Un Curso de Álgebra Lineal

con notación asociativa

Edición curso 2020/2021

Versión: 2 de marzo de 2021

Marcos Bujosa

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

Este libro ha sido escrito con LaTeX.

Nunca hubiera podido escribir este libro sin la ayuda de mi hermano Andrés.

“Notation is everything”

Charles F. Van Loan

Prefacio

Aunque muchos perciben las matemáticas como una batería de procedimientos mecánicos para resolver problemas tipo, en realidad las matemáticas son un lenguaje. La principal intención de este curso es que el estudiante lo perciba y juegue con las reglas de dicho lenguaje.

También es frecuente escuchar que *“las matemáticas están en todas las cosas”*. Yo discrepo. Creo que las matemáticas son algo al margen de la realidad material del mundo, y comparto la frase de Eduardo Sáenz de Cabezón de que *“las matemáticas son el lenguaje en el que nosotros leemos el mundo”* (en su charla “Las matemáticas nos hacen más libres y menos manipulables”; minuto 6:40). Es decir, son el lenguaje que usamos para describir la realidad. A las descripciones matemáticas del mundo las llamamos *modelos*...y cuando nos “ponemos estupendos” incluso las denominamos *modelos científicos*. Las matemáticas nos sirven para expresar ideas y describir relaciones, pero no necesariamente relacionadas con el mundo real. Por eso considero que las matemáticas son algo distinto al mundo material que denominamos “realidad”.

Las matemáticas son un *lenguaje formal*. Consisten en un conjunto de símbolos y unas reglas para su manipulación. Así es como entiendo la cita de Charles F. Van Loan “Notation is everything”. Las matemáticas son notación. No obstante, solemos dar una interpretación subjetiva a dicha notación. Las interpretaciones nos dotan de un esquema mental que nos ayuda a pensar, y por eso son importantes. Pero las interpretaciones corresponden a un nivel distinto. De hecho, una misma notación puede tener interpretaciones alternativas. Por ejemplo, utilizamos el Teorema de Pitágoras cuando trabajamos con distancias euclídeas, con números complejos, al deducir algunas identidades trigonométricas, al analizar señales o series temporales, para tener una interpretación geométrica de la esperanza y la varianza, etc... y en todos estos ejemplos el formalismo del Teorema de Pitágoras es el mismo. El formalismo es el pilar sobre el que se apoyan las distintas interpretaciones. Así pues, las matemáticas son un conjunto de símbolos y unas reglas de uso para dichos símbolos. *En matemáticas LA NOTACIÓN LO ES TODO.*

Fiel a este espíritu, este texto presta una especial importancia a la notación y todo es un juego de manipulación de símbolos. Por supuesto que también mostraremos algunas interpretaciones, como la representación gráfica de un vector o la proyección sobre un subespacio de $\mathbb{R}^n$, pero el peso recaerá en la manipulación de los símbolos, y no en la interpretación. Así pues, he tenido un especial cuidado en elegir una notación que evite la ambigüedad[1] y que facilite operar con ella (por ejemplo, explotando la asociatividad logramos expresiones más simples).

Para mostrar la potencia que la notación tiene como lenguaje, la he implementado en una librería de Python.

¡Qué mejor muestra de que las matemáticas son un lenguaje, que usar dicho lenguaje con un PC para “enseñar” al PC a resolver problemas matemáticos y que todo funcione como se espera!

El código es una implementación literal de la notación empleada en el libro. Así se puede verificar que siguiendo las reglas de reescritura podemos lograr que el ordenador sea capaz de operar con vectores, matrices, resolver sistemas de ecuaciones, diagonalizar matrices, calcular determinantes, etc.

[1] por ejemplo, hay una clara distinción entre vectores y matrices,...aquí no encontrara ¡la “transpuesta” de ningún vector!

La librería se instala ejecutando[2] `pip install nacal`. También puede descargar el código fuente y la documentación desde

https://github.com/mbujosab/nacallib

Pero incluso sin instalar nada, puede usar la librería si abre los Notebooks de Jupyter en su navegador:

https://mybinder.org/v2/gh/mbujosab/nacal-Jupyter-Notebooks/master

Aprenderá mucho si experimenta con el ordenador. No se desanime si le cuesta un poco al principio.

Otros aspectos del libro

Todos los resultados están demostrados (salvo el *Teorema Fundamental del Álgebra*[3]). No obstante, la mayoría de las demostraciones son tan sencillas que están propuestas como ejercicios (aunque con frecuencia se ofrecen pistas que indican los pasos necesarios en la demostración). Así que muchas de las demostraciones aparecen en la sección de soluciones a los ejercicios. El propósito es doble: por una parte se aligera el texto. Y por otra permite al estudiante demostrar la veracidad de los resultados por su cuenta (sin tener la solución a la vista).

El curso gira en torno al Método de Eliminación y las transformaciones elementales. Hay numerosos libros de texto y manuales que siguen este modelo; por ejemplo Cullen (1972); Larson et al. (2004); Lay (2007); Poole (2004); Strang (2007, 2003), o el libro de problemas Arvesú Carballo et al. (2005). Sin embargo este texto también presenta una diferencia operativa respecto a los manuales indicados más arriba. Por ejemplo, aunque el Profesor Strang indica que *"el sistema de ecuaciones $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ tiene solución si el vector $\boldsymbol{b}$ es una combinación lineal de las* columnas *de $\mathbf{A}$"*, aplica el método de eliminación gaussiana operando con las *filas*, por lo que realmente acaba resolviendo un sistema distinto aunque equivalente a $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$. Y así es como se hace en casi todos los manuales de Álgebra Lineal. Sin embargo, aquí emplearé el método de eliminación de Gauss por columnas, operando siempre en el espacio columna de $\mathbf{A}$ para encontrar las combinaciones de las columnas que son iguales a $\boldsymbol{b}$. Es más, con básicamente un único algoritmo (la eliminación por columnas y sin necesidad de la sustitución hacia atrás) calcularemos casi todo lo necesario (la inversa de una matriz, el conjunto de soluciones a un sistema de ecuaciones lineales, realizaremos operaciones con subespacios de $\mathbb{R}^n$, calcularemos determinantes, etc.). Solo al diagonalizar matrices necesitaremos tanto la eliminación por filas como por columnas (¡ambas!).

Mi experiencia es que los estudiantes de economía no suelen tener una especial afinidad por las matemáticas. Y en aquellos casos en que 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.[4] Tirando del hilo de los algoritmos irán surgiendo distintas estructuras conceptuales, cuya generalización permitirá reconsiderar lo ya visto, pero con un mayor grado de abstracción. No obstante, si uno analiza las demostraciones con perspectiva, se evidencia que de nuevo todo acaba reduciéndose a unas reglas precisas en la manipulación de unos símbolos (*"La notación lo es todo"*).

Instrucciones para mis alumnos

Antes y durante las clases en el aula. Este libro está dividido en lecciones. Para poder aprovechar adecuadamente las clases recibidas en el aula es IMPRESCINDIBLE leer con antelación la lección correspondiente. El profesor dará por supuesto que usted así lo ha hecho.

En clase se dará una exposición general de lo que aparece en este libro. Por tanto, los detalles de cada lección deben ser preparados por el alumno estudiando las secciones correspondientes.

Recuerde que usted debe haber leído las secciones de referencia correspondientes antes de cada clase.

[2] Previamente debe tener instalado python3 en su ordenador
[3] cuya demostración está fuera del alcance de este curso
[4] el desarrollo en paralelo de la librería NAcAL para Python me ha ayudado en este sentido.

Para ayudar a tener una idea general de cada lección, hay un resumen por lección en el Apéndice A. (**por hacer**). También hay unas transparencias que sirven de apoyo al profesor en sus clases (junto con un resumen previo de cada lección (**por hacer**). Le puede venir bien llevarlas a clase para no tomar tantos apuntes.

El tiempo dedicado a resolver problemas en clase es un buen momento para resolver dudas con su profesor.

También me gustaría acompañar las lecciones con vídeos (**por hacer**)…aunque no creo que nunca puedan ser tan buenos como los de Grant Sanderson de su canal 3blue1brown.

Al estudiar Recuerde que el esfuerzo y estudio individual (o en tándem) son imprescindibles; limitarse a atender las explicaciones en clase no es en modo alguno suficiente.

Al final del libro encontrará las soluciones a los problemas, pero no debe mirarlas hasta que haya dado con "su" solución. Consultar la solución de otro sin haber resuelto el ejercicio por cuenta propia sirve de muy poco cuando se estudia. **Recuerde que el aprendizaje es una tarea activa, es decir, usted debe encontrar la solución activamente, y no consultar la solución de otro** (la mía en particular).

Instrucciones para los profesores

Este libro se acompaña de unas transparencias para proyectar en clase. Son las que yo uso, y las acompaño de un resumen para orientar a quien quiera usarlas en clase (**añadir dirección de fichero con las transparencias**)

Índice general

Parte I

Álgebra Matricial

Operaciones con vectores y operaciones con matrices

En matemáticas los objetos tienen definiciones muy precisas. Comencemos con una que usaremos mucho:

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

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

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

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

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

Distinguimos conjuntos y sistemas del siguiente modo: por una parte encerramos los elementos de un conjunto entre llaves, separados por comas; mientras que los elementos de la lista correspondiente a un sistema genérico se encierran entre corchetes, y se separan con el símbolo de "punto y coma".

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

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

Denotamos por $\mathbb{R}$ al *conjunto de números reales* (por ejemplo los números 7, $\frac{-3}{11}$, $\sqrt{2}$, ó $-\pi$). A los números reales también los llamaremos *escalares*.

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 componentes que hay en la lista.

Recuerde que los sistemas están ordenados y que las componentes se identifican por su posición: hay una primera componente, una segunda componente, etc. La forma de escribir la lista no es importante siempre que el orden de las componentes quede claro (es decir, siempre que quede claro quien es la primera componente, quien la segunda, etc.). Por eso podemos escribir el *mismo* vector tanto en horizontal como en vertical:

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

Ambas son representaciones del vector de $\mathbb{R}^4$ cuya primera componente es π, la segunda -1, la tercera 0 y la cuarta 1.

[1] con "objeto" siempre nos referiremos a *objetos matemáticos*, es decir, números, funciones, ecuaciones, variables, etc.

Notación. Aunque los vectores de $\mathbb{R}^n$ son sistemas, nosotros usaremos una notación específica con ellos[2]: siempre escribiremos un vector como una lista entre ***paréntesis*** (y con sus *elementos separados por comas* cuando lo escribamos en horizontal).

Además, denotaremos los vectores de $\mathbb{R}^n$ con letras minúsculas en negrita cursiva: $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ etc. Y nos referiremos a los componentes genéricos de un vector con la misma letra con la que denotamos al vector completo (pero en cursiva y con un subíndice que indica la posición del componente dentro del vector). Por ejemplo:

$$\boldsymbol{a} = \begin{pmatrix} a_1, & a_2, & a_3, & a_4, & a_5, & a_6 \end{pmatrix};$$

es un vector de 6 componentes, y cuyo tercer componente es a_3; y para indicar que el segundo componente de $\boldsymbol{b} = \begin{pmatrix} 0, & 4, & 6 \end{pmatrix}$ es un cuatro, escribimos $b_2 = 4$.

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: $\quad x_i = y_i \;$ para $\; i = 1 : n \quad$ (i.e. si son iguales componente a componente). [3]

Por ejemplo, con los números 5, 10, 15 y 20, podemos denotar dos vectores *distintos* del siguiente modo:

$$\boldsymbol{x} = \begin{pmatrix} 20, & 10, & 15, & 5 \end{pmatrix} \quad \text{y} \quad \boldsymbol{z} = \begin{pmatrix} 5, & 20, & 15, & 10 \end{pmatrix}.$$

Tenga en cuenta que dos vectores son iguales solo cuando lo son las listas correspondientes a ambos vectores. Por eso los vectores $\boldsymbol{x}$ y $\boldsymbol{z}$ NO son iguales.

El operador selector de componentes " | ". Para homogeneizar la notación que emplearemos con vectores y matrices, también denotaremos la $i$ésima componente de $\boldsymbol{a}$ añadiendo un subíndice en el que aparece el símbolo "|" junto al índice de la componente: $\boldsymbol{a}_{|i}$. De esta manera indicamos que el operador selector "|" *actúa* sobre el vector que está a su izquierda para seleccionar la componente $i$ésima. Así, las siguientes son formas equivalentes de denotar la tercera componente de $\boldsymbol{b}$:

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

☞ Durante el curso usaremos la librería 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. La librería es una transcripción literal de los objetos y procedimientos que veremos en el libro.

Uso de NAcAL

Para usar la librería, el primer paso es importar la librería en memoria

Librería NAcAL para Python

```python
from nacal import *  # importamos la librería en memoria
```

Una vez importada la librería, tecleando `Vector( [0, 4, 6] )` en un terminal obtenemos:

terminal de Python

```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])
```

[2] si usáramos la notación genérica de sistema, deberíamos escribir $\begin{bmatrix}\pi; & -1; & 0; & 1\end{bmatrix}$ (una lista entre corchetes con sus elementos separados por punto y coma como en los ejemplos dados tras la Definición 1.1)

[3] Si $n \leq m \in \mathbb{N}$ denotamos con "$n : m$" la secuencia $n, n+1, \ldots, m$, es decir, la lista ordenada de números naturales que van de n a m: $\{k \in \mathbb{N} \mid n \leq k \leq m\}$. Así, con $i = 1 : n$ indicamos que el índice i recorre la sucesión de enteros que va de 1 a n, es decir $1, 2, \ldots, n$. Cuando el primer índice (n) sea mayor que el segundo (m) entenderemos que ($n : m$) 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 el tercer componente de $\boldsymbol{b}$ tecleamos:

Librería NAcAL para Python

```
b|3
```

de manera que tanto en un terminal de Python como en un Notebook de Jupyter obtenemos 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 tercera componente de $\boldsymbol{a}$ es:

$$\boldsymbol{a}_{|3} = a_3 = 9.$$

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 $\mathbb{R}^1$, $\mathbb{R}^2$, de $\mathbb{R}^3$ y de $\mathbb{R}^5$ son respectivamente:*

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

¡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)!

Librería 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 cuyos componentes son los opuestos de los de $\boldsymbol{a}$; es decir $\big(-\boldsymbol{a}\big)_{|i} = -\big(\boldsymbol{a}_{|i}\big)$. Por ejemplo:*

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

```
-b              # vector opuesto a b
```

Librería NAcAL para Python

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

1.2. Sumas de vectores y productos de vectores por escalares

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 $\mathbb{R}^n$ en el que cada componente iésima es la suma de las componentes $\boldsymbol{a}_{|i}$ y $\boldsymbol{b}_{|i}$. Es decir, sumamos componente a componente:*

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

Por tanto:

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

suma que también podríamos haber escrito de manera horizontal: $\boldsymbol{a}+\boldsymbol{b}=\Big((a_1+b_1),\ldots,(a_n+b_n)\Big)$.

Observación. Nótese que el operador " $|i$ " es *distributivo* 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 λ es el vector de $\mathbb{R}^n$ en el que cada componente iésima es el producto de λ por $\boldsymbol{b}_{|i}$. Es decir, multiplicamos cada componente de $\boldsymbol{b}$ por λ:*

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

Al vector $\lambda\boldsymbol{b}$ lo denominamos múltiplo *(o múltiplo escalar) de $\boldsymbol{b}$.*

Por tanto

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

múltiplo que también podríamos haber escrito horizontalmente: $\lambda\boldsymbol{b}=\lambda\Big(b_1,\ldots,b_n\Big)=\Big(\lambda b_1,\ldots,\lambda b_n\Big)$.

Observación. El operador "$|$" es *asociativo* para el producto por escalares: $(\lambda\boldsymbol{b})_{|i}=\lambda(\boldsymbol{b}_{|i})$. Fíjese como la región cubierta por el paréntesis se desplaza... en un lado del la igualdad "λ" está dentro del paréntesis e " $|i$ " fuera, pero en el otro lado ocurre lo contrario[4]. Por tanto, con el producto por escalares podemos escribir sencillamente $\lambda\boldsymbol{b}_{|i}$.

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

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

[4]de manera similar a la propiedad asociativa del producto de escalares: $(ab)c=a(bc)$.

☞ 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:

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

Observación. La expresión $2\boldsymbol{b}$ significa multiplicar el vector por 2, es decir

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

donde "·" indica la operación producto. Estamos acostumbrados a omitir el símbolo de la operación producto (con ello logramos expresiones más compactas), y no nos genera ningún problema la ausencia del símbolo de la operación producto. Sin embargo Python no funciona como un humano. El intérprete de Python requiere "explícitamente" el símbolo de la operación producto (en su caso el asterísco "*"):

Librería NAcAL para Python

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

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

A la operación de multiplicar un vector por un número también se la denomina "escalar el vector" (...es decir, cambiarlo de escala: piense que 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 intérprete de Python tiene establecidas una serie de precedencias entre operadores. En el siguiente cuadro de operadores, los de arriba preceden a los de abajo:

Precedencia en las operaciones

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

Cuadro 1.1: Precedencia de operadores en Python (más detalles en aquí, en las secciones 6.16. Orden de evaluación y 6.17. Prioridad de operador)

Como la precedencia no es fácil de recordar, por claridad en el código y para asegurar que realizamos las operaciones en el orden deseado, es recomendable usar paréntesis en Python.

Ejercicio 1. ¿Cómo evalúa la librería NAcAL las siguientes expresiones: `(3*b)|2` ; `3*(b|2)` y `3*b|2` ?

Ejercicio 2. ¿Qué falla en la expresión `b+b|2`?

1.2.1. Propiedades de la suma y del producto por un escalar

☞ Explotando las propiedades de la notación (junto con algunas propiedades de los *números reales*) demostraremos ocho propiedades que verifican la *suma de vectores* y el *producto de un vector por un escalar* (Proposición 1.2.1). (ocho propiedades que nos permitirán definir el *espacio vectorial* $\mathbb{R}^n$ en el Tema 2).

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

Recordatorio de algunas propiedades del conjunto de numeros reales

Sean a, b y c números reales. Entonces:

1. $a + b = b + a$.
2. $a + (b + c) = (a + b) + c$.
3. $a + 0 = a$.
4. $a + (-a) = 0$.
5. $ab = ba$.
6. $a(b + c) = ab + ac$.
7. $a(bc) = (ab)c$.
8. $1a = a$.

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 la librería NAcAL):

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 números del ejemplo? o ¿es una regla general para cualquier par de vectores?

2. ¿Es verdad que
$$\begin{pmatrix}1\\2\end{pmatrix} + \left(\begin{pmatrix}4\\5\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix}\right) = \left(\begin{pmatrix}1\\2\end{pmatrix} + \begin{pmatrix}4\\5\end{pmatrix}\right) + \begin{pmatrix}1\\0\end{pmatrix}?$$
El paréntesis significa que primero hay que hacer la suma dentro del paréntesis, y al vector resultante sumarle el vector de fuera.

¿Es cierto solo para los números 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 los vectores del ejemplo? o ¿es una regla general para cualquier par de vectores?

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 los vectores del ejemplo? o ¿es una regla general para cualquier par de vectores?

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 de dentro del 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 de dentro del paréntesis) ¿Es una regla general para cualquier par de vectores? ¿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 más arriba.

8. ¿Es verdad que si $a = 1$
$$a\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}1\\2\\3\end{pmatrix}?$$
¿Es cierto solo para el escalar $a = 1$?

En el EJERCICIO 3 se le pide que demuestre las ocho propiedades que acabamos de revisar[5]. Su solución aparece al final de este documento ¡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 los componentes de los vectores (que son números reales) aplique la propiedad de las operaciones entre números reales que necesite (de entre las del recordatorio de la Página 8).

EJERCICIO 3. 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)* λ *y* η*, se verifican las siguientes ocho propiedades:*

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

[5] 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 o 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 son sencillas es posible que le cuesten (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* (cuando acabe el curso se dará cuenta de que ha aprendido mucho con ello).

- **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 “ $_{|i}$ ”).
  - $(a\boldsymbol{a})_{|i}$ y $a(\boldsymbol{a}_{|i})$ (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.

Nota 1. *Estas ocho propiedades nos permitirán definir el* espacio vectorial $\mathbb{R}^n$ *en el Tema 2.*

1.3. Vectores como representación de datos

Como los vectores son un *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), R. Unido (4) y EEUU (5); entonces podemos interpretar los distintos componentes del siguiente vector en función de su posición en la lista:

$$\boldsymbol{datos(2003)} = \begin{pmatrix} 2.9 \\ 0.9 \\ -0.2 \\ 2.5 \\ 1.8 \end{pmatrix}.$$

Puesto que el número 1.8 está en la quinta posición, sabemos que corresponde al quinto país, y por tanto sabemos que EEUU creció a una tasa del 1.8 % en el año 2003[6]. ¿Qué país no creció en el año 2003?

En algunos manuales verá una interpretación de los vectores como segmentos con norma y dirección. Ésta es una interpretación frecuente en aplicaciones físicas, pero poco natural para los economistas. Para un economista la interpretación más natural es que los componentes de los vectores de $\mathbb{R}^n$ *son datos (precios, cantidades, número de parados, tasa de inflación, etc.) y la posición de cada componente indica a quien corresponde el dato (como en el ejemplo anterior). Ésta será la interpretación que use en la asignatura de Econometría dentro de un par de cursos.*

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

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

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

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

[6] fuente: Informe ICAE sobre la economía española (Octubre 2006), publicado por el Instituto Complutense de Análisis Económico

Podemos escribir todos los datos juntos y dar al resultado el nombre de PIB. Así tendremos recogidos en PIB los datos de crecimiento del PIB de los cinco países en los años 2003–2006. Esto se hace construyendo una *matriz*.

