *Si* $\mathbf{K}$ *está pre-escalonada, las filas con pivote son linealmente independientes. Además, cada fila sin pivote es combinación lineal de las filas con pivote que la preceden.*

Por tanto, la selección de las filas con pivote de una forma pre-escalonada $\mathbf{K}$ de la matriz $\mathbf{A}$ forma una base del espacio fila de la matriz $\mathbf{K}$ (que generalmente es distinto del espacio fila de $\mathbf{A}$). Veamos que la misma selección aplicada a las filas de $\mathbf{A}$ forma una base del espacio fila de $\mathbf{A}$.

**Encontrando una base para el espacio fila de A**

Sea $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ tal que $\mathbf{AE}$ es pre-escalonada; y sea $\boldsymbol{\varphi} = (\varphi_1, \dots \varphi_r,)$ la lista de índices de las filas con pivote. Entonces la matriz ${}_{\boldsymbol{\varphi}|}\mathbf{I}$ selecciona las filas pivote de $\mathbf{AE}$ que forman una base del espacio fila de $\mathbf{AE}$:

$$\big({}_{(1,2,4,)|}\mathbf{I}\big)(\mathbf{AE}) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & -4 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -6 & -7 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & -4 & 0 & 0 \\ -1 & -1 & 1 & 0 \end{bmatrix} = {}_{(1,2,4,)|}(\mathbf{AE}).$$

Como ${}_{\boldsymbol{\varphi}|}\mathbf{I}$ selecciona una base del espacio fila de $\mathbf{AE}$, entonces para cualquier vector $\boldsymbol{u}$ existe un único vector $\boldsymbol{x}$ (una única combinación lineal de las filas con pivote) tal que

$$\boldsymbol{x}\big({}_{\boldsymbol{\varphi}|}\mathbf{AE}\big) = \boldsymbol{u}\mathbf{AE}.$$

Pero entonces también la ecuación $\boldsymbol{x}\big({}_{\boldsymbol{\varphi}|}\mathbf{A}\big) = \boldsymbol{u}\mathbf{A}$ tiene una única solución ya que, al ser $\mathbf{E}$ invertible,

$$\boldsymbol{x}\big({}_{\boldsymbol{\varphi}|}\mathbf{A}\big) = \boldsymbol{u}\mathbf{A} \iff \boldsymbol{x}\big({}_{\boldsymbol{\varphi}|}\mathbf{AE}\big) = \boldsymbol{u}\mathbf{AE}.$$

Luego ${}_{\boldsymbol{\varphi}|}\mathbf{A}$ (la misma selección, pero de las filas de $\mathbf{A}$) selecciona una base del espacio de fila de $\mathbf{A}$.

Consecuentemente, la dimensión del espacio fila de una matriz coincide con la dimensión del espacio fila de cualquiera de sus formas pre-escalonadas, que a su vez coincide con el rango.

☞ **Algoritmo para encontrar una base de** $\mathcal{C}\left(\mathbf{A}^{\mathsf{T}}\right)$

- Pre-escalonamos $\mathbf{A}$
- Identificamos los índices $\boldsymbol{\varphi}$ de las filas con pivote de la forma escalonada $\mathbf{K}$
- Las filas de ${}_{\boldsymbol{\varphi}|}\mathbf{A}$ forman una base $\mathcal{C}\left(\mathbf{A}^{\mathsf{T}}\right)$

*Ejemplo* 26.

$$\left[\begin{array}{c}\mathbf{A}\\\hline\mathbf{I}\end{array}\right] = \left[\begin{array}{cccc} 1 & -2 & -1 & -1 \\ -2 & 0 & 1 & 2 \\ 4 & -4 & -3 & -4 \\ -1 & 1 & 1 & 1 \\ -6 & 5 & 5 & 6 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{[(1)\mathbf{1}+\mathbf{4}]\\[(1)\mathbf{1}+\mathbf{3}]\\[(2)\mathbf{1}+\mathbf{2}]}]{\boldsymbol{\tau}} \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ -2 & -4 & -1 & 0 \\ 4 & 4 & 1 & 0 \\ -1 & -1 & 0 & 0 \\ -6 & -7 & -1 & 0 \\ \hline 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[[(-1)\mathbf{2}+\mathbf{3}]]{\boldsymbol{\tau}[(4)\mathbf{3}]} \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ -2 & -4 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -6 & -7 & 3 & 0 \\ \hline 1 & 2 & 2 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{c}\mathbf{K}\\\hline\mathbf{E}\end{array}\right]$$

De este proceso de eliminación, deducimos que:

$$\mathcal{C}\left(\mathbf{A}\right) = \mathcal{L}\Big(\big[\mathbf{K}_{|1};\, \mathbf{K}_{|2};\, \mathbf{K}_{|3};\big]\Big), \qquad \mathcal{C}\left(\mathbf{A}^{\mathsf{T}}\right) = \mathcal{L}\Big(\big[{}_{1|}\mathbf{A};\, {}_{2|}\mathbf{A};\, {}_{4|}\mathbf{A};\big]\Big) \quad \text{y} \quad \mathcal{N}\left(\mathbf{A}\right) = \mathcal{L}\Big(\big[\mathbf{E}_{|4};\big]\Big).$$

**Encontrando una base para el espacio nulo por la izquierda.**

Nos falta un método para encontrar una base de $\mathcal{N}\left(\mathbf{A}^{\intercal}\right)$.

Por la Proposición 10.3.1 sabemos que $\mathcal{N}\left(\mathbf{A}^{\intercal}\right) = \mathcal{N}\left((\mathbf{AE})^{\intercal}\right)$. Así que la estrategia es simplificar al máximo alguna forma pre-escalonada de $\mathbf{A}$, y observar ahí qué combinaciones de las filas son nulas. Por tanto, aplicaremos la eliminación Gauss-Jordan hasta llegar a una forma escalonada *reducida* $\mathbf{R}$ (donde los pivotes son unos con ceros a la derecha e izquierda). Continuando el ejemplo anterior:

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & -1 & -1 \\ -2 & 0 & 1 & 2 \\ 4 & -4 & -3 & -4 \\ -1 & 1 & 1 & 1 \\ -6 & 5 & 5 & 6 \end{bmatrix} \xrightarrow[\substack{[(4)\mathbf{3}] \\ [(-1)\mathbf{2}+\mathbf{3}]}]{\substack{\boldsymbol{\tau} \\ [(2)\mathbf{1}+\mathbf{2}] \\ [(1)\mathbf{1}+\mathbf{3}] \\ [(1)\mathbf{1}+\mathbf{4}]}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & -4 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -6 & -7 & 3 & 0 \end{bmatrix} \xrightarrow[[(-1)\mathbf{2}+\mathbf{1}]]{\boldsymbol{\tau}\ [(2)\mathbf{1}]} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -5 & -7 & 3 & 0 \end{bmatrix}$$

$$\xrightarrow[[(1)\mathbf{3}+\mathbf{2}]]{\boldsymbol{\tau}\ [(1)\mathbf{3}+\mathbf{1}]} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -2 & -4 & 3 & 0 \end{bmatrix} \xrightarrow[[(-\frac{1}{4})\mathbf{2}]]{\boldsymbol{\tau}\ [(\frac{1}{2})\mathbf{1}]} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 1 & 3 & 0 \end{bmatrix} = \mathbf{R}.$$

Ahora es inmediato ver en $\mathbf{R}$ que cada fila sin pivote es combinación de las filas con pivote que la anteceden:

$$\begin{cases} {}_{3|}\mathbf{R} = 2\left({}_{1|}\mathbf{R}\right) - 1\left({}_{2|}\mathbf{R}\right) & \Rightarrow \quad (-2,\ 1,\ 1,\ 0,\ 0,)\mathbf{R} = \mathbf{0} \\ \quad \text{y} & \\ {}_{5|}\mathbf{R} = -1\left({}_{1|}\mathbf{R}\right) + 1\left({}_{2|}\mathbf{R}\right) + 3\left({}_{3|}\mathbf{R}\right) & \Rightarrow \quad (1,\ -1,\ 0,\ -3,\ 1,)\mathbf{R} = \mathbf{0} \end{cases}$$

Puesto que $\mathcal{N}\left(\mathbf{A}^{\intercal}\right) = \mathcal{N}\left(\mathbf{R}^{\intercal}\right)$, las mismas relaciones se verifican en la matriz original $\mathbf{A}$;

$$\begin{cases} {}_{3|}\mathbf{A} = 2\left({}_{1|}\mathbf{A}\right) - 1\left({}_{2|}\mathbf{A}\right) & \Rightarrow \quad (-2,\ 1,\ 1,\ 0,\ 0,)\mathbf{A} = \mathbf{0} \\ \quad \text{y} & \\ {}_{5|}\mathbf{A} = -1\left({}_{1|}\mathbf{A}\right) + 1\left({}_{2|}\mathbf{A}\right) + 3\left({}_{3|}\mathbf{A}\right) & \Rightarrow \quad (1,\ -1,\ 0,\ -3,\ 1,)\mathbf{A} = \mathbf{0} \end{cases}$$

☞ En el proceso de eliminación $\mathbf{A} \to \mathbf{R}$ "se ven" bases para los cuatro subespacios fundamentales de $\mathbf{A}$

## 10.4. Ejercicios

Ejercicio 57. Demuestre que para toda $\underset{m\times n}{\mathbf{A}}$

(a) $\mathcal{C}\left(\mathbf{A}(\mathbf{A}^{\intercal})\right) = \mathcal{C}\left(\mathbf{A}\right)$.
(b) $\mathcal{C}\left((\mathbf{A}^{\intercal})\mathbf{A}\right) = \mathcal{C}\left(\mathbf{A}^{\intercal}\right)$.
(c) El rango de $\mathbf{A}(\mathbf{A}^{\intercal})$ es igual que el rango de $\mathbf{A}$.
(d) El rango de $(\mathbf{A}^{\intercal})\mathbf{A}$ es igual que el rango de $\mathbf{A}$.

(Puede encontrar una demostración alternativa para el último apartado en el ejercicio 78 en la página 170).

# Parte IV

# Ortogonalidad

# Vectores ortogonales. Subespacios ortogonales

☞ Una importante aplicación de las matemáticas es la medición de distancias y ángulos. Los productos escalares definidos en espacios vectoriales crean el entorno que permite dichas mediciones.

Comenzamos con los entornos más familiares: $\mathbb{R}^2$ y $\mathbb{R}^3$. Luego generalizaremos a $\mathbb{R}^n$ con el producto punto. (*En la Sección 11.3 generalizamos más aún con productos escalares en espacios vectoriales genéricos*).

## 11.1. Geometría

### 11.1.1. Longitud de un vector en $\mathbb{R}^2$ y en $\mathbb{R}^3$

El cálculo de la longitud de un vector $\boldsymbol{x}$ en $\mathbb{R}^2$ o de un vector $\boldsymbol{y}$ en $\mathbb{R}^3$ se basa en el teorema de Pitágoras. En ambos casos, el vector se corresponde con la hipotenusa de un triángulo rectángulo, de manera que el cuadrado de la longitud del vector es igual a la suma del cuadrado de las longitudes de los catetos. Veámoslo.

En el caso de $\mathbb{R}^2$, la longitud de cada cateto es el valor absoluto de cada componente, $|x_i|$. Por tanto,

$$\|\boldsymbol{x}\|^2 = |x_1|^2 + |x_2|^2 = x_1^2 + x_2^2 = \boldsymbol{x}\cdot\boldsymbol{x} \qquad \Rightarrow \qquad \|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2} = \sqrt{\boldsymbol{x}\cdot\boldsymbol{x}};$$

donde $\|\boldsymbol{x}\|$ denota la longitud o *norma* de un vector $\boldsymbol{x}$ y donde $|x_i|^2 = x_i^2$.

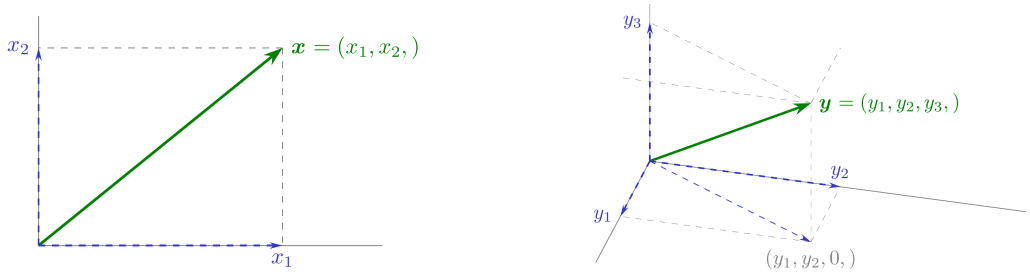



Figura 11.1: Longitud de un vector en $\mathbb{R}^2$ y en $\mathbb{R}^3$

En $\mathbb{R}^3$, la longitud del cateto vertical es el valor absoluto de la tercera componente; sin embargo, la longitud del cateto horizontal ha de ser calculada. Como el cateto horizontal es la hipotenusa de un triángulo rectángulo (como se observa en la figura del caso en $\mathbb{R}^2$), tenemos que el cuadrado de la longitud de dicho cateto es $(y_1^2 + y_2^2)$:

$$\|\boldsymbol{y}\|^2 = (y_1^2 + y_2^2) + y_3^2 = \boldsymbol{y}\cdot\boldsymbol{y} \qquad \Rightarrow \qquad \|\boldsymbol{y}\| = \sqrt{y_1^2 + y_2^2 + y_3^2} = \sqrt{\boldsymbol{y}\cdot\boldsymbol{y}}.$$

Así, la longitud de un vector (en $\mathbb{R}^2$ o $\mathbb{R}^3$) es la raíz cuadrada de la suma de los cuadrados de sus componentes.

Decimos que un vector es *unitario* cuando su longitud es 1. Por tanto, al dividir un vector $\boldsymbol{a}$ (no nulo) por su longitud obtendremos un vector unitario: $\left\| \frac{1}{\|\boldsymbol{a}\|}\boldsymbol{a} \right\| = 1$. La operación de dividir un vector no nulo por su longitud se denomina *normalizar* el vector.

### 11.1.2. Ángulo formado por dos vectores tanto en $\mathbb{R}^2$ como en $\mathbb{R}^3$

Un modo de indicar el grado de abertura del ángulo formado entre dos segmentos unidos por un vértice es mediante la función *coseno* (que toma valores entre 1 y $-1$). Cuando la abertura es nula (cuando el ángulo es cero), el coseno toma el valor 1 (marcado con un rombo sobre el eje horizontal de la figura de la izquierda). El coseno disminuye a medida que el ángulo se abre... cuando el ángulo es recto, su coseno toma el valor 0 (figura central); y conforme la abertura aumenta, el coseno continúa disminuyendo, hasta que se alcanza la abertura máxima, en cuyo caso el coseno alcanza su mínimo valor, es decir, $-1$ (figura de la derecha).

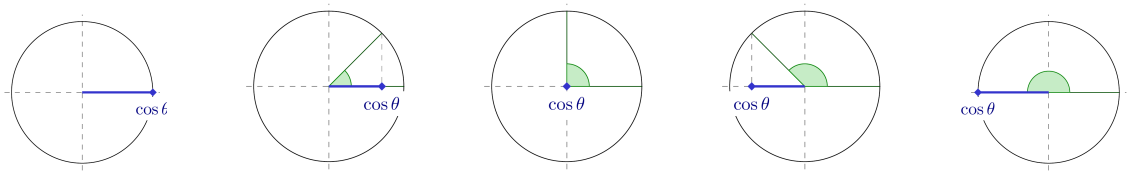



Figura 11.2: Relación entre el ángulo $\theta$ y su coseno (el círculo exterior es de radio uno)

Ahora, piense que tiene tres listones de madera y que ninguno de ellos tiene longitud mayor que la suma de los otros dos. Entonces, con ellos se puede formar un triángulo cuyos ángulos están unívocamente determinados por las longitudes de los listones de madera. Es decir, conociendo los lados de un triángulo, podemos calcular sus ángulos. Para ello echaremos mano, una vez más, del teorema de Pitágoras.

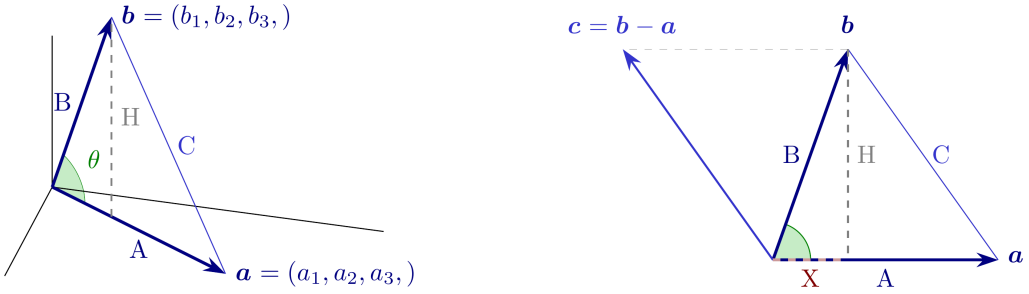



Figura 11.3: Triángulo formado por $\mathbf{0}$ y dos vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ no nulos.

A la izquierda de la Figura 11.3 podemos ver un triángulo formado por los segmentos A, B y C (donde asociamos el segmento A con el vector $\boldsymbol{a}$ y el segmento B con el vector $\boldsymbol{b}$). Para poder calcular el ángulo formado por los segmentos A y B (que es el mismo que el formado por los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$), dividimos el triángulo en dos triángulos rectángulos mediante el segmento vertical H.

A la derecha de la Figura 11.3 aparece esquemáticamente representado el mismo triángulo; pero mostrando que la longitud del segmento C es igual a la longitud del vector $\boldsymbol{b} - \boldsymbol{a}$. Además, también se muestra el cateto X de uno de los triángulos rectángulos resultantes de dividir el triángulo ABC con el segmento vertical H.

*Asumamos inicialmente que los segmentos* A *y* B *son de longitud* 1; entonces el coseno del ángulo $\theta$ que

forman los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ es la longitud $x$ del cateto X (véanse los tres primeros casos de la Figura 11.2):[1]

$$\cos\theta = x. \tag{11.1}$$

Así, si $h$ es la longitud del cateto H, por el T$^{ma}$. de Pitágoras:

$$x^2 + h^2 = 1^2. \tag{11.2}$$

Además, aplicando el Teorema de Pitágoras también en el segundo triángulo rectángulo, deducimos que

$$\begin{aligned} (1-x)^2 + h^2 &= c^2 \\ 1 - 2x + x^2 + h^2 &= c^2 \qquad \text{desarrollando el cuadrado.} \end{aligned} \tag{11.3}$$

(donde $c$ es la longitud del segmento C). Restando (11.2) de (11.3) tenemos $2-2x = c^2$, es decir, $x = (2-c^2)/2$. Así, $\cos\theta = x = \frac{2-c^2}{2}$. Y como

$$\begin{aligned} c^2 =& \|\boldsymbol{a}-\boldsymbol{b}\|^2 \;=\; (\boldsymbol{a}-\boldsymbol{b})\cdot(\boldsymbol{a}-\boldsymbol{b}) \\ =& \boldsymbol{a}\cdot\boldsymbol{a} - 2\boldsymbol{a}\cdot\boldsymbol{b} + \boldsymbol{b}\cdot\boldsymbol{b} \\ =& 2 - 2\boldsymbol{a}\cdot\boldsymbol{b} \qquad\qquad \text{pues } \|\boldsymbol{a}\| = \|\boldsymbol{b}\| = 1, \end{aligned}$$

tenemos que cuando $\boldsymbol{a}$ y $\boldsymbol{b}$ son vectores de norma 1, el coseno del ángulo que forman es su producto punto:

$$\cos\theta = \boldsymbol{a}\cdot\boldsymbol{b} \quad \text{cuando } \|\boldsymbol{a}\| = \|\boldsymbol{b}\| = 1.$$

*Ahora asumamos el caso general en el que los segmentos* A *y* B *son de longitud mayor que cero*, entonces $\frac{1}{\|\boldsymbol{a}\|}\boldsymbol{a}$ y $\frac{1}{\|\boldsymbol{b}\|}\boldsymbol{b}$ son vectores unitarios y, en virtud de lo anterior:

$$\cos\theta \;=\; \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|\,\|\boldsymbol{b}\|}; \qquad (\text{con } \boldsymbol{a}\neq\boldsymbol{0}\neq\boldsymbol{b}).$$

Dado que cuando $\cos\theta = 0$ los vectores son perpendiculares (véase la Figura 11.2), diremos que $\boldsymbol{a}$ y $\boldsymbol{b}$ son perpendiculares cuando $\boldsymbol{a}\cdot\boldsymbol{b} = 0$.

### 11.1.3. Generalizando al espacio euclídeo $\mathbb{R}^n$

☞ Comenzamos la generalización con el concepto de perpendicularidad en el espacio euclídeo[a] $\mathbb{R}^n$.

[a] Para una definición de *espacio euclídeo* véase la Sección 11.3.4.

**Definición 11.1.** *Decimos que* $\boldsymbol{a}$ *y* $\boldsymbol{b}$ *son ortogonales* o perpendiculares ($\boldsymbol{a}\perp\boldsymbol{b}$) *cuando* $\boldsymbol{a}\cdot\boldsymbol{b} = 0$.

Consecuentemente, $\boldsymbol{0}$ es ortogonal a todos los vectores (incluido él mismo).

☞ La generalización del concepto de longitud a vectores del espacio vectorial $\mathbb{R}^n$ es inmediata.

**Definición 11.2.** *La longitud* (o norma) *de un vector* $\boldsymbol{a}$ *es la raíz cuadrada de* $\boldsymbol{a}\cdot\boldsymbol{a}$*:*

$$\textit{longitud de } \boldsymbol{a} \;=\; \|\boldsymbol{a}\| \;=\; \sqrt{\boldsymbol{a}\cdot\boldsymbol{a}}.$$

[1] Como en un triángulo rectángulo, ningún ángulo tiene más de 90 grados, el coseno del ángulo de cualquiera de sus vértices siempre es mayor o igual a cero.

Por ejemplo, la longitud de $\boldsymbol{x} = \begin{pmatrix} 6 \\ 0 \\ -2 \\ 3 \end{pmatrix}$ es $\sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{\begin{pmatrix} 6 \\ 0 \\ -2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 6 \\ 0 \\ -2 \\ 3 \end{pmatrix}} = \sqrt{6^2 + 0^2 + (-2)^2 + 3^2} = \sqrt{49} = 7.$

Puesto que $\boldsymbol{x} \cdot \boldsymbol{x} = 0$ si y solo si $\boldsymbol{x} = \boldsymbol{0}$, resulta que: $\|\boldsymbol{x}\| = 0$ si y solo si $\boldsymbol{x} = \boldsymbol{0}$.

Multiplicando $\boldsymbol{x}$ por $\lambda$, obtenemos un múltiplo de $\boldsymbol{x}$ cuya longitud es $|\lambda|$ veces su longitud original.

$$\|\lambda \boldsymbol{x}\| = \sqrt{\lambda^2 (\boldsymbol{x} \cdot \boldsymbol{x})} = |\lambda| \cdot \|\boldsymbol{x}\|.$$

**Definición 11.3** (Vector unitario)**.** *Se dice que un vector es unitario si tiene longitud uno.*

Es sencillo obtener un múltiplo unitario (i.e., de norma uno) de $\boldsymbol{x} \neq \boldsymbol{0}$; basta dividir el vector por su longitud. Por ejemplo, si $\boldsymbol{x} = (1, \quad 1, \quad 1, \quad 1,)$ entonces $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{4} = 2$. Así, el vector

$$\boldsymbol{y} = \frac{1}{\|\boldsymbol{x}\|}\boldsymbol{x} = \frac{1}{2}\boldsymbol{x} = (\tfrac{1}{2}, \quad \tfrac{1}{2}, \quad \tfrac{1}{2}, \quad \tfrac{1}{2},) \quad \text{es unitario. ¡Compruébelo!}$$

☞ Una vez definidas la perpendicularidad y la longitud en $\mathbb{R}^n$; vamos a generalizar la propiedad más importante que verifican los vectores perpendiculares: el **Teorema de Pitágoras**.

(recuerde las propiedades del *producto punto* en la Página 27)

**Teorema 11.1.1** (Teorema de Pitágoras en $\mathbb{R}^n$)**.** *Dos vectores $\boldsymbol{x}$ e $\boldsymbol{y}$ de $\mathbb{R}^n$* son perpendiculares *si y solo si*

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2.$$

*Demostración.* Como $\boldsymbol{x}$ e $\boldsymbol{y}$ son *perpendiculares* cuando $\boldsymbol{x} \cdot \boldsymbol{y} = 0$:

$$\begin{aligned} \|\boldsymbol{x} + \boldsymbol{y}\|^2 = (\boldsymbol{x} + \boldsymbol{y}) \cdot (\boldsymbol{x} + \boldsymbol{y}) &= \boldsymbol{x} \cdot (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{y} \cdot (\boldsymbol{x} + \boldsymbol{y}) \\ &= (\boldsymbol{x} + \boldsymbol{y}) \cdot \boldsymbol{x} + (\boldsymbol{x} + \boldsymbol{y}) \cdot \boldsymbol{y} \\ &= \boldsymbol{x} \cdot \boldsymbol{x} + \boldsymbol{y} \cdot \boldsymbol{x} + \boldsymbol{x} \cdot \boldsymbol{y} + \boldsymbol{y} \cdot \boldsymbol{y} \\ &= \boldsymbol{x} \cdot \boldsymbol{x} + \quad 2(\boldsymbol{x} \cdot \boldsymbol{y}) \quad + \boldsymbol{y} \cdot \boldsymbol{y} \qquad \text{donde } \boldsymbol{x} \cdot \boldsymbol{y} = 0 \text{ cuando } \boldsymbol{x} \perp \boldsymbol{y} \\ &= \boldsymbol{x} \cdot \boldsymbol{x} + \qquad 0 \qquad + \boldsymbol{y} \cdot \boldsymbol{y} \\ &= \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2. \end{aligned}$$

□

Fíjese que en el anterior teorema, $\boldsymbol{x}$ e $\boldsymbol{y}$ corresponden a los catetos y el vector suma $\boldsymbol{x} + \boldsymbol{y}$ a la hipotenusa.

☞ Para generalizar el *coseno del ángulo formado por dos vectores* a $\mathbb{R}^n$, antes necesitamos probar que $\boldsymbol{a} \cdot \boldsymbol{b}$ es siempre menor en valor absoluto que $\|\boldsymbol{a}\|\|\boldsymbol{b}\|$, de manera que $\frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}$ solo toma valores entre $-1$ y $1$. Dicho resultado se denomina *Desigualdad de Cauchy-Schwarz*.

**Proposición 11.1.2** (Desigualdad de Cauchy-Schwarz)**.** *Para $\boldsymbol{a}$ y $\boldsymbol{b}$ de $\mathbb{R}^n$ se verifica que* $|\boldsymbol{a} \cdot \boldsymbol{b}| \leq \|\boldsymbol{a}\|\|\boldsymbol{b}\|$.

*Demostración.* **Caso trivial:** *Si $\boldsymbol{a}$ y $\boldsymbol{b}$ son linealmente dependientes.* Entonces, o bien $\boldsymbol{a} = \boldsymbol{0}$ o bien existe $\lambda \in \mathbb{R}$ tal que $\lambda\boldsymbol{a} = \boldsymbol{b}$. Si ocurre lo primero, $\boldsymbol{a} \cdot \boldsymbol{b} = 0$ y $\|\boldsymbol{a}\|\|\boldsymbol{b}\| = 0$. Si ocurre lo segundo,

$$|\boldsymbol{a} \cdot \boldsymbol{b}| = |\boldsymbol{a} \cdot \lambda\boldsymbol{a}| = |\lambda| \cdot \|\boldsymbol{a}\|^2 = |\lambda| \cdot \|\boldsymbol{a}\|\|\boldsymbol{a}\| = \|\lambda\boldsymbol{a}\|\|\boldsymbol{a}\| = \|\boldsymbol{b}\|\|\boldsymbol{a}\|.$$

**Caso no trivial:** *Si $\boldsymbol{a}$ y $\boldsymbol{b}$ son linealmente independientes.* En tal caso, veamos que existe un vector $\boldsymbol{h}$ ortogonal a $\boldsymbol{a}$ tal que $\mathcal{L}\big([\boldsymbol{a}; \boldsymbol{h};]\big) = \mathcal{L}\big([\boldsymbol{a}; \boldsymbol{b};]\big)$.

Para encontrarlo, definimos el vector como $\boldsymbol{h} = \boldsymbol{b} - \alpha\boldsymbol{a}$; así garantizamos que $\mathcal{L}\big([\boldsymbol{a}; \boldsymbol{h};]\big) = \mathcal{L}\big([\boldsymbol{a}; \boldsymbol{b};]\big)$. Además, vamos a imponer la condición de que $\boldsymbol{a} \perp \boldsymbol{h}$; así

$$0 = \boldsymbol{a} \cdot \boldsymbol{h} = \boldsymbol{a} \cdot (\boldsymbol{b} - \alpha\boldsymbol{a}) = \boldsymbol{a} \cdot \boldsymbol{b} - \alpha\boldsymbol{a} \cdot \boldsymbol{a} \quad \text{por tanto} \quad \boxed{\alpha = \tfrac{\boldsymbol{a}\cdot\boldsymbol{b}}{\boldsymbol{a}\cdot\boldsymbol{a}}}. \tag{11.4}$$

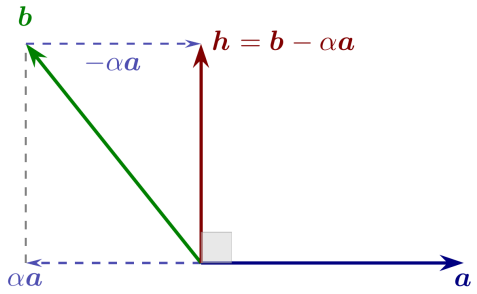



Figura 11.4: Representación esquemática donde $\boldsymbol{h}$ es combinación lineal de $\boldsymbol{a}$ y $\boldsymbol{b}$ y perpendicular a $\boldsymbol{a}$.

Por tanto, $\|\boldsymbol{b}\|^2 = \boldsymbol{b} \cdot \boldsymbol{b} = (\boldsymbol{h} + \alpha\boldsymbol{a}) \cdot (\boldsymbol{h} + \alpha\boldsymbol{a}) = \boldsymbol{h} \cdot \boldsymbol{h} + 2\alpha\boldsymbol{h} \cdot \boldsymbol{a} + \alpha^2\boldsymbol{a} \cdot \boldsymbol{a} = \|\boldsymbol{h}\|^2 + \alpha^2\|\boldsymbol{a}\|^2$.
(*Fíjese en el esquema de la figura: $\boldsymbol{b}$ es la hipotenusa y $\boldsymbol{h}$ y $\alpha\boldsymbol{a}$ son los catetos de un triángulo rectángulo*).

Como $\boldsymbol{a}$ y $\boldsymbol{b}$ son linealmente independientes, necesariamente $\boldsymbol{h} \neq \boldsymbol{0}$, es decir, $\|\boldsymbol{h}\| > 0$; y entonces

$$\|\boldsymbol{b}\|^2 = \|\boldsymbol{h}\|^2 + \alpha^2\|\boldsymbol{a}\|^2 > \alpha^2\|\boldsymbol{a}\|^2 = \frac{(\boldsymbol{a} \cdot \boldsymbol{b})^2}{(\boldsymbol{a} \cdot \boldsymbol{a})^2}\|\boldsymbol{a}\|^2 = \frac{(\boldsymbol{a} \cdot \boldsymbol{b})^2}{\big(\|\boldsymbol{a}\|^2\big)^2}\|\boldsymbol{a}\|^2 = \frac{(\boldsymbol{a} \cdot \boldsymbol{b})^2}{\|\boldsymbol{a}\|^2},$$

de donde se deduce que $\big(\|\boldsymbol{a}\|^2\big)\big(\|\boldsymbol{b}\|^2\big) > (\boldsymbol{a} \cdot \boldsymbol{b})^2$ y, por consiguiente, $\|\boldsymbol{a}\|\|\boldsymbol{b}\| > |\boldsymbol{a} \cdot \boldsymbol{b}|$. □

Fíjese que, en un extremo tenemos $\boldsymbol{a} \cdot \boldsymbol{b} = 0$ cuando $\boldsymbol{a} \perp \boldsymbol{b}$; y en el otro $|\boldsymbol{a} \cdot \boldsymbol{b}| = \|\boldsymbol{a}\|\|\boldsymbol{b}\|$ cuando $\boldsymbol{b} = \lambda\boldsymbol{a}$.

☞ Ahora, y gracias a que sabemos que se verifica la desigualdad de Cauchy-Schwarz, también podemos generalizar a $\mathbb{R}^n$ el *coseno del ángulo formado por dos vectores no nulos.*

**Definición 11.4.** *El coseno del ángulo $\theta$ formado por los vectores no nulos $\boldsymbol{a}$ y $\boldsymbol{b}$ de $\mathbb{R}^n$ es*

$$\cos\theta = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}; \qquad -1 \leq \cos\theta \leq 1.$$

Con esta generalización, la parte derecha de la Figura 11.3 se puede reinterpretar como una representación esquemática de un triángulo formado por vectores de $\mathbb{R}^n$.

☞ Por experiencia sabemos que la distancia entre dos puntos $\boldsymbol{a}$ y $\boldsymbol{c}$ siempre es menor o igual que la suma de las distancias del camino indirecto que va de $\boldsymbol{a}$ a $\boldsymbol{b}$ y luego de $\boldsymbol{b}$ a $\boldsymbol{c}$. Para extender el concepto de *distancia* debemos comprobar que ocurre lo mismo en $\mathbb{R}^n$. Para ello, demostraremos la *desigualdad triangular*.

**Proposición 11.1.3** (Desigualdad triangular)**.** *Sean $\boldsymbol{a}$ y $\boldsymbol{b}$ vectores de $\mathbb{R}^n$, entonces $\|\boldsymbol{a}+\boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$.*

*Demostración.*

$$\begin{aligned}\|\boldsymbol{a}+\boldsymbol{b}\|^2 = (\boldsymbol{a}+\boldsymbol{b})\cdot(\boldsymbol{a}+\boldsymbol{b}) &= \boldsymbol{a}\cdot\boldsymbol{a} + 2(\boldsymbol{a}\cdot\boldsymbol{b}) + \boldsymbol{b}\cdot\boldsymbol{b} \\ &= \|\boldsymbol{a}\|^2 + 2(\boldsymbol{a}\cdot\boldsymbol{b}) + \|\boldsymbol{b}\|^2 \\ &\leq \|\boldsymbol{a}\|^2 + 2\|\boldsymbol{a}\|\|\boldsymbol{b}\| + \|\boldsymbol{b}\|^2 \quad \text{puesto que } |\boldsymbol{a}\cdot\boldsymbol{b}| \leq \|\boldsymbol{a}\|\|\boldsymbol{b}\|;\ \text{(Prop. 11.1.2)} \\ &= (\|\boldsymbol{a}\| + \|\boldsymbol{b}\|)^2\end{aligned}$$

Consecuentemente $\|\boldsymbol{a}+\boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$. □

La *Desigualdad Triangular* con normas de vectores da lugar a la noción de *distancia entre vectores de* $\mathbb{R}^n$:

$$\text{distancia entre } \boldsymbol{x} \text{ y } \boldsymbol{y}: \quad d(\boldsymbol{x},\boldsymbol{y}) = \|\boldsymbol{x}-\boldsymbol{y}\| = \|\boldsymbol{y}-\boldsymbol{x}\| = d(\boldsymbol{y},\boldsymbol{x}),$$

que verifica otra *Desigualdad Triangular* (ahora referida a las distancias):

$$d(\boldsymbol{a},\boldsymbol{c}) = \|\boldsymbol{a}-\boldsymbol{c}\| = \|(\boldsymbol{a}-\boldsymbol{b})+(\boldsymbol{b}-\boldsymbol{c})\| \leq \|\boldsymbol{a}-\boldsymbol{b}\| + \|\boldsymbol{b}-\boldsymbol{c}\| = d(\boldsymbol{a},\boldsymbol{b}) + d(\boldsymbol{b},\boldsymbol{c}).$$

Puesto que $\|\boldsymbol{x}\| = 0$ si y solo si $\boldsymbol{x} = \boldsymbol{0}$, tenemos que $d(\boldsymbol{a},\boldsymbol{b}) = \|\boldsymbol{a}-\boldsymbol{b}\| = 0$ si y solo si $\boldsymbol{a} = \boldsymbol{b}$.

**Vectores alineados.** Cuando $\boldsymbol{y} = \lambda\boldsymbol{x}$ (cuando $\boldsymbol{y}$ es un múltiplo de $\boldsymbol{x}$), el producto punto entre ambos es

$$\boldsymbol{y}\cdot\boldsymbol{x} = \lambda\boldsymbol{x}\cdot\boldsymbol{x} = \lambda\|\boldsymbol{x}\|^2.$$

En tal caso (y si $\lambda \neq 0$), el ángulo que forman $\boldsymbol{x}$ e $\boldsymbol{y}$ es cero o bien $180^{\text{o}}$ (es decir, 0 ó $\pi$ radianes) ya que:

$$\cos\theta = \frac{\boldsymbol{x}\cdot\boldsymbol{y}}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} = \frac{\lambda\|\boldsymbol{x}\|^2}{|\lambda|\cdot\|\boldsymbol{x}\|\|\boldsymbol{x}\|} = \frac{\lambda}{|\lambda|} = \pm 1,$$

y entonces se dice que los vectores están alineados.

Cuando $\boldsymbol{a}$ y $\boldsymbol{b}$ están alineados: si $\lambda > 0$ entonces $\boldsymbol{a}\cdot\boldsymbol{b} = \|\boldsymbol{a}\|\|\boldsymbol{b}\|$; y si $\lambda < 0$, entonces $\boldsymbol{a}\cdot\boldsymbol{b} = -\|\boldsymbol{a}\|\|\boldsymbol{b}\|$.

Consecuentemente, cuando $\boldsymbol{a}$ y $\boldsymbol{b}$ están alineados y $\lambda > 0$ (cuando ambos "apuntan" en el mismo sentido), resulta que: $\|\boldsymbol{a}+\boldsymbol{b}\| = \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$ (véase la demostración de la Desigualdad Triangular).

## 11.2. Subespacios ortogonales en $\mathbb{R}^n$

**Definición 11.5** (Conjunto ortogonal al subconjunto $A$ de vectores de $\mathbb{R}^n$)**.** *El conjunto de todos los vectores de $\mathbb{R}^n$ que son perpendiculares a los vectores de $A \subset \mathbb{R}^n$ se denomina* conjunto ortogonal *(o perpendicular) a $A$, y se denota con $A^\perp$:*

$$A^\perp = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x}\cdot\boldsymbol{a} = 0 \ \text{ para cada } \ \boldsymbol{a} \in A\}$$

Ejercicio 58. Demuestre que $A^\perp$ es un subespacio.

¡Fíjese que, aunque $A$ no sea un subespacio, el conjunto ortogonal $A^\perp$ sí lo es!

Ejercicio 59.

(a) Demuestre que si $B$ es un subconjunto de vectores de $\mathbb{R}^n$ y $A \subset B$, entonces $B^\perp \subset A^\perp$.

(b) Demuestre que para todo subconjunto $A$ de vectores de $\mathbb{R}^n$ se verifica que $A^\perp = \mathcal{L}(A)^\perp$.
(*donde $\mathcal{L}(A)$ corresponde a la Definición 9.11 en la página 128*)

(c) Demuestre que para todo subespacio $\mathcal{A}$ de $\mathbb{R}^n$ se verifica que $\mathcal{A} \cap \mathcal{A}^\perp = \{\boldsymbol{0}\}$.

**Proposición 11.2.1** ($\mathcal{A} \subset \mathbb{R}^m$ y $\mathcal{A}^\perp$ son suplementarios)**.** [2] *Sean $\mathcal{A} \subset \mathbb{R}^m$ y $\mathcal{A}^\perp$, entonces $\mathcal{A} \oplus \mathcal{A}^\perp = \mathbb{R}^m$.*

*Demostración.* Sea el sistema $\mathsf{A}$ de $r$ vectores una base de $\mathcal{A}$. Por el apartado (b) del ejercicio anterior sabemos que $\mathcal{A}^\perp = \mathcal{L}(\mathsf{A})^\perp = \mathsf{A}^\perp$.

Si llamamos $\mathbf{A}$ al sistema $\mathsf{A}$ de vectores linealmente independientes de $\mathbb{R}^m$, entonces $\mathcal{A}^\perp$ es el subconjunto de vectores perpendiculares a las columnas de $\mathbf{A}$; es decir $\mathcal{A}^\perp = \{\boldsymbol{x} \in \mathbb{R}^m \mid \boldsymbol{x}\mathbf{A} = \mathbf{0}\} = \mathcal{N}\left(\mathbf{A}^\intercal\right)$.

Consecuentemente, $\dim \mathcal{A}^\perp = \dim \mathcal{N}\left(\mathbf{A}^\intercal\right) = m - r$. Como $\dim \mathcal{A} = r$ y $\dim \mathcal{A}^\perp = m - r$, y como $\dim \mathcal{A} \cap \mathcal{A}^\perp = \dim\{\mathbf{0}\} = 0$, por la Proposición 10.1.4 en la página 133, necesariamente $\mathcal{A} \oplus \mathcal{A}^\perp = \mathbb{R}^m$. □

*Observación.* Hay muchos subespacios suplementarios a uno dado: al observar la Figura 6.1 en la página 94 es fácil notar que cualquier función lineal $g\colon \mathcal{V} \to \mathcal{W}$ es un espacio suplementario de $\{\vec{0}\} \times \mathcal{W}$. Sin embargo, dado un subespacio, el correspondiente subespacio suplementario y ortogonal *es único*. En la citada figura corresponde al eje horizontal: $(\{\vec{0}\} \times \mathcal{W})^\perp = \mathcal{V} \times \{\vec{0}\}$. Esto da lugar a la siguiente

**Definición 11.6** (Complemento ortogonal de $\mathcal{A} \subset \mathbb{R}^n$)**.** *En el caso de que $\mathcal{A} \subset \mathbb{R}^n$ sea un subespacio, el subespacio $\mathcal{A}^\perp$* (conjunto ortogonal de $\mathcal{A}$) *se llama* complemento ortogonal *de $\mathcal{A}$.*

Así podemos reescribir el último apartado del Ejercicio 59 en la página anterior como: *'Demuestre que la intersección entre un subespacio de $\mathbb{R}^n$ y su* complemento ortogonal *es $\{\mathbf{0}\}$".*

### 11.2.1. Ortogonalidad de los 4 subespacios fundamentales de A

**Definición 11.7** (Subconjuntos ortogonales)**.** *Dos* subconjuntos *$A$ y $B$ de $\mathbb{R}^n$ son ortogonales ($A \perp B$), si todo vector de $A$ es perpendicular a todo vector de $B$; es decir, si $\boldsymbol{a} \cdot \boldsymbol{b} = 0$ para cualesquiera $\boldsymbol{a} \in A$ y $\boldsymbol{b} \in B$.*

Fíjese que un *subespacio* $\mathcal{A}$ y su complemento ortogonal $\mathcal{A}^\perp$ son un caso particular de subconjuntos ortogonales, pues $\mathcal{A} \perp \left(\mathcal{A}^\perp\right)$.

Ejercicio 60. Demuestre la siguiente

**Proposición 11.2.2.** *Sean $A$ y $B$ subconjuntos de $\mathbb{R}^n$, entonces $A \perp B$ si y solo si $\mathcal{L}(A) \perp \mathcal{L}(B)$.*

La siguiente proposición muestra que los subespacios fundamentales de $\mathbf{A}$ son ortogonales dos a dos.

**Proposición 11.2.3.** *Sea $\mathbf{A}$ de orden $m$ por $n$, entonces*

- *El espacio fila y el espacio nulo son ortogonales:* $\boxed{\mathcal{C}\left(\mathbf{A}^\intercal\right) \perp \mathcal{N}\left(\mathbf{A}\right)}$.
- *El espacio columna y el espacio nulo por la izquierda son ortogonales:* $\boxed{\mathcal{C}\left(\mathbf{A}\right) \perp \mathcal{N}\left(\mathbf{A}^\intercal\right)}$.

*Demostración.* El espacio fila es el conjunto de combinaciones lineales de las filas de $\mathbf{A}$, es decir, es el conjunto de vectores de la forma $\boldsymbol{f} = \boldsymbol{y}\mathbf{A}$ con $\boldsymbol{y} \in \mathbb{R}^m$; y los vectores del espacio nulo son los vectores $\boldsymbol{x}$ tales que $\mathbf{A}\boldsymbol{x} = \mathbf{0}$. Evidentemente, los vectores $\boldsymbol{f}$ y $\boldsymbol{x}$ son perpendiculares puesto que:

$$\boldsymbol{f} \cdot \boldsymbol{x} = \boldsymbol{y}\mathbf{A}\boldsymbol{x} = \boldsymbol{y} \cdot \mathbf{0} = 0.$$

Del mismo modo, como el espacio columna es el conjunto de combinaciones lineales de las columnas de $\mathbf{A}$, es decir, es el conjunto de vectores de la forma $\boldsymbol{b} = \mathbf{A}\boldsymbol{x}$ con $\boldsymbol{x} \in \mathbb{R}^n$; y como los vectores del espacio nulo por la izquierda son los vectores $\boldsymbol{y}$ tales que $\boldsymbol{y}\mathbf{A} = \mathbf{0}$; los vectores $\boldsymbol{y}$ y $\boldsymbol{b}$ también son perpendiculares:

$$\boldsymbol{y} \cdot \boldsymbol{b} = \boldsymbol{y}\mathbf{A}\boldsymbol{x} = \mathbf{0} \cdot \boldsymbol{x} = 0.$$

□

[2] En espacios vectoriales de dimensión infinita este resultado es falso en general (véase el ejemplo de la Página 156 donde $\mathcal{A}$ y $\mathcal{A}^\perp$ no son suplementarios); pero es verdadero cuando $\mathcal{A}$ es de dimensión finita (Proposición 11.3.4 en la página 155).

Módulo NAcAL para Python

```
a  = Vector( [1, 1, 1] )
A  = Matrix( [a, -a, a, -a] )

C  = SubEspacio(    A.sis() ) # Espacio columna de A
F  = SubEspacio( (~A).sis() ) # Espacio fila de A
N  = SubEspacio(  A )         # Espacio nulo de A
NI = SubEspacio( ~A )         # Espacio nulo por la izquierda de A

print( F  == ~N )       # ¿Es cierto que F  es igual a complemento ortogonal de N?
print( NI == ~C )       # ¿Es cierto que NI es igual a complemento ortogonal de C?
```

```
True
True
```

### El método de eliminación como generador de bases del complemento ortogonal

En la lección anterior vimos que con el *método de eliminación* podemos encontrar bases de los cuatro espacios fundamentales de una matriz. Consecuentemente, podemos encontrar una base del *complemento ortogonal* de cualquier subespacio de $\mathbb{R}^m$: si tenemos un sistema de vectores $\mathsf{Z}$ de $\mathbb{R}^m$, basta escribir dichos vectores como filas de una matriz y aplicar el método de eliminación "de izquierda a derecha"; las "soluciones especiales" serán base del complemento ortogonal del espacio $\mathcal{L}(\mathsf{Z})$ generado por $\mathsf{Z}$.

*Ejemplo* 27. Busquemos una base para el complemento ortogonal del espacio generado por los vectores

$$\left[\begin{pmatrix}1\\-3\\0\\-1\end{pmatrix};\begin{pmatrix}0\\-1\\1\\1\end{pmatrix};\begin{pmatrix}1\\-4\\1\\0\end{pmatrix};\right].$$

Aplicando la eliminación sobre una matriz cuyas filas son los vectores dados, tenemos:

$$\left[\begin{array}{cccc}1&-3&0&-1\\0&-1&1&1\\1&-4&1&0\\\hline 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right]\xrightarrow[\substack{[(3)\mathbf{1}+\mathbf{2}]\\[(1)\mathbf{1}+\mathbf{4}]\\[(1)\mathbf{2}+\mathbf{3}]\\[(1)\mathbf{2}+\mathbf{4}]}]{\tau}\left[\begin{array}{cccc}1&0&0&0\\0&-1&0&0\\1&-1&0&0\\\hline 1&3&3&4\\0&1&1&1\\0&0&1&0\\0&0&0&1\end{array}\right]\implies \text{Base del complemento ortogonal: }\left[\begin{pmatrix}3\\1\\1\\0\end{pmatrix};\begin{pmatrix}4\\1\\0\\1\end{pmatrix};\right].$$

Módulo NAcAL para Python

```
S = Sistema( [Vector([1,-3,0,-1]), Vector([0,-1,1,1]), Vector([1,-4,1,0])] )
~Matrix(S)                              # Matriz cuyas filas son los vectores de S
Homogenea(~Matrix(S),1).sgen            # Sistema generador del complemento ortogonal
SubEspacio( Homogenea(~Matrix(S)).sgen ) == ~SubEspacio(S)      # Comprobación
```

## 11.3. Generalizaciones o visiones alternativas (★)

☞ Aunque hasta este momento nos hemos limitado a trabajar en $\mathbb{R}^n$, lo interesante es que las ideas de longitud, ortogonalidad, norma, distancia y ángulo entre vectores son "de algún modo" extrapolables a espacios vectoriales genéricos $\mathcal{V}$ cuando en ellos definimos *productos escalares*. De hecho, las aplicaciones más interesantes ocurren en espacios vectoriales de dimensión infinita, por ejemplo en la Probabilidad.

En esta sección definimos el concepto de producto escalar en un espacio vectorial abstracto. Luego generalizamos las nociones de longitud, ortogonalidad, norma, distancia y ángulo entre vectores. Para finalizar, trataremos los subespacios ortogonales.

### 11.3.1. Productos escalares en $\mathcal{V}$

Llamamos producto escalar entre dos vectores $\vec{x}$ e $\vec{y}$ de $\mathcal{V}$ a cualquier función $\langle _ | _ \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ que verifique los axiomas de

- **Simetría:** $\langle \vec{x} | \vec{y} \rangle = \langle \vec{y} | \vec{x} \rangle$.
- **Linealidad respecto al primer argumento:**
  1. $\langle \alpha \vec{x} | \vec{y} \rangle = \alpha \langle \vec{x} | \vec{y} \rangle$.
  2. $\langle \vec{x} + \vec{y} | \vec{z} \rangle = \langle \vec{x} | \vec{z} \rangle + \langle \vec{y} | \vec{z} \rangle$.
- **Positivo:** $\langle \vec{x} | \vec{x} \rangle \geq 0$.
- **Definido:** $\langle \vec{x} | \vec{x} \rangle = 0 \Leftrightarrow \vec{x} = \vec{0}$.

Por tanto, un producto escalar es una *forma bilineal* (es decir, una función lineal en ambos argumentos) positiva y definida.

☞ Como ilustración, veamos algunos ejemplos de productos escalares. Primero en $\mathbb{R}^n$ y luego en otros espacios.

**Ejemplos de productos escalares en $\mathbb{R}^n$**

Fíjese que el producto punto es un producto escalar en $\mathbb{R}^n$

$$\begin{array}{rcl} \langle _ | _ \rangle : \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ (\boldsymbol{a}, \boldsymbol{b}) & \longmapsto & \boldsymbol{a} \cdot \boldsymbol{b} \end{array},$$

de hecho, el *producto punto* también se denomina *producto escalar usual en* $\mathbb{R}^n$. Sin embargo, hay otros ejemplos de productos escalares en $\mathbb{R}^n$ de uso muy extendido.

*Ejemplo* 28. En estadística se emplea un producto escalar que garantiza que la norma del vector $\mathbf{1}$ (cuyas componentes son todas iguales a uno) es 1 independientemente del número de unos que contenga el vector.

$$\begin{array}{rcl} \langle _ | _ \rangle_s : \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ (\boldsymbol{a}, \boldsymbol{b}) & \longmapsto & \frac{1}{n}(\boldsymbol{a} \cdot \boldsymbol{b}) \end{array}.$$

(Véase la Sección **??** en la página ??).

En la Lección **??** veremos que, usando este producto escalar, la desviación típica y el valor absoluto de la media aritmética son longitudes, y que la correlación es el coseno del ángulo formado entre vectores perpendiculares a $\mathbf{1}$.

Ejercicio 61. Demuestre que la función $\langle _ | _ \rangle_s : \begin{array}{rcl} \mathcal{V} \times \mathcal{V} & \longrightarrow & \mathbb{R} \\ (\boldsymbol{a}, \boldsymbol{b}) & \longmapsto & \frac{1}{n}(\boldsymbol{a} \cdot \boldsymbol{b}) \end{array}$ es un producto escalar en $\mathbb{R}^n$.

*Ejemplo* 29. En estadística también se emplean las *medias ponderadas*, en las que unos datos tienen mayor peso que otros. Sin embargo, se mantiene la propiedad de que la media (ponderada) de un vector de $n$ datos todos iguales a uno es uno. Un ejemplo es

$$\begin{array}{rl} \left\langle _ \middle| _ \right\rangle_{sp} : \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow \mathbb{R} \\ (\boldsymbol{a}, \boldsymbol{b}) & \longmapsto \boldsymbol{a}\mathbf{D}\boldsymbol{b} \end{array},$$

donde $\mathbf{D}$ es una matriz diagonal de orden $n$ cuyos elementos en la diagonal son positivos y suman 1. Fíjese que el anterior producto escalar $\left\langle _ \middle| _ \right\rangle_s$ corresponde al caso particular $\mathbf{D} = \frac{1}{n}\mathbf{I}$.

Ejercicio 62. Considere la matriz diagonal $\mathbf{D}$ de orden $n$ y cuyos elementos en la diagonal son positivos y suman 1. Demuestre que la función $\begin{array}{rl} \eta : \mathcal{V} \times \mathcal{V} & \longrightarrow \mathbb{R} \\ (\boldsymbol{a}, \boldsymbol{b}) & \longmapsto \boldsymbol{a}\mathbf{D}\boldsymbol{b} \end{array}$, es un producto escalar en $\mathbb{R}^n$.

*Ejemplo* 30. Para toda matriz $\mathbf{A}$ de orden $n$ y rango completo, la función

$$\begin{array}{rl} \eta : \mathcal{V} \times \mathcal{V} & \longrightarrow \mathbb{R} \\ (\boldsymbol{a}, \boldsymbol{b}) & \longmapsto \boldsymbol{a}\mathbf{A}^\intercal\mathbf{A}\boldsymbol{b} \end{array},$$

es un producto escalar en $\mathbb{R}^n$. De hecho, para cualquier producto escalar definido en $\mathbb{R}^n$ se puede encontrar una matriz $\mathbf{A}$ tal que dicho producto escalar se puede expresar de esta manera (ha de tenerse en cuenta que dicha matriz $\mathbf{A}$ no es única —véase la diagonalización por congruencia de la Lección 19).

**Convenio de notación:** si la función $\eta$ es un producto escalar, normalmente escribiremos $\left\langle \boldsymbol{a} \middle| \boldsymbol{b} \right\rangle_\eta$ para denotar el producto escalar $\eta$ entre los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ (es decir, para denotar $\eta(\boldsymbol{a}, \boldsymbol{b})$).

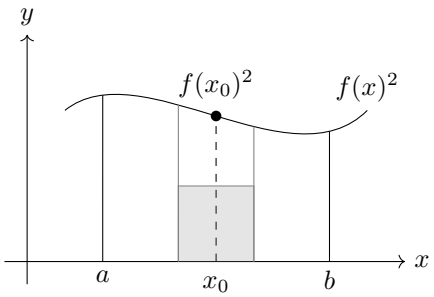
**Ejemplos de productos escalares en otros subespacios**

*Ejemplo* 31. El subconjunto $\mathcal{C}[a, b]$ de funciones reales y continuas con dominio en el intervalo $[a, b]$, junto con las operaciones habituales de suma de funciones y producto de funciones por escalares, es un subespacio del espacio $\mathbb{R}^{[a,b]}$ de funciones reales de variable real con dominio en el intervalo $[a, b]$. En $\mathcal{C}[a, b]$, la función

$$\begin{array}{rl} \eta : \mathcal{C}[0,1] \times \mathcal{C}[0,1] & \longrightarrow \mathbb{R} \\ (f, g) & \longmapsto \int_a^b f(x)g(x)dx \end{array}$$

es un producto escalar, pues verifica los axiomas:

- **Simetría:** $\int_a^b f(x)g(x)dx = \int_a^b g(x)f(x)dx$
- **Linealidad respecto al primer argumento:**
  1. $\int_a^b \big(f(x) + g(x)\big)h(x)dx = \int_a^b f(x)h(x)dx + \int_a^b g(x)h(x)dx$
  2. $\int_a^b \alpha f(x)g(x)dx = \alpha \int_a^b f(x)g(x)dx$
- **Positivo:** si $f$ no es nula, entonces $\int_a^b f(x)^2dx > 0$, ya que para el punto $x_0$ en el que la función es distinta de cero, existe un entorno tal que si $z \in (x_0 - \delta, x_0 + \delta)$ entonces $f(z)^2 \geq \frac{1}{2}\int_a^b f(x_0)^2dx$.



- **Definido:** $\int_a^b f(x)^2 dx = 0$, si y solo si $f$ es la función nula, es decir, $f(x) = 0$ para todo $x \in [a, b]$.

*Ejemplo* 32. Otro ejemplo que encontrará en estadística y econometría corresponde al conjunto de variables aleatorias de varianza finita definidas en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. En dicho subespacio, la esperanza $E(XY)$ del producto entre las variables aleatorias $X$ e $Y$ es un producto escalar.

### 11.3.2. Generalizando los conceptos geométricos a otros espacios vectoriales

☞ Aquí replicaremos la Sección 11.1.3 sustituyendo la notación de producto punto por la de producto escalar.

**Definición 11.8.** *Decimos que* $\vec{x}$ *e* $\vec{y}$ *son ortogonales* o perpendiculares ($\vec{x} \perp \vec{y}$) *cuando* $\langle \vec{x} | \vec{y} \rangle = 0$.

Ejercicio 63. ¿Qué vector es ortogonal a cualquier otro (incluido él mismo)?

**Definición 11.9.** *La longitud* (o norma) *de un vector* $\vec{x}$ *de* $\mathcal{V}$ *es la raíz cuadrada de* $\langle \vec{x} | \vec{x} \rangle$*:*

$$\textit{longitud de } \vec{x} \;=\; \|\vec{x}\| \;=\; \sqrt{\langle \vec{x} | \vec{x} \rangle}.$$

Ejercicio 64. Si $\|\vec{x}\| = 0$, ¿quién es $\vec{x}$?

Multiplicando $\vec{x}$ por $\lambda$ obtenemos un múltiplo de $\vec{x}$ cuya longitud es $|\lambda|$ veces su longitud original.

**Definición 11.10** (Vector unitario)**.** *Se dice que un vector es unitario si su longitud es uno.*

Para obtener un múltiplo unitario de $\vec{x}$ basta dividir el vector $\vec{x}$ por su longitud.

**Teorema 11.3.1** (Teorema de Pitágoras en $\mathcal{V}$)**.** *Sean* $\vec{x}$ *e* $\vec{y}$ *dos vectores de* $\mathcal{V}$*; entonces* $\langle \vec{x} | \vec{y} \rangle = 0$ (son perpendiculares) *si y solo si*

$$\|\vec{x} + \vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2.$$

Ejercicio 65. Demuestre el teorema de Pitágoras en $\mathcal{V}$ con producto escalar $\langle _ | _ \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$.

**Proposición 11.3.2** (Desigualdad de Cauchy-Schwarz)**.** *Para* $\vec{x}$ *e* $\vec{y}$ *de* $\mathcal{V}$ *y el producto escalar* $\langle _ | _ \rangle$ *definido en* $\mathcal{V}$ *se verifica que*

$$|\langle \vec{x} | \vec{y} \rangle| \leq \|\vec{x}\| \|\vec{y}\|.$$

Ejercicio 66. Demuestre la desigualdad de Cauchy-Schwarz.

Así, en un extremo $\langle \vec{x} | \vec{y} \rangle = 0$ cuando $\vec{x} \perp \vec{y}$; y en el otro $|\langle \vec{x} | \vec{y} \rangle| = \|\vec{x}\| \|\vec{y}\|$ cuando $\vec{y} = \lambda \vec{x}$.

**Definición 11.11.** *El coseno del ángulo* $\theta$ *formado por los vectores no nulos* $\vec{x}$ *y* $\vec{y}$ *de* $\mathcal{V}$ *es*

$$\cos\theta \;=\; \frac{\langle \vec{x} | \vec{y} \rangle}{\|\vec{x}\| \|\vec{y}\|}; \qquad -1 \leq \cos\theta \leq 1. \tag{11.5}$$

**Proposición 11.3.3** (Desigualdad triangular)**.** *Sean* $\vec{x}$ *y* $\vec{y}$ *vectores de* $\mathcal{V}$*, entonces* $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$.

Ejercicio 67. Demuestre la desigualdad triangular en $\mathcal{V}$.

La desigualdad triangular da lugar a la noción de *distancia entre vectores de* $\mathcal{V}$:

$$\text{distancia entre } \vec{x} \text{ y } \vec{y}\text{:} \quad d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = \|\vec{y} - \vec{x}\| = d(\vec{y}, \vec{x}),$$

Puesto que $\|\vec{x}\| = 0$ si y solo si $\vec{x} = \vec{0}$, tenemos que $d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = 0$ si y solo si $\vec{x} = \vec{y}$.

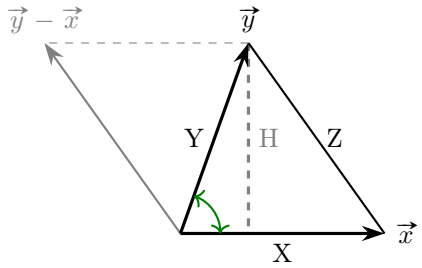



Figura 11.5: Representación esquemática del triángulo formado por $\vec{0}$ y dos vectores $\vec{x}$ y $\vec{y}$ no nulos de $\mathcal{V}$.

**Vectores alineados.** Cuando $\vec{y} = \lambda\vec{x}$ (cuando $\vec{y}$ es un múltiplo de $\vec{x}$) el producto punto entre ambos es

$$\langle \vec{y} \,|\, \vec{x} \rangle = \lambda\big(\langle \vec{x} \,|\, \vec{x} \rangle\big) = \lambda \|\vec{x}\|^2$$

En tal caso (y si $\lambda \neq 0$) el ángulo que forman $\vec{x}$ e $\vec{y}$ es cero o bien $180^{\text{o}}$ (es decir, 0 ó $\pi$ radianes) ya que:

$$\cos\theta = \frac{\langle \vec{x} \,|\, \vec{y} \rangle}{\|\vec{x}\|\,\|\vec{y}\|} = \frac{\lambda\|\vec{x}\|^2}{|\lambda| \cdot \|\vec{x}\|\,\|\vec{x}\|} = \frac{\lambda}{|\lambda|} = \pm 1.$$

y entonces se dice que los vectores están alineados.

### 11.3.3. Subespacios ortogonales en espacios vectoriales abstractos

**Definición 11.12** (Conjunto ortogonal al subconjunto $A$ de vectores de $\mathcal{V}$)**.** *Llamamos* conjunto ortogonal (o perpendicular) *a $A$, que denotamos como $A^\perp$, al subconjunto de vectores de $\mathcal{V}$ que son perpendiculares a los vectores de $A$.*

$$A^\perp = \big\{ \vec{x} \in \mathcal{V} \mid \langle \vec{x} \,|\, \vec{a} \rangle = 0 \;\; \textit{para cada} \;\; \vec{a} \in A \big\}$$

Ejercicio 68. Demuestre que $A^\perp$ es un subespacio.

¡Fíjese que aunque $A$ no sea un subespacio, $A^\perp$ sí lo es!

Ejercicio 69.

(a) Demuestre que si $B$ es un subconjunto de vectores de $\mathcal{V}$ y $A \subset B$, entonces $B^\perp \subset A^\perp$.

(b) Demuestre que para todo subconjunto $A$ de vectores de $\mathcal{V}$, se verifica que $A^\perp = \mathcal{L}\big(A\big)^\perp$.
(*donde $\mathcal{L}\big(A\big)$ corresponde a la Definición 9.11 en la página 128*)

(c) Demuestre que $\mathcal{A} \cap \mathcal{A}^\perp = \{\vec{0}\}$.

### 11.3.4. Espacios pre-Hilbert

**Definición 11.13.** *Se denomina espacio pre-Hilbert al par $\big(\mathcal{V}, \langle _ \,|\, _ \rangle\big)$, donde $\mathcal{V}$ es un espacio vectorial y $\langle _ \,|\, _ \rangle$ es un producto escalar definido en $\mathcal{V}$.*

**Definición 11.14.** *Cuando $\mathcal{V}$ es de dimensión finita, el espacio $\big(\mathcal{V}, \langle _ \,|\, _ \rangle\big)$ se denomina espacio euclídeo.*

*Observación.* Cuando nos referimos al *espacio euclídeo* $\mathbb{R}^n$ se sobreentiende que es aquél definido sobre $\mathbb{R}^n$ junto con el producto punto, es decir, al par: $\big(\mathbb{R}^n, _ \cdot _\big)$.

**Complementos ortogonales**

☞ Con la Proposición 11.2.1 en la página 149 ya hemos visto que en el espacio euclídeo $\big(\mathbb{R}^m, _ \cdot _\big)$ todo subespacio $\mathcal{A} \subset \mathbb{R}^m$ y su conjunto ortogonal $\mathcal{A}^\perp$ son suplementarios.

Ahora vamos a demostrar que en cualquier espacio pre-Hilbert $\big(\mathcal{V}, \langle _ \,|\, _ \rangle\big)$, si el subespacio $\mathcal{H}$ es de dimensión *finita*, entonces $\mathcal{V} = \mathcal{H} \oplus \mathcal{H}^\perp$, es decir, que los subespacios $\mathcal{H}$ y $\mathcal{H}^\perp$ son suplementarios cuando $\mathcal{H}$ es de dimensión *finita*.

**Teorema 11.3.4.** *Si $\big(\mathcal{V}, \langle _ | _ \rangle\big)$ es un espacio pre-Hilbert y $\mathcal{H}$ es un subespacio de $\mathcal{V}$ de dimensión finita, entonces necesariamente $\mathcal{H} + \mathcal{H}^\perp = \mathcal{V}$. Por tanto, si $\mathcal{H}$ es de dimensión finita se cumple que $\mathcal{H} \oplus \mathcal{H}^\perp = \mathcal{V}$.*

*Demostración.* Probar que $\mathcal{H} + \mathcal{H}^\perp = \mathcal{V}$ consiste en demostrar que para cualquier $\vec{v} \in \mathcal{V}$ existe $\vec{h} \in \mathcal{H}$ tal que $\vec{v} - \vec{h} \in \mathcal{H}^\perp$. Para ello, razonamos por inducción sobre la dimensión de $\mathcal{H}$.

Si $\dim(\mathcal{H}) = 0$, entonces $\mathcal{H} = \{\vec{0}\}$ y $\mathcal{H}^\perp = \mathcal{V}$, con lo que $\mathcal{H} + \mathcal{H}^\perp = \mathcal{V}$.

Ahora supongamos que el resultado es cierto para subespacios de dimensión $n$ y que $\dim(\mathcal{H}) = n + 1$. Sea entonces $\big[\vec{w}_1; \dots \vec{w}_n;\big]$ un sistema de $n$ vectores linealmente independientes de $\mathcal{H}$ y denotemos con $\mathcal{W}$ al subespacio $\mathcal{L}\big(\big[\vec{w}_1; \dots \vec{w}_n;\big]\big)$. Obviamente, $\dim(\mathcal{W}) = n$ y $\mathcal{W} \subset \mathcal{H}$; y como $\dim(\mathcal{H}) = n+1$, existe $\vec{h} \in \mathcal{H} - \mathcal{W}$ tal que $\big[\vec{w}_1; \dots \vec{w}_n; \vec{h};\big]$ es una base de $\mathcal{H}$. Por hipótesis de inducción sabemos que existe $\vec{h}_w \in \mathcal{W}$ tal que $\vec{h} - \vec{h}_w \in \mathcal{W}^\perp$. Y como $\vec{h}_w \in \mathcal{L}\big(\big[\vec{w}_1; \dots \vec{w}_n;\big]\big)$, tendremos que también $\big[\vec{w}_1; \dots \vec{w}_n; \vec{h} - \vec{h}_w;\big]$ es una base de $\mathcal{H}$. Denotemos, pues, con $\vec{z}$ al vector $\vec{h} - \vec{h}_w$.

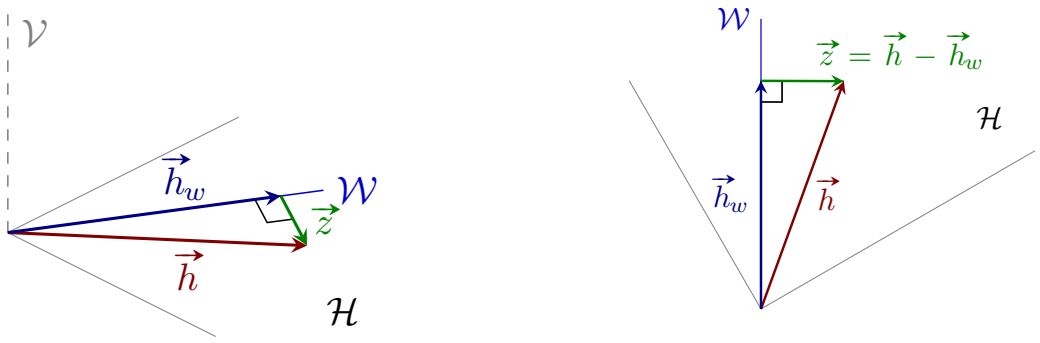



Figura 11.6: Representación esquemática desde dos ángulos distintos. El subespacio $\mathcal{W}$ es de dimensión $n$, el "suelo" $\mathcal{H}$ es un subespacio de dimensión $n+1$ que contiene a $\mathcal{W}$, y el espacio vectorial ambiente $\mathcal{V}$ (posiblemente de dimensión infinita) está representado por el espacio tridimensional (a la derecha se muestra la visión desde la vertical sobre $\mathcal{H}$).