Llamamos *matriz* a una lista ordenada de vectores *con el mismo número de componentes* (... es decir...¡un vector de vectores!). De esta manera podremos definir la matriz PIB como una *sistema* de vectores con los datos de los cinco años disponibles:

$$\Big[\boldsymbol{datos}(\mathbf{2003});\quad \boldsymbol{datos}(\mathbf{2004});\quad \boldsymbol{datos}(\mathbf{2005});\quad \boldsymbol{datos}(\mathbf{2006})\Big].$$

Ahora bien, como los vectores se pueden escribir tanto en vertical como en horizontal, esta escritura puede arrojar una lista muy larga y visualmente difícil de leer:

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

Sabemos que las cuartas componentes de cada uno de los vectores dentro de la matriz corresponden al Reino Unido...pero esta disposición de los datos no es fácil de leer [7]. Para facilitar la lectura dispondremos los datos según el siguiente convenio: disponemos los datos referidos al año 2003 en 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 ese modo se obtiene el siguiente arreglo rectangular de los datos:

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

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

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

$$\big[\boldsymbol{a};\ \boldsymbol{b};\ \dots\ \boldsymbol{v}\big] \qquad \textit{donde } \boldsymbol{a},\boldsymbol{b},\dots\boldsymbol{v} \textit{ son } n \textit{ vectores de } \mathbb{R}^m.$$

Cada uno de los vectores del sistema es una columna *de la matriz.*

Tampoco con las matrices usaremos la representación genérica de los sistemas ya que, aunque mantendremos los corchetes, omitiremos los puntos y comas que separan las columnas, escribiendo sencillamente:

$$\big[\boldsymbol{a}\ \boldsymbol{b}\ \cdots\ \boldsymbol{v}\big].$$

Por tanto, seguiremos 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

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

Y encerrando entre corchetes una lista de vectores (todos con el mismo número de componentes), escribimos la matriz cuyas columnas son los vectores de la lista y en el mismo orden en el que aparecen en la lista. Por ejemplo, si:

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

y encerramos “entre corchetes” algunos de estos vectores, creamos matrices cuyas columnas son los vectores de la lista escritos en vertical; por ejemplo,

$$\begin{bmatrix}\boldsymbol{a} & \boldsymbol{b} & \boldsymbol{0} & \boldsymbol{b}\end{bmatrix} = \begin{bmatrix}1&2&0&2\\0&2&0&2\\3&2&0&2\\0&2&0&2\end{bmatrix}; \qquad \begin{bmatrix}\boldsymbol{0} & \boldsymbol{0}\end{bmatrix} = \begin{bmatrix}0&0\\0&0\\0&0\\0&0\end{bmatrix}.$$

[7] aunque un ordenador no tiene ningún problema con esta representación, los humanos si lo tenemos...¡imagine una lista de 12 vectores con 20 componentes cada uno. Algo similar ocurre si escribimos varios vectores en columna, unos debajo de otros.

Definición 1.9. *El* orden *de la matriz indica su número de filas y columnas, y está compuesto por dos números: el primero corresponde al número de filas (m) y el segundo al número de columnas (n). El orden se suele expresar cómo "$m \times n$" (y se lee "m por n").*

En el ejemplo con datos económicos, el orden de la matriz $\mathbf{PIB}$ es 5 por 4. Nótese que siempre *se indica primero el número de filas y luego el de columnas.* Frecuentemente expresaremos el orden de la matriz debajo de su nombre; por ejemplo: $\underset{5\times 4}{\mathbf{PIB}}$.

A cada número dentro de la matriz lo denominamos *componente* (o elemento), y al igual que en el caso de los vectores, identificamos los componentes de la matriz por la posición que ocupan. De esta manera podemos decir que el componente de $\mathbf{PIB}$ ubicado en la primera fila y segunda columna es 3.1.

Notación. Denotamos las matrices de $\mathbb{R}^{m\times n}$ con letras mayúsculas y en negrita: $\mathbf{A}, \mathbf{B}, \mathbf{C}$; y a los componentes genéricos de la matriz con la misma letra con la que denotamos a la matriz completa, pero en minúscula cursiva y con dos subíndices que indican la posición que ocupa el componente dentro de la matriz, comenzando siempre por la fila, y luego por la columna (por tanto el componente a_{ij} de $\mathbf{A}$ se encuentra en la fila $i$ésima y columna $j$ésima). Otras formas que usaremos para denotar dicho componente son:

$$a_{ij} \;\equiv\; \text{elem}_{ij}(\mathbf{A}) \;\equiv\; {}_{i|}\mathbf{A}_{|j}.$$

Por ejemplo, escribimos una matriz genérica $\mathbf{A}$ de orden 2 por 3 como:

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

Y de la matriz

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

podemos afirmar que su orden es 2×3 y que su componente b_{21} es cero, y su componente ${}_{2|}\mathbf{B}_{|2}$ es -1.

terminal de Python

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

Nótese que usamos un solo índice para identificar las componentes de un vector $\boldsymbol{x}$, pero que necesitamos *dos* índices para identificar las componentes de una matriz $\mathbf{A}$. Así, cuando nos queremos referir a la componente $i$ésima del vector $\boldsymbol{x}$ escribimos:

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

y cuando queremos referirnos al componente situado en la fila $i$ésima y columna $j$ésima de la matriz $\mathbf{A}$:

$$\text{elem}_{ij}(\mathbf{A}), \qquad \text{ó} \qquad a_{ij}, \qquad \text{ó} \qquad {}_{i|}\mathbf{A}_{|j}.$$

Pero ¿qué pasa si nos queremos referir a toda la fila $i$ésima o toda la columna $j$ésima de $\mathbf{A}$?

1.4.1. Filas de una matriz

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

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

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

El operador selector de filas: "$i|$". Al escribir "$i|$" como subíndice a la *izquierda* de la matriz seleccionamos la *fila* $i$ésima de la matriz.

$$\text{Por ejemplo, si}\quad \mathbf{B} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 1 \end{bmatrix}; \qquad \text{entonces}\quad {}_{1|}\mathbf{B} = \begin{pmatrix} 1, & 2, & 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Librería NAcAL para Python

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

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

1.4.2. Columnas de una matriz

La columna $j$ésima de $\mathbf{A}$ (de orden $m \times n$) es el vector de $\mathbb{R}^m$ con componentes a_{kj}, donde el índice k recorre todas las filas ($k = 1 : m$). Denotaremos dicha columna como $\text{col}_j(\mathbf{A})$ o bien como $\mathbf{A}_{|j}$:

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

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

El operador selector de columnas: "$|j$". Al escribir "$|\,j$" como subíndice a la *derecha* de la matriz, seleccionamos la *columna* $j$ésima de la matriz.

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

Librería NAcAL para Python

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

¿Y cómo escribir la submatriz de $\mathbf{B}$ de orden $m \times 1$ cuya única columna es la $j$ésima de $\mathbf{B}$? Basta encerrar entre "corchetes" dicha columna:

$$\begin{bmatrix} \mathbf{B}_{|2} \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Librería 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* (cuya lista contiene dos números que se puede escribir en vertical u horizontal), pero que $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ es una *matriz* cuya lista contiene una única *columna*. Por tanto son objetos matemáticos distintos (lista de números *vs* lista de vectores), y su escritura es fácil de distinguir: los vectores se encierran entre "paréntesis" y matrices entre "corchetes".

Así $\boldsymbol{a} = \begin{pmatrix}1, & 2, & 3\end{pmatrix}$ es un vector y $\mathsf{A} = \begin{bmatrix}1 & 2 & 3\end{bmatrix}$ es una matriz (y por tanto **son objetos distintos**).[8]

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

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

Diremos que dos matrices del mismo orden son iguales si y solo si lo son sus correspondientes sistemas de vectores, es decir: $\mathsf{A} = \mathsf{B}$ si y solo si $\mathsf{A}_{|j} = \mathsf{B}_{|j}$ para $j = 1:n$ (i.e. si son iguales columna a columna).

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

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

Notaciones alternativas

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

En el ejemplo con datos económicos, ${}_{1|}\mathsf{PIB}$ corresponde a los datos de España y ${}_{5|}\mathsf{PIB}$ a los de EEUU; y por otra parte, $\mathsf{PIB}_{|3}$ corresponde a los datos del año 2005.

[8] **Nota para profesores de matemáticas y/o lectores "puntillosos"**. Tradicionalmente se sigue el siguiente convenio

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

Sin embargo, aquí sigo un convenio alternativo acorde con el funcionamiento interno de Python:

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

(donde $\emptyset$ denota al conjunto vacío). De tal manera que $\mathbb{R}^1 = \mathbb{R}\times\{\emptyset\} = \{(\alpha,\emptyset) \mid \alpha\in\mathbb{R}\}$. Así pasa en Python, y por ello una tupla con un único elemento se escribe con `(a,)`. Nótese la coma detrás de la `a` (que se puede entender que separa el primer elemento, `a`, del último, que es vacío). De este modo, la selección de los elementos sigue la siguiente función recursiva: si $\boldsymbol{v} = (v_1, \boldsymbol{w}) \in \mathbb{R}^n$, donde $\boldsymbol{w}\in\mathbb{R}^{n-1}$ entonces

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

Por tanto, $\boldsymbol{a} = \begin{pmatrix}1, & 2, & 3\end{pmatrix}$ es un sistema de tres *números* de $\mathbb{R}$, y $\mathsf{A} = \begin{bmatrix}1 & 2 & 3\end{bmatrix}$ es un sistema de tres *vectores* de $\mathbb{R}^1$, a saber $(1,)$, $(2,)$ y $(3,)$.

☞ *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 el que aparecen en la lista dentro del corchete. Por ejemplo

$$\boldsymbol{a} = \begin{pmatrix}1, & 2\end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix}1, & 0\end{pmatrix}, \quad \boldsymbol{c} = \begin{pmatrix}9, & 2\end{pmatrix}; \qquad \mathbf{A} = \begin{bmatrix}\boldsymbol{a} \; \boldsymbol{b} \; \boldsymbol{c}\end{bmatrix} = \begin{bmatrix}1 & 1 & 9\\ 2 & 0 & 2\end{bmatrix}.$$

Con las matrices NO hay la 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; $\begin{bmatrix}\mathbf{A}_{|3}\end{bmatrix}$ es una *matriz columna* cuya única columna es igual que la tercera columna de $\mathbf{A}$.

$$\mathbf{A}_{|3} = \begin{pmatrix}9 & 2\end{pmatrix} = \begin{pmatrix}9\\ 2\end{pmatrix} \quad \neq \quad \begin{bmatrix}\mathbf{A}_{|3}\end{bmatrix} = \begin{bmatrix}9\\ 2\end{bmatrix} \quad \neq \quad \begin{bmatrix}9 & 2\end{bmatrix}.$$

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.

Nótese que ${}_{i|}\mathbf{A}_{|j}$ es la componente $j$ésima de la fila i, es decir $\left({}_{i|}\mathbf{A}\right)_{|j}$, pero también es la componente $i$ésima de la columna j, por lo que sería conveniente poder escribir ${}_{i|}\left(\mathbf{A}_{|j}\right)$, de manera que

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

Para ello aceptaremos que el operador "$|$" también pueda operar por la izquierda de un *vector*. Así

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

Con esta flexibilización de la notación, y puesto que tanto las filas como las columnas de una matriz son vectores, cuando seleccionemos las *componentes* de $\mathbf{A}$, nos dará igual operar dos veces por el mismo lado o una vez por cada lado (ambas operaciones arrojan necesariamente el mismo resultado). Así

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

Para no perder la *asociatividad* de la notación cuando "$|$" opera por la izquierda de los *vectores*, también aceptaremos que los escalares aparezcan multiplicando por la derecha.[9] Es decir, $\lambda\boldsymbol{b} = \boldsymbol{b}\lambda$. Esto da lugar a dos nuevas reglas de reescritura:

[9]Como en estas notas propongo una notación propia, aquí aceptaremos una expresión como "$\boldsymbol{a}2$"...¡expresión que nunca aparece en los textos! El **modo habitual de escribir** dicho producto es "$2\boldsymbol{a}$" con "el coeficiente primero". Aquí lo aceptaremos para mantener la asociatividad por la izquierda, pero tenga en cuenta que no es una forma empleada habitualmente.

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

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

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

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

De hecho, 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 el paréntesis y escribir sencillamente $\lambda\boldsymbol{b}_{|i}$ o también cualquiera de las siguientes reordenaciones $\lambda\boldsymbol{b}_{|i} = \lambda_{i|}\boldsymbol{b} = \boldsymbol{b}_{|i}\lambda = {}_{i|}\boldsymbol{b}\lambda$.

No obstante a lo anterior, en las demostraciones siempre usaré expresiones con paréntesis (aunque sea innecesario), ya que ayuda a identificar la propiedad empleada en cada paso de la demostración.

Recuerde que en Python también es coveniente usar paréntesis para evitar errores y sorpresas.

Ejercicio 4. ¿Cómo evalúa la librería NAcAL las siguientes expresiones? (Revise la Tabla 1.2 en la página 7)

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

1.4.5. La transposición

Encerrando $\boldsymbol{a}$ entre corchetes denotamos una matriz con una única columna; para denotar una matriz cuya única única fila es $\boldsymbol{a}$ necesitamos una operación adicional: la *transposición*, $(^\intercal)$,...

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

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

Por tanto, la transpuesta de $\underset{m\times n}{\mathbf{A}}$ es la matriz n por m cuyas columnas son las filas de $\mathbf{A}$ escritas en vertical, es decir,

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

Ejemplo: $\quad \mathbf{A} = \begin{bmatrix} 1\\0\\5\\1 \end{bmatrix} \Rightarrow \mathbf{A}^\intercal = \begin{bmatrix} 1 & 0 & 5 & 1 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 1 & 2 & 3\\ 0 & -1 & 1 \end{bmatrix} \Rightarrow \mathbf{B}^\intercal = \begin{bmatrix} 1 & 0\\ 2 & -1\\ 3 & 1 \end{bmatrix}.$

Librería NAcAL para Python

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

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

Escribimos las *matrices columna* (matrices con una única columna) poniendo un vector entre corchetes; ahora podemos escribir *matrices fila* (matrices con una única fila) trasponiendo matrices columna.

Así $\big[{}_{i|}\mathbf{A}\big]^{\intercal}$ es una matriz cuya única fila es igual a la fila $i$ésima de $\mathbf{A}$; y si

$$\mathbf{B}=\begin{bmatrix}1&2&3\\0&-1&1\end{bmatrix};\quad\text{entonces}\quad\begin{cases}\big[{}_{1|}\mathbf{B}\big]^{\intercal}=\begin{bmatrix}1&2&3\end{bmatrix}\\[2ex]\big[\mathbf{B}_{|3}\big]^{\intercal}=\begin{bmatrix}3&1\end{bmatrix}\end{cases}.$$

Librería NAcAL para Python

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

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

Librería NAcAL para Python

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

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

EJERCICIO 5. Demuestre las siguientes propiedades elementales de la transposición:

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

(b) Si $\underset{m\times n}{\mathbf{A}}$ entonces $\big(\mathbf{A}^{\intercal}\big)^{\intercal}=\mathbf{A}$.

(c) Si $\underset{m\times n}{\mathbf{A}}$ y $1\leq i\leq n$ entonces $\mathbf{A}_{|i}={}_{i|}\big(\mathbf{A}^{\intercal}\big)$.

☞ Pese a que las expresiones ${}_{i|}\big(\mathbf{A}^{\intercal}\big)$, $\big(\mathbf{A}^{\intercal}\big)_{|j}$ y ${}_{i|}\big(\mathbf{A}^{\intercal}\big)_{|j}$ 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 escribiré el paréntesis. Así la notación es más clara.

Fíjese además en las reglas de reescritura: "*puedo transponer (o quitar la transpuesta) solo si además cambio de lado los subíndices*" (fíjese en su aplicación en las demostraciones de las propiedades anteriores).

1.4.6. Definiciones de algunas matrices especiales

☞ Al igual que hicimos con los vectores, añadiremos algunas definiciones relativas a las matrices. *Preste atención al uso de la notación.*

Definición 1.11. *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 $j=1:n$.*

Librería NAcAL para Python

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

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

Definición 1.12. *Sea la matriz $\mathbf{A}$. Llamamos matriz opuesta de $\mathbf{A}$ a la matriz $-\mathbf{A}$, del mismo orden que $\mathbf{A}$ y cuyas columnas son las de $\mathbf{A}$ multiplicadas por -1, es decir,* $\big(-\mathbf{A}\big)_{|j}=-\big(\mathbf{A}_{|j}\big)$.

Librería NAcAL para Python

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

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

También daremos nombre a ciertas matrices en función de la disposición de sus componentes:

Definición 1.13. *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 1.14. *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.

Definición 1.15 (Matriz simétrica)**.** *Decimos que la matriz $\mathbf{A}$ es simétrica cuando $\mathbf{A} = \mathbf{A}^{\intercal}$.*

Así, $\mathbf{A}_{|j} = \left(\mathbf{A}^{\intercal}\right)_{|j} = {}_{j|}\mathbf{A}$. Por tanto, toda matriz simétrica es necesariamente cuadrada. Y puesto que la "*columna*" $j$ésima es igual a la "*fila*" $j$ésima, sus componentes $i$ésimas son iguales: ${}_{i|}\mathbf{A}_{|j} = {}_{j|}\mathbf{A}_{|i}$.

Definición 1.16. *Decimos que una matriz es diagonal cuando todos los componentes fuera de la diagonal principal son nulos, es decir, si $d_{ij} = 0$ cuando $i \neq j$.*

Por ejemplo

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

Fíjese que de las tres matrices diagonales, solo $\mathbf{D}$ es cuadrada. Tenga en cuenta además, que los componentes de la diagonal principal pueden tomar cualquier valor. Así pues, ¡toda matriz nula $\mathbf{0}$ es diagonal! Además, *toda matriz diagonal y cuadrada es simétrica*, ya que por una parte

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

Definición 1.17. *Llamamos matriz identidad (y la denotamos con $\mathbf{I}$) a la matriz cuadrada y diagonal cuyas componentes en la diagonal principal son unos y el resto de componentes son cero, 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: $\underset{2\times 2}{\mathbf{I}} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}; \quad \underset{1\times 1}{\mathbf{I}} = \begin{bmatrix}1\end{bmatrix}.$

Librería NAcAL para Python

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

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

Nótese que las matrices identidad son simétricas (pues son cuadradas y diagonales).

1.5. Suma de matrices y producto de matrices por escalares

Definición 1.18 (suma de matrices)**.** *Definimos 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* $\mathbf{B}$*:*

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

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

Definición 1.19 (producto de una matriz por un escalar)**.** *Definimos el producto de una matriz* $\underset{m\times n}{\mathbf{A}}$ *por un escalar* λ *como la matriz resultante de multiplicar las columnas de* $\mathbf{A}$ *por el escalar* λ*:*

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

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

Definición 1.20. *Para cualquier* λ *decimos que* $\lambda\mathbf{A}$ *es* un múltiplo *de* $\mathbf{A}$*.*

☞ De nuevo la definición de las operaciones de suma y producto por un escalar convierten al operador "$_{|j}$" en un **operador lineal**; es decir, *distributivo* respecto a la suma de matrices

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

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

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

1.5.1. Propiedades de la suma de matrices y del producto por un escalar

La suma de matrices y producto de matrices por escalares verifican propiedades análogas a las de la Proposición 1.2.1; y su demostración es idéntica a la del Ejercicio 3…

EJERCICIO 6. 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* λ *y* η*, 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. $\mathbf{0}+\mathbf{A} = \mathbf{A}$.
4. $\mathbf{A}+(-\mathbf{A}) = \mathbf{0}$.
5. $\lambda(\mathbf{A}+\mathbf{B}) = \lambda\mathbf{A}+\lambda\mathbf{B}$.
6. $(\lambda+\eta)\mathbf{A} = \lambda\mathbf{A}+\eta\mathbf{A}$.
7. $\lambda(\eta\mathbf{A}) = (\lambda\eta)\mathbf{A}$.
8. $1\mathbf{A} = \mathbf{A}$.

Pista. Explote que el operador "$_{|j}$" es distributivo para la suma de matrices y asociativo para el producto por escalares. Use las propiedades de las operaciones con vectores de la Proposición 1.2.1, pues $\mathbf{A}_{|j}$, $\mathbf{B}_{|j}$ y $\mathbf{C}_{|j}$ son vectores de $\mathbb{R}^n$.

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 el Tema 2 el espacio vectorial de matrices* $\mathbb{R}^{m\times n}$*.*

1.5.2. Operaciones componente a componente

La mayoría de manuales definen la suma y el producto por escalares componente a componente:

> *"la suma de dos matrices* $\mathbf{A}$ *y* $\mathbf{B}$*, del mismo orden, es la matriz que resulta de sumar los componentes de* $\mathbf{A}$ *a los componentes de* $\mathbf{B}$*".*

y

> *"el producto de una matriz* $\mathbf{A}$ *por un escalar* λ*: es la matriz resultante de multiplicar los componentes de* $\mathbf{A}$ *por el escalar* λ*".*

EJERCICIO 7. Demuestre que las definiciones 1.18 y 1.19 implican que las operaciones de suma y producto por escalares se pueden calcular componente a componente:

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

Es decir,

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

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

y

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

1.5.3. La transposición es un operador lineal

Antes de operar por filas necesitamos demostrar las propiedades de linealidad de la transpuesta.[10]

EJERCICIO 8. Demuestre las siguientes proposiciones.

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

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

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

1.5.4. Operaciones fila a fila

Y ahora veamos las operaciones por filas:

EJERCICIO 9. Demuestre que las definiciones 1.18 y 1.19 implican que se puede operar "por filas", es decir:

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

[10]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 $\mathbf{AB}$.

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

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

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

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

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

Si de nuevo aceptamos el producto de un escalar por el lado derecho, $\mathbf{A}\lambda = \lambda\mathbf{A}$; entonces logramos unas reglas de reescritura similares a las que obtuvimos en el caso del los vectores.

☞ **Reglas de re-escritura:** como **el operador selector de filas es lineal**, podemos

- "*distribuir* el operador" entre sumandos

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

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

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

Y además también podemos

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

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

Como las cuatro últimas expresiones de la mitad izquierda del anterior recuadro arrojan el mismo resultado (pese al distinto orden de ejecución de las operaciones) podemos omitir el paréntesis y escribir $\lambda\mathbf{A}_{|j}$. Y como las cuatro últimas expresiones de la mitad derecha del recuadro arrojan el mismo resultado (pese al distinto orden de ejecución de operaciones) también podemos omitir el paréntesis y escribir $\lambda_{i|}\mathbf{A}$. Por motivos didácticos *en las demostraciones siempre usaré los paréntesis para aclarar qué regla o propiedad se aplica en cada paso*, pero lo más práctico es omitirlos si no son necesarios (¡esa es la ventaja de la **notación asociativa**!).[11]

La Propiedad 1.5.2 sobre la "traspuesta de un múltiplo" (que volvemos a copiar como primer punto del recuadro de más abajo) junto a permitir multiplicar por la derecha, $\mathbf{A}\lambda = \lambda\mathbf{A}$; nos dota de otra regla de reescritura que nos permite intercambiar la posición entre el escalar y el símbolo de trasposición.

[11] Pero recuerde que dada la precedencia establecida en la ejecución de las operaciones en Python, dicha ventaja desaparece en la librería NAcAL.

☞

Reglas de re-escritura: podemos

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

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

Además, como **el operador transposición es lineal**, también podemos

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

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

- "*distribuir* el operador" en una suma

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

Combinaciones lineales

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

Antes de definir las combinaciones lineales, 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*

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

Este producto será muy importante en la segunda parte del curso para definir la ortogonalidad. Pero su introducción en este momento nos permitirá dotar de propiedades muy potentes a la notación del producto de matrices con vectores, además de lograr demostraciones mucho más "limpias" (pues en lugar de sumatorios podremos usar productos punto) y con pasos más sencillos (tan solo habrá que aplicar repetidamente la linealidad tanto del producto punto como del operador selector).

Librería NAcAL para Python

```
a = Vector( (1, 2, 3, 4) )
b = Vector( (1,-1, 1,-1) )
a * b          # producto punto entre los vectores a y b
```

$$-2$$

Propiedades del producto punto

El producto punto (o producto escalar usual de $\mathbb{R}^n$) satisface los siguientes axiomas para cualesquiera vectores $\boldsymbol{x}$, $\boldsymbol{y}$ y $\boldsymbol{z}$ de $\mathbb{R}^n$ y para cualquier escalar a de $\mathbb{R}$.

Ejercicio 10. Demuestre que el producto punto cumple con los siguientes axiomas:

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

(b) **Linealidad respecto al primer argumento:**
   1. $(a\boldsymbol{x})\cdot\boldsymbol{y} = a(\boldsymbol{x}\cdot\boldsymbol{y})$
   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:** $\boldsymbol{x}\cdot\boldsymbol{x} = 0 \Leftrightarrow \boldsymbol{x} = 0$.

Fíjese que como el producto punto es simétrico, también es lineal en el segundo argumento

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

y además se verifica que $\boldsymbol{x}\cdot(a\boldsymbol{y}) = (a\boldsymbol{x})\cdot\boldsymbol{y}$.

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) iésima de la matriz identidad de orden n, se seleccciona la componente iésima del vector,* $\left({}_{i|}\mathbf{I}\right) \cdot \boldsymbol{x} \;=\; {}_{i|}\boldsymbol{x}$:

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

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 veamos 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 los vectores* $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$*. Llamamos combinación lineal a cualquier suma de múltiplos de dichos vectores:*

$$a_1\boldsymbol{b}_1 + a_2\boldsymbol{b}_2 + \cdots + a_n\boldsymbol{b}_n$$

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

Ejercicio 11. 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 1. 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 los componentes de* $\boldsymbol{b}$*:*

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

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

Fíjese 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!)

Librería 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 12. Escriba como producto de una matriz por un vector la combinación lineal del ejercicio 11.

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 1 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 resulta evidente 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.

EJERCICIO 13. 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 14. Demuestre que la componente $i$ésima del vector $\mathbf{A}\boldsymbol{b}$ es el producto punto $\left({}_{i|}\mathbf{A}\right)\cdot\boldsymbol{b}$:

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

Librería 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:

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

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

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 importantísimas propiedades: las *propiedades de linealidad*. La primera respecto a la suma de vectores, $\mathsf{A}\big(\boldsymbol{b}+\boldsymbol{c}\big) = \mathsf{A}\boldsymbol{b}+\mathsf{A}\boldsymbol{c}$, y la segunda respecto al producto por un escalar, $\mathsf{A}(\lambda\boldsymbol{b}) = \lambda(\mathsf{A}\boldsymbol{b})$. Veámoslo.

Ejercicio 15. Demuestre las siguientes proposiciones (inténtelo primero sin mirar las pistas).

(a) **Proposición 2.2.1.** *Sea la matriz* $\underset{m\times n}{\mathsf{A}}$ *y los vectores* $\boldsymbol{b}$ *y* $\boldsymbol{c}$ *de* $\mathbb{R}^n$*, entonces:* $\mathsf{A}\big(\boldsymbol{b}+\boldsymbol{c}\big) = \mathsf{A}\boldsymbol{b}+\mathsf{A}\boldsymbol{c}$.

Pista. Demuestre que ambos vectores son iguales componente a componente, es decir, que $_{i|}\big(\mathsf{A}(\boldsymbol{b}+\boldsymbol{c})\big)$ es igual a $_{i|}\big(\mathsf{A}\boldsymbol{b}+\mathsf{A}\boldsymbol{c}\big)$. Para ello emplee que cada componente es un producto punto y que los productos escalares también son lineales en el segundo argumento. Use también que el operador selector es lineal.

(b) **Proposición 2.2.2.** *Sea la matriz* $\underset{m\times n}{\mathsf{A}}$ *, el vector* $\boldsymbol{b}$ *de* $\mathbb{R}^n$ *y el escalar* λ*, entonces:* $\mathsf{A}\big(\lambda\boldsymbol{b}\big) = \lambda\big(\mathsf{A}\boldsymbol{b}\big)$.

Pista. Siga la misma estrategia que en el apartado anterior.

☞ El producto de una matriz por un vector es un **operador lineal**; es decir, *distributivo* respecto a la suma de vectores

$$\mathsf{A}\big(\boldsymbol{b}+\boldsymbol{c}\big) = \mathsf{A}\boldsymbol{b}+\mathsf{A}\boldsymbol{c}.$$

y *asociativo* respecto al producto de un vector por un escalar

$$\mathsf{A}\big(\lambda\boldsymbol{b}\big) = \lambda\big(\mathsf{A}\boldsymbol{b}\big).$$

Ejercicio 16. Demuestre el siguiente corolario (inténtelo primero sin mirar las pistas).

Corolario 2.2.3. *Sea la matriz* $\underset{p\times m}{\mathsf{A}}$ *, la matriz* $\underset{m\times n}{\mathsf{B}}$ *, y el vector* $\boldsymbol{c}$ *de* $\mathbb{R}^n$*, entonces:*

$$\mathsf{A}\big(\mathsf{B}\boldsymbol{c}\big) = \Big[\mathsf{A}\big(\mathsf{B}_{|1}\big)\ \dots\ \ \mathsf{A}\big(\mathsf{B}_{|n}\big)\Big]\,\boldsymbol{c},$$

donde $\mathsf{B}_{|i}$ *es la iésima columna de* B*.*

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 17. Demuestre las siguientes proposiciones (inténtelo sin mirar las pistas).

(a) **Proposición 2.2.4.** *Sea la matriz* $\underset{m\times n}{\mathsf{A}}$ *, el vector* $\boldsymbol{b}$ *de* $\mathbb{R}^n$ *y el escalar* λ*, entonces:* $\mathsf{A}(\lambda\boldsymbol{b}) = (\lambda\mathsf{A})\boldsymbol{b}$.

Pista. Recuerde que el producto punto es lineal en el primer argumento: $\boldsymbol{x}\cdot(\lambda\boldsymbol{y}) = (\lambda\boldsymbol{x})\cdot\boldsymbol{y}$.

(b) **Proposición 2.2.5.** *Sean las matrices* $\underset{m\times n}{\mathsf{A}}$ *y* $\underset{m\times n}{\mathsf{B}}$ *y el vector* $\boldsymbol{c}$ *de* $\mathbb{R}^n$ *entonces:* $(\mathsf{A}+\mathsf{B})\boldsymbol{c} = \mathsf{A}\boldsymbol{c}+\mathsf{B}\boldsymbol{c}$.

Más reglas de reescritura.

De las proposiciones 2.2.2 y 2.2.4 se deduce que

$$(\lambda\mathsf{A})\boldsymbol{b} = \mathsf{A}(\lambda\boldsymbol{b}) = \lambda(\mathsf{A}\boldsymbol{b})$$

y por tanto no son necesarios los paréntesis. Así que sin ninguna ambigüedad podemos escribir:

$$\lambda\mathsf{A}\boldsymbol{b} = \mathsf{A}\lambda\boldsymbol{b}.$$

Observe además que (como aceptamos que el escalar multiplique 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***

Propiedades de linealidad

- $\mathbf{A}\big(\boldsymbol{b}+\boldsymbol{c}\big) = \mathbf{A}\boldsymbol{b} + \mathbf{A}\boldsymbol{c}$
- $\mathbf{A}(\lambda\boldsymbol{b}) = \lambda(\mathbf{A}\boldsymbol{b})$

Otras propiedades

- $\mathbf{A}\big(\mathbf{B}\boldsymbol{c}\big) = \Big[\mathbf{A}\big(\mathbf{B}_{|1}\big), \quad \dots, \quad \mathbf{A}\big(\mathbf{B}_{|n}\big)\Big]\,\boldsymbol{c}$
- $(\mathbf{A}+\mathbf{B})\boldsymbol{c} = \mathbf{A}\boldsymbol{c} + \mathbf{B}\boldsymbol{c}$
- $(\lambda\mathbf{A})\boldsymbol{b} = \lambda(\mathbf{A}\boldsymbol{b})$

2.3. Producto de un vector por una matriz (vector a la izquierda)

☞ Hasta aquí hemos visto que matriz por vector ($\mathbf{A}\boldsymbol{b}$) es una combinación lineal de las *columnas* de la matriz.

A continuación veremos que vector por matriz ($\boldsymbol{a}\mathbf{B}$) es una combinación lineal de las *filas* de la matriz (con propiedades análogas).

Definición 2.4 (producto de un vector por una matriz)**.** *El producto de un vector $\boldsymbol{a} \in \mathbb{R}^m$ que multiplica por la izquierda a una matriz $\mathbf{B}$ de m filas es la combinación lineal de las m filas de $\mathbf{B}$ cuyos coeficientes son los componentes de $\boldsymbol{a}$:*

$$\boldsymbol{a}\mathbf{B} = \big(\mathbf{B}^{\intercal}\big)\boldsymbol{a}$$

Puesto que el producto $\boldsymbol{a}\mathbf{B}$ es una combinación lineal de las filas de $\mathbf{B}$, dicho producto es un vector de $\mathbb{R}^n$.

Para recordar que estamos operando con las filas de $\mathbf{B}$, escribiremos $\boldsymbol{a}$ en forma de fila (en horizontal):

$$\boldsymbol{a}\mathbf{B} = \big(a_1, \quad a_2, \quad \dots \quad a_m\big)\begin{bmatrix} b_{11} & \dots & b_{1n} \\ b_{21} & \dots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{bmatrix} = a_1\big({}_{1|}\mathbf{B}\big) + \cdots + a_m\big({}_{m|}\mathbf{B}\big) \quad \in \mathbb{R}^n.$$

Librería NAcAL para Python

```
print( b*A == (~A)*b )
b*A
```

True

$$\big(0, \quad -2, \quad -1\big)$$

Cada componente $j$ésima del vector $\boldsymbol{a}\mathbf{B}$ es el producto punto entre $\boldsymbol{a}$ y la columna $j$ésima de $\mathbf{B}$:

$$\big(\boldsymbol{a}\mathbf{B}\big)_{|j} = \Big(\big(\mathbf{B}^{\intercal}\big)\boldsymbol{a}\Big)_{|j} = {}_{j|}\Big(\big(\mathbf{B}^{\intercal}\big)\boldsymbol{a}\Big) = \Big({}_{j|}\big(\mathbf{B}^{\intercal}\big)\Big)\cdot\boldsymbol{a} = \big(\mathbf{B}_{|j}\big)\cdot\boldsymbol{a} = \boldsymbol{a}\cdot\big(\mathbf{B}_{|j}\big), \quad \text{donde } j = 1:n.$$

Librería 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 18. Demuestre las siguientes propiedades del producto $\boldsymbol{a}\mathbf{B}$.

(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 como $\mathbf{I}$ es simétrica, $\mathbf{I}^{\intercal} = \mathbf{I}$.

(b) Sea $\left({}_{i|}\mathbf{I}\right)$ la fila *i*ésima de $\underset{m\times m}{\mathbf{I}}$ y la matriz $\underset{m\times n}{\mathbf{A}}$, entonces: $\boxed{\left({}_{i|}\mathbf{I}\right)\mathbf{A} = {}_{i|}\mathbf{A}.}$

Pista. Recuerde el caso especial 3 en la página 25 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 recuerde que $\boldsymbol{v}\mathbf{A} = \left(\mathbf{A}^{\intercal}\right)\boldsymbol{v}$.

(d) Sea el escalar λ, 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 λ 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.

☞ 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 $\mathbf{B}$:

$$\boldsymbol{a}\mathbf{B} \;=\; a_1\left({}_{1|}\mathbf{B}\right) + \cdots + a_m\left({}_{m|}\mathbf{B}\right) \quad \in \mathbb{R}^n.$$

Cada componente del vector $\boldsymbol{a}\mathbf{B}$ se puede expresar como un producto punto:

$$\left(\boldsymbol{a}\mathbf{B}\right)_{|j} = \boldsymbol{a}\cdot\left(\mathbf{B}_{|j}\right), \quad \text{donde } j = 1:n.$$

El producto $\boldsymbol{a}\mathbf{B}$ tiene propiedades análogas a las del producto $\mathbf{A}\boldsymbol{b}$.

- $\boldsymbol{a}\mathbf{I} = \boldsymbol{a}$
- $\left({}_{i|}\mathbf{I}\right)\mathbf{A} = {}_{i|}\mathbf{A}$;

Propiedades de linealidad

- $\left(\boldsymbol{a}+\boldsymbol{b}\right)\mathbf{C} = \boldsymbol{a}\mathbf{C} + \boldsymbol{b}\mathbf{C}$
- $(\lambda\boldsymbol{a})\mathbf{B} = \lambda(\boldsymbol{a}\mathbf{B})$

Otras propiedades

- $\boldsymbol{a}\left(\mathbf{B}+\mathbf{C}\right) = \boldsymbol{a}\mathbf{B} + \boldsymbol{a}\mathbf{C}$
- $\boldsymbol{a}(\lambda\mathbf{B}) = \lambda(\boldsymbol{a}\mathbf{B})$

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 $\mathbf{B}$).

$$\mathbf{AB} \equiv \Big[\mathbf{A}(\mathbf{B}_{|1}), \quad \dots, \quad \mathbf{A}(\mathbf{B}_{|n})\Big].$$

Ejemplo 2. 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

$$\mathbf{AB} = \Big[\mathbf{A}(\mathbf{B}_{|1}); \quad \mathbf{A}(\mathbf{B}_{|2});\Big] = \left[\begin{pmatrix} -3 \\ -3 \\ 6 \end{pmatrix}; \quad \begin{pmatrix} 6 \\ -3 \\ -3 \end{pmatrix};\right] = \begin{bmatrix} -3 & 6 \\ -3 & -3 \\ 6 & -3 \end{bmatrix};$$

nótese que calculamos multiplicando $\mathbf{A}$ por cada columna de $\mathbf{B}$.[1]

Deduzcamos algunas propiedades que verifica el producto de matrices…

EJERCICIO 19. 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 $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$; también el producto de matrices es asociativo y por tanto podemos escribir sencillamente $\mathbf{ABC}$.

[1] Si conoce otro método para calcular el producto, no lo emplee por el momento. AHORA ES MUY IMPORTANTE que se acostúmbre a considerar las columnas de $\mathbf{AB}$ como combinaciones de las columnas de $\mathbf{A}$.

☞ Hemos definido 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 en la definición del producto!

El producto solo está definido si la primera matriz tiene tantas columnas como filas tiene la segunda.

EJERCICIO 20. 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* $\big(\mathbf{A}+\mathbf{B}\big)\mathbf{C} = \mathbf{AC}+\mathbf{BC}$.

(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* $\mathbf{A}\big(\mathbf{B}+\mathbf{C}\big) = \mathbf{AB}+\mathbf{AC}$.

(c) **Proposición 3.1.5.** *Sean las matrices* $\underset{m\times p}{\mathbf{A}}$, $\underset{p\times n}{\mathbf{B}}$ *y el escalar* λ, *entonces* $\mathbf{A}(\lambda\mathbf{B}) = \lambda(\mathbf{AB}) = (\lambda\mathbf{A})\mathbf{B}$.

Pista. Recuerde que $\mathbf{A}\big(\lambda\boldsymbol{b}\big)\begin{cases} = \lambda\big(\mathbf{A}\boldsymbol{b}\big) \\ \\ = \big(\lambda\mathbf{A}\big)\boldsymbol{b}\end{cases}$; (Proposiciones 2.2.2 y 2.2.4 de la Página 26).

(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 25).

(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 $\mathbf{A}\big(\mathbf{I}_{|j}\big) = \mathbf{A}_{|j}$ (Caso Especial 3 en la página 25).

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 $\mathbf{BA}$.[2] Por ejemplo:

$$\begin{bmatrix}1&1\\0&1\end{bmatrix}\begin{bmatrix}1&0\\0&3\end{bmatrix} \neq \begin{bmatrix}1&0\\0&3\end{bmatrix}\begin{bmatrix}1&1\\0&1\end{bmatrix}.$$

Por otra parte: entre los escalares si $\lambda^2 = \lambda$ entonces λ 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 λ 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:

- $\mathbf{A}(\mathbf{B}\boldsymbol{c}) = (\mathbf{AB})\boldsymbol{c}$
- $\mathbf{A}\big(\mathbf{BC}\big) = \big(\mathbf{AB}\big)\mathbf{C}$
- $\big(\mathbf{A}+\mathbf{B}\big)\mathbf{C} = \mathbf{AC}+\mathbf{BC}$
- $\mathbf{A}\big(\mathbf{B}+\mathbf{C}\big) = \mathbf{AB}+\mathbf{AC}$
- $\mathbf{A}(\lambda\mathbf{B}) = \lambda(\mathbf{AB}) = (\lambda\mathbf{A})\mathbf{B}$
- $\mathbf{IA} = \mathbf{A}$
- $\mathbf{AI} = \mathbf{A}$

Pese a las dos últimas propiedades, $\boxed{\text{en general } \mathbf{AB} \neq \mathbf{BA}.}$

Por tanto...

☞ El producto de matrices es un **operador lineal**; pues es *distributivo* respecto a la suma de matrices

$$\mathbf{A}\big(\mathbf{B}+\mathbf{C}\big) = \mathbf{AB}+\mathbf{AC}.$$

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

$$\mathbf{A}\big(\lambda\mathbf{B}\big) = \lambda\big(\mathbf{AB}\big).$$

Y también es lineal por la izquierda, pues

$$\big(\mathbf{A}+\mathbf{B}\big)\mathbf{C} = \mathbf{AC}+\mathbf{BC} \qquad \text{y} \qquad (\lambda\mathbf{A})\mathbf{B} = \lambda(\mathbf{AB}).$$

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 de matrices; y así completar la visión del producto de matrices (la primera forma de calcular el producto seguramente le resultará familiar...).

Cálculo del producto de matrices componente a componente (filas por columnas)

EJERCICIO 21. 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:*

$$\boxed{{}_{i|}\big(\mathbf{AB}\big)_{|j} = \big({}_{i|}\mathbf{A}\big)\cdot\big(\mathbf{B}_{|j}\big)}$$

Verifíquelo con el producto del Ejemplo 2 en la página 29 (hágalo a mano y con la librería NAcAL).

Pero hay más formas de calcular el producto...

Cálculo del producto de matrices operando con las filas

EJERCICIO 22. Demuestre la siguiente proposición:

Proposición 3.1.9. *Sean* $\underset{m\times p}{\mathbf{A}}$ *y* $\underset{p\times n}{\mathbf{B}}$, *entonces:* $\boxed{{}_{i|}\big(\mathbf{AB}\big) = \big({}_{i|}\mathbf{A}\big)\mathbf{B}.}$

Pista. Recuerde que $\boldsymbol{a}\cdot\left(\mathbf{B}_{|j}\right)=\left(\boldsymbol{a}\mathbf{B}\right)_{|j}$, para $j=1:n$.

Verifíquelo con el producto del Ejemplo 2 en la página 29 (hágalo a mano y con la librería 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*.

$$\boxed{{}_{i|}\left(\mathbf{AB}\right)_{|j}=\left({}_{i|}\mathbf{A}\right)\cdot\left(\mathbf{B}_{|j}\right)}.$$

La segunda alternativa calcula el producto fila a fila:

$$\boxed{{}_{i|}\left(\mathbf{AB}\right)=\left({}_{i|}\mathbf{A}\right)\mathbf{B}},$$

de manera que las filas de $\mathbf{AB}$ son combinaciones lineales de las filas de $\mathbf{B}$.

3.1.2. Nuevas reglas de reescritura.

Realmente no es necesario escribir el "punto" en los productos

$$\left({}_{i|}\mathbf{A}\right)\cdot\boldsymbol{b}\,,\qquad\qquad \boldsymbol{a}\cdot\left(\mathbf{B}_{|j}\right),\qquad\qquad \left({}_{i|}\mathbf{A}\right)\cdot\left(\mathbf{B}_{|j}\right)$$

(como tampoco lo es en el producto $2\cdot\boldsymbol{x}$). Si omitimos el "punto", recuperamos la asociatividad

$${}_{i|}\left(\mathbf{A}\boldsymbol{b}\right)=\left({}_{i|}\mathbf{A}\right)\boldsymbol{b}$$

$$\left(\boldsymbol{a}\mathbf{B}\right)_{|j}=\boldsymbol{a}\left(\mathbf{B}_{|j}\right)$$

$${}_{i|}\left(\mathbf{AB}\right)_{|j}={}_{i|}\Big(\left(\mathbf{AB}\right)_{|j}\Big)={}_{i|}\Big(\mathbf{A}\left(\mathbf{B}_{|j}\right)\Big)=\left({}_{i|}\mathbf{A}\right)\left(\mathbf{B}_{|j}\right).$$

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 arrojan 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 23. Recordando que $\left(\mathbf{A}^{\intercal}\right)_{|j}={}_{j|}\mathbf{A}$ y que $\boldsymbol{a}\cdot\boldsymbol{b}=\boldsymbol{b}\cdot\boldsymbol{a}$; demuestre que

$$\boxed{(\mathbf{AB})^{\intercal}=\left(\mathbf{B}^{\intercal}\right)\left(\mathbf{A}^{\intercal}\right)}.$$

Por tanto, para toda matriz $\mathbf{A}$ se verifica que las matrices $\mathbf{A}^{\intercal}\mathbf{A}$ y $\mathbf{A}\left(\mathbf{A}^{\intercal}\right)$ son simétricas:

$$\left(\mathbf{A}^{\intercal}\mathbf{A}\right)^{\intercal}=\mathbf{A}^{\intercal}\left(\mathbf{A}^{\intercal}\right)^{\intercal}=\mathbf{A}^{\intercal}\mathbf{A};\qquad\text{y}\qquad\left(\mathbf{A}\left(\mathbf{A}^{\intercal}\right)\right)^{\intercal}=\left(\mathbf{A}^{\intercal}\right)^{\intercal}\left(\mathbf{A}^{\intercal}\right)=\mathbf{A}\left(\mathbf{A}^{\intercal}\right).$$

☞ En próximas lecciones trabajaremos con matrices por bloques. De ello tratan los siguientes apéndices

pero por el momento puede saltarse las Secciones 3.A y 3.B

Apendices a la lección

3.A. Submatrices mediante selección de una lista de índices

Sea $\mathbf{I}$ de orden n y sea $\boldsymbol{\alpha}$ una lista de índices menores o iguales a n, por ejemplo, $n \geq 4$ y $\boldsymbol{\alpha} = (2,3,4)$. Definimos $\mathbf{I}_{|\boldsymbol{\alpha}}$ como la matriz (de n filas y tantas columnas como el *número de elementos de la lista*) cuyas columnas son las columnas de $\mathbf{I}$ (de orden n) indicadas en la lista. 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: ${}_{\boldsymbol{\alpha}|}\mathbf{I} = \left(\mathbf{I}_{|\boldsymbol{\alpha}}\right)^{\intercal}$.