Ahora tómenos $\vec{v} \in \mathcal{V}$. Por hipótesis de inducción existe $\vec{v}_w \in \mathcal{W}$ tal que $\vec{v} - \vec{v}_w$ pertenece a $\mathcal{W}^\perp$. Por tanto, $\{\vec{v} - \vec{v}_w, \vec{z}\} \perp \{\vec{w}_1, \dots, \vec{w}_n\}$, con lo que para cualquier $\alpha \in \mathbb{R}$ se verifica que

$$\left\{ \begin{array}{rcl} \vec{v} - \vec{v}_w - \alpha\vec{z} & \perp & \{\vec{w}_1, \dots, \vec{w}_n\} \\ \vec{v}_w + \alpha\vec{z} & \in & \mathcal{H} \end{array} \right. \tag{11.6}$$

Puesto que $\vec{z} \neq \vec{0}$ existe $\alpha_0 \in \mathbb{R}$ tal que $\vec{v} - \vec{v}_w - \alpha_0\vec{z} \perp \vec{z}$, ya que

$$\left. \begin{array}{rcl} 0 & = & \big\langle (\vec{v} - \vec{v}_w - \alpha_0\vec{z}) \big| \vec{z} \big\rangle \\ & = & \big\langle (\vec{v} - \vec{v}_w) \big| \vec{z} \big\rangle - \alpha_0 \|\vec{z}\|^2 \end{array} \right\} \quad \Leftrightarrow \quad \alpha_0 = \frac{\big\langle (\vec{v} - \vec{v}_w) \big| \vec{z} \big\rangle}{\|\vec{z}\|^2}$$

Pero entonces, si denotamos con $\vec{v}_h$ a $\vec{v}_w + \alpha_0\vec{z}$, tendremos que $\vec{v}_h \in \mathcal{H}$ y $\vec{v} - \vec{v}_h \in \mathcal{H}^\perp$ ya que $\vec{v} - \vec{v}_h \perp \{\underbrace{\vec{w}_1, \dots, \vec{w}_n}_{\text{por (11.6)}}, \vec{z}\}$. □

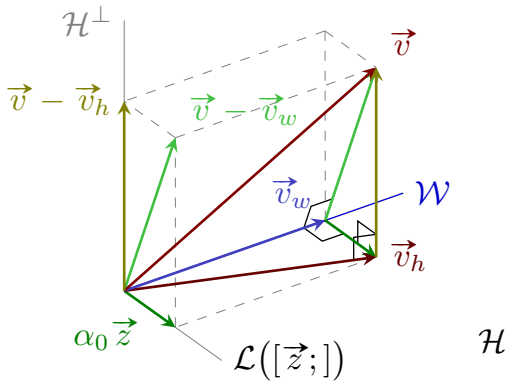



Figura 11.7: Representación esquemática. La recta $\mathcal{W}$ es un subespacio de dimensión $n$, el plano (el suelo) $\mathcal{H}$ es un subespacio de dimensión $n+1$ que contiene a $\mathcal{W}$, a $\mathcal{L}(\vec{z})$ (que es el único subespacio de la figura que con seguridad es de dimensión 1), y el espacio vectorial ambiente $\mathcal{V}$ (posiblemente de dimensión infinita) está representado por el espacio tridimensional (si $\mathcal{V}$ es de dimensión infinita, entonces $\mathcal{H}^\perp$ también).

### Un subespacio de dimensión infinita y su conjunto ortogonal no siempre son suplementarios

☞ Acabamos de ver que $\mathcal{H}$ y $\mathcal{H}^\perp$ son suplementarios cuando $\mathcal{H}$ es de *dimensión finita*.

Pero $\mathcal{H}$ y $\mathcal{H}^\perp$ pueden *no ser* suplementarios cuando el subespacio $\mathcal{H}$ es de *dimensión infinita*.[a] En esta sección veremos un ejemplo.

[a] Aunque si $\mathcal{H}$ es un subespacio cerrado de un espacio de Hilbert $\mathcal{V}$, entonces $\mathcal{V} = \mathcal{H} \oplus \mathcal{H}^\perp$ (independientemente de la dimensión de $\mathcal{H}$). En este caso, $\mathcal{H}^\perp$ se llama complemento ortogonal de $\mathcal{V}$... pero esto queda fuera del ámbito del curso.

*Ejemplo* 33. Considere el conjunto $\mathbb{R}[x]$ de sucesiones infinitas de números reales que se hacen cero a partir de una posición. Este conjunto también se describe como el conjunto de sucesiones infinitas de números reales con un número finito de componentes distintas de cero; por lo que dicho conjunto también se llama *conjunto de sucesiones casi nulas*. Algunos elementos de dicho conjunto son:

$$\begin{aligned} a =& (1,0,0,0,0,0,0,0\ldots)\\ b =& (1,2,0,0,0,0,0,0\ldots)\\ c =& (3,2,1,0,0,0,0,0\ldots)\\ d =& (0,1,0,1,0,0,0,0\ldots)\\ e =& (1,2,3,4,5,0,0,0\ldots) \end{aligned}$$

Definamos el producto escalar como la suma del producto entre componentes:

$$\begin{array}{rl} \langle _ | _ \rangle : \mathbb{R}[x] \times \mathbb{R}[x] & \longrightarrow \mathbb{R} \\ (p,q) & \longmapsto \sum_{i=0}^{\infty} p_i q_i \end{array}$$

(como el producto punto, pero con una suma infinita donde a partir de algún índice $k$ los sumandos son cero).

Así, con las sucesiones infinitas anteriores, tenemos que: $\langle a|b\rangle = 1$, $\langle c|d\rangle = 2$, y $\langle d|e\rangle = 6$.

Ahora consideremos el subespacio $\mathcal{H} \subset \mathbb{R}[x]$ de sucesiones cuyas componentes suman cero:

$$\mathcal{H} = \left\{ p \in \mathbb{R}[x] \;\middle|\; \sum_i p_i = 0 \right\}.$$

(fíjese en que ninguna de las sucesiones $a$, $b$, $c$, $d$ y $e$ pertenece a $\mathcal{H}$)

El conjunto ortogonal a $\mathcal{H}$ es

$$\mathcal{H}^\perp = \left\{ p \in \mathbb{R}[x] \;\middle|\; \langle p|q\rangle = 0 \;\text{ para todo }\; q \in \mathbb{R}[x] \right\}.$$

Consecuentemente, los elementos de $\mathcal{H}^\perp$ deben ser:

- perpendiculares a $(1, -1, 0, 0, 0, 0 \ldots)$; así que $p_0 = p_1$,
- perpendiculares a $(1, 1, -2, 0, 0, 0 \ldots)$; así que $p_0 + p_1 = 2p_2$ y, por tanto, $p_0 = p_1 = p_2$,
- perpendiculares a $(1, 1, 1, -3, 0, 0 \ldots)$; así que $p_0 + p_1 + p_2 = 3p_3$ y, por tanto, $p_0 = p_1 = p_2 = p_3$,
- perpendiculares a $(1, 1, 1, 1, -4, 0 \ldots)$; así que $p_0 + p_1 + p_2 + p_3 = 4p_4$ y, por tanto, $p_0 = p_1 = p_2 = p_3 = p_4$,
- ...

Por tanto, los elementos de $\mathcal{H}^\perp$ son sucesiones infinitas cuyas componentes son todas iguales; pero como las sucesiones de $\mathbb{R}$[x] son cero a partir de alguna posición, el único elemento posible de $\mathcal{H}^\perp$ es la sucesión nula.

Consecuentemente, el subespacio $\mathcal{H} + \mathcal{H}^\perp$ no contiene ninguna de las sucesiones $a$, $b$, $c$, $d$ y $e$; es decir, $\mathcal{H} + \mathcal{H}^\perp$ no es todo el espacio $\mathbb{R}[x]$. Así que, $\mathcal{H}$ y $\mathcal{H}^\perp$ no son suplementarios.

Fíjese en que el subespacio de dimensión infinita formado por las *sucesiones casi nulas* $\mathbb{R}[x]$ resulta ser el conjunto de polinomios en una variable[3], aunque habitualmente se escriben de otra manera... en el caso de los del Ejemplo 33 en la página anterior:

$$\begin{aligned}
a =& (1, 0, 0, 0, 0, 0, 0, 0 \ldots) = 1 \\
b =& (1, 2, 0, 0, 0, 0, 0, 0 \ldots) = 1 + 2x \\
c =& (3, 2, 1, 0, 0, 0, 0, 0 \ldots) = 3 + 2x + x^2 \\
d =& (0, 1, 0, 1, 0, 0, 0, 0 \ldots) = x + x^3 \\
e =& (1, 2, 3, 4, 5, 0, 0, 0 \ldots) = 1 + 2x + 3x^2 + 4x^3 + 5x^4
\end{aligned}$$

así, el exponente de cada "$x$" indica la posición de su correspondiente coeficiente dentro de la sucesión; donde se asocia la primera posición al índice 0, es decir, puesto que $y^0 = 1$, entonces $ax^0 = (a, 0, 0 \ldots)$.

## *Apéndices a la lección*

[3] Véase la definición de polinomio en la Sección 19 en la página 215

# Ecuaciones cartesianas y paramétricas

## 12.1. De las ecuaciones paramétricas a las ecuaciones cartesianas

En la Sección 7.4 de la Lección 7 aparecieron por primera vez tanto las ecuaciones paramétricas como las ecuaciones cartesianas (véase las notas a pie de página números 3 en la página 107 y 1 en la página 103).

Cada modo de representación tiene un propósito distinto: las ecuaciones paramétricas nos permiten generar ejemplos de vectores que pertenecen al conjunto (basta dar valores arbitrarios a los parámetros); las ecuaciones cartesianas nos permiten comprobar si un vector $\boldsymbol{y}$ pertenece al conjunto (basta comprobar si $\boldsymbol{y}$ es solución del sistema de ecuaciones indicado en las ecuaciones cartesianas).

Veamos ahora cómo pasar de unas a las otras; así dispondremos en todo momento de dos modos de expresar el mismo conjunto de vectores.

### 12.1.1. El camino recorrido hasta ahora: de las ecuaciones cartesianas a las paramétricas

Dado un sistema de ecuaciones lineales $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, donde $\mathbf{A}$ es de orden $m$ por $n$, podemos expresar el conjunto de sus soluciones con la siguiente *ecuación cartesiana*:

$$\{\boldsymbol{x} \in \mathbb{R}^n \,|\, \mathbf{A}\boldsymbol{x} = \boldsymbol{b}\}$$

Al resolver dicho sistema encontrando una solución particular $\boldsymbol{s}$ y una base $\left[\boldsymbol{n}_1; \ldots \boldsymbol{n}_k;\right]$ de $\mathcal{N}\left(\mathbf{A}\right)$, reexpresamos el anterior conjunto con unas *ecuaciones paramétricas* (denominadas *conjunto solución*):

$$\left\{\boldsymbol{x} \in \mathbb{R}^n \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^k \text{ tal que } \boldsymbol{x} = \boldsymbol{s} + \left[\boldsymbol{n}_1; \ldots \boldsymbol{n}_k;\right]\boldsymbol{p}\right\}.$$

Es decir, al resolver un sistema de ecuaciones, partimos de unas ecuaciones cartesianas, o implícitas, para obtener unas ecuaciones paramétricas (el conjunto solución). Veamos ahora cómo recorrer el camino inverso: partiendo de un conjunto solución, obtener un sistema de ecuaciones cuyas soluciones son dicho conjunto.

### 12.1.2. Recorriendo el camino inverso: de las ecuaciones paramétricas a las cartesianas

*En el camino inverso, dadas unas ecuaciones paramétricas para conjunto de soluciones de "un sistema desconocido, $\mathbf{D}\boldsymbol{x} = \boldsymbol{f}$", encontraremos las filas de una matriz de coeficientes $\mathbf{A}$ y un vector del lado derecho $\boldsymbol{b}$ que definan un sistema de ecuaciones $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ equivalente a $\mathbf{D}\boldsymbol{x} = \boldsymbol{f}$, es decir, cuyo conjunto de soluciones coincida con el del sistema de ecuaciones desconocido, $\mathbf{D}\boldsymbol{x} = \boldsymbol{f}$.*

**Procedimiento.** Cuando disponemos de unas ecuaciones paramétricas

$$\left\{\boldsymbol{x} \in \mathbb{R}^n \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^k \text{ tal que } \boldsymbol{x} = \boldsymbol{s} + \left[\boldsymbol{n}_1; \ldots \boldsymbol{n}_k;\right]\boldsymbol{p}\right\},$$

disponemos tanto de una solución $\boldsymbol{s}$ (un vector del conjunto de soluciones) como de una base $\left[\boldsymbol{n}_1; \ldots \boldsymbol{n}_k;\right]$ del espacio nulo de la matriz $\mathbf{A}$ que buscamos, es decir, $\mathcal{N}\left(\mathbf{A}\right) = \mathcal{L}\left(\left[\boldsymbol{n}_1; \ldots \boldsymbol{n}_k;\right]\right)$.

Sabemos que, por eliminación, podemos encontrar una base del complemento ortogonal de $\mathcal{L}\big(\big[\boldsymbol{n}_1;\dots\boldsymbol{n}_k;\big]\big)$, es decir, una base de $\mathcal{C}\left(\mathbf{A}^\intercal\right)$. Usemos los vectores de dicha base para formar las *filas* de $\mathbf{A}$. Los pasos son:

**Primero** encontrar una base del complemento ortogonal de $\mathcal{N}\left(\mathbf{A}\right) = \mathcal{L}\big(\big[\boldsymbol{n}_1;\dots\boldsymbol{n}_k;\big]\big)$. Para ello aplicamos la eliminación "de izquierda a derecha" sobre la matriz $\mathbf{Z}$ cuyas $k$ filas son los vectores $\boldsymbol{n}_k,\dots,\boldsymbol{n}_n$ siguiendo el esquema:

$$\begin{bmatrix}\mathbf{Z}\\\hline\mathbf{I}\end{bmatrix} \xrightarrow{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \begin{bmatrix}\mathbf{K}\\\hline\mathbf{E}\end{bmatrix}; \qquad \text{donde } \mathbf{Z} = \big[\boldsymbol{n}_1;\dots\boldsymbol{n}_k;\big]^\intercal \qquad \text{(nótese el símbolo de transposición).}$$

Como $\mathbf{ZE} = \mathbf{K}$; si $\boldsymbol{\nu}$ es la lista de índices de las columnas nulas de $\mathbf{K}$, entonces $\mathbf{ZE}_{|\boldsymbol{\nu}} = \mathbf{K}_{|\boldsymbol{\nu}} = \mathbf{0}$.

**Después**. Si usamos las columnas de $\mathbf{E}_{|\boldsymbol{\nu}}$ (es decir, las soluciones especiales) como *filas* para construir una matriz $\mathbf{A}$, es decir, si $\mathbf{E}_{|\boldsymbol{\nu}} = \mathbf{A}^\intercal$ tenemos que $\mathbf{Z}(\mathbf{A}^\intercal) = \mathbf{0}$, o transponiendo ambos lados, $\mathbf{A}\big[\boldsymbol{n}_1;\dots\boldsymbol{n}_k;\big] = \mathbf{0}$.

Si ahora multiplicamos por $\mathbf{A}$ la ecuación paramétrica $\boldsymbol{x} = \boldsymbol{s} + \big[\boldsymbol{n}_1;\dots\boldsymbol{n}_k;\big]\boldsymbol{p}$ tenemos que

$$\mathbf{A}\boldsymbol{x} \;=\; \mathbf{A}\boldsymbol{s} \;+\; \underbrace{\mathbf{A}\big[\boldsymbol{n}_1;\dots\boldsymbol{n}_k;\big]}_{\mathbf{0}}\boldsymbol{p} \quad\Rightarrow\quad \mathbf{A}\boldsymbol{x} \;=\; \mathbf{A}\boldsymbol{s}, \tag{12.1}$$

donde $\boldsymbol{b} = \mathbf{A}\boldsymbol{s}$ será el vector del lado derecho del sistema que buscamos. Este sistema es, por construcción, un sistema cuyas soluciones son las ecuaciones paramétricas de las que hemos partido, ya que

$$\begin{cases}\mathbf{A}\boldsymbol{x} = \boldsymbol{b} & \text{cuando} \quad \boldsymbol{x} = \boldsymbol{s}\\ \mathbf{A}\boldsymbol{x} = \mathbf{0} & \text{cuando} \quad \boldsymbol{x} \in \mathcal{L}\big(\big[\boldsymbol{n}_1;\dots\boldsymbol{n}_k;\big]\big) = \mathcal{N}\left(\mathbf{A}\right)\end{cases}.$$

**Procedimiento abreviado para usar con lápiz y papel.** No es necesario encontrar primero los vectores perpendiculares a $\boldsymbol{n}_k,\dots,\boldsymbol{n}_n$ para formar con ellos las filas de $\mathbf{A}$ y multiplicar la ecuación paramétrica para obtener el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$. Podemos hacerlo todo a la vez con *la eliminación de izquierda a derecha.*

Sea $\mathbf{M}$ la matriz que resulta tras añadir los vectores $\boldsymbol{x}$ y $\boldsymbol{s}$ como nuevas filas de $\mathbf{Z}$. Entonces, aplicando la secuencia de transformaciones $\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k$ que transforma $\mathbf{Z}$ en una forma preescalonada $\mathbf{K}$, tenemos que

$$\mathbf{M} = \begin{bmatrix}\mathbf{Z}\\\hline\boldsymbol{x}\\\hline\boldsymbol{s}\end{bmatrix} \xrightarrow{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \begin{bmatrix}\mathbf{K}\\\hline\boldsymbol{x}\mathbf{E}\\\hline\boldsymbol{s}\mathbf{E}\end{bmatrix}, \quad \text{donde } \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

Por la Ecuación 12.1 sabemos que $\mathbf{A}\boldsymbol{x} = \mathbf{A}\boldsymbol{s}$, es decir $\boldsymbol{x}(\mathbf{A}^\intercal) = \boldsymbol{s}(\mathbf{A}^\intercal)$, por tanto $\boldsymbol{x}\mathbf{E}_{|\boldsymbol{\nu}} = \boldsymbol{s}\mathbf{E}_{|\boldsymbol{\nu}}$; donde $\boldsymbol{\nu}$ es la lista de índices de las columnas nulas de $\mathbf{K}$. Así pues, si $\mathbf{K}_{|j} = \mathbf{0}$, entonces $\boldsymbol{x}\mathbf{E}_{|j} = \boldsymbol{s}\mathbf{E}_{|j}$. Dicho de otro modo, *bajo las columnas nulas de* $\mathbf{K}$ *podemos ver las ecuaciones del sistema buscado.*

*Ejemplo* 34. Encontremos unas ecuaciones cartesianas para el siguiente conjunto $P$ de vectores de $\mathbb{R}^4$:

$$P = \left\{ \boldsymbol{v}\in\mathbb{R}^4 \;\middle|\; \exists\boldsymbol{p}\in\mathbb{R}^2,\ \boldsymbol{v} = \begin{pmatrix}0\\2\\0\\5\end{pmatrix} + \begin{bmatrix}1&0\\2&0\\0&1\\-2&3\end{bmatrix}\boldsymbol{p} \right\}.$$

Aplicando la eliminación sobre la matriz $\mathbf{M}$, cuya tercera fila es $\boldsymbol{x} = (x,y,z,w,)$ y cuya última fila es la solución particular $\boldsymbol{s} = (0,2,0,5,)$, tenemos:

$$\left[\begin{array}{cccc}1&2&0&-2\\0&0&1&3\\\hline x&y&z&w\\\hline 0&2&0&5\end{array}\right] \xrightarrow[[(2)\mathbf{1}+\mathbf{4}]]{\boldsymbol{\tau}[(-2)\mathbf{1}+\mathbf{2}]} \left[\begin{array}{cccc}1&0&0&0\\0&0&1&3\\\hline x&-2x+y&z&2x+w\\\hline 0&5&0&5\end{array}\right] \xrightarrow{\boldsymbol{\tau}[(-3)\mathbf{3}+\mathbf{4}]} \left[\begin{array}{cccc}1&0&0&0\\0&0&1&0\\\hline x&-2x+y&z&2x-3z+w\\\hline 0&2&0&5\end{array}\right].$$

Mirando bajo las columnas nulas se vislumbran las ecuaciones que buscamos:

$$\begin{cases} -2x + y & = 2 \\ 2x & -3z + w = 5 \end{cases}.$$

Así pues, unas *ecuaciones cartesianas* para $P$ son: $\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix} -2 & 1 & 0 & 0 \\ 2 & 0 & -3 & 1 \end{bmatrix} \boldsymbol{v} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \right\}$.

## 12.2. Puntos, rectas, planos e hiperplanos de $\mathbb{R}^n$. Espacios afines

Dado un sistema de ecuaciones $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, su conjunto de soluciones posee una forma geométrica plana (que "no se curva"). En particular, si $\dim \mathcal{N}\left(\mathbf{A}\right) = 0$, el conjunto es *un punto*; si $\dim \mathcal{N}\left(\mathbf{A}\right) = 1$, el conjunto es *una recta*; cuando $\dim \mathcal{N}\left(\mathbf{A}\right) = 2$, es *un plano*, y si $\dim \mathcal{N}\left(\mathbf{A}\right) = d > 2$, es *un hiperplanos de dimensión* $d$.

Esto sugiere emplear nuevos nombres para conceptos ya conocidos. Podemos denominar ecuaciones de un punto, de una recta, de un plano, etc., al conjunto de soluciones del correspondiente sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$. Así, podemos preguntarnos por las ecuaciones paramétricas o cartesianas de una recta en $\mathbb{R}^n$ (y al decir *recta* ya indicamos implícitamente que el espacio nulo de la matriz de coeficientes de las ecuaciones cartesianas tiene dimensión uno).

A los subconjuntos de $\mathbb{R}^n$ que podemos describir como la suma de un vector particular $\boldsymbol{s}$ más las combinaciones lineales de un sistema de vectores $\left[\boldsymbol{n}_1; \ldots \boldsymbol{n}_k;\right]$ (es decir, a aquellos subconjuntos que podemos describir como el conjunto de soluciones de un sistema de ecuaciones $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$) se les denomina *espacios afines*.

La dimensión de un espacio afín es la dimensión del subespacio generado por las combinaciones lineales del citado sistema de vectores $\left[\boldsymbol{n}_1; \ldots \boldsymbol{n}_k;\right]$ (o, equivalentemente, la dimensión del espacio nulo de la matriz de coeficientes $\mathbf{A}$ del correspondiente sistema de ecuaciones).

NAcAL guarda el conjunto de soluciones de $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ como un espacio afín en el atributo `eafin`. Su representación muestra tanto las ecuaciones cartesianas como las paramétricas.

Módulo NAcAL para Python

```
A = Matrix( [ [-1,1,0,0], [-1,0,-1,1] ] )
b = Vector([0,1])
SEL(A,b).eafin  # conjunto de soluciones
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^2,\ \boldsymbol{v} = \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \begin{bmatrix} -1 & 1 \\ -1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & -1 & 1 \end{bmatrix} \boldsymbol{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

NAcAL permite generar espacios afines de varias maneras. Con una matriz $\mathbf{A}$ y un punto $\boldsymbol{v}$:

Módulo NAcAL para Python

```
A = Matrix( [ [-1,1,0,0], [-1,0,-1,1] ] )
v = Vector([0,0,0,1])
EAfin(A,v)     # contiene al vector v y es paralelo a N(A)
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^2,\ \boldsymbol{v} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \begin{bmatrix} -1 & 1 \\ -1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & -1 & 1 \end{bmatrix} \boldsymbol{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

o con un sistema de vectores $\mathsf{S}$ y un punto $\boldsymbol{v}$:

Módulo NAcAL para Python

```
S = Sistema([ Vector([-1,-1,1,0]), Vector([0,0,1,1]) ])
v = Vector([0,0,0,1])
EAfin(S,v)      # contiene al vector v y es paralelo a L(S)
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^2,\ \boldsymbol{v} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \begin{bmatrix} -1 & 0 \\ -1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & -1 & 1 \end{bmatrix} \boldsymbol{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

(Por hacer: *definir el conjunto de soluciones como un elemento de un espacio cociente.*)

## 12.3. Escogiendo ecuaciones paramétricas (despejando variables)

Por hacer: *describir la eliminación sobre los vectores que forman la solución para escoger una base del espacio nulo interesante; es decir, para despejar unas variables con el propósito de expresarlas en función de las otras (si ello es posible...)*

# Proyecciones sobre subespacios

## 13.1. Proyección ortogonal y mínima distancia

**Definición 13.1.** *Llamamos* proyección ortogonal de $\boldsymbol{y}$ sobre el subespacio $\mathcal{V} \subset \mathbb{R}^m$ *a la función*[1] $\mathrm{Prj}_{\mathcal{V}}(\boldsymbol{y})$ *tal que la diferencia* $\boldsymbol{y} - \mathrm{Prj}_{\mathcal{V}}(\boldsymbol{y})$ *es ortogonal a* $\mathcal{V}$*; es decir, tal que* $\big(\boldsymbol{y} - \mathrm{Prj}_{\mathcal{V}}(\boldsymbol{y})\big) \in \mathcal{V}^{\perp}$.

Para cada $\boldsymbol{y}$ de $\mathbb{R}^m$, el vector $\widehat{\boldsymbol{y}} = \mathrm{Prj}_{\mathcal{V}}(\boldsymbol{y})$ de $\mathcal{V}$ es la *proyección ortogonal de* $\boldsymbol{y}$ *sobre el subespacio* $\mathcal{V} \subset \mathbb{R}^m$. Como $\mathbb{R}^m = \mathcal{V} \oplus \mathcal{V}^{\perp}$, sabemos que el vector $\widehat{\boldsymbol{y}} \in \mathcal{V}$ existe y es único (Proposición 11.2.1)[2].

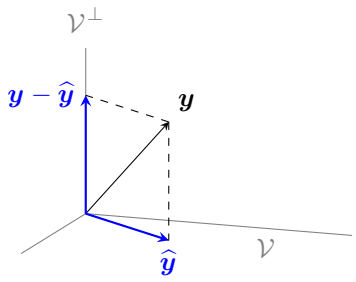



Figura 13.1: Representación esquemática de la proyección ortogonal de $\boldsymbol{y}$ sobre $\mathcal{V}$

Fíjese en la Figura 11.4 en la página 147. El vector $-\alpha\boldsymbol{a}$ es la proyección ortogonal de $\boldsymbol{b}$ sobre $\mathcal{L}(\boldsymbol{a})$. Es decir, aún sin nombrar la proyección ortogonal, la hemos empleado implícitamente en la lección anterior.

Lo más interesante de $\widehat{\boldsymbol{y}}$ se afirma en la siguiente

**Proposición 13.1.1.** *La proyección ortogonal de* $\boldsymbol{y} \in \mathbb{R}^m$ *sobre* $\mathcal{V} \subset \mathbb{R}^m$ *es el vector de* $\mathcal{V}$ *más próximo a* $\boldsymbol{y}$

*Demostración.* Sea $\widehat{\boldsymbol{y}}$ la proyección de $\boldsymbol{y}$ sobre $\mathcal{V}$ y tomemos un vector $\boldsymbol{v}$ cualquiera de $\mathcal{V}$. Veamos que $\boldsymbol{y}$ está más lejos de $\boldsymbol{v}$ que de su proyección ortogonal $\widehat{\boldsymbol{y}}$.

Como $\boldsymbol{v}$ y $\widehat{\boldsymbol{y}}$ están en $\mathcal{V}$, su diferencia $(\widehat{\boldsymbol{y}} - \boldsymbol{v})$ está en $\mathcal{V}$ (pues $\mathcal{V}$ es subespacio); consecuentemente $(\widehat{\boldsymbol{y}} - \boldsymbol{v})$ es ortogonal a $(\boldsymbol{y} - \widehat{\boldsymbol{y}}) \in \mathcal{V}^{\perp}$ (por ser $\widehat{\boldsymbol{y}}$ la proyección ortogonal sobre $\mathcal{V}$). Y como la suma de ambos vectores perpendiculares es $(\boldsymbol{y} - \widehat{\boldsymbol{y}}) + (\widehat{\boldsymbol{y}} - \boldsymbol{v}) = (\boldsymbol{y} - \boldsymbol{v})$, por el Teorema de Pitágoras concluimos que

$$\|\boldsymbol{y} - \boldsymbol{v}\|^2 = \|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|^2 + \|\widehat{\boldsymbol{y}} - \boldsymbol{v}\|^2 \geq \|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|^2.$$

Por tanto, $\boxed{\|\boldsymbol{y} - \boldsymbol{v}\| \geq \|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|}$ para todo $\boldsymbol{v} \in \mathcal{V}$. □

[1] Al ser un caso particular del apartado (b) del ejercicio 55 en la página 133, la función proyección es lineal.
[2] Por otra parte, el vector $\boldsymbol{y} - \mathrm{Prj}_{\mathcal{V}}(\boldsymbol{y})$ es la *proyección ortogonal de* $\boldsymbol{y}$ *sobre* $\mathcal{V}^{\perp}$.

☞ Por la Proposición 11.2.1 sabemos que para toda matriz $\mathbf{A}$ de $m$ filas y un vector $\boldsymbol{y}$ de $\mathbb{R}^m$, existen $\widehat{\boldsymbol{y}} \in \mathcal{C}(\mathbf{A})$ y $\boldsymbol{e} \in \mathcal{N}(\mathbf{A}^\intercal)$ tales que $\boldsymbol{y} = \widehat{\boldsymbol{y}} + \boldsymbol{e}$. Pero ¿cómo encontrar dichos vectores $\widehat{\boldsymbol{y}}$ y $\boldsymbol{e}$? Veámoslo...

### 13.1.1. Sistema de ecuaciones normales

Para encontrar la proyección ortogonal $\widehat{\boldsymbol{y}}$ del vector $\boldsymbol{y}$ sobre $\mathcal{C}(\mathbf{A})$ combinaremos lo que sabemos:

- $\widehat{\boldsymbol{y}}$ es una combinación lineal de las columnas de $\mathbf{A}$, es decir, existe $\boldsymbol{x} \in \mathbb{R}^n$ tal que $\mathbf{A}\boldsymbol{x} = \widehat{\boldsymbol{y}}$.
- $\widehat{\boldsymbol{y}}$ es la proyección ortogonal de $\boldsymbol{y}$ sobre $\mathcal{C}(\mathbf{A})$, es decir $(\boldsymbol{y} - \widehat{\boldsymbol{y}}) \in \mathcal{N}(\mathbf{A}^\intercal)$.

Así pues, una combinación lineal $\mathbf{A}\boldsymbol{x}$ resulta ser $\widehat{\boldsymbol{y}}$ si y solo si $\mathbf{A}^\intercal(\boldsymbol{y} - \mathbf{A}\boldsymbol{x}) = \mathbf{0}$:

$$\mathbf{A}\boldsymbol{x} = \widehat{\boldsymbol{y}} \iff \mathbf{A}^\intercal(\boldsymbol{y} - \mathbf{A}\boldsymbol{x}) = \mathbf{A}^\intercal\boldsymbol{y} - (\mathbf{A}^\intercal\mathbf{A})\boldsymbol{x} = \mathbf{0} \iff \boxed{(\mathbf{A}^\intercal\mathbf{A})\boldsymbol{x} = \mathbf{A}^\intercal\boldsymbol{y}}. \tag{13.1}$$

El sistema de ecuaciones $(\mathbf{A}^\intercal\mathbf{A})\boldsymbol{x} = \mathbf{A}^\intercal\boldsymbol{y}$ se denomina *sistema de ecuaciones normales*; y para obtener la proyección de $\boldsymbol{y}$ sobre $\mathcal{C}(\mathbf{A})$ basta multiplicar $\mathbf{A}$ por cualquier vector $\boldsymbol{x}$ que sea solución a dicho sistema.

Lo llamativo es que para resolver $\mathbf{A}\boldsymbol{x} = \widehat{\boldsymbol{y}}$ (para encontrar la combinación lineal de las columnas de $\mathbf{A}$ requerida) resolvemos un sistema diferente (el *sistema de ecuaciones normales*) donde no aparece $\widehat{\boldsymbol{y}}$. Este procedimiento indirecto funciona porque el *sistema de ecuaciones normales* y el sistema de ecuaciones $\mathbf{A}\boldsymbol{x} = \widehat{\boldsymbol{y}}$ tienen el mismo conjunto de soluciones (fíjese en las implicaciones "si y solo si" de (13.1)).

**Definición 13.2.** *Dos sistemas de ecuaciones son* equivalentes *si poseen el mismo conjunto de soluciones.*

Así, como $\mathbf{A}\boldsymbol{x} = \widehat{\boldsymbol{y}}$ tiene solución (pues $\widehat{\boldsymbol{y}} \in \mathcal{C}(\mathbf{A})$), el *sistema de ecuaciones normales* también tiene solución (por ser equivalente). Es más, por ser equivalentes, ambos sistemas tienen solución única si y solo si las columnas de $\mathbf{A}$ son linealmente independientes.

### 13.1.2. Alternativa a un sistema de ecuaciones lineales sin solución

Imagine que tiene el siguiente sistema de ecuaciones

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{y} \qquad \text{donde} \quad \boldsymbol{y} \notin \mathcal{C}(\mathbf{A}).$$

Como ninguna combinación de las columnas de $\mathbf{A}$ es igual a $\boldsymbol{y}$, no es posible encontrar una solución.

Pero podemos plantearnos el siguiente *problema alternativo*: ¿Qué combinación lineal de las columnas de $\mathbf{A}$ está más próxima a $\boldsymbol{y}$? Es decir,

$$\mathbf{A}\boldsymbol{x} = \widehat{\boldsymbol{y}},$$

donde $\widehat{\boldsymbol{y}}$ es la proyección ortogonal de $\boldsymbol{y}$ sobre $\mathcal{C}(\mathbf{A})$. Este segundo problema siempre tiene solución, pues $\widehat{\boldsymbol{y}}$ existe y pertenece al subespacio $\mathcal{C}(\mathbf{A})$.

*Ésta es la idea detrás del ajuste mínimo cuadrático en estadística y econometría* (Véase la Parte **??**)

*Observación.* Cuando el *sistema de ecuaciones normales* tiene solución única, ésta *suele ser interpretada como la mejor "respuesta"* al sistema de ecuaciones irresoluble

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{y}; \quad \text{donde} \quad \boldsymbol{y} \notin \mathcal{C}(\mathbf{A});$$

pero recuerde que, *en realidad*, es la solución a un sistema de ecuaciones completamente distinto:

$$\mathbf{A}\boldsymbol{x} = \widehat{\boldsymbol{y}}; \quad \text{donde } \widehat{\boldsymbol{y}} \text{ es la proyección ortogonal de } \boldsymbol{y} \text{ sobre } \mathcal{C}(\mathbf{A}).$$

## 13.2. Expresión matricial de la proyección ortogonal

☞ Por último, vamos a encontrar una expresión matricial de la función proyección sobre un subespacio $\mathcal{V}$. Dicha expresión es la que se emplea en los manuales de econometría.

Como $\mathbf{A}\boldsymbol{x} = \widehat{\boldsymbol{y}}$ y $\left(\mathbf{A}^\intercal\mathbf{A}\right)\boldsymbol{x} = \mathbf{A}^\intercal\boldsymbol{y}$ son *sistemas de ecuaciones lineales equivalentes*, necesariamente sus matrices de coeficientes tienen el mismo espacio nulo $\mathcal{N}\left(\mathbf{A}\right) = \mathcal{N}\left(\mathbf{A}^\intercal\mathbf{A}\right)$; y como ambas matrices tienen $n$ columnas, también tienen el mismo rango.

Consecuentemente, si las $n$ columnas de $\mathbf{A}$ son linealmente independientes, entonces la matriz cuadrada $\mathbf{A}^\intercal\mathbf{A}$ es de rango completo (invertible) y *la solución* al sistema de ecuaciones normales $(\mathbf{A}^\intercal\mathbf{A})\boldsymbol{x} = \mathbf{A}^\intercal\boldsymbol{b}$ es

$$\widehat{\boldsymbol{x}} = \left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}\mathbf{A}^\intercal\boldsymbol{y}.$$

En tal caso, el vector *proyección* de $\boldsymbol{y}$ sobre $\mathcal{C}\left(\mathbf{A}\right)$ resulta ser

$$\widehat{\boldsymbol{y}} \;=\; \mathbf{A}\widehat{\boldsymbol{x}} \;=\; \mathbf{A}\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}\mathbf{A}^\intercal\boldsymbol{y}; \tag{13.2}$$

y el vector *proyección* de $\boldsymbol{y}$ sobre $\mathcal{N}\left(\mathbf{A}^\intercal\right)$ (que denominaremos $\boldsymbol{e}$) es

$$\boldsymbol{e} \;=\; \boldsymbol{y} - \widehat{\boldsymbol{y}} \;=\; \mathbf{I}\boldsymbol{y} - \mathbf{A}\widehat{\boldsymbol{x}} \;=\; \Big(\mathbf{I} - \mathbf{A}\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}(\mathbf{A}^\intercal)\Big)\boldsymbol{y}.$$

### 13.2.1. Matrices proyección

Ejercicio 70. Sea $\mathcal{A}$ un subespacio vectorial de $\mathbb{R}^n$. Encuentre $\mathbf{P}$ tal que $\mathbf{P}\boldsymbol{x} \in \mathcal{A}$ para todo $\boldsymbol{x} \in \mathbb{R}^n$.

Si $\mathbf{A}$ es de rango completo por columnas y $\mathbf{P}$ es la matriz $\mathbf{A}\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}\mathbf{A}^\intercal$ de la Ecuación 13.2, entonces

| | |
|---|---|
| $\widehat{\boldsymbol{y}} = \mathbf{P}\boldsymbol{y}$ | es la proyección de $\boldsymbol{y}$ sobre $\mathcal{C}\left(\mathbf{A}\right)$ y |
| $\boldsymbol{e} = \left(\mathbf{I} - \mathbf{P}\right)\boldsymbol{y}$ | es la proyección de $\boldsymbol{y}$ sobre $\mathcal{N}\left(\mathbf{A}^\intercal\right)$. |

Tanto $\mathbf{P}$ como $(\mathbf{I} - \mathbf{P})$ poseen dos propiedades íntimamente relacionadas con la proyección ortogonal:

- Ambas matrices son simétricas. Por una parte

$$\mathbf{P}^\intercal \;=\; \left(\mathbf{A}\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}\mathbf{A}^\intercal\right)^\intercal \;=\; \left(\mathbf{A}^\intercal\right)^\intercal\left((\mathbf{A}^\intercal\mathbf{A})^{-1}\right)^\intercal\mathbf{A}^\intercal \;=\; \mathbf{A}\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}\mathbf{A}^\intercal \;=\; \mathbf{P},$$

  pues la inversa de una matriz simétrica es simétrica: $\left((\mathbf{A}^\intercal\mathbf{A})^{-1}\right)^\intercal = \left((\mathbf{A}^\intercal\mathbf{A})^\intercal\right)^{-1} = \left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}$.

  Ejercicio 71. Por otra parte, demuestre que $\left(\mathbf{I} - \mathbf{P}\right)^\intercal = \mathbf{I} - \mathbf{P}$.

- Ambas matrices son idempotentes, es decir, son iguales a su cuadrado. Por una parte

$$\mathbf{P}^2 \;=\; \mathbf{P}\mathbf{P} \;=\; \mathbf{A}\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}\underbrace{\mathbf{A}^\intercal\mathbf{A}\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}}_{\mathbf{I}}\mathbf{A}^\intercal \;=\; \mathbf{A}\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}\mathbf{A}^\intercal = \mathbf{P}.$$

  Ejercicio 72. Por otra parte, demuestre que $(\mathbf{I} - \mathbf{P})^2 = \mathbf{I} - \mathbf{P}$.

De hecho, en la Sección 13.4 se demuestra que si una matriz $\underset{m\times m}{\mathbf{B}}$ es *simétrica* e *idempotente*, entonces

$$\begin{aligned} f_{\mathbf{B}} : \mathbb{R}^m &\longrightarrow \mathcal{C}\left(\mathbf{B}\right) \\ \boldsymbol{y} &\longmapsto \mathbf{B}\boldsymbol{y} \end{aligned}$$

es la función proyección ortogonal sobre $\mathcal{C}(\mathbf{B})$, es decir, $f_{\mathbf{B}}(\boldsymbol{y}) = \mathrm{Prj}_{\mathcal{C}(\mathbf{B})}(\boldsymbol{y})$. Por ello, las matrices simétricas e idempotentes se denominan *matrices proyección*. Consecuentemente $\mathbf{P}$ y $(\mathbf{I}-\mathbf{P})$ son las matrices proyección correspondientes a las respectivas proyecciones ortogonales sobre $\mathcal{C}(\mathbf{A})$ y $\mathcal{N}(\mathbf{A}^\intercal)$ (véase el siguiente ejercicio).

EJERCICIO 73. Sea $\mathbf{P} = \mathbf{A}\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}\mathbf{A}^\intercal$ y sea $\mathbf{M} = \mathbf{I} - \mathbf{P}$.

(a) Demuestre que $\mathcal{C}(\mathbf{P}) = \mathcal{C}(\mathbf{A})$.

(b) Demuestre que $\mathcal{C}(\mathbf{M}) = \mathcal{N}(\mathbf{A}^\intercal)$.

Resumiendo, si las columnas de $\mathbf{A}$ (de orden $m$ por $n$) son una base de $\mathcal{C}(\mathbf{A})$ y si $\mathbf{P} = \mathbf{A}\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}\mathbf{A}^\intercal$, entonces la función *proyección ortogonal* sobre $\mathcal{C}(\mathbf{A})$ es:

$$\begin{aligned} f_{\mathbf{P}} : \mathbb{R}^m &\longrightarrow \mathcal{C}(\mathbf{A}). \\ \boldsymbol{y} &\longmapsto \mathbf{P}\boldsymbol{y} \end{aligned}$$

Fíjese que, como la proyección $f_{\mathbf{P}}(\boldsymbol{y})$ de $\boldsymbol{y}$ sobre $\mathcal{C}(\mathbf{A})$ pertenece a $\mathcal{C}(\mathbf{A})$, el vector de $\mathcal{C}(\mathbf{A})$ más próximo a $f_{\mathbf{P}}(\boldsymbol{y})$ resulta ser él mismo; es decir, $f_{\mathbf{P}}(\widehat{\boldsymbol{y}}) = f_{\mathbf{P}}\left(f_{\mathbf{P}}(\boldsymbol{y})\right) = \mathbf{P}\mathbf{P}\boldsymbol{y} = \mathbf{P}\boldsymbol{y} = f_{\mathbf{P}}(\boldsymbol{y}) = \widehat{\boldsymbol{y}}$.

## 13.3. Bases ortogonales y proyección ortogonal. Coeficientes de Fourier

Considere $k$ vectores ortogonales $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots \boldsymbol{w}_k$ de $\mathbb{R}^n$. Encontrar la proyección $\widehat{\boldsymbol{b}}$ de $\boldsymbol{b} \in \mathbb{R}^n$ sobre el subespacio engendrado por los $k$ vectores $\boldsymbol{w}_i$ es muy sencillo cuando se cuenta con una base ortogonal.

Como $\widehat{\boldsymbol{b}} \in \mathcal{L}\left([\boldsymbol{w}_1; \ldots \boldsymbol{w}_k;]\right)$, sabemos que la proyección es de la forma

$$\widehat{\boldsymbol{b}} = c_1\boldsymbol{w}_1 + \cdots + c_k\boldsymbol{w}_k;$$

y, al ser $\widehat{\boldsymbol{b}}$ la proyección ortogonal, también sabemos que la diferencia $\boldsymbol{b} - \widehat{\boldsymbol{b}}$ es ortogonal a todo vector perteneciente a $\mathcal{L}\left([\boldsymbol{w}_1; \ldots \boldsymbol{w}_k;]\right)$. En consecuencia, para $i = 1 : k$ tenemos que

$$\begin{aligned} 0 =& (\boldsymbol{b} - \widehat{\boldsymbol{b}}) \cdot \boldsymbol{w}_i \\ =& \Big(\boldsymbol{b} - (c_1\boldsymbol{w}_1 + \cdots + c_k\boldsymbol{w}_k)\Big) \cdot \boldsymbol{w}_i \\ =& \boldsymbol{b}\cdot\boldsymbol{w}_i - c_1\boldsymbol{w}_1\cdot\boldsymbol{w}_i - \cdots - c_k\boldsymbol{w}_k\cdot\boldsymbol{w}_i \\ =& \boldsymbol{b}\cdot\boldsymbol{w}_i - c_i\boldsymbol{w}_i\cdot\boldsymbol{w}_i \qquad \text{puesto que } \boldsymbol{w}_p\cdot\boldsymbol{w}_i = 0 \text{ si } q \neq i; \end{aligned}$$

y despejando, vemos que los coeficientes $c_i$ de la proyección ortogonal $\widehat{\boldsymbol{b}} = c_1\boldsymbol{w}_1 + \cdots + c_k\boldsymbol{w}_k$ son de la forma

$$c_i = \frac{\boldsymbol{b}\cdot\boldsymbol{w}_i}{\boldsymbol{w}_i\cdot\boldsymbol{w}_i} = \frac{\boldsymbol{b}\cdot\boldsymbol{w}_i}{\|\boldsymbol{w}_i\|^2}.$$

Dichos coeficientes se denominan *coeficientes de Fourier*.[3]

## 13.4. Generalizaciones o visiones alternativas (★)

☞ Hemos definido las proyecciones ortogonales sobre subespacios de $\mathbb{R}^m$; también hemos visto que si las columnas de $\mathbf{A}$ son linealmente independientes, podemos definir la matriz *simétrica* e *idempotente* $\mathbf{P} = \mathbf{A}\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}\mathbf{A}^\intercal$. Entonces el producto $\mathbf{P}\boldsymbol{y}$ es la proyección ortogonal de $\boldsymbol{y}$ sobre $\mathcal{C}(\mathbf{A})$. De manera similar, podemos definir la matriz *simétrica* e *idempotente* $(\mathbf{I}-\mathbf{P})$. En este caso $(\mathbf{I}-\mathbf{P})\boldsymbol{y}$ es la proyección ortogonal de $\boldsymbol{y}$ sobre el complemento ortogonal de $\mathcal{C}(\mathbf{A})$.

Aquí vamos a generalizar todo esto a subespacios abstractos. Para lograrlo, primero debemos definir propiedades "equivalentes" a la idempotencia y simetría de matrices, pero para funciones.

[3] Fíjese que el parámetro $\alpha$ en la Ecuación 11.4 en la página 147 es un coeficiente de Fourier.

**Funciones idempotentes**

**Definición 13.3.** *Una función $f\colon V \to V$ se dice que es* idempotente *si verifica:*

$$f \circ f = f,$$

*es decir, para todo $x \in V$, $f\big(f(x)\big) = f(x)$.*

Dicho de otro modo, $f$ es idempotente si

$$f(x) = x \quad \text{para todo} \quad x \in \mathit{imagen}(f).$$

Fíjese que si $f_{\mathbf{B}}\colon \mathbb{R}^m \longrightarrow \mathbb{R}^m$ es *idempotente* implica que $\mathbf{B}$ es idempotente, pues
$\boldsymbol{x} \longmapsto \mathbf{B}\boldsymbol{x}$

$$f\big(f(\boldsymbol{x})\big) = f(\boldsymbol{x}) \quad \Leftrightarrow \quad \mathbf{B}\mathbf{B}\boldsymbol{x} = \mathbf{B}\boldsymbol{x}.$$

**Funciones autoadjuntas**

La propiedad equivalente a la simetría tiene un nombre diferente en el caso de las funciones.

**Definición 13.4.** *Una función $f\colon \mathcal{V} \to \mathcal{V}$, donde en $\mathcal{V}$ hay definido un producto escalar $\langle _ | _ \rangle$, se dice que es* autoadjunta *si verifica que*

$$\text{para todo } \vec{x}, \vec{y} \in \mathcal{V} \quad \langle \vec{x} | f(\vec{y}) \rangle = \langle f(\vec{x}) | \vec{y} \rangle.$$

En particular, para $\mathbb{R}^n$ con el producto punto, si $f_{\mathbf{B}}\colon \mathbb{R}^m \longrightarrow \mathbb{R}^m$ es *autoadjunta*, entonces:
$\boldsymbol{x} \longmapsto \mathbf{B}\boldsymbol{x}$

$$\boldsymbol{x}\mathbf{B}\boldsymbol{y} = \langle \boldsymbol{x} | \mathbf{B}\boldsymbol{y} \rangle = \langle \mathbf{B}\boldsymbol{x} | \boldsymbol{y} \rangle = \langle \boldsymbol{y} | \mathbf{B}\boldsymbol{x} \rangle = \boldsymbol{y}\mathbf{B}\boldsymbol{x};$$

de donde se deduce que $\mathbf{B}$ es necesariamente simétrica, pues si $\boldsymbol{x} = {}_{i|}\mathbf{I} = \mathbf{I}_{|i}$ e $\boldsymbol{y} = \mathbf{I}_{|j} = {}_{j|}\mathbf{I}$,

$${}_{i|}\mathbf{B}_{|j} = {}_{i|}\mathbf{I}\mathbf{B}\mathbf{I}_{|j} = {}_{j|}\mathbf{I}\mathbf{B}\mathbf{I}_{|i} = {}_{j|}\mathbf{B}_{|i}.$$

El recíproco también es cierto:

Ejercicio 74. Demuestre que si $\mathbf{B}$ es simétrica, entonces $f_{\mathbf{B}}\colon \mathbb{R}^m \longrightarrow \mathbb{R}^m$ es autoadjunta.
$\boldsymbol{x} \longmapsto \mathbf{B}\boldsymbol{x}$

### 13.4.1. Proyección ortogonal

☞ Vayamos ahora con las funciones proyección ortogonal en espacios pre-Hilbert abstractos.

**Definición 13.5.** *Sea $\big(\mathcal{V}, \langle _ | _ \rangle\big)$ un espacio pre-Hilbert. Llamamos proyección ortogonal a toda función $f\colon \mathcal{V} \to \mathcal{V}$ que además es* idempotente *y* autoadjunta*, es decir:* $\begin{cases} f \circ f = f \\ \langle \vec{x} | f(\vec{y}) \rangle = \langle f(\vec{x}) | \vec{y} \rangle \quad \text{para todo } \vec{x}, \vec{y} \in \mathcal{V}. \end{cases}$

A diferencia de la Definición 13.1, aquí no se requiere especificar el subespacio sobre el que se proyecta; dicho espacio es la imagen de la función $f$. Además, $f$ es lineal como se prueba en el siguiente ejercicio.

Ejercicio 75. (**Opcional**) Sea un espacio pre-Hilbert $\big(\mathcal{V}, \langle _ | _ \rangle\big)$ y sea una función $f\colon \mathcal{V} \to \mathcal{V}$ tal que $f \circ f = f$ (idempotente) y $\forall \vec{u}, \vec{v} \in \mathcal{V}$, $\langle f(\vec{u}) | \vec{v} \rangle = \langle \vec{u} | f(\vec{v}) \rangle$ (autoadjunta).

(a) Demuestre que $\langle f(\vec{u}) | f(\vec{w}) \rangle = \langle f(\vec{u}) | \vec{w} \rangle$.

(b) Demuestre que $f$ es lineal.

**Proposición 13.4.1.** *$f\colon \mathcal{V} \to \mathcal{V}$ es una proyección ortogonal (i.e., idempotente y autoadjunta) si y solo si para todo $\vec{h} \in \mathcal{V}$ y todo $\vec{y} \in imagen(f)$ se verifica que $\langle \vec{h} - f(\vec{h}) | \vec{y} \rangle = 0$.*

*Demostración.* Empecemos por demostrar que:

$$\boxed{f \text{ es proyección ortogonal } \Rightarrow \langle \vec{h} - f(\vec{h}) | \vec{y} \rangle = 0 \text{ para todo } \vec{h} \in \mathcal{V} \text{ y todo } \vec{y} \in imagen(f)}$$

$$\begin{aligned}
\langle \vec{h} - f(\vec{h}) | \vec{y} \rangle &= \langle \vec{h} | \vec{y} \rangle - \langle f(\vec{h}) | \vec{y} \rangle \\
&= \langle \vec{h} | \vec{y} \rangle - \langle \vec{h} | f(\vec{y}) \rangle && \text{por ser } f \text{ autoadjunta} \\
&= \langle \vec{h} | \vec{y} - f(\vec{y}) \rangle
\end{aligned}$$

y como $\vec{y} \in imagen(f)$ sabemos que $\vec{y} = f(\vec{z})$ para algún $\vec{z} \in \mathcal{V}$

$$\begin{aligned}
&= \langle \vec{h} | f(\vec{z}) - f\big(f(\vec{z})\big) \rangle && \text{sustituyendo } \vec{y} \text{ por } f(\vec{z}) \\
&= \langle \vec{h} | f(\vec{z}) - f(\vec{z}) \rangle && \text{por ser } f \text{ idempotente} \\
&= \langle \vec{h} | \vec{0} \rangle = 0.
\end{aligned}$$

Ahora demostremos que

$$\boxed{\langle \vec{h} - f(\vec{h}) | \vec{y} \rangle = 0 \text{ para todo } \vec{h} \in \mathcal{V} \text{ y todo } \vec{y} \in imagen(f) \Rightarrow f \text{ es proyección ortogonal}}$$

Demostremos que si $\langle \vec{h} - f(\vec{h}) | \vec{y} \rangle = 0$ para todo $\vec{h} \in \mathcal{V}$ y todo $\vec{y} \in imagen(f)$ entonces $f$ es idempotente y autoadjunta.

- Vamos a demostrar que $f(\vec{p}) = \vec{p}$ para todo $\vec{p} \in imagen(f)$, y que por tanto $f$ es idempotente.

  Reemplazando en la hipótesis la variable $\vec{h}$ por un vector $\vec{p} \in imagen(f)$, tenemos $\langle \vec{p} - f(\vec{p}) | \vec{y} \rangle = 0$ para todo $\vec{y} \in imagen(f)$, en particular también para $\vec{y} = \vec{p} - f(\vec{p})$:

  $$\langle \vec{p} - f(\vec{p}) | \vec{p} - f(\vec{p}) \rangle = 0,$$

  ya que $\big(\vec{p} - f(\vec{p})\big) \in imagen(f)$, pues $imagen(f)$ es subespacio por ser $f$ lineal, y $\vec{p}, f(\vec{p}) \in imagen(f)$.

  Por tanto $\vec{p} - f(\vec{p}) = \vec{0}$, es decir, $f(\vec{p}) = \vec{p}$; consecuentemente $f$ es <u>idempotente</u>.

- Para demostrar que es autoadjunta, por un lado tenemos que

  $$\begin{aligned}
  \langle \vec{v} - f(\vec{v}) | \vec{w} - f(\vec{w}) \rangle &= \langle \vec{v} | \vec{w} - f(\vec{w}) \rangle - \underbrace{\langle f(\vec{v}) | \vec{w} - f(\vec{w}) \rangle}_{=0} && \text{pues } f(\vec{v}) \in imagen(f) \\
  &= \langle \vec{v} | \vec{w} \rangle - \langle \vec{v} | f(\vec{w}) \rangle
  \end{aligned}$$

  por otro lado

  $$\begin{aligned}
  \langle \vec{v} - f(\vec{v}) | \vec{w} - f(\vec{w}) \rangle; &= \langle \vec{v} - f(\vec{v}) | \vec{w} \rangle - \underbrace{\langle \vec{v} - f(\vec{v}) | f(\vec{w}) \rangle}_{=0} && \text{pues } f(\vec{w}) \in imagen(f) \\
  &= \langle \vec{v} | \vec{w} \rangle - \langle f(\vec{v}) | \vec{w} \rangle.
  \end{aligned}$$

  Así que necesariamente $\langle \vec{v} | f(\vec{w}) \rangle = \langle f(\vec{v}) | \vec{w} \rangle$; es decir, $f$ es <u>autoadjunta</u>.

□

☞ Conocemos una condición suficiente para asegurar que existe la proyección ortogonal sobre $\mathcal{H}$, Veámoslo...

Por la Proposición 10.1.4 en la página 133, sabemos que si el subespacio $\mathcal{H} \subset \mathcal{V}$ y su conjunto ortogonal $\mathcal{H}^\perp$ son suplementarios, entonces para todo $\vec{v} \in \mathcal{V}$ existe una descomposición única $\vec{v} = \vec{h} + \vec{c}$ con $\vec{h} \in \mathcal{H}$ y $\vec{c} \in \mathcal{H}^\perp$. Es decir, si $\mathcal{V} = \mathcal{H} \oplus \mathcal{H}^\perp$ entonces existe proyección ortogonal tanto sobre $\mathcal{H}$ como sobre $\mathcal{H}^\perp$.

Así pues, por el Teorema 11.3.4 sabemos que si $\mathcal{H} \subset \big(\mathcal{V}, \langle _ | _ \rangle\big)$ es un subespacio de dimensión *finita* entonces $\mathcal{V} = \mathcal{H} \oplus \mathcal{H}^\perp$ y por tanto existe proyección ortogonal sobre $\mathcal{H}$ y sobre $\mathcal{H}^\perp$.

☞ Veamos ahora que, en el caso de que exista la función proyección ortogonal de $\vec{y} \in \mathcal{V}$ sobre $\mathcal{H} \subset \mathcal{V}$, la proyección del vector $\vec{y}$ es el vector de $\mathcal{H}$ "más próximo" a $\vec{y}$.

Después, y para finalizar la lección, veremos un ejemplo de subespacio de dimensión *infinita* en el que no existe la función proyección ortogonal, pues no existe el vector "más próximo" a uno dado.

Ejercicio 76. Demuestre la siguiente proposición:

**Proposición 13.4.2.** *Sea $\big(\mathcal{V}, \langle _ | _ \rangle\big)$ un espacio pre-Hilbert. En el caso de que exista la proyección ortogonal de $\vec{y} \in \mathcal{V}$ sobre $\mathcal{H} \subset \mathcal{V}$, dicha proyección de $\vec{y}$ es el vector de $\mathcal{H}$ más próximo a $\vec{y}$.*

**Sobre algunos subespacios de dimensión infinita no existe la proyección ortogonal**

Hay subespacios $\mathcal{H}$ de dimensión infinita en espacios pre-Hilbert $\mathcal{V}$ donde no hay función proyección ortogonal... es decir, donde no existe "el vector más próximo" en $\mathcal{H}$ a un vector $\vec{y} \in \mathcal{V}$ dado.

Un ejemplo son las sucesiones casi nulas $\mathbb{R}[x]$ de la Página 156 y el subespacio $\mathcal{H} \subset \mathbb{R}[x]$ de sucesiones cuyas componentes suman cero:

$$\mathcal{H} = \left\{ p \in \mathbb{R}[x] \;\middle|\; \sum_i p_i = 0 \right\}.$$

Es fácil ver que, aunque podemos aproximarnos tanto como queramos al vector $\vec{a} = (1,0,0,0,0,0,0,0\ldots)$ con vectores del subespacio $\mathcal{H}$, no existe un vector en $\mathcal{H}$ que sea "el más próximo" a $\vec{a}$.

Para $\vec{b} = (1,-1,0,0,0,0,0,0\ldots)$, tenemos que

$$d(\vec{a},\vec{b}) = \|\vec{a}-\vec{b}\|^2 = \|(0,-1,0,0,0,0,0,0\ldots)\|^2 = (-1)^2 = 1.$$

Para $\vec{c} = \big(1,-\frac{1}{2},-\frac{1}{2},0,0,0,0,0\ldots\big)$, tenemos que

$$d(\vec{a},\vec{c}) = \|\vec{a}-\vec{c}\|^2 = \|\big(0,-\tfrac{1}{2},-\tfrac{1}{2},0,0,0,0,0\ldots\big)\|^2 = \tfrac{1}{2^2}+\tfrac{1}{2^2} = \tfrac{1}{2}.$$

Para $\vec{d} = \big(1,-\frac{1}{3},-\frac{1}{3},-\frac{1}{3},0,0,0,0\ldots\big)$, tenemos que

$$d(\vec{a},\vec{d}) = \|\vec{a}-\vec{d}\|^2 = \|\big(0,-\tfrac{1}{3},-\tfrac{1}{3},-\tfrac{1}{3},0,0,0,0\ldots\big)\|^2 = \tfrac{1}{3^2}+\tfrac{1}{3^2}+\tfrac{1}{3^2} = \tfrac{1}{3}.$$

Para $\vec{e} = \big(1,-\frac{1}{4},-\frac{1}{4},-\frac{1}{4},-\frac{1}{4},0,0,0\ldots\big)$, tenemos que

$$d(\vec{a},\vec{e}) = \|\vec{a}-\vec{e}\|^2 = \|\big(0,-\tfrac{1}{4},-\tfrac{1}{4},-\tfrac{1}{4},-\tfrac{1}{4},0,0,0\ldots\big)\|^2 = \tfrac{1}{4^2}+\tfrac{1}{4^2}+\tfrac{1}{4^2}+\tfrac{1}{4^2} = \tfrac{1}{4}.$$

Y, en general, para $\vec{z} = \big(1,\underbrace{-\frac{1}{n},-\frac{1}{n},\ldots,-\frac{1}{n}}_{n \text{ componentes}},0\ldots\big)$, tenemos que

$$d(\vec{a},\vec{z}) = \|\vec{a}-\vec{z}\|^2 = \|\big(0,\underbrace{-\tfrac{1}{n},-\tfrac{1}{n},\ldots,-\tfrac{1}{n}}_{n \text{ componentes}},0\ldots\big)\|^2 = \tfrac{1}{n}.$$

Es decir, podemos aproximarnos al vector $\vec{a}$ tanto como queramos. Sin embargo, la distancia nunca puede ser cero, pues $d(\vec{a},\vec{z})=0 \;\Rightarrow\; \vec{a}=\vec{z}$, y esto es imposible, pues entonces $\vec{a}$ pertenecería a $\mathcal{H}$, algo que no puede ocurrir puesto que sus componentes no suman cero.

Como no existe el vector $\vec{z}$ más próximo a $\vec{a}$ en $\mathcal{H}$, no existe la proyección ortogonal sobre $\mathcal{H}$.[4]

### 13.4.2. Bases ortogonales y proyección ortogonal. Coeficientes de Fourier

Sea $\big(\mathcal{V},\langle_|_\rangle\big)$ un espacio pre-Hilbert y considere $k$ vectores ortogonales $\vec{w}_1,\vec{w}_2,\dots\vec{w}_k$ de $\mathcal{V}$. Encontrar la proyección $\widehat{\vec{a}}$ de $\vec{a}\in\mathcal{V}$ sobre el subespacio engendrado por los $k$ vectores $\vec{w}_i$ es tan sencillo como en $\mathbb{R}^n$: como $\widehat{\vec{a}}\in\mathcal{L}\big([\vec{w}_1;\dots\vec{w}_k;]\big)$, sabemos que la proyección es de la forma

$$\widehat{\vec{a}}=c_1\vec{w}_1+\cdots+c_k\vec{w}_k;$$

y, por ser $\widehat{\vec{a}}$ la proyección ortogonal, también sabemos que la diferencia $\vec{a}-\widehat{\vec{a}}$ es ortogonal a todo vector perteneciente a $\mathcal{L}\big(\{\vec{w}_1,\dots\vec{w}_k\}\big)$. En consecuencia, para $i=1:k$ tenemos que

$$\begin{aligned}0=\Big\langle \vec{a}-\widehat{\vec{a}}\,\Big|\,\vec{w}_i\Big\rangle &=\Big\langle \vec{a}-(c_1\vec{w}_1+\cdots+c_k\vec{w}_k)\,\Big|\,\vec{w}_i\Big\rangle\\ &=\langle\vec{a}|\vec{w}_i\rangle-c_1\langle\vec{w}_1|\vec{w}_i\rangle-\cdots-c_k\langle\vec{w}_k|\vec{w}_i\rangle\\ &=\langle\vec{a}|\vec{w}_i\rangle-c_i\langle\vec{w}_i|\vec{w}_i\rangle \qquad\qquad \text{puesto que } \langle\vec{w}_q|\vec{w}_i\rangle=0 \text{ si } q\neq i;\end{aligned}$$

y despejando, vemos que los coeficientes $c_i$ de la proyección ortogonal $\widehat{\vec{a}}=c_1\vec{w}_1+\cdots+c_k\vec{w}_k$ son de la forma

$$c_i=\frac{\langle\vec{a}|\vec{w}_i\rangle}{\langle\vec{w}_i|\vec{w}_i\rangle}=\frac{\langle\vec{a}|\vec{w}_i\rangle}{\|\vec{w}_i\|^2}.$$

Dichos coeficientes se denominan *coeficientes de Fourier* de la proyección ortogonal sobre la base ortogonal.

## 13.5. Ejercicios

EJERCICIO 77. Sean $\mathbf{P}=\mathbf{A}\big(\mathbf{A}^\intercal\mathbf{A}\big)^{-1}\mathbf{A}^\intercal$ y $\mathbf{M}=\mathbf{I}-\mathbf{P}$, con $\mathbf{A}$ de orden $m$ por $n$ y rango $n$.

(a) Demuestre que para todo $\boldsymbol{y}\in\mathbb{R}^m$, la diferencia $\boldsymbol{y}-\mathbf{P}\boldsymbol{y}$ es perpendicular a $\mathcal{C}(\mathbf{A})$.
(b) Demuestre que $f_{\mathbf{P}}$ actúa como la función identidad con todo $\boldsymbol{y}\in\mathcal{C}(\mathbf{A})$.
(c) Demuestre que $f_{\mathbf{P}}$ actúa como la función nula con todo $\boldsymbol{y}\in\mathcal{N}\big(\mathbf{A}^\intercal\big)$.
(d) Demuestre que $\begin{array}{rcl} f_{\mathbf{M}}:\mathbb{R}^m & \longrightarrow & \mathcal{N}\big(\mathbf{A}^\intercal\big)\\ \boldsymbol{y} & \longmapsto & \mathbf{M}\boldsymbol{y}\end{array}$ es la función *proyección ortogonal* sobre $\mathcal{N}\big(\mathbf{A}^\intercal\big)$.

*Pista.* Fíjese en que $f_{\mathbf{M}}(\boldsymbol{y})=\mathbf{M}\boldsymbol{y}=(\mathbf{I}-\mathbf{P})\boldsymbol{y}=\boldsymbol{y}-\mathbf{P}\boldsymbol{y}=\boldsymbol{y}-f_{\mathbf{P}}(\boldsymbol{y})$ y también en que $\mathcal{N}\big(\mathbf{A}^\intercal\big)^\perp=\mathcal{C}(\mathbf{A})$.

(e) Demuestre también que $f_{\mathbf{M}}$ actúa como la función identidad con todo $\boldsymbol{y}\in\mathcal{N}\big(\mathbf{A}^\intercal\big)$ y que actúa como la función nula con todo $\boldsymbol{y}\in\mathcal{C}(\mathbf{A})$.

*Aquí hemos explotado la equivalencia de dos sistemas de ecuaciones y la definición de proyección ortogonal. En otros manuales se sigue un argumento distinto que básicamente consiste en demostrar las siguientes proposiciones* (que propongo como ejercicio *opcional*, pues son resultados que ya hemos visto más arriba).

EJERCICIO 78. (*Opcional*) Demuestre las siguientes proposiciones:

(a) **Proposición 13.5.1.** $\mathcal{N}(\mathbf{A})=\mathcal{N}\big(\mathbf{A}^\intercal\mathbf{A}\big)$.
(b) **Corolario 13.5.2.** *El rango de* $\mathbf{A}^\intercal\mathbf{A}$ *es igual al rango de* $\mathbf{A}$.
(c) **Proposición 13.5.3.** *El sistema de ecuaciones normales* $\big(\mathbf{A}^\intercal\mathbf{A}\big)\boldsymbol{x}=\mathbf{A}^\intercal\boldsymbol{y}$ *tiene solución.*

[4] Para garantizar que existe la proyección ortogonal en un espacio pre-Hilbert $\mathcal{V}$, se necesita que además sea *completo*; es decir, que toda sucesión de Cauchy de elementos de $\mathcal{V}$ converja en norma a algún elemento del propio espacio $\mathcal{V}$. En tal caso $\mathcal{V}$ es un espacio de Hilbert (los espacios de Hilbert son muy importantes en Probabilidad, Estadística, Econometría o Series Temporales).

*Pista.* Compruebe que la matriz de coeficientes y la matriz ampliada tienen el mismo rango.

*En econometría se emplea el siguiente resultado*

EJERCICIO 79. Si una matriz $\mathbf{Q}$ de orden $n$ es idempotente, entonces su rango es igual a su traza. (encontrará el significado de *traza* en la Definición 3.7 en la página 45).

# Parte V

# Determinantes

# Propiedades de los determinantes

## 14.1. Función determinante y función volumen

### 14.1.1. Tres propiedades de la función volumen de un paralelogramo

La función volumen[1] (*base* × *altura*) de un paralelepípedo de dimensión $n$, cuyas aristas corresponden a las columnas de una matriz cuadrada $\mathbf{A}$, tiene las siguientes tres propiedades:

- Dado que las columnas de la matriz identidad tienen norma uno ($\|\mathbf{I}_{|j}\| = 1$), el volumen (o área) del hipercubo de dimensión $n$ descrito por las columnas de $\mathbf{I}$ es siempre uno:

$$\text{Volumen}\Big(\underset{n\times n}{\mathbf{I}}\Big) = 1.$$

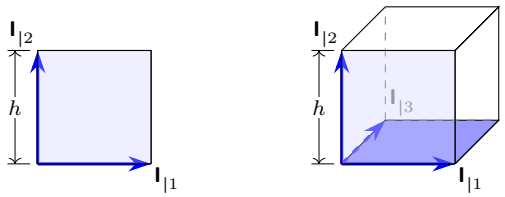



Figura 14.1: Área$\big[\mathbf{I}_{|1};\ \mathbf{I}_{|2};\big] = 1$ (izquierda) y Volumen$\big[\mathbf{I}_{|1};\ \mathbf{I}_{|2};\ \mathbf{I}_{|3};\big] = 1$ (derecha).

- Al sumar a una de las aristas (columnas) un múltiplo de otra arista (columna), el volumen (o el área) del paralelogramo resultante no cambia (ya que se mantiene la misma base y la misma altura $h$). Por tanto, aplicar una transformación elemental de Tipo I no modifica el volumen:

$$\text{Vol}(\mathbf{A}) \ = \ \text{Vol}\Big(\mathbf{A}\underset{[(\alpha)\boldsymbol{k}+\boldsymbol{j}]}{\boldsymbol{\tau}}\Big) \text{ para } i \neq k.$$

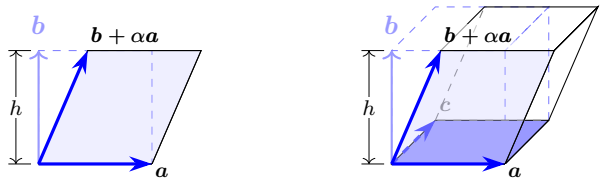



Figura 14.2: Área$\big[\boldsymbol{a};\ \boldsymbol{b};\big] =$ Área$\big[\boldsymbol{a};\ (\alpha\boldsymbol{a}+\boldsymbol{b});\big]$ (izquierda) y Vol$\big[\boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c};\big] =$ Vol$\big[\boldsymbol{a};\ (\alpha\boldsymbol{a}+\boldsymbol{b});\ \boldsymbol{c};\big]$ (derecha).

[1] en dimensión 2 se denomina área

- Al multiplicar una de las aristas (columnas) por un escalar, el volumen (o área) queda multiplicado por el valor absoluto de dicho escalar (por tanto no hay áreas o volúmenes negativos).

$$|\alpha| \cdot \mathrm{Vol}(\mathbf{A}) \;=\; |\alpha| \cdot \mathrm{Vol}\big[\ldots \mathbf{A}_{|k}; \ldots\big] \;=\; \mathrm{Vol}\big[\ldots \alpha\mathbf{A}_{|k}; \ldots\big]$$

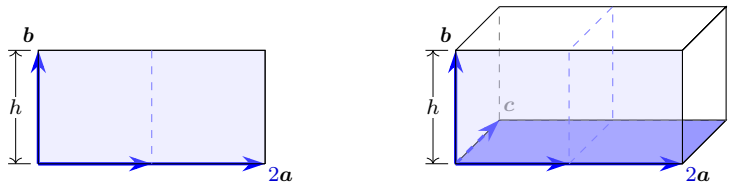



Figura 14.3: $|\alpha| \cdot \text{Área}\big[\boldsymbol{a};\ \boldsymbol{b};\big] = \text{Área}\big[\alpha\boldsymbol{a};\ \boldsymbol{b};\big]$ (izquierda) y $|\alpha| \cdot \mathrm{Vol}\big[\boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c};\big] = \mathrm{Vol}\big[\alpha \cdot \boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c};\big]$ (derecha).

En consecuencia, aplicar la transformación, $\underset{[(\alpha)\boldsymbol{j}]}{\boldsymbol{\tau}}$, donde $\alpha \neq 0$, multiplica el volumen por $|\alpha|$. Y por otra parte, multiplicar una arista por 0 multiplica el volumen por 0.

### 14.1.2. Las tres primeras propiedades (P-1 a P-3)

Por analogía con la función volumen, definimos la función determinante del siguiente modo

**Definición 14.1.** *Denominamos* función determinante *a toda función que asigna a cada sistema de n vectores de* $\mathbb{R}^n$ *(a cada matriz cuadrada de orden n) un número real*

$$\det: \quad \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$$

*si verifica las siguientes tres propiedades:*

***P-1*** **P-1** **El determinante de las matrices identidad es uno:**

$$\det \mathbf{I} = 1$$

***P-2*** **P-2** **Sumar a una columna un múltiplo de *otra* no altera el determinante:**

$$\det \mathbf{A} = \det\left(\mathbf{A}\ \underset{[(\alpha)\boldsymbol{k}+\boldsymbol{j}]}{\boldsymbol{\tau}}\right) \quad \textit{para } i \neq k.$$

***P-3*** **P-3** **Multiplicar una columna por $\alpha$ multiplica el determinante por $\alpha$**

$$\alpha \cdot \det \mathbf{A} \;=\; \det\big[\ldots \alpha\mathbf{A}_{|k}; \ldots\big] \textit{ para cualquier } k \in \{1:n\} \textit{ y } \alpha \in \mathbb{R}$$

(consecuentemente, los determinantes pueden tomar valores negativos).

Así, por ***P-3***, aplicar una transformación elemental de Tipo II, $\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}$, multiplica el determinante por $\alpha \neq 0$.

**La relación entre la función *Volumen* y la función *Determinante***

Nótese que la única diferencia con la función *Volumen* es que los determinantes pueden tomar valores negativos. De hecho, la relación entre la función *Volumen* y la función *Determinante* es tan estrecha que podemos definir la función *Volumen* del paralelogramo cuyas aristas son las columnas de la matriz cuadrada $\mathbf{A}$ como el valor absoluto del *Determinante* de $\mathbf{A}$:

$$\boxed{\mathrm{Vol}\,(\mathbf{A}) = \text{Valor absoluto del determinante de } \mathbf{A}}.$$

**Notación** Emplearemos dos notaciones alternativas para el determinante de la matriz $\mathbf{A}$:

$$\text{determinante de } \mathbf{A} \equiv \det(\mathbf{A}) \equiv |\mathbf{A}|$$

de manera que las tres propiedades anteriores se pueden expresar para matrices de orden 3 como:

**P-1:** $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1;$ $\qquad \det\left[\mathbf{I}_{|1};\ \mathbf{I}_{|2};\ \mathbf{I}_{|3};\right] = 1.$

**P-2:** $\begin{vmatrix} a_1 & (b_1+\alpha c_1) & c_1 \\ a_2 & (b_2+\alpha c_2) & c_2 \\ a_3 & (b_3+\alpha c_3) & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix};$ $\qquad \det\left[\boldsymbol{a};\ (\boldsymbol{b}+\alpha\boldsymbol{c});\ \boldsymbol{c};\right] = \det\left[\boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c};\right]$

**P-3:** $\begin{vmatrix} a_1 & (\beta b_1) & c_1 \\ a_2 & (\beta b_2) & c_2 \\ a_3 & (\beta b_3) & c_3 \end{vmatrix} = \beta \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix};$ $\qquad \det\left[\boldsymbol{a};\ \beta\boldsymbol{b};\ \boldsymbol{c};\right] = \beta\, \det\left[\boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c};\right].$

**Advertencia:** Una barra vertical a cada lado de una matriz $|\mathbf{A}|$ denota el determinante de $\mathbf{A}$. Una barra vertical a cada lado de un número $|a|$ significa valor absoluto del número. Es decir, el significado de las barras verticales depende del objeto situado entre ellas: cuando es un número denota el *valor absoluto* y cuando es una matriz denota su *determinante*. Jugando con esto, podemos decir que

$$\text{Vol}\left(\mathbf{A}\right) = \text{Valor absoluto de } \det\left(\mathbf{A}\right) = |\det \mathbf{A}| = \big||\mathbf{A}|\big|.$$

☞ Las propiedades ***P-1***, ***P-2*** y ***P-3*** definen la función determinante. Ahora queda deducir todo lo demás...

## 14.2. Seis propiedades del determinante

### 14.2.1. Determinante de una matriz con una columna de ceros

EJERCICIO 80. Demuestre la siguiente propiedad:

**P-4** **Determinante de una matriz con una columna nula.** **P-4**

*Si* $\mathbf{A}$ tiene una columna nula, entonces:

$$\boxed{\det(\mathbf{A}) = 0}$$

Esta propiedad es análoga a considerar el volumen de un paralelogramo con una arista de longitud cero.

### 14.2.2. Determinantes de matrices elementales

Ya sabemos que

$$\det\left(\mathbf{A}_{\ \boldsymbol{\tau}\ [(\alpha)\boldsymbol{k}+\boldsymbol{j}]}\right) = |\mathbf{A}|; \qquad \det\left(\mathbf{A}_{\ \boldsymbol{\tau}\ [(\alpha)\boldsymbol{k}]}\right) = \alpha|\mathbf{A}|; \tag{14.1}$$

así, y puesto que $\det\left(\mathbf{I}\right) = 1$, en particular tenemos que:

$$\det\left(\mathbf{I}_{\ \boldsymbol{\tau}\ [(\alpha)\boldsymbol{k}+\boldsymbol{j}]}\right) = 1 \qquad \text{y} \qquad \det\left(\mathbf{I}_{\ \boldsymbol{\tau}\ [(\alpha)\boldsymbol{j}]}\right) = \alpha.$$

Recordando que $\mathbf{A}_{\boldsymbol{\tau}} = \mathbf{A}(\mathbf{I}_{\boldsymbol{\tau}})$, para los dos casos de (14.1), concluimos que

$$|\mathbf{A}(\mathbf{I}_{\boldsymbol{\tau}})| = |\mathbf{A}| \cdot |\mathbf{I}_{\boldsymbol{\tau}}|; \tag{14.2}$$

es decir, no importa de qué tipo sea la matriz elemental $\mathbf{I}_{\boldsymbol{\tau}}$.

### 14.2.3. Sucesión de transformaciones elementales por columnas

Veamos más resultados relacionados con las matrices elementales.

EJERCICIO 81. Demuestre las siguientes proposiciones.

(a) $\det\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = |\mathbf{A}|\cdot|\mathbf{I}_{\boldsymbol{\tau}_1}|\cdots|\mathbf{I}_{\boldsymbol{\tau}_k}|$ .

(b) Si $\mathbf{B}$ es de rango completo, es decir, si $\mathbf{B} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$, entonces $|\mathbf{B}| = |\mathbf{I}_{\boldsymbol{\tau}_1}|\cdots|\mathbf{I}_{\boldsymbol{\tau}_k}|$, y por tanto $|\mathbf{B}| \neq 0$.

(c) *Si* $\mathbf{A}$ *y* $\mathbf{B}$ *son de orden* $n$*, y* $\mathbf{B}$ *es de rango completo entonces*

$$\det(\mathbf{AB}) = |\mathbf{A}|\cdot|\mathbf{B}| \tag{14.3}$$

*Ejemplo* 35. Una sucesión $\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k$ de transformaciones *Tipo I* de la matriz $\mathbf{A}$ no altera el determinante.

$$|\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}| = |\mathbf{A}\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}| = |\mathbf{A}|\cdot|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}| = |\mathbf{A}|\cdot 1 = |\mathbf{A}|$$

*Ejemplo* 36. Pero una sucesión de transformaciones *Tipo II* sí puede modificar el valor del determinante:

$$\begin{vmatrix} 2a & 3c \\ 2b & 3d \end{vmatrix} = 6\begin{vmatrix} a & c \\ b & d \end{vmatrix}$$

#### Intercambio de columnas. Propiedad antisimétrica

EJERCICIO 82. Demuestre la siguiente propiedad:

**P-5** P-5 **[Propiedad antisimétrica]** *Intercambiar dos columnas cambia el signo del determinante.*

Por ejemplo,

$$\begin{vmatrix} c_1 & b_1 & a_1 \\ c_2 & b_2 & a_2 \\ c_3 & b_3 & a_3 \end{vmatrix} = (-1)\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

### 14.2.4. Matriz singular, matriz inversa y producto de matrices

#### Determinante de matrices singulares y determinante de matrices inversas

Vamos a demostrar conjuntamente dos propiedades más de la función determinante:

**P-6** P-6 **Si $\mathbf{A}$ es singular, entonces $|\mathbf{A}| = 0$.**

**P-7** P-7 $\det(\mathbf{A}^{-1}) = \left(\det\mathbf{A}\right)^{-1}$.

*Demostración.* Sea $\underset{n\times n}{\mathbf{A}}$ y sea $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ tal que $\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \mathbf{R}$ es una forma escalonada reducida. Entonces $|\mathbf{A}|\cdot|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}| = |\mathbf{R}|$ y solo se pueden dar dos casos:

$$\begin{cases} \mathbf{A} \text{ singular } (\mathbf{R}_{|n} = \mathbf{0}\,): & |\mathbf{R}| = |\mathbf{A}|\cdot|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}| = 0 \quad \Rightarrow \quad |\mathbf{A}| = 0 \\ \mathbf{A} \text{ invertible } (\mathbf{R} = \mathbf{I}): & |\mathbf{R}| = |\mathbf{A}|\cdot|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}| = 1 \quad \Rightarrow \quad |\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}| = |\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|} \end{cases}.$$

□

Así, podemos calcular el determinante de la matriz inversa mediante la eliminación gaussiana.

- Cuando la matriz es singular (rango $< n$) el determinante es cero;
- Cuando la matriz es de rango completo, basta con mirar qué transformaciones elementales *Tipo II* han ido cambiando el valor del determinante de $\mathbf{A}$ hasta llegar a la matriz $\mathbf{I}$ (las de *Tipo I* no importan!...)

*Ejemplo* 37. Sea $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}$, entonces

$$\left[\begin{array}{cc} 1 & 2 \\ 2 & 2 \\ \hline 1 & 0 \\ 0 & 1 \end{array}\right] \xrightarrow[(TipoI)]{\tau_{[(-2)\mathbf{1}+\mathbf{2}]}} \left[\begin{array}{cc} 1 & 0 \\ 2 & -2 \\ \hline 1 & -2 \\ 0 & 1 \end{array}\right] \xrightarrow[(TipoII)]{\tau_{[(-1/2)\mathbf{2}]}} \left[\begin{array}{cc} 1 & 0 \\ 2 & 1 \\ \hline 1 & 1 \\ 0 & -1/2 \end{array}\right] \xrightarrow[(TipoI)]{\tau_{[(-2)\mathbf{2}+\mathbf{1}]}} \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline -1 & 1 \\ 1 & -1/2 \end{array}\right]$$

Por tanto, $|\mathbf{A}^{-1}| = |\mathbf{I}_{\tau_{[(-2)\mathbf{1}+\mathbf{2}]}}| \cdot |\mathbf{I}_{\tau_{[(-1/2)\mathbf{2}]}}| \cdot |\mathbf{I}_{\tau_{[(-2)\mathbf{2}+\mathbf{1}]}}| = 1 \cdot \dfrac{-1}{2} \cdot 1 = \dfrac{-1}{2}$; es decir $\boxed{|\mathbf{A}| = -2.\,.}$

**Determinante del producto.**

Puesto que para cualquier $\mathbf{B}$ de orden $n$, existe una sucesión de transformaciones elementales tal que $\mathbf{B}\big(\mathbf{I}_{\tau_1 \cdots \tau_k}\big) = \mathbf{L}$ es una forma escalonada de $\mathbf{B}$, si $\mathbf{B}$ es singular, entonces $\mathbf{L}_{|n} = \mathbf{0}$. Aplicando las mismas transformaciones elementales a las columnas de $\mathbf{AB}$ tenemos

$$\big(\mathbf{AB}(\mathbf{I}_{\tau_1 \cdots \tau_k})\big)_{|n} = \big(\mathbf{AL}\big)_{|n} = \mathbf{0};$$

y como $\big(\mathbf{I}_{\tau_1 \cdots \tau_k}\big)_{|n} \neq \mathbf{0}$, necesariamente $(\mathbf{AB})$ es singular (véase Corolario 8.1.2 en la página 114).

Este resultado nos permite enunciar una nueva propiedad de la función determinante.

**P-8** **[Determinante del producto de matrices]** **P-8**

$$\boxed{\det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B}).} \tag{14.4}$$

*Demostración.*
$$\begin{cases} \text{Si } \mathbf{B} \text{ es singular, también lo es } \mathbf{AB} \Rightarrow & \det(\mathbf{AB}) = 0 = \det(\mathbf{A}) \cdot \det(\mathbf{B}), \quad \text{pues } |\mathbf{B}| = 0 \\ \text{Si } \mathbf{B} = \mathbf{I}_{\tau_1 \cdots \tau_k} \Rightarrow & \det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B}) \qquad \text{(Ecuación (14.3))} \end{cases}$$
□

### 14.2.5. Determinante de la matriz transpuesta.

EJERCICIO 83.

(a) ¿Qué relación existe entre el determinante de $\mathbf{I}_\tau$ y el determinante de su transpuesta ${}_\tau\mathbf{I}$?
(b) Sea $\mathbf{B}$ de rango completo, demuestre que $|\mathbf{B}| = |\mathbf{B}^\intercal|$.

Así llegamos a la última propiedad que veremos en esta lección:

**P-9** **Determinante de la transpuesta** **P-9**

$$|\mathbf{A}| = |\mathbf{A}^\intercal|.$$

*Demostración.* Por una parte, si $\mathbf{A}$ es singular, $\mathbf{A}^\intercal$ también es singular (véase Proposición 9.5.13 en la página 127); y por tanto ambas matrices tienen determinante nulo. Y por otra, si $\mathbf{A}$ es de rango completo, acabamos de ver en el ejercicio anterior que su determinante es igual al de su transpuesta. □

# Fórmulas para el cálculo del determinante y aplicaciones

**Definición 15.1.** *Llamaremos* matriz extendida *de* $\underset{n\times n}{\mathbf{A}}$ *a la matriz* $\begin{bmatrix}\mathbf{A} & \\ & 1\end{bmatrix}$ *de orden* $n+1$.

Si $\mathbf{B}_{\boldsymbol{\tau}}$ es de orden $n$, entonces $\boldsymbol{\tau}$ realiza una transformación elemental sobre alguna de las $n$ columnas de $\mathbf{B}$; es decir, tanto si $\boldsymbol{\tau}$ es de *Tipo I*: $\underset{[(\alpha)\boldsymbol{p}+\boldsymbol{s}]}{\boldsymbol{\tau}}$ como si es de *Tipo II*: $\underset{[(\alpha)\boldsymbol{p}]}{\boldsymbol{\tau}}$, los índices $\boldsymbol{p}$ y $\boldsymbol{s}$ son menores o iguales a $n$.

Por tanto, para las matrices extendidas de $\mathbf{B}_{\boldsymbol{\tau}}$ y de $\mathbf{B}$ necesariamente se verifica que

$$\begin{bmatrix}\mathbf{B}_{\boldsymbol{\tau}} & \\ & 1\end{bmatrix} = \begin{bmatrix}\mathbf{B} & \\ & 1\end{bmatrix}_{\boldsymbol{\tau}}. \tag{15.1}$$

Además, tanto $\begin{bmatrix}\mathbf{I} & \\ & 1\end{bmatrix}_{\boldsymbol{\tau}}$ como $\mathbf{I}_{\boldsymbol{\tau}}$ *tienen idéntico determinante* por ser matrices elementales del mismo tipo. En consecuencia, si $\mathbf{B}$ es invertible, repitiendo $k$ veces el paso dado en (15.1), deducimos que

$$\left|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \quad \begin{matrix}\\ 1\end{matrix}\right| = \det\left(\begin{bmatrix}\mathbf{I} & \\ & 1\end{bmatrix}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \det\left(\begin{bmatrix}\mathbf{I} & \\ & 1\end{bmatrix}_{\boldsymbol{\tau}_1}\cdots\begin{bmatrix}\mathbf{I} & \\ & 1\end{bmatrix}_{\boldsymbol{\tau}_k}\right) = |\mathbf{I}_{\boldsymbol{\tau}_1}|\cdots|\mathbf{I}_{\boldsymbol{\tau}_k}| = |\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}|. \tag{15.2}$$

Vamos a emplear este resultado para relacionar el determinante de dos matrices de distinto orden:

**Proposición 15.0.1.** *Si* $\mathbf{A}$ *es la matriz extendida de* $\mathbf{B}$*, entonces* $\boxed{\det(\mathbf{A}) = \det(\mathbf{B}).}$

*Demostración.* Solo caben dos posibilidades. Si $\mathbf{B}$ es de rango completo, es decir, si $\mathbf{B} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$, por (15.2) ya sabemos que

$$|\mathbf{A}| = \left|\mathbf{B} \quad \begin{matrix}\\ 1\end{matrix}\right| = \left|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \quad \begin{matrix}\\ 1\end{matrix}\right| = |\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}| = |\mathbf{B}|.$$

Y si $\mathbf{B}$ es singular, como las columnas de $\mathbf{B}$ son linealmente dependientes, también lo son las de $\mathbf{A}$. Por tanto

$$|\mathbf{A}| = 0 = |\mathbf{B}|.$$

□

## 15.1. Determinante de matrices triangulares

Ejercicio 84.

(a) ¿A qué es igual el determinante de una matriz triangular inferior $\mathbf{L}$ de rango completo?
(b) ¿A qué es igual el determinante de una matriz triangular con algún elemento nulo en su diagonal?
(c) ¿A qué es igual el determinante de una matriz triangular superior $\mathbf{U}$?

## 15.2. Cálculo del determinante por eliminación Gaussiana

Ahora sabemos que el determinante de una matriz cuadrada y triangular es igual al producto de los elementos de la diagonal principal. Este descubrimiento nos permite calcular de manera sencilla el determinante mediante eliminación: basta con encontrar una forma escalonada de la matriz extendida, compensando en su última fila las transformaciones elementales Tipo II aplicadas sobre las $n$ primeras columnas.

*Ejemplo* 38. Considere la matriz $\left[\begin{array}{cc} 1 & 5 \\ 2 & 3 \end{array}\right]$ entonces

$$\left[\begin{array}{cc|c} 1 & 5 & 0 \\ 2 & 3 & 0 \\ \hline 0 & 0 & 1 \end{array}\right] \xrightarrow{\boldsymbol{\tau}\,[(-5)\mathbf{1}+\mathbf{2}]} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 2 & -7 & 0 \\ \hline 0 & 0 & 1 \end{array}\right]$$

Multiplicando los elementos de la diagonal, obtenemos el valor del determinante de $\mathbf{A}$. En este caso $\det \mathbf{A} = -7$.

*Ejemplo* 39. Para la matriz $\left[\begin{array}{ccc} 0 & 2 & 1 \\ 9 & 6 & 3 \\ 0 & 1 & 1 \end{array}\right]$ tenemos

$$\left[\begin{array}{ccc|c} 0 & 2 & 1 & 0 \\ 9 & 6 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{[(-1)\mathbf{1}+\mathbf{3}] \\ \boldsymbol{\tau}\,[(\frac{1}{9})\mathbf{4}]}]{\substack{\boldsymbol{\tau}\,[(3)\mathbf{2}] \\ [(-2)\mathbf{1}+\mathbf{2}] \\ [(3)\mathbf{3}]}} \left[\begin{array}{ccc|c} 0 & 6 & 3 & 0 \\ 9 & 0 & 0 & 0 \\ 0 & 3 & 3 & 0 \\ \hline 0 & 0 & 0 & \frac{1}{9} \end{array}\right] \xrightarrow[\boldsymbol{\tau}\,[(\frac{1}{2})\mathbf{4}]]{\boldsymbol{\tau}\,[(2)\mathbf{3}]\ [(-1)\mathbf{2}+\mathbf{3}]} \left[\begin{array}{ccc|c} 0 & 6 & 0 & 0 \\ 9 & 0 & 0 & 0 \\ 0 & 3 & 3 & 0 \\ \hline 0 & 0 & 0 & \frac{1}{18} \end{array}\right] \xrightarrow[\boldsymbol{\tau}\,[(-1)\mathbf{4}]]{\boldsymbol{\tau}\,[\mathbf{1}\rightleftharpoons\mathbf{2}]} \left[\begin{array}{ccc|c} 6 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ \hline 0 & 0 & 0 & -\frac{1}{18} \end{array}\right]$$

Y por tanto $\det \mathbf{A} = (6)(9)(1)(\frac{-1}{6}) = -9$.

## 15.3. Determinante de matrices diagonales por bloques

Puesto que para cualquier matriz cuadrada $\mathbf{B}$ se verifica que $\begin{vmatrix} \mathbf{B} & \\ & 1 \end{vmatrix} = |\mathbf{B}|$; si $\mathbf{A}$ es de orden $m$ e $\mathbf{I}$ de orden $n$, entonces, aplicando la anterior igualdad repetidamente $n$ veces, deducimos que

$$\begin{vmatrix} \mathbf{A} & \underset{m\times n}{\mathbf{0}} \\ \underset{n\times m}{\mathbf{0}} & \underset{n\times n}{\mathbf{I}} \end{vmatrix} = \begin{vmatrix} \mathbf{A} & \underset{m\times(n-1)}{\mathbf{0}} \\ \underset{(n-1)\times m}{\mathbf{0}} & \underset{(n-1)\times(n-1)}{\mathbf{I}} \end{vmatrix} = \begin{vmatrix} \mathbf{A} & \underset{m\times(n-2)}{\mathbf{0}} \\ \underset{(n-2)\times m}{\mathbf{0}} & \underset{(n-2)\times(n-2)}{\mathbf{I}} \end{vmatrix} = \ldots = \begin{vmatrix} \mathbf{A} & 0 \\ 0 & 1 \end{vmatrix} = |\mathbf{A}|.$$

Pero todo lo realizado en las secciones anteriores se podría haber desarrollado de manera similar si, en lugar de extender la matriz $\mathbf{B}$ con la última fila y columna de la matriz identidad $n+1$, se hubiera hecho anteponiendo la primera fila y columna de la matriz identidad $n+1$. Así, para cualquier matriz cuadrada $\mathbf{B}$ también se verifica que

$$\begin{vmatrix} 1 & \\ & \mathbf{B} \end{vmatrix} = |\mathbf{B}|; \quad \text{por lo que también se deduce que} \quad \begin{vmatrix} \mathbf{I} & \\ & \mathbf{A} \end{vmatrix} = |\mathbf{A}|.$$

**Producto de matrices por bloques.** En la Sección 3.C se vio cómo calcular el producto de matrices por bloques siguiendo la fórmula de la Ecuación 3.7 en la página 44.

Ejercicio 85. Demuestre las siguientes propiedades:

(a) Sean $\mathbf{A}$ de orden $m$ y $\mathbf{B}$ de orden $n$, entonces $\begin{vmatrix} \mathbf{A} & \underset{m\times n}{\mathbf{0}} \\ \underset{n\times m}{\mathbf{0}} & \mathbf{B} \end{vmatrix} = \big|\mathbf{A}\big| \cdot \big|\mathbf{B}\big|$.

(b) Sean $\mathbf{A}$ de orden $m$, $\mathbf{B}$ de orden $n$ y $\mathbf{C}$ de orden $n$ por $m$, entonces

$$\begin{vmatrix} \mathbf{A} & \underset{m\times n}{\mathbf{0}} \\ \mathbf{C} & \mathbf{B} \end{vmatrix} = \big|\mathbf{A}\big| \cdot \big|\mathbf{B}\big|.$$

## 15.4. Fórmulas sencillas para matrices de orden menor que 4

☞ Para matrices de orden 1, 2 o 3, hay fórmulas sencillas. ¡Pero ojo! NO HAY REGLAS SENCILLAS para matrices de orden 4 o más. No extrapole que lo que funciona para una matriz 3 por 3 se puede hacer de manera similar para una matriz de orden 4 o mayor.

**Para matrices de orden 1:** Para $\mathbf{A} = \begin{bmatrix} a \end{bmatrix}$ tenemos

$$\left[\begin{array}{c|c} a & 0 \\ \hline 0 & 1 \end{array}\right] \quad \Rightarrow \quad \boxed{\det \mathbf{A} = a.}$$

**Para matrices de orden 2.** Deduzcamos la fórmula general para una matriz de la forma $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

- Caso $a \neq 0$:

$$\left[\begin{array}{cc|c} a & b & 0 \\ c & d & 0 \\ \hline 0 & 0 & 1 \end{array}\right] \xrightarrow{[(-\frac{b}{a})\boldsymbol{\tau}\mathbf{1+2}]} \left[\begin{array}{cc|c} a & 0 & 0 \\ c & \frac{ad-bc}{a} & 0 \\ \hline 0 & 0 & 1 \end{array}\right] \quad \Rightarrow \quad \det \mathbf{A} = ad - bc.$$

- Caso $a = 0$ y $b \neq 0$:

$$\left[\begin{array}{cc|c} a & b & \\ c & d & \\ \hline & & 1 \end{array}\right] \xrightarrow[[(-1)\boldsymbol{\tau}\mathbf{3}]]{[\mathbf{1}\rightleftharpoons\mathbf{2}]\boldsymbol{\tau}} \left[\begin{array}{cc|c} b & a & \\ d & c & \\ \hline & & -1 \end{array}\right] \quad \Rightarrow \quad \det \mathbf{A} = (bc - ad)(-1) = ad - bc.$$

Es decir, en todos los casos $\boxed{\det \mathbf{A} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.}$

**Para matrices de orden 3.** Deduzcamos la fórmula general para una matriz de la forma $\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

- Caso $a \neq 0$ y $(ae - bd) \neq 0$:

$$\left[\begin{array}{ccc|c} a & b & c & 0 \\ d & e & f & 0 \\ g & h & i & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[[(-\frac{c}{a})\boldsymbol{\tau}\mathbf{1+3}]]{[(-\frac{b}{a})\boldsymbol{\tau}\mathbf{1+2}]} \left[\begin{array}{ccc|c} a & 0 & 0 & 0 \\ d & e-\frac{bd}{a} & f-\frac{cd}{a} & 0 \\ g & h-\frac{bg}{a} & i-\frac{cg}{a} & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{[(\frac{-af+cd}{ae-bd})\boldsymbol{\tau}\mathbf{2+3}]}$$

$$\left[\begin{array}{ccc|c} a & 0 & 0 & 0 \\ d & e-\frac{bd}{a} & 0 & 0 \\ g & h-\frac{bg}{a} & \frac{aei-afh-bdi+bfg+cdh-ceg}{ae-bd} & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \Rightarrow \det \mathbf{A} = (bc-ad)(-1) = aei-afh-bdi+bfg+cdh-ceg.$$

- Caso $a = 0$, $b \neq 0$ y $(bd - ae) \neq 0$:

$$\left[\begin{array}{ccc|c} a & b & c & \\ d & e & f & \\ g & h & i & \\ \hline & & & 1 \end{array}\right] \xrightarrow[[(-1)\boldsymbol{\tau}\mathbf{4}]]{[\mathbf{1}\rightleftharpoons\mathbf{2}]\boldsymbol{\tau}} \left[\begin{array}{ccc|c} b & a & c & \\ e & d & f & \\ h & g & i & \\ \hline & & & -1 \end{array}\right] \Rightarrow \det \mathbf{A} = aei - afh - bdi + bfg + cdh - ceg.$$

- ... y de modo similar para el caso $a = 0$, $b = 0$, $c \neq 0$ y $(ce - bf) \neq 0$.

Reordenando los términos, en todos los casos resulta la misma fórmula que se conoce comúnmente como "Regla de Sarrus".

$$\det \mathbf{A} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + dhc - ceg - bdi - fha$$

## 15.5. No hay fórmulas sencillas para matrices de orden mayor a 3

A modo de ejemplo, tratemos de calcular el determinante de una matriz genérica de orden 4 (asumiremos que las componentes de la matriz son tales que todos los denominadores son distintos de cero). ¡Las expresiones son tan largas que la última matriz se sale de los márgenes del libro!

$$\left[\begin{array}{cccc|c} a & b & c & d & 0 \\ e & f & h & i & 0 \\ k & l & m & n & 0 \\ o & p & q & r & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\substack{\boldsymbol{\tau}\\ \left[\left(-\frac{b}{a}\right)\mathbf{1+2}\right]\\ \left[\left(-\frac{c}{a}\right)\mathbf{1+3}\right]\\ \left[\left(-\frac{d}{a}\right)\mathbf{1+4}\right]}} \left[\begin{array}{cccc|c} a & 0 & 0 & 0 & 0 \\ e & \frac{af-be}{a} & \frac{ah-ce}{a} & \frac{ai-de}{a} & 0 \\ k & \frac{al-bk}{a} & \frac{am-ck}{a} & \frac{an-dk}{a} & 0 \\ o & \frac{ap-bo}{a} & \frac{aq-co}{a} & \frac{ar-do}{a} & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\substack{\boldsymbol{\tau}\\ \left[\left(-\frac{ah-ce}{af-be}\right)\mathbf{2+3}\right]\\ \left[\left(-\frac{ai-de}{af-be}\right)\mathbf{2+4}\right]}}$$

$$\left[\begin{array}{cccc|c} a & 0 & 0 & 0 & 0 \\ e & \frac{af-be}{a} & 0 & 0 & 0 \\ k & \frac{al-bk}{a} & \frac{afm-ahl-bem+bhk+cel-cfk}{af-be} & \frac{afn-ail-ben+bik+del-dfk}{af-be} & 0 \\ o & \frac{ap-bo}{a} & -\frac{-afq+ahp+beq-bho-cep+cfo}{af-be} & -\frac{-afr+aip+ber-bio-dep+dfo}{af-be} & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\left[\left(-\frac{afn-ail-ben+bik+del-dfk}{afm-ahl-bem+bhk+cel-cfk}\right)\mathbf{3+4}\right]}$$

$$\left[\begin{array}{cccc|c} a & 0 & 0 & 0 & 0 \\ e & \frac{af-be}{a} & 0 & 0 & 0 \\ k & \frac{al-bk}{a} & \frac{afm-ahl-bem+bhk+cel-cfk}{af-be} & 0 & 0 \\ o & \frac{ap-bo}{a} & -\frac{-afq+ahp+beq-bho-cep+cfo}{af-be} & \frac{afmr-afnq-ahlr+ahnp+ailq-aimp-bemr+benq+bhkr-bhno-bikq+bimo+celr-cenp-cfkr+cfno+cikp-cilo-del}{afm-ahl-bem+bhk+cel-cfk} & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

Calculando el producto de la diagonal y simplificando llegamos a la expresión final: $\det \mathbf{A} = \begin{vmatrix} a & b & c & d \\ e & f & h & i \\ k & l & m & n \\ o & p & q & r \end{vmatrix} =$

$afmr - afnq - ahlr + ahnp + ailq - aimp - bemr + benq + bhkr - bhno - bikq + bimo + celr - cenp - cfkr + cfno + cikp - cilo - delq + demp + dfkq - dfmo - dhkp + dhlo.$

Está claro que la fórmula no es ni sencilla ni fácil de recordar... *¡No puedo imaginar lo espantosa que tiene que ser la expresión para una matriz de orden 5 o mayor!* Es importante que sepa que:

*No hay expresiones sencillas para calcular determinantes de orden mayor que 3*

## 15.6. Expansión de Laplace (desarrollo por cofactores).

Hay otra forma de calcular el determinante de una matriz de orden $n$ empleando los determinantes de submatrices de orden $n-1$. Es una descripción recursiva de la función determinante que no utiliza la eliminación gaussiana y se llama *Expansión de Laplace*. Sin embargo, tenga en cuenta que, con cualquier procedimiento, el cálculo del determinante de una matriz genérica de orden elevado es computacionalmente intenso.

☞ Para deducir la citada *Expansión de Laplace* necesitamos previamente enunciar la *propiedad multilineal* de la función determinante y también definir los *menores* y los *cofactores*.

### 15.6.1. Propiedad multilineal

**P-10** **Propiedad multilineal** **P-10**

$$\det\left[\dots(\beta\boldsymbol{a}+\psi\boldsymbol{b});\dots\right]=\beta\det\left[\dots\boldsymbol{a};\dots\right]+\psi\det\left[\dots\boldsymbol{b};\dots\right]$$

*Demostración.* Puesto que se verifica la propiedad **P-3** del producto, basta demostrar que

$$\det\left[\dots(\boldsymbol{a}+\boldsymbol{b});\dots\right]=\det\left[\dots\boldsymbol{a};\dots\right]+\det\left[\dots\boldsymbol{b};\dots\right].$$

Haremos la demostración para la primera columna. La demostración para el resto de columnas es similar.

Hay dos casos: si $\boldsymbol{y}_1,\dots,\boldsymbol{y}_{n-1}$ son linealmente dependientes, entonces la demostración es inmediata pues

$$\det\left[(\boldsymbol{a}+\boldsymbol{b});\ \boldsymbol{y}_1;\dots\boldsymbol{y}_{n-1};\right]=0=\det\left[\boldsymbol{a};\ \boldsymbol{y}_1;\dots\boldsymbol{y}_{n-1};\right]+\det\left[\boldsymbol{b};\ \boldsymbol{y}_1;\dots\boldsymbol{y}_{n-1};\right]=0+0.$$

Si, por el contrario, existe $\boldsymbol{x}$ tal que $\left[\boldsymbol{x};\ \boldsymbol{y}_1;\dots\boldsymbol{y}_{n-1};\right]$ es una base de $\mathbb{R}^n$, entonces

$$\begin{aligned}\det\left[(\alpha\boldsymbol{x}+\psi_1\boldsymbol{y}_1+\dots+\psi_{n-1}\boldsymbol{y}_{n-1});\ \boldsymbol{y}_1;\dots\boldsymbol{y}_{n-1};\right]&\overset{*}{=}\det\left[\alpha\boldsymbol{x};\ \boldsymbol{y}_1;\dots\boldsymbol{y}_{n-1};\right] &(15.3)\\ &=\alpha\det\left[\boldsymbol{x};\ \boldsymbol{y}_1;\dots\boldsymbol{y}_{n-1};\right] &(15.4)\end{aligned}$$

(*) ya que, con una secuencia de transformaciones elementales, podemos reducir la primera columna a $\alpha\boldsymbol{x}$.

Ahora, si

$$\begin{aligned}\boldsymbol{a}&=\alpha\boldsymbol{x}+\psi_1\boldsymbol{y}_1+\dots+\psi_{n-1}\boldsymbol{y}_{n-1}\\ \boldsymbol{b}&=\beta\boldsymbol{x}+\gamma_1\boldsymbol{y}_1+\dots+\gamma_{n-1}\boldsymbol{y}_{n-1}\end{aligned}$$

(pues $\left[\boldsymbol{x};\ \boldsymbol{y}_1;\dots\boldsymbol{y}_{n-1};\right]$ es una base de $\mathbb{R}^n$) tendremos que

$$\boldsymbol{a}+\boldsymbol{b}=(\alpha+\beta)\boldsymbol{x}+(\psi_1+\gamma_1)\boldsymbol{y}_1+\dots+(\psi_{n-1}+\gamma_{n-1})\boldsymbol{y}_{n-1}$$

y, consecuentemente,

$$\begin{aligned}\det\left[\boldsymbol{a}+\boldsymbol{b};\ \boldsymbol{y}_1;\dots;\boldsymbol{y}_{n-1};\right]&=(\alpha+\beta)\det\left[\boldsymbol{x};\ \boldsymbol{y}_1;\dots;\boldsymbol{y}_{n-1};\right] && \text{por (15.4)}\\ &=\alpha\det\left[\boldsymbol{x};\ \boldsymbol{y}_1;\dots;\boldsymbol{y}_{n-1};\right]+\beta\det\left[\boldsymbol{x};\ \boldsymbol{y}_1;\dots;\boldsymbol{y}_{n-1};\right] && \text{pues det } \mathbf{A} \text{ es un número}\\ &=\det\left[\alpha\boldsymbol{x};\ \boldsymbol{y}_1;\dots;\boldsymbol{y}_{n-1};\right]+\det\left[\beta\boldsymbol{x};\ \boldsymbol{y}_1;\dots;\boldsymbol{y}_{n-1};\right] && \text{por P-3}\\ &=\det\left[\boldsymbol{a};\ \boldsymbol{y}_1;\dots;\boldsymbol{y}_{n-1};\right]+\det\left[\boldsymbol{b};\ \boldsymbol{y}_1;\dots;\boldsymbol{y}_{n-1};\right] && \text{por (15.3)}\end{aligned}$$

□

*Ejemplo* 40. Para $\mathbb{R}^2$

$$\begin{vmatrix}a+\alpha & c\\ b+\beta & d\end{vmatrix}=\begin{vmatrix}a & c\\ b & d\end{vmatrix}+\begin{vmatrix}\alpha & c\\ \beta & d\end{vmatrix}.$$

O también

$$\begin{vmatrix}a & c\\ b & d\end{vmatrix}=\begin{vmatrix}a & c\\ 0 & d\end{vmatrix}+\begin{vmatrix}0 & c\\ b & d\end{vmatrix}.$$

### 15.6.2. Menores y cofactores

☞ En esta sección veremos cómo calcular los determinantes mediante fórmulas más sencillas. Fórmulas que expresan determinantes de matrices de orden $n$ como sumas de determinantes de matrices de orden $n-1$.

**Proposición 15.6.1.** *Si* $\underset{n\times n}{\mathbf{A}}$ *es de la forma* $\left[\begin{array}{ccc|c} & \mathbf{B} & & 0\\ & {\scriptstyle (n-1)\times(n-1)} & & \vdots\\ & & & 0\\ \hline a_{11} & \cdots & a_{1n-1} & 1 \end{array}\right]$ *entonces* $\boxed{\det(\mathbf{A})=\det(\mathbf{B}).}$

*Demostración.* Mediante operaciones elementales por columnas de *Tipo I*, podemos reducir $\mathbf{A}$ de manera que

$$\mathbf{A}=\left[\begin{array}{ccc|c} & \mathbf{B} & & 0\\ & & & \vdots\\ & & & 0\\ \hline a_{11} & \cdots & a_{1n-1} & 1 \end{array}\right]\xrightarrow{\text{transformaciones } \textit{Tipo I}}\left[\begin{array}{ccc|c} & \mathbf{B} & & 0\\ & & & \vdots\\ & & & 0\\ \hline 0 & \cdots & 0 & 1 \end{array}\right]\equiv\begin{bmatrix}\mathbf{B} & \\ & 1\end{bmatrix}$$

Y, al emplear solo transformaciones *Tipo I*, deducimos que $|\mathbf{A}|=\begin{vmatrix}\mathbf{B} & \\ & 1\end{vmatrix}=|\mathbf{B}|$ (Proposición 15.0.1). □

**Nueva notación y definición de menores y cofactores**

Sea $\boldsymbol{q}=(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n,)$ de $\mathbb{R}^n$, al quitar la $j$-ésima componente, $q_j$, obtenemos un vector de $\mathbb{R}^{(n-1)}$ que denotamos con $\boldsymbol{q}^{\not j}$:

$$\boldsymbol{q}^{\not j}=(q_1,\ldots,\ q_{j-1},\ q_{j+1},\ \ldots,q_n,)\in\mathbb{R}^{n-1}.$$

Análogamente, si $\mathbf{A}$ es de orden $m$ por $n$, denotamos la submatriz que resulta de quitar la $i$-ésima fila con ${}^{i\not}\mathbf{A}$, y la submatriz que resulta de quitar la $j$-ésima columna con $\mathbf{A}^{\not j}$. Así, ${}^{i\not}\mathbf{A}^{\not j}$, es la submatriz de orden $m-1$ por $n-1$ que resulta de quitar la fila $i$-ésima y la columna $j$-ésima.

**Definición 15.2** (menores y cofactores)**.** *Sea* ${}^{i\not}\mathbf{A}^{\not j}$ *la submatriz de* $\mathbf{A}$ *resultante de eliminar la fila* $i$ *y la columna* $j$*. Llamamos* menor *del elemento* ${}_{i|}\mathbf{A}_{|j}$ *al determinante*

$$\det\left({}^{i\not}\mathbf{A}^{\not j}\right).$$

*Los menores con los signos alternados en función de si* $(i+j)$ *es par (en cuyo caso el signo no cambia) o impar (en cuyo caso se invierte el signo) se denominan* cofactores*. Así pues, el* cofactor *de* ${}_{i|}\mathbf{A}_{|j}$ *es*

$$\operatorname{cof}_{ij}\left(\mathbf{A}\right)=(-1)^{i+j}\det\left({}^{i\not}\mathbf{A}^{\not j}\right).$$

*Ejemplo* 41. Para $\mathbf{A}=\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}$, tenemos

$${}^{1\not}\mathbf{A}^{\not 2}=\begin{bmatrix}4&6\\7&9\end{bmatrix},\qquad {}^{3\not}\mathbf{A}^{\not 3}=\begin{bmatrix}1&2\\4&5\end{bmatrix},$$

y, por tanto,

$$\operatorname{cof}_{12}\left(\mathbf{A}\right)=(-1)^{1+2}\det\left({}^{1\not}\mathbf{A}^{\not 2}\right)=(-1)\begin{vmatrix}4&6\\7&9\end{vmatrix}.$$

y

$$\operatorname{cof}_{33}\left(\mathbf{A}\right)=(-1)^{3+3}\det\left({}^{3\not}\mathbf{A}^{\not 3}\right)=\begin{vmatrix}1&2\\4&5\end{vmatrix}.$$

### 15.6.3. Desarrollo del determinante por cofactores (Expansión de Laplace).

**Teorema 15.6.2 (Expansión de Laplace).** *Para* $\mathbf{A}$ *de orden* $n$, $\det(\mathbf{A})$ *se puede expresar como suma de los productos de los elementos de cualquier columna (o fila) de* $\mathbf{A}$ *por sus correspondientes cofactores:*

$$\det(\mathbf{A}) = \sum_{i=1}^{n} a_{ij}\,\mathrm{cof}_{ij}\left(\mathbf{A}\right), \qquad \textit{expansión por la columna } j\textit{-ésima};$$

*o bien:* $\det(\mathbf{A}) = \sum_{j=1}^{n} a_{ij}\,\mathrm{cof}_{ij}\left(\mathbf{A}\right)$, *expansión por la fila* $i$*-ésima.*

*Demostración.* Probaremos la expansión por la columna $j$-ésima (la expansión por filas es similar). Primero movemos la columna $j$-ésima a la última posición con una secuencia de intercambios:

$$\mathbf{A}_{\underset{[j\rightleftharpoons(j+1)]}{\boldsymbol{\tau}}\ \underset{[(j+1)\rightleftharpoons(j+2)]}{\boldsymbol{\tau}}\ \underset{[(j+2)\rightleftharpoons(j+3)]}{\boldsymbol{\tau}}\ \cdots\ \underset{[(n-1)\rightleftharpoons n]}{\boldsymbol{\tau}}} = \left[\begin{array}{c|c} & \\ \mathbf{A}^{\overset{\curvearrowright}{j}} & \mathbf{A}_{|j} \\ & \end{array}\right].$$

Puesto que hay $n-j$ permutaciones en la sucesión, tenemos que: $\det\mathbf{A} = (-1)^{n-j}\cdot\det\left[\begin{array}{c|c} & \\ \mathbf{A}^{\overset{\curvearrowright}{j}} & \mathbf{A}_{|j} \\ & \end{array}\right]$.

Escribiendo $\mathbf{A}_{|j}$ como $\begin{pmatrix} a_{1j}\\ a_{2j}\\ \vdots\\ a_{nj}\end{pmatrix} = \begin{pmatrix} a_{1j}\\ 0\\ \vdots\\ 0\end{pmatrix} + \begin{pmatrix} 0\\ a_{2j}\\ \vdots\\ 0\end{pmatrix} + \cdots + \begin{pmatrix} 0\\ \vdots\\ 0\\ a_{nj}\end{pmatrix}$, y empleando la propiedad multilineal (***P-10***):

$$\det\mathbf{A} = (-1)^{n-j}\left(\det\left[\begin{array}{c|c} & a_{1j}\\ \mathbf{A}^{\overset{\curvearrowright}{j}} & 0\\ & \vdots\\ & 0\end{array}\right] + \det\left[\begin{array}{c|c} & 0\\ \mathbf{A}^{\overset{\curvearrowright}{j}} & a_{2j}\\ & \vdots\\ & 0\end{array}\right] + \cdots + \det\left[\begin{array}{c|c} & 0\\ \mathbf{A}^{\overset{\curvearrowright}{j}} & \vdots\\ & 0\\ & a_{nj}\end{array}\right]\right).$$

Si "hundimos" la primera fila de la matriz del primer determinante hasta la última posición mediante una sucesión de $(n-1)$ intercambios de filas: $\underset{[1\rightleftharpoons 2]}{\boldsymbol{\tau}}\ \underset{[2\rightleftharpoons 3]}{\boldsymbol{\tau}}\ \cdots\ \underset{[(n-1)\rightleftharpoons n]}{\boldsymbol{\tau}}$; y la segunda fila de la matriz del segundo determinante mediante $(n-2)$ intercambios: $\underset{[2\rightleftharpoons 3]}{\boldsymbol{\tau}}\ \underset{[3\rightleftharpoons 4]}{\boldsymbol{\tau}}\ \cdots\ \underset{[(n-1)\rightleftharpoons n]}{\boldsymbol{\tau}}$; —y en general, "hundimos" la fila $i$-ésima de la matriz $i$-ésima con $(n-i)$ intercambios: $\underset{[i\rightleftharpoons(i+1)]}{\boldsymbol{\tau}}\ \underset{[(i+1)\rightleftharpoons(i+2)]}{\boldsymbol{\tau}}\ \cdots\ \underset{[(n-1)\rightleftharpoons n]}{\boldsymbol{\tau}}$, obtenemos:

$$|\mathbf{A}| = (-1)^{n-j}\left((-1)^{n-1}\cdot\det\left[\begin{array}{c|c} {}^{1\urcorner}\mathbf{A}^{\overset{\curvearrowright}{j}} & \mathbf{0}\\ \hline {}_{1|}\mathbf{A}^{\overset{\curvearrowright}{j}} & a_{1j}\end{array}\right] + (-1)^{n-2}\cdot\det\left[\begin{array}{c|c} {}^{2\urcorner}\mathbf{A}^{\overset{\curvearrowright}{j}} & \mathbf{0}\\ \hline {}_{2|}\mathbf{A}^{\overset{\curvearrowright}{j}} & a_{2j}\end{array}\right] + \cdots + \det\left[\begin{array}{c|c} {}^{n\urcorner}\mathbf{A}^{\overset{\curvearrowright}{j}} & \mathbf{0}\\ \hline {}_{n|}\mathbf{A}^{\overset{\curvearrowright}{j}} & a_{nj}\end{array}\right]\right),$$

donde ${}_{i|}\mathbf{A}^{\overset{\curvearrowright}{j}}$ es la fila $i$-ésima de la submatriz $\mathbf{A}^{\overset{\curvearrowright}{j}}$ (que en cada matriz hemos "hundido" hasta la última posición junto con el elemento $a_{ij}$). Así, por la Proposición 15.6.1

$$\det(\mathbf{A}) \;=\; (-1)^{n-j}\sum_{i=1}^{n} a_{ij}(-1)^{n-i}\det\left({}^{i\urcorner}\mathbf{A}^{\overset{\curvearrowright}{j}}\right) \;=\; \sum_{i=1}^{n} a_{ij}\,\mathrm{cof}_{ij}\left(\mathbf{A}\right);$$

puesto que $(-1)^{n-j}\cdot(-1)^{n-i} = (-1)^{2n-(i+j)} = (-1)^{-(i+j)} = (-1)^{i+j}$. □

EJERCICIO 86. Calcule el siguiente determinante: $\det\mathbf{A} = \begin{vmatrix} 2 & 0 & 3 & 2\\ 5 & 1 & 2 & 4\\ 3 & 0 & 1 & 2\\ 5 & 3 & 2 & 1\end{vmatrix}$.

EJERCICIO 87. Calcule $\begin{vmatrix} a & b & c\\ d & e & f\\ g & h & i\end{vmatrix}$ desarrollado por los cofactores de la segunda columna.

## 15.7. Aplicación de los determinantes

### 15.7.1. Regla de Cramer para la resolución de un sistema de ecuaciones

Suponga el sistema de $n$ ecuaciones $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, donde $\mathbf{A}$ es cuadrada y $|\mathbf{A}| \neq 0$. Sabemos que dicho sistema tiene solución única para cualquier vector $\boldsymbol{b} \in \mathbb{R}^n$. Dicha solución es

$$\boldsymbol{x} = \mathbf{A}^{-1}\,\boldsymbol{b};$$

de manera que

$$\boldsymbol{b} = (\mathbf{A}_{|1})x_1 + \cdots + (\mathbf{A}_{|j})x_j + \cdots + (\mathbf{A}_{|n})x_n, \tag{15.5}$$

donde los coeficientes $x_i$ son los componentes[1] del vector solución $\boldsymbol{x}$.

Calculemos el determinante de una nueva matriz, idéntica a $\mathbf{A}$ excepto por su $j$-ésima columna $\mathbf{A}_{|j}$, que ha sido sustituida por el vector $\boldsymbol{b}$:

$$\det\Big[\mathbf{A}_{|1};\ \mathbf{A}_{|2};\ \ldots\ \overbrace{\boldsymbol{b}}^{\text{pos. } j};\ \ldots\ \mathbf{A}_{|n};\Big].$$

Si nos fijamos en la Ecuación (15.5), es fácil constatar que para pasar de $\mathbf{A}$ a la nueva matriz, hemos multiplicado $\mathbf{A}_{|j}$ por $x_j$ y le hemos sumado una combinación lineal del resto de columnas; por tanto:

$$\det\Big[\mathbf{A}_{|1};\ \ldots\ \overbrace{\boldsymbol{b}}^{\text{pos. } j};\ \ldots\ \mathbf{A}_{|n};\Big] = x_j \cdot \det(\mathbf{A}).$$

Despejando $x_j$, obtenemos la *"regla de Cramer"* para la resolución de sistemas de ecuaciones determinados, que afirma que: *cada componente $x_j$ de la solución $\boldsymbol{x}$ del sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ se puede calcular del siguiente modo*

$$x_j = \frac{\det\Big[\mathbf{A}_{|1};\ \ldots\ \overbrace{\boldsymbol{b}}^{\text{pos. } j};\ \ldots\ \mathbf{A}_{|n};\Big]}{\det(\mathbf{A})}.$$

### 15.7.2. Cálculo de la inversa de una matriz

**Definición 15.3.** *Para $\mathbf{A}$, de orden $n$, la matriz* $\mathbf{Adj(A)}$ *(la* matriz adjunta *de $\mathbf{A}$) es la transpuesta de la matriz resultante de sustituir cada elemento ${}_{i|}\mathbf{A}_{|j}$ por su correspondiente cofactor:* $\operatorname{cof}_{ij}(\mathbf{A})$,

$$\mathbf{Adj(A)} = \begin{bmatrix} \operatorname{cof}_{11}(\mathbf{A}) & \operatorname{cof}_{12}(\mathbf{A}) & \cdots & \operatorname{cof}_{1n}(\mathbf{A}) \\ \operatorname{cof}_{21}(\mathbf{A}) & \operatorname{cof}_{22}(\mathbf{A}) & \cdots & \operatorname{cof}_{2n}(\mathbf{A}) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{cof}_{n1}(\mathbf{A}) & \operatorname{cof}_{n2}(\mathbf{A}) & \cdots & \operatorname{cof}_{nn}(\mathbf{A}) \end{bmatrix}^{\mathsf{T}}.$$

Si $\mathbf{A}$ es *invertible*, ¿qué obtenemos si multiplicamos la matriz adjunta de $\mathbf{A}$ por la matriz $\mathbf{A}$?

$$\underbrace{\begin{bmatrix} \operatorname{cof}_{11}(\mathbf{A}) & \operatorname{cof}_{21}(\mathbf{A}) & \cdots & \operatorname{cof}_{n1}(\mathbf{A}) \\ \operatorname{cof}_{12}(\mathbf{A}) & \operatorname{cof}_{22}(\mathbf{A}) & \cdots & \operatorname{cof}_{n2}(\mathbf{A}) \\ \vdots & \vdots & \ddots & \vdots \\ \operatorname{cof}_{1n}(\mathbf{A}) & \operatorname{cof}_{2n}(\mathbf{A}) & \cdots & \operatorname{cof}_{nn}(\mathbf{A}) \end{bmatrix}}_{\mathbf{Adj(A)}} \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}}_{\mathbf{A}}$$

El primer elemento de la diagonal de la matriz resultante es el producto punto entre la primera fila de $\mathbf{Adj(A)}$ y la primera columna de $\mathbf{A}$; y, por tanto, es el desarrollo de la expansión de Laplace de $|\mathbf{A}|$ por la

[1] las coordenadas de $\boldsymbol{b}$ en la base formada por las columnas de $\mathbf{A}$

primera columna de $\mathbf{A}$. El segundo elemento de la diagonal es el desarrollo de $|\mathbf{A}|$ por la segunda columna, etc. Y en general, el componente $j$-ésimo de la diagonal es la expansión por la columna $j$-ésima:

$$_{j|}\mathbf{Adj(A)A}_{|j} = \operatorname{cof}_{1j}(\mathbf{A})\, a_{1j} + \operatorname{cof}_{2j}(\mathbf{A})\, a_{2j} + \operatorname{cof}_{3j}(\mathbf{A})\, a_{3j} + \cdots + \operatorname{cof}_{nj}(\mathbf{A})\, a_{nj} = \sum_{i=1}^{n} a_{ij} \operatorname{cof}_{ij}(\mathbf{A}) = \det(\mathbf{A}).$$

Así pues, la diagonal de la matriz resultante está compuesta por el valor del determinante de $\mathbf{A}$.

Los elementos fuera de la diagonal $_{i|}\big(\mathbf{Adj(A)A}\big)_{|j}$ (con $i \neq j$) son determinantes de matrices singulares por la particularidad de que las columnas $i$ y $j$ son iguales, es decir, que $\mathbf{A}_{|i} = \mathbf{A}_{|j}$. Veámoslo.

Por ejemplo, $_{1|}\big(\mathbf{Adj(A)A}\big)_{|2}$ es el desarrollo por la primera columna del determinante de una matriz cuya primera columna es igual que la segunda, de manera que el desarrollo por la primera columna es cero, pues:

$$\begin{vmatrix} a_{12} & a_{12} & \dots & a_{1n} \\ a_{22} & a_{22} & \dots & a_{2n} \\ a_{32} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & & \vdots \\ a_{n2} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{12} \operatorname{cof}_{11}(\mathbf{A}) + a_{22} \operatorname{cof}_{21}(\mathbf{A}) + a_{32} \operatorname{cof}_{31}(\mathbf{A}) + \cdots + a_{n2} \operatorname{cof}_{n1}(\mathbf{A}) = \sum_{i=1}^{n} a_{i2} \operatorname{cof}_{i1}(\mathbf{A}) = 0.$$

Y el elemento $k$-ésimo ($k \neq 1$) de la primera fila, $_{1|}\big(\mathbf{Adj(A)A}\big)_{|k}$, es el desarrollo por la primera columna del determinante de una matriz cuya primera y $k$-ésima columnas son iguales:

$$\begin{vmatrix} a_{1k} & a_{12} & \dots & a_{1n} \\ a_{2k} & a_{22} & \dots & a_{2n} \\ a_{3k} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & & \vdots \\ a_{nk} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{1k} \operatorname{cof}_{11}(\mathbf{A}) + a_{2k} \operatorname{cof}_{21}(\mathbf{A}) + a_{3k} \operatorname{cof}_{31}(\mathbf{A}) + \cdots + a_{nk} \operatorname{cof}_{n1}(\mathbf{A}) = \sum_{i=1}^{n} a_{ik} \operatorname{cof}_{i1}(\mathbf{A}) = 0.$$

En general, $_{i|}\big(\mathbf{Adj(A)A}\big)_{|j} = \begin{cases} |\mathbf{A}| & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$; es decir

$$\mathbf{Adj(A)A} = \begin{bmatrix} |\mathbf{A}| & 0 & \cdots & 0 \\ 0 & |\mathbf{A}| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |\mathbf{A}| \end{bmatrix} = |\mathbf{A}| \cdot \mathbf{I}.$$

Dividiendo por $|\mathbf{A}|$, se concluye que $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\big(\mathbf{Adj(A)}\big)$.

# Parte VI

# Autovalores y autovectores. Diagonalización y formas cuadráticas

# Autovalores y autovectores

Considere la siguiente ecuación donde $\mathbf{A}$ es cuadrada y de orden $n$:

$$\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x} \tag{16.1}$$

**Definición 16.1** (Autovalor, autovector y espectro)**.** *Un* autovalor *de una matriz cuadrada* $\mathbf{A}$ *es cualquier número* $\lambda$ *tal que* (16.1) *tiene soluciones no nulas* ($\boldsymbol{x} \neq \mathbf{0}$)*. En tal caso, los vectores* $\boldsymbol{x}$ *se llaman* autovectores *de* $\mathbf{A}$ *correspondientes al autovalor* $\lambda$*. El conjunto de autovalores de* $\mathbf{A}$ *se denomina* espectro *de* $\mathbf{A}$*.*

Por ejemplo, es fácil observar que

$$\boldsymbol{x}_1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \text{ y } \boldsymbol{x}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{son autovectores de} \quad \mathbf{A} = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$$

correspondientes a los autovalores 6 y 1, respectivamente (basta multiplicar $\mathbf{A}(\boldsymbol{x}_i)$ para comprobarlo). El espectro de $\mathbf{A}$ es $\{6, 1\}$.

Ejercicio 88. Demuestre la siguiente proposición:

**Proposición 16.0.1.** *Una combinación lineal no nula de autovectores correspondientes al autovalor* $\lambda$ *es otro autovector correspondiente al mismo autovalor* $\lambda$*.*

El anterior resultado justifica la siguiente definición:

**Definición 16.2** (Autoespacio)**.** *Sea* $\mathbf{A}$ *de orden* $n$*. Denominamos* autoespacio *correspondiente al autovalor* $\lambda$ *al subespacio formado por los autovectores correspondientes al autovalor* $\lambda$ *de* $\mathbf{A}$ *junto con el vector nulo:*

$$\mathcal{E}_\lambda(\mathbf{A}) = \left\{ \boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x} \right\}.$$

## 16.1. Cálculo de los autovalores y los autovectores

¿Qué podemos decir acerca de la existencia de autovalores de una matriz en general? ¿Y cómo podemos encontrar dichos autovalores? Para responder, vamos a escribir la Ecuación (16.1) de manera diferente

$$\begin{aligned} \mathbf{A}\boldsymbol{x} &= \lambda\boldsymbol{x} \\ \mathbf{A}\boldsymbol{x} &= \lambda\mathbf{I}\boldsymbol{x} \\ \mathbf{A}\boldsymbol{x} - \lambda\mathbf{I}\boldsymbol{x} &= \mathbf{0} \\ (\mathbf{A} - \lambda\mathbf{I})\boldsymbol{x} &= \mathbf{0}, \end{aligned} \tag{16.2}$$

es decir, el autoespacio correspondiente a $\lambda$ es el espacio nulo de $(\mathbf{A} - \lambda\mathbf{I})$:

$$\mathcal{E}_\lambda(\mathbf{A}) = \mathcal{N}\left(\mathbf{A} - \lambda\mathbf{I}\right).$$

Como $\mathcal{E}_\lambda(\mathbf{A})$ contiene vectores no nulos (los autovectores), $\mathcal{N}\left(\mathbf{A}-\lambda\mathbf{I}\right)$ también los contiene, es decir, la matriz cuadrada $(\mathbf{A}-\lambda\mathbf{I})$ es singular, y por tanto:

$$\det(\mathbf{A}-\lambda\mathbf{I})=0. \tag{16.3}$$

Así pues, el problema reside en encontrar los valores de $\lambda$ para los que $(\mathbf{A}-\lambda\mathbf{I})$ es singular ($|\mathbf{A}-\lambda\mathbf{I}|=0$). Como veremos a continuación, al desarrollar el determinante se comprueba que existen $n+1$ coeficientes $p_0,\ldots,p_n$ de $\mathbb{R}$, tales que para todo $\lambda$

$$|\mathbf{A}-\lambda\mathbf{I}|= p_0+p_1\lambda+\cdots+p_{n-1}\lambda^{n-1}+p_n\lambda^n.$$

$|\mathbf{A}-\lambda\mathbf{I}|$ es un polinomio en $\lambda$ que denominamos *polinomio característico* de $\mathbf{A}$ y que denotamos con $P_{\mathbf{A}}(\lambda)$.

En el ejemplo de más arriba:

$$P_{\mathbf{A}}(\lambda)=\det(\mathbf{A}-\lambda\mathbf{I})=\begin{vmatrix}5-\lambda & 4\\ 1 & 2-\lambda\end{vmatrix}=6-7\lambda+\lambda^2=0.$$

Vamos a demostrar que efectivamente, para cualquier matriz $\mathbf{A}$ cuadrada, $P_{\mathbf{A}}(\lambda)$ siempre es un polinomio.

*Demostración.* Vamos a demostrarlo por inducción sobre el orden $n$.

El resultado es evidente para cualquier matriz de orden uno, pues: $\det\left(a_{11}-\lambda\right)=a_{11}-\lambda^1$. Supuesto que es cierto para cualquier matriz de orden $(n-1)$, veamos que también es cierto para cualquier matriz de orden $n$:

$$\begin{aligned}\det(\mathbf{A}-\lambda\mathbf{I}) &= \det\left[\begin{array}{ccc|c} & {}^{n\urcorner}\mathbf{A}^{\urcorner n}-\lambda\mathbf{I} & & \begin{matrix}a_{1n}\\ \vdots\\ a_{(n-1)n}\end{matrix}\\ \hline * & \cdots & * & a_{nn}-\lambda\end{array}\right]\\ &= \det\left[\begin{array}{ccc|c} & {}^{n\urcorner}\mathbf{A}^{\urcorner n}-\lambda\mathbf{I} & & \begin{matrix}a_{1n}\\ \vdots\\ a_{(n-1)n}\end{matrix}\\ \hline * & \cdots & * & a_{nn}\end{array}\right]+\det\left[\begin{array}{ccc|c} & {}^{n\urcorner}\mathbf{A}^{\urcorner n}-\lambda\mathbf{I} & & \begin{matrix}0\\ \vdots\\ 0\end{matrix}\\ \hline * & \cdots & * & -\lambda\end{array}\right]\\ &= \det\left[\begin{array}{ccc|c} & {}^{n\urcorner}\mathbf{A}^{\urcorner n}-\lambda\mathbf{I} & & \begin{matrix}a_{1n}\\ \vdots\\ a_{(n-1)n}\end{matrix}\\ \hline * & \cdots & * & a_{nn}\end{array}\right]-\lambda\det\left({}^{n\urcorner}\mathbf{A}^{\urcorner n}-\lambda\mathbf{I}\right),\end{aligned} \tag{16.4}$$

donde ${}^{n\urcorner}\mathbf{A}^{\urcorner n}$ es la submatriz de $\mathbf{A}$ que resulta tras quitar la fila y columna $n$-ésimas. Por hipótesis de inducción, $\det\left({}^{n\urcorner}\mathbf{A}^{\urcorner n}-\lambda\mathbf{I}\right)$ es un polinomio de grado $(n-1)$; por tanto, el término que resta en (16.4) es un polinomio de grado $n$.

Ahora se presentan dos casos. Si $a_{nn}\neq 0$, mediante transformaciones elementales de *Tipo I* se pueden anular todos los componentes de la última fila que están a la izquierda de $a_{nn}$. Entonces

$$\det\left[\begin{array}{ccc|c} & {}^{n\urcorner}\mathbf{A}^{\urcorner n}-\lambda\mathbf{I} & & \begin{matrix}a_{1n}\\ \vdots\\ a_{(n-1)n}\end{matrix}\\ \hline * & \cdots & * & a_{nn}\end{array}\right]=\det\left[\begin{array}{ccc|c} & \mathbf{B}-\lambda\mathbf{I} & & \begin{matrix}a_{1n}\\ \vdots\\ a_{(n-1)n}\end{matrix}\\ \hline 0 & \cdots & 0 & a_{nn}\end{array}\right]=a_{nn}\cdot\det(\mathbf{B}-\lambda\mathbf{I}),$$

que por hipótesis de inducción es un polinomio de grado $(n-1)$, por lo que la diferencia de polinomios en (16.4) es un polinomio de grado $n$. En el segundo caso, $a_{nn} = 0$, por tanto

$$\det\left[\begin{array}{ccc|c} & \overset{n}{\mathbf{A}} - \lambda\mathbf{I} & & \begin{matrix} a_{1n} \\ \vdots \\ a_{(n-1)n} \end{matrix} \\ \hline * & \cdots & * & 0 \end{array}\right] = \det\left[\begin{array}{ccc|c} & \overset{n}{\mathbf{A}} - \lambda\mathbf{I} & & \begin{matrix} a_{1n} \\ \vdots \\ a_{(n-1)n} \end{matrix} \\ \hline * & \cdots & * & -1 \end{array}\right] + \det\left[\begin{array}{ccc|c} & \overset{n}{\mathbf{A}} - \lambda\mathbf{I} & & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline * & \cdots & * & 1 \end{array}\right].$$

Repitiendo el argumento de más arriba, constatamos que ambos determinantes son polinomios de grado $(n-1)$, por lo que su diferencia es un polinomio de grado menor o igual que $(n-1)$. Así que de nuevo la diferencia de polinomios en (16.4) es un polinomio de grado $n$. □

**Corolario 16.1.1.** *Todo autovalor de* $\mathbf{A}$ *es raíz del polinomio característico* $P_{\mathbf{A}}(\lambda)$.

Para el ejemplo de más arriba, puesto que $\det(\mathbf{A} - \lambda\mathbf{I}) = \lambda^2 - 7\lambda + 6 = 0$; tenemos que

$$\lambda = \frac{7 \pm \sqrt{49-24}}{2} = \frac{7 \pm 5}{2} = \begin{cases} 6 \\ 1 \end{cases} \quad \Longrightarrow \quad \text{El } \textit{espectro} \text{ de } \mathbf{A} \text{ es } \{6, 1\}.$$

El *Teorema Fundamental del Álgebra*[1] establece que un polinomio $P(\lambda)$ con coeficientes complejos [2] y de grado $n > 0$ se puede factorizar como

$$P_{\mathbf{A}}(\lambda) \;=\; p_0 + p_1\lambda + \cdots + p_{n-1}\lambda^{n-1} + p_n\lambda^n \;=\; \alpha(\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda), \quad \text{donde } \alpha, \lambda_1, \ldots, \lambda_n \in \mathbb{C}.$$

y, por tanto, *tiene como mínimo una raíz* y como máximo $n$ raíces complejas distintas.

☞ *A partir de ahora, y para poder hacer uso del* Teorema Fundamental del Álgebra, *asumiremos que tanto los vectores como las matrices están formadas por números complejos:* $\mathbb{C}^n$ *y* $\mathbb{C}^{n\times m}$ *respectivamente.*

Como consecuencia, tenemos el siguiente resultado:

**Teorema 16.1.2.** *Los autovalores de una matriz* $\mathbf{A}$ *de orden* $n$ *son las raíces del* Polinomio Característico $P_{\mathbf{A}}(\lambda)$. *Por tanto,* $\mathbf{A}$ *tiene como máximo* $n$ *autovalores distintos.*

Se denomina *multiplicidad* de una raíz al número de veces que aparece en la factorización del polinomio. Se extiende esta nomenclatura a los autovalores añadiendo la "coletilla" *algebraica*:

**Definición 16.3** (Multiplicidad algebráica de un autovalor)**.** *Si* $\lambda$ *es una raíz de* $P_{\mathbf{A}}$ *de multiplicidad* $k$, *diremos que* $\lambda$ *es un autovalor de* $\mathbf{A}$ *de* multiplicidad algebraica $k$; *que denotamos con* $\mu(\lambda) = k$.

Siguiendo con el ejemplo de más arriba, la multiplicidad algebraica tanto del autovalor $\lambda_1 = 6$ como de $\lambda_2 = 1$ es uno; y los autoespacios de $\mathbf{A}$ correspondientes a 6 y a 1 son respectivamente los espacios nulos

$$\mathcal{N}\left(\mathbf{A} - 6\mathbf{I}\right); \qquad \text{y} \qquad \mathcal{N}\left(\mathbf{A} - 1\mathbf{I}\right);$$

por tanto, los autovectores de $\mathbf{A}$ correspondientes a 6 y 1 son las soluciones no nulas de los respectivos sistemas

$$\begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{y} \qquad \begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

**Definición 16.4** (Multiplicidad geométrica de un autovalor)**.** *La dimensión del autoespacio* $\mathcal{N}\left(\mathbf{A} - \lambda\mathbf{I}\right)$ *se denomina* multiplicidad geométrica del autovalor $\lambda$; *que denotamos con* $\gamma(\lambda)$.

[1] Cuya demostración está fuera del alcance de este curso.
[2] El conjunto de números complejos se denota con $\mathbb{C}$.

☞ El problema de encontrar los autovalores y autovectores de una matriz requiere los siguientes pasos:

1. Encontrar el polinomio característico $P_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$.
2. Encontrar las raíces $\lambda_i$ de la ecuación característica $P_{\mathbf{A}}(\lambda) = 0$.
3. Resolver los sistemas homogéneos $(\mathbf{A} - \lambda_i\mathbf{I})\boldsymbol{x} = \mathbf{0}$ para encontrar los autovectores.

La multiplicidad *algebraica* de $\lambda$ es el número de veces que se repite la raíz $\lambda$ en el polinomio característico $P_{\mathbf{A}}(\lambda)$.

La multiplicidad *geométrica* de $\lambda$ es la dimensión del correspondiente autoespacio $\mathcal{E}_{\lambda}(\mathbf{A})$.

## 16.2. Autovalores y autovectores de potencias de una matriz

Sea $\mathbf{A}$ de orden $n$ y sea $\lambda$ un autovalor de $\mathbf{A}$ con autovector asociado $\boldsymbol{v} \neq \mathbf{0}$, es decir,

$$\mathbf{A}\boldsymbol{v} = \lambda\boldsymbol{v}.$$

Demostraremos que para todo entero $m$ para el cual $\mathbf{A}^m$ esté definido[3]

$$\mathbf{A}^m\boldsymbol{v} = \lambda^m\boldsymbol{v}.$$

En consecuencia los autovectores de $\mathbf{A}$ son autovectores de $\mathbf{A}^m$ y

$$\text{Espectro}(\mathbf{A}^m) = \{\lambda^m : \lambda \in \text{Espectro}(\mathbf{A})\}.$$

**Caso** $m \geq 0$ Procedemos por inducción sobre $m$.

**Base de inducción:** $m = 0$**.** Por definición $\mathbf{A}^0 = \mathbf{I}$ (véase 3.F.1 en la página 46). Por tanto,

$$\mathbf{A}^0\boldsymbol{v} = \mathbf{I}\boldsymbol{v} = \boldsymbol{v}.$$

Como $\lambda^0 = 1$, concluimos que

$$\mathbf{A}^0\boldsymbol{v} = \lambda^0\boldsymbol{v}.$$

**Paso inductivo.** Supongamos que para algún $n \geq 0$ se cumple

$$\mathbf{A}^n\boldsymbol{v} = \lambda^n\boldsymbol{v}.$$

Entonces

$$\mathbf{A}^{n+1}\boldsymbol{v} = \mathbf{A}(\mathbf{A}^n\boldsymbol{v});$$

y, usando la hipótesis inductiva, tenemos:

$$\mathbf{A}^{n+1}\boldsymbol{v} = \mathbf{A}(\lambda^n\boldsymbol{v}).$$

Como $\lambda^n$ es escalar,

$$\mathbf{A}^{n+1}\boldsymbol{v} = \lambda^n\mathbf{A}\boldsymbol{v}.$$

Finalmente, usando que $\mathbf{A}\boldsymbol{v} = \lambda\boldsymbol{v}$,

$$\mathbf{A}^{n+1}\boldsymbol{v} = \lambda^n(\lambda\boldsymbol{v}) = \lambda^{n+1}\boldsymbol{v}.$$

En consecuencia,

$$\mathbf{A}^m\boldsymbol{v} = \lambda^m\boldsymbol{v} \qquad \text{para todo } m \geq 0.$$

[3] Fíjese que no estará definido cuando $m$ es negativo y $\mathbf{A}$ es singular.

**Caso** $m < 0$ Supongamos ahora que $\mathbf{A}$ es invertible. Entonces $\lambda \neq 0$. Como

$$\mathbf{A}\boldsymbol{v} = \lambda \boldsymbol{v},$$

multiplicando por $\mathbf{A}^{-1}$ tenemos

$$(\mathbf{A}^{-1})\mathbf{A}\boldsymbol{v} = (\mathbf{A}^{-1})(\lambda \boldsymbol{v}).$$

Por tanto,

$$\boldsymbol{v} = \lambda(\mathbf{A}^{-1})\boldsymbol{v}.$$

Como $\lambda \neq 0$,

$$(\mathbf{A}^{-1})\boldsymbol{v} = \lambda^{-1}\boldsymbol{v}.$$

Aplicando ahora el resultado anterior a $\mathbf{A}^{-1}$, obtenemos para todo $k \geq 0$:

$$(\mathbf{A}^{-1})^k\boldsymbol{v} = (\lambda^{-1})^k\boldsymbol{v}.$$

Pero $(\mathbf{A}^{-1})^k = \mathbf{A}^{-k}$, así que

$$\mathbf{A}^{-k}\boldsymbol{v} = \lambda^{-k}\boldsymbol{v}.$$

Por tanto, $\mathbf{A}^m\boldsymbol{v} = \lambda^m\boldsymbol{v}$ para todo entero $m$ para el cual $\mathbf{A}^m$ esté definido.

**Conclusión** .

- Si $m \geq 0$, el resultado vale para cualquier matriz cuadrada.
- Si $m < 0$, se requiere que $\mathbf{A}$ sea invertible.
- En todos los casos, los autovectores de $\mathbf{A}$ son autovectores de $\mathbf{A}^m$ y

$$\text{Espectro}(\mathbf{A}^m) = \{\lambda^m : \lambda \in \text{Espectro}(\mathbf{A})\}.$$

# Diagonalización en bloques triangulares por semejanza

☞ En esta lección veremos que toda matriz cuadrada se puede transformar en una matriz que es diagonal **por bloques triangulares**, es decir, tal que cada submatriz (cada bloque) en la diagonal principal es una matriz triangular; además, cada uno de dichos bloques tendrá un mismo autovalor repetido en su correspondiente diagonal principal. También veremos la condición necesaria para que los bloques sean de tamaño 1, es decir, para que la matriz sea *diagonalizable*.

Pero primero es *fundamental* que usted lea la Sección 4.B en la página 65 sobre las transformaciones *"espejo"*.

## 17.1. Matrices semejantes

**Definición 17.1.** *Decimos que* $\mathbf{A}$ *y* $\mathbf{C}$ *(del mismo orden) son* semejantes *si existe* $\mathbf{S}$*, invertible, tal que*

$$\mathbf{A} = \mathbf{S}^{-1}\mathbf{C}\mathbf{S}.$$

Las *matrices semejantes* comparten muchas características; por ejemplo, *tienen el mismo determinante*:

$$\det \mathbf{A} \;=\; \det\left(\mathbf{S}^{-1}\mathbf{C}\mathbf{S}\right) \;=\; \det \mathbf{S}^{-1}\cdot\det\mathbf{C}\cdot\det\mathbf{S} \;=\; \det\mathbf{C}; \qquad \text{ya que } \det(\mathbf{S}^{-1}) = \left(\det\mathbf{S}\right)^{-1}.$$

Las matrices semejantes también tienen idéntico polinomio característico:

**Proposición 17.1.1.** *Si* $\mathbf{A}$ *y* $\mathbf{C}$ *son semejantes, entonces tienen el mismo polinomio característico.*

*Demostración.* Puesto que son semejantes, existe una matriz invertible $\mathbf{S}$ tal que $\mathbf{C} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$, entonces

$$\begin{aligned}
P_{\mathbf{C}}(\lambda) = \det(\mathbf{C}-\lambda\mathbf{I}) &= \det\left(\mathbf{S}^{-1}\mathbf{A}\mathbf{S} - \lambda\mathbf{S}^{-1}\mathbf{S}\right) && \text{pues } \mathbf{C} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S} \text{ y } \mathbf{S}^{-1}\mathbf{S} = \mathbf{I}\\
&= \det\left(\mathbf{S}^{-1}(\mathbf{A}-\lambda\mathbf{I})\mathbf{S}\right) && \text{sacando factor común}\\
&= \det\mathbf{S}^{-1}\cdot\det(\mathbf{A}-\lambda\mathbf{I})\cdot\det\mathbf{S} && \text{determinante de un producto de matrices}\\
&= \det(\mathbf{A}-\lambda\mathbf{I}) = P_{\mathbf{A}}(\lambda) && \text{ya que } \det(\mathbf{S}^{-1}) = \left(\det\mathbf{S}\right)^{-1}.
\end{aligned}$$

□

Y puesto que cada raíz del polinomio característico es un autovalor de la matriz, y que la multiplicidad de cada raíz es la multiplicidad *algebraica* de cada autovalor, concluimos que *dos matrices semejantes tienen los mismos autovalores y con la misma multiplicidad algebraica.*

Además, *los autovalores de dos matrices semejantes también tienen la misma multiplicidad geométrica.* Es fácil deducirlo. Sea $\mathbf{A}$ de orden $n$; sabemos que, si $\mathbf{S}$ es invertible y del mismo orden, entonces $\mathcal{C}\left(\mathbf{A}\mathbf{S}\right) = \mathcal{C}\left(\mathbf{A}\right)$ , y por tanto las formas escalonadas de las matrices $\mathbf{A}\mathbf{S}$ y $\mathbf{A}$ tienen el mismo número de columnas y

de pivotes; por tanto $\dim\mathcal{N}\left(\mathbf{AS}\right) = \dim\mathcal{N}\left(\mathbf{A}\right)$. Por otra parte, sabemos que si $\mathbf{S}$ es invertible, $\mathcal{N}\left(\mathbf{SA}\right) = \mathcal{N}\left(\mathbf{A}\right)$.

Combinando ambos resultados y teniendo en cuenta que

$$\mathbf{S}^{-1}\mathbf{AS} - \lambda_i\mathbf{I} \quad = \quad \mathbf{S}^{-1}\mathbf{AS} - \lambda_i\mathbf{S}^{-1}\mathbf{S} \quad = \quad \mathbf{S}^{-1}(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{S},$$

tenemos que para cada autovalor $\lambda_i$ de $\mathbf{A}$,

$$\dim\mathcal{N}\left(\mathbf{S}^{-1}\mathbf{AS} - \lambda_i\mathbf{I}\right) \quad = \quad \dim\mathcal{N}\left(\mathbf{S}^{-1}(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{S}\right) \quad = \quad \dim\mathcal{N}\left(\mathbf{A} - \lambda_i\mathbf{I}\right).$$

Recordemos que la *traza* de $\underset{m\times n}{\mathbf{A}}$ es la suma de los elementos de su diagonal principal, es decir:

$$\operatorname{tr}\left(\mathbf{A}\right) = a_{11} + a_{22} + \cdots + a_{kk}, \qquad \text{donde } k = \min(m, n).$$

Pues bien, *dos matrices semejantes también tienen la misma traza.* Para demostrarlo comenzamos multiplicando $\mathbf{A}$ por la matriz elemental $\mathbf{I}_{\boldsymbol{\tau}}$ por la derecha y por su inversa por la izquierda. Evidentemente, la matriz resultante es similar a $\mathbf{A}$. Veamos que la traza no cambia [1]...

**Proposición 17.1.2.** *Si* $\mathbf{A} = (\mathbf{I}_{\boldsymbol{\tau}^{-1}})\mathbf{B}(\mathbf{I}_{\boldsymbol{\tau}})$, *es decir, si* $\mathbf{A} = {}_{\operatorname{esp}(\boldsymbol{\tau}^{-1})}\mathbf{B}_{\boldsymbol{\tau}}$, *entonces* $\operatorname{tr}\left(\mathbf{A}\right) = \operatorname{tr}\left(\mathbf{B}\right)$.

*Demostración.* Veamos que la traza se mantiene tanto con transformaciones *Tipo I* como *Tipo II*:

**Si $\boldsymbol{\tau}$ es de Tipo II** y $\boldsymbol{\tau}$ multiplica por $\alpha$ a la columna $\boldsymbol{i}$-ésima, entonces $\operatorname{esp}(\boldsymbol{\tau}^{-1})$ dividirá por $\alpha$ la fila $\boldsymbol{i}$-ésima. Por tanto, el $\boldsymbol{i}$-ésimo componente de la diagonal no cambiará. Veámoslo:

$$\mathbf{A}\Big(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{ji}]}{\boldsymbol{\tau}}}\Big)_{|j} \;=\; \Big(\mathbf{A}_{\underset{[(\alpha)\boldsymbol{ji}]}{\boldsymbol{\tau}}}\Big)_{|j} \;=\; \alpha(\mathbf{A}_{|j});$$

pero al aplicar la inversa también tenemos

$${}_{j|}\Big(\underset{\left[\left(\frac{1}{\alpha}\right)\boldsymbol{ji}\right]}{\boldsymbol{\tau}}\mathbf{I}\Big)\mathbf{A} \;=\; {}_{j|}\Big(\underset{\left[\left(\frac{1}{\alpha}\right)\boldsymbol{ji}\right]}{\boldsymbol{\tau}}\mathbf{A}\Big) \;=\; \frac{1}{\alpha}({}_{j|}\mathbf{A}).$$

Como la primera operación multiplica la componente ${}_{j|}\mathbf{A}_{|j}$ por $\alpha$ y la segunda la divide por $\alpha$, la diagonal no cambia; y por tanto tampoco cambia la traza.

**Si $\boldsymbol{\tau}$ es de Tipo I** y $\boldsymbol{\tau}$ suma $\alpha$ veces la columna $\boldsymbol{i}$-ésima a la $\boldsymbol{j}$-ésima, entonces la inversa de la correspondiente matriz elemental $\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}$ es la matriz elemental $\mathbf{I}_{\underset{[(-\alpha)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}}$; y por tanto

$$\Big(\mathbf{A}\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\Big)_{|j} = \alpha(\mathbf{A}_{|i}) + \mathbf{A}_{|j} \quad \Rightarrow \quad {}_{j|}\Big(\mathbf{A}\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\Big)_{|j} = \alpha({}_{j|}\mathbf{A}_{|i}) + {}_{j|}\mathbf{A}_{|j},$$

pero, por otra parte,

$${}_{i|}\Big(\underset{[(-\alpha)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}\mathbf{I}\mathbf{A}\Big) = -\alpha({}_{j|}\mathbf{A}) + {}_{i|}\mathbf{A} \quad \Rightarrow \quad {}_{i|}\Big(\underset{[(-\alpha)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}\mathbf{I}\mathbf{A}\Big)_{|i} = -\alpha({}_{j|}\mathbf{A}_{|i}) + {}_{i|}\mathbf{A}_{|i}.$$

Así pues, estas transformaciones cambian los elementos $\boldsymbol{j}$-ésimo e $\boldsymbol{i}$-ésimo de la diagonal, a uno se le suma $\alpha_{\,j|}\mathbf{A}_{|i}$ y al otro se le resta $\alpha_{\,j|}\mathbf{A}_{|i}$. Por tanto (aunque cambia la diagonal), la traza no cambia. □

... y puesto que las matrices invertibles son producto de matrices elementales, tenemos el siguiente

**Corolario 17.1.3.** *Si* $\mathbf{A}$ *y* $\mathbf{B}$ *son semejantes, entonces tienen la misma traza.*

[1] repase las transformaciones *espejo* en la Sección 4.B en la página 65.

Módulo NAcAL para Python