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:

$$\mathbf{A}_{|\boldsymbol{\alpha}} = \mathbf{A}\left(\mathbf{I}_{|\boldsymbol{\alpha}}\right), \qquad \text{y} \qquad {}_{\boldsymbol{\alpha}|}\mathbf{A} = \left({}_{\boldsymbol{\alpha}|}\mathbf{I}\right)\mathbf{A}.$$

Consideremos de nuevo la matriz $\mathbf{PIB}$ de la Página 10. La submatriz con las filas de 2 a 4 de $\mathbf{PIB}$ (datos de Francia, Alemania y Reino Unido) o la submatriz con sus dos primeras columnas (años 2003 y 2004) son

$${}_{(2,3,4)|}\mathbf{PIB} = \begin{bmatrix} 0.9 & 2.0 & 1.6 & 2.0 \\ -0.2 & 1.1 & 0.7 & 0.9 \\ 2.5 & 3.2 & 1.9 & 1.5 \end{bmatrix}; \quad \text{y} \quad \mathbf{PIB}_{|(1,2)} = \begin{bmatrix} 2.9 & 3.1 \\ 0.9 & 2.0 \\ -0.2 & 1.1 \\ 2.5 & 3.2 \\ 1.8 & 3.1 \end{bmatrix}.$$

Y podemos crear submatrices seleccionando simultáneamente filas y columnas; si $\boldsymbol{\beta}$ es otra lista de índices:

$${}_{\boldsymbol{\alpha}|}\left(\mathbf{A}_{|\boldsymbol{\beta}}\right) = \left({}_{\boldsymbol{\alpha}|}\mathbf{I}\right)\mathbf{A}\left(\mathbf{I}_{|\boldsymbol{\beta}}\right) = \left({}_{\boldsymbol{\alpha}|}\mathbf{A}\right)_{|\boldsymbol{\beta}},$$

que, aprovechando la asociatividad, podemos escribir sencillamente como: ${}_{\boldsymbol{\alpha}|}\mathbf{A}_{|\boldsymbol{\beta}}$.

Así, la submatriz con los datos de Francia, Alemania y Reino Unido de los años 2003 y 2004 es

$${}_{(2,3,4)|}\mathbf{PIB}_{|(1,2)} = \begin{bmatrix} 0.9 & 2.0 \\ -0.2 & 1.1 \\ 2.5 & 3.2 \end{bmatrix}.$$

Nótese que

$$\left({}_{\boldsymbol{\alpha}|}\mathbf{A}\right)\left(\mathbf{B}_{|\boldsymbol{\beta}}\right) = \left({}_{\boldsymbol{\alpha}|}\mathbf{I}\right)\mathbf{A}\mathbf{B}\left(\mathbf{I}_{|\boldsymbol{\beta}}\right) = {}_{\boldsymbol{\alpha}|}\left(\mathbf{A}\mathbf{B}\right)_{|\boldsymbol{\beta}},$$

y en particular

$$\left(\mathbf{A}\mathbf{B}\right)_{|\boldsymbol{\beta}} = \mathbf{A}\left(\mathbf{B}_{|\boldsymbol{\beta}}\right) \qquad \text{y} \qquad {}_{\boldsymbol{\alpha}|}\left(\mathbf{A}\mathbf{B}\right) = \left({}_{\boldsymbol{\alpha}|}\mathbf{A}\right)\mathbf{B};$$

por lo que podemos escribir sencillamente: ${}_{\boldsymbol{\alpha}|}\mathbf{A}\mathbf{B}_{|\boldsymbol{\beta}}$. Y también se verifica que

$${}_{\boldsymbol{\alpha}|}\left(\mathbf{A}+\mathbf{B}\right)_{|\boldsymbol{\beta}} = {}_{\boldsymbol{\alpha}|}\mathbf{A}_{|\boldsymbol{\beta}} + {}_{\boldsymbol{\alpha}|}\mathbf{B}_{|\boldsymbol{\beta}} \qquad \text{y} \qquad \lambda\left({}_{\boldsymbol{\alpha}|}\mathbf{A}_{|\boldsymbol{\beta}}\right) = {}_{\boldsymbol{\alpha}|}\left(\lambda\mathbf{A}\right)_{|\boldsymbol{\beta}}. \tag{3.1}$$

3.B. Matrices particionadas

Recuerde que si $p \leq q \in \mathbb{N}$, entonces $(p:q)$ denota la secuencia $p, p+1, \ldots, q$. Ahora considere la matriz

$$\underset{5\times 6}{\mathbf{B}} = \begin{bmatrix} 1&1&1&1&2&2\\ 1&1&1&1&2&2\\ 1&1&1&1&2&2\\ 3&3&3&3&4&4\\ 3&3&3&3&4&4 \end{bmatrix},$$

junto con las siguientes particiones de índices: por una parte los índices correspondientes a las tres primeras filas, $(1:3)$, y los correspondientes a las dos últimas, $(4:5)$. Por otra, los índices de las cuatro primeras columnas, $(1:4)$, y los de las dos últimas, $(5:6)$. Con dichas particiones de índices partimos $\mathbf{B}$ en 4 bloques:

$$\boldsymbol{\mathcal{B}} = \left[\begin{array}{cccc|cc} 1&1&1&1&2&2\\ 1&1&1&1&2&2\\ 1&1&1&1&2&2\\ \hline 3&3&3&3&4&4\\ 3&3&3&3&4&4 \end{array}\right] = \left[\begin{array}{c|c} {}_{(1:3)|}\mathbf{B}_{|(1:4)} & {}_{(1:3)|}\mathbf{B}_{|(5:6)} \\ \hline {}_{(4:5)|}\mathbf{B}_{|(1:4)} & {}_{(4:5)|}\mathbf{B}_{|(5:6)} \end{array}\right];$$

donde

$${}_{(1:3)|}\mathbf{B}_{|(1:4)} = \begin{bmatrix} 1&1&1&1\\ 1&1&1&1\\ 1&1&1&1 \end{bmatrix}, \qquad {}_{(1:3)|}\mathbf{B}_{|(5:6)} = \begin{bmatrix} 2&2\\ 2&2\\ 2&2 \end{bmatrix},$$

$${}_{(4:5)|}\mathbf{B}_{|(1:4)} = \begin{bmatrix} 3&3&3&3\\ 3&3&3&3 \end{bmatrix}, \qquad {}_{(4:5)|}\mathbf{B}_{|(5:6)} = \begin{bmatrix} 4&4\\ 4&4 \end{bmatrix}.$$

Así, al cruzar una matriz con lineas que van de lado a lado o de arriba a abajo, delimitamos los bloques o particiones (i.e., submatrices dentro de una matriz). Denotaremos las *matrices particionadas* con: $\boldsymbol{\mathcal{A}}$, $\boldsymbol{\mathcal{B}}$, $\boldsymbol{\mathcal{C}}$, etc. pues al ser objetos distintos necesitamos una grafía diferente para distinguirlas de las matrices.[3]

Nótese que dos bloques adyacentes, uno a la derecha del otro, comparten la misma lista de filas; y que dos bloques adyacentes, uno encima del otro, comparten la misma lista de columnas.

Una notación más compacta para particionar una matriz. Con un conjunto de índices vamos a indicar por debajo de qué filas pintamos una línea horizontal (si particionamos por filas), y a la derecha de qué columnas pintamos una línea vertical (si particionamos por columnas). De este modo, la matriz particionada del ejemplo de más arriba se puede expresar del siguiente modo:

$$\boldsymbol{\mathcal{B}} = {}_{\{3\}|}\mathbf{B}_{|\{4\}} = \left[\begin{array}{cccc|cc} 1&1&1&1&2&2\\ 1&1&1&1&2&2\\ 1&1&1&1&2&2\\ \hline 3&3&3&3&4&4\\ 3&3&3&3&4&4 \end{array}\right]$$

Librería NAcAL para Python

```
a = Vector( [1,1,1,3,3] )
b = Vector( [2,2,2,4,4] )
B = Matrix( [a,a,a,a,b,b] )
{3}|B|{4}
```

[3] de hecho, para distinguir una matriz particionada en un único bloque $\boldsymbol{\mathcal{A}}$, de una matriz $\mathbf{A}$, la librería NAcAL dibuja un recuadro alrededor para indicar que la matriz está particionada: $\boldsymbol{\mathcal{A}} = \boxed{\begin{bmatrix} 4&4\\ 4&4 \end{bmatrix}}$.

Una misma matriz admite distintas particiones; y con esta notación es sencillo describirlas. Por ejemplo:

$$\boldsymbol{\mathcal{B}}' = {}_{\{1\}|}\mathbf{B}_{|\{1,4\}} = \left[\begin{array}{c|ccc|cc} 1 & 1 & 1 & 1 & 2 & 2 \\ \hline 1 & 1 & 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 & 2 & 2 \\ 3 & 3 & 3 & 3 & 4 & 4 \\ 3 & 3 & 3 & 3 & 4 & 4 \end{array}\right] = \left[\begin{array}{c|c|c} {}_{(1,)|}\mathbf{B}_{|(1,)} & {}_{(1,)|}\mathbf{B}_{|(2:4)} & {}_{(1,)|}\mathbf{B}_{|(5:6)} \\ \hline {}_{(2:5)|}\mathbf{B}_{|(1,)} & {}_{(2:5)|}\mathbf{B}_{|(2:4)} & {}_{(2:5)|}\mathbf{B}_{|(5:6)} \end{array}\right],$$

donde ahora

$${}_{(1,)|}\mathbf{B}_{|(1,)} = \begin{bmatrix}1\end{bmatrix}, \qquad {}_{(1,)|}\mathbf{B}_{|(2:4)} = \begin{bmatrix}1 & 1 & 1\end{bmatrix}, \qquad {}_{(1,)|}\mathbf{B}_{|(5:6)} = \begin{bmatrix}2 & 2\end{bmatrix},$$

$${}_{(2:5)|}\mathbf{B}_{|(1,)} = \begin{bmatrix}1\\1\\3\\3\end{bmatrix}, \qquad {}_{(2:5)|}\mathbf{B}_{|(2:4)} = \begin{bmatrix}1 & 1 & 1\\1 & 1 & 1\\3 & 3 & 3\\3 & 3 & 3\end{bmatrix}, \qquad {}_{(2:5)|}\mathbf{B}_{|(5:6)} = \begin{bmatrix}2 & 2\\2 & 2\\4 & 4\\4 & 4\end{bmatrix},$$

☞ A continuación vamos a ver la suma ($\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{B}}$) y el producto ($\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{B}}$) de matrices particionadas.

3.B.1. Suma de matrices particionadas.

Considere dos matrices del mismo orden, $\mathbf{A}$ y $\mathbf{B}$, y las matrices particionadas $\boldsymbol{\mathcal{A}}$ y $\boldsymbol{\mathcal{B}}$, obtenidas usando idénticas particiones para las filas y columnas de $\mathbf{A}$ y $\mathbf{B}$. Por ejemplo

$$\boldsymbol{\mathcal{A}} = {}_{\{h\}|}\mathbf{A}_{|\{k\}} = \left[\begin{array}{c|c} \mathbf{A_{11}} & \mathbf{A_{12}} \\ \hline \mathbf{A_{21}} & \mathbf{A_{22}} \end{array}\right]; \qquad \boldsymbol{\mathcal{B}} = {}_{\{h\}|}\mathbf{B}_{|\{k\}} = \left[\begin{array}{c|c} \mathbf{B_{11}} & \mathbf{B_{12}} \\ \hline \mathbf{B_{21}} & \mathbf{B_{22}} \end{array}\right],$$

entonces la suma se puede calcular sumando bloque a bloque (como en la parte izquierda de la Ecuación 3.1), donde $\mathbf{A}_{ij}$ y $\mathbf{B}_{ij}$ son las submatrices que constituyen cada uno de los bloques. Por ejemplo

$$\left[\begin{array}{cc|c} 1 & 2 & 3 \\ \hline 3 & 2 & 1 \end{array}\right] + \left[\begin{array}{cc|c} -1 & 1 & 0 \\ \hline -1 & 1 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 0 & 3 & 3 \\ \hline 2 & 3 & 1 \end{array}\right].$$

Es decir, $\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{B}} = {}_{\{h\}|}\mathbf{A}_{|\{k\}} + {}_{\{h\}|}\mathbf{B}_{|\{k\}} = {}_{\{h\}|}\big(\mathbf{A} + \mathbf{B}\big)_{|\{k\}}$.

3.B.2. Producto de matrices particionadas

☞ En esta sección describiré el producto por bloques, aunque la justificación de este modo de operar se hará adelante (en este mismo apéndice).

Sean $\boldsymbol{\mathcal{A}}$ y $\boldsymbol{\mathcal{B}}$ tales que la partición $\{h_1, \ldots, h_{s-1}\}$ de las columnas de $\boldsymbol{\mathcal{A}}$ y las filas $\boldsymbol{\mathcal{B}}$ es común:

$$\boldsymbol{\mathcal{A}} = {}_{\{f_1,\ldots,f_{m-1}\}|}\mathbf{A}_{|\{h_1,\ldots,h_{s-1}\}} = \left[\begin{array}{c|c|c|c} \mathbf{A_{11}} & \mathbf{A_{12}} & \cdots & \mathbf{A_{1s}} \\ \hline \mathbf{A_{21}} & \mathbf{A_{22}} & \cdots & \mathbf{A_{2s}} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{A_{m1}} & \mathbf{A_{m2}} & \cdots & \mathbf{A_{ms}} \end{array}\right]$$

$$\boldsymbol{\mathcal{B}} = {}_{\{h_1,\ldots,h_{s-1}\}|}\mathbf{B}_{|\{c_1,\ldots,c_{n-1}\}} = \left[\begin{array}{c|c|c|c} \mathbf{B_{11}} & \mathbf{B_{12}} & \cdots & \mathbf{B_{1n}} \\ \hline \mathbf{B_{21}} & \mathbf{B_{22}} & \cdots & \mathbf{B_{2n}} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{B_{s1}} & \mathbf{B_{s2}} & \cdots & \mathbf{B_{sn}} \end{array}\right],$$

entonces el bloque $(\mathbf{AB})_{\boldsymbol{ij}}$ se puede calcular mediante el producto de la fila $\boldsymbol{i}$ésima de bloques de $\boldsymbol{\mathcal{A}}$ por la columna $\boldsymbol{j}$ésima de bloques de $\boldsymbol{\mathcal{B}}$ (de manera análoga al producto de las matrices). Por ejemplo:

$$\boldsymbol{\mathcal{AB}} = \left({}_{\{1\}|}\mathbf{A}_{|\{2\}}\right)\left({}_{\{2\}|}\mathbf{B}_{|\{2\}}\right) = \left[\begin{array}{cc|cc} 2 & -1 & -1 & 1 \\ \hline 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{cc|c} 1 & 2 & 0 \\ -1 & 0 & 0 \\ \hline 0 & 1 & 1 \\ 1 & -1 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 4 & 2 & -1 \\ \hline 3 & 1 & 1 \\ -1 & 1 & 1 \\ 2 & 1 & 0 \end{array}\right] = {}_{\{1\}|}(\mathbf{AB})_{|\{2\}}.$$

donde el primer bloque se calcula del siguiente modo

$${}_{1|}(\boldsymbol{\mathcal{AB}})_{|1} = \left[\begin{array}{cc|cc} 2 & -1 & -1 & 1 \end{array}\right] \left[\begin{array}{cc} 1 & 2 \\ -1 & 0 \\ \hline 0 & 1 \\ 1 & -1 \end{array}\right] = \begin{bmatrix} 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 2 \end{bmatrix},$$

y donde he empleado la notación con selectores para seleccionar las filas y columnas de bloques en matrices particionadas. En general,

$${}_{i|}(\boldsymbol{\mathcal{AB}})_{|j} = \left[\begin{array}{c|c|c|c} \mathbf{A}_{\boldsymbol{i1}} & \mathbf{A}_{\boldsymbol{i2}} & \cdots & \mathbf{A}_{\boldsymbol{is}} \end{array}\right] \left[\begin{array}{c} \mathbf{B}_{\boldsymbol{1j}} \\ \hline \mathbf{B}_{\boldsymbol{2j}} \\ \hline \vdots \\ \hline \mathbf{B}_{\boldsymbol{sj}} \end{array}\right] = \sum_{\boldsymbol{k=1}}^{\boldsymbol{s}} \mathbf{A}_{\boldsymbol{ik}}\mathbf{B}_{\boldsymbol{kj}}. \tag{3.2}$$

De (3.2) se deduce que el producto de matrices particionadas también se puede calcular por columnas de bloques, de manera que la $\boldsymbol{j}$ésima columna de bloques es:

$$(\boldsymbol{\mathcal{AB}})_{|j} = \left[\begin{array}{c} \mathbf{A}_{\boldsymbol{11}} \\ \hline \mathbf{A}_{\boldsymbol{21}} \\ \hline \vdots \\ \hline \mathbf{A}_{\boldsymbol{q1}} \end{array}\right] \mathbf{B}_{\boldsymbol{1j}} + \left[\begin{array}{c} \mathbf{A}_{\boldsymbol{12}} \\ \hline \mathbf{A}_{\boldsymbol{22}} \\ \hline \vdots \\ \hline \mathbf{A}_{\boldsymbol{q2}} \end{array}\right] \mathbf{B}_{\boldsymbol{2j}} + \cdots + \left[\begin{array}{c} \mathbf{A}_{\boldsymbol{1s}} \\ \hline \mathbf{A}_{\boldsymbol{2s}} \\ \hline \vdots \\ \hline \mathbf{A}_{\boldsymbol{qs}} \end{array}\right] \mathbf{B}_{\boldsymbol{sj}} = \sum_{k=1}^{s} \left(\boldsymbol{\mathcal{A}}_{|k}\right)\left({}_{k|}\boldsymbol{\mathcal{B}}_{|j}\right). \tag{3.3}$$

Si piensa en ello se dará cuenta de que también es posible calcular el producto por filas de bloques del siguiente modo: ${}_{i|}(\boldsymbol{\mathcal{AB}}) = \sum_{k=1}^{s}\left({}_{i|}\boldsymbol{\mathcal{A}}_{|k}\right)\left({}_{k|}\boldsymbol{\mathcal{B}}\right)$.

3.C. Producto matricial como suma de productos de submatrices

La justificación de por qué (3.2) funciona se asienta en una propiedad curiosa que, además, nos permitirá generalizar el producto de matrices particionadas y trabajar con otros tipos de submatrices. Así pues, comencemos viendo dicha propiedad.

Considere la lista $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_q)$ de q índices no repetidos, y la submatriz ${}_{\boldsymbol{\alpha}|}\mathbf{I}$, de la matriz identidad $\mathbf{I}$ de orden mayor o igual al mayor de los índices en $\boldsymbol{\alpha}$. Pues si α_j es la j*ésima* componente de la lista $\boldsymbol{\alpha}$, entonces la columna $\left({}_{\boldsymbol{\alpha}|}\mathbf{I}\right)_{|\alpha_j}$ es igual $\mathbf{I}_{|j} \in \mathbb{R}^q$ (donde q es la longitud de la lista $\boldsymbol{\alpha}$):

$${}_{\boldsymbol{\alpha}|}\mathbf{I}_{|k} = {}_{(\alpha_1,\ldots,\alpha_q)|}\mathbf{I}_{|k} = {}_{(\alpha_1,\ldots,\alpha_q)|}\left(\mathbf{I}_{|k}\right) = \begin{cases} \mathbf{I}_{|j} \in \mathbb{R}^q & \text{si } k = \alpha_j \\ & \\ \mathbf{0} \in \mathbb{R}^q & \text{si } k \notin \boldsymbol{\alpha} \end{cases} \in \mathbb{R}^q, \tag{3.4}$$

pues 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: $\mathbf{0} \in \mathbb{R}^q$.

Por ejemplo, consideremos la lista de 3 índices $\boldsymbol{\alpha} = (4,1,2)$ (cuyo mayor índice es el 4) y una matriz identidad de orden n mayor o igual que 4, por ejemplo $n = 4$. Entonces

$$_{\boldsymbol{\alpha}|}\mathbf{I} = {}_{(4,1,2)|}\mathbf{I} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Y ahora fijémonos en las columnas de la anterior matriz:

- 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

$${}_{(4,1,2)|}\mathbf{I}_{|3} = {}_{(4,1,2)|}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Librería NAcAL para Python

```
(4,1,2)|I(5)|4    # Primera columna de la Matriz identidad de orden 3
(4,1,2)|I(5)|2    # Segunda columna de la Matriz identidad de orden 3
(4,1,2)|I(5)|1    # Tercera columna de la Matriz identidad de orden 3
(4,1,2)|I(5)|3    # Vector nulo de tres componentes (R3)
```

Considere las columnas del producto de matrices $\left(\mathbf{I}_{|\boldsymbol{\alpha}}\right)\left({}_{\boldsymbol{\alpha}|}\mathbf{I}\right)$. Usando (3.4) se deduce que

$$\Big(\left(\mathbf{I}_{|\boldsymbol{\alpha}}\right)\left({}_{\boldsymbol{\alpha}|}\mathbf{I}\right)\Big)_{|k} = \left(\mathbf{I}_{|\boldsymbol{\alpha}}\right)\left({}_{\boldsymbol{\alpha}|}\mathbf{I}_{|k}\right) = \begin{cases} \left(\mathbf{I}_{|\boldsymbol{\alpha}}\right)\mathbf{I}_{|j} = \left(\mathbf{I}_{|\boldsymbol{\alpha}}\right)_{|j} = \mathbf{I}_{|k} & \text{si } k = \alpha_j \\ \left(\mathbf{I}_{|\boldsymbol{\alpha}}\right)\mathbf{0} \;= \mathbf{0} & \text{si } k \notin \boldsymbol{\alpha} \end{cases}; \qquad (3.5)$$

Así, para $\boldsymbol{\alpha} = (4,1,2)$ y $\underset{4\times 4}{\mathbf{I}}$

$$\big(\mathbf{I}_{|(4,1,2)}\big)\big({}_{(4,1,2)|}\mathbf{I}\big) = \begin{bmatrix} 0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 \end{bmatrix},$$

de manera que solo hay unos en las posiciones 1, 2 y 4 de la diagonal.

Librería NAcAL para Python

```
(I(4)|(4,1,2)) * ((4,1,2)|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 $\boldsymbol{\gamma}_3 = (2,4)$; entonces

$$\mathbf{A}_{|\boldsymbol{\gamma}_1 \# \boldsymbol{\gamma}_2 \# \boldsymbol{\gamma}_3} \;=\; \mathbf{A}_{|(3,1,5,4,2)};$$

donde $\boldsymbol{\gamma}_1 \# \boldsymbol{\gamma}_2 \# \boldsymbol{\gamma}_3 = (3,1,5,4,2)$ es la concatenación de las listas $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$ y $\boldsymbol{\gamma}_3$. Así pues, $\mathbf{A}_{|\boldsymbol{\gamma}_1 \# \boldsymbol{\gamma}_2 \# \boldsymbol{\gamma}_3}$ es un reordenamiento (o permutación) de las columnas, pues todas ellas aparecen una, y solo una, vez.

Si empleamos las listas $\boldsymbol{\gamma}_1,\ldots,\boldsymbol{\gamma}_h$ para reordenar tanto las columnas de $\mathbf{A}$ y como las filas de $\mathbf{B}$, el producto de las matrices resultantes resulta ser igual al producto $\mathbf{AB}$:

$$\big(\mathbf{A}_{|\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h}\big)\big({}_{\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h|}\mathbf{B}\big) = \mathbf{A}\big(\mathbf{I}_{|\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h}\big)\big({}_{\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h|}\mathbf{I}\big)\mathbf{B} = \mathbf{AIB} = \mathbf{AB},$$

pues como todos los índices están incluidos, por (3.5), sabemos que $\big(\mathbf{I}_{|\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h}\big)\big({}_{\boldsymbol{\gamma}_1\#\cdots\#\boldsymbol{\gamma}_h|}\mathbf{I}\big) = \mathbf{I}$.

Consecuentemente podemos calcular $\mathbf{AB}$ mediante la suma de k productos de submatrices:

$$\sum_{k=1}^{h}\big(\mathbf{A}_{|\boldsymbol{\gamma}_k}\big)\big({}_{\boldsymbol{\gamma}_k|}\mathbf{B}\big) = \sum_{k=1}^{h}\Big(\mathbf{A}\big(\mathbf{I}_{|\boldsymbol{\gamma}_k}\big)\big({}_{\boldsymbol{\gamma}_k|}\mathbf{I}\big)\mathbf{B}\Big) = \mathbf{A}\left(\sum_{k=1}^{h}\big(\mathbf{I}_{|\boldsymbol{\gamma}_k}\big)\big({}_{\boldsymbol{\gamma}_k|}\mathbf{I}\big)\right)\mathbf{B} = \mathbf{AIB} = \mathbf{AB},$$

donde, por (3.5), cada sumando de $\sum_{k=1}^{h}\big(\mathbf{I}_{|\boldsymbol{\gamma}_k}\big)\big({}_{\boldsymbol{\gamma}_k|}\mathbf{I}\big)$ 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.

Por ejemplo, si usamos las listas $\boldsymbol{\gamma}_1 = (3,1)$ y $\boldsymbol{\gamma}_2 = (4,2)$ con $\underset{4\times 4}{\mathbf{I}}$, entonces

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

Librería NAcAL para Python

```
g1=(3,1,); g2=(4,2,)
(I(4)|g1)*(g1|I(4)) + (I(4)|g2)*(g2|I(4))
```

Así, para $\mathbf{A} = \begin{bmatrix} 1 & 2 & -1\\ 2 & -2 & 0 \end{bmatrix}$ y $\mathbf{B} = \begin{bmatrix} 2 & 1\\ 1 & 1\\ -1 & 1 \end{bmatrix}$, y las listas $\boldsymbol{\gamma}_1 = (3,1,)$ y $\boldsymbol{\gamma}_2 = (2,)$; tenemos que

$$\begin{aligned}\big(\mathbf{A}_{|\boldsymbol{\gamma}_1\#\boldsymbol{\gamma}_2}\big)\big({}_{\boldsymbol{\gamma}_1\#\boldsymbol{\gamma}_2|}\mathbf{B}\big) &= \big(\mathbf{A}_{|(3,1)}\big)\big({}_{(3,1)|}\mathbf{B}\big) + \big(\mathbf{A}_{|(2,)}\big)\big({}_{(2,)|}\mathbf{B}\big)\\ &= \begin{bmatrix} -1 & 1\\ 0 & 2 \end{bmatrix}\begin{bmatrix} -1 & 1\\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 2\\ -2 \end{bmatrix}\begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0\\ 4 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 2\\ -2 & -2 \end{bmatrix} = \begin{bmatrix} 5 & 2\\ 2 & 0 \end{bmatrix} = \mathbf{AB}.\end{aligned}$$

Librería 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 *numero* 2 entre paréntesis de la *lista* que contiene al 2. Por claridad, y puesto que la distinción es importante, aquí sigo el mismo criterio de notación: 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 como suma de matrices

Como caso particular, consideremos 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 matrices:

$$\Big(\underset{m\times p}{\mathbf{A}}\Big)\Big(\underset{p\times n}{\mathbf{B}}\Big) = \sum_{k=1}^{p} \big(\mathbf{A}_{|(k,)}\big)\big({}_{(k,)|}\mathbf{B}\big),$$

donde cada sumando $\underset{m\times 1}{\big(\mathbf{A}_{|(k,)}\big)}\ \underset{1\times n}{\big({}_{(k,)|}\mathbf{B}\big)}$ es una matriz de orden m por n. Ejemplo:

$$\begin{aligned}\mathbf{AB} &= \begin{bmatrix}1&2&3\\1&1&1\\0&-1&2\end{bmatrix}\begin{bmatrix}-1&2\\0&0\\1&0\end{bmatrix} = \big(\mathbf{A}_{|(1,)}\big)\big({}_{(1,)|}\mathbf{B}\big) + \big(\mathbf{A}_{|(2,)}\big)\big({}_{(2,)|}\mathbf{B}\big) + \big(\mathbf{A}_{|(3,)}\big)\big({}_{(3,)|}\mathbf{B}\big)\\ &= \begin{bmatrix}1\\1\\0\end{bmatrix}\begin{bmatrix}-1&2\end{bmatrix} + \begin{bmatrix}2\\1\\-1\end{bmatrix}\begin{bmatrix}0&0\end{bmatrix} + \begin{bmatrix}3\\1\\2\end{bmatrix}\begin{bmatrix}1&0\end{bmatrix} = \begin{bmatrix}-1&2\\-1&2\\0&0\end{bmatrix} + \begin{bmatrix}0&0\\0&0\\0&0\end{bmatrix} + \begin{bmatrix}3&0\\1&0\\2&0\end{bmatrix} = \begin{bmatrix}2&2\\0&2\\2&0\end{bmatrix}.\end{aligned}$$

Librería NAcAL para Python

```
(A|(1,)) * ((1,)|B) + (A|(2,)) * ((2,)|B) + (A|(3,)) * ((3,)|B)
```

3.C.2. Submatriz de un producto como suma de productos de submatrices

Y ahora vayamos con la generalización objetivo de este apéndice. Considere una lista de listas $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_h$, donde cada índice $j = 1:p$ está incluido en una y solo una de las listas $\boldsymbol{\gamma}_k$ entonces:

$$\mathbf{AB} \;=\; \sum_{k=1}^{h} \big(\mathbf{A}_{|\boldsymbol{\gamma}_k}\big)\big({}_{\boldsymbol{\gamma}_k|}\mathbf{B}\big), \tag{3.6}$$

Seleccionando algunas filas de $\underset{m\times p}{\mathbf{A}}$ y algunas columnas de $\underset{p\times n}{\mathbf{B}}$, la submatriz ${}_{\boldsymbol{\alpha}|}\big(\mathbf{AB}\big)_{|\boldsymbol{\beta}}$ resulta ser

$${}_{\boldsymbol{\alpha}|}\big(\mathbf{AB}\big)_{|\boldsymbol{\beta}} \;=\; \sum_{k=1}^{h} \big({}_{\boldsymbol{\alpha}|}\mathbf{A}_{|\boldsymbol{\gamma}_k}\big)\big({}_{\boldsymbol{\gamma}_k|}\mathbf{B}_{|\boldsymbol{\beta}}\big),$$

que de alguna manera es una generalización de (3.2), pues aquí no se requiere que las submatrices estén formadas por vectores de índices consecutivos.

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 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 puede 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) una matriz.* Cada transformación en la secuencia (cada paso enla secuencia) hace una de estas dos posibles operaciones:

- *sumar* un múltiplo de un vector a *otro* vector (Tipo I). *Ejemplo:* $\big[\boldsymbol{a};\ \boldsymbol{b}\big] \to \big[\boldsymbol{a};\ \boldsymbol{b}+\lambda\boldsymbol{a}\big]$.
- *multiplicar* un vector por un número *distinto de cero* (Tipo II). *Ejemplo:* $\big[\boldsymbol{a};\ \boldsymbol{b}\big] \to \big[\boldsymbol{a};\ \lambda\boldsymbol{b}\big]$.

A estas dos operaciones las denominaremos *transformaciones elementales.* [1]

4.1. Transformaciones y matrices elementales

Denotamos con "$\boldsymbol{\tau}$" al operador correspondiente a la transformación aplicada, y lo situamos como subíndice a la derecha de la matriz para indicar que opera sobre las columnas: $\mathbf{A}_{\boldsymbol{\tau}}$. Entre corchetes, por debajo de cada "$\boldsymbol{\tau}$", indicaremos los detalles de cada transformación.

La matriz resultante tras aplicar una sola transformación elemental sobre $\mathbf{I}$ tiene un nombre especial:

Definición 4.1. *Una* 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 se puede aplicar una transformación elemental $\boldsymbol{\tau}$ mediante un producto de matrices. Vayamos con los detalles.

4.1.1. Transformaciones y matrices elementales de Tipo I

Transformación elemental de Tipo I: $\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}$ suma λ 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*).

Librería NAcAL para Python

```
T( (8,2,3) )  # Suma ocho veces el segundo vector al tercero
```

$$\underset{[(8)\mathbf{2}+\mathbf{3}]}{\tau}$$

[1] En muchos manuales consideran el intercambio como una tercera transformación elemental, pero aquí no lo haremos.

Matriz elemental de Tipo I. Cualquier matriz obtenida *tras aplicar una sola transformación elemental* de *Tipo I* sobre una matriz identidad:

$$\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}$$

Por ejemplo, si sobre las columnas de la matriz identidad de orden 3 aplicamos la transformación $\underset{[(\lambda)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}$ (*que suma "λ veces" la segunda columna a la tercera*) obtenemos:

$$\mathbf{I}_{\underset{[(\lambda)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}} = \begin{bmatrix} 1 & & \\ & 1 & \lambda \\ & & 1 \end{bmatrix};$$

que es como la matriz identidad de orden 3 salvo porque la segunda componente de su tercera columna ha cambiado y ahora es λ (en lugar de cero).

Fíjese que la notación $[(\lambda)\,\boldsymbol{i}+\boldsymbol{j}]$ indica que sumamos λ 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 $\mathbf{I}$ que ha cambiado y entre paréntesis su nuevo valor (λ).

Librería 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)
```

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 8 \\ 0 & 0 & 1 \end{bmatrix}$$

En general, si aplicamos $\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}$ (*que suma "λ veces" la $i$ésima columna a la $j$ésima* (con $j \neq i$)) sobre las columnas de $\mathbf{I}$ (de orden n), cambiamos la $j$ésima columna de manera las columnas de la nueva matriz son:

$$\left(\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)_{|k} = \begin{cases} \text{Cuando } k = j: & \left(\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)_{|j} = \lambda\mathbf{I}_{|i} + \mathbf{I}_{|j} \text{ (nueva columna } j\text{ésima)} \\ \\ \text{Cuando } k \neq j: & \left(\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)_{|k} = \mathbf{I}_{|k} \text{ (pero el resto de columnas no cambian)} \end{cases} \tag{4.1}$$

donde $k = 1:n$. Veamos qué pasa si aplicamos una transformación elemental Tipo I sobre una matriz $\mathbf{A}$.

Transformación de las columnas de una matriz $\mathbf{A}$: En general, si aplicamos $\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}$ a las columnas de $\underset{m\times n}{\mathbf{A}}$ obtenemos una matriz cuyas columnas son:

$$\left(\mathbf{A}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)_{|k} = \begin{cases} \lambda\mathbf{A}_{|i} + \mathbf{A}_{|j} = \mathbf{A}\left(\lambda\mathbf{I}_{|i} + \mathbf{I}_{|j}\right) = \mathbf{A}\left(\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)_{|j} & \text{para } k = j \\ \\ \mathbf{A}_{|k} = \mathbf{A}\left(\mathbf{I}_{|k}\right) = \mathbf{A}\left(\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)_{|k} & \text{para } k \neq j \end{cases} = \mathbf{A}\left(\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)_{|k},$$

donde $k = 1:n$; y donde hemos usado (4.1).

Por tanto, aplicar la transformación $\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}$ sobre las columnas de $\mathbf{A}$ de orden $m \times n$ es equivalente a multiplicar $\mathbf{A}$ por la correspondiente matriz elemental $\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}$ de orden n por la derecha,

$$\mathbf{A}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}} = \mathbf{A}\left(\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right).$$

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:

$$\mathbf{A}_{\underset{[(\lambda)\mathbf{2}+\mathbf{3}]}{\tau}} = \begin{bmatrix} 0 & 0 & (\lambda 0 + c) \\ 0 & b & (\lambda b + c) \\ a & b & (\lambda b + c) \end{bmatrix} = \begin{bmatrix} 0 & 0 & c \\ 0 & b & c \\ a & b & c \end{bmatrix} \begin{bmatrix} 1 & & \\ & 1 & \lambda \\ & & 1 \end{bmatrix} = \mathbf{A}\Big(\mathbf{I}_{\underset{[(\lambda)\mathbf{2}+\mathbf{3}]}{\tau}}\Big).$$

4.1.2. Transformaciones y matrices elementales de Tipo II

Transformación elemental de Tipo II: $\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}$ multiplica por α el $\boldsymbol{i}$ésimo vector (con $\alpha \neq 0$).

(*solo modifica el vector cuyo índice aparece en última posición en el corchete*).

Librería NAcAL para Python

```
T( (5, 1) )    # multiplica por 5 el primer vector
```

$$\underset{[(5)\mathbf{1}]}{\tau}$$

Matriz elemental de Tipo II Cualquier matriz obtenida *tras aplicar una sola transformación elemental* de *Tipo II* sobre 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 $\underset{[(\alpha)\mathbf{2}]}{\boldsymbol{\tau}}$ (*que multiplica la segunda columna por* $\alpha \neq 0$) obtenemos:

$$\mathbf{I}_{\underset{[(\alpha)\mathbf{2}]}{\boldsymbol{\tau}}} = \begin{bmatrix} 1 & & \\ & \alpha & \\ & & 1 \end{bmatrix};$$

que es como la matriz $\underset{3\times3}{\mathbf{I}}$ salvo porque la segunda componente de la diagonal es α.

Librería 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)
```

$$\begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

En general, si aplicamos $\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}$ (*que multiplica la columna iésima por* $\alpha \neq 0$) sobre las columnas de $\mathbf{I}$ (de orden n), modificamos su $i$ésima columna, de manera las columnas de la nueva matriz son:

$$\Big(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\Big)_{|k} = \begin{cases} \text{Cuando } k = i: & \Big(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\Big)_{|i} = \alpha\mathbf{I}_{|i} \text{ (nueva columna } i\text{ésima)} \\ \\ \text{Cuando } k \neq i: & \Big(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\Big)_{|k} = \mathbf{I}_{|k} \text{ (pero el resto de columnas no cambian)} \end{cases}, \text{ donde } k = 1:n.$$

Las matrices elementales Tipo II son como la identidad salvo la $\boldsymbol{i}$ésima componente de la diagonal, que es α.

Transformación de las columnas de una matriz A: En general, si aplicamos $\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}$ a las columnas de $\mathbf{A}$, de orden $m \times n$, obtenemos una matriz cuyas columnas son:

$$\left(\mathbf{A}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\right)_{|k} = \begin{cases} \alpha\mathbf{A}_{|k} & = \mathbf{A}\left(\alpha\mathbf{I}_{|k}\right) \qquad \text{si } k = i \\ \\ \mathbf{A}_{|k} & = \mathbf{A}\left(\mathbf{I}_{|k}\right) \qquad \text{para el resto de columnas} \end{cases} = \mathbf{A}\left(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\right)_{|k}, \quad \text{donde } k = 1:n.$$

De nuevo, aplicar una transformación elemental de *Tipo II* sobre la matriz $\mathbf{A}$ (de n columnas) es equivalente a multiplicar dicha matriz $\mathbf{A}$ por la correspondiente matriz elemental *Tipo II* de orden n por la derecha.

$$\mathbf{A}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}} = \mathbf{A}\left(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\right).$$

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 α resulta que:

$$\mathbf{A}_{\underset{[(\alpha)\boldsymbol{2}]}{\boldsymbol{\tau}}} = \begin{bmatrix} 0 & 0 & c \\ 0 & \alpha b & c \\ a & \alpha b & c \end{bmatrix} = \begin{bmatrix} 0 & 0 & c \\ 0 & b & c \\ a & b & c \end{bmatrix} \begin{bmatrix} 1 & & \\ & \alpha & \\ & & 1 \end{bmatrix} = \mathbf{A}\left(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{2}]}{\boldsymbol{\tau}}}\right).$$

4.1.3. Transformaciones de las columnas

Como al aplicar cualquier transformación elemental $\boldsymbol{\tau}$ sobre las columnas de $\mathbf{A}$ siempre se verifica que:

$$\boxed{\mathbf{A}_{\boldsymbol{\tau}} = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}}\right)};$$

disponemos de una nueva regla de reescritura asociativa:

$$\boxed{\left(\mathbf{AB}\right)_{\boldsymbol{\tau}} \;=\; \mathbf{AB}\left(\mathbf{I}_{\boldsymbol{\tau}}\right) \;=\; \mathbf{A}\left(\mathbf{B}_{\boldsymbol{\tau}}\right)}.$$

Por tanto no son necesarios lo paréntesis y basta escribir: $\mathbf{AB}_{\boldsymbol{\tau}}$

Comentario. *Fíjese que*

- ***multiplicar** una columna **por cero no es una transformación elemental**; y **restar una columna a ella misma tampoco lo es**.*
- *la matriz identidad es simultáneamente una matriz elemental **Tipo I** y **Tipo II**:* $\mathbf{I}_{\underset{[(0)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}} = \mathbf{I} = \mathbf{I}_{\underset{[(1)\boldsymbol{i}]}{\boldsymbol{\tau}}}$.
- *la notación para denotar las matrices elementales las describe casi completamente (aunque no del todo, pues la notación no indica el orden de la matriz).*

  Por ejemplo, $\mathbf{I}_{\underset{[(-7)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}$ *(de orden n), es como la matriz identidad (de orden n) pero la componente de la fila 1 y columna 2 es un "-7", es decir*

$$\mathbf{I}_{\underset{[(-7)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}} = \begin{bmatrix} 1 & -7 \\ 0 & 1 \end{bmatrix}, \quad \textit{pero también} \quad \mathbf{I}_{\underset{[(-7)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}} = \begin{bmatrix} 1 & -7 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \textit{ó} \quad \mathbf{I}_{\underset{[(-7)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}} = \begin{bmatrix} 1 & -7 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \textit{ó...}$$

Por otra parte, $\mathbf{I}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}}$ *indica que la primera componente de la diagonal es un* 3*; es decir,*

$$\mathbf{I}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}} = \begin{bmatrix}3\end{bmatrix}; \quad \textit{pero también,} \quad \mathbf{I}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}} = \begin{bmatrix}3&0\\0&1\end{bmatrix}; \quad \textit{ó} \quad \mathbf{I}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}} = \begin{bmatrix}3&0&0\\0&1&0\\0&0&1\end{bmatrix}, \quad \textit{ó...}$$

- *las matrices elementales (de orden n) se obtienen aplicando las correspondientes transformaciones elementales sobre la matriz identidad (de orden n)*
- *La transpuesta de una matriz elemental es otra matriz elemental. Las de Tipo II son simétricas y para las de Tipo I resulta evidente que* $\left(\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)^{\intercal} = \mathbf{I}_{\underset{[(\lambda)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}}$ *es otra matriz elemental de Tipo I (la que suma λ veces la $j$ésima columna a la $i$ésima).*

☞ Solo hay dos tipos de *transformaciones elementales*

Tipo I. Suman a un vector un múltiplo de *otro* vector: $\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}$

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 a las columnas de la matriz identidad, $\mathbf{I}_{\boldsymbol{\tau}}$. Hay dos tipos:

Tipo I. $\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}$, donde $j \neq i$

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:

$$\boxed{(\mathbf{AB})_{\boldsymbol{\tau}} = \mathbf{A}(\mathbf{B}_{\boldsymbol{\tau}}) = \mathbf{AB}_{\boldsymbol{\tau}}.}$$

La transpuesta de una matriz elemental también es elemental (y del mismo tipo).

EJERCICIO 24. 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 25. ¿Qué operaciones realizan las matrices anteriores si multiplican por la derecha a $\mathbf{B}$?

EJERCICIO 26. 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.

4.2. Secuencias de transformaciones elementales. Parte I

Considere la aplicación de una secuencia de k transformaciones elementales sobre las columnas de $\mathbf{A}$:

$$\Big(\dots\big((\mathbf{A}_{\boldsymbol{\tau}_1})_{\boldsymbol{\tau}_2}\dots\big)_{\boldsymbol{\tau}_{(k-1)}}\Big)_{\boldsymbol{\tau}_k}$$

Denotamos la aplicación 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:

$$\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \;=\; \Big(\dots\big((\mathbf{A}_{\boldsymbol{\tau}_1})_{\boldsymbol{\tau}_2}\dots\big)_{\boldsymbol{\tau}_{(k-1)}}\Big)_{\boldsymbol{\tau}_k} \;=\; \Big(\cdots\big((\mathbf{A}\mathbf{I}_{\boldsymbol{\tau}_1})\mathbf{I}_{\boldsymbol{\tau}_2}\big)\cdots\mathbf{I}_{\boldsymbol{\tau}_{(k-1)}}\Big)\mathbf{I}_{\boldsymbol{\tau}_k} \;=\; \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1}\right)\cdots\left(\mathbf{I}_{\boldsymbol{\tau}_k}\right).$$

En particular, si $\mathbf{A}=\mathbf{I}$ entonces

$$\boxed{\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}=\left(\mathbf{I}_{\boldsymbol{\tau}_1}\right)\cdots\left(\mathbf{I}_{\boldsymbol{\tau}_k}\right)} \tag{4.2}$$

es decir, *la matriz* $\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)$ *es producto de matrices elementales*; y por tanto

$$\boxed{\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}=\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)}$$

EJERCICIO 27. 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:

$$\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \;=\; \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1}\right)\cdots\left(\mathbf{I}_{\boldsymbol{\tau}_k}\right) \;=\; \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)$$

la aplicación de una secuencia de transformaciones elementales de las columnas es un **operador lineal**.

Notación Denotaremos una secuencia de k trasformaciones de las *columnas* de $\mathbf{A}$ con el siguiente esquema:

$$\mathbf{A}\xrightarrow{\boldsymbol{\tau}_1}\mathbf{A}_{\boldsymbol{\tau}_1}\xrightarrow{\boldsymbol{\tau}_2}\mathbf{A}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2}\xrightarrow{\boldsymbol{\tau}_3}\cdots\xrightarrow{\boldsymbol{\tau}_k}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \qquad \text{o agregando varios pasos} \qquad \mathbf{A}\xrightarrow{\substack{\boldsymbol{\tau}_1\\ \vdots\\ \boldsymbol{\tau}_p}}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p}\xrightarrow{\substack{\boldsymbol{\tau}_{p+1}\\ \vdots\\ \boldsymbol{\tau}_k}}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

Por ejemplo, la sucesión de transformaciones $\mathbf{A}_{\underset{[(3)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(2)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}$ donde $\mathbf{A}=\begin{bmatrix}0&1&5\\7&0&4\end{bmatrix}$, se escribe como:

$$\begin{bmatrix}0&1&5\\7&0&4\end{bmatrix}\xrightarrow{\underset{[(3)\mathbf{2}]}{\boldsymbol{\tau}}}\begin{bmatrix}0&3&5\\7&0&4\end{bmatrix}\xrightarrow{\underset{[(2)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}\begin{bmatrix}0&3&5\\7&14&4\end{bmatrix}; \qquad \text{o bien como} \qquad \begin{bmatrix}0&1&5\\7&0&4\end{bmatrix}\xrightarrow{\substack{\boldsymbol{\tau}\\ [(3)\mathbf{2}]\\ [(2)\mathbf{1}+\mathbf{2}]}}\begin{bmatrix}0&3&5\\7&14&4\end{bmatrix}.$$

Alternativamente podemos expresar una secuencia de transformaciones elementales mediante productos de matrices elementales. Por ejemplo, para la secuencia anterior:

$$\begin{bmatrix}0&1&5\\7&0&4\end{bmatrix}_{\substack{\boldsymbol{\tau}\\ [(3)\mathbf{2}]\\ [(2)\mathbf{1}+\mathbf{2}]}} = \begin{bmatrix}0&1&5\\7&0&4\end{bmatrix}\Big(\mathbf{I}_{\underset{[(3)\mathbf{2}]}{\boldsymbol{\tau}}}\Big)\Big(\mathbf{I}_{\underset{[(2)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}\Big) = \begin{bmatrix}0&3&5\\7&14&4\end{bmatrix}.$$

Librería NAcAL para Python