```
A  = Matrix([[1,1,1],[1,1,1],[1,1,1]])
Tr = T((5,2)) & T((-2,3,1))
(Tr**-1).espejo() & Matrix(A) & Tr              # No cambia la traza
```

☞ Y ahora veamos el resultado más importante de la lección: que dada $\mathbf{A}$ de orden $n$, y *conocidos sus autovalores*, siempre es posible transformar $\mathbf{A}$ en una matriz diagonal por bloques triangulares similar a $\mathbf{A}$ y cuya diagonal contiene todos los autovalores de $\mathbf{A}$.

De propina, deduciremos que la suma de los autovalores es la traza y que su producto es el determinante.

**Nota 4.** *¡Recuerde que para obtener matrices semejantes hay que operar tanto con las filas como con las columnas! (... si multiplicamos por la derecha con $\mathbf{S}$ debemos multiplicar por la izquierda con $\mathbf{S}^{-1}$).*

## 17.2. Diagonalización por bloques triangulares

**Teorema 17.2.1.** *Para toda matriz cuadrada $\mathbf{A} \in \mathbb{C}^{n\times n}$ existe una matriz invertible $\mathbf{S}$ tal que*

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \begin{bmatrix} \boxed{\mathbf{T}_{\lambda_k}} & & \\ & \ddots & \\ & & \boxed{\mathbf{T}_{\lambda_1}} \end{bmatrix} \text{ donde } \mathbf{T}_{\lambda_i} = \begin{bmatrix} \lambda_i & & & \\ * & \lambda_i & & \\ * & * & \ddots & \\ * & * & \cdots & \lambda_i \end{bmatrix} \text{ es triangular y de orden igual a la multiplicidad}$$

*algebraica de $\lambda_i$, y donde $\{\lambda_1, \ldots, \lambda_k\}$ es el* espectro de $\mathbf{A}$*, es decir, el conjunto*[2] *de autovalores de $\mathbf{A}$.*

☞ Antes de demostrar este importante teorema (la demostración aparece en la página 205), veamos algunos resultados previos que nos ayudarán a entender los pasos del algoritmo de diagonalización por bloques.

**Lema 17.2.2** (De paso inicial)**.** *Si $\mathbf{A} \in \mathbb{C}^{n\times n}$ es de la forma*

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{C} & \\ \hline * & \mathbf{L} \end{array}\right],$$

*donde $\mathbf{C}$ (de orden $m$) es singular y $\mathbf{L}$ es triangular inferior sin ceros en la diagonal principal, entonces existe una matriz invertible $\mathbf{S}$ tal que*

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \left[\begin{array}{c|c|c} \mathbf{C'} & & \\ \hline & \boxed{0} & \\ \hline * & & \mathbf{L} \end{array}\right], \quad \text{donde } \mathbf{C'} \text{ es de orden } (m-1).$$

*Y si $\mathbf{A}$ es simplemente singular, entonces* $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \left[\begin{array}{c|c} \mathbf{C'} & \\ \hline * & 0 \end{array}\right]$, *donde $\mathbf{C'}$ es de orden $(m-1)$.*

*Demostración. Etapa 1* [*Anulando la última columna de $\mathbf{C}$ (su columna $m$-ésima)*]. Como $\mathbf{C}$ (de orden $m$) es singular, podemos anular su última columna por eliminación Gaussiana usando una sucesión de

[2] Asumimos que en el conjunto $\{\lambda_1, \ldots, \lambda_k\}$ no hay elementos repetidos.

transformaciones elementales, $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathbf{R}$, que involucran únicamente a las $m$ primeras columnas de $\mathbf{A}$. Por tanto, al aplicar las correspondientes transformaciones inversas *"espejo"* a las filas, ${}_{\mathrm{esp}(\boldsymbol{\tau}_1^{-1}\ldots\boldsymbol{\tau}_k^{-1})}\mathbf{I} = \mathbf{R}^{-1}$, únicamente modificaremos las primeras $m$ filas de $\mathbf{A}$ (todas ellas con un cero en la posición $m$-ésima). Así obtenemos una matriz de la forma

$$ {}_{\mathrm{esp}(\boldsymbol{\tau}_1^{-1}\ldots\boldsymbol{\tau}_k^{-1})}\mathbf{A}_{(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)} = \mathbf{R}^{-1}\mathbf{A}\mathbf{R} = \left[\begin{array}{c|c|ccccc} & 0 & & & & & \\ \ast & \vdots & & & & & \\ {\scriptstyle m\times(m-1)} & 0 & & & & & \\ \hline & d_{m+1} & \beta_{m+1} & & & & \\ & d_{m+2} & \ast & \beta_{m+2} & & & \\ \ast\ast & \vdots & \ast & \ast & \ddots & & \\ & d_n & \ast & \ast & \cdots & \beta_n \end{array}\right]. $$

*Etapa 2* [*Anulando el resto de coeficientes de la columna $m$-ésima de $\mathbf{A}$ (aquellos a la izquierda de $\mathbf{L}$)*]. Gracias a que las componentes $\beta_j$ de la diagonal principal de $\mathbf{L}$ son pivotes, mediante una sucesión de transformaciones elementales $\mathbf{I}_{\boldsymbol{\tau}_{(k+1)}\cdots\boldsymbol{\tau}_p} = \mathbf{P}$, $\Big($del tipo $\underset{[(\alpha_j)\boldsymbol{j}+\boldsymbol{m}]}{\boldsymbol{\tau}}$ con $j > m\Big)$, se pueden anular las componentes $d_{m+1},\ldots,d_n$ de la columna $m$-ésima. Al aplicar la sucesión de las correspondientes transformaciones inversas "*espejo*", ${}_{\mathrm{esp}(\boldsymbol{\tau}_{k+1}^{-1}\cdots\boldsymbol{\tau}_p^{-1})}\mathbf{I} = \mathbf{P}^{-1}$, $\Big($del tipo $\underset{[(-\alpha_j)\boldsymbol{m}+\boldsymbol{j}]}{\boldsymbol{\tau}}$ con $j > m\Big)$, solo varían las columnas correspondientes a los asteriscos "$\ast\ast$", ya que la fila $m$-ésima contiene únicamente ceros a partir de la posición $m$; resultando la matriz

$$ {}_{\mathrm{esp}(\boldsymbol{\tau}_1^{-1}\ldots\boldsymbol{\tau}_k^{-1},\boldsymbol{\tau}_{(k+1)}^{-1}\cdots\boldsymbol{\tau}_p^{-1})}\mathbf{A}_{(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k,\boldsymbol{\tau}_{k+1}\cdots\boldsymbol{\tau}_p)} = \big((\mathbf{P}^{-1})(\mathbf{R}^{-1})\big)\mathbf{A}\big(\mathbf{R}\mathbf{P}\big) = \left[\begin{array}{c|cc} \mathbf{C'} & & \\ \hline & 0 & \\ \ast & \hline 0 & \mathbf{L} \end{array}\right], \quad \text{donde } \mathbf{P} = \mathbf{I}_{\boldsymbol{\tau}_{(k+1)}\cdots\boldsymbol{\tau}_p}. $$

Llamando $\mathbf{S}$ a la matriz $\mathbf{R}\mathbf{P} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p}$ hemos terminado la demostración del primer caso.

Demostrar el caso en que $\mathbf{A}$ es simplemente singular es más sencillo... basta con aplicar la *Etapa 1*. □

**Lema 17.2.3** (Paso de continuación)**.** *Si $\mathbf{A} \in \mathbb{C}^{n\times n}$ es de la forma*

$$ \mathbf{A} = \left[\begin{array}{c|cc} \mathbf{C} & & \\ \hline \ast & \mathbf{T} & \\ \hline \ast & & \mathbf{L} \end{array}\right], $$

*donde $\mathbf{C}$ (de orden $k$) es singular, donde $\mathbf{L}$ (de orden $n-m$) es triangular inferior sin ceros en la diagonal principal y donde $\mathbf{T}$ es triangular inferior con la diagonal principal llena de ceros, entonces existe $\mathbf{S}$ (invertible) tal que*

$$ \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \left[\begin{array}{c|cc} \mathbf{C'} & & \\ \hline \ast & \mathbf{T'} & \\ \hline \ast & & \mathbf{L} \end{array}\right], $$

*donde $\mathbf{C'}$ es de orden $(k-1)$ y $\mathbf{T'}$ es triangular inferior con la diagonal principal llena de ceros.*

*Y si $\mathbf{A}$ es simplemente de la forma* $\left[\begin{array}{c|c} \mathbf{C} & \\ \hline \ast & \mathbf{T} \end{array}\right]$, *entonces* $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \left[\begin{array}{c|c} \mathbf{C'} & \\ \hline \ast & \mathbf{T'} \end{array}\right]$, *donde $\mathbf{C'}$ es de orden $k-1$ y $\mathbf{T'}$ es triangular inferior con la diagonal principal llena de ceros.*

*Demostración. Etapa 1* [*Anulando la última columna de* $\mathbf{C}$ (*su columna* $k$*ésima*)]. Aplicando la eliminación Gaussiana como en la *Etapa 1* del lema anterior obtenemos una matriz de la forma

$$\left[\begin{array}{c|c|c|c} & 0 & & \\ \ast & \vdots & & \\ {\scriptstyle k\times(k-1)} & 0 & & \\ \hline \ast & \ast & \mathbf{T} & \\ \hline & d_{m+1} & & \beta_{m+1} \\ & d_{m+2} & & \ast \quad \beta_{m+2} \\ \ast & \vdots & & \ast \quad \ast \quad \ddots \\ & d_n & & \ast \quad \ast \quad \cdots \quad \beta_n \end{array}\right], \quad \text{donde} \quad \mathbf{T'} = \left[\begin{array}{c|c} 0 & \\ \hline \ast & \mathbf{T} \end{array}\right].$$

*Etapa 2* [*Anulando coeficientes de la columna* $k$*ésima que están a la izquierda de* $\mathbf{L}$]. Hacemos lo mismo que en la *Etapa 2* del lema anterior quedando una matriz de la forma

$$\left[\begin{array}{c|c|c|c} \mathbf{C'} & & & \\ \hline & 0 & & \\ \ast & \ast & \mathbf{T} & \\ \hline & \mathbf{0} & & \mathbf{L} \end{array}\right], \quad \text{donde} \quad \mathbf{T'} = \left[\begin{array}{c|c} 0 & \\ \hline \ast & \mathbf{T} \end{array}\right].$$

Y si $\mathbf{A}$ es simplemente de la forma $\left[\begin{array}{c|c} \mathbf{C'} & \\ \hline \ast & \mathbf{L} \end{array}\right]$, entonces basta aplicar la *Etapa 1*. □

Antes de pasar al siguiente corolario, recuérdese que para una *matriz triangular por bloques* (véase el Ejercicio 85 en la página 182)

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{B} & \mathbf{0} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right] \qquad \text{se verifica que} \qquad |\mathbf{A}| = |\mathbf{B}| \cdot |\mathbf{D}|;$$

por tanto, el polinomio característico de la matriz triangular por bloques $\mathbf{A}$ es igual al producto de los polinomios característicos de $\mathbf{B}$ y $\mathbf{D}$; es decir $P_{\mathbf{A}}(\lambda) = P_{\mathbf{B}}(\lambda) \cdot P_{\mathbf{D}}(\lambda)$, o expresado con determinantes:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det(\mathbf{B} - \lambda\mathbf{I}) \cdot \det(\mathbf{D} - \lambda\mathbf{I}).$$

☞ Usaremos este resultado sobre polinomios característicos en la última parte de la demostración del siguiente corolario, que nos indica cómo iniciar el algoritmo de diagonalización para generar un primer bloque:

**Corolario 17.2.4.** *Si* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *y* $\lambda_1$ *es un autovalor de* $\mathbf{A}$, *entonces existe* $\mathbf{S}$ *invertible tal que*

$$o\ bien \qquad \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \left[\begin{array}{c|c} \mathbf{C'} & \\ \hline \ast & \mathbf{T}_{\lambda_1} \end{array}\right]; \qquad o\ bien \qquad \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \mathbf{T}_{\lambda_1};$$

*donde* $\mathbf{T}_{\lambda_1} = \begin{bmatrix} \lambda_1 & & & \\ \ast & \lambda_1 & & \\ \ast & \ast & \ddots & \\ \ast & \ast & \cdots & \lambda_1 \end{bmatrix}$ *es de orden igual a la multiplicidad* algebraica *de* $\lambda_1$.

*Demostración.* Como $(\mathbf{A}-\lambda_1\mathbf{I})$ es singular, aplicando el Lema *de paso inicial* e iterando el Lema *de paso de continuación* mientras sea posible (mientras la submatriz de la esquina superior izquierda sea singular), llegamos a

$$(\mathbf{S}_k^{-1}\cdots\mathbf{S}_1^{-1})(\mathbf{A}-\lambda_1\mathbf{I})(\mathbf{S}_1\cdots\mathbf{S}_k) = \begin{cases} \left[\mathbf{T}\right] & \text{si pudimos continuar hasta el final} \\ \left[\begin{array}{c|c} \mathbf{C'} & \\ \hline * & \mathbf{T} \end{array}\right] & \text{si topamos con una submatriz } \mathbf{C'} \text{ invertible} \end{cases}, \quad (17.1)$$

donde en ambos casos $\mathbf{T}$ es triangular inferior de orden $k$, con la diagonal llena de ceros (donde $k$ es el número de *pasos* que hemos dado). Como para cualquier $\mathbf{S}$ invertible se verifica que $\mathbf{S}^{-1}(\mathbf{A}-\lambda_1\mathbf{I})\mathbf{S} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}-\lambda_1\mathbf{I}$, sumando $\lambda_1\mathbf{I}$ en (17.1) tenemos

$$(\mathbf{S}_k^{-1}\cdots\mathbf{S}_1^{-1})\mathbf{A}(\mathbf{S}_1\cdots\mathbf{S}_k) = \begin{cases} \left[\mathbf{T}+\lambda_1\mathbf{I}\right] & \\ \left[\begin{array}{c|c} \mathbf{C'}+\lambda_1\mathbf{I} & \\ \hline * & \mathbf{T}+\lambda_1\mathbf{I} \end{array}\right] & \end{cases} ; \quad \text{donde } \mathbf{T}_{\lambda_1} = \left[\mathbf{T}+\lambda_1\mathbf{I}\right].$$

En el primer caso $k=n$ y el polinomio característico de la matriz $\mathbf{A}$ coincide con el polinomio característico de $\mathbf{T}_{\lambda_1}$, que es $(\lambda_1-\lambda)^n$, donde $n$ es el orden tanto de la matriz $\mathbf{T}$ como de $\mathbf{A}$.

En el segundo caso, el polinomio característico de $\mathbf{A}$ es igual al producto de los polinomios característicos de las dos submatrices, es decir, $\big(P_{(\mathbf{C'}+\lambda_1\mathbf{I})}(\lambda)\big)\cdot(\lambda_1-\lambda)^k$. Ahora bien, como $\lambda_1$ no es un autovalor[3] de $(\mathbf{C'}+\lambda_1\mathbf{I})$, entonces $\boxed{k \text{ es la multiplicidad del autovalor } \lambda_1}$. □

☞ … el último corolario nos dice cómo continuar el algoritmo para seguir generando el resto de bloques en la diagonal (note que es casi idéntico al anterior)…

**Corolario 17.2.5.** *Si* $\mathbf{A}\in\mathbb{C}^{n\times n}$ *de la forma* $\left[\begin{array}{c|c} \mathbf{C} & \\ \hline & \boxed{\mathbf{T}_{\lambda_r}} \\ * & \quad\ddots \\ & \qquad\boxed{\mathbf{T}_{\lambda_1}} \end{array}\right]$, *donde* $\mathbf{T}_{\lambda_i} = \begin{bmatrix} \lambda_i & & & \\ * & \lambda_i & & \\ * & * & \ddots & \\ * & * & \cdots & \lambda_i \end{bmatrix}$ *es de orden igual a la multiplicidad* algebraica *del autovalor* $\lambda_i$ *y donde* $\lambda_{r+1}$ *es un autovalor*[4] *de* $\mathbf{C}$, *entonces existe* $\mathbf{S}$ *invertible tal que*

$$o\ bien \quad \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \left[\begin{array}{c|c} \mathbf{C'} & \\ \hline & \boxed{\mathbf{T}_{\lambda_{r+1}}} \\ * & \quad\ddots \\ & \qquad\boxed{\mathbf{T}_{\lambda_1}} \end{array}\right], \quad o\ bien \quad \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \left[\begin{array}{c|c} \mathbf{T}_{\lambda_{r+1}} & \\ \hline & \boxed{\mathbf{T}_{\lambda_r}} \\ * & \quad\ddots \\ & \qquad\boxed{\mathbf{T}_{\lambda_1}} \end{array}\right],$$

*donde* $\mathbf{T}_{\lambda_{r+1}} = \begin{bmatrix} \lambda_{r+1} & & & \\ * & \lambda_{r+1} & & \\ * & * & \ddots & \\ * & * & \cdots & \lambda_{r+1} \end{bmatrix}$ *es de orden igual a la multiplicidad* algebraica *de* $\lambda_{r+1}$.

[3] si fuera autovalor, entonces $\mathbf{C'}$ sería singular.
[4] por tanto también es un autovalor de $\mathbf{A}$ distinto de $\lambda_1,\ldots,\lambda_r$

*Demostración.* Como $(\mathbf{C}-\lambda_{r+1}\mathbf{I})$ es singular, repitiendo los pasos de la anterior demostración:

$$(\mathbf{S}_k^{-1}\cdots\mathbf{S}_1^{-1})\mathbf{A}(\mathbf{S}_1\cdots\mathbf{S}_k) = \begin{cases} \left[\begin{array}{c|c} \mathbf{T}_{\lambda_{r+1}} & \\ \hline & \begin{array}{ccc}\boxed{\mathbf{T}_{\lambda_r}} & & \\ & \ddots & \\ & & \boxed{\mathbf{T}_{\lambda_1}}\end{array}\end{array}\right] & \text{si pudimos continuar hasta el final} \\ \left[\begin{array}{c|c} \mathbf{C}'+\lambda_{r+1}\mathbf{I} & \\ \hline & \begin{array}{ccc}\boxed{\mathbf{T}_{\lambda_{r+1}}} & & \\ & \ddots & \\ & & \boxed{\mathbf{T}_{\lambda_1}}\end{array}\end{array}\right] & \text{si topamos con una submatriz } \mathbf{C}' \text{ invertible} \end{cases},$$

donde en ambos casos $\mathbf{T}_{\lambda_{r+1}}$ es triangular inferior de orden $k$, con $\lambda_{r+1}$ en la diagonal principal y donde $k$ es el número de *pasos* que hemos dado.

En el primer caso el polinomio característico de la matriz $\mathbf{A}$ es el producto de los polinomios característicos

$$P_{\mathbf{A}}(\lambda) = P_{\mathbf{T}_{\lambda_1}}(\lambda)\cdots P_{\mathbf{T}_{\lambda_r}}(\lambda)\cdot P_{\mathbf{T}_{\lambda_{r+1}}}(\lambda) = (\lambda_1-\lambda)^{\mu(\lambda_1)}\cdots(\lambda_r-\lambda)^{\mu(\lambda_r)}\cdot(\lambda_{r+1}-\lambda)^k,$$

donde $\mu(\lambda_i)$ es la multiplicidad *algebraica* del autovalor $\lambda_i$. En el segundo caso

$$P_{\mathbf{A}}(\lambda) = \left(P_{(\mathbf{C}'+\lambda_{r+1}\mathbf{I})}(\lambda)\right)\cdot(\lambda_1-\lambda)^{\mu(\lambda_1)}\cdots(\lambda_r-\lambda)^{\mu(\lambda_r)}\cdot(\lambda_{r+1}-\lambda)^k.$$

Y como $\lambda_{r+1}$ no es un autovalor de $(\mathbf{C}'+\lambda_{r+1}\mathbf{I})$, entonces $\boxed{k \text{ es la multiplicidad del autovalor } \lambda_{r+1}}$. □

Ya solo resta demostrar el Teorema 17.2.1, pero...

*Demostración del Teorema 17.2.1.* Ya no hay nada que demostrar. Basta aplicar el primer corolario e iterar el segundo hasta finalizar la generación de bloques. □

*Ejemplo* 42. Sea $\mathbf{A}=\begin{bmatrix}1&-1&0\\0&0&0\\0&-2&1\end{bmatrix}$ con autovalores 1, 1 y 0. Vamos a diagonalizar por bloques triangulares:

$$\left[\begin{array}{c}\mathbf{A}\\\hline\mathbf{I}\end{array}\right] \xrightarrow{(-)} \left[\begin{array}{ccc|ccc}1&-1&0&1&0&0\\0&0&0&0&1&0\\0&-2&1&0&0&1\\\hline1&0&0&1&0&0\\0&1&0&0&0&1\\0&0&1&0&1&2\end{array}\right] \xrightarrow[\substack{[(1)\mathbf{1}+\mathbf{2}]\\[(2)\mathbf{3}+\mathbf{2}]\\[\mathbf{2}\rightleftharpoons\mathbf{3}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccc|ccc}1&0&0&1&0&0\\0&0&0&0&1&0\\0&1&0&0&0&1\\\hline1&0&1&1&0&0\\0&0&1&0&1&0\\0&1&2&0&0&1\end{array}\right] \xrightarrow[\substack{[\mathbf{2}\rightleftharpoons\mathbf{3}]\\[(-2)\mathbf{2}+\mathbf{3}]\\[(-1)\mathbf{2}+\mathbf{1}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccc|ccc}1&0&0&1&0&0\\0&1&0&0&0&1\\0&0&0&0&1&0\\\hline1&0&1&1&0&0\\0&0&1&0&1&0\\0&1&2&0&0&1\end{array}\right] \xrightarrow{(+)} \left[\begin{array}{ccc|ccc}1&0&0&1&0&0\\0&1&0&0&0&1\\0&0&0&0&1&0\\\hline1&0&1&1&0&0\\0&0&1&0&1&0\\0&1&2&0&0&1\end{array}\right] = \left[\begin{array}{c}\mathbf{D}\\\hline\mathbf{S}\end{array}\right]$$

Así tenemos la matriz diagonal $\mathbf{D}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix} = \begin{bmatrix}1&0&1\\0&0&1\\0&1&2\end{bmatrix}^{-1}\begin{bmatrix}1&-1&0\\0&0&0\\0&-2&1\end{bmatrix}\begin{bmatrix}1&0&1\\0&0&1\\0&1&2\end{bmatrix} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$.

**Módulo NAcAL para Python**

```
A = Matrix([[1,-1,0],[0,0,0],[0,-2,1]]); L=[1,1,0];
DiagonalizaS(A, L , 1)
```

*Ejemplo* 43. Sea la matriz $\mathbf{A} = \begin{bmatrix} -2 & 0 & 3 \\ 3 & -2 & -9 \\ -1 & 2 & 6 \end{bmatrix}$ con autovalores 1, 1 y 0. Vamos a diagonalizar por bloques.

$$\left[\begin{array}{c}\mathbf{A}\\\hline \mathbf{I}\end{array}\right] \xrightarrow[1\mathbf{I}]{(-)} \left[\begin{array}{ccc} -3 & 0 & 3 \\ 3 & -3 & -9 \\ -1 & 2 & 5 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow[[(-2)\mathbf{2}+\mathbf{3}]]{\tau\ [(1)\mathbf{1}+\mathbf{3}]} \left[\begin{array}{ccc} -3 & 0 & 0 \\ 3 & -3 & 0 \\ -1 & 2 & 0 \\ \hline 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow[[(-1)\mathbf{3}+\mathbf{1}]]{\tau\ [(2)\mathbf{3}+\mathbf{2}]} \left[\begin{array}{ccc} -2 & -2 & 0 \\ 1 & 1 & 0 \\ -1 & 2 & 0 \\ \hline 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow[1\mathbf{I}]{(+)} \left[\begin{array}{ccc} -1 & -2 & 0 \\ 1 & 2 & 0 \\ -1 & 2 & 1 \\ \hline 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[1\mathbf{I}]{(-)} \left[\begin{array}{ccc} -2 & -2 & 0 \\ 1 & 1 & 0 \\ -1 & 2 & 0 \\ \hline 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow[[(-1)\mathbf{1}+\mathbf{2}]]{\tau} \left[\begin{array}{ccc} -2 & 0 & 0 \\ 1 & 0 & 0 \\ -1 & 3 & 0 \\ \hline 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow[[(1)\mathbf{2}+\mathbf{1}]]{\tau} \left[\begin{array}{ccc} -1 & 0 & 0 \\ 1 & 0 & 0 \\ -1 & 3 & 0 \\ \hline 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow[1\mathbf{I}]{(+)} \left[\begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 3 & 1 \\ \hline 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[0\mathbf{I}]{(-)} \left[\begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 3 & 1 \\ \hline 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow[[(4)\mathbf{3}+\mathbf{1}]]{\tau\ [(-1)\mathbf{2}+\mathbf{1}]} \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \\ \hline 6 & -1 & 1 \\ -9 & 1 & -2 \\ 4 & 0 & 1 \end{array}\right] \xrightarrow[[(1)\mathbf{1}+\mathbf{2}]]{\tau\ [(-4)\mathbf{1}+\mathbf{3}]} \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \\ \hline 6 & -1 & 1 \\ -9 & 1 & -2 \\ 4 & 0 & 1 \end{array}\right] \xrightarrow[0\mathbf{I}]{(+)} \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \\ \hline 6 & -1 & 1 \\ -9 & 1 & -2 \\ 4 & 0 & 1 \end{array}\right] = \left[\begin{array}{c}\mathbf{C}\\\hline \mathbf{S}\end{array}\right]$$

Así pues,

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 6 & -1 & 1 \\ -9 & 1 & -2 \\ 4 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -2 & 0 & 3 \\ 3 & -2 & -9 \\ -1 & 2 & 6 \end{bmatrix} \begin{bmatrix} 6 & -1 & 1 \\ -9 & 1 & -2 \\ 4 & 0 & 1 \end{bmatrix} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}.$$

Por lo que hemos llegado a una matriz diagonal por bloques triangulares.

Módulo NAcAL para Python

```
A = Matrix([[-2,0,3],[3,-2,-9],[-1,2,6]]); L=[1,1,0];
C = DiagonalizaS (A, L, 1) # Matriz diagonal por bloques triangulares
S = C.S                    # C = (S**-1) * A * S  (La matriz S es un atributo de C)
(S**-1) * A * S            # Comprobación
```

Como se ha visto, el algoritmo es bastante pesado para ser calculado con "papel y lápiz"... ¡pero para eso se inventaron los ordenadores!

*Ejemplo* 44. $\mathbf{A} = \begin{bmatrix} 3 & 0 & -1 & 1 \\ 3 & 2 & -2 & 2 \\ 1 & -2 & 2 & 0 \\ -3 & -2 & 3 & -1 \end{bmatrix}$ con autovalores 2 (doble) y 1 (doble). Diagonalicemos por bloques:

Módulo NAcAL para Python

```
A  = Matrix([[3,0,-1,1],[3,2,-2,2],[1,-2,2,0],[-3,-2,3,-1]]); L=[1,1,2,2];
C = DiagonalizaS (A, L, 1) # Quitando el tercer argumento no se visualizan los pasos
C.S                        # La matriz S se guarda como atributo S
```

Por tanto, $\mathbf{C} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -2 \\ 1 & 0 & 2 & -2 \\ 1 & -1 & 2 & -2 \\ 1 & -1 & 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 3 & 0 & -1 & 1 \\ 3 & 2 & -2 & 2 \\ 1 & -2 & 2 & 0 \\ -3 & -2 & 3 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & -2 \\ 1 & 0 & 2 & -2 \\ 1 & -1 & 2 & -2 \\ 1 & -1 & 0 & 2 \end{bmatrix} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}.$

### 17.2.1. Autovalores, determinante y traza de matrices semejantes

Como las matrices semejantes tienen el mismo determinante, la misma traza y los mismos autovalores (con la misma multiplicidad aritmética y geométrica); y puesto que por el Teorema 17.2.1 sabemos que para toda matriz cuadrada podemos encontrar otra similar a ella, que es diagonal por bloques *con sus autovalores en la diagonal principal*, concluimos que:

**Corolario 17.2.6.** *La suma de los autovalores de* $\mathbf{A}$ *es igual a la traza de* $\mathbf{A}$.

**Corolario 17.2.7.** *El producto de los autovalores de* $\mathbf{A}$ *es igual al determinante de* $\mathbf{A}$.

### 17.2.2. Autovectores

Hemos visto que para toda matriz $\mathbf{A}$ cuadrada, existe una matriz $\mathbf{S}$ invertible tal que

$$\mathbf{C} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$$

es diagonal por bloques triangulares, y donde cada bloque triangular tiene repetido el mismo autovalor $\lambda_i$ en su diagonal principal. Fijémonos ahora en aquellas columnas de $\mathbf{C}$ en las que aparezca un autovalor $\lambda_i$ con únicamente ceros por debajo, es decir, $\mathbf{C}_{|j} = \lambda_i \mathbf{I}_{|j}$ (por ejemplo, las columnas $\mathbf{C}_{|2}$ y $\mathbf{C}_{|4}$ del último ejemplo). Puesto que $\mathbf{C} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$, tenemos que

$$\mathbf{A}\mathbf{S} = \mathbf{S}\mathbf{C} \quad \Longrightarrow \quad \mathbf{A}\mathbf{S}_{|j} = \mathbf{S}\mathbf{C}_{|j},$$

y como para dichas columnas $\mathbf{C}_{|j} = \lambda_i \mathbf{I}_{|j}$, tenemos que $\mathbf{S}\mathbf{C}_{|j} = \lambda_i \mathbf{S}\mathbf{I}_{|j} = \lambda_i \mathbf{S}_{|j}$ y, por tanto,

$$\mathbf{A}\big(\mathbf{S}_{|j}\big) = \lambda_i\big(\mathbf{S}_{|j}\big).$$

**Corolario 17.2.8.** *Si* $\mathbf{C} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ *es diagonal por bloques triangulares y* $\mathbf{C}_{|j} = \lambda_i \mathbf{I}_{|j}$, *entonces* $\mathbf{S}_{|j}$ *es un autovector asociado al autovalor* $\lambda_i$.

Así, usando el módulo NAcAL de Python con el último ejemplo, llegamos a $\begin{bmatrix} \mathbf{C} \\ \hline \mathbf{S} \end{bmatrix} = \left[\begin{array}{cccc} 2 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ \hline 0 & 0 & 0 & -2 \\ 1 & 0 & 2 & -2 \\ 1 & -1 & 2 & -2 \\ 1 & -1 & 0 & 2 \end{array}\right]$,

y deducimos que $\big(0, \quad 0, \quad -1, \quad -1,\big)$ es un autovector para $\lambda = 2$; y $\big(-2, \quad -2, \quad -2, \quad 2,\big)$ lo es para $\lambda = 1$.

Nótese que tras diagonalizar por bloques triangulares $\begin{bmatrix} \mathbf{A} \\ \hline \mathbf{I} \end{bmatrix} \xrightarrow[\text{esp}(\boldsymbol{\tau}_p^{-1} \ldots \boldsymbol{\tau}_1^{-1})]{\boldsymbol{\tau}_1 \ldots \boldsymbol{\tau}_p} \begin{bmatrix} \mathbf{C} \\ \hline \mathbf{S} \end{bmatrix}$ donde $\mathbf{S} = \mathbf{I}_{\boldsymbol{\tau}_1 \ldots \boldsymbol{\tau}_p}$, como $\mathbf{S}$ es invertible, necesariamente *los autovectores correspondientes a autovalores distintos son linealmente independientes*.

## 17.3. Matrices diagonalizables

**Definición 17.2** (Matriz diagonalizable)**.** *Se dice que* $\mathbf{A}$ *(de orden* $n$*) es* diagonalizable *si existe una matriz* $\mathbf{S}$ *tal que* $\mathbf{S}^{-1}\mathbf{AS}$ *es diagonal.*

*Ejemplo* 45.

$$\left[\begin{array}{c}\mathbf{A}\\\hline \mathbf{I}\end{array}\right] \xrightarrow{(-)} \left[\begin{array}{ccc|ccc} 3&4&2&1&0&0\\ 0&1&2&0&1&0\\ 0&0&0&0&0&1\end{array}\right] \xrightarrow[[(2)\mathbf{1}+\mathbf{3}]]{\boldsymbol{\tau}_{[(-2)\mathbf{2}+\mathbf{3}]}} \left[\begin{array}{ccc|ccc} 3&4&0&1&0&0\\ 0&1&0&0&1&0\\ 0&0&0&0&0&1\end{array}\right] \xrightarrow[[(2)\mathbf{3}+\mathbf{2}]]{\boldsymbol{\tau}_{[(-2)\mathbf{3}+\mathbf{1}]}} \left[\begin{array}{ccc|ccc} 3&4&0&1&0&0\\ 0&1&0&0&1&0\\ 0&0&0&0&0&1\end{array}\right] \xrightarrow{(+)} \left[\begin{array}{ccc|ccc} 3&4&0&1&0&0\\ 0&1&0&0&1&0\\ 0&0&0&0&0&1\end{array}\right]$$

$$\xrightarrow{(-)} \left[\begin{array}{ccc|ccc} 2&4&0&1&0&0\\ 0&0&0&0&1&0\\ \hline 0&0&-1&0&0&1\\ \hline 1&0&2&1&0&0\\ 0&1&-2&0&1&0\\ 0&0&1&0&0&1\end{array}\right] \xrightarrow[[(2)\mathbf{1}+\mathbf{2}]]{\boldsymbol{\tau}_{[(-2)\mathbf{1}+\mathbf{2}]}} \left[\begin{array}{ccc|ccc} 2&0&0&1&-2&0\\ 0&0&0&0&1&0\\ \hline 0&0&-1&0&0&1\\ \hline 1&-2&2&1&0&0\\ 0&1&-2&0&1&0\\ 0&0&1&0&0&1\end{array}\right] \xrightarrow[[(2)\mathbf{2}+\mathbf{1}]]{\boldsymbol{\tau}_{[(2)\mathbf{2}+\mathbf{1}]}} \left[\begin{array}{ccc|ccc} 2&0&0&1&-2&0\\ 0&0&0&0&1&0\\ \hline 0&0&-1&0&0&1\\ \hline 1&-2&2&1&0&0\\ 0&1&-2&0&1&0\\ 0&0&1&0&0&1\end{array}\right] \xrightarrow{(+)} \left[\begin{array}{ccc|ccc} 3&0&0&1&-2&0\\ 0&1&0&0&1&0\\ 0&0&0&0&0&1\\ \hline 1&-2&2&1&0&0\\ 0&1&-2&0&1&0\\ 0&0&1&0&0&1\end{array}\right] = \left[\begin{array}{c}\mathbf{D}\\\hline \mathbf{S}\end{array}\right]$$

Fíjese que

$$\mathbf{AS}_{|1} = \begin{bmatrix}3&4&2\\0&1&2\\0&0&0\end{bmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix} = 3\mathbf{S}_{|1};\quad \mathbf{AS}_{|2} = \begin{bmatrix}3&4&2\\0&1&2\\0&0&0\end{bmatrix}\begin{pmatrix}-2\\1\\0\end{pmatrix} = 1\mathbf{S}_{|2};\quad \mathbf{AS}_{|3} = \begin{bmatrix}3&4&2\\0&1&2\\0&0&0\end{bmatrix}\begin{pmatrix}2\\-2\\1\end{pmatrix} = 0\mathbf{S}_{|3}.$$

Cuando la matriz es diagonalizable, todas las columnas de $\mathbf{S}$ son autovectores (linealmente independientes), y por tanto, la multiplicidad *algebraica* de cada autovalor (el número de veces que aparece cada $\lambda_i$ en la diagonal) necesariamente coincide con la multiplicidad *geométrica* (el número de autovectores linealmente independientes asociados a dicho autovalor).

**Corolario 17.3.1.** *Si para cada autovalor* $\lambda_i$ *de* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *las multiplicidades algebraica y geométrica coinciden, la matriz es diagonalizable (es decir, los bloques de la diagonalización por bloques son diagonales).*

**Corolario 17.3.2.** *Si todos los autovalores de* $\mathbf{A}$ *son distintos entre sí, entonces* $\mathbf{A}$ *es diagonalizable.*

*Demostración.* Si la multiplicidad algebraica de cada autovalor es uno, entonces al diagonalizar por bloques, cada bloque resultará de orden 1. □

☞ Si $\mathbf{A}$ es diagonalizable e invertible, tenemos que

$$\mathbf{A}^{-1} = \left(\mathbf{SD}(\mathbf{S}^{-1})\right)^{-1} = \mathbf{SD}^{-1}(\mathbf{S}^{-1}).$$

Además, si seguimos el siguiente convenio

$$\mathbf{A}^0 = \mathbf{A}^{(1-1)} = \mathbf{A}^1(\mathbf{A}^{-1}) = \mathbf{I};$$

entonces, para toda $\mathbf{A}$ diagonalizable e invertible y todo número entero $n$: $\mathbf{A}^n = \mathbf{SD}^n(\mathbf{S}^{-1})$.

# Diagonalización ortogonal de matrices simétricas

## 18.1. Método de Gram-Schmidt

**Definición 18.1.** *Un sistema de vectores es ortogonal si cada uno de los vectores es ortogonal al resto.*

**Proposición 18.1.1.** *Un sistema ortogonal sin vectores nulos es linealmente independiente.*

*Demostración.* Sea $\mathbf{Z} = \begin{bmatrix}\boldsymbol{z}_1;\dots\boldsymbol{z}_n;\end{bmatrix}$ un sistema ortogonal sin vectores nulos. Si $\mathbf{0} = a_1\boldsymbol{z}_1 + \cdots + a_n\boldsymbol{z}_n$ entonces para cada $j = 1:n$

$$0 \;=\; \mathbf{0}\cdot\boldsymbol{z}_j \;=\; (a_1\boldsymbol{z}_1 + \cdots + a_n\boldsymbol{z}_n)\cdot\boldsymbol{z}_j \;=\; a_j\boldsymbol{z}_j\cdot\boldsymbol{z}_j,$$

y como $\boldsymbol{z}_j\cdot\boldsymbol{z}_j \neq 0$ (pues $\boldsymbol{z}_j \neq \mathbf{0}$), necesariamente $a_j = 0$. □

**Teorema 18.1.2** (Método de Gram-Schmidt)**.** *Dado un sistema de vectores, existe una sucesión de transformaciones elementales "de izquierda a derecha" que transforman el sistema en otro equivalente*[1] *ortogonal.*

*Demostración.* Lo demostraremos por inducción sobre el número de vectores del sistema.

Si el sistema contiene un único vector no hay nada que hacer. Veamos que si el resultado es cierto para sistemas de $n$ vectores, entonces también es cierto para sistemas de $n+1$ vectores.

Sea el sistema $\begin{bmatrix}\boldsymbol{y}_1;\dots\ \boldsymbol{y}_n;\ \boldsymbol{y}_{(n+1)}\end{bmatrix}$, aplicando la hipótesis de inducción, existe una sucesión, $\boldsymbol{\tau}_k,\dots,\boldsymbol{\tau}_n$, de transformaciones elementales de "izquierda a derecha" que transforman el sub-sistema formado por los $n$ primeros vectores en otro sistema $\begin{bmatrix}\boldsymbol{z}_1;\dots\ \boldsymbol{z}_n;\end{bmatrix}$ que es equivalente, pero además es ortogonal; por tanto, aplicando las transformaciones elementales sobre el sistema completo tenemos

$$\begin{bmatrix}\boldsymbol{y}_1;\dots\ \boldsymbol{y}_n;\ \boldsymbol{y}_{(n+1)};\end{bmatrix}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \begin{bmatrix}\boldsymbol{z}_1;\dots\ \boldsymbol{z}_n;\ \boldsymbol{y}_{(n+1)};\end{bmatrix}.$$

Ahora, con una segunda sucesión de transformaciones elementales "de izquierda a derecha", $\boldsymbol{\tau}_{(k+1)},\ \dots,\ \boldsymbol{\tau}_p$, podemos transformar el último vector en $\boldsymbol{z}_{(n+1)} = \boldsymbol{y}_{(n+1)} - a_1\boldsymbol{z}_1 - a_2\boldsymbol{z}_2\cdots - a_n\boldsymbol{z}_n$. Pues bien, hagámoslo de manera que el nuevo vector $\boldsymbol{z}_{(n+1)}$ sea ortogonal a todos los vectores $\boldsymbol{z}_j$ que le anteceden, es decir, de manera que

$$\boldsymbol{z}_{(n+1)}\cdot\boldsymbol{z}_j \;=\; 0, \qquad \text{para } j = 1,\dots,n.$$

Para lograrlo fijémonos en lo siguiente. Para cada $j = 1,\dots,n$ tenemos dos casos posibles:

- Si $\boldsymbol{z}_j$ es cero, entonces $\boldsymbol{z}_j$ y $\boldsymbol{z}_{(n+1)}$ ya son ortogonales (y no hay nada que hacer).
- Si $\boldsymbol{z}_j \neq \mathbf{0}$, entonces

$$0 \;=\; \boldsymbol{z}_{(n+1)}\cdot\boldsymbol{z}_j \;=\; \big(\boldsymbol{y}_{(n+1)} - \widehat{a_1}\boldsymbol{z}_1 - \cdots - \widehat{a_n}\boldsymbol{z}_n\big)\cdot\boldsymbol{z}_j \;=\; \big(\boldsymbol{y}_{(n+1)}\cdot\boldsymbol{z}_j\big) - \widehat{a_j}\,\big(\boldsymbol{z}_j\cdot\boldsymbol{z}_j\big) \;\Leftrightarrow\; \widehat{a_j} = \frac{\boldsymbol{y}_{(n+1)}\cdot\boldsymbol{z}_j}{\boldsymbol{z}_j\cdot\boldsymbol{z}_j}.$$

[1] decimos que dos sistemas son equivalentes si generan el mismo espacio (véase la Definición 9.4 en la página 120)

Por tanto, la segunda sucesión de transformaciones elementales realiza la siguiente transformación.

$$\boldsymbol{z}_{(n+1)} = \boldsymbol{y}_{(n+1)} - \sum_{\boldsymbol{z}_j \neq \mathbf{0}} \widehat{a_j}\, \boldsymbol{z}_j, \qquad \text{donde} \quad \widehat{a_j} \;=\; \frac{\boldsymbol{y}_{(n+1)} \cdot \boldsymbol{z}_j}{\boldsymbol{z}_j \cdot \boldsymbol{z}_j}, \tag{18.1}$$

de manera que tras las dos sucesiones de transformaciones elementales tenemos

$$\big[\boldsymbol{y}_1; \dots\ \boldsymbol{y}_n;\ \boldsymbol{y}_{(n+1)};\big] \xrightarrow{\boldsymbol{\tau}_1, \dots \boldsymbol{\tau}_k} \big[\boldsymbol{z}_1; \dots\ \boldsymbol{z}_n;\ \boldsymbol{y}_{(n+1)};\big] \xrightarrow{\boldsymbol{\tau}_{(k+1)}, \dots \boldsymbol{\tau}_p} \big[\boldsymbol{z}_1; \dots\ \boldsymbol{z}_n;\ \boldsymbol{z}_{(n+1)};\big],$$

que es un sistema *ortogonal* equivalente al de partida. □

Compare el Teorema 18.1.2 en la página anterior con el Teorema 11.3.4 en la página 155. Fíjese además que $\sum_{\boldsymbol{z}_j \neq \mathbf{0}} \widehat{a_j}\, \boldsymbol{z}_j$ en la Ecuación 18.1 es la proyección ortogonal de $\boldsymbol{y}_{(n+1)}$ sobre $\mathcal{L}\big(\big[\boldsymbol{z}_1; \dots\ \boldsymbol{z}_n;\big]\big)$. Fíjese en la Figura 11.4.

**Corolario 18.1.3.** *Cualquier sistema ortogonal de vectores de $\mathbb{R}^n$ no nulos se puede extender hasta formar una base ortogonal de $\mathbb{R}^n$.*

*Demostración.* Sea $\big[\boldsymbol{z}_1; \dots\ \boldsymbol{z}_r;\big]$ un sistema ortogonal de vectores de $\mathbb{R}^n$ y sea $\big[\boldsymbol{y}_1; \dots\ \boldsymbol{y}_m;\big]$ un sistema generador de $\mathbb{R}^n$. Entonces, aplicando Gram-Schmidt sobre el sistema ampliado

$$\big[\boldsymbol{z}_1; \dots\ \boldsymbol{z}_r;\ \boldsymbol{y}_1; \dots\ \boldsymbol{y}_m;\big]$$

comenzando sobre el vector $(r+1)$-ésimo, obtenemos un nuevo sistema equivalente al ampliado y ortogonal[2]

$$\big[\boldsymbol{z}_1; \dots\ \boldsymbol{z}_r;\ \boldsymbol{z}_{r+1}; \dots\ \boldsymbol{z}_{r+m};\big].$$

Por ser un sistema equivalente al anterior, es un sistema generador de $\mathbb{R}^n$. Si de este sistema quitamos los vectores nulos, seguirá siendo generador y ortogonal, pero además será linealmente independiente. □

## 18.2. Matrices ortogonales

**Definición 18.2.** *Un sistema de vectores es ortonormal si es ortogonal y cada vector es de norma uno.*

En la demostración del Teorema espectral que enunciaremos más adelante, usaremos el siguiente resultado.

Ejercicio 89. Demuestre el siguiente corolario:

**Corolario 18.2.1.** *Dado $\boldsymbol{q}$ de $\mathbb{R}^n$ y unitario, existe una base ortonormal de $\mathbb{R}^n$ cuyo último vector es $\boldsymbol{q}$.*

Habitualmente, denotaremos con $\mathbf{Q}$ a las matrices cuyas columnas forman un sistema *ortonormal*.

Ejercicio 90. Demuestre la siguiente proposición:

**Proposición 18.2.2.** *Las columnas de $\underset{m\times n}{\mathbf{Q}}$ son ortonormales si y solo si $(\mathbf{Q}^\intercal)\mathbf{Q} = \underset{n\times n}{\mathbf{I}}$.*

Nótese que cuando $m > n = r$ entonces $\underset{n\times n}{(\mathbf{Q}^\intercal)\mathbf{Q}} = \underset{n\times n}{\mathbf{I}} \neq \underset{m\times m}{\mathbf{Q}(\mathbf{Q}^\intercal)}$. Por ejemplo, si $\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ entonces

$$(\mathbf{Q}^\intercal)\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \underset{2\times 2}{\mathbf{I}}; \quad \text{pero} \quad \mathbf{Q}(\mathbf{Q}^\intercal) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq \underset{3\times 3}{\mathbf{I}}.$$

[2] En realidad se puede empezar desde la posición 1, pues los vectores $\boldsymbol{y}_i$ ya son ortogonales entre sí, y por tanto el método no los modifica.

**Definición 18.3.** *Decimos que* $\mathbf{Q}$ *es* matriz ortogonal *si es cuadrada y sus columnas son ortonormales.*

**Corolario 18.2.3.** $\mathbf{Q}$ *es ortogonal si y solo si* $\mathbf{Q}^\intercal = \mathbf{Q}^{-1}$.

Nótese, por tanto, que las columnas de una matriz ortogonal $\mathbf{Q}$ de orden $n$ forman una base de $\mathbb{R}^n$.

Ejercicio 91. Demuestre las siguientes proposiciones:

(a) **Proposición 18.2.4.** *El producto de matrices ortogonales es ortogonal.*
(b) **Proposición 18.2.5.** *Si* $\mathbf{A}$ *es simétrica y* $\mathbf{Q}$ *es ortogonal, entonces* $(\mathbf{Q}^{-1})\mathbf{A}\mathbf{Q}$ *es simétrica.*

## 18.3. Nota sobre la conjugación de números complejos

Un número complejo se expresa de la forma $a+bi$, donde $a$ y $b$ son números reales, y donde $i$ es la solución de la ecuación $x^2 = -1$ (es decir, $i^2 = -1$).

Puesto que ningún número real satisface dicha ecuación, a $i$ se le denomina *número imaginario*. Para el número complejo $a + bi$, denominamos *parte real* a "$a$" y *parte imaginaria* a "$b$".

- Para un número complejo $a + bi$, su conjugado es: $\overline{a + bi} = a - bi$ (es decir, se cambia el signo de la parte imaginaria).
- Un número complejo es real si y solo si es igual a su conjugado: $\overline{x} = x$.
- El producto del número $a + bi$ por su conjugado es $(a + bi) \cdot (\overline{a + bi}) = a^2 + b^2$.
- El conjugado de una suma es la suma de los conjugados: $\overline{x + y} = \overline{x} + \overline{y}$.
- El conjugado de un producto es el producto de los conjugados: $\overline{x \cdot y} = \overline{x} \cdot \overline{y}$.

## 18.4. Diagonalización de matrices simétricas

**Proposición 18.4.1.** *Los autovalores* $\lambda$ *de una matriz* $\mathbf{A}$ *real y simétrica son reales.*

*Demostración.* Suponga $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ (con $\boldsymbol{x} \neq \mathbf{0}$); y donde tanto $\lambda$ como $\boldsymbol{x}$ son complejos. Entonces, multiplicando $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ por el conjugado $\overline{\boldsymbol{x}}$, tenemos

$$\overline{\boldsymbol{x}}\mathbf{A}\boldsymbol{x} \;=\; \overline{\boldsymbol{x}} \cdot (\lambda\boldsymbol{x}) \;=\; \lambda\big(\overline{\boldsymbol{x}} \cdot \boldsymbol{x}\big), \tag{18.2}$$

donde $\overline{\boldsymbol{x}} \cdot \boldsymbol{x} \neq 0$ por ser $\boldsymbol{x} \neq \mathbf{0}$. Puesto que el conjugado de $\mathbf{A}$ es $\mathbf{A}$, por ser una matriz real; tomando el conjugado de $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, tenemos $\mathbf{A}\overline{\boldsymbol{x}} = \overline{\lambda}\overline{\boldsymbol{x}}$. Como $\mathbf{A}$ es simétrica, para cualquier $\boldsymbol{y}$ tenemos que $\mathbf{A}\boldsymbol{y} = \boldsymbol{y}\mathbf{A}$, en particular $\mathbf{A}\overline{\boldsymbol{x}} \;=\; \overline{\boldsymbol{x}}\mathbf{A} \;=\; \overline{\lambda}\overline{\boldsymbol{x}}$. Y multiplicando por $\boldsymbol{x}$ a ambos lados, tenemos

$$\overline{\boldsymbol{x}}\mathbf{A}\boldsymbol{x} \;=\; \big(\overline{\lambda}\overline{\boldsymbol{x}}\big) \cdot \boldsymbol{x} \;=\; \overline{\lambda}\big(\overline{\boldsymbol{x}} \cdot \boldsymbol{x}\big). \tag{18.3}$$

Como en (18.2) y (18.3) los lados izquierdos son idénticos, necesariamente los lados derechos son iguales; por tanto

$$\overline{\lambda} = \lambda,$$

algo que sólo es posible si la parte imaginaria es cero. Por tanto, los autovalores son reales. □

Además, como los autovectores provienen de resolver la ecuación real $(\mathbf{A} - \lambda\mathbf{I})\boldsymbol{x} = \mathbf{0}$, necesariamente son reales; por tanto, obtenemos el siguiente corolario:

**Corolario 18.4.2.** *Los autovectores* $\boldsymbol{x}$ *de una matriz real y simétrica son reales.*

**Teorema 18.4.3.** *Si* $\underset{n\times n}{\mathbf{A}}$ *es real y simétrica, entonces existe* $\mathbf{Q}$ *ortogonal y tal que* $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ *es diagonal.*

*Demostración.* Lo demostraremos describiendo un algoritmo que construye dicha base en dos etapas.

*Paso inicial.* Sea $\boldsymbol{q}_n$ un autovector *unitario* de $\mathbf{A}$ correspondiente al autovalor $\lambda_n$. Aplicando el Corolario 18.2.1, extendemos el sistema $\left[\boldsymbol{q}_n;\right]$ hasta obtener una base ortonormal de $\mathbb{R}^n$

$$\mathbf{Q}_n = \left[\boldsymbol{q}_1; \dots\ \boldsymbol{q}_n;\right].$$

Por tanto, $\mathbf{Q}_n$ es ortogonal, y $\mathbf{Q}_n^{-1}\mathbf{A}\mathbf{Q}_n$ es simétrica y de la forma $\mathbf{Q}_n^{-1}\mathbf{A}\mathbf{Q}_n = \begin{bmatrix} & & 0 \\ & \mathbf{A'}_{(n-1)\times(n-1)} & 0 \\ & & \vdots \\ 0 & 0 \ \cdots & \lambda_n \end{bmatrix}$, ya que $\left(\mathbf{Q}_n^{-1}\mathbf{A}\mathbf{Q}_n\right)_{|n} = \left(\mathbf{Q}_n^{\intercal}\mathbf{A}\mathbf{Q}_n\right)_{|n} = \mathbf{Q}_n^{\intercal}\mathbf{A}\boldsymbol{q}_n = \lambda_n \cdot \mathbf{Q}_n^{\intercal}\boldsymbol{q}_n = \lambda_n \begin{pmatrix} \boldsymbol{q}_1 \cdot \boldsymbol{q}_n \\ \boldsymbol{q}_2 \cdot \boldsymbol{q}_n \\ \vdots \\ \boldsymbol{q}_n \cdot \boldsymbol{q}_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_n \cdot \mathbf{I}_{|n}$.

*Paso de continuación.* Supongamos que tras $n-k$ pasos tenemos

$$\mathbf{Q}_{k+1}^{-1} \cdots \mathbf{Q}_n^{-1} \cdot \mathbf{A} \cdot \mathbf{Q}_n \cdots \mathbf{Q}_{k+1} = \left[\begin{array}{c|ccc} \mathbf{A'}_{k\times k} & & & \\ \hline & \lambda_{k+1} & & \\ & & \ddots & \\ & & & \lambda_n \end{array}\right], \tag{18.4}$$

como $\mathbf{A'}$ es simétrica, tomamos un autovector $\boldsymbol{x}_k \in \mathbb{R}^k$ de $\mathbf{A'}$ de norma 1 y correspondiente al autovalor $\lambda_k$. Aplicando el Corolario 18.2.1 formamos una base ortonormal $\mathbf{R} = \left[\boldsymbol{x}_1; \dots\ \boldsymbol{x}_k;\right]$ de $\mathbb{R}^k$. Ahora, si consideramos la matriz por bloques

$$\mathbf{Q}_k = \left[\begin{array}{c|c} \mathbf{R} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I}_{(n-k)\times(n-k)} \end{array}\right] = \left[\boldsymbol{y}_1; \dots\ \boldsymbol{y}_n;\right],$$

entonces $\mathbf{Q}_k$ es ortogonal y sus últimas columnas $\boldsymbol{y}_k, \dots, \boldsymbol{y}_n$ son autovectores de la matriz de la Ecuación (18.4); es decir, llamando $\mathbf{B}$ a la matriz de dicha ecuación tenemos que

$$\mathbf{B}\boldsymbol{y}_j = \lambda_j \boldsymbol{y}_j; \quad \text{para} \quad j = k, \dots, n.$$

Consecuentemente, multiplicando (18.4) por $\mathbf{Q}_k^{-1}$ y por $\mathbf{Q}_k$ tenemos

$$\mathbf{Q}_k^{-1} \cdot \left(\mathbf{Q}_{k+1}^{-1} \cdots \mathbf{Q}_n^{-1} \cdot \mathbf{A} \cdot \mathbf{Q}_n \cdots \mathbf{Q}_{k+1}\right) \cdot \mathbf{Q}_k = \mathbf{Q}_k^{-1}\mathbf{B}\mathbf{Q}_k = \left[\begin{array}{c|ccc} \mathbf{A''}_{k-1\times k-1} & & & \\ \hline & \lambda_k & & \\ & & \ddots & \\ & & & \lambda_n \end{array}\right],$$

ya que, $\left(\mathbf{Q}_k^{-1}\mathbf{B}\mathbf{Q}_k\right)_{|j} = \mathbf{Q}_k^{\intercal}\left(\mathbf{B}\mathbf{Q}_k\right)_{|j} = \mathbf{Q}_k^{\intercal}\left(\mathbf{B}\boldsymbol{y}_j\right) = \mathbf{Q}_k^{\intercal}\left(\lambda_j \boldsymbol{y}_j\right) = \lambda_j \begin{pmatrix} \boldsymbol{y}_1 \cdot \boldsymbol{y}_j \\ \boldsymbol{y}_2 \cdot \boldsymbol{y}_j \\ \vdots \\ \boldsymbol{y}_n \cdot \boldsymbol{y}_j \end{pmatrix} = \lambda_j \left(\mathbf{I}_{|j}\right)$.

Así, el producto de matrices $\mathbf{Q} = \mathbf{Q}_n \cdots \mathbf{Q}_1$ es ortogonal y $\mathbf{D} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ es diagonal con los autovalores $\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_n$ de $\mathbf{A}$ correspondientes a las columnas de $\mathbf{Q}$ en su diagonal principal. □

**Definición 18.4.** $\underset{n\times n}{\mathbf{A}}$ *es diagonalizable ortogonalmente si existe* $\mathbf{Q}$ *ortogonal tal que* $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ *es diagonal.*

El anterior teorema tiene un importante corolario que se conoce como Teorema Espectral:

**Corolario 18.4.4** (Teorema espectral)**.** *Si* $\mathbf{A}$ *(de orden n) es real y simétrica, entonces existe una base ortonormal de* $\mathbb{R}^n$ *formada por autovectores de* $\mathbf{A}$*.*

Para finalizar, demostramos un último resultado acerca de la diagonalización de matrices simétricas:

**Proposición 18.4.5.** *Los autovectores correspondientes a autovalores distintos, de una matriz simétrica, son ortogonales entre sí.*

*Demostración.* Considere $\boldsymbol{x}$ e $\boldsymbol{y}$, autovectores correspondientes a dos autovalores distintos: $\mathbf{A}\boldsymbol{x} = \lambda_1\boldsymbol{x}$ y $\mathbf{A}\boldsymbol{y} = \lambda_2\boldsymbol{y}$. Entonces

$$\lambda_1\boldsymbol{x}\cdot\boldsymbol{y} \;=\; \mathbf{A}\boldsymbol{x}\cdot\boldsymbol{y} \;=\; \boldsymbol{x}(\mathbf{A}^\intercal)\boldsymbol{y} \;=\; \boldsymbol{x}\mathbf{A}\boldsymbol{y} \;=\; \lambda_2(\boldsymbol{x}\cdot\boldsymbol{y}).$$

Como $\lambda_1 \neq \lambda_2$; necesariamente:

$$\lambda_1(\boldsymbol{x}\cdot\boldsymbol{y}) - \lambda_2(\boldsymbol{x}\cdot\boldsymbol{y}) = 0 \quad\Longrightarrow\quad (\lambda_1-\lambda_2)\boldsymbol{x}\cdot\boldsymbol{y} = 0 \quad\Longrightarrow\quad \boldsymbol{x}\cdot\boldsymbol{y} = 0.$$

□

Ejercicio 92. Demuestre la siguiente

**Proposición 18.4.6.** *Si una matriz es diagonalizable ortogonalmente, entonces es simétrica.*

Para finalizar, la demostración del Teorema 18.4.3 nos describe un algoritmo para diagonalizar $\mathbf{A}$ ortogonalmente, obteniendo una matriz ortogonal $\mathbf{Q}$ cuyas columnas son autovectores de $\mathbf{A}$:

Módulo NAcAL para Python

```
A = Matrix([[1,1,1],[1,0,0],[1,0,0]]); # Matriz simétrica
L = [-1,0,2];                          # Espectro de A
A
```

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Módulo NAcAL para Python

```
D=DiagonalizaO(A,L); # Diagonaliza Ortogonalmente
D
```

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

El atributo `Q` de la matriz diagonalizada es la matriz ortogonal de autovectores:

Módulo NAcAL para Python

```
Q=D.Q          # Matriz de autovectores
```

$$\begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \end{bmatrix}$$

Comprobemos que efectivamente $\mathbf{Q}$ es ortogonal:

Módulo NAcAL para Python

```
~Q*Q          # Matriz de autovectores es ortogonal
```

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

y que diagonaliza a $\mathbf{A}$:

Módulo NAcAL para Python

```
~Q*A*Q        # Matriz de autovectores diagonaliza A
```

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

No obstante, como toda matriz simétrica es diagonalizable, también podemos obtener una matriz $\mathbf{Q}$ aplicando el algoritmo descrito en la demostración del Teorema 17.2.1 de diagonalización por bloques triangulares: primero aplicamos la diagonalización por semejanza

$$\begin{bmatrix} \mathbf{A} \\ \hline \mathbf{I} \end{bmatrix} \xrightarrow[\operatorname{esp}(\boldsymbol{\tau}_1^{-1} \ldots \boldsymbol{\tau}_p^{-1})]{\boldsymbol{\tau}_1 \ldots \boldsymbol{\tau}_p} \begin{bmatrix} \mathbf{C} \\ \hline \mathbf{S} \end{bmatrix} \qquad \text{donde} \quad \mathbf{S} = \mathbf{I}_{\boldsymbol{\tau}_1 \ldots \boldsymbol{\tau}_p},$$

Módulo NAcAL para Python

```
D = DiagonalizaS(A, L) # D matriz diagonalizada por semejanza
S = D.S                # S matriz de autovectores
S
```

$$\begin{bmatrix} \frac{2}{3} & 0 & -2 \\ \frac{1}{3} & 1 & 2 \\ \frac{1}{3} & -1 & 2 \end{bmatrix}$$

A continuación, aplicamos Gram-Schmidt sobre las columnas de $\mathbf{S}$ para asegurarnos de que las columnas de $\mathbf{S}$ que corresponden a autovalores repetidos sean perpendiculares, y normalizamos todas las columnas para obtener una matriz $\mathbf{Q}$.

Módulo NAcAL para Python

```
Q = S.GS().normalizada()  # Q matriz ortogonal de autovectores
Q
```

$$\begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} \end{bmatrix}$$

# Diagonalización por congruencia de matrices simétricas. Formas cuadráticas

## 19.1. Formas cuadráticas y matrices definidas positivas

Un polinomio es una expresión que contiene variables y coeficientes, y en la que únicamente están involucradas las operaciones de suma, resta de *términos* que contienen números, productos y exponentes no negativos de las variables. Un ejemplo de polinomio en $x$ con tres términos es: $x^2 - 4x + 7$. Y un ejemplo de polinomio en tres variables y cuatro términos es: $x^3 + 2xyz^2 - yz + 1$.

El grado de cada término es la suma de los exponentes de las variables que aparecen en él y, por tanto, nunca es negativo. El grado del polinomio es el mayor de los grados de sus términos con coeficientes no nulos. Así, para los ejemplos anteriores, el grado del polinomio $x^2 - 4x^1 + 7x^0$ es dos; y el grado del polinomio $x^3 + 2x^1y^1z^2 - y^1z^1 + 1$ es cuatro: $4 = 1 + 1 + 2$ por el término $2x^1y^1z^2$.

**Definición 19.1.** *Una forma cuadrática es un polinomio en el que todos sus términos son de grado dos.*

Por ejemplo, $4x^2 + 2xy - 3y^2$ es una forma cuadrática en las variables $x$ e $y$.

### 19.1.1. Formas cuadráticas reales

Toda matriz $\mathbf{A}$ simétrica de orden $n$ define una forma cuadrática $q_{\mathbf{A}}$ en $n$ variables mediante la fórmula

$$q_{\mathbf{A}}(\boldsymbol{x}) = \boldsymbol{x}\mathbf{A}\boldsymbol{x},$$

donde $\boldsymbol{x} = (x, y, \ldots)$ es el vector que contiene las variables de la forma cuadrática.

Recíprocamente, dada una forma cuadrática en $n$ variables, sus coeficientes se pueden organizar en una matriz simétrica de orden $n$. Volviendo al ejemplo de más arriba

$$4x^2 + 2xy - 3y^2 = \begin{pmatrix} x, & y, \end{pmatrix} \begin{bmatrix} 4 & 1 \\ 1 & -3 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

y con la matriz $\begin{bmatrix} 1 & -2 & \\ -2 & 3 & \\ & & 5 \end{bmatrix}$ formamos la forma cuadrática

$$\begin{pmatrix} x, & y, & z, \end{pmatrix} \begin{bmatrix} 1 & -2 & \\ -2 & 3 & \\ & & 5 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 - 4xy + 3y^2 + 5z^2.$$

### 19.1.2. Matrices definidas positivas

**Definición 19.2.** *Una matriz simétrica* $\mathbf{A} \in \mathbb{R}^{n\times n}$ *es* definida positiva *si para todo* $\boldsymbol{x} \neq \mathbf{0}$ *de* $\mathbb{R}^n$, $\boldsymbol{x}\mathbf{A}\boldsymbol{x} > 0$.

*Entonces también se dice que la forma cuadrática* $\boldsymbol{x}\mathbf{A}\boldsymbol{x}$ *es* definida positiva.

Como consecuencia de la definición tenemos que

**Proposición 19.1.1.** *Si* $\underset{m\times n}{\mathbf{A}}$ *tiene rango* $n$*, entonces* $\mathbf{A}^\intercal\mathbf{A}$ *es definida positiva.*

*Demostración.* Por una parte,

$$\boldsymbol{x}\mathbf{A}^\intercal\mathbf{A}\boldsymbol{x} = \big(\mathbf{A}\boldsymbol{x}\big)\cdot\big(\mathbf{A}\boldsymbol{x}\big) = \|\mathbf{A}\boldsymbol{x}\|^2 \geq 0.$$

Como $\mathbf{A}$ es de rango completo por columnas, si $\boldsymbol{x}\neq\mathbf{0}$, entonces $\mathbf{A}\boldsymbol{x}\neq\mathbf{0}$; y por tanto $\boldsymbol{x}\mathbf{A}\boldsymbol{x} = \|\mathbf{A}\boldsymbol{x}\|^2 > 0$. □

Ejercicio 93. Demuestre las siguientes proposiciones:

(a) Si $\mathbf{A}$ es definida positiva, los elementos de la diagonal son positivos.
(b) Si $\mathbf{A}$ y $\mathbf{B}$ son definidas positivas, entonces la suma $(\mathbf{A}+\mathbf{B})$ también es definida positiva.

La primera de las anteriores proposiciones se puede generalizar a todas las submatrices principales de $\mathbf{A}$ (véase el Corolario 19.1.5 en el Ejercicio 95 más abajo).

Hay una estrecha relación entre los signos de los autovalores de una matriz $\mathbf{A}$ y el signo de la forma cuadrática asociada: $q_{\mathbf{A}}(\boldsymbol{x}) = \boldsymbol{x}\mathbf{A}\boldsymbol{x}$.

**Proposición 19.1.2.** *Una matriz real y simétrica de orden* $n$ *es definida positiva si y sólo si son positivos todos sus autovalores.*

*Demostración.* La demostración se basa en que toda matriz real y simétrica es diagonalizable ortogonalmente. Puesto que $\mathbf{A}$ es real y simétrica, tenemos que $\mathbf{A} = \mathbf{Q}\mathbf{D}(\mathbf{Q}^{-1})$ con $\mathbf{Q}^{-1} = \mathbf{Q}^\intercal$; y por tanto

$$\boldsymbol{x}\mathbf{A}\boldsymbol{x} \;=\; \boldsymbol{x}\mathbf{Q}\mathbf{D}(\mathbf{Q}^\intercal)\boldsymbol{x} \;=\; \big(\mathbf{Q}^\intercal\boldsymbol{x}\big)\mathbf{D}\big(\mathbf{Q}^\intercal\boldsymbol{x}\big) \;=\; \boldsymbol{y}\mathbf{D}\boldsymbol{y}, \qquad (\text{donde}\quad \boldsymbol{y} = \mathbf{Q}^\intercal\boldsymbol{x}).$$

De esta expresión es evidente que $\boldsymbol{x}\mathbf{A}\boldsymbol{x}$ es una suma ponderada de cuadrados:

$$\boldsymbol{x}\mathbf{A}\boldsymbol{x} = \boldsymbol{y}\mathbf{D}\boldsymbol{y} = \begin{pmatrix} y_1, & \cdots, & y_n, \end{pmatrix}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n\end{pmatrix} = \sum_{j=1}^{n}\lambda_j\big(y_j^2\big),$$

donde las ponderaciones $\lambda_j$ son los autovalores de $\mathbf{A}$. Puesto que $\mathbf{Q}$ es invertible, $\boldsymbol{y} = \mathbf{Q}^\intercal\boldsymbol{x}$ es distinto de cero siempre que $\boldsymbol{x}\neq\mathbf{0}$, la suma $\sum_{j=1}^{n}\lambda_j(\boldsymbol{y}_{|j})^2$ es positiva cuando los autovalores son positivos.

Por otra parte, la forma cuadrática no puede ser positiva si algún autovalor $\lambda_j$ es negativo o cero; para verlo, basta elegir $\boldsymbol{y} = \mathbf{I}_{|j}$ pues $\boldsymbol{y}\mathbf{D}\boldsymbol{y} = {}_{j|}\mathbf{I}\mathbf{D}\mathbf{I}_{|j} = {}_{j|}\mathbf{D}_{|j} = \lambda_j$. □

Así, puesto que el determinante de $\mathbf{A}$ es igual al producto de sus autovalores, llegamos al siguiente

**Corolario 19.1.3.** *Si una matriz es definida positiva, su determinante es positivo.*

Ejercicio 94. Demuestre que si $\mathbf{A}$ es simétrica y definida positiva, entonces $\mathbf{A}^{-1}$ también es definida positiva.

Para finalizar esta sección:

Ejercicio 95. Revise la Definición 3.5 en la página 40 y demuestre las siguientes proposiciones:

(a) **Proposición 19.1.4.** *Si* $\underset{n\times n}{\mathbf{A}}$ *es definida positiva y* $\underset{n\times k}{\mathbf{X}}$ *de rango* $k$*, entonces* $(\mathbf{X}^\intercal)\mathbf{A}\mathbf{X}$ *es definida positiva.*
(b) **Corolario 19.1.5.** *Si* $\mathbf{A}$ *es definida positiva, todas sus submatrices principales son definidas positivas.*
(c) **Proposición 19.1.6.** *Si* $\mathbf{A}$ *es definida positiva, entonces la mayor componente de* $\mathbf{A}$ *está en la diagonal.*
*Pista.* Demuestre antes que $a_{ij}^2 < a_{ii}a_{jj}$ usando el determinante de ${}_{(i,j,)|}\mathbf{A}_{|(i,j,)}$.

### 19.1.3. Algunos tipos de formas cuadráticas

Decimos que una forma cuadrática (y su correspondiente matriz de orden $n$) es

- Definida positiva si para cualquier $\boldsymbol{x} \in \mathbb{R}^n$ no nulo se verifica que $\boldsymbol{x}\mathbf{A}\boldsymbol{x} > 0$.
- Semi-definida positiva si para cualquier $\boldsymbol{x} \in \mathbb{R}^n$ no nulo se verifica que $\boldsymbol{x}\mathbf{A}\boldsymbol{x} \geq 0$.
- Definida negativa si para cualquier $\boldsymbol{x} \in \mathbb{R}^n$ no nulo se verifica que $\boldsymbol{x}\mathbf{A}\boldsymbol{x} < 0$.
- Semi-definida negativa si para cualquier $\boldsymbol{x} \in \mathbb{R}^n$ no nulo se verifica que $\boldsymbol{x}\mathbf{A}\boldsymbol{x} \leq 0$.
- Indefinida si no es ni semi-definida positiva ni semi-definida negativa.

Nótese que si $\mathbf{A}$ es (*semi-*) definida positiva, entonces $(-\mathbf{A})$ es (*semi-*) definida negativa.

EJERCICIO 96. Demuestre que si $\mathbf{A}$ es semi-definida positiva, los elementos de la diagonal son no negativos.

## 19.2. Diagonalización de matrices simétricas por congruencia

☞ Vamos a ver que podemos diagonalizar cualquier matriz por congruencia (mediante eliminación Gaussiana de filas y columnas). En general, la matriz obtenida con este procedimiento no contiene los autovalores de la matriz original en su diagonal, pero permite expresar cualquier forma cuadrática como sumas y/o restas de términos al cuadrado. Así podremos comprobar si una matriz es definida positiva sin necesidad de calcular sus autovalores (bastará con comprobar que la expresión solo contiene sumas de cuadrados).

Este resultado tiene importancia práctica. No es posible encontrar las raíces de un polinomio cualquiera (solamente está asegurado encontrar las raíces de polinomios de grado menor o igual a 4). Afortunadamente, la diagonalización por congruencia nos revelará *los signos* de los autovalores de *cualquier matriz real y simétrica*. Este resultado se llama *"Ley de inercia"*.

**Definición 19.3.** *Dos matrices* $\mathbf{A}$ *y* $\mathbf{C}$ *son congruentes si existe una matriz* $\mathbf{B}$ *invertible tal que*

$$\mathbf{C} = (\mathbf{B}^\intercal)\mathbf{A}\mathbf{B}.$$

En esta sección veremos que siempre es posible encontrar una matriz diagonal que sea congruente con una matriz simétrica $\mathbf{A}$. La demostración describirá los pasos a seguir para diagonalizar por congruencia (paso de inicio y paso de continuación). Pero antes de exponer la demostración, veamos tres ejemplos de diagonalización por congruencia. El método consiste en tratar de escalonar una matriz, pero aplicando a las filas todas las operaciones que hayamos aplicado a las columnas.

*Ejemplo* 46. Vamos a diagonalizar la matriz $\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ mediante eliminación, pero cada operación sobre las columnas la repetiremos también sobre las filas:[1]

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \overset{\substack{\boldsymbol{\tau} \\ [(2)\mathbf{2}] \\ [(1)\mathbf{1}+\mathbf{2}]}}{\longrightarrow} \begin{bmatrix} 2 & 0 & 0 \\ -1 & 3 & -1 \\ 0 & -2 & 2 \end{bmatrix} \underset{\substack{\boldsymbol{\tau} \\ [(1)\mathbf{1}+\mathbf{2}] \\ [(2)\mathbf{2}]}}{\longrightarrow} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & -2 \\ 0 & -2 & 2 \end{bmatrix} \overset{\substack{\boldsymbol{\tau} \\ [(3)\mathbf{3}] \\ [(1)\mathbf{2}+\mathbf{3}]}}{\longrightarrow} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & -2 & 4 \end{bmatrix} \underset{\substack{\boldsymbol{\tau} \\ [(1)\mathbf{2}+\mathbf{3}] \\ [(3)\mathbf{3}]}}{\longrightarrow} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{bmatrix}.$$

Aunque la matriz obtenida es diagonal, ni 6 ni 12 son autovalores de $\mathbf{A}$ (*este método NO encuentra los autovalores*)... Pero como los componentes de la diagonal son positivos, $\mathbf{A}$ es definida positiva.

[1] Recuerde que este esquema indica que el orden en el que se aplican las transformaciones *de las filas* es de abajo hacia arriba, es decir, primero las más "alejadas" de las flechas y luego las más "cercanas" a las flechas (véase el Apéndice 4.A en la página 62.

Y cuando en la diagonal tenemos un cero, ¿qué podemos hacer?... Podemos sumar a la primera columna otra columna (en un ejemplo posterior veremos que hay que hacerlo con cuidado, pues no siempre funciona).

*Ejemplo* 47. Considere $\begin{bmatrix}0&1&1\\1&0&1\\1&1&0\end{bmatrix}$. La primera componente de la primera fila es un cero, así que no tenemos un pivote con el que anular todo lo que queda a la derecha de la diagonal. Como la segunda columna sí tiene un pivote en la primera fila, podemos sumar la segunda columna a la primera columna y lograr tener un pivote donde nos interesa (pero *recuerde que cada operación sobre las columnas se repite sobre las filas*):

$$\begin{bmatrix}0&1&1\\1&0&1\\1&1&0\end{bmatrix} \xrightarrow[[(1)\mathbf{2}+\mathbf{1}]]{\boldsymbol{\tau}} \begin{bmatrix}1&1&1\\1&0&1\\2&1&0\end{bmatrix} \xrightarrow[\substack{\boldsymbol{\tau}\\[(1)\mathbf{2}+\mathbf{1}]}]{} \begin{bmatrix}2&1&2\\1&0&1\\2&1&0\end{bmatrix} \xrightarrow[\substack{[(-1)\mathbf{1}+\mathbf{2}]\\[(-1)\mathbf{1}+\mathbf{3}]}]{\overset{\boldsymbol{\tau}}{[(2)\mathbf{2}]}} \begin{bmatrix}2&0&0\\1&-1&0\\2&0&-2\end{bmatrix} \xrightarrow[\substack{\boldsymbol{\tau}\\[(-1)\mathbf{1}+\mathbf{3}]\\[(-1)\mathbf{1}+\mathbf{2}]\\[(2)\mathbf{2}]}]{} \begin{bmatrix}2&0&0\\0&-2&0\\0&0&-2\end{bmatrix}.$$

Un tercer ejemplo nos indica que debemos ser cautelosos cuando encontramos un cero en la diagonal.

*Ejemplo* 48. Considere la matriz $\left[\begin{smallmatrix}0&1\\1&a\end{smallmatrix}\right]$, que tiene un cero en donde querríamos tener un pivote con el que eliminar todo lo que queda a su derecha (y por debajo). ¿Qué podemos hacer?

**Primer intento:** podemos intentar la estrategia del ejemplo anterior, sumando la segunda columna a la primera (y repetir la operación con las filas):

$$\begin{bmatrix}0&1\\1&a\end{bmatrix} \xrightarrow[[(1)\mathbf{2}+\mathbf{1}]]{\boldsymbol{\tau}} \begin{bmatrix}1&1\\1+a&a\end{bmatrix} \xrightarrow[\substack{\boldsymbol{\tau}\\[(1)\mathbf{2}+\mathbf{1}]}]{} \begin{bmatrix}2+a&1+a\\1+a&a\end{bmatrix},$$

y ahora con el pivote $(2+a)$ podemos anular todo lo que está a su derecha (y por debajo), pero... ¿estamos seguros de que $(2+a)$ es un pivote? Si $a=-2$ entonces estaríamos como al principio. En el anterior ejemplo la estrategia funcionó porque sumamos una columna que tenía un cero en la diagonal... pero en este ejemplo la columna que sumamos tiene el número $a$ en la diagonal, y esto podría hacer fallar el algoritmo.

**Segundo intento:** para que la estrategia anterior no falle debemos asegurarnos de que la columna que sumamos tiene un cero en la diagonal. Como en este caso no es así, la solución pasa por intercambiar dos columnas (y filas) para "mover" alguna componente no nula de la diagonal (que esté situada a la derecha y por debajo del cero que queremos sustituir) y colocarla en la posición del cero que ocasiona el problema, y luego continuar como siempre:

$$\begin{bmatrix}0&1\\1&-2\end{bmatrix} \xrightarrow[[\mathbf{1}\rightleftharpoons\mathbf{2}]]{\boldsymbol{\tau}} \begin{bmatrix}1&0\\-2&1\end{bmatrix} \xrightarrow[\substack{\boldsymbol{\tau}\\[\mathbf{1}\rightleftharpoons\mathbf{2}]}]{} \begin{bmatrix}-2&1\\1&0\end{bmatrix} \xrightarrow[[(1)\mathbf{1}+\mathbf{2}]]{\overset{\boldsymbol{\tau}}{[(2)\mathbf{2}]}} \begin{bmatrix}-2&0\\1&1\end{bmatrix} \xrightarrow[\substack{\boldsymbol{\tau}\\[(1)\mathbf{1}+\mathbf{2}]\\[(2)\mathbf{2}]}]{} \begin{bmatrix}-2&0\\0&2\end{bmatrix}.$$

Vayamos con la demostración de los pasos que nos permiten diagonalizar por congruencia:

**Proposición 19.2.1** (Paso de inicio)**.** *Dada una matriz* $\mathbf{A}$*, simétrica y de orden* $n$*, existe una matriz no singular* $\mathbf{B}$ *tal que* $\mathbf{B}^\intercal\mathbf{A}\mathbf{B}$ *es una matriz simétrica de la forma*

$$\left[\begin{array}{c|c} * & \\ \hline & \mathbf{A'}_{(n-1)\times(n-1)} \end{array}\right].$$

*Demostración.* Tenemos tres casos:

Caso trivial. Si $\mathbf{A}$ ya tiene la forma de más arriba, es decir, si a la derecha y por debajo de $a_{11}$ son todo ceros, entonces basta que $\mathbf{B}$ sea la matriz identidad.

En caso contrario, hay dos posibilidades: que $a_{11}$ sea cero o que no lo sea.

Caso 1 ($a_{11} \neq 0$). Cuando $a_{11}$ es un pivote, se pueden anular los componentes situados a su derecha por eliminación de izquierda a derecha, y aplicando las mismas operaciones sobre las filas (de arriba hacia abajo) llegamos a la siguiente matriz congruente con $\mathbf{A}$:

$$\boldsymbol{\tau}_k \cdots \boldsymbol{\tau}_1 \mathbf{A} \boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k = \mathbf{B}^\intercal \mathbf{A} \mathbf{B} = \left[\begin{array}{c|c} a_{11} & \\ \hline & \mathbf{A'}_{(n-1)\times(n-1)} \end{array}\right]; \quad \text{donde} \quad \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} = \mathbf{B}.$$

Caso 2 ($a_{11} = 0$). En este caso hay dos posibilidades: que algún otro elemento de la diagonal principal sea distinto de cero o que todos los elementos de la diagonal sean cero.

- Si el elemento $a_{jj}$ (con $j > 1$) es distinto de cero, intercambiamos la primera columna con la $j$-ésima, de manera que ahora el elemento no nulo se encuentra en la posición $(j, 1)$ y a continuación se intercambia la fila $j$-ésima por la primera, con lo que el componente *no nulo* termina por situarse en la posición $(1, 1)$. Así hemos llegado a una matriz del Caso 1.
- Si todos los elementos de la diagonal principal son nulos —y puesto que algún elemento $a_{1j}$ es distinto de cero (pues no estamos en el Caso Trivial)— sumamos $\mathbf{A}_{\lfloor j}$ a la primera columna, y por tanto, también sumamos ${}_{j\rfloor}\mathbf{A}$ a la primera fila. Entonces, de nuevo habremos llegado al Caso 1 ya que

$$\underset{[(1)\boldsymbol{\tau}_{j+1}]}{\mathbf{A}}\underset{[(1)\boldsymbol{\tau}_{j+1}]}{} = \left[\begin{array}{c|c} 2a_{1j} & * \ \ \ldots \ \ * \\ \hline * & \\ \vdots & \mathbf{C}_{(n-1)\times(n-1)} \\ * & \end{array}\right]; \quad \text{donde} \quad 2a_{1j} \neq 0.$$

□

**Proposición 19.2.2** (Paso de continuación)**.** *Dada una matriz simétrica de orden $n$ de la forma*

$$\left[\begin{array}{ccc|c} * & & & \\ & \ddots & & \\ & & * & \\ \hline & & & \mathbf{A'}_{k\times k} \end{array}\right],$$

*existe una matriz no singular $\mathbf{B}$ tal que $\mathbf{B}^\intercal\mathbf{A}\mathbf{B}$ es una matriz simétrica de la forma*

$$\left[\begin{array}{ccc|c} * & & & \\ & \ddots & & \\ & & * & \\ \hline & & & \mathbf{A''}_{(k-1)\times(k-1)} \end{array}\right].$$

*Demostración.* Basta hacer como en el *Paso de Inicio*, pero trabajando sobre las filas y columnas de $\mathbf{A'}$. □

Combinando las dos proposiciones anteriores, llegamos al siguiente

**Corolario 19.2.3.** *Para toda $\mathbf{A}$ simétrica de orden $n$, existe $\mathbf{B}$ invertible tal que $\mathbf{B}^\intercal\mathbf{A}\mathbf{B}$ es diagonal.*

*Ejemplo* 49. Diagonalice por congruencia la matriz

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Piense en qué pasos debe sumar una columna (y una fila) para generar un pivote en la diagonal, y cuándo debe intercambiar columnas (y filas).

Módulo NAcAL para Python