```
T( (2,3,4) ) & T( (-7,1) ) & T( (-1,3,1) )
```

$$\substack{\boldsymbol{\tau}\\ [(2)\mathbf{3}+\mathbf{4}]\\ [(-7)\mathbf{1}]\\ [(-1)\mathbf{3}+\mathbf{1}]}$$

4.2.1. 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: $\underset{[\boldsymbol{i}\rightleftharpoons\boldsymbol{j}]}{\boldsymbol{\tau}}$ → intercambia los vectores $\boldsymbol{i}$ésimo y $\boldsymbol{j}$ésimo.

Librería NAcAL para Python

```
T( {1, 3} )    # intercambia los vectores primero y tercero
```

$$\underset{[\mathbf{1}\rightleftharpoons\mathbf{3}]}{\tau}$$

Definición 4.2 (Matriz de intercambio)**.** *Cualquier matriz obtenida tras intercambiar las columnas p y q de la matriz identidad de orden n:* $\mathbf{I}_{\underset{[\boldsymbol{p}\rightleftharpoons\boldsymbol{q}]}{\boldsymbol{\tau}}}$

Puesto que en $\mathbf{I}_{\underset{[\boldsymbol{p}\rightleftharpoons\boldsymbol{q}]}{\boldsymbol{\tau}}}$ solo han cambiado las columnas p y q de $\mathbf{I}$, se deduce que $\mathbf{I}_{\underset{[\boldsymbol{p}\rightleftharpoons\boldsymbol{q}]}{\boldsymbol{\tau}}}$ es simétrica:

$$\left(\mathbf{I}_{\underset{[\boldsymbol{p}\rightleftharpoons\boldsymbol{q}]}{\boldsymbol{\tau}}}\right)_{|j} = \begin{cases} \mathbf{I}_{|q} & \text{si } j=p \\ \mathbf{I}_{|p} & \text{si } j=q \\ \mathbf{I}_{|j} & \text{si } j\neq p,q \end{cases} \implies {}_{k|}\left(\mathbf{I}_{\underset{[\boldsymbol{p}\rightleftharpoons\boldsymbol{q}]}{\boldsymbol{\tau}}}\right)_{|j} = \begin{cases} 1 & \text{si} \quad k=q \quad \text{y} \quad j=p \\ 1 & \text{si} \quad k=p \quad \text{y} \quad j=q \\ 1 & \text{si} \quad k=j\neq p,q \\ 0 & \text{en el resto de casos} \end{cases} = {}_{j|}\left(\mathbf{I}_{\underset{[\boldsymbol{p}\rightleftharpoons\boldsymbol{q}]}{\boldsymbol{\tau}}}\right)_{|k},$$

es decir, transponiendo (intercambiando los índices k y j) la matriz no cambia. Por tanto:

$$\boxed{\textit{toda matriz intercambio } \mathbf{I}_{\underset{[\boldsymbol{p}\rightleftharpoons\boldsymbol{q}]}{\boldsymbol{\tau}}} \textit{ es simétrica.}}$$

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 28.

(a) Mediante una secuencia de transformaciones elementales Tipo I y II transforme $\underset{3\times 3}{\mathbf{I}}$ 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 $\mathbf{I}_{\underset{[\boldsymbol{i}\rightleftharpoons\boldsymbol{j}]}{\boldsymbol{\tau}}}$ de orden n.

Librería NAcAL para Python

```
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
```

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

☞ Toda matriz intercambio $\mathbf{I}_{\underset{[\boldsymbol{i}\rightleftharpoons\boldsymbol{j}]}{\boldsymbol{\tau}}}$ se puede factorizar como producto de varias matrices elementales **Tipo I** y una matriz elemental **Tipo II**, que multiplica por -1 una de las columnas intercambiadas.

Intercambio de las columnas de $\mathbf{A}$:

$$\left(\mathbf{A}_{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}}\right)_{|k} = \begin{cases} \mathbf{A}_{|j} = \mathbf{A}\mathbf{I}_{|j} & \text{si } k=i \text{ (es decir, en la posición } i \text{ se coloca la columna } j) \\ \mathbf{A}_{|i} = \mathbf{A}\mathbf{I}_{|i} & \text{si } k=j \text{ (es decir, en la posición } j \text{ se coloca la columna } i) \\ \mathbf{A}_{|k} = \mathbf{A}\mathbf{I}_{|k} & \text{en el resto de casos (se mantienen las columnas en su sitio)} \end{cases} = \mathbf{A}\left(\mathbf{I}_{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}}\right)_{|k}.$$

Si $\mathbf{A} = \begin{bmatrix} a & 0 & 0 \\ b & b & 0 \\ c & c & c \end{bmatrix}$, entonces $\mathbf{A}_{\underset{[1\rightleftharpoons 2]}{\boldsymbol{\tau}}} = \mathbf{A}\left(\mathbf{I}_{\underset{[1\rightleftharpoons 2]}{\boldsymbol{\tau}}}\right) = \begin{bmatrix} a & 0 & 0 \\ b & b & 0 \\ c & c & c \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ intercambia las columnas 1 y 2.

Librería 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
```

$$\begin{bmatrix} 0 & a & 0 \\ b & a & 0 \\ b & a & c \end{bmatrix}$$

4.3. Permutaciones

Mediante una sucesión intercambios se logra un reordenamiento o *permutación* de vectores.

Permutación: $\underset{[\circlearrowleft]}{\boldsymbol{\tau}}$ $\rightarrow$ reordena una lista mediante una sucesión de intercambios.

Definición 4.3 (Matriz permutación)**.** *Cualquier matriz obtenida tras realizar un número arbitrario de intercambios entre las columnas de la matriz identidad de orden n,* $\mathbf{I}_{\underset{[\circlearrowleft]}{\boldsymbol{\tau}}}$.

Librería 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 $\mathbf{I}_{\underset{[p\rightleftharpoons q]}{\boldsymbol{\tau}}}$ son simétricas, las matrices permutación $\mathbf{I}_{\underset{[\circlearrowleft]}{\boldsymbol{\tau}}}$ *no* son simétricas en general:[2]

$$\left(\mathbf{I}_{\underset{[1\rightleftharpoons 3]}{\boldsymbol{\tau}}}\right)\left(\mathbf{I}_{\underset{[1\rightleftharpoons 2]}{\boldsymbol{\tau}}}\right) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

[2] el producto de dos matrices simétricas es simétrico si y solo si dichas matrices conmutan, es decir $\mathbf{AB} = \mathbf{BA}$

☞ El intercambio, $\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}$, es una secuencia de transformaciones *Tipo I* junto a una de *Tipo II* (*que cambia el signo de uno de los vectores*). Las matrices intercambio, $\mathbf{I}_{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}}}$, son simétricas y su cuadrado es la matriz identidad ("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 $\mathbf{I}_{\underset{[\circlearrowleft]}{\boldsymbol{\tau}}}$ no son simétricas.

4.4. 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 está el pivote.*

Definición 4.5. *Una matriz* $\mathbf{K}$ es pre-escalonada *si las componentes a la derecha de cada pivote son nulas.*

Ejemplo 3.

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 \\ 0 & 0 & 0 & 6 \\ 3 & 0 & 0 & 0 \\ 6 & 0 & 1 & 0 \end{bmatrix}$$

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 a de la última es 2 (segunda fila).

4.4.1. 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:

$$\begin{bmatrix} 0 & 1 & 1 \\ 2 & -2 & 2 \\ 2 & -8 & 1 \end{bmatrix} \xrightarrow[{[(-1)\mathbf{2}+\mathbf{3}]}]{\boldsymbol{\tau}} \begin{bmatrix} 0 & 1 & 0 \\ 2 & -2 & 4 \\ 2 & -8 & 9 \end{bmatrix} \xrightarrow[{\substack{[(1)\mathbf{1}+\mathbf{2}] \\ [(-2)\mathbf{1}+\mathbf{3}]}}]{\boldsymbol{\tau}} \begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & 0 \\ 2 & -6 & 5 \end{bmatrix}$$

Toda matriz se puede *pre-escalonar* mediante transformaciones elementales. *¡El prácticamente resto del curso descansa sobre en este teorema!* Lo demostraremos por inducción; es decir, demostrando que si es cierto para cualquier matriz de n columnas, también lo es para cualquier matriz de $n+1$ columnas.

Teorema 4.4.1 (Eliminación)**.** *Para toda* $\mathbf{A}$ *existe una secuencia* $\boldsymbol{\tau}_1 \dots \boldsymbol{\tau}_k$ *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 trasformaciones elementales de las columnas tal que

$$\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} = \left[\begin{array}{ccc} & \mathbf{K} & \\ \hline b_{m1} & \cdots & b_{mn} \end{array}\right]; \qquad \text{y donde la submatriz } \mathbf{K} \text{ es } \textit{pre-escalonada}.$$

Entonces $\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$ es pre-escalonada si se da alguno de los siguientes *supuestos* para su última fila:

- *Supuesto 1.* La última fila es nula.
- *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 3).

- *Supuesto 3.* De los componentes de la última fila situados bajo las columnas nulas de K, solo uno es distinto de cero, y las componentes a su derecha son nulas (como en la última fila del Ejemplo 3, 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.

Así pues, si b_{mr} es primer pivote de los situados en la última fila (i.e., si de aquellos componentes situados bajo las columnas nulas de K, es el situado más a la izquierda); entonces podemos *eliminar* los componentes a su derecha restando de cada columna $j$ésima (con $j \geq r$) la columna $r$ésima multiplicada por $\left(\frac{b_{mj}}{b_{mr}}\right)$. Es decir, 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 K, es decir con una matriz A con una sola fila y donde $\mathsf{A}_{|r}$ es la primera columna no nula. □

Librería NAcAL para Python

```
A = Matrix( [ [0,2,2], [1,-2,-8], [1,1,1] ] )
Elim(A,1)                                          # Eliminación por columnas
```

$$\begin{bmatrix}0&2&2\\1&-2&-8\\1&1&1\end{bmatrix} \xrightarrow[{[(-1)\mathbf{2}+\mathbf{3}]}]{\boldsymbol{\tau}} \begin{bmatrix}0&2&0\\1&-2&-6\\1&1&0\end{bmatrix} \xrightarrow[\substack{[(2)\mathbf{1}+\mathbf{2}]\\ [(6)\mathbf{1}+\mathbf{3}]}]{\boldsymbol{\tau}} \begin{bmatrix}0&2&0\\1&0&0\\1&3&6\end{bmatrix}$$

Aunque el método de eliminación está implementado en NAcAL, es muy conveniente que usted sepa aplicar el algoritmo con lápiz y papel (o bien, que sepa programarlo). [4]

4.4.2. Método de eliminación Gaussiano

Definición 4.6. *Decimos que la matriz* L es escalonada*, si toda columna que precede a una no nula* $\mathsf{L}_{|k}$ *es no nula y su posición de pivote es anterior a la posición de pivote de* $\mathsf{L}_{|k}$.

(Esta definición implica que a la derecha de cualquier pivote solo puede 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").

Ejemplo 4.

$$\mathsf{L} = \begin{bmatrix}0&0&0&0\\7&0&0&0\\6&0&0&0\\0&3&0&0\\0&6&1&0\end{bmatrix}$$

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.

[3] que no modifican K puesto que por encima del pivote b_{mr} todo son ceros

[4] Use la librería solo para verificar los resultados que usted ha obtenido previamente usando lápiz y papel.

Fíjese que aunque toda matriz *escalonada* es *pre-escalonada*, las matrices *pre-escalonadas* no suelen ser *escalonadas*. (Ejemplo 3 en la página 51).

Toda matriz se puede escalonar mediante transformaciones elementales tal como afirma el siguiente

Corolario 4.4.2 (Eliminación Gaussiana)**.** *Para toda* $\mathbf{A}$ *existen* $\boldsymbol{\tau}_1 \ldots \boldsymbol{\tau}_k$ *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 de Gaussiano*.

Librería NAcAL para Python

```
A = Matrix( [ [0,2,2], [1,-2,-8], [1,1,1] ] )
L = ElimG(A,1)                               # Eliminación Gaussiana por columnas
```

$$\begin{bmatrix}0&2&2\\1&-2&-8\\1&1&1\end{bmatrix} \xrightarrow[{[(-1)\mathbf{2}+\mathbf{3}]}]{\boldsymbol{\tau}} \begin{bmatrix}0&2&0\\1&-2&-6\\1&1&0\end{bmatrix} \xrightarrow[{[(6)\mathbf{1}+\mathbf{3}]}]{\substack{\boldsymbol{\tau}\\ [(2)\mathbf{1}+\mathbf{2}]}} \begin{bmatrix}0&2&0\\1&0&0\\1&3&6\end{bmatrix} \xrightarrow[{[\mathbf{1}\rightleftharpoons\mathbf{2}]}]{\boldsymbol{\tau}} \begin{bmatrix}2&0&0\\0&1&0\\3&1&6\end{bmatrix}$$

4.4.3. Método de eliminación Gauss-Jordan

La eliminación gaussiana se pude 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 derecha e izquierda de los pivotes son cero.*

Ejemplo 5.

$$\mathbf{R} = \begin{bmatrix}0&0&0&0\\1&0&0&0\\\frac{6}{7}&0&0&0\\0&1&0&0\\0&0&1&0\end{bmatrix}$$

Corolario 4.4.3 (Eliminación Gauss-Jordan)**.** *Para toda* $\mathbf{A}$ *existen* $\boldsymbol{\tau}_1 \ldots \boldsymbol{\tau}_k$ *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*.

Librería NAcAL para Python

```
A = Matrix( [ [0,2,2], [1,-2,-8], [1,1,1] ] )
R = ElimGJ(A,1)                              # Forma escalonada reducida
```

$$\begin{bmatrix}0&2&2\\1&-2&-8\\1&1&1\end{bmatrix} \xrightarrow[{[(6)\mathbf{1}+\mathbf{3}]}]{\substack{\boldsymbol{\tau}\\ [(-1)\mathbf{2}+\mathbf{3}]\\ [(2)\mathbf{1}+\mathbf{2}]}} \begin{bmatrix}0&2&0\\1&0&0\\1&3&6\end{bmatrix} \xrightarrow[{[\mathbf{1}\rightleftharpoons\mathbf{2}]}]{\boldsymbol{\tau}} \begin{bmatrix}2&0&0\\0&1&0\\3&1&6\end{bmatrix} \xrightarrow[{[(-1)\mathbf{3}+\mathbf{2}]}]{\substack{\boldsymbol{\tau}\\ [(2)\mathbf{1}]\\ [(-1)\mathbf{3}+\mathbf{1}]\\ [(6)\mathbf{2}]}} \begin{bmatrix}4&0&0\\0&6&0\\0&0&6\end{bmatrix} \xrightarrow[{[(\frac{1}{6})\mathbf{3}]}]{\substack{\boldsymbol{\tau}\\ [(\frac{1}{4})\mathbf{1}]\\ [(\frac{1}{6})\mathbf{2}]}} \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

☞ Mediante transformaciones elementales se puede *pre-escalonar*, escalonar, o *escalonar y reducir* cualquier matriz.

- 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).

4.5. Nota sobre las transformaciones elementales de las filas.

En casi todos los manuales de Álgebra Lineal se aplican las transformaciones elementales sobre las filas. Aquí usaremos dichas transformaciones para demostrar la Proposición 5.4.4, para encontrar el espacio nulo por la izquierda de $\mathbf{A}$ (Capítulo V) y para diagonalizar matrices (Capítulo VI). Vamos a comentar brevemente las transformaciones sobre las filas de una matriz y su relación con las transformaciones sobre las columnas.

- Podemos aplicar una transformación elemental a las filas de $\mathbf{A}$ transformando las columnas de $\mathbf{A}^\intercal$, y transponiendo el resultado de nuevo. Y operando concluimos que

$$\left((\mathbf{A}^\intercal)_{\boldsymbol{\tau}}\right)^\intercal = \left((\mathbf{A}^\intercal)\mathbf{I}_{\boldsymbol{\tau}}\right)^\intercal = ((\mathbf{I}_{\boldsymbol{\tau}})^\intercal)\left(\mathbf{A}^\intercal\right)^\intercal = ((\mathbf{I}_{\boldsymbol{\tau}})^\intercal)\mathbf{A},$$

  donde $(\mathbf{I}_{\boldsymbol{\tau}})^\intercal$ es una matriz elemental (por ser la transpuesta de una elemental).

- Por tanto, multiplicar por una matriz elemental *por la izquierda* transforma las filas de $\mathbf{A}$.
- Ello sugiere la notación: ${}_{\boldsymbol{\tau}}\mathbf{A}$ para la transformación $\boldsymbol{\tau}$ de las filas de $\mathbf{A}$;

$${}_{\boldsymbol{\tau}}\mathbf{A} = \left((\mathbf{I}_{\boldsymbol{\tau}})^\intercal\right)\mathbf{A}$$

  y consecuentemente: ${}_{\boldsymbol{\tau}}\mathbf{I}$ para la transformación $\boldsymbol{\tau}$ de las filas de $\mathbf{I}$; es decir,

$${}_{\boldsymbol{\tau}}\mathbf{I} = (\mathbf{I}_{\boldsymbol{\tau}})^\intercal.$$

  (nuevamente la transposición supone cambiar de lado el subíndice). Así pues,

$${}_{\boldsymbol{\tau}}\mathbf{A} = \left({}_{\boldsymbol{\tau}}\mathbf{I}\right)\mathbf{A}.$$

  ¡No solo eso!...Puesto que $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = (\mathbf{I}_{\boldsymbol{\tau}_1})\cdots(\mathbf{I}_{\boldsymbol{\tau}_k})$ y puesto que el producto de matrices es asociativo, deducimos que la transpuesta de $\mathbf{I}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2\cdots\boldsymbol{\tau}_k}$ es

$$\left(\mathbf{I}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2\cdots\boldsymbol{\tau}_k}\right)^\intercal \;=\; \left((\mathbf{I}_{\boldsymbol{\tau}_1})\cdots(\mathbf{I}_{\boldsymbol{\tau}_k})\right)^\intercal \;=\; (\mathbf{I}_{\boldsymbol{\tau}_k})^\intercal\cdots(\mathbf{I}_{\boldsymbol{\tau}_i})^\intercal \;=\; {}_{\boldsymbol{\tau}_k}\mathbf{I}\cdots{}_{\boldsymbol{\tau}_1}\mathbf{I} \;=\; {}_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_2\boldsymbol{\tau}_1}\mathbf{I}$$

  Nótese cómo al transponer no solo cambiamos de lado los subíndices, sino también invertimos el orden de la secuencia de transformaciones (de manera similar a cuando se transpone un producto de matrices).

- Esto sugiere denotar a la operación de invertir el orden de las transformaciones como una transposición:

$$(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^\intercal = \boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1;$$

- Fíjese que aunque la secuencia $\boldsymbol{\tau}_1\boldsymbol{\tau}_2\cdots\boldsymbol{\tau}_k$ 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.