```
A = Matrix([[0,0,1],[0,0,0],[1,0,0]])
D = DiagonalizaC(A,1)
B = D.B
~B*A*B                    # D es congruente con A
```

☞

**Nota 5.** *Fíjese que el Teorema Espectral muestra una forma muy especial de diagonalización, pues*

$$\mathbf{D} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{Q}^{\intercal}\mathbf{A}\mathbf{Q},$$

*es simultáneamente una diagonalización por semejanza y por congruencia.*

### 19.2.1. Ley de inercia

**Teorema 19.2.4** (Ley de inercia)**.** *Sea* $\mathbf{A}$ *real y simétrica; y sean* $\mathbf{B}$ *y* $\mathbf{C}$ *matrices invertibles tales que*

$$\mathbf{D} = (\mathbf{B}^{\intercal})\mathbf{A}\mathbf{B} \quad y \quad \mathbf{D}^{*} = (\mathbf{C}^{\intercal})\mathbf{A}\mathbf{C}$$

*son diagonales. Entonces la diagonal de* $\mathbf{D}$ *tiene tantas componentes positivas, tantas componentes nulas y tantas componentes negativas como la diagonal de* $\mathbf{D}^{*}$*.*

*Demostración.* Comencemos con el número de componentes nulas. Puesto que $\mathbf{B}$ y $\mathbf{C}$ son de rango completo, el rango de $\mathbf{D}$ y el rango de $\mathbf{D}^{*}$ deben ser iguales al rango de $\mathbf{A}$. Así, puesto que $\mathbf{D}$ y $\mathbf{D}^{*}$ son diagonales y tienen el mismo rango, necesariamente tienen el mismo número de ceros. Pasemos a las componentes positivas:

Sean $p$ y $p^{*}$ el número de componentes *positivas* en la diagonal de $\mathbf{D}$ y de $\mathbf{D}^{*}$ respectivamente. Por tanto $q = n - p$ y $q^{*} = n - p^{*}$ son el número de componentes menores o iguales a cero en las respectivas diagonales. Veamos que $p = p^{*}$.

Definamos el vector $\boldsymbol{p} = (i \mid d_{ii} > 0)$ con la lista de índices de las $p$ componentes positivas en $\mathbf{D}$; y el vector $\boldsymbol{p}^{*} = (i \mid d^{*}_{ii} > 0)$ con la correspondiente lista de las $p^{*}$ componentes positivas en $\mathbf{D}^{*}$. Entonces

$$\boldsymbol{x}\left({}_{\boldsymbol{p}|}\mathbf{D}_{|\boldsymbol{p}}\right)\boldsymbol{x} = \sum d_{ii}\left(\boldsymbol{x}_{|i}\right)^{2} > 0; \qquad \text{para todo } \boldsymbol{x} \in \mathbb{R}^{p} \text{ tal que } \boldsymbol{x} \neq \mathbf{0}.$$

De manera similar, definamos $\boldsymbol{q} = (i \mid d_{ii} \leq 0)$ con la lista de índices de las $q$ componentes negativas o nulas en la diagonal de $\mathbf{D}$; y $\boldsymbol{q}^{*} = (i \mid d^{*}_{ii} \leq 0)$ con la correspondiente lista para las $q^{*}$ componentes negativas o nulas en la diagonal de $\mathbf{D}^{*}$. Entonces

$$\boldsymbol{y}\left({}_{\boldsymbol{q}^{*}|}\mathbf{D}^{*}{}_{|\boldsymbol{q}^{*}}\right)\boldsymbol{y} = \sum d^{*}_{ii}\left(\boldsymbol{y}_{|i}\right)^{2} \leq 0; \qquad \text{para todo } \boldsymbol{y} \in \mathbb{R}^{q^{*}} \text{ tal que } \boldsymbol{y} \neq \mathbf{0}.$$

Ahora veamos que necesariamente

$$\mathcal{C}\left(\mathbf{B}_{|\boldsymbol{p}}\right) \cap \mathcal{C}\left(\mathbf{C}_{|\boldsymbol{q}^{*}}\right) = \{\mathbf{0}\}.$$

Para probarlo, supondremos que $\boldsymbol{v} \neq \mathbf{0}$ pertenece a dicha intersección de subespacios y llegaremos a la contradicción de que $\boldsymbol{v}\mathbf{A}\boldsymbol{v}$ es simultáneamente mayor que cero y también menor o igual a cero. Por pertenecer a $\mathcal{C}\left(\mathbf{B}_{|\boldsymbol{p}}\right)$ y ser $\mathbf{B}$ invertible, el vector $\boldsymbol{v}$ es de la forma $\boldsymbol{v} = \left(\mathbf{B}_{|\boldsymbol{p}}\right)\boldsymbol{x}$ para algún $\boldsymbol{x} \in \mathbb{R}^p$ no nulo. Así pues,

$$\boldsymbol{v}\mathbf{A}\boldsymbol{v} \;=\; \boldsymbol{x}\big({}_{\boldsymbol{p}|}(\mathbf{B}^{\intercal})\big)\mathbf{A}\left(\mathbf{B}_{|\boldsymbol{p}}\right)\boldsymbol{x} \;=\; \boldsymbol{x}\big({}_{\boldsymbol{p}|}(\mathbf{B}^{\intercal})\mathbf{A}\mathbf{B}_{|\boldsymbol{p}}\big)\boldsymbol{x} \;=\; \boldsymbol{x}\big({}_{\boldsymbol{p}|}\mathbf{D}_{|\boldsymbol{p}}\big)\boldsymbol{x} > 0.$$

Análogamente, por pertenecer a $\mathcal{C}\left(\mathbf{C}_{|\boldsymbol{q}^*}\right)$ y ser $\mathbf{C}$ invertible, $\boldsymbol{v}$ es de la forma $\boldsymbol{v} = \left(\mathbf{C}_{|\boldsymbol{q}^*}\right)\boldsymbol{y}$ para algún $\boldsymbol{y} \in \mathbb{R}^{q^*}$ no nulo. Por tanto

$$\boldsymbol{v}\mathbf{A}\boldsymbol{v} \;=\; \boldsymbol{y}\big({}_{\boldsymbol{q}^*|}(\mathbf{C}^{\intercal})\big)\mathbf{A}\left(\mathbf{C}_{|\boldsymbol{q}^*}\right)\boldsymbol{y} \;=\; \boldsymbol{y}\big({}_{\boldsymbol{q}^*|}(\mathbf{C}^{\intercal})\mathbf{A}\mathbf{C}_{|\boldsymbol{q}^*}\big)\boldsymbol{y} \;=\; \boldsymbol{y}\big({}_{\boldsymbol{q}^*|}\mathbf{D}^*{}_{|\boldsymbol{q}^*}\big)\boldsymbol{y} \leq 0.$$

Así, como ambos son subespacios de $\mathbb{R}^n$ con intersección $\{\mathbf{0}\}$, tenemos que $\mathcal{C}\left(\mathbf{B}_{|\boldsymbol{p}}\right) + \mathcal{C}\left(\mathbf{C}_{|\boldsymbol{q}^*}\right) \subset \mathbb{R}^n$; consecuentemente (Corolario 10.1.3 en la página 132)

$$n = \dim \mathbb{R}^n \geq \dim\left(\mathcal{C}\left(\mathbf{B}_{|\boldsymbol{p}}\right)+\mathcal{C}\left(\mathbf{C}_{|\boldsymbol{q}^*}\right)\right) = \dim \mathcal{C}\left(\mathbf{B}_{|\boldsymbol{p}}\right)+\dim \mathcal{C}\left(\mathbf{C}_{|\boldsymbol{q}^*}\right)-\dim\left(\mathcal{C}\left(\mathbf{B}_{|\boldsymbol{p}}\right)\cap\mathcal{C}\left(\mathbf{C}_{|\boldsymbol{q}^*}\right)\right) = p+(n-p^*)+0,$$

es decir, $n \geq n + (p - p^*)$, y por tanto $\boxed{p \leq p^*}$.

De igual manera, si repetimos el razonamiento pero con las submatrices principales ${}_{\boldsymbol{p}^*|}\mathbf{D}^*{}_{|\boldsymbol{p}^*}$ y ${}_{\boldsymbol{q}|}\mathbf{D}_{|\boldsymbol{q}}$ y los subespacios $\mathcal{C}\left(\mathbf{C}_{|\boldsymbol{p}^*}\right)$ y $\mathcal{C}\left(\mathbf{B}_{|\boldsymbol{q}}\right)$ llegaremos a la conclusión de que $\boxed{p^* \leq p}$.

Así, $p = p^*$, es decir, ambas matrices tienen igual número de componentes positivas y, como tienen el mismo número de ceros en la diagonal, también tienen el mismo número de componentes negativas. □

## 19.3. Completar el cuadrado para clasificar las formas cuadráticas

☞ Ahora usaremos el signo de los pivotes de la matriz diagonalizada por congruencia (y su rango) para decidir el signo de las formas cuadráticas.

Se denomina *"completar el cuadrado"* a expresar una forma cuadrática $\boldsymbol{x}\mathbf{A}\boldsymbol{x}$ como sumas (y/o restas) de términos al cuadrado. Pues bien, la diagonalización por congruencia nos permite encontrar diversas formas de completar el cuadrado. Basta darse cuenta de que dada $\mathbf{A}$ (simétrica), para toda $\mathbf{B}$ invertible tal que $\mathbf{B}^{\intercal}\mathbf{A}\mathbf{B}$ es diagonal, y dado que $(\mathbf{B}^{-1})^{\intercal} = \left(\mathbf{B}^{\intercal}\right)^{-1}$, tenemos que

$$\mathbf{D} = (\mathbf{B}^{\intercal})\mathbf{A}\mathbf{B} \qquad \Rightarrow \qquad \mathbf{A} = \big((\mathbf{B}^{-1})^{\intercal}\big)\mathbf{D}(\mathbf{B}^{-1})$$

y, por tanto, denotando con $\boldsymbol{y}$ al vector $(\mathbf{B}^{-1})\boldsymbol{x}$, la forma cuadrática se puede expresar como

$$\boldsymbol{x}\mathbf{A}\boldsymbol{x} = \boldsymbol{y}\mathbf{D}\boldsymbol{y} = \sum d_{jj}\left(y_j\right)^2;$$

es decir, como una suma de los cuadrados de los elementos de $\boldsymbol{y}$ ponderados por los elementos de la diagonal de $\mathbf{D}$. Así, si todos los pivotes de $\mathbf{D}$ son positivos, la correspondiente forma cuadrática es definida positiva. Por tanto, hemos demostrado la siguiente proposición:

**Proposición 19.3.1.** *Una matriz real y simétrica de orden n es definida positiva si y sólo si es congruente con una matriz diagonal cuyos pivotes son todos mayores que cero.*

**Corolario 19.3.2.** *Si una matriz real y simétrica de orden n es congruente con una matriz diagonal cuyos pivotes son todos mayores que cero, entonces sus autovalores también son mayores que cero.*

Es lo que ocurría con la matriz del Ejemplo 46 en la página 217. Vamos a repetir su diagonalización, pero ahora realizando solo transformaciones *Tipo I* (así minimizamos el número de operaciones y, además, no modificamos su determinante, por lo que el producto de los sucesivos pivotes de $\mathbf{D}$ calcula los sucesivos *menores* de la matriz):

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \xrightarrow[}{\tau\,[(\frac{1}{2})\mathbf{1}+\mathbf{2}]} \begin{bmatrix} 2 & 0 & 0 \\ -1 & \frac{3}{2} & -1 \\ 0 & -1 & 2 \end{bmatrix} \xrightarrow[\tau\,[(\frac{1}{2})\mathbf{1}+\mathbf{2}]]{} \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & -1 \\ 0 & -1 & 2 \end{bmatrix} \xrightarrow[}{\tau\,[(\frac{2}{3})\mathbf{2}+\mathbf{3}]} \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & -1 & \frac{4}{3} \end{bmatrix} \xrightarrow[\tau\,[(\frac{2}{3})\mathbf{2}+\mathbf{3}]]{} \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{3}{2} & 0 \\ 0 & 0 & \frac{4}{3} \end{bmatrix}.$$

Como

$$\mathbf{B} = \mathbf{I}\,{}_{\tau\,[(\frac{1}{2})\mathbf{1}+\mathbf{2}]\,\tau\,[(\frac{2}{3})\mathbf{2}+\mathbf{3}]} \quad\Longrightarrow\quad \mathbf{B}^{-1} = \mathbf{I}\,{}_{\tau\,[(\frac{-2}{3})\mathbf{2}+\mathbf{3}]\,\tau\,[(\frac{-1}{2})\mathbf{1}+\mathbf{2}]} = \begin{bmatrix} 1 & -1/2 & 0 \\ 0 & 1 & -2/3 \\ 0 & 0 & 1 \end{bmatrix}.$$

Así, la forma cuadrática $\boldsymbol{x}\mathbf{A}\boldsymbol{x}$ se puede reescribir como suma de cuadrados:

$$\begin{aligned} \boldsymbol{x}\mathbf{A}\boldsymbol{x} &= \begin{pmatrix} x, & y, & z, \end{pmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -2/3 & 1 \end{bmatrix} \begin{bmatrix} 2 & & \\ & 3/2 & \\ & & 4/3 \end{bmatrix} \begin{bmatrix} 1 & -1/2 & 0 \\ 0 & 1 & -2/3 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &&= \boldsymbol{x}\big((\mathbf{B}^{-1})^{\mathsf{T}}\big)\mathbf{D}(\mathbf{B}^{-1})\boldsymbol{x} \\ &= \Big( \big(x-\tfrac{1}{2}y\big), \ \big(y-\tfrac{2}{3}z\big), \ z, \Big) \begin{bmatrix} 2 & & \\ & 3/2 & \\ & & 4/3 \end{bmatrix} \begin{pmatrix} (x-(1/2)y) \\ (y-(2/3)z) \\ z \end{pmatrix} &&= \big(\mathbf{B}^{-1}\boldsymbol{x}\big)\mathbf{D}\big(\mathbf{B}^{-1}\boldsymbol{x}\big) \\ &= 2\left(x-\tfrac{1}{2}y\right)^2 + \tfrac{3}{2}\left(y-\tfrac{2}{3}z\right)^2 + \tfrac{4}{3}(z)^2 &&= \sum \big({}_{j|}\mathbf{D}_{|j}\big)\big(w_j\big)^2. \end{aligned}$$

donde $\boldsymbol{w} = (\mathbf{B}^{-1})\boldsymbol{x}$. Así, $\boldsymbol{x}\mathbf{A}\boldsymbol{x} = \boldsymbol{w}\mathbf{D}\boldsymbol{w}$ es suma de tres términos al cuadrado (los componentes de $\boldsymbol{w}$) multiplicados por los correspondientes pivotes de $\mathbf{D}$. Como dichos pivotes son positivos, la forma cuadrática es definida positiva; algo que podemos verificar calculando los autovalores (las raíces de $|\mathbf{A}-\lambda\mathbf{I}|=0$):

$$\begin{vmatrix} 2-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 2-\lambda \end{vmatrix} = (2-\lambda)^3 - 2(2-\lambda) = 0 \quad\Longrightarrow\quad \begin{cases} \lambda = 2 \\ (2-\lambda)^2 - 2 = \lambda^2 - 4\lambda + 2 \begin{cases} \lambda = 2+\sqrt{2} \\ \lambda = 2-\sqrt{2} \end{cases} \end{cases}.$$

Por el contrario, la matriz diagonal del Ejemplo 47 en la página 218 tiene dos pivotes negativos y uno positivo, así que la correspondiente forma cuadrática es indefinida (sus autovalores son $-1$ (doble) y 2).

# *Apéndices a la lección*

## 19.A. Factorización de Cholesky de una matriz definida positiva

De los Teoremas 4.2.1 y 5.1.16 y del Corolario 5.1.17 se deduce el siguiente

**Corolario 19.A.1.** *Toda matriz $\mathbf{A}$ se puede factorizar en $\mathbf{A} = \mathbf{K}\mathbf{U}$ donde $\mathbf{K}$ es pre-escalonada y $\mathbf{U}$ es triangular superior unitaria.*

*Demostración.* Sea $\mathbf{K} = \mathbf{A}(\mathbf{I}_{\tau_1\cdots\tau_k})$ una forma pre-escalonada obtenida mediante eliminación "de izquierda a derecha". Por el Teorema 4.2.1 en la página 59 necesariamente $\mathbf{I}_{\tau_1\cdots\tau_k}$ es triangular superior unitaria. Si denotamos por $\mathbf{U}$ a $(\mathbf{I}_{\tau_1\cdots\tau_k})^{-1}$, que también es triangular superior unitaria, entonces $\mathbf{A} = \mathbf{K}\mathbf{U}$. □

Más interesante es el caso de las matrices cuyas submatrices principales superiores son de rango completo.

**Proposición 19.A.2** (Factorización LU)**.** *Sea $\mathbf{A}$ de orden $n$. Si para $p = 1:(n-1)$ las submatrices ${}_{(1:p)|}\mathbf{A}_{|(1:p)}$ son invertibles, entonces $\mathbf{A}$ se puede factorizar en la forma $\mathbf{A} = \mathbf{L}\mathbf{U}$, donde $\mathbf{L}$ es triangular inferior y $\mathbf{U}$ es triangular superior unitaria. Si $\mathbf{A}$ es invertible, esta factorización es única.*

*Demostración.* Por el anterior corolario sabemos que podemos factorizar $\mathbf{A}$ en $\mathbf{KU}$ donde $\mathbf{K} = \mathbf{A}(\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_h})$ es una forma pre-escalonada obtenida mediante eliminación "de izquierda a derecha" y $\mathbf{U}$ (triangular superior unitaria) es la inversa de $\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_h}$. Ahora necesitamos demostrar que *si las submatrices* ${}_{(1:p)|}\mathbf{A}_{|(1:p)}$ *son invertibles, entonces los* $p$ *primeros componentes de la diagonal de* $\mathbf{K}$ *son pivotes* (i.e., no son nulos); y por tanto $\mathbf{K}$ es escalonada (i.e, triangular inferior).

Para la matriz principal superior de orden 1 es trivial: como ${}_{(1,)|}\mathbf{A}_{|(1,)}$ ya está escalonada, entonces ${}_{(1,)|}\mathbf{K}_{|(1,)} = {}_{(1,)|}\mathbf{A}_{|(1,)}$; y como además es invertible, su único componente, ${}_{1|}\mathbf{A}_{|1}$, es distinto de cero.

Vamos a demostrar que si la forma pre-escalonada de la matriz principal superior de orden $p$, obtenida por eliminación "de izquierda a derecha", es triangular inferior y sin ceros en la diagonal, entonces la forma pre-escalonada obtenida por eliminación "de izquierda a derecha" de la matriz principal superior de orden $p+1$ también es triangular inferior, y tampoco tiene ceros en la diagonal si es de orden menor o igual a $n-1$.

Sea $\left({}_{(1:p)|}\mathbf{A}_{|(1:p)}\right)_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ la forma pre-escalonada de ${}_{(1:p)|}\mathbf{A}_{|(1:p)}$ obtenida por eliminación de "izquierda a derecha" y que, por hipótesis, tiene $p$ pivotes en la diagonal. Si aplicamos la misma secuencia de transformaciones elementales sobre la siguiente submatriz principal superior, la de orden $(p+1)$, entonces $\left({}_{(1:p+1)|}\mathbf{A}_{|(1:p+1)}\right)_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ tendrá exactamente los mismos $p$ pivotes en la diagonal (pues las $p$ primeras componentes de las $p$ primeras filas de ${}_{(1:p+1)|}\mathbf{A}_{|(1:p+1)}$ coinciden con las de ${}_{(1:p)|}\mathbf{A}_{|(1:p)}$). Por tanto, podemos usar esos $p$ pivotes para, mediante eliminación de "izquierda a derecha", anular las $p$ primeras componentes de la última columna, para obtener una matriz escalonada (i.e., triangular inferior).

Para finalizar la primera parte de la demostración: tras el proceso de eliminación de "izquierda a derecha" por el que hemos obtenido una matriz escalonada, ¿puede ser cero la última componente de la última columna? Nunca en el caso de las $(n-1)$ primeras submatrices principales superiores, puesto que son de rango completo (y tampoco al aplicar la eliminación sobre toda la matriz $\mathbf{A}$ si ésta es invertible).

En cuanto a la unicidad. Si $\mathbf{A} = \mathbf{L}_1\mathbf{U}_1$ y $\mathbf{A} = \mathbf{L}_2\mathbf{U}_2$ son dos factorizaciones LU de una matriz $\mathbf{A}$ de rango completo, entonces, multiplicando por $\mathbf{L}_2^{-1}$ por la izquierda y $\mathbf{U}_1^{-1}$ por la derecha tenemos que $(\mathbf{L}_2^{-1})\mathbf{L}_1 = \mathbf{U}_2(\mathbf{U}_1^{-1})$. Puesto que el producto de dos matrices triangulares superiores unitarias también es triangular superior unitario (Corolario 3.A.4) y el producto de dos matrices triangulares inferiores también es triangular inferior, necesariamente las matrices a ambos lados de la igualdad anterior son la matriz identidad (es la única matriz cuadrada que simultáneamente es triangular inferior y superior unitaria). Por tanto, $\mathbf{L}_1 = \mathbf{L}_2$ y $\mathbf{U}_1 = \mathbf{U}_2$. □

Ejercicio 97. Demuestre el siguiente

**Corolario 19.A.3** (Factorización LDU)**.** *Sea* $\mathbf{A}$ *de orden* $n$*. Tal que todas sus submatrices principales superiores son invertibles, entonces existe una única matriz triangular inferior unitaria* $\mathbf{L}$*, una única matriz triangular superior unitaria* $\mathbf{U}$ *y una única matriz diagonal* $\mathbf{D}$ *tales que* $\mathbf{A} = \mathbf{LDU}$*.*

La factorización LDU es incluso más simple cuando, además, la matriz es simétrica.

Ejercicio 98. Demuestre el siguiente

**Corolario 19.A.4.** $\mathbf{L} = \mathbf{U}^\intercal$ *cuando* $\mathbf{A} = \mathbf{LDU}$ *es la factorización LDU de una matriz simétrica e invertible.*

Ahora ya estamos en disposición de presentar la Factorización de Cholesky.

**Teorema 19.A.5** (Factorización de Cholesky)**.** *Si* $\mathbf{A}$ *de orden* $n$ *es simétrica y definida positiva, entonces existe una única matriz triangular superior* $\mathbf{C}$ *con componentes positivas en la diagonal tal que* $\mathbf{A} = \mathbf{C}(\mathbf{C}^\intercal)$*.*

*Demostración.* Por el corolario anterior, existe una matriz triangular inferior unitaria $\mathbf{L}$ y una matriz diagonal $\mathbf{D}$ tal que $\mathbf{A} = \mathbf{LD}(\mathbf{L}^\intercal)$. Puesto que las componentes de la diagonal de $\mathbf{D}$ son positivas, podemos definir la matriz diagonal $\mathbf{D}^*$ cuyas componentes son la raíz cuadrada de las componentes de $\mathbf{D}$, es decir, ${}_{i|}\mathbf{D}^*{}_{|i} =$

$\sqrt{_{i|}\mathbf{D}_{|i}}$. Entonces $\mathbf{C} = \mathbf{L}\mathbf{D}^*$ es una matriz real y triangular inferior con componentes positivas en la diagonal. Además, satisface $\mathbf{A} = \mathbf{C}(\mathbf{C}^\intercal)$. La unicidad proviene de la unicidad de la factorización LDU. □

## 19.B. Productos escalares en $\mathbb{R}^n$

Considere la siguiente función $\langle _ | _ \rangle_{\mathbf{A}} : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ , donde $\mathbf{A}$ es definida positiva.
$(\boldsymbol{x}, \boldsymbol{y}) \longmapsto \boldsymbol{x}\mathbf{A}\boldsymbol{y}$

Ejercicio 99. Demuestre que la anterior función es un producto escalar en $\mathbb{R}^n$.

*Pista.* Debe verificar los cuatro axiomas indicados en la Sección 11.3.1 en la página 151, que son los mismos que en el Ejercicio 11 en la página 27.

La relación entre productos escalares en $\mathbb{R}^n$ y matrices definidas positivas es mucho más estrecha de lo que se pueda pensar...

Sea $\mathcal{V}$ un subespacio de $\mathbb{R}^m$ de dimensión $n$ y sea $\mathbf{B}$ ($m$ por $n$) tal que sus columnas son una base de $\mathcal{V}$. Considere dos vectores $\boldsymbol{x}$ e $\boldsymbol{y}$ de $\mathcal{V}$, que por tanto son de la forma $\boldsymbol{x} = \mathbf{B}\boldsymbol{a}$ e $\boldsymbol{y} = \mathbf{B}\boldsymbol{c}$. Entonces su producto escalar se puede expresar como

$$\boldsymbol{x} \cdot \boldsymbol{z} = \mathbf{B}\boldsymbol{a} \cdot \mathbf{B}\boldsymbol{c} = \boldsymbol{a}(\mathbf{B}^\intercal)\mathbf{B}\boldsymbol{c} = \boldsymbol{a}\mathbf{A}\boldsymbol{c},$$

donde $\mathbf{A} = (\mathbf{B}^\intercal)\mathbf{B}$ es definida positiva (Proposición 19.1.1). Por tanto, el producto punto entre $\boldsymbol{x}$ e $\boldsymbol{y}$ se puede expresar con un producto escalar entre las coordenadas de $\boldsymbol{x}$ e $\boldsymbol{y}$ respecto a la base formada por las columnas de $\mathbf{B}$.

# Resumen de los temas por lecciones

## A.1. Resumen del Tema 1

El tema consta de cinco lecciones que resumimos a continuación.

### Lección 1.

En la primera lección solo se establece la notación para vectores, matrices, y operadores selectores junto con algunas reglas de reescritura. Por ejemplo, estas dos reglas de reescritura

$$\boxed{(\boldsymbol{a}+\boldsymbol{b})_{|i} = \boldsymbol{a}_{|i} + \boldsymbol{b}_{|i}} \qquad \text{y} \qquad \boxed{(\lambda \boldsymbol{b})_{|i} = \lambda(\boldsymbol{b}_{|i})}$$

definen la suma de vectores y su producto por escalares. Jugando con estas reglas se deducen las propiedades de la suma de vectores y producto por escalares que permitirán definir el *espacio vectorial* $\mathbb{R}^n$ en el Tema 2. Las reglas

$$\boxed{(\mathbf{A}+\mathbf{B})_{|j} = \mathbf{A}_{|j} + \mathbf{B}_{|j}} \qquad \text{y} \qquad \boxed{(\lambda \mathbf{B})_{|j} = \lambda(\mathbf{B}_{|j})}$$

definen la suma de matrices y su producto por escalares (operando con las columnas). Como esta reglas son esencialmente iguales a las anteriores, se obtienen propiedades análogas, lo que permitirá definir el espacio vectorial $\mathbb{R}^{n\times m}$. ¡El juego con las reglas de reescritura, independientemente del significado que tengan los símbolos $\boldsymbol{a}$ o $\mathbf{A}$ da lugar a las mismas propiedades! Con la transposición de una matriz

$$\boxed{(\mathbf{A}^\intercal)_{|j} = {}_{j|}\mathbf{A}}$$

mostraremos que las operaciones con matrices también se pueden definir como operaciones entre filas ${}_{i|}\mathbf{A}$ (y también entre componentes ${}_{i|}\mathbf{A}_{|j}$). Teniendo en cuenta que $\boxed{\lambda \boldsymbol{a} = \boldsymbol{a}\lambda}$ y que $\boxed{\lambda \mathbf{A} = \mathbf{A}\lambda}$, y jugando con la notación se deducen una serie de reglas de reescritura que se usarán en las siguientes lecciones: [1]

**Reglas distributivas**

$$\begin{aligned} (\boldsymbol{a}+\boldsymbol{b})_{|i} &= \boldsymbol{a}_{|i} + \boldsymbol{b}_{|i} & {}_{i|}(\boldsymbol{a}+\boldsymbol{b}) &= {}_{i|}\boldsymbol{a} + {}_{i|}\boldsymbol{b} \\ (\mathbf{A}+\mathbf{B})_{|j} &= \mathbf{A}_{|j} + \mathbf{B}_{|j} & {}_{i|}(\mathbf{A}+\mathbf{B}) &= {}_{i|}\mathbf{A} + {}_{i|}\mathbf{B} \end{aligned}$$

**Reglas asociativas (desplazando el paréntesis)**

$$\begin{aligned} (\lambda \boldsymbol{b})_{|i} &= \lambda(\boldsymbol{b}_{|i}) & {}_{i|}(\boldsymbol{b}\lambda) &= ({}_{i|}\boldsymbol{b})\lambda \\ (\lambda \mathbf{A})_{|j} &= \lambda(\mathbf{A}_{|j}) & {}_{i|}(\mathbf{A}\lambda) &= ({}_{i|}\mathbf{A})\lambda \end{aligned}$$

[1] De hecho, la librería de Python solo se implementan las reglas recuadradas... y entonces ¡todo lo demás funciona automáticamente!

**Intercambio entre el escalar y el operador**

$$\begin{aligned} (\boldsymbol{b}\lambda)_{|i} &= (\boldsymbol{b}_{|i})\lambda & {}_{i|}(\lambda\boldsymbol{b}) &= \lambda({}_{i|}\boldsymbol{b}) \\ (\mathbf{A}\lambda)_{|j} &= (\mathbf{A}_{|j})\lambda & {}_{i|}(\lambda\mathbf{A}) &= \lambda({}_{i|}\mathbf{A}) \end{aligned}$$

## Lección 2

Comienza con el *producto punto* (o producto escalar usual) de dos vectores de $\mathbb{R}^n$ y sus propiedades.[2]

$$\boxed{\boldsymbol{a}\cdot\boldsymbol{b} = \textstyle\sum_i a_i b_i}$$

*El uso del producto punto nos dota de una notación muy compacta (sin sumatorios).*

La Lección 2 trata de *combinaciones lineales* de vectores de $\mathbb{R}^n$ y su notación matricial $\mathbf{A}\boldsymbol{b}$ (combinación de las columnas de $\mathbf{A}$) y $\boldsymbol{a}\mathbf{B}$ (combinación de las filas de $\mathbf{B}$).

$$\boxed{\mathbf{A}\boldsymbol{b} = \textstyle\sum_j (\mathbf{A}_{|j})b_j} \qquad \text{y} \qquad \boxed{\boldsymbol{a}\mathbf{B} = (\mathbf{B}^\intercal)\boldsymbol{a}}$$

Usando las reglas de reescritura de la lección anterior, se deducen nuevas reglas para el producto de una matriz por un vector a su derecha $\mathbf{A}\boldsymbol{b}$:

**Propiedades de linealidad**

- $\mathbf{A}(\boldsymbol{b}+\boldsymbol{c}) = \mathbf{A}\boldsymbol{b} + \mathbf{A}\boldsymbol{c}$
- $\mathbf{A}(\lambda\boldsymbol{b}) = \lambda(\mathbf{A}\boldsymbol{b})$

**Otras propiedades**

- $\mathbf{A}(\lambda\boldsymbol{b}) = (\lambda\mathbf{A})\boldsymbol{b}$
- $(\mathbf{A}+\mathbf{B})\boldsymbol{c} = \mathbf{A}\boldsymbol{c} + \mathbf{B}\boldsymbol{c}$
- $\mathbf{A}(\mathbf{B}\boldsymbol{c}) = \left[\mathbf{A}(\mathbf{B}_{|1});\quad \ldots \quad \mathbf{A}(\mathbf{B}_{|n});\right]\boldsymbol{c}$

y propiedades análogas para el producto de un vector por una matriz $\boldsymbol{a}\mathbf{B}$.

Adicionalmente, y solo en las transparencias de la clase correspondiente a esta lección, se adelanta la interpretación geométrica de un sistema de ecuaciones $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ ... ¿qué combinaciones lineales de las columnas ($\mathbf{A}\boldsymbol{x}$) son iguales al vector del lado derecho $\boldsymbol{b}$?

## Lección 3.

Trata sobre el *producto de matrices*.

$$\boxed{(\mathbf{AB})_{|j} = \mathbf{A}(\mathbf{B}_{|j})}$$

(*cada columna de* $\mathbf{AB}$ *es una combinación de las columnas de* $\mathbf{A}$). Jugando con la definición y con las reglas de reescritura de las lecciones anteriores se deducen las siguientes propiedades

- $\mathbf{A}(\mathbf{B}\boldsymbol{c}) = (\mathbf{AB})\boldsymbol{c}$
- $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
- $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
- $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
- $\mathbf{A}(\lambda\mathbf{B}) = \lambda(\mathbf{AB})$
- $\mathbf{A}(\lambda\mathbf{B}) = (\lambda\mathbf{A})\mathbf{B}$
- $\mathbf{IA} = \mathbf{A}$
- $\mathbf{AI} = \mathbf{A}$

[2] Como en este material se hace una marcada distinción entre vectores y matrices, aquí carece de sentido "transponer un vector". Así evitamos el frecuente abuso de notación del que hacen uso otros textos.

Y continuado con el mismo juego de manipulación de símbolos, se deducen nuevas interpretaciones del producto

$$ {}_{i|}(\mathbf{AB}) = ({}_{i|}\mathbf{A})\mathbf{B} \qquad \text{y} \qquad {}_{i|}(\mathbf{AB})_{|j} = ({}_{i|}\mathbf{A}) \cdot (\mathbf{B}_{|j}) $$

es decir, las filas de $\mathbf{AB}$ son combinaciones lineales de las filas de $\mathbf{B}$, y los elementos de $\mathbf{AB}$ son productos punto de las filas de $\mathbf{A}$ con las columnas de $\mathbf{B}$. Aquí se evidencian las bondades de la notación. Una expresión como

$$ {}_{i|}\mathbf{AB}_{|j} $$

se puede interpretar como el elemento de la fila $i$, y columna $j$ de $\mathbf{AB}$, como el elemento $i$-ésimo de la combinación de las columnas $\mathbf{A}(\mathbf{B}_{|j})$, como el elemento $j$-ésimo de una combinación de las filas $({}_{i|}\mathbf{A})\mathbf{B}$, o como el producto escalar de la fila ${}_{i|}\mathbf{A}$ con la columna $\mathbf{B}_{|j}$. Todas estas interpretaciones son correctas, y todas ellas están sugeridas en la expresión, ${}_{i|}\mathbf{AB}_{|j}$. Es destacable la potencia computacional de la notación (véase la demostración de $(\mathbf{AB})^\intercal = \mathbf{B}^\intercal(\mathbf{A}^\intercal)$ así como su implementación en la librería de Python).

## Lección 4

Trata sobre las *transformaciones elementales*, y su uso en el *método de eliminación*.

Aquí solo consideramos dos tipos de transformaciones elementales:

- $\boxed{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}$ suma $\lambda$ veces el vector $i$-ésimo al $j$-ésimo.
- $\boxed{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}$ multiplica por $\alpha$ el vector $i$-ésimo.

Llamamos *matriz elemental* a la matriz resultante de aplicar una única transformación elemental sobre las columnas (o bien sobre las filas) de una matriz identidad

$$ \mathbf{I}_{\boldsymbol{\tau}} \qquad \text{o} \qquad {}_{\boldsymbol{\tau}}\mathbf{I} $$

Dada la transformación elemental, $\boldsymbol{\tau}$, las dos matrices elementales de más arriba son una la transpuesta de la otra. Por tanto, la transpuesta de una matriz elemental es otra matriz elemental. Aplicar una transformación elemental a una matriz es equivalente a multiplicarla por la correspondiente matriz elemental, es decir

$$ \mathbf{A}_{\boldsymbol{\tau}} = \mathbf{A}(\mathbf{I}_{\boldsymbol{\tau}}) \qquad \text{y} \qquad {}_{\boldsymbol{\tau}}\mathbf{A} = ({}_{\boldsymbol{\tau}}\mathbf{I})\mathbf{A}. $$

Para describir la aplicación de la secuencia $\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k$ de $k$ trasformaciones elementales de las *columnas* de $\mathbf{A}$ usamos el esquema. [3]

$$ \mathbf{A} \xrightarrow{\boldsymbol{\tau}_1} \left(\mathbf{A}_{\boldsymbol{\tau}_1}\right) \xrightarrow{\boldsymbol{\tau}_2} \left(\mathbf{A}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2}\right) \cdots \xrightarrow{\boldsymbol{\tau}_k} \left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) \quad \text{o agregando varios pasos} \quad \mathbf{A} \xrightarrow[\boldsymbol{\tau}_p]{\overset{\boldsymbol{\tau}_1}{\vdots}} \left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p}\right) \xrightarrow[\boldsymbol{\tau}_k]{\overset{\boldsymbol{\tau}_{p+1}}{\vdots}} \left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right). $$

Cuando se aplica la sucesión $\boldsymbol{\tau}_1 \ldots \boldsymbol{\tau}_k$ de $k$ transformaciones elementales sobre las *columnas*, o bien la sucesión $\boldsymbol{\tau}_k \ldots \boldsymbol{\tau}_1$ de $k$ transformaciones elementales sobre las *filas* (nótese el distinto orden en las sucesiones) se obtienen relaciones similares a las que encontramos al aplicar una única transformación:

$$ \mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathbf{A}(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}) \qquad \text{y} \qquad {}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\mathbf{A} = ({}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\mathbf{I})\mathbf{A}. $$

[3] De manera análoga, para describir una secuencia de trasformaciones elementales de las *filas* de $\mathbf{A}$ (donde $\boldsymbol{\tau}_1$ es la primera que se aplica, luego $\boldsymbol{\tau}_1$, ... y por último $\boldsymbol{\tau}_k$) se debe usar el siguiente esquema

$$ \mathbf{A} \xrightarrow[\boldsymbol{\tau}_1]{} ({}_{\boldsymbol{\tau}_1}\mathbf{A}) \xrightarrow[\boldsymbol{\tau}_2]{} ({}_{\boldsymbol{\tau}_2\boldsymbol{\tau}_1}\mathbf{A}) \cdots \xrightarrow[\boldsymbol{\tau}_k]{} ({}_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{A}) \quad \text{o agregando varios pasos} \quad \mathbf{A} \xrightarrow[\boldsymbol{\tau}_1]{\overset{\boldsymbol{\tau}_p}{\vdots}} ({}_{\boldsymbol{\tau}_p\cdots\boldsymbol{\tau}_1}\mathbf{A}) \xrightarrow[\boldsymbol{\tau}_{p+1}]{\overset{\boldsymbol{\tau}_k}{\vdots}} ({}_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{A}); $$

donde la secuencia $\boldsymbol{\tau}_k \ldots \boldsymbol{\tau}_1$ es la transpuesta de la secuencia $\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k$, es decir, $(\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k)^\intercal = \boldsymbol{\tau}_k \ldots \boldsymbol{\tau}_1$, pues al actuar por la izquierda, las primeras transformaciones que se aplican son la que están más a la derecha de la secuencia. Por tanto ${}_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{I}$ es la transpuesta de $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ (cuando ambas tienen el mismo orden).

Es posible realizar un *intercambio* de posición entre dos vectores mediante una sucesión de transformaciones elementales:

- $\boxed{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}}$ intercambia de posición los vectores $i$-ésimo y $j$-ésimo.

Llamamos matriz de intercambio a la matriz resultante de aplicar un único intercambio entre dos columnas (o bien dos filas) de una matriz identidad.

$$\mathbf{I}_{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}} \qquad \text{o} \qquad {}_{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}}\mathbf{I}$$

Las matrices intercambio son simétricas (por lo que estas dos de arriba son iguales si tienen el mismo orden). Aplicar un intercambio a una matriz es equivalente a multiplicarla por la correspondiente matriz de intercambio, es decir

$$\mathbf{A}_{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}} = \mathbf{A}\left(\mathbf{I}_{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}}\right) \qquad \text{y} \qquad {}_{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}}\mathbf{A} = \left({}_{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}}\mathbf{I}\right)\mathbf{A}.$$

La aplicación de una sucesión de intercambios da lugar a un reordenamiento de los vectores. Denominamos *matriz permutación* a la matriz que resulta tras una sucesión de intercambios en las columnas (o en las filas) de la matriz identidad. De nuevo tenemos que

$$A_{\underset{[\circlearrowleft]}{\boldsymbol{\tau}}} = \mathbf{A}\left(I_{\underset{[\circlearrowleft]}{\boldsymbol{\tau}}}\right) \qquad \text{y} \qquad {}_{\underset{[\circlearrowleft]}{\boldsymbol{\tau}}}A = \left({}_{\underset{[\circlearrowleft]}{\boldsymbol{\tau}}}I\right)\mathbf{A}$$

(la flecha circular denota un reordenamiento de las columnas (o de las filas) de la matriz).

Mediante una sucesión de transformaciones elementales es posible pre-escalonar cualquier matriz. La demostración de este importante teorema describe la implementación del método en Python. A este procedimiento se le llama *Método de eliminación*. Hay dos extensiones más: la *eliminación Gaussiana* que escalona la matriz reordenando las columnas y la eliminación *Gauss-Jordan*, que reduce la matriz escalonada.

## Lección 5.

Trata sobre la *inversión de las transformaciones elementales* y la *inversión de matrices* (cuando es posible) aplicando el Método de eliminación Gauss-Jordan.

Definimos la inversa de $\mathbf{A}$ de orden $n$ (por tanto *cuadrada*) como aquella matriz $\mathbf{B}$ tal que $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$. Y demostramos que, si existe $\mathbf{B}$, es única y la denotamos por $\mathbf{A}^{-1}$. A continuación se demuestra que:

- Si $\mathbf{A}$ y $\mathbf{B}$ tienen inversa, entonces $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}(\mathbf{A}^{-1})$.
- Si $\mathbf{B}$ es invertible, entonces $\mathbf{AB}$ es invertible si y solo si $\mathbf{A}$ es invertible.
- Si $\mathbf{A}$ es invertible, entonces $\mathbf{AB}$ es invertible si y solo si $\mathbf{B}$ es invertible.
- Si $\mathbf{A}$ es invertible, entonces $\left(\mathbf{A}^{\intercal}\right)^{-1} = (\mathbf{A}^{-1})^{\intercal}$.
- Si $\mathbf{A}$ tiene alguna columna (o fila) nula entonces no tiene inversa.

Todas *las transformaciones elementales* se pueden deshacer (todas *son invertibles*), lo que implica que *todas las matrices elementales son invertibles*. En la segunda parte de lección relacionamos la invertibilidad de una matriz con las transformaciones elementales, pues cualquier matriz de la forma $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ es invertible por ser producto de matrices elementales:

$$\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)\left(\mathbf{I}_{\boldsymbol{\tau}_k^{-1}\cdots\boldsymbol{\tau}_1^{-1}}\right) = \mathbf{E}_1\cdots\mathbf{E}_k\cdot\mathbf{E}_k^{-1}\cdots\mathbf{E}_1^{-1} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k\cdot\boldsymbol{\tau}_k^{-1}\cdots\boldsymbol{\tau}_1^{-1}} = \mathbf{I}.$$

Siguiendo esta idea, se demuestra que

- Si $\mathbf{A}$ es producto de matrices elementales, entonces es invertible.

- Si $\mathbf{A}$ tiene columnas que son combinación lineal del resto, entonces no es invertible.
- $\mathbf{A}$ es invertible si y solo si sus formas escalonadas $\mathbf{L}$ son invertibles.

A continuación se demuestra que toda matriz cuadrada $\mathbf{L}$ escalonada y sin columnas nulas se puede transformar en la matriz identidad mediante transformaciones elementales (de nuevo la demostración indica los pasos para implementar el algoritmo en Python). Esto da lugar a un corolario que establece que las siguientes propiedades son equivalentes

1. El resultado de escalonar $\mathbf{A}$ no tiene columnas nulas.
2. $\mathbf{A}$ es producto de matrices elementales.
3. $\mathbf{A}$ tiene inversa.

A continuación se demuestra que si $\mathbf{A}$ y $\mathbf{B}$ son cuadradas y del mismo orden, y $\mathbf{AB} = \mathbf{I}$, entonces $\mathbf{A}$ y $\mathbf{B}$ son una la inversa de la otra.

En la lección anterior vimos que el Método de eliminación permite encontrar una forma pre-escalonada de toda matriz (Teorema 4.2.2 en la página 61)

$$\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_p} = \mathbf{K},$$

y, si $\mathbf{K}$ no tiene columnas nulas, se puede continuar con las transformaciones elementales hasta transformar $\mathbf{K}$ en $\mathbf{I}$ (Teorema 5.1.10 en la página 72)

$$\mathbf{K}_{\boldsymbol{\tau}_{(p+1)} \cdots \boldsymbol{\tau}_k} = \mathbf{I}.$$

Por tanto, la sucesión de $k$ transformaciones elementales $\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_p$ , $\boldsymbol{\tau}_{(p+1)} \cdots \boldsymbol{\tau}_k$ transforma $\mathbf{A}$ en $\mathbf{I}$, es decir,

$$\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}\right) = \mathbf{I};$$

así, si $\mathbf{L}$ no tiene columnas nulas, entonces $\boxed{\mathbf{A}^{-1} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}}$. Por otra parte, $\mathbf{A}^{-1}$ no existe cuando $\mathbf{L}$ tiene columnas nulas.

La lección finaliza proponiendo un método para encontrar la inversa a la vez que se transforma $\mathbf{A}$ en $\mathbf{I}$, y dando una definición de *rango* de una matriz (el número de columnas no nulas de cualquiera de sus formas escalonadas).

**Aquí finaliza el primer tema del curso.**

## Soluciones a los Ejercicios

**Ejercicio 1.** En `(3*b)|2`, primero multiplica el vector por 3. Después selecciona la segunda componente. En `3*(b|2)`, primero se selecciona la segunda componente. Luego multiplica dicha componente por 3. `3*b|2` hace exactamente lo mismo que con la primera expresión (pues `*` tiene precedencia sobre `|`).

□

**Ejercicio 2.** En `b+b*2` lo primero es el producto. En `b+b|2` lo primero es la suma.

□

**Ejercicio 3.** Como en Python el operador `+` tiene precedencia sobre `|`, en la expresión `b+b|2` primero se suman los vectores, y luego se selecciona la segunda componente del resultado.

Pero para nosotros la expresión es confusa, ya que podríamos entender $\boldsymbol{b} + \boldsymbol{b}_{|2}$ (la suma del vector $\boldsymbol{b}$ con la segunda componente de $\boldsymbol{b}$); y la operación suma entre un vector y un número *no está definida*.

Para nosotros, el modo correcto de escribir la operación es $\big(\boldsymbol{b} + \boldsymbol{b}\big)_{|2}$. Así que, por claridad, es recomendable escribir `(b+b)|2` también en Python.

□

**Ejercicio 4.** Mostraremos ambas estrategias en cada caso. En el caso de la segunda, en azul aparecen las operaciones entre números reales.

1. Estrategia 1:

$$\boldsymbol{a} + \boldsymbol{b} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix} = \begin{pmatrix} b_1 + a_1 \\ \vdots \\ b_n + a_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} + \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \boldsymbol{b} + \boldsymbol{a}.$$

   Estrategia 2: $(\boldsymbol{a} + \boldsymbol{b})_{|i} = \boldsymbol{a}_{|i} + \boldsymbol{b}_{|i} = \boldsymbol{b}_{|i} + \boldsymbol{a}_{|i} = (\boldsymbol{b} + \boldsymbol{a})_{|i}$.

2. Estrategia 1:

$$\boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c}) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \left( \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} + \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \right) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 + c_1 \\ \vdots \\ b_n + c_n \end{pmatrix} = \begin{pmatrix} a_1 + (b_1 + c_1) \\ \vdots \\ a_n + (b_n + c_n) \end{pmatrix} =$$

$$= \begin{pmatrix} (a_1 + b_1) + c_1 \\ \vdots \\ (a_n + b_n) + c_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix} + \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \left( \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \right) + \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = (\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c}.$$

   Estrategia 2:

$$\Big(\boldsymbol{a} + (\boldsymbol{b} + \boldsymbol{c})\Big)_{|i} = \boldsymbol{a}_{|i} + (\boldsymbol{b} + \boldsymbol{c})_{|i} = \boldsymbol{a}_{|i} + \Big(\boldsymbol{b}_{|i} + \boldsymbol{c}_{|i}\Big) = \Big(\boldsymbol{a}_{|i} + \boldsymbol{b}_{|i}\Big) + \boldsymbol{c}_{|i}$$
$$= (\boldsymbol{a} + \boldsymbol{b})_{|i} + \boldsymbol{c}_{|i} = \Big((\boldsymbol{a} + \boldsymbol{b}) + \boldsymbol{c}\Big)_{|i}.$$

3. Estrategia 1:

$$\boldsymbol{a} + \boldsymbol{0} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 + 0 \\ \vdots \\ a_n + 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \boldsymbol{a}.$$

   Estrategia 2: $(\boldsymbol{a} + \boldsymbol{0})_{|i} = \boldsymbol{a}_{|i} + \boldsymbol{0}_{|i} = \boldsymbol{a}_{|i} + 0 = \boldsymbol{a}_{|i}$.

4. Estrategia 1:
$$\boldsymbol{a}+(-\boldsymbol{a})=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}+\begin{pmatrix}-a_1\\ \vdots\\ -a_n\end{pmatrix}=\begin{pmatrix}a_1+(-a_1)\\ \vdots\\ a_n+(-a_n)\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}=\boldsymbol{0}.$$
Estrategia 2: $\Big(\boldsymbol{a}+(-\boldsymbol{a})\Big)_{|i}=\boldsymbol{a}_{|i}+\big(-\boldsymbol{a}\big)_{|i}=\boldsymbol{a}_{|i}+\big(-\big(\boldsymbol{a}_{|i}\big)\big)=0=\boldsymbol{0}_{|i}$.

5. Estrategia 1:
$$\lambda(\boldsymbol{a}+\boldsymbol{b})=\lambda\begin{pmatrix}a_1+b_1\\ \vdots\\ a_n+b_n\end{pmatrix}=\begin{pmatrix}\lambda(a_1+b_1)\\ \vdots\\ \lambda(a_n+b_n)\end{pmatrix}=\begin{pmatrix}\lambda a_1+\lambda b_1\\ \vdots\\ \lambda a_n+\lambda b_n\end{pmatrix}=\begin{pmatrix}\lambda a_1\\ \vdots\\ \lambda a_n\end{pmatrix}+\begin{pmatrix}\lambda b_1\\ \vdots\\ \lambda b_n\end{pmatrix}=\lambda\boldsymbol{a}+\lambda\boldsymbol{b}.$$
Estrategia 2: $\Big(\lambda(\boldsymbol{a}+\boldsymbol{b})\Big)_{|i}=\lambda\Big(\big(\boldsymbol{a}+\boldsymbol{b}\big)_{|i}\Big)=\lambda\Big(\boldsymbol{a}_{|i}+\boldsymbol{b}_{|i}\Big)=\lambda\big(\boldsymbol{a}_{|i}\big)+\lambda\big(\boldsymbol{b}_{|i}\big)=\big(\lambda\boldsymbol{a}\big)_{|i}+\big(\lambda\boldsymbol{b}\big)_{|i}=\big(\lambda\boldsymbol{a}+\lambda\boldsymbol{b}\big)_{|i}$.

6. Estrategia 1:
$$(\lambda+\varphi)\boldsymbol{a}=\begin{pmatrix}(\lambda+\varphi)a_1\\ \vdots\\ (\lambda+\varphi)a_n\end{pmatrix}=\begin{pmatrix}a_1(\lambda+\varphi)\\ \vdots\\ a_n(\lambda+\varphi)\end{pmatrix}=\begin{pmatrix}a_1\lambda+a_1\varphi\\ \vdots\\ a_n\lambda+a_n\varphi\end{pmatrix}=$$
$$=\begin{pmatrix}\lambda a_1+\varphi a_1\\ \vdots\\ \lambda a_n+\varphi a_n\end{pmatrix}=\begin{pmatrix}\lambda a_1\\ \vdots\\ \lambda a_n\end{pmatrix}+\begin{pmatrix}\varphi a_1\\ \vdots\\ \varphi a_n\end{pmatrix}=\lambda\boldsymbol{a}+\varphi\boldsymbol{a}.$$
Estrategia 2: $\Big((\lambda+\varphi)\boldsymbol{a}\Big)_{|i}=(\lambda+\varphi)\big(\boldsymbol{a}_{|i}\big)=\big(\boldsymbol{a}_{|i}\big)(\lambda+\varphi)=\big(\boldsymbol{a}_{|i}\big)\lambda+\big(\boldsymbol{a}_{|i}\big)\varphi=\lambda\big(\boldsymbol{a}_{|i}\big)+\varphi\big(\boldsymbol{a}_{|i}\big)=\big(\lambda\boldsymbol{a}\big)_{|i}+\big(\varphi\boldsymbol{a}\big)_{|i}=\big(\lambda\boldsymbol{a}+\varphi\boldsymbol{a}\big)_{|i}$. (Como la propiedad distributiva de los números reales del recordatorio tiene el factor común a la izquierda, en ambas estrategias hacen falta tres pasos para pasar de $(\lambda+\varphi)a_i$ a $\lambda a_i+\varphi a_i$: conmutativa del producto, distributiva y de nuevo conmutativa del producto en cada sumando.)

7. Estrategia 1:
$$\lambda(\varphi\boldsymbol{a})=\lambda\begin{pmatrix}\varphi a_1\\ \vdots\\ \varphi a_n\end{pmatrix}=\begin{pmatrix}\lambda(\varphi a_1)\\ \vdots\\ \lambda(\varphi a_n)\end{pmatrix}=\begin{pmatrix}(\lambda\varphi)a_1\\ \vdots\\ (\lambda\varphi)a_n\end{pmatrix}=(\lambda\varphi)\boldsymbol{a}.$$
Estrategia 2: $\Big(\lambda(\varphi\boldsymbol{a})\Big)_{|i}=\lambda\Big((\varphi\boldsymbol{a})_{|i}\Big)=\lambda\Big(\varphi\big(\boldsymbol{a}_{|i}\big)\Big)=(\lambda\varphi)\big(\boldsymbol{a}_{|i}\big)=\big((\lambda\varphi)\boldsymbol{a}\big)_{|i}$.

8. Estrategia 1:
$$1\boldsymbol{a}=\begin{pmatrix}1a_1\\ \vdots\\ 1a_n\end{pmatrix}=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}=\boldsymbol{a}.$$
Estrategia 2: $\big(1\boldsymbol{a}\big)_{|i}=1\big(\boldsymbol{a}_{|i}\big)=\boldsymbol{a}_{|i}$.

Cada cadena de la Estrategia 2 empieza en la componente $i$-ésima del vector del lado izquierdo de la igualdad y termina en la componente $i$-ésima del vector del lado derecho. Como esto vale para cualquier componente $i$, en los ocho casos los dos vectores son iguales componente a componente; es decir, son iguales.

□

**Ejercicio 5(a)** Primero multiplica el vector por 3 y luego selecciona la primera componente.

□

**Ejercicio 5(b)** Primero multiplica 3 por 1; así que selecciona la tercera componente del vector.

□

**Ejercicio 5(c)** Primero multiplica 1 por 3; así que selecciona la tercera componente del vector.

□

**Ejercicio 5(d)** Como en la primera expresión, primero multiplica el vector por 3 y luego selecciona la primera componente.

□

**Ejercicio 6(a)** Por la definición de transposición, $(\mathbf{A}^\intercal)_{|j} = {}_{j|}\mathbf{A}$. Como son dos vectores iguales, sus componentes $i$-ésimas son iguales (en la columna seleccionamos la componente por la izquierda, y en la fila por la derecha):

$$ {}_{i|}\big((\mathbf{A}^\intercal)_{|j}\big) = \big({}_{j|}\mathbf{A}\big)_{|i}. $$

Como la selección de componentes es asociativa, podemos quitar los paréntesis: ${}_{i|}(\mathbf{A}^\intercal)_{|j} = {}_{j|}\mathbf{A}_{|i}$.

□

**Ejercicio 6(b)** Volviendo a poner los paréntesis en el apartado anterior,

$$ \big({}_{i|}(\mathbf{A}^\intercal)\big)_{|j} = {}_{i|}(\mathbf{A}^\intercal)_{|j} = {}_{j|}\mathbf{A}_{|i} = {}_{j|}\big(\mathbf{A}_{|i}\big). $$

Como esto vale para cualquier componente $j$, los vectores ${}_{i|}(\mathbf{A}^\intercal)$ y $\mathbf{A}_{|i}$ son iguales.

□

**Ejercicio 6(c)** Columna a columna, por la definición de transposición y por el apartado anterior:

$$ \big((\mathbf{A}^\intercal)^\intercal\big)_{|i} = {}_{i|}\big(\mathbf{A}^\intercal\big) = \mathbf{A}_{|i}. $$

Como esto vale para cualquier columna $i$, las matrices $(\mathbf{A}^\intercal)^\intercal$ y $\mathbf{A}$ son iguales.

*Otra demostración.* Como dos matrices del mismo orden son iguales si lo son componente a componente, basta aplicar dos veces el apartado (a): ${}_{i|}\big((\mathbf{A}^\intercal)^\intercal\big)_{|j} = {}_{j|}(\mathbf{A}^\intercal)_{|i} = {}_{i|}\mathbf{A}_{|j}$.

□

**Ejercicio 7.** Comenzamos por la suma.

$$ {}_{i|}(\mathbf{A}+\mathbf{B})_{|j} = {}_{i|}\Big((\mathbf{A}+\mathbf{B})_{|j}\Big) = {}_{i|}\Big(\mathbf{A}_{|j}+\mathbf{B}_{|j}\Big) = {}_{i|}\big(\mathbf{A}_{|j}\big) + {}_{i|}\big(\mathbf{B}_{|j}\big) = {}_{i|}\mathbf{A}_{|j} + {}_{i|}\mathbf{B}_{|j}. $$

En el primer paso ponemos los paréntesis (la selección de componentes es asociativa, así que podemos seleccionar primero la columna); el segundo es la definición de suma de matrices; el tercero, que el selector de componentes es distributivo respecto a la suma de vectores (aquí con el selector por la izquierda: Sección 1.4.4); y en el cuarto quitamos los paréntesis.

Y ahora el producto.

$$ {}_{i|}(\lambda\mathbf{A})_{|j} = {}_{i|}\Big((\lambda\mathbf{A})_{|j}\Big) = {}_{i|}\Big(\lambda(\mathbf{A}_{|j})\Big) = \lambda\Big({}_{i|}\big(\mathbf{A}_{|j}\big)\Big) = \lambda\Big({}_{i|}\mathbf{A}_{|j}\Big). $$

Los pasos son los mismos: ponemos los paréntesis; aplicamos la definición de producto de una matriz por un escalar; sacamos el escalar, porque el selector de componentes es asociativo respecto al producto de un vector por un escalar (aquí con el selector por la izquierda del vector: es la segunda de las reglas de reescritura de la Página 17); y quitamos los paréntesis de la selección.

□

**Ejercicio 8.** Cada cadena es la cadena de la Estrategia 2 del Ejercicio 4 con el mismo número, con tres sustituciones: la matriz $\mathbf{A}$ en lugar del vector $\boldsymbol{a}$; el selector de columnas "$|\,j$" en lugar del selector de componentes "$|\,i$"; y, en cada paso en el que allí se aplicaba una propiedad de los números reales, una propiedad de los vectores. En azul aparecen las operaciones entre vectores; para pasar de una expresión en azul a la siguiente se aplica la propiedad de los vectores que lleva el mismo número que la propiedad que se está demostrando. Las cadenas 1 a 5, 7 y 8 coinciden con las de los vectores símbolo a símbolo; la 6 es más corta.

1. $(\mathbf{A}+\mathbf{B})_{|j} = \mathbf{A}_{|j} + \mathbf{B}_{|j} = \mathbf{B}_{|j} + \mathbf{A}_{|j} = (\mathbf{B}+\mathbf{A})_{|j}$.
2. $\Big(\mathbf{A}+(\mathbf{B}+\mathbf{C})\Big)_{|j} = \mathbf{A}_{|j} + \big(\mathbf{B}+\mathbf{C}\big)_{|j} = \mathbf{A}_{|j} + (\mathbf{B}_{|j}+\mathbf{C}_{|j}) = (\mathbf{A}_{|j}+\mathbf{B}_{|j})+\mathbf{C}_{|j} = \big(\mathbf{A}+\mathbf{B}\big)_{|j} + \mathbf{C}_{|j} = \Big((\mathbf{A}+\mathbf{B})+\mathbf{C}\Big)_{|j}$.
3. $\big(\mathbf{A}+\mathbf{0}\big)_{|j} = \mathbf{A}_{|j} + \mathbf{0}_{|j} = \mathbf{A}_{|j} + 0 = \mathbf{A}_{|j}$.

   (El segundo paso es la definición de matriz nula: todas sus columnas son vectores nulos, $\mathbf{0}_{|j} = \mathbf{0}$.)
4. $\Big(\mathbf{A}+\big(-\mathbf{A}\big)\Big)_{|j} = \mathbf{A}_{|j} + \big(-\mathbf{A}\big)_{|j} = \mathbf{A}_{|j} + \big(-(\mathbf{A}_{|j})\big) = \mathbf{0} = \mathbf{0}_{|j}$.

   (El segundo paso es la definición de matriz opuesta: sus columnas son las opuestas de las de $\mathbf{A}$, $\big(-\mathbf{A}\big)_{|j} = -\big(\mathbf{A}_{|j}\big)$. El último es, de nuevo, la definición de matriz nula: el vector nulo es la columna $j$-ésima de $\mathbf{0}$.)
5. $\Big(\lambda(\mathbf{A}+\mathbf{B})\Big)_{|j} = \lambda\Big((\mathbf{A}+\mathbf{B})_{|j}\Big) = \lambda(\mathbf{A}_{|j}+\mathbf{B}_{|j}) = \lambda(\mathbf{A}_{|j}) + \lambda(\mathbf{B}_{|j}) = \big(\lambda\mathbf{A}\big)_{|j} + \big(\lambda\mathbf{B}\big)_{|j} = \big(\lambda\mathbf{A}+\lambda\mathbf{B}\big)_{|j}$.
6. $\Big((\lambda+\varphi)\mathbf{A}\Big)_{|j} = (\lambda+\varphi)(\mathbf{A}_{|j}) = \lambda(\mathbf{A}_{|j}) + \varphi(\mathbf{A}_{|j}) = \big(\lambda\mathbf{A}\big)_{|j} + \big(\varphi\mathbf{A}\big)_{|j} = \big(\lambda\mathbf{A}+\varphi\mathbf{A}\big)_{|j}$.

   (Es la única cadena que no coincide símbolo a símbolo con la de los vectores. Allí hicieron falta tres pasos con números reales para pasar de $(\lambda+\varphi)a_i$ a $\lambda a_i + \varphi a_i$: conmutativa del producto, distributiva y conmutativa del producto. Aquí basta uno, porque la propiedad 6 de los vectores ya está escrita con la suma de escalares a la izquierda.)
7. $\Big(\lambda\big(\varphi\mathbf{A}\big)\Big)_{|j} = \lambda\Big(\big(\varphi\mathbf{A}\big)_{|j}\Big) = \lambda\Big(\varphi\big(\mathbf{A}_{|j}\big)\Big) = (\lambda\varphi)(\mathbf{A}_{|j}) = \big((\lambda\varphi)\mathbf{A}\big)_{|j}$.
8. $\big(1\mathbf{A}\big)_{|j} = 1(\mathbf{A}_{|j}) = \mathbf{A}_{|j}$.

Cada cadena empieza en la columna $j$-ésima de la matriz del lado izquierdo de la igualdad y termina en la columna $j$-ésima de la matriz del lado derecho. Como esto vale para cualquier columna $j$, en los ocho casos las dos matrices son iguales columna a columna; es decir, son iguales.

□

**Ejercicio 9.** Comenzamos con la suma. La componente $j$-ésima de la fila $i$-ésima de $\mathbf{A}+\mathbf{B}$ es

$$\Big({}_{i|}\big(\mathbf{A}+\mathbf{B}\big)\Big)_{|j} = {}_{i|}\Big(\big(\mathbf{A}+\mathbf{B}\big)_{|j}\Big) = {}_{i|}\Big(\mathbf{A}_{|j}+\mathbf{B}_{|j}\Big) = {}_{i|}\mathbf{A}_{|j} + {}_{i|}\mathbf{B}_{|j} = \Big({}_{i|}\mathbf{A} + {}_{i|}\mathbf{B}\Big)_{|j}.$$

En el primer paso cambiamos el orden de selección (la selección de componentes es asociativa, así que da igual seleccionar primero la fila o la columna); el segundo es la definición de suma de matrices; en el tercero distribuimos el selector de la izquierda, que es distributivo respecto a la suma de vectores (Sección 1.4.4); y en el cuarto sacamos el selector de la derecha como factor común, por la definición de suma de vectores. Como esto vale para cualquier componente $j$, los vectores ${}_{i|}\big(\mathbf{A}+\mathbf{B}\big)$ y ${}_{i|}\mathbf{A}+{}_{i|}\mathbf{B}$ son iguales.

Y continuamos con el producto, con los mismos pasos:

$$\Big({}_{i|}\big(\lambda\mathbf{A}\big)\Big)_{|j} = {}_{i|}\Big(\big(\lambda\mathbf{A}\big)_{|j}\Big) = {}_{i|}\Big(\lambda\big(\mathbf{A}_{|j}\big)\Big) = \lambda\Big({}_{i|}\mathbf{A}_{|j}\Big) = \Big(\lambda\big({}_{i|}\mathbf{A}\big)\Big)_{|j}.$$

(El segundo paso es ahora la definición de producto de una matriz por un escalar; en el tercero sacamos el escalar, porque el selector de la izquierda es asociativo respecto al producto por escalares; y el cuarto es la definición de producto de un vector por un escalar.) Como esto vale para cualquier componente $j$, los vectores ${}_{i|}\big(\lambda\mathbf{A}\big)$ y $\lambda\big({}_{i|}\mathbf{A}\big)$ son iguales.

□

**Ejercicio 10(a)** $\Big(\big(\lambda\mathbf{A}\big)^{\mathsf{T}}\Big)_{|j} = {}_{j|}\big(\lambda\mathbf{A}\big) = \lambda\Big({}_{j|}\mathbf{A}\Big) = \lambda\Big(\big(\mathbf{A}^{\mathsf{T}}\big)_{|j}\Big) = \Big(\lambda\big(\mathbf{A}^{\mathsf{T}}\big)\Big)_{|j}.$

El primer paso es la definición de transposición (la columna $j$-ésima de la transpuesta es la fila $j$-ésima); el segundo, que el selector de filas es asociativo respecto al producto por escalares (Ejercicio 9); el tercero, otra vez la definición de transposición; y el cuarto, la definición de producto de una matriz por un escalar. Como esto vale para cualquier columna $j$, las matrices $\big(\lambda\mathbf{A}\big)^{\mathsf{T}}$ y $\lambda(\mathbf{A}^{\mathsf{T}})$ son iguales.

□

**Ejercicio 10(b)** $\big((\mathbf{A}+\mathbf{B})^{\mathsf{T}}\big)_{|j} = {}_{j|}(\mathbf{A}+\mathbf{B}) = {}_{j|}\mathbf{A}+{}_{j|}\mathbf{B} = (\mathbf{A}^{\mathsf{T}})_{|j}+(\mathbf{B}^{\mathsf{T}})_{|j} = \big(\mathbf{A}^{\mathsf{T}}+\mathbf{B}^{\mathsf{T}}\big)_{|j}.$

Los pasos son los mismos: la definición de transposición; que el selector de filas es distributivo respecto a la suma (Ejercicio 9); otra vez la definición de transposición, en cada sumando; y la definición de suma de matrices. Como esto vale para cualquier columna $j$, las matrices $\big(\mathbf{A}+\mathbf{B}\big)^{\mathsf{T}}$ y $\mathbf{A}^{\mathsf{T}}+\mathbf{B}^{\mathsf{T}}$ son iguales.

□

**Ejercicio 11(a)** $\boldsymbol{x}\cdot\boldsymbol{y} \;=\; x_1y_1+\cdots+x_ny_n \;=\; y_1x_1+\cdots+y_nx_n \;=\; \boldsymbol{y}\cdot\boldsymbol{x}$

□

**Ejercicio 11(b)**

- Primera igualdad:

$$\begin{aligned}
(\lambda\boldsymbol{x})\cdot\boldsymbol{y} &= \big(\lambda\boldsymbol{x}\big)_{|1}y_1+\cdots+\big(\lambda\boldsymbol{x}\big)_{|n}y_n && \text{definición de producto punto}\\
&= (\lambda x_1)y_1+\cdots+(\lambda x_n)y_n && \text{el selector es asociativo respecto al producto por escalares}\\
&= \lambda(x_1y_1)+\cdots+\lambda(x_ny_n) && \text{propiedad asociativa del producto}\\
&= \lambda\big(x_1y_1+\cdots+x_ny_n\big) && \text{propiedad distributiva}\\
&= \lambda(\boldsymbol{x}\cdot\boldsymbol{y}) && \text{definición de producto punto.}
\end{aligned}$$

- Segunda igualdad:

$$\begin{aligned}
(\boldsymbol{x}+\boldsymbol{y})\cdot\boldsymbol{z} &= \big(\boldsymbol{x}+\boldsymbol{y}\big)_{|1}z_1+\cdots+\big(\boldsymbol{x}+\boldsymbol{y}\big)_{|n}z_n && \text{definición de producto punto}\\
&= (x_1+y_1)z_1+\cdots+(x_n+y_n)z_n && \text{el selector es distributivo respecto a la suma}\\
&= z_1(x_1+y_1)+\cdots+z_n(x_n+y_n) && \text{propiedad conmutativa del producto}\\
&= (z_1x_1+z_1y_1)+\cdots+(z_nx_n+z_ny_n) && \text{propiedad distributiva}\\
&= (x_1z_1+y_1z_1)+\cdots+(x_nz_n+y_nz_n) && \text{propiedad conmutativa del producto}\\
&= (x_1z_1+\cdots+x_nz_n)+(y_1z_1+\cdots+y_nz_n) && \text{asociativa y conmutativa de la suma}\\
&= \boldsymbol{x}\cdot\boldsymbol{z}+\boldsymbol{y}\cdot\boldsymbol{z} && \text{definición de producto punto.}
\end{aligned}$$

□

**Ejercicio 11(c)** $\boldsymbol{x}\cdot\boldsymbol{x} = x_1^2+\cdots+x_n^2 \geq 0.$

Esta solución y la siguiente usan dos propiedades de los números reales que seguro conoce, aunque no estén entre las ocho del recordatorio de la Página 8: el cuadrado de un número real no es negativo; y una suma de números no negativos es cero solo si todos los sumandos son cero.

□

**Ejercicio 11(d)** $\boldsymbol{x}\cdot\boldsymbol{x} = x_1^2+\cdots+x_n^2 = 0 \iff x_i = 0$ para $i = 1:n$; es decir, si y solo si $\boldsymbol{x}=\mathbf{0}$.

□

**Ejercicio 12.**

$$2\begin{pmatrix}1\\1\\-2\end{pmatrix}+2\begin{pmatrix}1\\-2\\1\end{pmatrix}+2\begin{pmatrix}-2\\1\\1\end{pmatrix}=\begin{pmatrix}2\\2\\-4\end{pmatrix}+\begin{pmatrix}2\\-4\\2\end{pmatrix}+\begin{pmatrix}-4\\2\\2\end{pmatrix}=\begin{pmatrix}2+2-4\\2-4+2\\-4+2+2\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}=\mathbf{0}.$$

□

**Ejercicio 13.** $\mathbf{A}\boldsymbol{b} = \begin{bmatrix}1&1&-2\\1&-2&1\\-2&1&1\end{bmatrix}\begin{pmatrix}2\\2\\2\end{pmatrix} = 2\begin{pmatrix}1\\1\\-2\end{pmatrix}+2\begin{pmatrix}1\\-2\\1\end{pmatrix}+2\begin{pmatrix}-2\\1\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$

□

**Ejercicio 14.** $\mathbf{A}\big(\mathbf{I}_{|j}\big) = \mathbf{A}\begin{pmatrix}0\\\vdots\\1\\\vdots\\0\end{pmatrix} = \big(\mathbf{A}_{|1}\big)0+\cdots+\big(\mathbf{A}_{|j}\big)1+\cdots+\big(\mathbf{A}_{|n}\big)0 = \mathbf{A}_{|j}.$

□

**Ejercicio 15.** Sea $\mathbf{A}$, de $n$ columnas y $\boldsymbol{b}$ de $\mathbb{R}^n$, entonces

$$
\begin{aligned}
{}_{i|}(\mathbf{A}\boldsymbol{b}) &= {}_{i|}\Big((\mathbf{A}_{|1})b_1 + \cdots + (\mathbf{A}_{|n})b_n\Big) && \text{Definición 2.3}\\
&= {}_{i|}\Big((\mathbf{A}_{|1})b_1\Big) + \cdots + {}_{i|}\Big((\mathbf{A}_{|n})b_n\Big) && \text{el selector es distributivo respecto a la suma}\\
&= \Big({}_{i|}(\mathbf{A}_{|1})\Big)b_1 + \cdots + \Big({}_{i|}(\mathbf{A}_{|n})\Big)b_n && \text{se desplazan los paréntesis (escalar a la derecha)}\\
&= \Big({}_{i|}(\mathbf{A})_{|1}\Big)b_1 + \cdots + \Big({}_{i|}(\mathbf{A})_{|n}\Big)b_n && \text{la selección de componentes es asociativa}\\
&= {}_{i|}(\mathbf{A}) \cdot \boldsymbol{b} && \text{definición de producto punto.}
\end{aligned}
$$

□

**Ejercicio 16(a)** ${}_{i|}(\mathbf{A}(\boldsymbol{b}+\boldsymbol{c})) = {}_{i|}(\mathbf{A}) \cdot (\boldsymbol{b}+\boldsymbol{c}) = {}_{i|}(\mathbf{A}) \cdot \boldsymbol{b} + {}_{i|}(\mathbf{A}) \cdot \boldsymbol{c} = {}_{i|}(\mathbf{A}\boldsymbol{b}) + {}_{i|}(\mathbf{A}\boldsymbol{c}) = {}_{i|}(\mathbf{A}\boldsymbol{b} + \mathbf{A}\boldsymbol{c})$.

□

**Ejercicio 16(b)** ${}_{i|}(\mathbf{A}(\lambda\boldsymbol{b})) = {}_{i|}(\mathbf{A}) \cdot (\lambda\boldsymbol{b}) = \lambda\Big({}_{i|}(\mathbf{A}) \cdot \boldsymbol{b}\Big) = \lambda\Big({}_{i|}(\mathbf{A}\boldsymbol{b})\Big) = {}_{i|}(\lambda(\mathbf{A}\boldsymbol{b}))$.

□

**Ejercicio 17.**

$$
\begin{aligned}
\mathbf{A}(\mathbf{B}\boldsymbol{c}) =& \mathbf{A}\Big((\mathbf{B}_{|1})c_1 + \cdots + (\mathbf{B}_{|n})c_n\Big) && \text{por la Definición 2.3 de producto}\\
=& \mathbf{A}\Big((\mathbf{B}_{|1})c_1\Big) + \cdots + \mathbf{A}\Big((\mathbf{B}_{|n})c_n\Big) && \text{Proposición 2.2.1}\\
=& \Big(\mathbf{A}(\mathbf{B}_{|1})\Big)c_1 + \cdots + \Big(\mathbf{A}(\mathbf{B}_{|n})\Big)c_n && \text{Proposición 2.2.2 con el escalar a la derecha}\\
=& \Big[\mathbf{A}(\mathbf{B}_{|1}); \quad \ldots \quad \mathbf{A}(\mathbf{B}_{|n});\Big]\,\boldsymbol{c} && \text{por la Definición 2.3 de producto.}
\end{aligned}
$$

La Proposición 2.2.1 está enunciada para dos sumandos. Aquí se aplica varias veces, hasta separar los $n$ sumandos.

□

**Ejercicio 18(a)** ${}_{i|}(\mathbf{A}(\lambda\boldsymbol{b})) = {}_{i|}(\mathbf{A}) \cdot (\lambda\boldsymbol{b}) = \lambda\Big({}_{i|}(\mathbf{A}) \cdot \boldsymbol{b}\Big) = \Big(\lambda({}_{i|}(\mathbf{A}))\Big) \cdot \boldsymbol{b} = \Big({}_{i|}(\lambda\mathbf{A})\Big) \cdot \boldsymbol{b} = {}_{i|}((\lambda\mathbf{A})\boldsymbol{b})$

□

**Ejercicio 18(b)**
${}_{i|}((\mathbf{A}+\mathbf{B})\boldsymbol{c}) = {}_{i|}(\mathbf{A}+\mathbf{B}) \cdot \boldsymbol{c} = {}_{i|}\mathbf{A} + {}_{i|}\mathbf{B} \cdot \boldsymbol{c} = {}_{i|}(\mathbf{A}) \cdot \boldsymbol{c} + {}_{i|}(\mathbf{B}) \cdot \boldsymbol{c} = {}_{i|}(\mathbf{A}\boldsymbol{c}) + {}_{i|}(\mathbf{B}\boldsymbol{c}) = {}_{i|}(\mathbf{A}\boldsymbol{c} + \mathbf{B}\boldsymbol{c})$.

□

**Ejercicio 19(a)** $\boldsymbol{a}\mathbf{I} = (\mathbf{I}^\intercal)\boldsymbol{a} = \mathbf{I}\boldsymbol{a} = \boldsymbol{a}$.

□

**Ejercicio 19(b)** ${}_{i|}(\mathbf{I})\mathbf{A} = (\mathbf{A}^\intercal)({}_{i|}\mathbf{I}) = (\mathbf{A}^\intercal)(\mathbf{I}_{|i}) = (\mathbf{A}^\intercal)_{|i} = {}_{i|}\mathbf{A}$.

□

**Ejercicio 19(c)** $(\boldsymbol{a}+\boldsymbol{b})\mathbf{C} = (\mathbf{C}^\intercal)(\boldsymbol{a}+\boldsymbol{b}) \underset{(*)}{=} (\mathbf{C}^\intercal)\boldsymbol{a} + (\mathbf{C}^\intercal)\boldsymbol{b} = \boldsymbol{a}\mathbf{C} + \boldsymbol{b}\mathbf{C}$.

□

**Ejercicio 19(d)** $(\lambda \boldsymbol{a})\mathbf{B} = (\mathbf{B}^\intercal)(\lambda \boldsymbol{a}) \underset{(*)}{=} \lambda\Big((\mathbf{B}^\intercal)\boldsymbol{a}\Big) = \lambda(\boldsymbol{a}\mathbf{B})$

□

**Ejercicio 19(e)** $\boldsymbol{a}(\lambda\mathbf{B}) = \big((\lambda\mathbf{B})^\intercal\big)\boldsymbol{a} = \Big(\lambda(\mathbf{B}^\intercal)\Big)\boldsymbol{a} \underset{(*)}{=} (\mathbf{B}^\intercal)(\lambda\boldsymbol{a}) = (\lambda\boldsymbol{a})\mathbf{B}.$

□

**Ejercicio 19(f)** $\boldsymbol{a}(\mathbf{B}+\mathbf{C}) = (\mathbf{B}+\mathbf{C})^\intercal\boldsymbol{a} = \Big((\mathbf{B}^\intercal)+(\mathbf{C}^\intercal)\Big)\boldsymbol{a} \underset{(*)}{=} (\mathbf{B}^\intercal)\boldsymbol{a} + (\mathbf{C}^\intercal)\boldsymbol{a} = \boldsymbol{a}\mathbf{B} + \boldsymbol{a}\mathbf{C}.$

□

**Ejercicio 20(a)** Recordando el Corolario 2.2.3 en la página 31 (*) y usando la Definición 3.1 de producto:

$$\mathbf{A}(\mathbf{B}\boldsymbol{c}) \underset{(*)}{=} \Big[\mathbf{A}\big(\mathbf{B}_{|1}\big), \quad \ldots, \quad \mathbf{A}\big(\mathbf{B}_{|n}\big)\Big]\,\boldsymbol{c} = (\mathbf{AB})\boldsymbol{c}.$$

□

**Ejercicio 20(b)** Aplicando repetidamente la definición de producto y una vez $(*)$ la proposición anterior tenemos:

$$\Big(\mathbf{A}\big(\mathbf{BC}\big)\Big)_{|j} = \mathbf{A}\Big(\big(\mathbf{BC}\big)_{|j}\Big) = \mathbf{A}\Big(\mathbf{B}\big(\mathbf{C}_{|j}\big)\Big) \underset{(*)}{=} \Big(\mathbf{AB}\Big)\big(\mathbf{C}_{|j}\big) = \Big(\big(\mathbf{AB}\big)\mathbf{C}\Big)_{|j}, \quad j = 1:n.$$

□

**Ejercicio 21(a)** Usando la Definición 3.1 de producto matricial y recordando que $\big(\mathbf{A}+\mathbf{B}\big)\boldsymbol{c} = \mathbf{A}\boldsymbol{c} + \mathbf{B}\boldsymbol{c}$; para $j = 1:n$ tenemos

$$\Big(\big(\mathbf{A}+\mathbf{B}\big)\mathbf{C}\Big)_{|j} = \big(\mathbf{A}+\mathbf{B}\big)\big(\mathbf{C}_{|j}\big) = \mathbf{A}\big(\mathbf{C}_{|j}\big) + \mathbf{B}\big(\mathbf{C}_{|j}\big) = \Big(\mathbf{AC}\Big)_{|j} + \Big(\mathbf{BC}\Big)_{|j} = \Big(\mathbf{AC}+\mathbf{BC}\Big)_{|j}.$$

□

**Ejercicio 21(b)** Usando la Definición 3.1 de producto matricial y la propiedad de linealidad: $\mathbf{A}(\boldsymbol{u}+\boldsymbol{v}) = \mathbf{A}\boldsymbol{u} + \mathbf{A}\boldsymbol{v}$; para $j = 1:n$ tenemos

$$\Big(\mathbf{A}\big(\mathbf{B}+\mathbf{C}\big)\Big)_{|j} = \mathbf{A}\Big(\big(\mathbf{B}+\mathbf{C}\big)_{|j}\Big) = \mathbf{A}\Big(\big(\mathbf{B}_{|j}\big)+\big(\mathbf{C}_{|j}\big)\Big) = \mathbf{A}\big(\mathbf{B}_{|j}\big)+\mathbf{A}\big(\mathbf{C}_{|j}\big) = \big(\mathbf{AB}\big)_{|j}+\big(\mathbf{AC}\big)_{|j} = \Big(\mathbf{AB}+\mathbf{AC}\Big)_{|j}.$$

□

**Ejercicio 21(c)**

$$\Big(\mathbf{A}(\lambda\mathbf{B})\Big)_{|j} = \mathbf{A}\Big(\big(\lambda\mathbf{B}\big)_{|j}\Big) = \mathbf{A}\Big(\lambda\big(\mathbf{B}_{|j}\big)\Big) \begin{cases} = \big(\lambda\mathbf{A}\big)\big(\mathbf{B}_{|j}\big) = \Big(\big(\lambda\mathbf{A}\big)\mathbf{B}\Big)_{|j} \\ \\ = \lambda\Big(\mathbf{A}\big(\mathbf{B}_{|j}\big)\Big) = \lambda\Big(\big(\mathbf{AB}\big)_{|j}\Big) = \Big(\lambda\big(\mathbf{AB}\big)\Big)_{|j} \end{cases}, \quad j = 1:n.$$

□

**Ejercicio 21(d)** $\quad \big(\mathbf{IA}\big)_{|j} = \mathbf{I}\big(\mathbf{A}_{|j}\big) = \mathbf{A}_{|j}, \quad j = 1:n.$

□

**Ejercicio 21(e)** $\quad \big(\mathbf{AI}\big)_{|j} = \mathbf{A}\big(\mathbf{I}_{|j}\big) = \mathbf{A}_{|j}, \quad j = 1:n.$

□

**Ejercicio 22.** $\quad {}_{i|}\big(\mathbf{AB}\big)_{|j} = {}_{i|}\Big(\big(\mathbf{AB}\big)_{|j}\Big) = {}_{i|}\Big(\mathbf{A}\big(\mathbf{B}_{|j}\big)\Big) = \big({}_{i|}\mathbf{A}\big)\cdot\big(\mathbf{B}_{|j}\big)$

□

**Ejercicio 23.** Veamos que las componentes $j$-ésimas son iguales:

$$\Big({}_{i|}(\mathbf{AB})\Big)_{|j} = {}_{i|}\big(\mathbf{AB}\big)_{|j} = \big({}_{i|}\mathbf{A}\big)\cdot\big(\mathbf{B}_{|j}\big) = \Big(\big({}_{i|}\mathbf{A}\big)\mathbf{B}\Big)_{|j}; \quad \text{por tanto } {}_{i|}\big(\mathbf{AB}\big) = \big({}_{i|}\mathbf{A}\big)\mathbf{B}.$$

□

**Ejercicio 24.** $\quad (\mathbf{AB})^{\intercal}{}_{|j} \;=\; {}_{j|}\mathbf{AB} \;=\; (\mathbf{B}^{\intercal})\big({}_{j|}\mathbf{A}\big) \;=\; (\mathbf{B}^{\intercal})(\mathbf{A}^{\intercal})_{|j}; \quad$ por tanto, $\big(\mathbf{AB}\big)^{\intercal} = (\mathbf{B}^{\intercal})\ (\mathbf{A}^{\intercal}).$

□

**Ejercicio 25(a)** Sean las matrices $\underset{m\times p}{\mathbf{A}}$ y $\underset{p\times n}{\mathbf{B}}$ tales que ${}_{i|}\mathbf{A}_{|j} = 0$ cuando $i > j$ y ${}_{i|}\mathbf{B}_{|j} = 0$ cuando $i > j$. Tenemos que demostrar que ${}_{i|}\mathbf{AB}_{|j} = 0$ cuando $\boxed{i > j}$:

$${}_{i|}\mathbf{AB}_{|j} \;=\; \sum_{k=1}^{p}\Big({}_{i|}\mathbf{A}_{|k}\Big)\Big({}_{k|}\mathbf{B}_{|j}\Big) \;=\; \sum_{k=1}^{j}\underbrace{\Big({}_{i|}\mathbf{A}_{|k}\Big)}_{0}\Big({}_{k|}\mathbf{B}_{|j}\Big) \;+\; \sum_{k=j+1}^{p}\Big({}_{i|}\mathbf{A}_{|k}\Big)\underbrace{\Big({}_{k|}\mathbf{B}_{|j}\Big)}_{0} \;=\; 0.$$

□

**Ejercicio 25(b)** Sean las matrices $\underset{m\times p}{\mathbf{A}}$ y $\underset{p\times n}{\mathbf{B}}$ tales que por una parte ${}_{i|}\mathbf{A}_{|j} = 0$ cuando $i > j$, y ${}_{j|}\mathbf{A}_{|j} = 1$, y por otra ${}_{i|}\mathbf{B}_{|j} = 0$ cuando $i > j$. Por la proposición anterior, ya sabemos que $\mathbf{AB}$ es triangular superior, así que nos basta demostrar que ${}_{i|}\mathbf{AB}_{|j} = {}_{j|}\mathbf{B}_{|j}$ cuando $\boxed{i = j}$:

$${}_{i|}\mathbf{AB}_{|j} \;=\; \sum_{k=1}^{p}\Big({}_{i|}\mathbf{A}_{|k}\Big)\Big({}_{k|}\mathbf{B}_{|j}\Big) \;=\; \sum_{k=1}^{j-1}\underbrace{\Big({}_{i|}\mathbf{A}_{|k}\Big)}_{0}\Big({}_{k|}\mathbf{B}_{|j}\Big) + \underbrace{\Big({}_{j|}\mathbf{A}_{|j}\Big)}_{1}\Big({}_{j|}\mathbf{B}_{|j}\Big) + \sum_{k=j+1}^{p}\Big({}_{i|}\mathbf{A}_{|k}\Big)\underbrace{\Big({}_{k|}\mathbf{B}_{|j}\Big)}_{0} \;=\; {}_{j|}\mathbf{B}_{|j}.$$

□

**Ejercicio 25(c)** Sean las matrices $\underset{m\times p}{\mathbf{A}}$ y $\underset{p\times n}{\mathbf{B}}$ tales que por una parte ${}_{i|}\mathbf{A}_{|j} = 0$ cuando $i > j$, y por otra ${}_{i|}\mathbf{B}_{|j} = 0$ cuando $i > j$ y ${}_{j|}\mathbf{B}_{|j} = 1$. Como antes, nos basta demostrar que ${}_{i|}\mathbf{AB}_{|j} = {}_{j|}\mathbf{A}_{|j}$ cuando $\boxed{i = j}$:

$${}_{i|}\mathbf{AB}_{|j} \;=\; \sum_{k=1}^{p}\Big({}_{i|}\mathbf{A}_{|k}\Big)\Big({}_{k|}\mathbf{B}_{|j}\Big) \;=\; \sum_{k=1}^{j-1}\underbrace{\Big({}_{i|}\mathbf{A}_{|k}\Big)}_{0}\Big({}_{k|}\mathbf{B}_{|j}\Big) + \Big({}_{j|}\mathbf{A}_{|j}\Big)\underbrace{\Big({}_{j|}\mathbf{B}_{|j}\Big)}_{1} + \sum_{k=j+1}^{p}\Big({}_{i|}\mathbf{A}_{|k}\Big)\underbrace{\Big({}_{k|}\mathbf{B}_{|j}\Big)}_{0} \;=\; {}_{j|}\mathbf{B}_{|j}.$$

□

**Ejercicio 26(a)**

- Sea $a$ un escalar, entonces $\operatorname{tr}(a\mathbf{A}) = \sum_{i=j}^{k} {}_{j|}(a\mathbf{A})_{|j} = \sum_{j=1}^{k} a({}_{j|}\mathbf{A}_{|j}) = a\sum_{j=1}^{k} {}_{j|}\mathbf{A}_{|j} = a\operatorname{tr}(\mathbf{A})$.
- Sean $\mathbf{A}$ y $\mathbf{B}$ del mismo orden: $\operatorname{tr}(\mathbf{A}+\mathbf{B}) = \sum_{j=1}^{k} {}_{j|}(\mathbf{A}+\mathbf{B})_{|j} = \sum_{j=1}^{k} {}_{j|}\mathbf{A}_{|j} + \sum_{j=1}^{k} {}_{j|}\mathbf{B}_{|j} = \operatorname{tr}(\mathbf{A}) + \operatorname{tr}(\mathbf{B})$,

□

**Ejercicio 26(b)** $\operatorname{tr}(\mathbf{A}) = \sum_{j=1}^{k} {}_{j|}\mathbf{A}_{|j} = \sum_{j=1}^{k} {}_{j|}(\mathbf{A}^{\intercal})_{|j} = \operatorname{tr}(\mathbf{A}^{\intercal})$.

□

**Ejercicio 26(c)** Recordando la Ecuación 3.6 en la página 43 respecto al producto de matrices; que para cualesquiera vectores $\boldsymbol{a}$ y $\boldsymbol{b}$, se verifica que $\boldsymbol{a}\cdot\boldsymbol{b} = \boldsymbol{b}\cdot\boldsymbol{a}$ y que son iguales los vectores ${}_{(p,)|}\mathbf{A}_{|j} = {}_{p|}\mathbf{A}_{|(j,)}$:

$$\begin{aligned}\operatorname{tr}(\mathbf{A}^{\intercal}\mathbf{A}) &= \sum_{j=1}^{n} {}_{j|}(\mathbf{A}^{\intercal}\mathbf{A})_{|j} = \sum_{j=1}^{n} {}_{j|}\left(\sum_{p=1}^{m}\left((\mathbf{A}^{\intercal})_{|(p,)}\right)\left({}_{(p,)|}\mathbf{A}\right)\right)_{|j} = \sum_{j=1}^{n}\sum_{p=1}^{m}\left({}_{j|}(\mathbf{A}^{\intercal})_{|(p,)}\right)\cdot\left({}_{(p,)|}\mathbf{A}_{|j}\right)\\ &= \sum_{j=1}^{n}\sum_{p=1}^{m}\left({}_{(p,)|}\mathbf{A}_{|j}\right)\cdot\left({}_{j|}(\mathbf{A}^{\intercal})_{|(p,)}\right) = \sum_{p=1}^{m}\sum_{j=1}^{n} {}_{p|}\left(\mathbf{A}_{|(j,)}\right)\left({}_{(j,)|}(\mathbf{A}^{\intercal})\right)_{|p} = \sum_{p=1}^{m} {}_{p|}(\mathbf{A}\mathbf{A}^{\intercal})_{|p} = \operatorname{tr}(\mathbf{A}\mathbf{A}^{\intercal}).\end{aligned}$$

□

**Ejercicio 26(d)** Dado que $\operatorname{tr}(\mathbf{A}) = \operatorname{tr}(\mathbf{A}^{\intercal})$, basta con demostrar que $\operatorname{tr}(\mathbf{A}^{\intercal}\mathbf{B}) = \operatorname{tr}(\mathbf{B}\mathbf{A}^{\intercal})$. Repitiendo los pasos del ejercicio anterior:

$$\begin{aligned}\operatorname{tr}(\mathbf{A}^{\intercal}\mathbf{B}) &= \sum_{j=1}^{n} {}_{j|}(\mathbf{A}^{\intercal}\mathbf{B})_{|j} = \sum_{j=1}^{n}\sum_{p=1}^{m}\left({}_{j|}(\mathbf{A}^{\intercal})_{|(p,)}\right)\cdot\left({}_{(p,)|}\mathbf{B}_{|j}\right) = \sum_{j=1}^{n}\sum_{p=1}^{m}\left({}_{(p,)|}\mathbf{B}_{|j}\right)\cdot\left({}_{j|}(\mathbf{A}^{\intercal})_{|(p,)}\right)\\ &= \sum_{p=1}^{m}\sum_{j=1}^{n} {}_{p|}\left(\mathbf{B}_{|(j,)}\right)\left({}_{(j,)|}(\mathbf{A}^{\intercal})\right)_{|p} = \sum_{p=1}^{m} {}_{p|}\left(\sum_{j=1}^{n}\left(\mathbf{B}_{|(j,)}\right)\left({}_{(j,)|}(\mathbf{A}^{\intercal})\right)\right)_{|p} = \sum_{p=1}^{m} {}_{p|}(\mathbf{B}\mathbf{A}^{\intercal})_{|p} = \operatorname{tr}(\mathbf{B}\mathbf{A}^{\intercal}).\end{aligned}$$

□

**Ejercicio 27(a)** $\begin{bmatrix}1 & 0 & -7 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$.

□

**Ejercicio 27(b)** $\begin{bmatrix}1 & 0 & 0 & 1\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$.

□

**Ejercicio 27(c)** $\begin{bmatrix}1 & 0 & 0 & 0\\ 3 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$.

□

**Ejercicio 27(d)** $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -10 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

□

**Ejercicio 28.** a) $\underset{[(-7)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}$ transforma $\mathbf{B}_{|3}$, restándole siete veces $\mathbf{B}_{|1}$, es decir: $\left(\mathbf{B}_{\underset{[(-7)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}}\right)_{|3} = -7\mathbf{B}_{|1} + \mathbf{B}_{|3}$.

b) $\underset{[(1)\mathbf{1}+\mathbf{4}]}{\boldsymbol{\tau}}$ transforma la cuarta columna de $\mathbf{B}$, sumándole la primera columna: $\left(\mathbf{B}_{\underset{[(1)\mathbf{1}+\mathbf{4}]}{\boldsymbol{\tau}}}\right)_{|4} = \mathbf{B}_{|1} + \mathbf{B}_{|4}$.

c) $\underset{[(3)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}$ transforma la primera columna de $\mathbf{B}$, sumándole tres veces la segunda: $\left(\mathbf{B}_{\underset{[(3)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}}\right)_{|1} = 3\mathbf{B}_{|2} + \mathbf{B}_{|1}$.

d) $\underset{[(-10)\mathbf{3}]}{\boldsymbol{\tau}}$ transforma la tercera columna de $\mathbf{B}$, multiplicándola por $-10$, es decir: $\left(\mathbf{B}_{\underset{[(-10)\mathbf{3}]}{\boldsymbol{\tau}}}\right)_{|3} = -10\mathbf{B}_{|3}$.

□

**Ejercicio 29(a)** $\mathbf{A}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}} = \mathbf{A}\left(\mathbf{I}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}}\right) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 3 & 2 & 3 & 4 \end{bmatrix}$.

□

**Ejercicio 29(b)** $\mathbf{A}_{\underset{[(-1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}} = \mathbf{A}\left(\mathbf{I}_{\underset{[(-1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}\right) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 3 & 4 \end{bmatrix}$.

□

**Ejercicio 30.** Por una parte,

$$(\lambda\mathbf{A})_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = (\lambda\mathbf{A})\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \lambda(\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)) = \lambda(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}),$$

y por otra,

$$(\mathbf{A}+\mathbf{B})_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = (\mathbf{A}+\mathbf{B})\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) + \mathbf{B}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} + \mathbf{B}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

□

**Ejercicio 31(a)** Una posible sucesión de transformaciones elementales por columnas es:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}]{\text{Tipo I}} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[\underset{[(-1)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}]{\text{Tipo I}} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}]{\text{Tipo I}} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[\underset{[(-1)\mathbf{1}]}{\boldsymbol{\tau}}]{\text{Tipo II}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

es decir, $\mathbf{I}_{\tau_{[\mathbf{1}\rightleftharpoons\mathbf{2}]}} = \mathbf{I}_{\tau_{[(1)\mathbf{1}+\mathbf{2}][(-1)\mathbf{2}+\mathbf{1}][(1)\mathbf{1}+\mathbf{2}][(-1)\mathbf{1}]}}$.

□

**Ejercicio 31(b)** Hay muchas combinaciones posibles. Por ejemplo: $\mathbf{I}_{\tau_{[i\rightleftharpoons j]}} = \mathbf{I}_{\tau_{[(-1)j][(-1)j+i][(1)i+j][(-1)j+i]}}$.

□

**Ejercicio 32(a)** Por una parte $\mathbf{ABB^{-1}A^{-1}} = \mathbf{AIA^{-1}} = \mathbf{I}$, y por otra $\mathbf{B^{-1}A^{-1}AB} = \mathbf{B^{-1}IB} = \mathbf{I}$.

□

**Ejercicio 32(b)** Basta aplicar iterativamente la proposición anterior:

$$\big(\mathbf{A}_1 \cdots \mathbf{A}_k\big)^{-1} = \big(\mathbf{A}_2 \cdots \mathbf{A}_k\big)^{-1}\mathbf{A}_1^{-1} = \big(\mathbf{A}_3 \cdots \mathbf{A}_k\big)^{-1}\mathbf{A}_2^{-1}\mathbf{A}_1^{-1} = \cdots = \big(\mathbf{A}_{k-1}\mathbf{A}_k\big)^{-1}\mathbf{A}_{k-2}^{-1}\cdots\mathbf{A}_1^{-1} = \mathbf{A}_k^{-1}\cdots\mathbf{A}_1^{-1}.$$

□

**Ejercicio 32(c)** Si $\mathbf{A}$ es invertible, entonces $\mathbf{AB}$ es invertible (Proposición 5.1.2 en la página 69).

Supongamos que $\mathbf{AB}$ es invertible. Entonces sabemos que existe $\mathbf{C}$ tal que

$$\big(\mathbf{AB}\big)\mathbf{C} = \mathbf{I}, \qquad \text{y} \qquad \mathbf{C}\big(\mathbf{AB}\big) = \mathbf{I}.$$

De la primera igualdad tenemos que $\mathbf{A}$ es invertible por la derecha, pues

$$\mathbf{A}\big(\mathbf{BC}\big) = \mathbf{I}.$$

Solo nos falta comprobar que también es invertible por la izquierda, es decir, que $\big(\mathbf{BC}\big)\mathbf{A} = \mathbf{I}$. Para ello, nos fijamos en la segunda igualdad y, multiplicando por $\mathbf{B}$ por la izquierda y por $\mathbf{B^{-1}}$ por la derecha (a ambos lados de la ecuación) tenemos

$$\mathbf{BC}\big(\mathbf{AB}\big)\mathbf{B^{-1}} = \mathbf{BIB^{-1}} \qquad \Rightarrow \qquad \big(\mathbf{BC}\big)\mathbf{A} = \mathbf{I}.$$

□

**Ejercicio 32(d)** Si $\mathbf{B}$ es invertible, entonces $\mathbf{AB}$ es invertible (Proposición 5.1.2 en la página 69).

Supongamos que $\mathbf{AB}$ es invertible. Entonces sabemos que existe $\mathbf{C}$ tal que

$$\big(\mathbf{AB}\big)\mathbf{C} = \mathbf{I}, \qquad \text{y} \qquad \mathbf{C}\big(\mathbf{AB}\big) = \mathbf{I}.$$

De la segunda igualdad, tenemos que $\mathbf{B}$ es invertible por la izquierda, pues

$$\big(\mathbf{CA}\big)\mathbf{B} = \mathbf{I}.$$

Solo nos falta comprobar que también es invertible por la derecha, es decir, que $\mathbf{B}\big(\mathbf{CA}\big) = \mathbf{I}$. Para ello, nos fijamos en la primera igualdad y, multiplicando por $\mathbf{A^{-1}}$ por la izquierda y por $\mathbf{A}$ por la derecha (a ambos lados de la ecuación) tenemos

$$\mathbf{A^{-1}}\big(\mathbf{AB}\big)\mathbf{CA} = \mathbf{A^{-1}IA}, \qquad \Rightarrow \qquad \big(\mathbf{BC}\big)\mathbf{A} = \mathbf{I}.$$

□

**Ejercicio 32(e)** Aplicando la pista, tenemos por un lado: $\mathbf{A}^\intercal(\mathbf{A}^{-1})^\intercal = \left(\mathbf{A}^{-1}\mathbf{A}\right)^\intercal = \mathbf{I}^\intercal = \mathbf{I}$ y por el otro $(\mathbf{A}^{-1})^\intercal\mathbf{A}^\intercal = \left(\mathbf{A}\mathbf{A}^{-1}\right)^\intercal = \mathbf{I}^\intercal = \mathbf{I}$. Es decir, $(\mathbf{A}^{-1})^\intercal = \left(\mathbf{A}^\intercal\right)^{-1}$.

□

**Ejercicio 32(f)** Sea $\mathbf{A}$ de orden $n$ una matriz cuya $j$-ésima columna es nula: $\mathbf{A}_{|j} = \mathbf{0}$. Si asumimos que existe $\mathbf{A}^{-1}$, llegamos a una contradicción. Para verlo, multipliquemos la ecuación por $\mathbf{A}^{-1}$:

$$\mathbf{A}^{-1}\left(\mathbf{A}_{|j}\right) = \mathbf{A}^{-1}(\mathbf{0}) \qquad \Rightarrow \qquad \left(\mathbf{A}^{-1}\mathbf{A}\right)_{|j} = \mathbf{0} \qquad \Rightarrow \qquad \mathbf{I}_{|j} = \mathbf{0}.$$

Pero sabemos que $\mathbf{I}_{|j}$ (la columna $j$-ésima de $\mathbf{I}$) no es $\mathbf{0}$, así que hemos llegado a una contradicción. Por lo tanto, $\mathbf{A}^{-1}$ no puede existir: así, $\mathbf{A}$ *es singular* (no tiene inversa) *si tiene alguna columna nula.*

Para las filas, basta repetir lo anterior con $\mathbf{A}^\intercal$, pues, si ${}_{i|}\mathbf{A} = (\mathbf{A}^\intercal)_{|i} = \mathbf{0}$, llegamos a la misma contradicción.

□

**Ejercicio 33.** Si para alguna columna $\mathbf{A}_{|j}$ se verifica que $\mathbf{A}_{|j} = \mathbf{A}\boldsymbol{x}$ (con $x_j = 0$), es decir, $\mathbf{A}_{|j}$ es una combinación lineal del resto de columnas, entonces mediante $(n-1)$ transformaciones elementales $\underset{[(-x_k)\boldsymbol{k}+\boldsymbol{j}]}{\boldsymbol{\tau}}$ (con $k = 1:n$ excepto $k = j$) es posible transformar la columna $\mathbf{A}_{|j}$ en un vector de ceros, es decir, hay transformaciones elementales tales que $\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_{(n-1)}}\right)_{|j} = \mathbf{0}$, pues a la columna $\mathbf{A}_{|j}$ se le resta $\mathbf{A}\boldsymbol{x}$:

$$\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_{(n-1)}}\right) = \mathbf{A}\underset{\substack{[(-x_1)\mathbf{1}+\mathbf{j}]\\ [(-x_2)\mathbf{2}+\mathbf{j}]\\ \vdots \\ [(-x_n)\mathbf{n}+\mathbf{j}]}}{\boldsymbol{\tau}} = \mathbf{A}\begin{bmatrix} 1 & & & -x_1 & & \\ & 1 & & -x_2 & & \\ & & \ddots & \vdots & & \\ & & & 1 & & \\ & & & \vdots & \ddots & \\ & & & -x_n & & 1 \end{bmatrix} = \left[\mathbf{A}_{|1};\ \mathbf{A}_{|2}; \cdots \mathbf{0}; \cdots \mathbf{A}_{|n};\right],$$

Como $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_{(n-1)}}$ es invertible y el resultado es una matriz singular (por ser la columna $j$-ésima nula), $\mathbf{A}$ es singular (Proposición 5.1.4 en la página 70).

Para las filas, basta transponer $\mathbf{A}$ y repetir el argumento con las columnas de $\mathbf{A}^\intercal$.

□

**Ejercicio 34(a)** En particular, también es cierto para $\vec{x} = \vec{0}$; así tenemos que

$$\begin{aligned} \vec{z} + \vec{0} &= \vec{0} && \text{por la propiedad que hemos pedido a } \vec{z} \\ \vec{z} &= \vec{0} && \text{por la propiedad que tiene } \vec{0}. \end{aligned}$$

□

**Ejercicio 34(b)** $\vec{z} = \vec{z} + \vec{0} = \vec{z} + \left(\vec{x} + (-\vec{x})\right) = \underbrace{(\vec{z} + \vec{x})}_{\vec{0}} + (-\vec{x}) = \vec{0} + (-\vec{x}) = -\vec{x}.$

□

**Ejercicio 35.**

1. $\vec{x} + \vec{y} = xy = yx = \vec{y} + \vec{x}$

2. $\vec{x} + (\vec{y} + \vec{z}) = x(yz) = (xy)z = (\vec{x} + \vec{y}) + \vec{z}$

3. Si $\vec{0} = 1$; entonces $\vec{x} + \vec{0} = x1 = x = \vec{x}$; es decir, $\vec{0} = 1$.

4. Si $-\vec{x} = 1/x$; entonces $\vec{x} + (-\vec{x}) = x/x = 1 = \vec{0}$.

5. $1\vec{x} = x^1 = x = \vec{x}$.

6. $(ab)\vec{x} = x^{(ab)} = (x^b)^a = (b\vec{x})^a = a(b\vec{x})$.

7. $a(\vec{x} + \vec{y}) = (xy)^a = x^a \cdot y^a = a\vec{x} + a\vec{y}$.

8. $(a + b)\vec{x} = (x)^{a+b} = x^a \cdot x^b = a\vec{x} + b\vec{x}$.

Por tanto, en este espacio vectorial el vector nulo es $\vec{0} = 1$ y el vector opuesto de $\vec{x}$ es $-\vec{x} = 1/x$.

□

**Ejercicio 36(a)** Esquema de la función lineal: $\begin{array}{rcl} {}_{-|j}: \mathbb{R}^{m\times n} & \longrightarrow & \mathbb{R}^m \\ \mathbf{A} & \longmapsto & (a_{1j}, \ldots, a_{mj},) \end{array}$.

Es lineal puesto que

$$\begin{aligned} \big(\mathbf{A} + \mathbf{B}\big)_{|j} &= \mathbf{A}_{|j} + \mathbf{B}_{|j} \\ \big(\lambda\mathbf{A}\big)_{|j} &= \lambda\big(\mathbf{A}_{|j}\big). \end{aligned}$$

□

**Ejercicio 36(b)** Esquema de la función lineal: $\begin{array}{rcl} {}_{i|-|j}: \mathbb{R}^{m\times n} & \longrightarrow & \mathbb{R} \\ \mathbf{A} & \longmapsto & a_{ij} \end{array}$.

Es lineal puesto que

$$\begin{aligned} {}_{i|}\big(\mathbf{A} + \mathbf{B}\big)_{|j} &= {}_{i|}\mathbf{A}_{|j} + {}_{i|}\mathbf{B}_{|j} \\ {}_{i|}\big(\lambda\mathbf{A}\big)_{|j} &= \lambda\Big({}_{i|}\mathbf{A}_{|j}\Big). \end{aligned}$$

□

**Ejercicio 36(c)** Esquema de la función lineal: $\begin{array}{rcl} {}_{i|-}: \mathbb{R}^{m\times n} & \longrightarrow & \mathbb{R}^n \\ \mathbf{A} & \longmapsto & (a_{i1}, \ldots, a_{in},) \end{array}$.

Es lineal puesto que

$$\begin{aligned} {}_{i|}\big(\mathbf{A} + \mathbf{B}\big) &= {}_{i|}\mathbf{A} + {}_{i|}\mathbf{B} \\ {}_{i|}\big(\lambda\mathbf{A}\big) &= \lambda\big({}_{i|}\mathbf{A}\big). \end{aligned}$$

□

**Ejercicio 36(d)** Esquema de la función lineal: $\begin{array}{rcl} _^{\intercal}\colon \mathbb{R}^{m\times n} & \longrightarrow & \mathbb{R}^{n\times m} \\ \mathbf{A} & \longmapsto & [{}_{1|}\mathbf{A};\ \ldots, {}_{m|}\mathbf{A}] \end{array}$.

Es lineal puesto que

$$\begin{aligned} (\mathbf{A}+\mathbf{B})^{\intercal} &= \mathbf{A}^{\intercal}+\mathbf{B}^{\intercal} \\ \big(\lambda\mathbf{A}\big)^{\intercal} &= \lambda(\mathbf{A}^{\intercal}). \end{aligned}$$

□

**Ejercicio 36(e)** Esquema de la función lineal: $\begin{array}{rcl} _\boldsymbol{b}\colon \mathbb{R}^{m\times n} & \longrightarrow & \mathbb{R}^{n} \\ \mathbf{A} & \longmapsto & \big(\mathbf{A}_{|1}\big)b_1 + \big(\mathbf{A}_{|2}\big)b_2 + \cdots + \big(\mathbf{A}_{|n}\big)b_n \end{array}$.

Es lineal puesto que

$$\begin{aligned} (\mathbf{A}+\mathbf{C})\boldsymbol{b} &= \mathbf{A}\boldsymbol{b}+\mathbf{C}\boldsymbol{b} \\ (\lambda\mathbf{A})\boldsymbol{b} &= \lambda(\mathbf{A}\boldsymbol{b}). \end{aligned}$$

□

**Ejercicio 36(f)** Esquema de la función lineal: $\begin{array}{rcl} \mathbf{A}_\colon \mathbb{R}^{n} & \longrightarrow & \mathbb{R}^{n} \\ \boldsymbol{b} & \longmapsto & \big(\mathbf{A}_{|1}\big)b_1 + \big(\mathbf{A}_{|2}\big)b_2 + \cdots + \big(\mathbf{A}_{|n}\big)b_n \end{array}$.

Es lineal puesto que

$$\begin{aligned} \mathbf{A}(\boldsymbol{b}+\boldsymbol{c}) &= \mathbf{A}\boldsymbol{b}+\mathbf{A}\boldsymbol{c} \\ \mathbf{A}(\lambda\boldsymbol{b}) &= \lambda(\mathbf{A}\boldsymbol{b}). \end{aligned}$$

□

**Ejercicio 36(g)** Esquema de la función lineal: $\begin{array}{rcl} \boldsymbol{a}_\colon \mathbb{R}^{m\times n} & \longrightarrow & \mathbb{R}^{m} \\ \mathbf{B} & \longmapsto & a_1\big({}_{1|}\mathbf{B}\big) + \cdots + a_m\big({}_{m|}\mathbf{B}\big) \end{array}$.

Es lineal puesto que

$$\begin{aligned} \boldsymbol{a}(\mathbf{B}+\mathbf{C}) &= \boldsymbol{a}\mathbf{B}+\boldsymbol{a}\mathbf{C} \\ \boldsymbol{a}(\lambda\mathbf{B}) &= \lambda(\boldsymbol{a}\mathbf{B}). \end{aligned}$$

□

**Ejercicio 36(h)** Esquema de la función lineal: $\begin{array}{rcl} _\mathbf{B}\colon \mathbb{R}^{m} & \longrightarrow & \mathbb{R}^{m} \\ \boldsymbol{a} & \longmapsto & a_1\big({}_{1|}\mathbf{B}\big) + \cdots + a_m\big({}_{m|}\mathbf{B}\big) \end{array}$.

Es lineal puesto que

$$\begin{aligned} \big(\boldsymbol{a}+\boldsymbol{c}\big)\mathbf{B} &= \boldsymbol{a}\mathbf{B}+\boldsymbol{c}\mathbf{B} \\ (\lambda\boldsymbol{a})\mathbf{B} &= \lambda(\boldsymbol{a}\mathbf{B}). \end{aligned}$$

□

**Ejercicio 36(i)** Esquema de la función lineal: $\begin{array}{rcl} \mathbf{A}_\colon \mathbb{R}^{m\times n} & \longrightarrow & \mathbb{R}^{m\times p} \\ \mathbf{A} & \longmapsto & [\mathbf{A}\mathbf{B}_{|1};\ \ldots, \mathbf{A}\mathbf{B}_{|p}] \end{array}$.

Es lineal puesto que

$$(\mathbf{A}+\mathbf{C})\mathbf{B} = \mathbf{AB}+\mathbf{CB}$$
$$(\lambda\mathbf{A})\mathbf{B} = \lambda(\mathbf{AB}).$$

□

**Ejercicio 36(j)** Esquema de la función lineal: $\begin{array}{rcl} _\,\mathbf{B}:\mathbb{R}^{m\times n} & \longrightarrow & \mathbb{R}^{k\times n} \\ \mathbf{B} & \longmapsto & [{}_{1|}\mathbf{AB};\ \ldots,{}_{k|}\mathbf{AB}]^{\intercal} \end{array}$.

Es lineal puesto que

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$$
$$\mathbf{A}(\lambda\mathbf{B}) = \lambda(\mathbf{AB}).$$

□

**Ejercicio 36(k)** Esquema de la función lineal: $\begin{array}{rcl} _{}_{\boldsymbol{\tau}}:\mathbb{R}^{m\times n} & \longrightarrow & \mathbb{R}^{m\times n} \\ \mathbf{A} & \longmapsto & \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}}\right) \end{array}$ donde $\mathbf{I}_{\boldsymbol{\tau}}$ es una matriz elemental del orden $n$. Es lineal puesto que

$$(\mathbf{A}+\mathbf{B})_{\boldsymbol{\tau}} = \mathbf{A}_{\boldsymbol{\tau}}+\mathbf{B}_{\boldsymbol{\tau}}$$
$$(\lambda\mathbf{A})_{\boldsymbol{\tau}} = \lambda(\mathbf{A}_{\boldsymbol{\tau}}).$$

□

**Ejercicio 36(l)** Esquema de la función lineal: $\begin{array}{rcl} {}_{\boldsymbol{\tau}}_:\mathbb{R}^{m\times n} & \longrightarrow & \mathbb{R}^{m\times n} \\ \mathbf{A} & \longmapsto & ({}_{\boldsymbol{\tau}}\mathbf{I})\mathbf{A} \end{array}$ donde ${}_{\boldsymbol{\tau}}\mathbf{I}$ es una matriz elemental del orden $m$. Es lineal puesto que

$${}_{\boldsymbol{\tau}}(\mathbf{A}+\mathbf{B}) = {}_{\boldsymbol{\tau}}\mathbf{A}+{}_{\boldsymbol{\tau}}\mathbf{B}$$
$${}_{\boldsymbol{\tau}}(\lambda\mathbf{A}) = \lambda({}_{\boldsymbol{\tau}}\mathbf{A}).$$

□

**Ejercicio 37(a)** Debemos comprobar que la función $[g\circ f]\colon \mathcal{D}\to\mathcal{W}$ es lineal.

Dados $\vec{x}$ e $\vec{y}$ de $\mathcal{D}$, se verifica que

$$[g\circ f]\big(\vec{x}+\vec{y}\big) = g\big(f(\vec{x}+\vec{y})\big) = g\big(f(\vec{x})+f(\vec{y})\big) = g\big(f(\vec{x})\big)+g\big(f(\vec{y})\big) = [g\circ f](\vec{x})+[g\circ f](\vec{y}).$$

Dados $\vec{x}$ de $\mathcal{D}$ y $\alpha\in\mathbb{R}$, se verifica que

$$[g\circ f](\alpha\vec{x}) = g\big(f(\alpha\vec{x})\big) = g\big(\alpha f(\vec{x})\big) = \alpha g\big(f(\vec{x})\big) = \alpha[g\circ f](\vec{x}).$$

□

**Ejercicio 37(b)** Tomemos $\vec{x}$ e $\vec{y}$ de $\mathcal{V}$, entonces

$$\begin{aligned} f^{-1}\Big(\vec{x}+\vec{y}\Big) &= f^{-1}\Big(f\big(f^{-1}(\vec{x})\big)+f\big(f^{-1}(\vec{y})\big)\Big) && \text{puesto que } \vec{x} = f\big(f^{-1}(\vec{x})\big) \text{ e } \vec{y} = f\big(f^{-1}(\vec{y})\big) \\ &= f^{-1}\Big(f\big(f^{-1}(\vec{x})+f^{-1}(\vec{y})\big)\Big) && \text{puesto que } f \text{ es lineal} \\ &= f^{-1}(\vec{x})+f^{-1}(\vec{y}) && \text{puesto que en general } \vec{z} = f^{-1}\big(f(\vec{z})\big). \end{aligned}$$

Tomemos $\vec{x} \in \mathcal{V}$ y $\alpha \in R$ entonces

$$\begin{aligned} f^{-1}\Big(\alpha\vec{x}\Big) =& f^{-1}\Big(\alpha f\big(f^{-1}(\vec{x})\big)\Big) && \text{puesto que } \vec{x} = f\big(f^{-1}(\vec{x})\big) \\ =& f^{-1}\Big(f\big(\alpha f^{-1}(\vec{x})\big)\Big) && \text{puesto que } f \text{ es lineal} \\ =& \alpha f^{-1}(\vec{x}) && \text{puesto que en general } \vec{z} = f^{-1}\big(f(z)\big). \end{aligned}$$

□

**Ejercicio 38.** Sí, pues la suma de dos combinaciones lineales es una combinación lineal; y cualquier múltiplo de una combinación lineal también es una combinación lineal. Por tanto, el conjunto de combinaciones lineales es cerrado para la suma y el producto por escalares.

□

**Ejercicio 39.** Hay que verificar que se cumplen las ocho propiedades de la Definición 6.1:

Sean los pares $(\vec{a}, \vec{x})$, $(\vec{b}, \vec{y})$ y $(\vec{c}, \vec{z})$, donde las primeras componentes son vectores de $\mathcal{A}$ y las segundas componentes son vectores de $\mathcal{X}$; y sean los escalares $\lambda$ y $\mu$. Entonces

1.

$$\begin{aligned} (\vec{a}, \vec{x}) + (\vec{b}, \vec{y}) =& (\vec{a} + \vec{b}, \vec{x} + \vec{y}) && \text{aplicando la definición de suma de pares} \\ =& (\vec{b} + \vec{a}, \vec{y} + \vec{x}) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales} \\ =& (\vec{b}, \vec{y}) + (\vec{a}, \vec{x}) && \text{aplicando la definición de suma de pares.} \end{aligned}$$

2.

$$\begin{aligned} (\vec{a}, \vec{x}) + \Big((\vec{b}, \vec{y}) + (\vec{c}, \vec{z})\Big) =& (\vec{a}, \vec{x}) + \Big((\vec{b} + \vec{c}),\ (\vec{y} + \vec{z})\Big) && \text{aplicando la definición de suma de pares} \\ =& \Big(\vec{a} + (\vec{b} + \vec{c}),\ \vec{x} + (\vec{y} + \vec{z})\Big) && \text{aplicando la definición de suma de pares} \\ =& \Big((\vec{a} + \vec{b}) + \vec{c},\ (\vec{x} + \vec{y}) + \vec{z}\Big) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales} \\ =& \Big((\vec{a} + \vec{b}),\ (\vec{x} + \vec{y})\Big) + (\vec{c}, \vec{z}) && \text{aplicando la definición de suma de pares} \\ =& \Big((\vec{a}, \vec{x}) + (\vec{b}, \vec{y})\Big) + (\vec{c}, \vec{z}). \end{aligned}$$

3.

$$\begin{aligned} (\vec{a}, \vec{x}) + (\vec{0}_{\mathcal{A}}, \vec{0}_{\mathcal{X}}) =& (\vec{a} + \vec{0}_{\mathcal{A}},\ \vec{x} + \vec{0}_{\mathcal{X}}) && \text{aplicando la definición de suma de funciones} \\ =& (\vec{a}, \vec{x}) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales} \\ & && \text{por tanto } \vec{0} = (\vec{0}_{\mathcal{A}}, \vec{0}_{\mathcal{X}}). \end{aligned}$$

4.

$$\begin{aligned} (\vec{a}, \vec{x}) + (-\vec{a}, -\vec{x}) =& (\vec{a} - \vec{a},\ \vec{x} - \vec{x}) && \text{aplicando la definición de suma de funciones} \\ =& (\vec{0}_{\mathcal{A}}, \vec{0}_{\mathcal{X}}) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales} \\ & && \text{por tanto } (-\vec{a}, -\vec{x}) = -(\vec{a}, \vec{x}). \end{aligned}$$

5.

$$\begin{aligned}
\lambda\big((\vec{a},\vec{x})+(\vec{b},\vec{y})\big) &= \lambda\big((\vec{a}+\vec{b}),\ (\vec{x}+\vec{y})\big) && \text{aplicando la definición de suma de pares}\\
&= \big(\lambda(\vec{a}+\vec{b}),\ \lambda(\vec{x}+\vec{y})\big) && \text{aplicando la definición de escalar por par}\\
&= (\lambda\vec{a}+\lambda\vec{b},\ \lambda\vec{x}+\lambda\vec{y}) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}\\
&= (\lambda\vec{a},\ \lambda\vec{x})+(\lambda\vec{b},\ \lambda\vec{y}) && \text{aplicando la definición de suma de pares}\\
&= \lambda(\vec{a},\vec{x})+\lambda(\vec{b},\vec{y}) && \text{aplicando la definición de escalar por par}
\end{aligned}$$

6.

$$\begin{aligned}
(\lambda+\mu)(\vec{a},\vec{x}) &= \big((\lambda+\mu)\vec{a},\ (\lambda+\mu)\vec{x}\big) && \text{aplicando la definición de escalar por par}\\
&= \big((\lambda\vec{a}+\mu\vec{a}),\ (\lambda\vec{x}+\mu\vec{x})\big) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}\\
&= (\lambda\vec{a},\ \lambda\vec{x})+(\mu\vec{a},\ \mu\vec{x}) && \text{aplicando la definición de suma de pares}\\
&= \lambda(\vec{a},\vec{x})+\mu(\vec{a},\vec{x}) && \text{aplicando la definición de escalar por par}
\end{aligned}$$

7.

$$\begin{aligned}
(\lambda\mu)(\vec{a},\vec{x}) &= \big((\lambda\mu)\vec{a},\ (\lambda\mu)\vec{x}\big) && \text{aplicando la definición de escalar por par}\\
&= \big(\lambda(\mu\vec{a}),\ \lambda(\mu\vec{x})\big) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}\\
&= \lambda(\mu\vec{a},\ \mu\vec{x}) && \text{aplicando la definición de escalar por par}
\end{aligned}$$

8.

$$\begin{aligned}
1(\vec{a},\vec{x}) &= (1\vec{a},1\vec{x}) && \text{aplicando la definición de escalar por par}\\
&= (\vec{a},\vec{x}) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}
\end{aligned}$$

□

**Ejercicio 40.** Hay que verificar que se cumplen las ocho propiedades de la Definición 6.1:

Sean las funciones $f\colon \mathcal{D}\to\mathcal{V}$, $g\colon \mathcal{D}\to\mathcal{V}$, y $h\colon \mathcal{D}\to\mathcal{V}$; y sean los escalares $\lambda$ y $\mu$. Entonces

1.

$$\begin{aligned}
[f+g](\vec{x}) &= f(\vec{x})+g(\vec{x}) && \text{aplicando la definición de suma de funciones}\\
&= g(\vec{x})+f(\vec{x}) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial}\\
&= [g+f](\vec{x}) && \text{aplicando la definición de suma de funciones}
\end{aligned}$$

2.

$$\begin{aligned}
[f+(g+h)](\vec{x}) &= f(\vec{x})+[g+h](\vec{x}) && \text{aplicando la definición de suma de funciones}\\
&= f(\vec{x})+\big(g(\vec{x})+h(\vec{x})\big) && \text{aplicando la definición de suma de funciones}\\
&= \big(f(\vec{x})+g(\vec{x})\big)+h(\vec{x}) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial}\\
&= [f+g](\vec{x})+h(\vec{x}) && \text{aplicando la definición de suma de funciones}\\
&= [(f+g)+h](\vec{x}) && \text{aplicando la definición de suma de funciones}
\end{aligned}$$

3. Sea $0: X \longrightarrow \mathcal{V}$, entonces
$\vec{x} \longmapsto \vec{0}$

$$\begin{aligned}
[f+0](\vec{x}) &= f(\vec{x}) + 0(\vec{x}) && \text{aplicando la definición de suma de funciones} \\
&= f(\vec{x}) + \vec{0} && \text{pues } 0(\vec{x}) = \vec{0} \\
&= f(\vec{x})
\end{aligned}$$

4. Sea $-f: X \longrightarrow \mathcal{V}$, entonces
$\vec{x} \longmapsto -f(\vec{x})$

$$\begin{aligned}
[f+(-f)](\vec{x}) &= f(\vec{x}) + \big(-f(\vec{x})\big) && \text{aplicando la definición de suma de funciones} \\
&= \vec{0} && \text{puesto que en } \mathcal{V} \text{ se verifica que } f(\vec{x}) + \big(-f(\vec{x})\big) = \vec{0} \\
&= 0(\vec{x}) && \text{pues } 0(\vec{x}) = \vec{0}
\end{aligned}$$

5.

$$\begin{aligned}
[\lambda(f+g)](\vec{x}) &= \lambda\big([f+g](\vec{x})\big) && \text{aplicando la definición de función por escalar} \\
&= \lambda\big(f(\vec{x}) + g(\vec{x})\big) && \text{aplicando la definición de suma de funciones} \\
&= \lambda f(\vec{x}) + \lambda g(\vec{x}) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial} \\
&= [\lambda f](\vec{x}) + [\lambda g](\vec{x}) && \text{aplicando la definición de función por escalar} \\
&= [\lambda f + \lambda g](\vec{x}) && \text{aplicando la definición de suma de funciones}
\end{aligned}$$

6.

$$\begin{aligned}
[(\lambda+\mu)f](\vec{x}) &= (\lambda+\mu)f(\vec{x}) && \text{aplicando la definición de función por escalar} \\
&= \lambda f(\vec{x}) + \mu f(\vec{x}) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial} \\
&= [\lambda f](\vec{x}) + [\mu f](\vec{x}) && \text{aplicando la definición de función por escalar} \\
&= [\lambda f + \mu f](\vec{x}) && \text{aplicando la definición de suma de funciones}
\end{aligned}$$

7.

$$\begin{aligned}
[(\lambda \cdot \mu)f](\vec{x}) &= (\lambda \cdot \mu)f(\vec{x}) && \text{aplicando la definición de función por escalar} \\
&= \lambda\big(\mu f(\vec{x})\big) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial} \\
&= \lambda\big([\mu f](\vec{x})\big) && \text{aplicando la definición de función por escalar} \\
&= [\lambda(\mu f)](\vec{x}) && \text{aplicando la definición de función por escalar}
\end{aligned}$$

8.

$$\begin{aligned}
[1 \cdot f](\vec{x}) &= 1 \cdot f(\vec{x}) && \text{aplicando la definición de función por escalar} \\
&= f(\vec{x}) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial}
\end{aligned}$$

□

**Ejercicio 41(a)** Tenemos que comprobar que la función $[g+f]: \mathcal{D} \to \mathcal{W}$ es lineal.

Dados $\vec{x}$ e $\vec{y}$ de $\mathcal{D}$, se verifica que

$$[g+f]\big(\vec{x}+\vec{y}\big) = g(\vec{x}+\vec{y}) + f(\vec{x}+\vec{y}) = g(\vec{x}) + g(\vec{y}) + f(\vec{x}) + f(\vec{y}) = \\ g(\vec{x}) + f(\vec{x}) + g(\vec{y}) + f(\vec{y}) = [g+f](\vec{x}) + [g+f](\vec{y}).$$

Dados $\vec{x}$ de $\mathcal{D}$ y $\alpha \in \mathbb{R}$, se verifica que

$$[g+f](\alpha\vec{x}) = g(\alpha\vec{x}) + f(\alpha\vec{x}) = \alpha g(\vec{x}) + \alpha f(\vec{x}) = \alpha\big(g(\vec{x}) + f(\vec{x})\big) = \alpha[g+f](\vec{x}).$$

□

**Ejercicio 41(b)** Tenemos que comprobar que la función $[\alpha \cdot f]\colon \mathcal{D} \to \mathcal{W}$ es lineal.

Dados $\vec{x}$ e $\vec{y}$ de $\mathcal{D}$, se verifica que

$$[\alpha \cdot f](\vec{x}+\vec{y}) = \alpha f(\vec{x}+\vec{y}) = \alpha\big(f(\vec{x}) + f(\vec{y})\big) = \alpha f(\vec{x}) + \alpha f(\vec{y}) = [\alpha \cdot f](\vec{x}) + [\alpha \cdot f](\vec{y}).$$

Dados $\vec{x}$ de $\mathcal{D}$ y $\gamma \in \mathbb{R}$, se verifica que

$$[\alpha \cdot f](\gamma\vec{x}) = \alpha f(\gamma\vec{x}) = \alpha\big(\gamma f(\vec{x})\big) = \gamma\big(\alpha f(\vec{x})\big) = \gamma[\alpha \cdot f](\vec{x}).$$

□

**Ejercicio 42.** Basta comprobar que cualquier combinación de soluciones de $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ es también solución.

Sean $\boldsymbol{x}$ y $\boldsymbol{y}$ dos vectores de $\mathcal{N}\left(\mathbf{A}\right)$, es decir, $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ y $\mathbf{A}\boldsymbol{y} = \mathbf{0}$, y sean $b$ y $c$ dos números reales; entonces

$$\mathbf{A}\big(b\boldsymbol{x} + c\boldsymbol{y}\big) = \mathbf{A}b\boldsymbol{x} + \mathbf{A}c\boldsymbol{y} = b\mathbf{A}\boldsymbol{x} + c\mathbf{A}\boldsymbol{y} = b\mathbf{0} + c\mathbf{0} = \mathbf{0} \quad \Rightarrow \quad \big(b\boldsymbol{x} + c\boldsymbol{y}\big) \in \mathcal{N}\left(\mathbf{A}\right).$$

□

**Ejercicio 43.** Sea $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_r,)$ la lista de índices de las $r$ columnas de $\mathbf{K}$ con pivote, y sea $\boldsymbol{\nu}$ la lista de índices de las $n-r$ columnas nulas, es decir, $\mathbf{K}_{|\boldsymbol{\nu}} = \mathbf{0}$.

Que en $\mathbf{K}\boldsymbol{x}$ sean nulos los coeficientes correspondientes a las columnas no nulas significa que el subvector $({}_{\boldsymbol{\rho}|}\boldsymbol{x})$ es $\mathbf{0}$. Así, usando la Ecuación 3.8 en la página 45 deducimos que necesariamente

$$\mathbf{K}\boldsymbol{x} \;=\; \big(\mathbf{K}_{|\boldsymbol{\rho}}\big)\underbrace{\big({}_{\boldsymbol{\rho}|}\boldsymbol{x}\big)}_{\mathbf{0}} \;+\; \underbrace{\big(\mathbf{K}_{|\boldsymbol{\nu}}\big)}_{\mathbf{0}}\big({}_{\boldsymbol{\nu}|}\boldsymbol{x}\big) \;=\; \mathbf{0}.$$

Recíprocamente, supongamos que $({}_{\boldsymbol{\rho}|}\boldsymbol{x}) \neq \mathbf{0}$, es decir, que algún coeficiente es distinto de cero para alguna columna no nula de $\mathbf{K}$; y sea $j$ el menor índice de $\boldsymbol{\rho}$ tal que ${}_{j|}\boldsymbol{x} \neq 0$. Entonces, para todo $h < j$, el producto $(\mathbf{K}_{|h})({}_{h|}\boldsymbol{x}) = \mathbf{0}$ (pues son nulos $\mathbf{K}_{|h}$ ó ${}_{h|}\boldsymbol{x}$). Por tanto, $\mathbf{K}\boldsymbol{x} = (\mathbf{K}_{|j})x_j + \cdots + (\mathbf{K}_{|n})x_n$. Si $i$ es la posición de pivote de $\mathbf{K}_{|j}$, tendremos que ${}_{i|}\Big((\mathbf{K}_{|j})x_j + \cdots + (\mathbf{K}_{|n})x_n\Big) = ({}_{i|}\mathbf{K}_{|j})x_j \neq \mathbf{0}$; pues todas las componentes a la derecha del pivote ${}_{i|}\mathbf{K}_{|j}$ son nulas por estar $\mathbf{K}$ preescalonada. Consecuentemente $\mathbf{K}\boldsymbol{x} \neq \mathbf{0}$.

□

**Ejercicio 44(a)**

$$\mathsf{Z}(\boldsymbol{b}+\boldsymbol{c}) = \sum_{k=1}^{n}\left(\mathsf{Z}_{|k}(b_k + c_k)\right) = \sum_{k=1}^{n}\left(\big(\mathsf{Z}_{|k}\big)b_k \;+\; \big(\mathsf{Z}_{|k}\big)c_k\right) = \sum_{k=1}^{n}\big(\mathsf{Z}_{|k}\big)b_k \;+\; \sum_{k=1}^{n}\big(\mathsf{Z}_{|k}\big)c_k = \mathsf{Z}\boldsymbol{b} + \mathsf{Z}\boldsymbol{c}.$$

□

**Ejercicio 44(b)** $\mathsf{Z}(\lambda \boldsymbol{b}) = \sum_{k=1}^{n} (\mathsf{Z}_{|k}) \lambda b_k = \lambda \sum_{k=1}^{n} (\mathsf{Z}_{|k}) b_k = \lambda(\mathsf{Z}\boldsymbol{b})$.

□

**Ejercicio 45(a)**

$$\mathsf{Z}(\mathbf{B}\boldsymbol{c}) = \mathsf{Z}\left(\sum_{k=1}^{n} (\mathbf{B}_{|k}) c_k\right) = \sum_{k=1}^{n} \Big(\mathsf{Z}(\mathbf{B}_{|k}) c_k\Big) = \sum_{k=1}^{n} \Big((\mathsf{Z}\mathbf{B})_{|k}\Big) c_k = \Big[\mathsf{Z}\mathbf{B}_{|1}; \ldots; \mathsf{Z}\mathbf{B}_{|n}\Big]\boldsymbol{c} = (\mathsf{Z}\mathbf{B})\boldsymbol{c}.$$

□

**Ejercicio 45(b)** Dado que $\mathbf{C}_{|j}$ es un vector de $\mathbb{R}^q$, aplicando las definiciones de producto de sistema por matriz, y matriz por matriz junto con la proposición anterior, obtenemos:

$$\Big(\mathsf{Z}(\mathbf{B}\mathbf{C})\Big)_{|j} = \mathsf{Z}(\mathbf{B}\mathbf{C})_{|j} = \mathsf{Z}\Big(\mathbf{B}(\mathbf{C}_{|j})\Big) \underset{\text{(Prop. 9.1.3)}}{=} \Big(\mathsf{Z}\mathbf{B}\Big)(\mathbf{C}_{|j}) = \Big((\mathsf{Z}\mathbf{B})\mathbf{C}\Big)_{|j}, \qquad \forall j = \{1, \ldots, n\}.$$

□

**Ejercicio 46.** Sea la lista de índices menores que $i$, $\boldsymbol{\gamma}_1 = \big(h \mid h < i\big)$ y la lista de índices mayores que $i$, $\boldsymbol{\gamma}_2 = \big(h \mid h > i\big)$, de manera que la lista de índices $(1:n)$ de 1 a $n$ es la concatenación de la lista $\boldsymbol{\gamma}_1$, con la lista de un solo elemento $(i,)$ y la lista $\boldsymbol{\gamma}_2$; es decir, $(1:n) = \big(\boldsymbol{\gamma}_1 +\!\!+ (i,) +\!\!+ \boldsymbol{\gamma}_2\big)$.

Por una parte, si $\vec{z}_i$ es combinación lineal del resto de vectores de $\mathsf{Z}$ (los índices menores que $i$ y los de índice mayor que $i$):

$$\mathsf{Z}_{|i} = (\mathsf{Z}_{|\boldsymbol{\gamma}_1})\boldsymbol{b} + (\mathsf{Z}_{|\boldsymbol{\gamma}_2})\boldsymbol{c};$$

entonces, restando $\mathsf{Z}_{|i}$ de ambos lados (y recordando la Ecuación 3.8 en la página 45) tenemos que

$$\vec{0} \;=\; (\mathsf{Z}_{|\boldsymbol{\gamma}_1})\boldsymbol{b} + \underbrace{(\mathsf{Z}_{|(i,)})(-1,)}_{-\mathsf{Z}_{|i}} + (\mathsf{Z}_{|\boldsymbol{\gamma}_2})\boldsymbol{c} \;=\; \mathsf{Z}\boldsymbol{a},$$

donde $\boldsymbol{a}$ es la concatenación de los vectores $\boldsymbol{b}$, $(-1,)$ y $\boldsymbol{c}$, es decir $\boldsymbol{a} = \boldsymbol{b} +\!\!+ (-1,) +\!\!+ \boldsymbol{c}$.

Por otra parte, si $\vec{0} = \mathsf{Z}\boldsymbol{a}$, con $\boldsymbol{a} \neq \mathbf{0}$, entonces existe un $i$ tal que ${}_{i|}\boldsymbol{a} \neq 0$, y, por lo tanto,

$$\vec{0} \;=\; \mathsf{Z}\boldsymbol{a} \;=\; (\mathsf{Z}_{|\boldsymbol{\gamma}_1})({}_{\boldsymbol{\gamma}_1|}\boldsymbol{a}) + \underbrace{(\mathsf{Z}_{|(i,)})({}_{(i,)|}\boldsymbol{a})}_{a_i \mathsf{Z}_{|i}} + (\mathsf{Z}_{|\boldsymbol{\gamma}_2})({}_{\boldsymbol{\gamma}_2|}\boldsymbol{a}).$$

Restando $(\mathsf{Z}_{|(i,)})({}_{(i,)|}\boldsymbol{a})$ de ambos lados, tenemos que

$$-a_i \mathsf{Z}_{|i} \;=\; (\mathsf{Z}_{|\boldsymbol{\gamma}_1})({}_{\boldsymbol{\gamma}_1|}\boldsymbol{a}) + (\mathsf{Z}_{|\boldsymbol{\gamma}_2})({}_{\boldsymbol{\gamma}_2|}\boldsymbol{a}) \quad\Longrightarrow\quad \mathsf{Z}_{|i} \;=\; \frac{-1}{a_i}(\mathsf{Z}_{|(\boldsymbol{\gamma}_1 +\!\!+ \boldsymbol{\gamma}_2)})({}_{(\boldsymbol{\gamma}_1 +\!\!+ \boldsymbol{\gamma}_2)|}\boldsymbol{a}).$$

□

**Ejercicio 47.** Basta expresar la imagen de $f_{\mathsf{Z}}$ con la notación funcional (6.6) y comprobar que coincide con la notación matricial de la Definición 9.2 de espacio engendrado por $\mathsf{Z}$ (9.1):

$$\mathit{imagen}(f_{\mathsf{Z}}) = \Big\{ \vec{v} \in \mathcal{V} \;\Big|\; \text{existe } \boldsymbol{a} \in \mathbb{R}^n \text{ tal que } \vec{v} = \mathsf{Z}\boldsymbol{a} \Big\} = \mathcal{L}(\mathsf{Z})$$

□

**Ejercicio 48.** Recordemos que $\boldsymbol{b} = f_{\mathbf{A}}(\boldsymbol{x})$ es equivalente a decir que $(\boldsymbol{x}, \boldsymbol{b}) \in f_{\mathbf{A}}$ (véase *notación funcional* en la Página 96). Así pues, la imagen de $f_{\mathbf{A}}$ es

$$\begin{aligned} imagen(f_{\mathbf{A}}) &= \left\{\boldsymbol{b} \mid \text{existe } \boldsymbol{x} \text{ tal que } (\boldsymbol{x}, \boldsymbol{b}) \in f_{\mathbf{A}}\right\} \\ &= \left\{\boldsymbol{b} \mid \text{existe } \boldsymbol{x} \in \mathbb{R}^n \text{ tal que } \boldsymbol{b} = f_{\mathbf{A}}(\boldsymbol{x}) = \mathbf{A}\boldsymbol{x}\right\} \;=\; \mathcal{C}\left(\mathbf{A}\right). \end{aligned}$$

□

**Ejercicio 49.** Sean $\vec{x}, \vec{y} \in \mathcal{N}\left(f\right)$, entonces dados $\alpha, \beta \in \mathbb{R}$

$$f(\alpha\vec{x} + \beta\vec{y}) = f(\alpha\vec{x}) + f(\beta\vec{y}) = \alpha f(\vec{x}) + \beta f(\vec{y}) = \alpha\vec{0} + \beta\vec{0} = \vec{0}.$$

Por tanto, $(\alpha\vec{x} + \beta\vec{y}) \in \mathcal{N}\left(f\right)$.

□

**Ejercicio 50.** Hay que demostrar que la imagen es cerrada para las combinaciones lineales.

Sea $f\colon \mathcal{V} \to \mathcal{W}$ lineal. Sabemos que la imagen de $f$ es no vacía, pues al ser $f$ lineal, su dominio $\mathcal{V}$ es subespacio y $f(\vec{0}) = \vec{0}$.

Sean $\vec{x}, \vec{y} \in imagen(f)$, entonces existen $\vec{v}, \vec{w} \in \mathcal{V}$ tales que $f(\vec{v}) = \vec{x}$ y $f(\vec{w}) = \vec{y}$. Así, dados $\alpha, \beta \in \mathbb{R}$

$$f(\alpha\vec{v} + \beta\vec{w}) = f(\alpha\vec{v}) + f(\beta\vec{w}) = \alpha f(\vec{v}) + \beta f(\vec{w}) = \alpha\vec{x} + \beta\vec{y} \in imagen(f).$$

□

**Ejercicio 51.** Supongamos que $\vec{z} \in imagen(g \circ f)$. Entonces tiene que existir $\vec{x}$ tal que $(\vec{x}, \vec{y}) \in f$ y $(\vec{y}, \vec{z}) \in g$. Por tanto, $\vec{z} \in imagen(g)$.

□

**Ejercicio 52.** Por una parte sabemos que $imagen(g \circ f) \subset imagen(g)$.

Por otra parte, puesto que $imagen(g) = imagen(g \circ f \circ f^{-1})$, aplicando la proposición anterior tenemos que

$$imagen(g) = imagen\big([g \circ f] \circ f^{-1}\big) \subset imagen(g \circ f).$$

□

**Ejercicio 53.** La función $f \circ f_{\mathsf{B}}$ es lineal y va de $\mathbb{R}^n$ a $\mathcal{W}$. Por tanto, existe un sistema $\mathsf{Z}$ de vectores de $\mathcal{W}$ tal que $f \circ f_{\mathsf{B}} = f_{\mathsf{Z}}$. Por tanto

$$f(\vec{v}) = f\Big(\mathsf{B}\left(\vec{v}_{/_{\mathsf{B}}}\right)\Big) = f\Big(f_{\mathsf{B}}\left(\vec{v}_{/_{\mathsf{B}}}\right)\Big) = f \circ f_{\mathsf{B}}\left(\vec{v}_{/_{\mathsf{B}}}\right) = f_{\mathsf{Z}}\left(\vec{v}_{/_{\mathsf{B}}}\right) = \mathsf{Z}\left(\vec{v}_{/_{\mathsf{B}}}\right).$$

□

**Ejercicio 54.**

Módulo NAcAL para Python

```
A = Sistema([Vector([1, 0, 1, 0]), Vector([0, -1, 0, -1])])
B = Sistema([Vector([1, 1, 1, 1]), Vector([1,  0, 0,  0])])
C = A.concatena(B)      # concatenación de A y B
Homogenea(C,1)
```

$$\left[\begin{array}{cccc} 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{[(-1)\mathbf{1}+\mathbf{3}]\\ [(-1)\mathbf{1}+\mathbf{4}]}]{\boldsymbol{\tau}} \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & -1 & 1 & 0 \\ \hline 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[[(1)\mathbf{2}+\mathbf{3}]]{\boldsymbol{\tau}} \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 \\ \hline 1 & 0 & -1 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

De donde deducimos que las dos primeras componentes de la única solución especial nos indican cómo combinar los vectores del sistema `A` (es decir, que la diferencia de vectores de dicho sistema es una base de la intersección). Fíjese también en que las últimas dos componentes de la única solución especial nos indican cómo combinar los vectores del sistema `B` (es decir, que el primer vector de dicho sistema es una base de la intersección). Comprobación:

Módulo NAcAL para Python

```
SubEspacio(A) & SubEspacio(B) # SubEspacio intersección
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^1,\ \boldsymbol{v} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

□

**Ejercicio 55(a)** Sean $\vec{x}, \vec{y} \in \mathcal{V}$, entonces

$$\begin{array}{ll} f(\vec{x}) = (\vec{x}_a, \vec{x}_b) & \text{con } \vec{x}_a \in \mathcal{A};\ \text{y}\ \ \vec{x}_b \in \mathcal{B},\ \text{ tales que } \ \vec{x} = \vec{x}_a + \vec{x}_b \\ f(\vec{y}) = (\vec{y}_a, \vec{y}_b) & \text{con } \vec{y}_a \in \mathcal{A};\ \text{e}\ \ \vec{y}_b \in \mathcal{B},\ \text{ tales que } \ \vec{y} = \vec{y}_a + \vec{y}_b. \end{array}$$

Consecuentemente

$$\alpha\vec{x} + \beta\vec{y} \ = \ \underbrace{(\alpha\vec{x}_a + \beta\vec{y}_a)}_{\in\mathcal{A}} + \underbrace{(\alpha\vec{x}_b + \beta\vec{y}_b)}_{\in\mathcal{B}}.$$

Y como, por la Proposición 10.1.4, sabemos que dicha descomposición es única, entonces sabemos que

$$\begin{array}{rll} f(\alpha\vec{x} + \beta\vec{y}) = (\alpha\vec{x}_a + \beta\vec{y}_a\ ,\ \alpha\vec{x}_b + \beta\vec{y}_b) & & \text{aplicando } f \text{ y dado que la descomposición es única} \\ = (\alpha\vec{x}_a, \alpha\vec{x}_b) + (\beta\vec{y}_a, \beta\vec{y}_b) & & \text{aplicando la suma de productos cartesianos} \\ = \alpha(\vec{x}_a, \vec{x}_b) + \beta(\vec{y}_a, \vec{y}_b) & & \text{aplicando el producto por escalar del producto cartesiano} \\ = \alpha f(\vec{x}) + \beta f(\vec{y}). & & \end{array}$$

□

**Ejercicio 55(b)** La demostración es como la anterior, pero considerando únicamente las primeras componentes de los productos cartesianos. Es decir: $f_{\mathcal{A}}(\alpha\vec{x} + \beta\vec{y}) = \alpha\vec{x}_a + \beta\vec{y}_a = \alpha f_{\mathcal{A}}(\vec{x}) + \beta f_{\mathcal{A}}(\vec{y})$.

□

**Ejercicio 55(c)** Debemos demostrar que $\mathcal{N}\left(f_{\mathcal{A}}\right) = \left\{\vec{v} \in \mathcal{V} \mid f(\vec{v}) = \vec{0}\right\} = \mathcal{B}$.

Por una parte, sea $\vec{x} = \vec{x}_a + \vec{x}_b$ con $\vec{x}_a \in \mathcal{A}$ y $\vec{x}_b \in \mathcal{B}$; entonces

$$\text{si } \vec{x} \in \mathcal{N}\left(f_{\mathcal{A}}\right) \quad \Rightarrow \quad f(\vec{x}) = f(\vec{x}_a + \vec{x}_b) = \vec{0} \quad \Rightarrow \quad \vec{x}_a = \vec{0} \quad \Rightarrow \quad \vec{x} = \vec{x}_b \in \mathcal{B}.$$

Por otra parte, sea $\vec{x} \in \mathcal{B}$, entonces $f(\vec{x}) = f(\vec{0} + \vec{x}) = \vec{0}$ por tanto $\vec{x} \in \mathcal{N}\left(f_{\mathcal{A}}\right)$.

□

**Ejercicio 56.** Sea $\boldsymbol{\varphi} = \left(\varphi_1, \dots \varphi_r, \right)$ la lista de posiciones de los $r$ pivotes de $\mathbf{K}$

$$\boldsymbol{\varphi} = \left(h \mid h \text{ es la posición de algún pivote de } \mathbf{K}\right),$$

donde no hay índices repetidos y los índices $\varphi_i$ están dispuestos de menor a mayor. Entonces la matriz ${}_{\boldsymbol{\varphi}|}\mathbf{K}$ de orden $r$ por $n$ es de rango completo por filas. Para demostrar que sus filas son linealmente independientes, vamos a demostrar que si combinación lineal de las filas es nula, $\boldsymbol{x}\left({}_{\boldsymbol{\varphi}|}\mathbf{K}\right) = \mathbf{0}$, entonces $\boldsymbol{x} = \mathbf{0} \in \mathbb{R}r$.

Sea $j$ la columna correspondiente al *último* pivote (posición $r$); como $\boldsymbol{x} \cdot \left({}_{\boldsymbol{\varphi}|}\mathbf{K}\right)_{|j} = \mathbf{0}_{|j}$ y dado que la única componente no nula de dicha columna es el pivote, la componente de $\boldsymbol{x}$ que multiplica al pivote tiene que ser cero: $\boldsymbol{x}_{|r} = 0$.

Sea $j$ la columna correspondiente al *penúltimo* pivote (posición $r-1$); como $\boldsymbol{x} \cdot \left({}_{\boldsymbol{\varphi}|}\mathbf{K}\right)_{|j} = \mathbf{0}_{|j}$ y dado que las componentes de $\boldsymbol{x}$ con posición mayor que $r-1$ son nulas, la componente que multiplica al pivote tiene que ser cero: $\boldsymbol{x}_{|r-1} = 0$.

Sea $j$ la columna correspondiente al pivote *ante-penúltimo* (posición $r-2$); entonces repitiendo el argumento se deduce que $\boldsymbol{x}_{|r-2} = 0 \dots$. Y así con todas las componentes de $\boldsymbol{x}$, es decir, $\boldsymbol{x}_{|h} = 0$ para $h = 1 : r$.

Para ver que cada fila sin pivote es combinación lineal de las filas con pivote que la preceden, basta aplicar la eliminación "de arriba a abajo" sobre las filas de $\mathbf{K}$ para anular las componentes por debajo de los pivotes. De esta manera, las filas sin pivote quedan anuladas, es decir, son combinación lineal de las filas con pivote[4].

□

**Ejercicio 57(a)**

$\boxed{\subset}$ Dado que las columnas de $\mathbf{A}(\mathbf{A}^\intercal)$ son combinaciones lineales de las de $\mathbf{A}$, ya que $\left(\mathbf{A}(\mathbf{A}^\intercal)\right)_{|j} = \mathbf{A}(\mathbf{A}^\intercal)_{|j}$, necesariamente $\mathcal{C}\left(\mathbf{A}(\mathbf{A}^\intercal)\right) \subset \mathcal{C}\left(\mathbf{A}\right)$.

$\boxed{\supset}$ A continuación, buscamos demostrar la inclusión en la otra dirección, es decir, que para todo $\boldsymbol{b} \in \mathcal{C}\left(\mathbf{A}\right)$ resulta que $\boldsymbol{b} \in \mathcal{C}\left(\mathbf{A}(\mathbf{A}^\intercal)\right)$; es decir, que existe un $\boldsymbol{y} \in \mathbb{R}^m$ tal que $\boldsymbol{b} = \mathbf{A}(\mathbf{A}^\intercal)\boldsymbol{y}$:

Sabemos que para todo $\boldsymbol{b} \in \mathcal{C}\left(\mathbf{A}\right)$ existe un $\boldsymbol{x} \in \mathbb{R}^n$ tal que $\boldsymbol{b} = \mathbf{A}\boldsymbol{x}$. Además, sabemos que $\mathbb{R}^n = \mathcal{C}\left(\mathbf{A}^\intercal\right) \oplus \mathcal{N}\left(\mathbf{A}\right)$, por lo tanto, existen $\boldsymbol{f} \in \mathcal{C}\left(\mathbf{A}^\intercal\right)$ y $\boldsymbol{n} \in \mathcal{N}\left(\mathbf{A}\right)$ tales que $\boldsymbol{x} = \boldsymbol{f} + \boldsymbol{n}$.

Por tanto, $\boldsymbol{b} = \mathbf{A}\boldsymbol{x} = \mathbf{A}(\boldsymbol{f} + \boldsymbol{n}) = \mathbf{A}\boldsymbol{f} + \mathbf{0} = \mathbf{A}\boldsymbol{f}$. Sin embargo, como $\boldsymbol{f} \in \mathcal{C}\left(\mathbf{A}^\intercal\right)$ (es decir, como $\boldsymbol{f}$ es combinación de las filas de $\mathbf{A}$), sabemos que existe un $\boldsymbol{y} \in \mathbb{R}^m$ tal que $\boldsymbol{f} = \boldsymbol{y}\mathbf{A} = (\mathbf{A}^\intercal)\boldsymbol{y}$. En consecuencia, $\boldsymbol{b} = \mathbf{A}\boldsymbol{f} = \mathbf{A}(\mathbf{A}^\intercal)\boldsymbol{y}$; y por tanto, $\boldsymbol{b} \in \mathcal{C}\left(\mathbf{A}(\mathbf{A}^\intercal)\right)$. Así que $\mathcal{C}\left(\mathbf{A}(\mathbf{A}^\intercal)\right) \supset \mathcal{C}\left(\mathbf{A}\right)$.

□

**Ejercicio 57(b)** Basta con renombrar $\mathbf{A}^\intercal$ como $\mathbf{B}$ y aplicar el apartado anterior.

□

**Ejercicio 57(c)** Puesto que $\mathbf{A}(\mathbf{A}^\intercal)$ y $\mathbf{A}$ tienen el mismo espacio columna, y que la dimensión de dicho espacio es el rango, ambas matrices tienen el mismo rango.

□

[4] Aplicar la eliminación "de arriba a abajo" sobre las filas es equivalente a aplicar transformaciones elementales *"de izquierda a derecha"* sobre el sistema de vectores $\left[{}_{1|}\mathbf{L}, \dots, {}_{m|}\mathbf{L}\right]$.

**Ejercicio 57(d)** Como el rango no cambia al transponer la matriz, tenemos que $\operatorname{rg}\big(\mathbf{A}(\mathbf{A}^\intercal)\big) = \operatorname{rg}\big((\mathbf{A}^\intercal)\mathbf{A}\big) = \operatorname{rg}\big(\mathbf{A}^\intercal\big) = \operatorname{rg}\big(\mathbf{A}\big)$.

□

**Ejercicio 58.** Evidentemente, $\mathbf{0}$ pertenece a $A^\perp$, así que el subconjunto $A^\perp$ no es vacío.

Sean $\boldsymbol{c}_1, \boldsymbol{c}_2 \in A^\perp$ y $x, y \in \mathbb{R}$. Entonces $\boldsymbol{c}_1 \cdot \boldsymbol{a} = 0$ y $\boldsymbol{c}_2 \cdot \boldsymbol{a} = 0$, para cualquier $\boldsymbol{a} \in A$. Consecuentemente

$$(x\boldsymbol{c}_1 + y\boldsymbol{c}_2) \cdot \boldsymbol{a} = x\big(\boldsymbol{c}_1 \cdot \boldsymbol{a}\big) + y\big(\boldsymbol{c}_2 \cdot \boldsymbol{a}\big) = x \cdot 0 + y \cdot 0 = 0.$$

Por tanto, $(x\boldsymbol{c}_1 + y\boldsymbol{c}_2) \in A^\perp$. Es decir, $A^\perp$ es cerrado para las combinaciones lineales.

□

**Ejercicio 59(a)** $\boldsymbol{x} \in B^\perp$ si y solo si $\boldsymbol{x} \cdot \boldsymbol{b} = 0$ para cada $\boldsymbol{b} \in B$; (y como $A \subset B$) en particular también $\boldsymbol{x} \cdot \boldsymbol{a} = 0$ para cada $\boldsymbol{a} \in A$, es decir, $\boldsymbol{x} \in A^\perp$.

□

**Ejercicio 59(b)** Tenemos que demostrar la inclusión de estos subespacios en ambos sentidos.

*Inclusión en un sentido*: como $A \subset \mathcal{L}\big(A\big)$, por el apartado anterior sabemos que $\boxed{\mathcal{L}\big(A\big)^\perp \subset A^\perp}$.

*Inclusión en el otro sentido*: sabemos que $\boldsymbol{x} \in A^\perp$ si y solo si $\boldsymbol{x} \cdot \boldsymbol{a} = 0$ para cada $\boldsymbol{a} \in A$. Sea $\boldsymbol{z} \in \mathcal{L}\big(A\big)$, entonces existen $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ y $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \in A$ tales que $\boldsymbol{z} = \alpha_1\boldsymbol{a}_1 + \cdots + \alpha_k\boldsymbol{a}_k$. Entonces $\boldsymbol{x} \cdot \boldsymbol{z} = \boldsymbol{x} \cdot (\alpha_1\boldsymbol{a}_1 + \cdots + \alpha_k\boldsymbol{a}_k) = \alpha_1\boldsymbol{x} \cdot \boldsymbol{a}_1 + \cdots + \alpha_k\boldsymbol{x} \cdot \boldsymbol{a}_k = \alpha_1 0 + \cdots + \alpha_k 0 = 0$; es decir, $\boldsymbol{x} \in \mathcal{L}\big(A\big)^\perp$. Por tanto $\boxed{A^\perp \subset \mathcal{L}\big(A\big)^\perp}$.

□

**Ejercicio 59(c)** Sea $\boldsymbol{a} \in \mathcal{A} \cap \mathcal{A}^\perp$; como $\boldsymbol{a} \in \mathcal{A}$ es perpendicular a todo vector en $\mathcal{A}^\perp$, entonces $\boldsymbol{a} \cdot \boldsymbol{a} = 0 \Rightarrow \boldsymbol{a} = \mathbf{0}$.

□

**Ejercicio 60.** Supongamos primero que tanto $A$ como $B$ contienen al menos un vector no nulo. Entonces, sea $\mathbf{A}$ una matriz cuyas $k$ columnas son vectores de $A$ que forman una base de $\mathcal{L}\big(A\big)$ y sea $\mathbf{B}$ otra matriz cuyas $p$ columnas son vectores de $B$ que forman una base de $\mathcal{L}\big(B\big)$. Entonces, para todo $\boldsymbol{a} \in \mathcal{L}\big(A\big)$, existe un único $\boldsymbol{x} \in \mathbb{R}^k$ tal que $\boldsymbol{a} = \mathbf{A}\boldsymbol{x}$, y para todo $\boldsymbol{b} \in \mathcal{L}\big(B\big)$ existe un único $\boldsymbol{y} \in \mathbb{R}^p$ tal que $\boldsymbol{b} = \mathbf{B}\boldsymbol{y}$.

Veamos que $\boxed{A \perp B \;\Rightarrow\; \mathcal{L}\big(A\big) \perp \mathcal{L}\big(B\big)}$.

$A \perp B \quad \Rightarrow \quad \mathbf{A}^\intercal\mathbf{B} = \mathbf{0} \quad \Rightarrow$ Para todo $\boldsymbol{a} \in \mathcal{L}\big(A\big)$ y $\boldsymbol{b} \in \mathcal{L}\big(B\big)$ tenemos que $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{x}\mathbf{A}^\intercal\mathbf{B}\boldsymbol{y} = \boldsymbol{a}\mathbf{0}\boldsymbol{y} = 0$.

Veamos que $\boxed{\mathcal{L}\big(A\big) \perp \mathcal{L}\big(B\big) \;\Rightarrow\; A \perp B}$.

$$\mathcal{L}\big(A\big) \perp \mathcal{L}\big(B\big) \quad \Rightarrow \quad \text{para todo } \boldsymbol{a} \in \mathcal{L}\big(A\big) \text{ y } \boldsymbol{b} \in \mathcal{L}\big(B\big),\ \boldsymbol{a} \cdot \boldsymbol{b} = 0.$$

En particular, también para todo $\boldsymbol{a} \in A$ y $\boldsymbol{b} \in B$.

Ahora veamos qué ocurre cuando alguno de los conjuntos es vacío (o solo contiene el vector cero). Como $A \perp B$ implica que todo vector en $A$ es perpendicular a todo vector en $B$. Entonces $A \perp B$ es falso solo cuando existe algún vector $\boldsymbol{a}$ en $A$ y algún vector $\boldsymbol{b}$ en $B$ que no son perpendiculares ($\boldsymbol{a} \cdot \boldsymbol{b} \neq 0$). Por lo

tanto, si se da el caso de que $A$ es vacío (o $\{\mathbf{0}\}$), tanto $A$ como $\mathcal{L}\big(A\big) = \{\mathbf{0}\}$ son ortogonales a cualquier otro subconjunto de vectores (pues no existe tal par de vectores $\boldsymbol{a}$ y $\boldsymbol{b}$). Lo mismo ocurre si $B$ es vacío (o $\{\mathbf{0}\}$).

□

**Ejercicio 61.** Basta comprobar que todas las propiedades del producto punto (Página 2.1) también se verifican si se divide el producto punto por $n$.

□

**Ejercicio 62.** Hay que demostrar que se verifican todos los axiomas. Denotemos $\eta(\boldsymbol{a}, \boldsymbol{b})$ como $\big\langle \boldsymbol{a} \big| \boldsymbol{b} \big\rangle_\eta$,

- **Simetría:** $\eta(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{a}\mathbf{D}\boldsymbol{b} = \boldsymbol{b}(\mathbf{D}^\intercal)\boldsymbol{a} = \boldsymbol{b}\mathbf{D}\boldsymbol{a} = \eta(\boldsymbol{b}, \boldsymbol{a})$, pues $\mathbf{D}$ es simétrica.
- **Linealidad respecto al primer argumento:**
  1. $\eta(\alpha\boldsymbol{a}, \boldsymbol{b}) = (\alpha\boldsymbol{a})\mathbf{D}\boldsymbol{b} = \alpha\big(\boldsymbol{a}\mathbf{D}\boldsymbol{b}\big) = \alpha\eta(\boldsymbol{a}, \boldsymbol{b})$.
  2. $\eta((\boldsymbol{a} + \boldsymbol{b}), \boldsymbol{c}) = (\boldsymbol{a} + \boldsymbol{b})\mathbf{D}\boldsymbol{c} = \boldsymbol{a}\mathbf{D}\boldsymbol{c} + \boldsymbol{b}\mathbf{D}\boldsymbol{c} = \eta(\boldsymbol{a}, \boldsymbol{c}) + \eta(\boldsymbol{b}, \boldsymbol{c})$.
- **Positivo:** $\eta(\boldsymbol{a}, \boldsymbol{a}) = \boldsymbol{a}\mathbf{D}\boldsymbol{a} = \sum_{i=1}^{n} a_i d_{ii} a_i = \sum_{i=1}^{n} d_{ii} a_i^2 \geq 0$, ya que $d_{ii} > 0$ y $a_i^2 \geq 0$.
- **Definido:** $\eta(\boldsymbol{a}, \boldsymbol{a}) = \sum_{i=1}^{n} d_{ii} a_i^2 = 0$, entonces $a_i^2 = 0$ ya que $d_{ii} > 0$, por tanto $\boldsymbol{a} = \mathbf{0}$.

□

**Ejercicio 63.** El vector nulo: $\vec{0}$.

□

**Ejercicio 64.** El vector nulo: $\vec{x} = \vec{0}$.

□

**Ejercicio 65.** Si $\vec{x}$ e $\vec{y}$ son *perpendiculares*, entonces $\big\langle \vec{x} \big| \vec{y} \big\rangle = 0$ y

$$\begin{aligned}
\|\vec{x} + \vec{y}\|^2 = \big\langle (\vec{x} + \vec{y}) \big| (\vec{x} + \vec{y}) \big\rangle &= \big\langle \vec{x} \big| (\vec{x} + \vec{y}) \big\rangle + \big\langle \vec{y} \big| (\vec{x} + \vec{y}) \big\rangle \\
&= \big\langle (\vec{x} + \vec{y}) \big| \vec{x} \big\rangle + \big\langle (\vec{x} + \vec{y}) \big| \vec{y} \big\rangle \\
&= \big\langle \vec{x} \big| \vec{x} \big\rangle + \big\langle \vec{y} \big| \vec{x} \big\rangle + \big\langle \vec{x} \big| \vec{y} \big\rangle + \big\langle \vec{y} \big| \vec{y} \big\rangle \\
&= \big\langle \vec{x} \big| \vec{x} \big\rangle + \quad 2\big(\big\langle \vec{x} \big| \vec{y} \big\rangle\big) \quad + \big\langle \vec{y} \big| \vec{y} \big\rangle \\
&= \big\langle \vec{x} \big| \vec{x} \big\rangle + \qquad 0 \qquad + \big\langle \vec{y} \big| \vec{y} \big\rangle \qquad\qquad = \|\vec{x}\|^2 + \|\vec{y}\|^2.
\end{aligned}$$

□

**Ejercicio 66.** Sean $\vec{x}$ y $\vec{y}$ vectores de $\mathcal{V}$ y sea el producto escalar $\big\langle _ \big| _ \big\rangle \colon \mathcal{V} \times \mathcal{V} \to \mathbb{R}$. Veamos que siempre se verifica que $|\big\langle \vec{x} \big| \vec{y} \big\rangle| \leq \|\vec{x}\| \|\vec{y}\|$

Caso trivial: *$\vec{x}$ y $\vec{y}$ son linealmente dependientes*

En tal caso, o bien $\vec{x} = \vec{0}$ o bien existe $\alpha \in \mathbb{R}$ tal que $\alpha\vec{x} = \vec{y}$. Si ocurre lo primero $\big\langle \vec{x} \big| \vec{y} \big\rangle = 0$ y $\|\vec{x}\| \|\vec{y}\| = 0$. Si ocurre lo segundo

$$|\big\langle \vec{x} \big| \vec{y} \big\rangle| = |\big\langle \vec{x} \big| \alpha\vec{x} \big\rangle| = |\alpha \|\vec{x}\|^2| = \|\vec{x}\| |\alpha| \|\vec{x}\| = \|\vec{x}\| \|\vec{y}\|.$$

Caso no trivial: *$\vec{x}$ e $\vec{y}$ son linealmente independientes*

En tal caso existe un vector $\vec{h}$ ortogonal a $\vec{x}$ tal que $L(\vec{x},\vec{h}) = L(\vec{x},\vec{y})$. Dicho vector lo construimos de la forma $\vec{h} = \vec{y} + \alpha\vec{x}$, con lo que garantizamos que $L(\vec{x},\vec{h}) = L(\vec{x},\vec{y})$ e imponemos que

$$0 = \langle \vec{x} | \vec{h} \rangle = \langle \vec{x} | \vec{y} + \alpha\vec{x} \rangle = \langle \vec{x} | \vec{y} \rangle + \alpha \langle \vec{x} | \vec{x} \rangle,$$

con lo que $\alpha = -\dfrac{\langle \vec{x} | \vec{y} \rangle}{\langle \vec{x} | \vec{x} \rangle}$.

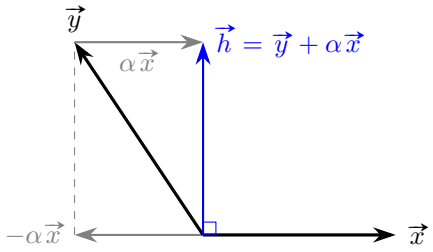



Figura A.1: Representación esquemática donde $\vec{h}$ es combinación lineal de $\vec{x}$ y $\vec{y}$ y perpendicular a $\vec{x}$.

Por tanto, tendremos que

$$\begin{aligned} \|\vec{y}\|^2 &= \langle \vec{y} | \vec{y} \rangle = \langle \vec{h} - \alpha\vec{x} | \vec{h} - \alpha\vec{x} \rangle = \langle \vec{h} | \vec{h} \rangle - 2\alpha \langle \vec{h} | \vec{x} \rangle + \alpha^2 \langle \vec{x} | \vec{x} \rangle \\ &= \|\vec{h}\|^2 + \alpha^2 \|\vec{x}\|^2 \quad > \quad \alpha^2 \|\vec{x}\|^2 = \frac{\langle \vec{x} | \vec{y} \rangle^2}{\langle \vec{x} | \vec{x} \rangle^2} \|\vec{x}\|^2 = \frac{\langle \vec{x} | \vec{y} \rangle^2}{\|\vec{x}\|^2}, \end{aligned}$$

de donde $\langle \vec{x} | \vec{y} \rangle^2 < \|\vec{x}\|^2 \|\vec{y}\|^2$ y, por consiguiente, $|\langle \vec{x} | \vec{y} \rangle| < \|\vec{x}\| \|\vec{y}\|$.

□

**Ejercicio 67.**

$$\begin{aligned} \|\vec{x} + \vec{y}\|^2 = \langle (\vec{x} + \vec{y}) | (\vec{x} + \vec{y}) \rangle &= \langle \vec{x} | \vec{x} \rangle + 2(\langle \vec{x} | \vec{y} \rangle) + \langle \vec{y} | \vec{y} \rangle \\ &= \|\vec{x}\|^2 + 2(\langle \vec{x} | \vec{y} \rangle) + \|\vec{y}\|^2 \\ &\leq \|\vec{x}\|^2 + 2\|\vec{x}\| \|\vec{y}\| + \|\vec{y}\|^2 \qquad \text{puesto que } |\langle \vec{x} | \vec{y} \rangle| \leq \|\vec{x}\| \|\vec{y}\| \\ &= (\|\vec{x}\| + \|\vec{y}\|)^2 \end{aligned}$$

Consecuentemente, $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$.

□

**Ejercicio 68.** Evidentemente $\vec{0}$ pertenece a $A^\perp$, así que el subconjunto $A^\perp$ no es vacío.

Sean $\vec{c}_1, \vec{c}_2 \in A^\perp$ y $x, y \in \mathbb{R}$. Entonces $\langle \vec{c}_1 | \vec{a} \rangle = 0$ y $\langle \vec{c}_2 | \vec{a} \rangle = 0$, para cualquier $\vec{a} \in A$. Consecuentemente

$$\langle (x\vec{c}_1 + y\vec{c}_2) | \vec{a} \rangle = x(\langle \vec{c}_1 | \vec{a} \rangle) + x(\langle \vec{c}_2 | \vec{a} \rangle) = x \cdot 0 + y \cdot 0 = 0.$$

y, por tanto, $(x\vec{c}_1 + y\vec{c}_2) \in A^\perp$. Es decir, $A^\perp$ es cerrado para las combinaciones lineales.

□

**Ejercicio 69(a)** $\vec{x} \in B^\perp$ si y solo si $\langle \vec{x} | \vec{b} \rangle = 0$ para cada $\vec{b} \in B$; (y como $A \subset B$) en particular también $\langle \vec{x} | \vec{b} \rangle = 0$ para cada $\vec{a} \in A$, es decir, $\vec{x} \in A^\perp$.

□

**Ejercicio 69(b)** Tenemos que demostrar la inclusión de estos subespacios en ambos sentidos.

*Inclusión en un sentido*: como $A \subset \mathcal{L}(A)$, por el apartado anterior sabemos que $\boxed{\mathcal{L}(A)^\perp \subset A^\perp}$.

*Inclusión en el otro sentido*: sabemos que $\vec{x} \in A^\perp$ si y solo si $\langle \vec{x} | \vec{a} \rangle = 0$ para cada $\vec{a} \in A$. Sea $\vec{z} \in \mathcal{L}(A)$, entonces existen $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ y $\vec{a}_1 \ldots \vec{a}_k \in A$ tales que $\vec{z} = \alpha_1 \vec{a}_1 + \cdots + \alpha_k \vec{a}_k$. Entonces $\langle \vec{x} | \vec{z} \rangle = \langle \vec{x} | (\alpha_1 \vec{a}_1 + \cdots + \alpha_k \vec{a}_k) \rangle = \alpha_1 \langle \vec{x} | \vec{a}_1 \rangle + \cdots + \alpha_k \langle \vec{x} | \vec{a}_k \rangle = \alpha_1 0 + \cdots + \alpha_k 0 = 0$; es decir, $\vec{x} \in \mathcal{L}(A)^\perp$. Por tanto $\boxed{A^\perp \subset \mathcal{L}(A)^\perp}$.

□

**Ejercicio 69(c)** Sea $\vec{a} \in \mathcal{A} \cap \mathcal{A}^\perp$; como $\vec{a} \in \mathcal{A}$ es perpendicular a todo vector en $\mathcal{A}^\perp$, entonces $\langle \vec{a} | \vec{a} \rangle = 0 \Rightarrow \vec{a} = \vec{0}$.

□

**Ejercicio 70.** Si tenemos un subespacio $\mathcal{A}$ de $\mathbb{R}^n$ y queremos construir la matriz $\mathbf{P}$ de la proyección ortogonal sobre $\mathcal{A}$, debemos tener en cuenta que:

1. Para que $\mathbf{P}\boldsymbol{x} \in \mathcal{A}$, las columnas de $\mathbf{P}$ deben pertenecer a $\mathcal{A}$. Luego, si $\mathbf{A}$ es una matriz cuyas columnas forman una base de $\mathcal{A}$, la matriz $\mathbf{P}$ tiene que ser de la forma $\mathbf{AB}$, donde $\mathbf{B}$ es una matriz a determinar.
2. Para que $(\mathbf{P}\boldsymbol{x} - \boldsymbol{x}) \perp \mathcal{A}$, tiene que ocurrir que las columnas de $\mathbf{P} - \mathbf{I}$ sean ortogonales a las columnas de $\mathbf{A}$. Dicho de otro modo
$$(\mathbf{P} - \mathbf{I})^\intercal \mathbf{A} = \mathbf{0}$$

Si lo juntamos todo, tenemos que

$$\mathbf{0} = (\mathbf{P} - \mathbf{I})^\intercal \mathbf{A} = (\mathbf{AB} - \mathbf{I})^\intercal \mathbf{A} = (\mathbf{B}^\intercal)(\mathbf{A}^\intercal)\mathbf{A} - \mathbf{A}$$

es decir, $\mathbf{B}^\intercal(\mathbf{A}^\intercal\mathbf{A}) = \mathbf{A}$. Pero $\mathbf{A}^\intercal\mathbf{A}$ es la matriz del producto escalar del subespacio $\mathcal{A}$ respecto a la base formada por las columnas de $\mathbf{A}$, por lo que es simétrica e invertible. Por tanto

$$\mathbf{B}^\intercal = \mathbf{A}(\mathbf{A}^\intercal\mathbf{A})^{-1} \quad \text{es decir,} \quad \mathbf{B} = (\mathbf{A}^\intercal\mathbf{A})^{-1}\mathbf{A}^\intercal;$$

con lo que

$$\mathbf{P} = \mathbf{AB} = \mathbf{A}(\mathbf{A}^\intercal\mathbf{A})^{-1}\mathbf{A}^\intercal.$$

□

**Ejercicio 71.** $(\mathbf{I} - \mathbf{P})^\intercal = \mathbf{I}^\intercal - \mathbf{P}^\intercal = \mathbf{I} - \mathbf{P}$.

□

**Ejercicio 72.** $(\mathbf{I} - \mathbf{P})^2 = (\mathbf{I} - \mathbf{P})(\mathbf{I} - \mathbf{P}) = \mathbf{I} - \mathbf{P} - \mathbf{P} + \mathbf{P}^2 = \mathbf{I} - \mathbf{P} - \mathbf{P} + \mathbf{P} = \mathbf{I} - \mathbf{P}$.

□

**Ejercicio 73(a)** Por una parte, $\mathcal{C}(\mathbf{P}) \subset \mathcal{C}(\mathbf{A})$, pues para todo $\boldsymbol{z} \in \mathcal{C}(\mathbf{P})$

$$\text{existe } \boldsymbol{y} \in \mathbb{R}^m \text{ tal que } \quad \boldsymbol{z} = \mathbf{P}\boldsymbol{y} = \mathbf{A}(\mathbf{A}^\intercal\mathbf{A})^{-1}\mathbf{A}^\intercal\boldsymbol{y} = \mathbf{A}\widehat{\boldsymbol{x}} \in \mathcal{C}(\mathbf{A}), \quad \text{donde } \widehat{\boldsymbol{x}} = (\mathbf{A}^\intercal\mathbf{A})^{-1}\mathbf{A}^\intercal\boldsymbol{y},$$

y por otra parte, $\mathcal{C}(\mathbf{A}) \subset \mathcal{C}(\mathbf{P})$, pues para todo $\boldsymbol{z} \in \mathcal{C}(\mathbf{A})$

$$\text{existe } \boldsymbol{x} \in \mathbb{R}^n\text{:} \quad \boldsymbol{z} = \mathbf{A}\boldsymbol{x} = \mathbf{A}\underbrace{\left(\mathbf{A}^\intercal\mathbf{A}\right)^{-1}\mathbf{A}^\intercal\mathbf{A}}_{\mathbf{I}}\boldsymbol{x} = \mathbf{P}\mathbf{A}\boldsymbol{x} = \mathbf{P}\boldsymbol{z} \in \mathcal{C}(\mathbf{P}), \quad \text{donde } \boldsymbol{z} = \mathbf{A}\boldsymbol{x}.$$

□

**Ejercicio 73(b)** Por una parte, $\mathcal{C}(\mathbf{M}) \subset \mathcal{N}\left(\mathbf{A}^\intercal\right)$, pues para todo $\boldsymbol{z} \in \mathcal{C}(\mathbf{M})$ existe $\boldsymbol{y} \in \mathbb{R}^m$ tal que

$$\boldsymbol{z} = \mathbf{M}\boldsymbol{y} = (\mathbf{I} - \mathbf{P})\boldsymbol{y} = \boldsymbol{y} - \mathbf{P}\boldsymbol{y} \ \in \mathcal{C}(\mathbf{A})^\perp = \mathcal{N}\left(\mathbf{A}^\intercal\right)$$

por ser $\mathbf{P}\boldsymbol{y}$ la proyección ortogonal sobre $\mathcal{C}(\mathbf{A})$. Por otra parte, $\mathcal{N}\left(\mathbf{A}^\intercal\right) \subset \mathcal{C}(\mathbf{M})$, pues para todo $\boldsymbol{y} \in \mathcal{N}\left(\mathbf{A}^\intercal\right)$ sabemos que $\boldsymbol{y} \perp \mathcal{C}(\mathbf{A})$, y por tanto $\mathbf{P}\boldsymbol{y} = \mathbf{0}$. Así pues

$$\boldsymbol{y} = \boldsymbol{y} + \mathbf{0} = \boldsymbol{y} + \mathbf{P}\boldsymbol{y} = (\mathbf{I} + \mathbf{P})\boldsymbol{y} = \mathbf{M}\boldsymbol{y} \quad \Rightarrow \quad \boldsymbol{y} \in \mathcal{C}(\mathbf{M}).$$

□

**Ejercicio 74.** Puesto que $\boldsymbol{x}\mathbf{B} = \mathbf{B}^\intercal\boldsymbol{x} = \mathbf{B}\boldsymbol{x}$, tenemos que $\langle \boldsymbol{x} | \mathbf{B}\boldsymbol{y} \rangle = \boldsymbol{x}\mathbf{B}\boldsymbol{y} = \langle \boldsymbol{x}\mathbf{B} | \boldsymbol{y} \rangle = \langle \mathbf{B}\boldsymbol{x} | \boldsymbol{y} \rangle$.

□

**Ejercicio 75(a)** Por ser autoadjunta e idempotente: $\langle f(\vec{u}) | f(\vec{w}) \rangle = \langle f\big(f(\vec{u})\big) | \vec{w} \rangle = \langle f(\vec{u}) | \vec{w} \rangle$.

□

**Ejercicio 75(b)** Para la demostración usaremos el resultado del apartado (a). Es sencillo pero pesado.

Por una parte tenemos que demostrar que $f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v})$, y lo haremos demostrando que la diferencia es $\vec{0}$, pues su norma es cero, es decir, $\|f(\vec{u} + \vec{v}) - f(\vec{u}) - f(\vec{v})\|^2 = 0$:

$$\begin{aligned}
\|f(\vec{u} + \vec{v}) - f(\vec{u}) - f(\vec{v})\|^2 =& \langle f(\vec{u} + \vec{v}) - f(\vec{u}) - f(\vec{v}) | f(\vec{u} + \vec{v}) - f(\vec{u}) - f(\vec{v}) \rangle \\
=& \langle f(\vec{u} + \vec{v}) | f(\vec{u} + \vec{v}) \rangle - \langle f(\vec{u} + \vec{v}) | f(\vec{u}) \rangle - \langle f(\vec{u} + \vec{v}) | f(\vec{v}) \rangle \\
& - \langle f(\vec{u}) | f(\vec{u} + \vec{v}) \rangle + \langle f(\vec{u}) | f(\vec{u}) \rangle + \langle f(\vec{u}) | f(\vec{v}) \rangle \\
& - \langle f(\vec{v}) | f(\vec{u} + \vec{v}) \rangle + \langle f(\vec{v}) | f(\vec{u}) \rangle + \langle f(\vec{v}) | f(\vec{v}) \rangle \quad \text{(desarrollando)} \\
=& \langle f(\vec{u} + \vec{v}) | \vec{u} + \vec{v} \rangle - \langle f(\vec{u} + \vec{v}) | \vec{u} \rangle - \langle f(\vec{u} + \vec{v}) | \vec{v} \rangle \\
& - \langle f(\vec{u}) | \vec{u} + \vec{v} \rangle + \langle f(\vec{u}) | \vec{u} \rangle + \langle f(\vec{u}) | \vec{v} \rangle \\
& - \langle f(\vec{v}) | \vec{u} + \vec{v} \rangle + \langle f(\vec{v}) | \vec{u} \rangle + \langle f(\vec{v}) | \vec{v} \rangle \quad \text{(por (a))} \\
=& \langle f(\vec{u} + \vec{v}) | \vec{u} \rangle + \langle f(\vec{u} + \vec{v}) | \vec{v} \rangle - \langle f(\vec{u} + \vec{v}) | \vec{u} \rangle - \langle f(\vec{u} + \vec{v}) | \vec{v} \rangle \\
& - \langle f(\vec{u}) | \vec{u} \rangle - \langle f(\vec{u}) | \vec{v} \rangle + \langle f(\vec{u}) | \vec{u} \rangle + \langle f(\vec{u}) | \vec{v} \rangle \\
& - \langle f(\vec{v}) | \vec{u} \rangle - \langle f(\vec{v}) | \vec{v} \rangle + \langle f(\vec{v}) | \vec{u} \rangle + \langle f(\vec{v}) | \vec{v} \rangle \quad \text{(desarrollando)} \\
=& 0 \qquad \text{(cancelando términos).}
\end{aligned}$$

Y, por otra parte, tenemos que demostrar que $f(\alpha\vec{u}) = \alpha f(\vec{u})$; y seguiremos una estrategia similar, es decir,

demostraremos que $\|f(\alpha\vec{u}) - \alpha f(\vec{u})\|^2 = 0$.

$$\begin{aligned}
\|f(\alpha\vec{u}) - \alpha f(\vec{u})\|^2 =& \big\langle f(\alpha\vec{u}) - \alpha f(\vec{u}) \big| f(\alpha\vec{u}) - \alpha f(\vec{u}) \big\rangle \\
=& \big\langle f(\alpha\vec{u}) \big| f(\alpha\vec{u}) \big\rangle - \big\langle f(\alpha\vec{u}) \big| \alpha f(\vec{u}) \big\rangle - \big\langle \alpha f(\vec{u}) \big| f(\alpha\vec{u}) \big\rangle + \big\langle \alpha f(\vec{u}) \big| \alpha f(\vec{u}) \big\rangle \\
=& \big\langle f(\alpha\vec{u}) \big| \alpha\vec{u} \big\rangle - \big\langle f(\alpha\vec{u}) \big| \alpha f(\vec{u}) \big\rangle - \big\langle \alpha f(\vec{u}) \big| f(\alpha\vec{u}) \big\rangle + \big\langle \alpha f(\vec{u}) \big| \alpha f(\vec{u}) \big\rangle \quad \text{(por (a))} \\
=& \alpha \big\langle f(\alpha\vec{u}) \big| \vec{u} \big\rangle - \alpha \big\langle f(\alpha\vec{u}) \big| f(\vec{u}) \big\rangle - \alpha \big\langle f(\vec{u}) \big| f(\alpha\vec{u}) \big\rangle + \alpha^2 \big\langle f(\vec{u}) \big| f(\vec{u}) \big\rangle \\
=& \alpha \big\langle f(\alpha\vec{u}) \big| \vec{u} \big\rangle - \alpha \big\langle f(\alpha\vec{u}) \big| \vec{u} \big\rangle - \alpha \big\langle f(\vec{u}) \big| \alpha\vec{u} \big\rangle + \alpha^2 \big\langle f(\vec{u}) \big| \vec{u} \big\rangle \quad \text{(por (a))} \\
=& \alpha \big\langle f(\alpha\vec{u}) \big| \vec{u} \big\rangle - \alpha \big\langle f(\alpha\vec{u}) \big| \vec{u} \big\rangle - \alpha^2 \big\langle f(\vec{u}) \big| \vec{u} \big\rangle + \alpha^2 \big\langle f(\vec{u}) \big| \vec{u} \big\rangle \\
=& 0 \quad \text{(cancelando términos).}
\end{aligned}$$

□

**Ejercicio 76.** Sea $\vec{p}_y$ la proyección de $\vec{y}$ sobre $\mathcal{H}$ y tomemos un vector $\vec{v}$ cualquiera de $\mathcal{H}$. Veamos que $\vec{y}$ está más lejos de $\vec{v}$ que de su proyección ortogonal $\vec{p}_y$.