- Así, $\left(\mathbf{A}_{(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)}\right)^\intercal = \left(\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)\right)^\intercal = \left({}_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{I}\right)(\mathbf{A})^\intercal = {}_{(\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1)}(\mathbf{A})^\intercal = {}_{(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^\intercal}(\mathbf{A})^\intercal.$

Librería NAcAL para Python

```
Tr = T( [(-7,1,2), {3,2}, (2,1,3), {1,2}], 'h' )
Tr
```

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

Librería NAcAL para Python

```
~Tr                     #Transposición invierte el orden
```

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

Librería NAcAL para Python

```
I(3) & Tr
```

$$\begin{bmatrix}0 & 1 & -5\\ 0 & 0 & 1\\ 1 & 0 & 0\end{bmatrix}$$

Librería NAcAL para Python

```
(~Tr) & I(3)              # transpuesta de la matriz anterior
```

$$\begin{bmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ -5 & 1 & 0\end{bmatrix}$$

- Para indicar que una secuencia de transformaciones actúa sobre las *filas* de una matriz, escribiremos las abreviaturas por debajo de las fechas:

Librería NAcAL para Python

```
A = Matrix([ [0,2,2], [1,-2,-8], [3,6,-6] ])
U = ElimGF(A,1)                # Escalonamiento operando con las filas
```

$$\begin{bmatrix}0 & 2 & 2\\ 1 & -2 & -8\\ 3 & 6 & -6\end{bmatrix} \xrightarrow[\underset{[(-3)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}]{} \begin{bmatrix}0 & 2 & 2\\ 1 & -2 & -8\\ 0 & 12 & 18\end{bmatrix} \xrightarrow[\substack{\underset{[(-6)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}\\ [(1)\mathbf{1}+\mathbf{2}]}]{} \begin{bmatrix}0 & 2 & 2\\ 1 & 0 & -6\\ 0 & 0 & 6\end{bmatrix} \xrightarrow[\underset{[\mathbf{1}\rightleftharpoons\mathbf{2}]}{\boldsymbol{\tau}}]{} \begin{bmatrix}1 & 0 & -6\\ 0 & 2 & 2\\ 0 & 0 & 6\end{bmatrix}$$

Definición 4.8. *Una matriz está* escalonada por filas *si su transpuesta está escalonada por columnas.*

La definición anterior sugiere otra para *pivote de una fila*; útil cuando se escalona operando por filas:

Definición 4.9. *Llamamos* pivote de una fila *no nula, a su primer componente no nulo, y* posición de pivote *al índice de la columna en la que está el pivote.*

- Sabemos que una secuencia de transformaciones sobre las columnas se puede expresar de varias formas:

$$\mathbf{B}_{\underset{[\mathbf{1}\rightleftharpoons\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-6)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-3)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}} \;=\; \mathbf{B}\Big(\mathbf{I}_{\underset{[\mathbf{1}\rightleftharpoons\mathbf{2}]}{\boldsymbol{\tau}}}\Big)\Big(\mathbf{I}_{\underset{[(-6)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}}\Big)\Big(\mathbf{I}_{\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}\Big)\Big(\mathbf{I}_{\underset{[(-3)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}}\Big) \;=\; \mathbf{B}_{\underset{\substack{[\mathbf{1}\rightleftharpoons\mathbf{2}]\\ [(-6)\mathbf{1}+\mathbf{3}]\\ [(1)\mathbf{1}+\mathbf{2}]\\ [(-3)\mathbf{2}+\mathbf{3}]}}{\boldsymbol{\tau}}}$$

  También disponemos de representaciones alternativas al transformar las *filas* de $\mathbf{A}$:

$${}_{\underset{[\mathbf{1}\rightleftharpoons\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-6)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-3)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}}\mathbf{A} \;=\; \Big({}_{\underset{[\mathbf{1}\rightleftharpoons\mathbf{2}]}{\boldsymbol{\tau}}}\mathbf{I}\Big)\Big({}_{\underset{[(-6)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}}\mathbf{I}\Big)\Big({}_{\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}\mathbf{I}\Big)\Big({}_{\underset{[(-3)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}}\mathbf{I}\Big)\mathbf{A} \;=\; {}_{\underset{\substack{[\mathbf{1}\rightleftharpoons\mathbf{2}]\\ [(-6)\mathbf{1}+\mathbf{3}]\\ [(1)\mathbf{1}+\mathbf{2}]\\ [(-3)\mathbf{2}+\mathbf{3}]}}{\boldsymbol{\tau}}}\mathbf{A}$$

- Al operar sobre las filas, las primeras transformaciones en actuar son las que quedan más a la derecha de $\underset{[\mathbf{1}\rightleftharpoons\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-6)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-3)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}$ o más abajo de $\underset{\substack{[\mathbf{1}\rightleftharpoons\mathbf{2}]\\ [(-6)\mathbf{1}+\mathbf{3}]\\ [(1)\mathbf{1}+\mathbf{2}]\\ [(-3)\mathbf{2}+\mathbf{3}]}}{\boldsymbol{\tau}}$. Por tanto, el proceso de eliminación por filas

$$\begin{bmatrix}0&2&2\\1&-2&-8\\3&6&-6\end{bmatrix}\xrightarrow[{\underset{[(-3)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}}]{}\begin{bmatrix}0&2&2\\1&-2&-8\\0&12&18\end{bmatrix}\xrightarrow[{\underset{\substack{[(-6)\mathbf{1}+\mathbf{3}]\\ [(1)\mathbf{1}+\mathbf{2}]}}{\boldsymbol{\tau}}}]{}\begin{bmatrix}0&2&2\\1&0&-6\\0&0&6\end{bmatrix}\xrightarrow[{\underset{[\mathbf{1}\rightleftharpoons\mathbf{2}]}{\boldsymbol{\tau}}}]{}\begin{bmatrix}1&0&-6\\0&2&2\\0&0&6\end{bmatrix}$$

  se puede representar de estas dos formas alternativas:

$${}_{\underset{[\mathbf{1}\rightleftharpoons\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-6)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-3)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}}\begin{bmatrix}0&2&2\\1&-2&-8\\3&6&-6\end{bmatrix} = {}_{\underset{\substack{[\mathbf{1}\rightleftharpoons\mathbf{2}]\\ [(-6)\mathbf{1}+\mathbf{3}]\\ [(1)\mathbf{1}+\mathbf{2}]\\ [(-3)\mathbf{2}+\mathbf{3}]}}{\boldsymbol{\tau}}}\begin{bmatrix}0&2&2\\1&-2&-8\\3&6&-6\end{bmatrix} = \begin{bmatrix}1&0&-6\\0&2&2\\0&0&6\end{bmatrix}.$$

- La transpuesta de una matriz elemental es una matriz elemental del mismo tipo.
- Aunque las matrices elementales *Tipo II* son simétricas ${}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\mathbf{I} = \mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}$, las matrices elementales *Tipo I* NO lo son (salvo para el caso $\lambda = 0$):

$$\mathbf{I}_{\underset{[(\lambda)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}} = {}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\mathbf{I} \neq \mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}} = {}_{\underset{[(\lambda)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}}\mathbf{I}.$$

- En general ${}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\mathbf{I} \neq \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ y ${}_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{I} \neq \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$.
- Pero sin embargo ${}_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{I} = {}_{(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{\intercal}}\mathbf{I} = \left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)^{\intercal}$ (fíjese en el orden de los subíndices).

Librería NAcAL para Python

```
(Tr & I(3)) == (~(~Tr) & I(3)) == (~(I(3) & ~Tr))
```

```
True
```

- Incluso cuando operamos con las filas, la notación sigue siendo asociativa:

$$\boxed{\left({}_{\boldsymbol{\tau}}\mathbf{A}\right)\mathbf{B} = {}_{\boldsymbol{\tau}}(\mathbf{A}\mathbf{B})}$$

  Así que podemos escribir sencillamente ${}_{\boldsymbol{\tau}}\mathbf{A}\mathbf{B}$.

Matrices inversas

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).

Pues bien, 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*

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I} \qquad \textit{y que} \qquad \mathbf{AB'} = \mathbf{B'A} = \mathbf{I}$$

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*

$$\mathbf{A}\left(\mathbf{A}^{-1}\right) = \left(\mathbf{A}^{-1}\right)\mathbf{A} = \mathbf{I}.$$

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 29. 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* $\mathbf{A}_1, \ldots, \mathbf{A}_k$ *tienen inversa, entonces* $\boxed{(\mathbf{A}_1 \cdots \mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1} \cdots \mathbf{A}_1^{-1}.}$

Pista. $\left(\mathbf{A}_1 \cdots \mathbf{A}_k\right)^{-1} = \left(\mathbf{A}_1(\mathbf{A}_2 \cdots \mathbf{A}_k)\right)^{-1} = \left(\mathbf{A}_2 \cdots \mathbf{A}_k\right)^{-1}\mathbf{A}_1^{-1}.$

(c) **Proposición 5.1.4.** *Sea* $\mathbf{A}$ *de orden* n *y sea* $\mathbf{B}$ *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 sea* $\mathbf{B}$ *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:* $\boxed{\left(\mathbf{A}^\intercal\right)^{-1} = \left(\mathbf{A}^{-1}\right)^\intercal.}$

Pista. Recuerde que $\mathbf{B}^\intercal\mathbf{A}^\intercal = \left(\mathbf{AB}\right)^\intercal$ y aplíquelo a $\mathbf{A}^\intercal\left(\mathbf{A}^{-1}\right)^\intercal$ y a $\left(\mathbf{A}^{-1}\right)^\intercal\mathbf{A}^\intercal$.

(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 es única y se denota con $\mathbf{A}^{-1}$.

- el producto de matrices invertibles es invertible: $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
- la transpuesta de una matriz invertible es invertible: $\left(\mathbf{A}^\intercal\right)^{-1} = \left(\mathbf{A}^{-1}\right)^\intercal$.
- Si alguna fila o columna de $\mathbf{A}$ es nula la matriz es *singular* (no tiene inversa).

5.1.1. Inversa de las matrices (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:

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

Algo similar ocurre cuando multiplicamos y dividimos una columna por $\alpha \neq 0$:

$$\mathbf{A}_{\underset{[(5)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(\frac{1}{5})\mathbf{2}]}{\boldsymbol{\tau}}} = \mathbf{A}\Big(\mathbf{I}_{\underset{[(5)\mathbf{2}]}{\boldsymbol{\tau}}}\Big)\Big(\mathbf{I}_{\underset{[(\frac{1}{5})\mathbf{2}]}{\boldsymbol{\tau}}}\Big) = \mathbf{A}\begin{bmatrix}1&0&0\\0&5&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&0&0\\0&\frac{1}{5}&0\\0&0&1\end{bmatrix} = \mathbf{A}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} = \mathbf{A}$$

En general,

$$\mathbf{A}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}\underset{[(-\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}} = \mathbf{A}_{\underset{[(-\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}} = \mathbf{A} \qquad \text{y} \qquad \mathbf{A}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}\underset{[(\frac{1}{\alpha})\boldsymbol{i}]}{\boldsymbol{\tau}}} = \mathbf{A}_{\underset{[(\frac{1}{\alpha})\boldsymbol{i}]}{\boldsymbol{\tau}}\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}} = \mathbf{A},$$

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* $\mathsf{I}_{\boldsymbol{\tau}}$*, existe otra matriz elemental* $\mathsf{I}_{\boldsymbol{\tau}^{-1}}$*, tal que*

$$\mathsf{I}_{\boldsymbol{\tau}}\mathsf{I}_{\boldsymbol{\tau}^{-1}} = \mathsf{I}_{\boldsymbol{\tau}^{-1}}\mathsf{I}_{\boldsymbol{\tau}} = \mathsf{I}.$$

En particular:

$$\boxed{\Big(\mathsf{I}_{\underset{[(-\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\Big)\Big(\mathsf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\Big) = \Big(\mathsf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\Big)\Big(\mathsf{I}_{\underset{[(-\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\Big) = \mathsf{I}} \qquad \text{y} \qquad \boxed{\Big(\mathsf{I}_{\underset{[(\frac{1}{\alpha})\boldsymbol{i}]}{\boldsymbol{\tau}}}\Big)\Big(\mathsf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\Big) = \Big(\mathsf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\Big)\Big(\mathsf{I}_{\underset{[(\frac{1}{\alpha})\boldsymbol{i}]}{\boldsymbol{\tau}}}\Big) = \mathsf{I}}.$$

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* A *es producto de matrices elementales, entonces existe* B *(también producto de matrices elementales) tal que* $\mathsf{A}\mathsf{B} = \mathsf{I} = \mathsf{B}\mathsf{A}$ *(es decir, si* $\mathsf{A} = \mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ *entonces es invertible).*

Pero hay que tener en cuenta que la última trasformación 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 trasformaciones elementales en determinado orden, se deshacen en el orden inverso —véase la Proposición 5.1.3.

☞ Cualquier matriz $\mathsf{E} = \mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ es invertible por ser producto de matrices elementales, en particular,

$$\big(\mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\big)\big(\mathsf{I}_{\boldsymbol{\tau}_k^{-1}\cdots\boldsymbol{\tau}_1^{-1}}\big) = \big(\mathsf{I}_{\boldsymbol{\tau}_1}\big)\cdots\big(\mathsf{I}_{\boldsymbol{\tau}_k}\big)\ \big(\mathsf{I}_{\boldsymbol{\tau}_k^{-1}}\big)\cdots\big(\mathsf{I}_{\boldsymbol{\tau}_1^{-1}}\big) = \mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k\,\boldsymbol{\tau}_k^{-1}\cdots\boldsymbol{\tau}_1^{-1}} = \mathsf{I};$$

por lo que podemos denotar $\boldsymbol{\tau}_k^{-1}\cdots\boldsymbol{\tau}_1^{-1}$ como $\big(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k\big)^{-1}$. Así pues,

$$\mathsf{E}^{-1} = \mathsf{I}_{\boldsymbol{\tau}_k^{-1}\cdots\boldsymbol{\tau}_1^{-1}} = \mathsf{I}_{(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{-1}}.$$

Librería NAcAL para Python

```python
Tr = T([ (3,2), (5,2,3), (10,1,2) ], 'h'); Tr # Secuencia de transformaciones
                                              # el parámetro 'h' es para mostrar
                                              # la secuencia en 'h'orizontal
```

$$\underset{[(3)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(5)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}\underset{[(10)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}$$

```python
Tr**(-1)                                      # Inversa de la secuencia de transf.
```

$$\underset{[(-10)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-5)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}\underset{[(\frac{1}{3})\mathbf{2}]}{\boldsymbol{\tau}}$$

```python
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 pre-escalonadas con inversa.

☞ Como primer paso para encontrar un algoritmo que calcule la inversa, demostraremos que una matriz K, de orden n, pre-escalonada y *sin columnas nulas*, se puede transformar en la matriz identidad I mediante una secuencia de transformaciones elementales

$$\mathsf{K}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathsf{K}\left(\mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \mathsf{I} \qquad \text{(Teorema 5.1.10).}$$

Consecuentemente, K^{-1} es el producto de las correspondientes matrices elementales; $\boxed{\mathsf{K}^{-1} = \mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}$.

Teorema 5.1.10. *Si una matriz* K, *de orden* n, *es pre-escalonada y sin columnas nulas, existen* k *trasformaciones elementales tales que* $\mathsf{K}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathsf{I}$.

Demostración. Como toda matriz pre-escalonada y *sin columnas nulas* tiene necesariamente pivotes en todas sus columnas, cualquier matriz *cuadrada*, pre-escalonada y sin columnas nulas, tiene pivotes en todas sus *filas* (pues cada pivote ocupa una fila y columna distintas, y todas sus columnas tienen pivote).

Por tanto, basta aplicar las k transformaciones $\boldsymbol{\tau}_1,\ldots\boldsymbol{\tau}_k$ de la eliminación Gauss-Jordan (Corolario 4.4.3 en la página 53) que transforman K en su forma escalonada reducida $\mathsf{K}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$; que en este caso resulta ser la matriz identidad de orden n (pues K, de orden n, tiene n pivotes). □

Librería NAcAL para Python

```
K = Matrix( [[0,5,0], [2,0,0], [6,9,3]] )
R = ElimGJ(K,1)                                  # Forma Escalonada Reducida
```

$$\begin{bmatrix}0&5&0\\2&0&0\\6&9&3\end{bmatrix} \xrightarrow[{[\mathbf{1}\rightleftharpoons\mathbf{2}]}]{\boldsymbol{\tau}} \begin{bmatrix}5&0&0\\0&2&0\\9&6&3\end{bmatrix} \xrightarrow[{\substack{[(-3)\mathbf{3}+\mathbf{1}]\\[(-2)\mathbf{3}+\mathbf{2}]}}]{\boldsymbol{\tau}} \begin{bmatrix}5&0&0\\0&2&0\\0&0&3\end{bmatrix} \xrightarrow[{\substack{[(\frac{1}{5})\mathbf{1}]\\[(\frac{1}{2})\mathbf{2}]\\[(\frac{1}{3})\mathbf{3}]}}]{\boldsymbol{\tau}} \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

Corolario 5.1.11. *Si* K *es pre-escalonada y cuadrada, las siguientes propiedades son equivalentes*

1. K *no tiene columnas nulas*
2. K *es un producto de matrices elementales*
3. K *tiene inversa*

Demostración.

1. ⇒ *2.* Si K no tiene columnas nulas, existen transf. elementales $\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k$ tales que $\mathsf{K}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathsf{I}$; así que

$$\begin{aligned}\mathsf{I} &= \mathsf{K}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathsf{K}\left(\mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) \quad\Rightarrow\\ \mathsf{I}\left(\mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)^{-1} &= \mathsf{K}\left(\mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)\left(\mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)^{-1} \quad\Rightarrow\\ \left(\mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)^{-1} &= \mathsf{K}\end{aligned}$$

donde $\left(\mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)^{-1} = \mathsf{I}_{(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{-1}} = \mathsf{I}_{\boldsymbol{\tau}_k^{-1}\cdots\boldsymbol{\tau}_1^{-1}}$ 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 58.

□

Librería NAcAL para Python

```
(R.TrC)**-1 # inversa de la secuencia de Transf. de las
            # Columnas en el proceso de eliminación K --> R
```

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

(compare con la salida del bloque de código anterior y fíjese cómo la primera transformación inversa deshace la última del bloque anterior, la segunda deshace la penúltima, etc.)

Librería NAcAL para Python

```
I(K.n) & ( R.TrC )**-1   # al aplicar la inversa de las
                         #Transf. 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 4.4.1 sabemos que existen $\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k$ tales que $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}=\mathbf{K}$ es pre-escalonada; y si $\mathbf{K}$ no tiene columnas nulas, entonces es producto de p matrices elementales, $\mathbf{K}=\mathbf{I}_{\boldsymbol{\tau}'_1\cdots\boldsymbol{\tau}'_p}$:

$$\begin{aligned}\mathbf{A}\big(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\big) &= \big(\mathbf{I}_{\boldsymbol{\tau}'_1\cdots\boldsymbol{\tau}'_p}\big)\\ \mathbf{A}\big(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\big)\big(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\big)^{-1} &= \big(\mathbf{I}_{\boldsymbol{\tau}'_1\cdots\boldsymbol{\tau}'_p}\big)\big(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\big)^{-1}\\ \mathbf{A} &= \big(\mathbf{I}_{(\boldsymbol{\tau}'_1\cdots\boldsymbol{\tau}'_p)}\big)\big(\mathbf{I}_{(\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k)^{-1}}\big) \;=\; \mathbf{I}_{\boldsymbol{\tau}'_1\cdots\boldsymbol{\tau}'_p\boldsymbol{\tau}_k^{-1}\cdots\boldsymbol{\tau}_1^{-1}},\end{aligned}$$

2. $\Rightarrow$ *3.* ¡Ya se sabe! Proposición 5.1.9 en la página 59

3. $\Rightarrow$ *1.* Sean $\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k$ tales que $\mathbf{A}\big(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\big)=\mathbf{K}$ es pre-escalonada. Entonces, puesto que $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ es invertible, $\mathbf{A}$ es invertible si y solo si $\mathbf{K}$ es invertible (Proposición 5.1.4 en la página 57).

□

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.

Librería NAcAL para Python

```python
A  = Matrix( [ [2,-1,0], [-1,2,-1], [0,-1,2] ] )
K  = Elim(A,1)                # una forma pre-escalonada de A
```

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

(Como $\mathbf{A}$ es cuadrada y la forma pre-escalonada $\mathbf{K}$ no tiene columnas nulas, $\mathbf{A}$ es invertible.)

Podemos ver las transformaciones que hemos aplicado para llegar a la forma pre-escalonada $\mathbf{K}$ con:

Librería NAcAL para Python

```python
Tr = K.TrC                    # Transformaciones aplicadas a las columnas
T(Tr, 'h')                    # Las representaremos en 'h'orizontal
```

$$\underset{[(2)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(3)\mathbf{3}]}{\boldsymbol{\tau}}\underset{[(1)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}$$

Aplicado la inversa de dichas transformaciones sobre $\mathbf{K}$ recuperamos la matriz original $\mathbf{A}$:

Librería NAcAL para Python

```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:

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

☞ La siguiente proposición nos ofrece una nueva característica de las matrices NO invertibles:

- Si $\mathbf{A}$ tiene columnas que son combinación lineal del resto, entonces es singular.

EJERCICIO 30. Demuestre la siguiente proposición.

Proposición 5.1.13. *Sea* $\mathbf{A}$ *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* $\mathbf{A}$ *es* singular*.*

☞ Y ahora veamos que si $\mathbf{A}$ y $\mathbf{B}$ son de orden n y $\mathbf{AB} = \mathbf{I}$, entonces $\mathbf{A}$ y $\mathbf{B}$ 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 $\mathbf{B}^{\intercal}\mathbf{A}^{\intercal} = \left(\mathbf{AB}\right)^{\intercal}$ 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:

$$\mathbf{A}^{-1}\left(\mathbf{AB}\right) = \mathbf{A}^{-1}\mathbf{I} \quad \Rightarrow \quad \mathbf{B} = \mathbf{A}^{-1}.$$

□

☞ ¡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 4.4.1 en la página 51)

$$\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p} = \mathbf{K},$$

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 60)

$$\mathbf{K}_{\boldsymbol{\tau}_{(p+1)}\cdots\boldsymbol{\tau}_k} = \mathbf{I}.$$

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 $\mathbf{I}$

$$\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \mathbf{I};$$

es decir, si $\mathbf{K}$ no tiene columnas nulas, entonces $\boxed{\mathbf{A}^{-1} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}$; y si las tiene entonces $\mathbf{A}^{-1}$ no existe.

Librería 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
```

$$\begin{bmatrix}a&b\\c&d\end{bmatrix}\xrightarrow{\underset{[(-\frac{b}{a})\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}\begin{bmatrix}a&0\\c&d-\frac{bc}{a}\end{bmatrix}\xrightarrow{\underset{\left[\left(-\frac{ac}{ad-bc}\right)\mathbf{2}+\mathbf{1}\right]}{\boldsymbol{\tau}}}\begin{bmatrix}a&0\\0&d-\frac{bc}{a}\end{bmatrix}\xrightarrow{\underset{\left[\left(\frac{a}{ad-bc}\right)\mathbf{2}\right]}{\overset{\boldsymbol{\tau}}{[(\frac{1}{a})\mathbf{1}]}}}\begin{bmatrix}1&0\\0&1\end{bmatrix}$$

(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

$$\begin{bmatrix}1&0\\0&1\end{bmatrix}_{\underset{[(-\frac{b}{a})\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-\frac{ac}{ad-bc})\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}\underset{[(\frac{1}{a})\mathbf{1}]}{\boldsymbol{\tau}}\underset{[(\frac{a}{ad-bc})\mathbf{2}]}{\boldsymbol{\tau}}}=\begin{bmatrix}\frac{d}{ad-bc}&-\frac{b}{ad-bc}\\-\frac{c}{ad-bc}&\frac{a}{ad-bc}\end{bmatrix}$$

Librería 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 $\mathbf{I}_{\boldsymbol{\tau}_i}$ puede ser bastante pesado.

¿Hay alguna una forma de realizar ambas operaciones a la vez: la transformación $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}=\mathbf{I}$ y la transformación $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}=\mathbf{A}^{-1}$?

!Si! y es sencillísimo... basta "alargar" las columnas de $\mathbf{A}$ con $\mathbf{I}$ y entonces aplicar $\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k$ a $\begin{bmatrix}\mathbf{A}\\\hline\mathbf{I}\end{bmatrix}$.

5.2. Secuencias de transformaciones elementales. Parte II

Concatenando $\mathbf{A}$ y $\mathbf{B}$ (ambas con el mismo número de columnas), es decir, alargando las columnas de $\mathbf{A}$ poniendo las de $\mathbf{B}$ por debajo $\begin{bmatrix}\mathbf{A}\\\hline\mathbf{B}\end{bmatrix}$; 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 concatenar luego las matrices resultantes:

$$\begin{bmatrix}\mathbf{A}\\\hline\mathbf{B}\end{bmatrix}\xrightarrow{\boldsymbol{\tau}}\begin{bmatrix}\mathbf{A}\\\hline\mathbf{B}\end{bmatrix}_{\boldsymbol{\tau}}=\begin{bmatrix}\mathbf{A}_{\boldsymbol{\tau}}\\\hline\mathbf{B}_{\boldsymbol{\tau}}\end{bmatrix},$$

de manera que la columna $j$ésima $\begin{bmatrix}\mathbf{A}_{\boldsymbol{\tau}}\\\hline\mathbf{B}_{\boldsymbol{\tau}}\end{bmatrix}_{|j}$ es la concatenación de las columnas $\left(\mathbf{A}_{\boldsymbol{\tau}}\right)_{|j}$ y $\left(\mathbf{B}_{\boldsymbol{\tau}}\right)_{|j}$.