Como $\vec{v}$ y $\vec{p}_y$ están en $\mathcal{H}$, su diferencia $(\vec{p}_y - \vec{x})$ está en $\mathcal{H}$ (pues $\mathcal{H}$ es un subespacio) y es ortogonal a $(\vec{y} - \vec{p}_y)$ (por ser $\vec{p}_y$ la proyección ortogonal sobre $\mathcal{H}$). Y como la suma de ambos vectores perpendiculares es $(\vec{y} - \vec{p}_y) + (\vec{p}_y - \vec{x}) = (\vec{y} - \vec{x})$, por el Teorema de Pitágoras, concluimos que

$$\|\vec{y} - \vec{x}\|^2 = \|\vec{y} - \vec{p}_y\|^2 + \|\vec{p}_y - \vec{x}\|^2 \geq \|\vec{y} - \vec{p}_y\|^2.$$

Por tanto, $\boxed{\|\vec{y} - \vec{x}\| \geq \|\vec{y} - \vec{p}_y\|}$ para todo $\vec{x} \in \mathcal{H}$.

□

**Ejercicio 77(a)** Puesto que $\mathbf{P}\boldsymbol{y} = \boldsymbol{y}\mathbf{P}$ (por ser $\mathbf{P}$ simétrica):

$$(\boldsymbol{y} - \mathbf{P}\boldsymbol{y})\mathbf{A} = (\boldsymbol{y} - \boldsymbol{y}\mathbf{P})\mathbf{A} = \boldsymbol{y}\mathbf{A} - \boldsymbol{y}\mathbf{A}\big(\mathbf{A}^\intercal\mathbf{A}\big)^{-1}\mathbf{A}^\intercal\mathbf{A} = \boldsymbol{y}\mathbf{A} - \boldsymbol{y}\mathbf{A} = \mathbf{0}.$$

□

**Ejercicio 77(b)** Si $\boldsymbol{y} = \mathbf{A}\boldsymbol{x}$ para algún $\boldsymbol{x} \in \mathbb{R}^n$:

$$f_{\mathbf{P}}(\boldsymbol{y}) = \mathbf{P}\boldsymbol{y} = \mathbf{P}\mathbf{A}\boldsymbol{x} = \mathbf{A}\big(\mathbf{A}^\intercal\mathbf{A}\big)^{-1}\mathbf{A}^\intercal\mathbf{A}\boldsymbol{x} = \mathbf{A}\boldsymbol{x} = \boldsymbol{y};$$

□

**Ejercicio 77(c)** Si $\mathbf{A}^\intercal\boldsymbol{y} = \mathbf{0}$, entonces

$$f_{\mathbf{P}}(\boldsymbol{y}) = \mathbf{P}\boldsymbol{y} = \mathbf{A}\big(\mathbf{A}^\intercal\mathbf{A}\big)^{-1}\mathbf{A}^\intercal\boldsymbol{y} = \mathbf{A}\big(\mathbf{A}^\intercal\mathbf{A}\big)^{-1}\mathbf{0} = \mathbf{0}.$$

□

**Ejercicio 77(d)** Como $\mathcal{C}(\mathbf{A})$ es el complemento ortogonal de $\mathcal{N}\left(\mathbf{A}^\intercal\right)$, demostrar que para todo $\boldsymbol{y} \in \mathbb{R}^m$ la diferencia $(\boldsymbol{y} - f_{\mathbf{M}}(\boldsymbol{y}))$ es perpendicular a $\mathcal{N}\left(\mathbf{A}^\intercal\right)$ es lo mismo que demostrar que dicha diferencia pertenece a $\mathcal{C}(\mathbf{A})$; así

$$\boldsymbol{y} - f_{\mathbf{M}}(\boldsymbol{y}) = \boldsymbol{y} - \big(\boldsymbol{y} - f_{\mathbf{P}}(\boldsymbol{y})\big) = f_{\mathbf{P}}(\boldsymbol{y}) \in \mathcal{C}(\mathbf{A}).$$

□

**Ejercicio 77(e)** Por una parte, $f_{\mathbf{M}}$ actúa como la función identidad con todo $\boldsymbol{y} \in \mathcal{N}\left(\mathbf{A}^{\intercal}\right)$, pues en tal caso $f_{\mathbf{P}}(\boldsymbol{y}) = \mathbf{0}$, y por tanto:

$$f_{\mathbf{M}}(\boldsymbol{y}) = \boldsymbol{y} - f_{\mathbf{P}}(\boldsymbol{y}) = \boldsymbol{y} - \mathbf{0} = \boldsymbol{y}.$$

Por otra parte, $f_{\mathbf{M}}$ actúa como la función nula con todo $\boldsymbol{y} \in \mathcal{C}\left(\mathbf{A}\right)$, pues en tal caso $f_{\mathbf{P}}(\boldsymbol{y}) = \boldsymbol{y}$, y por tanto:

$$f_{\mathbf{M}}(\boldsymbol{y}) = \boldsymbol{y} - f_{\mathbf{P}}(\boldsymbol{y}) = \boldsymbol{y} - \boldsymbol{y} = \mathbf{0}.$$

□

**Ejercicio 78(a)** Debemos demostrar que cada uno de los conjuntos está contenido en el otro.

$\boxed{\mathcal{N}\left(\mathbf{A}\right) \subset \mathcal{N}\left(\mathbf{A}^{\intercal}\mathbf{A}\right)}$: Si $\boldsymbol{x} \in \mathcal{N}\left(\mathbf{A}\right)$, entonces $\mathbf{A}\boldsymbol{x} = \mathbf{0}$. Multiplicando ambos lados de la igualdad por $\mathbf{A}^{\intercal}$, tenemos $\mathbf{A}^{\intercal}\mathbf{A}\boldsymbol{x} = \mathbf{A}^{\intercal}\mathbf{0} = \mathbf{0}$; y por tanto $\boldsymbol{x} \in \mathcal{N}\left(\mathbf{A}^{\intercal}\mathbf{A}\right)$.

$\boxed{\mathcal{N}\left(\mathbf{A}^{\intercal}\mathbf{A}\right) \subset \mathcal{N}\left(\mathbf{A}\right)}$: Si $\boldsymbol{x} \in \mathcal{N}\left(\mathbf{A}^{\intercal}\mathbf{A}\right)$, entonces $\mathbf{A}^{\intercal}\mathbf{A}\boldsymbol{x} = \mathbf{0}$. Multiplicando por $\boldsymbol{x}$, tenemos que $\boldsymbol{x}\mathbf{A}^{\intercal}\mathbf{A}\boldsymbol{x} = \boldsymbol{x} \cdot \mathbf{0} = 0$. Por tanto $\mathbf{A}\boldsymbol{x}$ tiene norma cero, ya que $\boldsymbol{x}\mathbf{A}^{\intercal}\mathbf{A}\boldsymbol{x} = (\mathbf{A}\boldsymbol{x}) \cdot (\mathbf{A}\boldsymbol{x}) = \|\mathbf{A}\boldsymbol{x}\|^2$. Como el único vector de norma cero es $\mathbf{0}$, tenemos que $\mathbf{A}\boldsymbol{x} = \mathbf{0}$; es decir, $\boldsymbol{x} \in \mathcal{N}\left(\mathbf{A}\right)$.

□

**Ejercicio 78(b)** Dado que $\mathbf{A}^{\intercal}\mathbf{A}$ y $\mathbf{A}$ tienen $n$ columnas e idéntico espacio nulo $\mathcal{N}\left(\mathbf{A}^{\intercal}\mathbf{A}\right) = \mathcal{N}\left(\mathbf{A}\right)$, ambas matrices tienen el mismo rango: $k = n - \dim \mathcal{N}\left(\mathbf{A}\right)$.

□

**Ejercicio 78(c)** Sea $\mathbf{A}$ de rango $k$. El sistema $(\mathbf{A}^{\intercal}\mathbf{A})\boldsymbol{x} = \mathbf{A}^{\intercal}\boldsymbol{b}$ tiene solución si el rango de la matriz de coeficientes $\mathbf{A}^{\intercal}\mathbf{A}$ es igual que el rango de la matriz ampliada $\left[\mathbf{A}^{\intercal}\mathbf{A} \,\middle|\, \mathbf{A}^{\intercal}\boldsymbol{b}\right]$; es decir, si $\mathbf{A}^{\intercal}\boldsymbol{b}$ es combinación lineal de las columnas de $\mathbf{A}^{\intercal}\mathbf{A}$. Veámoslo:

Como al añadir una columna a una matriz, el rango (el número de pivotes) no puede disminuir, sabemos que $\operatorname{rg}\left(\left[\mathbf{A}^{\intercal}\mathbf{A} \,\middle|\, \mathbf{A}^{\intercal}\boldsymbol{b}\right]\right) \geq \operatorname{rg}\left(\mathbf{A}^{\intercal}\mathbf{A}\right) = \operatorname{rg}\left(\mathbf{A}\right) = k$ (por el Corolario 13.5.2).

Por otra parte, $\left[\mathbf{A}^{\intercal}\mathbf{A} \,\middle|\, \mathbf{A}^{\intercal}\boldsymbol{b}\right] = \mathbf{A}^{\intercal}\left[\mathbf{A} \,\middle|\, \boldsymbol{b}\right]$ es el producto de $\mathbf{A}^{\intercal}$ por la matriz ampliada $\left[\mathbf{A} \,\middle|\, \boldsymbol{b}\right]$, y por la Proposición 5.2.6 en la página 81 sabemos que $\operatorname{rg}\left(\mathbf{A}^{\intercal}\left[\mathbf{A} \,\middle|\, \boldsymbol{b}\right]\right) \leq \operatorname{rg}\left(\mathbf{A}^{\intercal}\right) = \operatorname{rg}\left(\mathbf{A}\right) = k$.

Así, como por una parte $\operatorname{rg}\left(\left[\mathbf{A}^{\intercal}\mathbf{A} \,\middle|\, \mathbf{A}^{\intercal}\boldsymbol{b}\right]\right) \geq k$ y por la otra $\operatorname{rg}\left(\left[\mathbf{A}^{\intercal}\mathbf{A} \,\middle|\, \mathbf{A}^{\intercal}\boldsymbol{b}\right]\right) \leq k$, se deduce que $\operatorname{rg}\left(\left[\mathbf{A}^{\intercal}\mathbf{A} \,\middle|\, \mathbf{A}^{\intercal}\boldsymbol{b}\right]\right) = k$; que es el rango de la matriz de coeficientes $\mathbf{A}^{\intercal}\mathbf{A}$ del sistema $\mathbf{A}^{\intercal}\mathbf{A}\boldsymbol{x} = \mathbf{A}^{\intercal}\boldsymbol{b}$, y por tanto el sistema siempre tiene solución, es decir, para todo $\boldsymbol{b} \in \mathbb{R}^m$ y toda $\underset{m\times n}{\mathbf{A}}$ se verifica que $\mathbf{A}^{\intercal}\boldsymbol{b} \in \mathcal{C}\left(\mathbf{A}^{\intercal}\mathbf{A}\right)$.

□

**Ejercicio 79.** Dado que $\mathbf{Q}$ es de rango $r$, existen matrices $\mathbf{A} = {}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_i}\mathbf{I}$ y $\mathbf{B} = \mathbf{I}_{\boldsymbol{\tau}_{i+1} \cdots \boldsymbol{\tau}_j}$ de orden $n$ tales que

$$\mathbf{AQB} = \begin{bmatrix} \underset{r\times r}{\mathbf{I}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \text{y por tanto} \quad \mathbf{Q} = \mathbf{A}^{-1}\begin{bmatrix} \underset{r\times r}{\mathbf{I}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{B}^{-1} = \mathbf{A}^{-1}\begin{bmatrix} \underset{r\times r}{\mathbf{I}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \underset{r\times r}{\mathbf{I}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{B}^{-1} = \left({}_{|(1:r)}\mathbf{A}^{-1}\right)\left({}_{(1:r)|}\mathbf{B}^{-1}\right);$$

donde las dos últimas matrices son de rango $r$ y tienen órdenes $n \times r$ y $r \times n$ respectivamente. Consecuentemente, existen $\mathbf{C}$ y $\mathbf{D}$ invertibles y de orden $r$, tales que

$$\mathbf{C}\left({}_{|(1:r)}\mathbf{A}^{-1}\right) = \underset{r\times r}{\mathbf{I}} \qquad \text{y} \qquad \left({}_{(1:r)|}\mathbf{B}^{-1}\right)\mathbf{D} = \underset{r\times r}{\mathbf{I}}\,.$$

Como $\mathbf{QQ} = \mathbf{Q}$, si denotamos $\left(\mathbf{A}^{-1}{}_{|(1:r)}\right)$ con $\mathbf{A}_1$ y $\left({}_{(1:r)|}\mathbf{B}^{-1}\right)$ con $\mathbf{B}_1$, tenemos que $\mathbf{A}_1\mathbf{B}_1 = \mathbf{A}_1\mathbf{B}_1\mathbf{A}_1\mathbf{B}_1$. Y, multiplicando por $\mathbf{C}$ y $\mathbf{D}$ a izquierda y derecha respectivamente, también tenemos que $\underset{r\times r}{\mathbf{I}} = \mathbf{B}_1\mathbf{A}_1$, pues, como $\mathbf{CA}_1 = \underset{r\times r}{\mathbf{I}} = \mathbf{B}_1\mathbf{D}$,

$$\mathbf{CA}_1\mathbf{B}_1\mathbf{D} \;=\; \mathbf{CA}_1\mathbf{B}_1\mathbf{A}_1\mathbf{B}_1\mathbf{D} \quad\Longrightarrow\quad \underset{r\times r}{\mathbf{I}} \;=\; \mathbf{B}_1\mathbf{A}_1;$$

Así que, usando la Ecuación 3.9 en la página 45: $\operatorname{tr}(\mathbf{Q}) = \operatorname{tr}(\mathbf{A}_1\mathbf{B}_1) = \operatorname{tr}(\mathbf{B}_1\mathbf{A}_1) = \operatorname{tr}\left(\underset{r\times r}{\mathbf{I}}\right) = r = \operatorname{rg}(\mathbf{Q})$.

□

**Ejercicio 80.** El vector cero $\mathbf{0}$ de $\mathbb{R}^n$ es múltiplo de cualquier otro vector $\boldsymbol{x}$ de $\mathbb{R}^n$, pues $\mathbf{0} = 0\cdot\boldsymbol{x}$; y por tanto, por la propiedad ***P-3*** el determinante de $\mathbf{A}$ es cero:

$$\det\left[\mathbf{A}_{|1};\dots\ \mathbf{0};\dots\ \mathbf{A}_{|n};\right] = \det\left[\mathbf{A}_{|1};\dots\ 0\boldsymbol{x};\dots\ \mathbf{A}_{|n};\right] = 0\cdot\det\left[\mathbf{A}_{|1};\dots\ \boldsymbol{x};\dots\ \mathbf{A}_{|n};\right] = 0$$

□

**Ejercicio 81(a)** Como $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1}\right)\cdots\left(\mathbf{I}_{\boldsymbol{\tau}_k}\right)$, aplicando repetidamente (14.2) tenemos

$$\begin{aligned}
\det\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) =& |\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1}\right)\cdots\left(\mathbf{I}_{\boldsymbol{\tau}_k}\right)| \\
=& |\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1}\right)\cdots\left(\mathbf{I}_{\boldsymbol{\tau}_{(k-1)}}\right)|\cdot|\mathbf{I}_{\boldsymbol{\tau}_k}| \\
=& |\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1}\right)\cdots\left(\mathbf{I}_{\boldsymbol{\tau}_{(k-2)}}\right)|\cdot|\mathbf{I}_{\boldsymbol{\tau}_{(k-1)}}|\cdot|\mathbf{I}_{\boldsymbol{\tau}_k}| \\
& \vdots \\
=& |\mathbf{A}|\cdot|\mathbf{I}_{\boldsymbol{\tau}_1}|\cdots|\mathbf{I}_{\boldsymbol{\tau}_k}|.
\end{aligned}$$

□

**Ejercicio 81(b)**

$$|\mathbf{B}| = \det\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = |\mathbf{I}|\cdot|\mathbf{I}_{\boldsymbol{\tau}_1}|\cdots|\mathbf{I}_{\boldsymbol{\tau}_k}|.$$

Y como los determinantes de las matrices elementales son distintos de cero, necesariamente $|\mathbf{B}| \neq 0$.

□

**Ejercicio 81(c)** Si $\mathbf{B}$ es de rango completo entonces $\mathbf{B} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$. Así

$$\det(\mathbf{AB}) = \det\left(\mathbf{AI}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \det\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = |\mathbf{A}|\cdot\left(|\mathbf{I}_{\boldsymbol{\tau}_1}|\cdots|\mathbf{I}_{\boldsymbol{\tau}_k}|\right) = |\mathbf{A}|\cdot|\mathbf{B}|.$$

□

**Ejercicio 82.** Un intercambio de columnas es una sucesión de transformaciones elementales *Tipo I* y una única transformación elemental de *Tipo II* que multiplica por $-1$ una columna (Ejercicio 31 en la página 57).

□

**Ejercicio 83(a)** Puesto que ambas son matrices elementales del mismo tipo, $\det\left(\mathbf{I}_{\boldsymbol{\tau}}\right) = \det\left({}_{\boldsymbol{\tau}}\mathbf{I}\right)$.

□

**Ejercicio 83(b)** Puesto que $\mathbf{B} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} = \left(\mathbf{I}_{\boldsymbol{\tau}_1}\right) \cdots \left(\mathbf{I}_{\boldsymbol{\tau}_k}\right)$, su determinante es el producto de los determinantes $\det\left(\mathbf{I}_{\boldsymbol{\tau}_i}\right)$

$$|\mathbf{B}| = \det\left(\mathbf{I}_{\boldsymbol{\tau}_1}\right) \cdots \det\left(\mathbf{I}_{\boldsymbol{\tau}_k}\right) = \prod_{i=1}^{k} \det\left(\mathbf{I}_{\boldsymbol{\tau}_i}\right).$$

Pero también sabemos que $\mathbf{B}^\intercal = {}_{\boldsymbol{\tau}_k \cdots \boldsymbol{\tau}_1}\mathbf{I} = \left({}_{\boldsymbol{\tau}_k}\mathbf{I}\right) \cdots \left({}_{\boldsymbol{\tau}_1}\mathbf{I}\right)$, por lo que su determinante es

$$|\mathbf{B}^\intercal| \;=\; \prod_{i=1}^{k} \det\left({}_{\boldsymbol{\tau}_i}\mathbf{I}\right) \;=\; \prod_{i=1}^{k} \det\left(\mathbf{I}_{\boldsymbol{\tau}_i}\right) \;=\; |\mathbf{B}|.$$

□

**Ejercicio 84(a)** Como la matriz de orden $n$ es de rango completo, los $n$ elementos de la diagonal principal son pivotes (i.e., distintos de cero).

$$\mathbf{L} = \begin{bmatrix} *_1 & & & \\ \vdots & *_2 & & \\ \vdots & \vdots & \ddots & \\ \vdots & \vdots & \vdots & *_n \end{bmatrix} \qquad \text{donde } *_j \text{ son números distintos de cero.}$$

Como sabemos que $|\mathbf{L}| = \left|\begin{array}{c|c} \mathbf{L} & \\ \hline & 1 \end{array}\right|$, vamos a realizar una secuencia de transformaciones elementales sobre las $n$ primeras columnas de esta última matriz, pero aplicaremos transformaciones elementales sobre la última *fila* para mantener constante el valor del determinante:

Primero dividimos cada una de las $n$ primeras columnas por su correspondiente pivote $*_j$ $(j = 1 : n)$ para normalizar dichos pivotes al valor 1 (así que multiplicamos la última *fila* por cada uno de los $n$ primeros pivotes para compensar dichas transformaciones *Tipo II* y mantener constante el valor de el determinante). De esta manera logramos una matriz en la que todos los pivotes (salvo el de la última columna) son 1. Después aplicamos la eliminación de derecha a izquierda para anular todo lo que queda a la izquierda de los pivotes (como las transformaciones empleadas en este segundo paso son de *tipo I*, ahora basta con multiplicar la última fila por 1).

$$\left[\begin{array}{cccc|c} *_1 & & & & \\ \vdots & *_2 & & & \\ \vdots & \vdots & \ddots & & \\ \vdots & \vdots & \vdots & *_n & \\ \hline & & & & 1 \end{array}\right] \xrightarrow[\substack{\left[\left(\frac{1}{*_1}\right)\mathbf{1}\right] \\ \vdots \\ \left[\left(\frac{1}{*_n}\right)\mathbf{n}\right]}]{\boldsymbol{\tau}} \left[\begin{array}{cccc|c} 1 & & & & \\ \vdots & 1 & & & \\ \vdots & \vdots & \ddots & & \\ \vdots & \vdots & \vdots & 1 & \\ \hline & & & & *_1 \cdot *_2 \cdots *_n \end{array}\right] \xrightarrow[\substack{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_q \\ \text{(de Tipo I)}}]{\boldsymbol{\tau}} \left[\begin{array}{cccc|c} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ \hline & & & & *_1 \cdot *_2 \cdots *_n \end{array}\right]$$
$$[(*_1)(\mathbf{n+1})]$$
$$[(*_n)(\mathbf{n+1})]$$
$$[(1)(\mathbf{n+1})]$$

De esta manera hemos obtenido una matriz como la identidad salvo por su última columna, que está multiplicada por el producto de los pivotes de $\mathbf{L}$. Así pues, el determinante de esta última matriz es igual al producto de los pivotes de $\mathbf{L}$, y como en todo momento hemos mantenido el valor del determinante, concluimos que el determinante de una matriz triangular inferior de rango completo es igual al producto de sus pivotes; es decir, al producto de los elementos de la diagonal.

$$\det(\mathbf{L}) = \text{ producto de los elementos de la diagonal}$$

□

**Ejercicio 84(b)** Una matriz de orden $n$ y triangular solo puede ser de rango completo si los $n$ elementos de la diagonal son distintos de cero. Así, si hay algún cero en su diagonal, la matriz es singular. Consecuentemente,

su determinante es cero, es decir, es igual al producto de los elementos de la diagonal (donde uno de ellos es cero).

□

**Ejercicio 84(c)**

$$\det(\mathbf{U}) = \det\left(\mathbf{U}^\intercal\right) = \text{ producto de los elementos de la diagonal}$$

por ser $\mathbf{U}^\intercal$ triangular inferior.

□

**Ejercicio 85(a)** Puesto que $\begin{bmatrix}\mathbf{A} & \\ & \mathbf{B}\end{bmatrix} = \begin{bmatrix}\mathbf{A} & \\ & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \\ & \mathbf{B}\end{bmatrix}$ entonces $\begin{vmatrix}\mathbf{A} & \\ & \mathbf{B}\end{vmatrix} = \begin{vmatrix}\mathbf{A} & \\ & \mathbf{I}\end{vmatrix} \cdot \begin{vmatrix}\mathbf{I} & \\ & \mathbf{B}\end{vmatrix} = \left|\mathbf{A}\right| \cdot \left|\mathbf{B}\right|$.

□

**Ejercicio 85(b)** Puesto que

$$\begin{bmatrix}\mathbf{A} & \\ \mathbf{C} & \mathbf{B}\end{bmatrix} = \begin{bmatrix}\mathbf{A} & \\ \mathbf{C} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \\ & \mathbf{B}\end{bmatrix}$$

y puesto que mediante una sucesión $\boldsymbol{\tau}_1, \ldots \boldsymbol{\tau}_k$ de transformaciones elementales *Tipo I* es posible la transformación $\begin{bmatrix}\mathbf{A} & \\ \mathbf{C} & \mathbf{I}\end{bmatrix} \xrightarrow{\boldsymbol{\tau}_1, \ldots \boldsymbol{\tau}_k} \begin{bmatrix}\mathbf{A} & \\ & \mathbf{I}\end{bmatrix}$ tenemos que $\begin{vmatrix}\mathbf{A} & \\ \mathbf{C} & \mathbf{I}\end{vmatrix} = \begin{vmatrix}\mathbf{A} & \\ & \mathbf{I}\end{vmatrix}$; y por tanto

$$\begin{vmatrix}\mathbf{A} & \\ \mathbf{C} & \mathbf{B}\end{vmatrix} = \begin{vmatrix}\mathbf{A} & \\ \mathbf{C} & \mathbf{I}\end{vmatrix}\begin{vmatrix}\mathbf{I} & \\ & \mathbf{B}\end{vmatrix} = \begin{vmatrix}\mathbf{A} & \\ & \mathbf{I}\end{vmatrix}\begin{vmatrix}\mathbf{I} & \\ & \mathbf{B}\end{vmatrix} = \left|\mathbf{A}\right| \cdot \left|\mathbf{B}\right|.$$

□

**Ejercicio 86.** Desarrollando por la segunda columna, tenemos

$$\det \mathbf{A} = -0\begin{vmatrix}5 & 2 & 4\\ 3 & 1 & 2\\ 5 & 2 & 1\end{vmatrix} + \begin{vmatrix}2 & 3 & 2\\ 3 & 1 & 2\\ 5 & 2 & 1\end{vmatrix} - 0\begin{vmatrix}2 & 3 & 2\\ 5 & 2 & 4\\ 5 & 2 & 1\end{vmatrix} + 3\begin{vmatrix}2 & 3 & 2\\ 5 & 2 & 4\\ 3 & 1 & 2\end{vmatrix} = 0 + 17 - 0 + 3 \times 4 = 29$$

□

**Ejercicio 87.**

$$\begin{aligned}\begin{vmatrix}a & b & c\\ d & e & f\\ g & h & i\end{vmatrix} &= -b\begin{vmatrix}d & f\\ g & i\end{vmatrix} + e\begin{vmatrix}a & c\\ g & i\end{vmatrix} - h\begin{vmatrix}a & c\\ d & f\end{vmatrix} \\ &= -b(-fg + id) + e(-cg + ia) - h(-cd + fa) \\ &= bfg - bdi - ecg + eia + hcd - hfa && \text{desarrollando términos} \\ &= a(ei - hf) - d(bi - hc) + g(bf - ec) && \text{reordenando} \\ &= aei + bfg + dhc - gec - dbi - hfa && \text{reordenando.}\end{aligned}$$

(Esta fórmula se denomina regla de Sarrus.)

□

**Ejercicio 88.** Sea $\lambda$ un autovalor de $\mathbf{A}$ y sean $\boldsymbol{x}$ y $\boldsymbol{y}$ dos autovectores correspondientes al autovalor $\lambda$, es decir, tales que $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ y $\mathbf{A}\boldsymbol{y} = \lambda\boldsymbol{y}$; entonces

$$\mathbf{A}\big(a\boldsymbol{x} + b\boldsymbol{y}\big) = a\mathbf{A}\boldsymbol{x} + b\mathbf{A}\boldsymbol{y} = \lambda a\boldsymbol{x} + \lambda b\boldsymbol{y} = \lambda\big(a\boldsymbol{x} + b\boldsymbol{y}\big).$$

□

**Ejercicio 89.** Extendemos el sistema $\big[\boldsymbol{q}\big]$ hasta formar una base ortogonal de $\mathbb{R}^n$

$$\big[\boldsymbol{q}\big] \longrightarrow \big[\boldsymbol{q};\ \boldsymbol{z}_1;\dots\ \boldsymbol{z}_{(n-1)};\big];$$

y luego colocamos $\boldsymbol{q}$ en la última posición y normalizamos todos los vectores; obteniendo la siguiente base ortonormal de $\mathbb{R}^n$:

$$\left[\frac{1}{\|\boldsymbol{z}_1\|}(\boldsymbol{z}_1);\dots\ \frac{1}{\|\boldsymbol{z}_{(n-1)}\|}(\boldsymbol{z}_{(n-1)});\ \boldsymbol{q};\right].$$

□

**Ejercicio 90.** $\quad {}_{i|}(\mathbf{Q}^{\intercal})\mathbf{Q}_{|j} = \big(\mathbf{Q}_{|i}\big)\cdot\big(\mathbf{Q}_{|j}\big) = \begin{cases} 1 & \text{cuando } i=j \\ 0 & \text{cuando } i\neq j \end{cases}$, $\quad$ es decir $\quad (\mathbf{Q}^{\intercal})\mathbf{Q} = \underset{n\times n}{\mathbf{I}}$.

□

**Ejercicio 91(a)** $\quad \big(\mathbf{QR}\big)^{\intercal} = \mathbf{R}^{\intercal}(\mathbf{Q}^{\intercal}) = \mathbf{R}^{-1}(\mathbf{Q}^{-1}) = \big(\mathbf{QR}\big)^{-1}$.

□

**Ejercicio 91(b)** $\quad \big(\mathbf{Q}^{-1}\mathbf{AQ}\big)^{\intercal} = \mathbf{Q}^{\intercal}\mathbf{A}(\mathbf{Q}^{-1})^{\intercal} = \mathbf{Q}^{\intercal}\mathbf{A}(\mathbf{Q}^{\intercal})^{\intercal} = \mathbf{Q}^{-1}\mathbf{AQ}$.

□

**Ejercicio 92.** Si $\mathbf{Q}^{-1}\mathbf{AQ} = \mathbf{D}$ es diagonal, entonces $\mathbf{A} = \mathbf{QDQ}^{\intercal}$. Por tanto, $\mathbf{A}^{\intercal} = \Big(\mathbf{QDQ}^{\intercal}\Big)^{\intercal} = \mathbf{QDQ}^{\intercal} = \mathbf{A}$.

□

**Ejercicio 93(a)** Si $\mathbf{A}$ es definida positiva, entonces para cualquier $\boldsymbol{x}\in\mathbb{R}^n$ no nulo se verifica que $\boldsymbol{x}\mathbf{A}\boldsymbol{x} > 0$. En particular, si $\boldsymbol{x}$ es la columna $j$-ésima de la matriz identidad, tenemos

$${}_{j|}\mathbf{I}\mathbf{A}\mathbf{I}_{|j} = {}_{j|}\mathbf{A}_{|j} > 0.$$

□

**Ejercicio 93(b)** $\quad \boldsymbol{x}(\mathbf{A}+\mathbf{B})\boldsymbol{x} = (\boldsymbol{x}\mathbf{A}+\boldsymbol{x}\mathbf{B})\boldsymbol{x} = \boldsymbol{x}\mathbf{A}\boldsymbol{x} + \boldsymbol{x}\mathbf{B}\boldsymbol{x} > 0$; (suma de dos números positivos).

□

**Ejercicio 94.** $\mathbf{A}$ es invertible y, como es simétrica, es ortogonalmente diagonalizable $\mathbf{A} = \mathbf{QD}(\mathbf{Q}^{-1}) = \mathbf{QD}(\mathbf{Q}^{\intercal})$; así

$$\mathbf{A}^{-1} = \big(\mathbf{QD}(\mathbf{Q}^{-1})\big)^{-1} = \big(\mathbf{Q}^{-1}\big)^{-1}(\mathbf{D}^{-1})(\mathbf{Q}^{-1}) = \mathbf{Q}(\mathbf{D}^{-1})(\mathbf{Q}^{\intercal});$$

y como los autovalores de $\mathbf{A}$ son positivos, sus inversos (los autovalores de $\mathbf{A}^{-1}$ que se encuentran en la diagonal de $\mathbf{D}^{-1}$) también lo son.

□

**Ejercicio 95(a)** Sea $\boldsymbol{z}\in\mathbb{R}^k$ tal que $\boldsymbol{z}(\mathbf{X}^{\intercal})\mathbf{AX}\boldsymbol{z} = \big(\mathbf{X}\boldsymbol{z}\big)\mathbf{A}\big(\mathbf{X}\boldsymbol{z}\big) \leq 0$; entonces necesariamente $\mathbf{X}\boldsymbol{z} = \mathbf{0}$ (pues $\mathbf{A}$ es definida positiva), y como $\mathbf{X}$ es de rango completo por columnas, necesariamente $\boldsymbol{z} = \mathbf{0}$.

□

**Ejercicio 95(b)** Sea $\mathbf{X} = \mathbf{I}_{|\boldsymbol{\alpha}}$, donde los elementos del vector $\boldsymbol{\alpha}$ son números enteros tales que $1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_k \leq n$; entonces, por la proposición anterior, $(\mathbf{X}^\intercal)\mathbf{A}\mathbf{X} = {}_{\boldsymbol{\alpha}|}\mathbf{A}_{|\boldsymbol{\alpha}}$ es definida positiva.

□

**Ejercicio 95(c)** Comencemos con el resultado propuesto como pista. Consideremos la submatriz principal

$$ {}_{(i,j,)|}\mathbf{A}_{|(i,j,)} = \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ij} & a_{jj} \end{bmatrix}, $$

que es simétrica, pues $\mathbf{A}$ también lo es. Además, como $\mathbf{A}$ es definida positiva, esta submatriz también: así que su determinante es positivo; es decir: $a_{ii}a_{jj} - a_{ij}^2 > 0$. Por tanto $\boxed{a_{ii}a_{jj} > a_{ij}^2}$.

Ahora supongamos erróneamente que $a_{ij}$ (con $i \neq j$) es la componente más grande en valor absoluto, en particular $|a_{ij}| > a_{ii}$ y $|a_{ij}| > a_{jj}$; consecuentemente $a_{ij}^2 > a_{ii}a_{jj}$. Pero esto es imposible, pues contradice el resultado del recuadro de arriba.

□

**Ejercicio 96.** La demostración es análoga a la del Apartado (a) del Ejercicio 93 en la página 216.

□

**Ejercicio 97.** Sea $\mathbf{K} = \mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ la matriz pre-escalonada de $\mathbf{A}$ obtenida por eliminación de izquierda a derecha; entonces $\mathbf{U} = (\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k})^{-1}$ es triangular superior unitaria. Como todos los menores principales superiores de $\mathbf{A}$ son de rango completo, por la proposición anterior sabemos que $\mathbf{A} = \mathbf{K}\mathbf{U}$ es la factorización LU y que todas las componentes de la diagonal de $\mathbf{K}$ son pivotes.

Sea $\mathbf{D}$ la matriz diagonal con los pivotes de $\mathbf{K}$, es decir, tal que ${}_{i|}\mathbf{D}_{|i} = {}_{i|}\mathbf{K}_{|i}$. Puesto que la diagonal de $\mathbf{K}$ no tiene ceros, la matriz $\mathbf{D}$ es invertible y, por tanto, la matriz $\mathbf{L} = \mathbf{K}(\mathbf{D}^{-1})$ es triangular inferior unitaria. Así, $\mathbf{A} = \mathbf{K}\mathbf{U} = \mathbf{K}(\mathbf{D}^{-1})\mathbf{D}\mathbf{U} = \mathbf{L}\mathbf{D}\mathbf{U}$.

La unicidad de esta factorización proviene de la unicidad de la factorización LU del teorema anterior.

□

**Ejercicio 98.** Sea $\mathbf{K} = \mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ la matriz pre-escalonada de $\mathbf{A}$ obtenida por eliminación de izquierda a derecha; entonces $\mathbf{U} = (\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k})^{-1}$ es triangular superior unitaria. Sea $\mathbf{D}$ diagonal tal que ${}_{i|}\mathbf{D}_{|i} = {}_{i|}\mathbf{K}_{|i}$, entonces $\mathbf{L} = \mathbf{K}(\mathbf{D}^{-1})$ es triangular inferior unitaria. Como todos los menores principales superiores de $\mathbf{A}$ son de rango completo, por la proposición anterior sabemos que $\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{U}$ es la factorización LDU.

Consideremos la matriz

$$ \underbrace{\left(\mathbf{U}^\intercal\right)^{-1}\mathbf{A}(\mathbf{U}^{-1})}_{\text{simétrica}} = \underbrace{\left(\mathbf{U}^\intercal\right)^{-1}\mathbf{K}}_{\text{triangular inferior}} = \left(\mathbf{U}^\intercal\right)^{-1}\mathbf{L}\mathbf{D}; $$

que por ser simultáneamente simétrica y triangular, es diagonal. Como $\mathbf{D}$ es diagonal, $\left(\mathbf{U}^\intercal\right)^{-1}\mathbf{L}$ lo es, pero además es triangular inferior unitaria. Por tanto, $\left(\mathbf{U}^\intercal\right)^{-1}\mathbf{L} = \mathbf{I}$, lo que implica que $\mathbf{L} = \mathbf{U}^\intercal$.

□

**Ejercicio 99.**

**Linealidad respecto al primer argumento:**

1. $\langle \alpha\boldsymbol{x} | \boldsymbol{y} \rangle_{\mathbf{A}} = \alpha\boldsymbol{x}\mathbf{A}\boldsymbol{y} = \alpha\langle \boldsymbol{x} | \boldsymbol{y} \rangle_{\mathbf{A}}$.

2. $\langle (\boldsymbol{x}+\boldsymbol{y})|\boldsymbol{z}\rangle_{\mathbf{A}} = (\boldsymbol{x}+\boldsymbol{y})\mathbf{A}\boldsymbol{z} = \boldsymbol{x}\mathbf{A}\boldsymbol{z} + \boldsymbol{y}\mathbf{A}\boldsymbol{z} = \langle \boldsymbol{x}|\boldsymbol{z}\rangle_{\mathbf{A}} + \langle \boldsymbol{y}|\boldsymbol{z}\rangle_{\mathbf{A}}$.

Por ser $\mathbf{A}$ simétrica:

**Simetría:** $\langle \boldsymbol{x}|\boldsymbol{y}\rangle_{\mathbf{A}} = \boldsymbol{x}\mathbf{A}\boldsymbol{y} = \boldsymbol{x}(\mathbf{A}^{\intercal})\boldsymbol{y} = \boldsymbol{y}\mathbf{A}\boldsymbol{x} = \langle \boldsymbol{y}|\boldsymbol{x}\rangle_{\mathbf{A}}$.

Por ser $\mathbf{A}$ definida positiva:

**Positivo:** $\langle \boldsymbol{x}|\boldsymbol{x}\rangle_{\mathbf{A}} = \boldsymbol{x}\mathbf{A}\boldsymbol{x} \geq 0$

**Definido:** $\langle \boldsymbol{x}|\boldsymbol{x}\rangle_{\mathbf{A}} = \boldsymbol{x}\mathbf{A}\boldsymbol{x} = 0 \Leftrightarrow \boldsymbol{x} = \mathbf{0}$.

□

,

# Bibliografía


Abadir, K. M. y Magnus, J. R. (2005). *Matrix Algebra*. N.º 1 en Econometric Exercises. Cambridge University Press, New York. ISBN 978-0-521-53746-9.

Arvesú Carballo, J., Marcellán Español, F. y Sánchez Ruiz, J. (2005). *Problemas resueltos de Álgebra Lineal*. Thomson Learning, Madrid. España. ISBN 84-9732-284-3.

Cullen, C. G. (1990). *Matrices and Linear Transformations*. Dover publications, Inc., New York, USA., 2.ª ed. Reproduce la segunda edición de Addison-Wesley, de 1972.

Fraleigh, J. B. y Beauregard, R. A. (1989). *Álgebra lineal*. Addison-Wesley Iberoamericana, Wilmington, Delaware. Versión en español de *Linear Algebra*, Addison-Wesley, 1987. ISBN 0-201-64420-7.

Godement, R. (1968). *Algebra*. Hermann. ISBN 9780901665003.

Godement, R. (1974). *Álgebra*. Editorial Tecnos, Madrid. Traducción de Mario Meléndez Rolla de *Cours d'algèbre* (Hermann, París). Segunda reimpresión de la primera edición, de 1967.

Herstein, I. N. y Winter, D. J. (1989). *Álgebra lineal y teoría de matrices*. Grupo Editorial Iberoamérica, México, D.F. Versión en español de *Matrix Theory and Linear Algebra*, Macmillan, 1988. ISBN 968-7270-52-7.

Higham, N. J. (1998). *Handbook of Writing for the Mathematical Sciences*. SIAM, Philadelphia, 2.ª ed.

Larson, R., Edwards, B. H. y Falvo, D. C. (2004). *Álgebra lineal*. Ediciones Pirámide, Madrid. España, 5.ª ed. ISBN 84-368-1878-4. Titulo de la obra original: Elementary Linear Algebra. Houghton Miffin Company.

Lay, D. C. (2007). *Álgebra Lineal y sus aplicaciones*. Pearson Educación, Inc., 3.ª ed.

Polya, G. (1957). *How to Solve It. A New Aspect of Mathematical Method*. Princeton University Press, Princeton, New Jersey. USA, 2.ª ed.

Poole, D. (2004). *Álgebra lineal. Una introducción moderna*. Thomson Learning, Mexico D.F. ISBN 970-686-272-2.

Strang, G. (2003). *Introduction to Linear Algebra*. Wellesley-Cambridge Press, Wellesley, Massachusetts. USA, 3.ª ed. ISBN 0-9614088-9-8.

Strang, G. (2007). *Álgebra Lineal y sus Aplicaciones*. Thomsom Learning, Inc, Santa Fe, México, D. F., 4.ª ed. ISBN 970686609-4.

Strang, G. (2016). *Introduction to Linear Algebra*. Wellesley-Cambridge Press, Wellesley, Massachusetts. USA, 5.ª ed. ISBN 978-0-9802327-7-6.

# Glosario

**C**

**combinación lineal de las columas de $\mathbf{A}$**: "$\mathbf{A}\boldsymbol{b}$"
**combinación lineal de las filas de $\mathbf{A}$**: "$\boldsymbol{a}\mathbf{B}$"
**combinación lineal de los vectores de Z**: "$\mathsf{Z}\boldsymbol{a}$"
**complemento ortogonal del subespacio $\mathcal{A}$**: "$\mathcal{A}^{\perp}$"
**conjunto**: "$\{a,\ B,\ \mathbf{C}, \ldots\}$"
  **conjunto descrito mediante la propiedad que cumplen sus elementos**: "$\{x \mid \text{condición}\}$"
    **conjunto de matrices de orden $m \times n$**: "$\mathbb{R}^{m\times n}$"
    **conjunto de numeros complejos**: "$\mathbb{C}$"
    **conjunto de numeros reales**: "$\mathbb{R}$"
    **conjunto de sistemas formados por $n$ números reales**: "$\mathbb{R}^{n}$"
**conjunto ortogonal al conjunto $A$**: "$A^{\perp}$"

**D**

**determinante de $\mathbf{A}$** : "$|\mathbf{A}|$"

**E**

**espacio columna de $\mathbf{A}$** : "$\mathcal{C}\left(\mathbf{A}\right)$"
**espacio nulo de $\mathbf{A}$** : "$\mathcal{N}\left(\mathbf{A}\right)$"
**espacios vectoriales**: "$\mathcal{A}$, $\mathcal{B}$, $\ldots$, $\mathcal{Y}$, $\mathcal{Z}$"

**M**

**matriz de $\mathbb{R}^{m\times n}$ (lista de $n$ columnas de $\mathbb{R}^{m}$)**: "$\left[\boldsymbol{a}\ \boldsymbol{b} \ldots\ \boldsymbol{v}\right]$"
**matriz elemental**: "$\mathbf{I}_{\boldsymbol{\tau}}$"
  **Tipo I**: "$\mathbf{I}_{\underset{[(\lambda)i+j]}{\boldsymbol{\tau}}}$"
  **Tipo II**: "$\mathbf{I}_{\underset{[(\alpha)i]}{\boldsymbol{\tau}}}$"
**matriz escalonada reducida**: "$\mathbf{R}$"
**matriz intercambio**: "$\mathbf{I}_{\underset{[p\rightleftharpoons q]}{\boldsymbol{\tau}}}$"
**matriz invertible, de rango completo, o no singular**: "$\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$"
**matriz permutación**: "$\mathbf{I}_{\underset{[\mathfrak{S}]}{\boldsymbol{\tau}}}$"
**matriz pre-escalonada**: "$\mathbf{K}$"
**matrices de $\mathbb{R}^{m\times n}$**: "$\mathbf{A}, \mathbf{B}, \ldots, \mathbf{Y}, \mathbf{Z}$"
  **identidad**: "$\mathbf{I}$"
  **inversa (de $\mathbf{A}$)**: "$\mathbf{A}^{-1}$"
  **nula**: "$\mathbf{0}$"
  **transpuesta**: "$\mathbf{A}^{\mathsf{T}}$; transpuesta de $\mathbf{A}$"
  **triangular inferior**: "$\mathbf{L}$"

**triangular superior**: "$\mathbf{U}$"

**O**

**operador concatenación de listas o sistemas**: "$\#$"
**operador selector de componentes de un sistema**: " $|$ "
**por la derecha** (selecciona el elemento $i$-ésimo de un sistema): "$|i$"
**por la izquierda** (selecciona el elemento $i$-ésimo de un vector o la fila $i$-ésima de una matriz): "$i|$"
**operador transformación elemental**: "$\boldsymbol{\tau}$"
**elemental Tipo I; suma $\lambda$ veces el elemento $i$-ésimo al elemento $j$-ésimo**: "$\underset{[(\lambda)i+j]}{\boldsymbol{\tau}}$"
**elemental Tipo II; multiplica por $\alpha$ el elemento $i$-ésimo**: "$\underset{[(\alpha)i]}{\boldsymbol{\tau}}$"
**intercambio entre los elementos $i$-ésimo y $j$-ésimo**: "$\underset{[i \rightleftharpoons j]}{\boldsymbol{\tau}}$"
**inverso de $\boldsymbol{\tau}$** : "$\boldsymbol{\tau}^{-1}$"
**permutación (reordenamiento) de elementos**: "$\underset{[\mathfrak{S}]}{\boldsymbol{\tau}}$"
**operador transposición**: "$\mathsf{T}$"

**P**

**producto punto (producto escalar usual en $\mathbb{R}^n$)**: "$\boldsymbol{a} \cdot \boldsymbol{b}$"

**R**

**rango de $\mathbf{A}$** : "$\mathrm{rg}\,(\mathbf{A})$"

**S**

**sistema genérico (lista ordenada)**: "$[a;\ B;\ \mathbf{C}; \ldots;]$"
**Sistema de ecuaciones lineales**: "$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$"
**sistemas de vectores de un subespacio genérico $\mathcal{V}$**: "$\mathsf{A}, \mathsf{B}, \ldots, \mathsf{Y}, \mathsf{Z}$"
**concatenación de los sistemas $\mathsf{Y}$ y $\mathsf{Z}$**: "$\mathsf{Y} \# \mathsf{Z}$"
**subespacio engendrado por un sistema $\mathsf{Z}$**: "$\mathcal{L}(\mathsf{Z})$"
**suma directa de los subespacios $\mathcal{A}$ y $\mathcal{B}$**: "$\mathcal{A} \oplus \mathcal{B}$"

**T**

**transformación elemental de las columnas de $\mathbf{A}$** : "$\mathbf{A}_{\boldsymbol{\tau}}$"

**V**

**vector formado por $n$ números reales**: "$(a,\ b,\ c, \ldots,)$"
**vectores de $\mathbb{R}^n$**: "$\boldsymbol{a}, \boldsymbol{b}, \ldots, \boldsymbol{y}, \boldsymbol{z}$"
**columna $j$-ésima de $\mathbf{A}$**: "$\mathbf{A}_{|j}$"
**fila $i$-ésima de $\mathbf{A}$**: "$_{i|}\mathbf{A}$"
**nulo**: "$\mathbf{0}$"
**vectores de un espacio vectorial genérico $\mathcal{V}$**: "$\vec{x}, \vec{y}, \vec{z},$"

# Símbolos

**combinaciones lineales**

$\mathbf{A}\boldsymbol{b}$ combinación lineal de las columas de $\mathbf{A}$ 29, 273
$\boldsymbol{a}\mathbf{B}$ combinación lineal de las filas de $\mathbf{A}$ 31, 273
$\mathsf{Z}\boldsymbol{a}$ combinación lineal de los vectores de $\mathsf{Z}$ 117, 273

**conjuntos**

$\{a,\ B,\ \mathbf{C}, \ldots\}$ conjunto 273
  $\{x \mid \textbf{condición}\}$ conjunto descrito mediante la propiedad que cumplen sus elementos 273
    $\mathbb{C}$ conjunto de numeros complejos 195, 273
    $\mathbb{R}$ conjunto de numeros reales 3, 273
    $\mathbb{R}^{m\times n}$ conjunto de matrices de orden $m \times n$ 85, 273
    $\mathbb{R}^{n}$ conjunto de sistemas formados por $n$ números reales 85, 273

**espacios y subespacios vectoriales**

$\mathcal{N}\left(\mathbf{A}\right)$ espacio nulo de $\mathbf{A}$ 103, 273
$\mathcal{C}\left(\mathbf{A}\right)$ espacio columna de $\mathbf{A}$ 114, 273

**matrices**

$\left[\boldsymbol{a}\ \boldsymbol{b} \ldots\ \boldsymbol{v}\right]$ matriz de $\mathbb{R}^{m\times n}$ (lista de $n$ columnas de $\mathbb{R}^{m}$) 12, 273
$\mathbf{A}, \mathbf{B}, \ldots, \mathbf{Y}, \mathbf{Z}$ matrices de $\mathbb{R}^{m\times n}$ 12, 273
  $\mathbf{A}^{-1}$ inversa (de $\mathbf{A}$) 69, 273
  $\mathbf{I}$ matriz identidad 19, 273
  $\mathbf{0}$ matriz nula 19, 273
  $\mathbf{A}^{\intercal}$ transpuesta de $\mathbf{A}$ 17, 273
  $\mathbf{L}$ triangular inferior 39, 273
  $\mathbf{U}$ triangular superior 39, 274

**matrices elementales**

$\mathbf{I}_{\boldsymbol{\tau}}$ matriz elemental 51, 273
  $\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}$ Tipo I 52, 273
  $\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}$ Tipo II 53, 273

**matrices: propiedades**

$\mathrm{rg}\left(\mathbf{A}\right)$ rango de $\mathbf{A}$ 80, 274
$|\mathbf{A}|$ determinante de $\mathbf{A}$ 177, 273

**matrices y eliminación**

$\mathsf{A}_{\boldsymbol{\tau}}$ transformación elemental de las columnas de $\mathsf{A}$ 51, 274
$\mathsf{I}_{\substack{\boldsymbol{\tau}\\[\boldsymbol{p}\rightleftharpoons\boldsymbol{q}]}}$ matriz intercambio 57, 273
$\mathsf{I}_{\substack{\boldsymbol{\tau}\\[\mathfrak{S}]}}$ matriz permutación 58, 273
$\mathsf{K}$ matriz pre-escalonada 59, 273
$\mathsf{R}$ matriz escalonada reducida 61, 273
$\mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ matriz invertible, de rango completo, o no singular 71, 273

**operaciones**

$\boldsymbol{a}\cdot\boldsymbol{b}$ producto punto (producto escalar usual en $\mathbb{R}^n$) 27, 274

**operadores**

$|$ operador selector de componentes de un sistema 274
$\intercal$ operador transposición 17, 274
$\shortparallel$ operador concatenación de listas o sistemas 42, 274
$i|$ por la izquierda 13, 274
$|i$ por la derecha 7, 274

**Sistema de ecuaciones lineales**

$\mathsf{A}\boldsymbol{x}=\boldsymbol{b}$ Sistema de ecuaciones lineales 102, 274

**sistema de vectores**

$\mathsf{A}$, $\mathsf{B}$,..., $\mathsf{Y}$, $\mathsf{Z}$ sistemas de vectores de un subespacio genérico $\mathcal{V}$ 117, 274
$\mathsf{Y}\shortparallel\mathsf{Z}$ concatenación de los sistemas $\mathsf{Y}$ y $\mathsf{Z}$ 124, 274

**sistema genérico**

$\big[a;\ B;\ \mathbf{C};\ldots;\big]$ sistema genérico (lista ordenada) 3, 274

**subespacios vectoriales**

$\mathcal{A}$, $\mathcal{B}$, ..., $\mathcal{Y}$, $\mathcal{Z}$ espacios vectoriales 87, 273
$\mathcal{L}\big(\mathsf{Z}\big)$ subespacio engendrado por un sistema $\mathsf{Z}$ 119, 274
$\mathcal{A}\oplus\mathcal{B}$ suma directa de los subespacios $\mathcal{A}$ y $\mathcal{B}$ 133, 274
$A^{\perp}$ conjunto ortogonal al conjunto $A$ 148, 273
$\mathcal{A}^{\perp}$ complemento ortogonal del subespacio $\mathcal{A}$ 149, 273

**transformaciones elementales (o sucesiones de transformaciones elementales)**

$\boldsymbol{\tau}$ operador transformación elemental 51, 274
$\boldsymbol{\tau}_{[\boldsymbol{i}\rightleftharpoons\boldsymbol{j}]}$ intercambio entre los elementos $\boldsymbol{i}$-ésimo y $\boldsymbol{j}$-ésimo 57, 274
$\boldsymbol{\tau}_{[\mathfrak{S}]}$ permutación (reordenamiento) de elementos 58, 274
$\boldsymbol{\tau}_{[(\alpha)\boldsymbol{i}]}$ elemental Tipo II; multiplica por $\alpha$ el elemento $\boldsymbol{i}$-ésimo 53, 274
$\boldsymbol{\tau}_{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}$ elemental Tipo I; suma $\lambda$ veces el elemento $\boldsymbol{i}$-ésimo al elemento $\boldsymbol{j}$-ésimo 51, 274
$\boldsymbol{\tau}^{-1}$ inverso de $\boldsymbol{\tau}$ 70, 274

**vectores de un subespacio vectorial genérico**

$\vec{x}$, $\vec{y}$, $\vec{z}$, vectores de un espacio vectorial genérico $\mathcal{V}$ 86, 274

**vectores formados por $n$ números reales**

$(a,\ b,\ c,\dots,)$ vector formado por $n$ números reales 4, 274
$\boldsymbol{a}, \boldsymbol{b},\dots, \boldsymbol{y}, \boldsymbol{z}$ vectores de $\mathbb{R}^n$ 4, 274
$\mathbf{A}_{|j}$ vector columna $j$-ésima de $\mathbf{A}$ 13, 274
$_{i|}\mathbf{A}$ vector fila $i$-ésima de $\mathbf{A}$ 13, 274
$\mathbf{0}$ vector nulo 5, 274

# Índice alfabético


aplicación lineal, *véase* función lineal
autoespacio, 193, 213
autovalor, 193, 195, 216, 221
  espectro, 193
  multiplicidad algebraica, 195, 208
  multiplicidad geométrica, 195, 208
  potencias de una matriz, 196
  producto, 207
  suma, 207
autovector, 193

*Cauchy-Schwarz*, *véase* desigualdad, 153
combinación lineal, 28, 29, 31, 45, 88, 117
concatenación de sistemas, 42, 124
conjunto, 3
  de vectores linealmente dependiente, 128
    también, *véase* sistema de vectores
  generador de un subespacio, 128
    también, *véase* sistema de vectores
  vacío, 99
conjunto de sucesiones casi nulas, 156
  conjunto de polinomios de una variable, 157
conjunto ortogonal de vectores, 148, 169
coordenadas, 130

desigualdad
  *Cauchy-Schwarz*, 146, 153
  triangular, 148, 153
determinante, 176, 207
  cofactor, 186
  cálculo de la inversa, 188
  cálculo por eliminación Gaussiana, 182
  de matrices diagonales por bloques, 182
  desarrollo por cofactores, 184
  expansión de Laplace, 184, 187
  menor, 186, 222
  producto de pivotes, 222
  regla de Cramer, 188
  regla de Sarrus, 184
diagonalización
  matriz simétrica, 211
  ortogonal, 211
  por bloques triangulares, 201, 208
  por congruencia, 217, 220
  por semejanza, *véase* por bloques triangulares, 220

ecuaciones cartesianas, 103, 159
ecuaciones paramétricas, 107, 159
ecuación lineal, 101
eliminación, 59, 104
  de izquierda a derecha, 59, 124
    pre-escalonar, 59
  Gauss-Jordan, 61
    escalonar y reducir, 61
  gaussiana, 60, 61
    escalonar, 61
eliminación por filas, 62
envolvente lineal, *véase* subespacios
escalares, 3
  números complejos, 211
  números reales, 3
espacio afín, 161
  hiperplano en $\mathbb{R}^n$, 161
  plano en $\mathbb{R}^n$, 161
  punto en $\mathbb{R}^n$, 161
  recta en $\mathbb{R}^n$, 161
espacio euclídeo, 154
  $\mathbb{R}^n$, 154
espacio pre-Hilbert, 154, 167
espacio vectorial, 86
  subespacio, *véase* subespacios
espectro, 193, 201

factorización de una matriz, 222
  factorización KU, 222
factorización de una matriz cuadrada
  factorización LDU, 223
  factorización LU, 222
factorización de una matriz cuadrada y simétrica
  factorización de Cholesky, 223
  factorización de LDU, 223
forma bilineal, **151**
forma cuadrática, 215, 217
  completar el cuadrado, 221
  tipos, 217
    definida positiva, 215
función, 95

autoadjunta, 167
composición de funciones, 97
idempotente, 167
invertible, 97
notación, 96
producto por un escalar, 92
restricción, 98
suma de funciones, 92
función determinante, *véase* determinante
función lineal, 88
composición, 89
imagen, 130
invertible, 89
núcleo o espacio nulo, 129

Gram-Schmidt, *véase* método de Gram-Schmidt

inercia, *véase* ley de inercia
inversa
de una función, 97
de una matriz, 69
transpuesta, 70
triangular, 77
de una transformación elemental, 70
*espejo*, 67

ley de inercia, 220
linealidad, *véase* operador lineal

matriz, 12
adjunta, *véase* determinante
columna de una matriz, 13
columna libre de una matriz, 107
columna pivote de una matriz, 107
componente (o elemento) de una, 12
congruente, 217, 221
cuadrada, 19
definida positiva, 215, 221
producto escalar, 224
diagonal, 19, 221
diagonal principal de una, 19
diagonalizable ortogonalmente, 212, 213
diagonalizable por semejanza, 208
elemental, 51
elemental de Tipo I, 52
elemental de Tipo II, 53
escalonada, 60, 62
escalonada por filas, 65
escalonada reducida, 61, 62
extendida, 181
fila de una matriz, 13
fila pivote de una matriz, 137
idempotente, 165, 167
identidad, 19
igualdad, 14
intercambio, 57
inversa, 69
invertible, **69**, 71, 73, 81
invertible por la derecha, 69
invertible por la izquierda, 69
*matriz columna*, 12, 16, 18
*matriz fila*, 18
múltiplo de, 20
notación, 12
nula, 19
opuesta, 22
orden de una matriz, 12, 19
ortogonal, 211
permutación, 58
potencias de una matriz cuadrada, 46
potencia cero, 46
pre-escalonada, 59, 62
producto de matrices, **35**, 37, 40
producto por un escalar, 20
proyección, 166
rango, 80
rango completo, 81, 127
rango completo por columnas, 81
rango completo por filas, 81
rectangular, 19
selección de un elemento (o componente), 14, 15
selección de una columna, 14, 15
selección de una fila, 13, 15
selección de una submatriz, 40
semejante, 199, 207
simétrica, 19, 167, 213, 217
singular, 69
suma, 20
transpuesta, 17, 38
traza, 45, 200, 207
propiedades, 45
triangular
por bloques, 182, 203
triangular inferior, 39
triangular superior, 39
triangular unitaria, 39
método de eliminación, *véase* eliminación
método de Gram-Schmidt, 209, 214

nomenclatura de G. Strang, 107
números complejos, 211
parte imaginaria, 211
parte real, 211

operador lineal, **6**, 21, 23, 28, 30, 33, 37, 56, 118

permutación, 58
*Pitágoras* (Teorema), 146, 153
pivote, 59, 222
  columna pivote, 107
  posición de pivote, 59
polinomio, 157, **215**
  grado, 215
    raíces, 217
polinomio característico, 194
  matriz triangular por bloques, 203
  raíces, 195
pre-escalonar, 59
producto cartesiano, 91
producto escalar, 27, **151**, 224
producto punto, **27**
proyección ortogonal, 163, **167**
  coeficientes de Fourier, 166, 170
  en $\mathbb{R}^n$, 163
  en un espacio pre-Hilbert $\mathcal{V}$, 167

rango, *véase* matriz
*Rouché-Frobenius* (Teorema), 113

selector (*operador*), 39
selector (*operador*), 4, 13–15, 25, 35, 40, 117, 118, 124
sistema, 3, 12, **98**, 117
  vacío, 99, 122, 128
sistema de ecuaciones lineales, 101
  compatible determinado, 112
  compatible indeterminado, 112
  equivalentes, 159, 164
  homogéneos, 102
    algoritmo de resolución, 104, 107
    solución trivial, 103
  incompatible, 112
  matriz de coeficientes, 102
  matriz de coeficientes ampliada, 109
  no homogéneos, 109
  regla de Cramer, 188
  resolución de, 112
  resolver un sistema, 102
  soluciones especiales, 105, 107
  solución de un sistema, 102
  variables libres (o exógenas), 107
  variables pivote (o endógenas), 107
  vector de incógnitas, 102
  vector del lado derecho, 102
sistema de ecuaciones normales, 164
sistema de vectores de $\mathcal{V}$
  base ortogonal, 210
  generador de un subespacio, 119, 120
    sistemas equivalentes, 120, 209
  linealmente dependiente, 122
  linealmente independiente, 122, 123, 130, 209
  ortogonal, 209
  ortonormal, 210
  selección de varias componentes, 118
  sistemas acoplados, 124
  sistemas equivalentes, 121, 209
subespacios, 90
  base, 122
  complementos ortogonales, 149, 154, 156
  dimensión, 123, 132
    infinita, 156
  envolvente lineal, $\mathcal{L}(\mathsf{Z})$, 128
  espacio columna de $\mathbf{A}$, 114
  espacio fila de $\mathbf{A}$, 134
  espacio nulo de $\mathbf{A}$, 103
  espacio nulo de un sistema de vectores $\mathsf{Z}$, 122
  espacio nulo de una función lineal, 129
  espacio nulo por la izquierda de $\mathbf{A}$, 135
  generado por el conjunto vacío, 128
  generado por un conjunto, $\mathcal{L}(Z)$, 128
  generado por un sistema, $\mathcal{L}(\mathsf{Z})$, 119
  generado por un sistema, $\mathcal{L}(\mathsf{Z})$, 120
  iguales, 120, 123
  imagen de una función lineal, 130
  intersección de, 91
  núcleo de una función lineal, 129
  suma de subespacios, 131
  suma directa de subespacios, 133
  suplementarios, 133, 149, 156, 169
submatriz, 40
  principal, 40, 216
  principal superior, 40

Teorema Espectral, 213, 220
Teorema Fundamental del Álgebra, 195
transformación elemental, 51
  de las columnas de una matriz, 52, 54
  de las filas de una matriz, 62
  de los vectores de un sistema $\mathsf{Z}$, 121
  intercambio, 57
  inversa, 70
  permutación, 42, 58
  secuencia de transformaciones, 56
  Tipo I, 51
  Tipo II, 53
  transpuesta, 62, 63
traza, *véase* matriz

vectores de $\mathcal{V}$, 86
  alineados, 154
  combinación lineal, 88, 117

coseno del ángulo formado por, 153
distancia entre vectores, 153
longitud o norma, 153
múltiplo de, 86
ortogonales (o perpendiculares), 153
producto por escalares de, 86
suma de, 86
unitario, 153
vectores de $\mathbb{R}^n$, 3
alineados, 148
combinación lineal, 28, 31, 45
coseno del ángulo formado por, 147
distancia entre vectores, 148
igualdad, 4
longitud o norma, 145
múltiplo de, 6
normalización de, 144
notación, 4
nulo, 5
opuesto, 5
ortogonales (o perpendiculares), 145
producto escalar usual en $\mathbb{R}^n$, *véase* producto punto
producto por un escalar de, 6
resta, 10
selección de una componente, 4, 14, 15
selección de varias componentes, 45
suma, 6
unitario, 144, **146**, 210

# Índice de figuras


6.1. Representación esquemática de que $\mathcal{V} \times \mathcal{W} = f \oplus (\{\vec{0}\} \times \mathcal{W})$. . . . . 94
6.2. Representación de los conjuntos de pares $(x, x)$ en verde, $(x, x^2)$ en azul. . . . . 97
6.3. Representación de los conjuntos de pares $(x, x)$ en verde, $(x, x^2)$ en azul restringidos al subconjunto $A = \{x \mid x \in (-1, 0.5)\}$. Compárese con la Figura 6.2 . . . . . 98

10.1. Esquema de descomposición de los cuatro subespacios fundamentales de una matriz (parte 1) 135

11.1. Longitud de un vector en $\mathbb{R}^2$ y en $\mathbb{R}^3$ . . . . . 143
11.2. Relación entre el ángulo $\theta$ y su coseno (el círculo exterior es de radio uno) . . . . . 144
11.3. Triángulo formado por $\mathbf{0}$ y dos vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ no nulos. . . . . 144
11.4. Representación esquemática donde $\boldsymbol{h}$ es combinación lineal de $\boldsymbol{a}$ y $\boldsymbol{b}$ y perpendicular a $\boldsymbol{a}$. . . 147
11.5. Representación esquemática del triángulo formado por $\vec{0}$ y dos vectores $\vec{x}$ y $\vec{y}$ no nulos de $\mathcal{V}$. 154
11.6. Representación esquemática desde dos ángulos distintos. El subespacio $\mathcal{W}$ es de dimensión $n$, el "suelo" $\mathcal{H}$ es un subespacio de dimensión $n+1$ que contiene a $\mathcal{W}$, y el espacio vectorial ambiente $\mathcal{V}$ (posiblemente de dimensión infinita) está representado por el espacio tridimensional (a la derecha se muestra la visión desde la vertical sobre $\mathcal{H}$). . . . . 155
11.7. Representación esquemática. La recta $\mathcal{W}$ es un subespacio de dimensión $n$, el plano (el suelo) $\mathcal{H}$ es un subespacio de dimensión $n+1$ que contiene a $\mathcal{W}$, a $\mathcal{L}(\vec{z})$ (que es el único subespacio de la figura que con seguridad es de dimensión 1), y el espacio vectorial ambiente $\mathcal{V}$ (posiblemente de dimensión infinita) está representado por el espacio tridimensional (si $\mathcal{V}$ es de dimensión infinita, entonces $\mathcal{H}^\perp$ también). . . . . 156

13.1. Representación esquemática de la proyección ortogonal de $\boldsymbol{y}$ sobre $\mathcal{V}$ . . . . . 163

14.1. Área$\left[\mathbf{I}_{|1};\ \mathbf{I}_{|2};\right] = 1$ (izquierda) y Volumen$\left[\mathbf{I}_{|1};\ \mathbf{I}_{|2};\ \mathbf{I}_{|3};\right] = 1$ (derecha). . . . . 175
14.2. Área$\left[\boldsymbol{a};\ \boldsymbol{b};\right] =$ Área$\left[\boldsymbol{a};\ (\alpha\boldsymbol{a}+\boldsymbol{b});\right]$ (izquierda) y Vol$\left[\boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c};\right] =$ Vol$\left[\boldsymbol{a};\ (\alpha\boldsymbol{a}+\boldsymbol{b});\ \boldsymbol{c};\right]$ (derecha). 175
14.3. $|\alpha| \cdot$ Área$\left[\boldsymbol{a};\ \boldsymbol{b};\right] =$ Área$\left[\alpha\boldsymbol{a};\ \boldsymbol{b};\right]$ (izquierda) y $|\alpha| \cdot$ Vol$\left[\boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c};\right] =$ Vol$\left[\alpha \cdot \boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c};\right]$ (derecha). 176

A.1. Representación esquemática donde $\vec{h}$ es combinación lineal de $\vec{x}$ y $\vec{y}$ y perpendicular a $\vec{x}$. . 257

# Índice de cuadros

1. Equivalencias entre la notación del libro y la notación habitual. En las cuatro últimas filas $E$ y $E_i$ son matrices elementales. En las transformaciones de las columnas $E_i = \mathbf{I}_{\boldsymbol{\tau}_i}$. En las transformaciones de las filas $E_i = {}_{\boldsymbol{\tau}_i}\mathbf{I}$ . . . . . xxvii

1.1. Precedencia de operadores (detalles aquí: 6.16 *Orden de evaluación* y 6.17 *Prioridad de operador*) 7