Lo mismo ocurre si aplicamos una secuencia de k transformaciones elementales:

$$\begin{bmatrix}\mathbf{A}\\\hline\mathbf{B}\end{bmatrix}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}=\begin{bmatrix}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\\\hline\mathbf{B}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\end{bmatrix}.$$

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:

$$\begin{bmatrix}\mathbf{A}\\\hline\mathbf{I}\end{bmatrix}=\left[\begin{array}{ccc}0&1&5\\7&0&4\\\hline1&0&0\\0&1&0\\0&0&1\end{array}\right]\xrightarrow[{[(3)\mathbf{2}]}]{\boldsymbol{\tau}}\left[\begin{array}{ccc}0&3&5\\7&0&4\\\hline1&0&0\\0&3&0\\0&0&1\end{array}\right]\xrightarrow[{[(2)\mathbf{1}+\mathbf{2}]}]{\boldsymbol{\tau}}\left[\begin{array}{ccc}0&3&5\\7&14&4\\\hline1&2&0\\0&3&0\\0&0&1\end{array}\right]=\begin{bmatrix}\mathbf{A}_{\underset{[(3)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(2)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}\\\hline\mathbf{I}_{\underset{[(3)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(2)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}\end{bmatrix},$$

entonces sin necesidad de multiplicar sabemos por una parte que

$$\mathbf{I}_{\underset{[(3)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(2)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}=\begin{bmatrix}1&2&0\\0&3&0\\0&0&1\end{bmatrix};\qquad\text{y por la otra que}\qquad\mathbf{A}_{\underset{[(3)\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(2)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}=\begin{bmatrix}0&3&5\\7&14&4\end{bmatrix}.$$

Teniendo en cuenta lo anterior, podemos calcular fácilmente la inversa de $\mathbf{A}$: basta con "pegar" la matriz identidad debajo y tratar de transformar la submatriz superior (la matriz $\mathbf{A}$) en la matriz identidad $\mathbf{I}$ mediante transformaciones elementales.

Ejemplo 6.

$$\left[\begin{array}{c}\mathbf{A}\\ \hline \mathbf{I}\end{array}\right] = \left[\begin{array}{cc} 5 & 0\\ 2 & -2\\ \hline 1 & 0\\ 0 & 1\end{array}\right] \xrightarrow[{[(1)\mathbf{2}+1]}]{\boldsymbol{\tau}} \left[\begin{array}{cc} 5 & 0\\ 0 & -2\\ \hline 1 & 0\\ 1 & 1\end{array}\right] \xrightarrow[{\left[\left(\frac{-1}{2}\right)\mathbf{2}\right]}]{\substack{\boldsymbol{\tau}\\ [(\frac{1}{5})\mathbf{1}]}} \left[\begin{array}{cc} 1 & 0\\ 0 & 1\\ \hline 1/5 & 0\\ 1/5 & -1/2\end{array}\right] = \left[\begin{array}{c}\mathbf{A}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2\boldsymbol{\tau}_3}\\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2\boldsymbol{\tau}_3}\end{array}\right].$$

Puesto que $\mathbf{A}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2\boldsymbol{\tau}_3} = \mathbf{I}$, entonces $\mathbf{I}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2\boldsymbol{\tau}_3} = \begin{bmatrix}1/5 & 0\\ 1/5 & -1/2\end{bmatrix} = \mathbf{A}^{-1}$.

☞

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 $\left[\begin{array}{c}\mathbf{R}\\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\end{array}\right]$ la submatriz $\mathbf{R}$ tiene columnas nulas, entonces $\mathbf{A}$ no tiene inversa.
3. En caso contrario, $\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ es la inversa de $\mathbf{A}$.

Librería NAcAL para Python

```python
A = Matrix([ [1, -1, -3], [0, 1, 0], [0, -1, 1] ])
InvMat(A,1)          # Calcula la inversa por eliminación Gauss-Jordan
```

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

5.3. Inversa de una matriz triangular

Definición 5.4. *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* $i \leq j$.

Por ejemplo las matrices

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0\\ a_{21} & a_{22} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mm}\end{bmatrix}; \quad \mathbf{B} = \begin{bmatrix} b_{11} & 0 & \cdots & 0 & 0\\ b_{21} & b_{22} & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ b_{m1} & b_{m2} & \cdots & b_{mp} & 0\end{bmatrix}; \quad \text{y} \quad \mathbf{C} = \begin{bmatrix} c_{11} & 0 & \cdots & 0\\ c_{21} & c_{22} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ \vdots & \vdots & \cdots & c_{(m-1)n}\\ c_{m1} & c_{m2} & \cdots & c_{mn}\end{bmatrix}.$$

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... y también es triangular superior...

Definición 5.5. *Decimos que una matriz* $\mathbf{U}$ *es triangular superior cuando todos los componentes por debajo la diagonal principal son nulos, es decir, si* $u_{ij} = 0$ *cuando* $i \geq j$*.*

Por tanto, una matriz es triangular superior si su transpuesta es triangular inferior.

Las matrices *diagonales* son simultáneamente triangulares superiores e inferiores.

☞ Veamos ahora 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 5.3.1. *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 seria 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) □

Consecuentemente, una matriz triangular inferior e invertible está necesariamente escalonada, pues los elementos en la diagonal son distintos de cero. Así, para encontrar la inversa solo necesitamos aplicar las dos últimas etapas de la eliminación Gauss-Jordan: *la eliminación "de derecha a izquierda"* para anular (en cada fila) los componentes situados a la izquierda de la diagonal, y *normalizar los pivotes* dividiendo cada columna por el valor de su pivote (fíjese en el Ejemplo 6 en la página anterior).

$$\left[\frac{\mathbf{L}}{\mathbf{I}}\right] \xrightarrow{\boldsymbol{\tau}_1,\ldots,\boldsymbol{\tau}_n} \left[\frac{\mathbf{I}}{\mathbf{L}^{-1}}\right].$$

Como las transformaciones aplicadas en la eliminación "de derecha a izquierda" (y en la normalización) no modifican las componentes nulas a la derecha (y por encima) de la diagonal principal de la matriz identidad, llegamos a la siguiente conclusión:

Proposición 5.3.2. *La inversa de una matriz triangular inferior es triangular inferior.*

Además, puesto que la inversa de transpuesta es la transpuesta de la inversa, tenemos el siguiente corolario:

Corolario 5.3.3. *La inversa de una matriz triangular superior es triangular superior.*

☞ Vimos (Corolario 5.1.12) que si una matriz es invertible, ninguna de sus formas pre-escalonadas tiene columnas nulas. Una generalización de este resultado da lugar a la definición de rango que usaremos aquí.

5.4. Rango de una matriz

Puesto que una forma escalonada es tan solo una reordenación de las columnas de una pre-escalonada (Corolario 4.4.2), y como es mucho más sencillo visualizar la estructura de una forma escalonada; el siguiente razonamiento lo realizaremos usando matrices escalonadas sin pérdida de generalidad.

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 dos matrices invertibles tales que $\mathbf{AE} = \mathbf{L}$ y $\mathbf{AE'} = \mathbf{L'}$ son escalonadas. Veamos que las posiciones de los pivotes de $\mathbf{L}$ y $\mathbf{L'}$ son coincidentes.

Por simetría basta comprobar que la posición de pivote de cualquier columna no nula de $\mathbf{L'}$ coincide con la posición de pivote de alguna columna no nula de $\mathbf{L}$[1]. Para ello tendremos en cuenta que como

$$\mathbf{L'} = \mathbf{AE'} = \mathbf{LE}^{-1}\mathbf{E'}, \qquad \text{pues } \mathbf{A} = \mathbf{LE}^{-1},$$

por tanto las columnas de $\mathbf{L'} = \mathbf{L}\big(\mathbf{E}^{-1}\mathbf{E'}\big)$ son combinación lineal de las de $\mathbf{L}$. Además, por ser $\mathbf{L}$ escalonada, la posición del pivote de cualquier combinación lineal de columnas de $\mathbf{L}$ coincide con la posición de pivote de la primera columna que esté multiplicada por un número no nulo. Por ejemplo, si la primera componente no nula del vector está en la tercera posición, obtenemos un vector cuyo pivote tiene la misma posición que el pivote de la tercera columna (es decir, quinta fila):

$$\begin{bmatrix} * & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & * & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & * & & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & * & & 0 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ \star \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} = 0 \begin{pmatrix} * \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} + 0 \begin{pmatrix} 0 \\ * \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} + \star \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ * \\ \vdots \\ \vdots \end{pmatrix} + \cdots = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \star * \\ \vdots \\ \vdots \end{pmatrix}$$

donde "$*$" son pivotes y "$\star \neq 0$".

Así pues, la posición del pivote de cualquier combinación lineal de columnas de $\mathbf{L}$ coincide con la posición del pivote de alguna de columna de $\mathbf{L}$... y por tanto la posición de pivote de cada columna de $\mathbf{L'}$ (que es una combinación de las columnas de $\mathbf{L}$) coincide con la posición de pivote de alguna columna $\mathbf{L}$; y lo mismo se puede decir de las columnas de $\mathbf{L}$ respecto de las de $\mathbf{L'}$, pues $\mathbf{L} = \mathbf{L'}\big((\mathbf{E'})^{-1}\,\mathbf{E}\big)$. Concluimos que como las posiciones de los pivotes de $\mathbf{L}$ y $\mathbf{L'}$ son coincidentes, el numero de columnas no nulas de $\mathbf{L}$ coincide con el número de columnas no nulas de $\mathbf{L'}$. Este resultado nos permite dar la siguiente definición:

Definición 5.6 (Rango de una matriz)**.** *El rango de $\mathbf{A}$ es el número de columnas no nulas de cualquiera de sus formas pre-escalonadas, es decir, el número de pivotes de cualquiera de sus formas pre-escalonadas.*

5.4.1. Algunas propiedades del rango de una matriz

Proposición 5.4.1. *Si $\mathbf{E}$ es invertible, entonces* $\operatorname{rango}(\mathbf{A}) = \operatorname{rango}(\mathbf{AE})$.

Demostración. Con ambas matrices se pueden obtener idéntica forma pre-escalonada $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$:

$$\text{Sea } \ \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}'_1\cdots\boldsymbol{\tau}'_p}, \ \text{ entonces } \ \mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \big(\mathbf{AEE}^{-1}\big)_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \big(\mathbf{AE}\big)_{(\boldsymbol{\tau}'_1\cdots\boldsymbol{\tau}'_p)^{-1}\,\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

□

Proposición 5.4.2. *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

[1] por "simetría" me refiero a que siguiendo el mismo razonamiento también se llega a que la posición del pivote de cualquier columna no nula de $\mathbf{L}$ coincide con la posición del pivote el de alguna columna no nula de $\mathbf{L'}$

- Aplicando la proposición anterior tenemos que en particular: $\text{rango}\,(\mathbf{AB}) = \text{rango}\,\big((\mathbf{AB})\mathbf{E}\big)$.
- 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

$$\text{rango}\,(\mathbf{AB}) = \text{rango}\,(\mathbf{ABE}) = \text{rango}\,(\mathbf{AL}) \;\leq\; \text{rango}\,(\mathbf{L}) = \text{rango}\,(\mathbf{B}).$$

□

Proposición 5.4.3. *Si* $\mathbf{E}$ *es invertible, entonces* $\text{rango}\,(\mathbf{EA}) = \text{rango}\,(\mathbf{A})$.

Demostración. La proposición anterior nos indica que $\text{rango}\,(\mathbf{EA}) \leq \text{rango}\,(\mathbf{A})$, pero también nos indica que $\text{rango}\,\big(\mathbf{E}^{-1}(\mathbf{EA})\big) \leq \text{rango}\,(\mathbf{EA})$; por tanto $\text{rango}\,(\mathbf{A}) \leq \text{rango}\,(\mathbf{EA})$. Por tanto, $\text{rango}\,(\mathbf{EA}) = \text{rango}\,(\mathbf{A})$. □

Aplicando eliminación *gaussina* por filas 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 $\text{rango}\,(\mathbf{S}) = \text{rango}\,(\mathbf{S}^\intercal)$:

Librería 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
```

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix} \xrightarrow[{[(1)\mathbf{1}+\mathbf{2}]}]{\boldsymbol{\tau}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix} \xrightarrow[{[(-2)\mathbf{4}+\mathbf{5}]}]{\boldsymbol{\tau}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow[{\substack{[\mathbf{3}\rightleftharpoons\mathbf{4}] \\ [\mathbf{2}\rightleftharpoons\mathbf{3}]}}]{\boldsymbol{\tau}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathbf{S}$$

$$\mathbf{S}^\intercal = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Vamos a aprovechar este hecho para demostrar la siguiente proposición.

Proposición 5.4.4. *Si* $\mathbf{R}$ *es escalonada reducida, entonces* $\text{rango}\,(\mathbf{R}) = \text{rango}\,(\mathbf{R}^\intercal)$.

Demostración. Puesto que podemos encontrar una matriz invertible $\mathbf{E}$ tal que $\mathbf{ER} = \mathbf{S}$ es diagonal. Así, por las proposiciones 5.4.3 y 5.4.1, y puesto que $\mathbf{E}^\intercal$ también es invertible:

$$\text{rango}\,(\mathbf{R}) = \text{rango}\,(\mathbf{ER}) = \text{rango}\,(\mathbf{S}) = \text{rango}\,(\mathbf{S}^\intercal) = \text{rango}\,(\mathbf{R}^\intercal\mathbf{E}^\intercal) = \text{rango}\,(\mathbf{R}^\intercal).$$

□

Proposición 5.4.5. $\text{rango}\,(\mathbf{A}) = \text{rango}\,(\mathbf{A}^\intercal)$.

Demostración. Sea $\mathbf{E}$ invertible tal que $\mathbf{AE} = \mathbf{R}$ es escalonada reducida, entonces (aplicando la Proposición 5.4.3 en la última igualdad) tenemos que: $\operatorname{rango}(\mathbf{A}) = \operatorname{rango}(\mathbf{AE}) = \operatorname{rango}(\mathbf{R}) = \operatorname{rango}(\mathbf{R}^\intercal) = \operatorname{rango}(\mathbf{E}^\intercal\mathbf{A}^\intercal) = \operatorname{rango}(\mathbf{A}^\intercal)$. □

Y por último

Proposición 5.4.6. *El rango de* $\mathbf{AB}$ *es menor o igual que el de* $\mathbf{A}$.

Demostración. $\operatorname{rango}(\mathbf{AB}) = \operatorname{rango}(\mathbf{B}^\intercal\mathbf{A}^\intercal) \leq \operatorname{rango}(\mathbf{A}^\intercal) = \operatorname{rango}(\mathbf{A})$. □

Librería NAcAL para Python

```
Elim(A).rango    # rango es una 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 pre-escalonadas 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).

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 usaremos a partir de ahora y, de paso, podremos recapitular lo visto hasta aquí empleando la nueva terminología abstracta.

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 suespacios: mediantes 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

Para cualesquiera *números reales* a, b y c (y empleando la suma y producto habituales) se verifica que:

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. $1a = a$.

También sabemos que si definimos la suma y el producto por escalares para $\mathbb{R}^n$ como:

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

se verifican las siguientes propiedades (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. $(\lambda + \eta)\boldsymbol{a} = \lambda\boldsymbol{a} + \eta\boldsymbol{a}$
7. $\lambda(\eta\boldsymbol{a}) = (\lambda\eta)\boldsymbol{a}$
8. $1\boldsymbol{a} = \boldsymbol{a}$

Y si definimos la suma y el producto por escalares para $\mathbb{R}^{m\times n}$ (*matrices de orden m por n*) como

$$\big(\mathbf{A}+\mathbf{B}\big)_{|j} = \mathbf{A}_{|j} + \mathbf{B}_{|j} \qquad \text{y} \qquad \big(\lambda\mathbf{A}\big)_{|j} = \lambda\big(\mathbf{A}_{|j}\big),$$

se verifican las siguientes propiedades (Proposición 1.5.1 en la página 19):

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. $(\lambda + \eta)\mathbf{A} = \lambda\mathbf{A} + \eta\mathbf{A}$
7. $\lambda(\eta\mathbf{A}) = (\lambda\eta)\mathbf{A}$
8. $1\mathbf{A} = \mathbf{A}$

A primera vista es evidente que subyace una estructura común a $\mathbb{R}$, a $\mathbb{R}^n$ (con la suma y producto por escalares correspondiente) y al conjunto de matrices $\mathbb{R}^{m\times m}$ (y sus correspondientes operaciones de suma de matrices y producto por escalares). Esta 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:* $\vec{x}$, $\vec{y}$, $\vec{z}$,...

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 $\vec{x}+\vec{y}$. Esto se representa esquemáticamente con[2]

  $$\begin{aligned} \_+\_ \colon \mathcal{V}\times\mathcal{V} &\to \mathcal{V} \\ (\vec{x},\vec{y}) &\to \vec{x}+\vec{y} \end{aligned}$$

  (Puesto que sumando vectores de $\mathcal{V}$ se obtienen elementos $\mathcal{V}$, 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,*[3]*, 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 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 $a\cdot\vec{x}$. Esto se representa esquemáticamente con

  $$\begin{aligned} \_\cdot\_ \colon \mathbb{R}\times\mathcal{V} &\to \mathcal{V} \\ (a,\vec{y}) &\to a\cdot\vec{x} \end{aligned}$$

  (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. $(a\cdot b)\vec{x}=a(b\vec{x})$
  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 a cada conjunto) son espacios vectoriales.

Ejercicio 31. Demuestre que
(a) si $\vec{z}+\vec{x}=\vec{x}$ para cualquier $\vec{x}$, entonces necesariamente $\vec{z}=\vec{0}$.
(b) si $\vec{z}+\vec{x}=\vec{0}$ entonces $\vec{z}=-\vec{x}$.

[1] objetos del mismo tipo: o números, o vectores, o matrices, o funciones, o polinomios, o variables aleatorias, etc.
[2] Véase el Apéndice 6.3.4 en la página 81 de esta lección.
[3] y por tanto $\mathcal{V}\neq\emptyset$

Por tanto, en todo espacio vectorial los vectores $\vec{0}$ y, para cada $\vec{x}$, el vector opuesto son únicos.

Notación. En el caso particular de $\mathbb{R}^n$ usamos los símbolos $\boldsymbol{a}$, $\boldsymbol{x}$, $\mathbf{0}$, etc. para denotar las *listas ordenadas de números reales* (i.e., los vectores de $\mathbb{R}^n$); y en el caso de $\mathbb{R}^{m\times n}$ usamos $\mathbf{A}$, $\mathbf{X}$, $\mathbf{0}$, etc. para las matrices (i.e., los vectores de $\mathbb{R}^{m\times n}$). Pero en general denotamos los vectores de un espacio vectorial genérico $\mathcal{V}$ con $\vec{a}$, $\vec{x}$, $\vec{0}$, etc.

Para denotar espacios vectoriales genéricos usaremos caracteres como: $\mathcal{C}$, $\mathcal{N}$, $\mathcal{X}$, $\mathcal{V}$, $\mathcal{W}$...

☞ Fíjese que la definición de espacio vectorial es abstracta, pues nada se dice sobre la naturaleza de los objetos que constituyen el conjunto $\mathcal{V}$, como tampoco sobre la definición de las operaciones *suma de vectores* y *producto de un escalar por un vector*; tan solo se enumeran las propiedades que deben verificar las 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.
(véase el ejercicio a continuación del recuadro)

Ejercicio 32. 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 por el vector $\vec{x}$ es x^c, es decir,

$$\begin{array}{rl} \_+\_ \colon \mathbb{R}^+\times\mathbb{R}^+ \to \mathbb{R} \\ (\vec{x},\vec{y}) \to xy \end{array} \qquad \text{y} \qquad \begin{array}{rl} \_\cdot\_ \colon \mathbb{R}\times\mathbb{R}^+ \to \mathbb{R} \\ (c,\vec{x}) \to x^c \end{array}$$

Demuestre que $\mathbb{R}^+$ junto con las operaciones indicadas es un espacio vectorial. ¿Quien 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 más arriba: para el vector nulo $\vec{0}\in\mathcal{V}$ y cualquier escalar a:

$$a\vec{0} = a(\vec{0}+\vec{0}) = a\vec{0}+a\vec{0}.$$

Restando $-(a\vec{0})$ a ambos lados deducimos que $\vec{0}=a\vec{0}$. Y para el escalar 0 y cualquier vector $\vec{x}\in\mathcal{V}$:

$$0\vec{x} = (0+0)\vec{x} = 0\vec{x}+0\vec{x}$$

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 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 $\vec{0}$, pues $\vec{0}$ por cualquier escalar es $\vec{0}$. Así llegamos a que si $a\neq 0$ entonces $\vec{v}=\vec{0}$.

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. $\mathbf{0} = 0\mathsf{A}$ para $\mathsf{A}\in\mathbb{R}^{m\times n}$ y $0\in\mathbb{R}$.
3. $(-1)\mathsf{A} = -\mathsf{A}$ para todo $\mathsf{A}\in\mathbb{R}^{m\times n}$.
4. Si $a\mathsf{A} = \mathbf{0}$ o bien $a = 0$ o bien $\mathsf{A} = \mathbf{0}$.

☞ 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 para la suma y producto para 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 24 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* $\left[\vec{v}_1;\ \vec{v}_2;\ \dots\ \vec{v}_n\right]$ *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 muchos conjuntos que (con la definición de suma y producto por escalares adecuada) poseen estructura de espacio vectorial y con los que usted se encontrará en otras asignaturas (véase la Sección 6.3.3).

6.2. Funciones lineales

☞ El segundo término abstracto de la lección son las funciones lineales…

A lo largo de estas lecciones hemos indicado en numerosas ocasiones que ciertos operadores eran *lineales*.[4] En cada una de esas ocasiones lo que hemos visto es un ejemplo de función lineal.[5]

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 7. Las definiciones de suma y producto por escalares dadas en la Sección 1.2 en la página 6 convierten a la función $f(\boldsymbol{x}) = \boldsymbol{x}_{|i}$ que selecciona la componente $i$ésima de un vector de $\mathbb{R}^n$ en una función lineal que

[4]Consideraremos como sinónimos los términos: función lineal, aplicación lineal, transformación lineal u operador lineal.

[5]Se supone que usted conoce las funciones, pero dispone de un Apéndice 6.3.4 en la página 81 con todo lo necesario.

podemos describir con el siguiente esquema

$$\begin{aligned} \_|_i : \mathbb{R}^n &\to \mathbb{R} \\ \boldsymbol{x} &\to x_i \end{aligned}$$

Y sabemos que la función que selecciona la $i$ésima componente $\boldsymbol{x} \to \boldsymbol{x}_{|i}$ 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 33. Describa con un esquema (al modo de 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: $\mathsf{A} \to \mathsf{A}_{|j}$ (Página 19)
(b) La selección de la componente de la fila $i$ésima y columna $j$ésima de una matriz: $\mathsf{A} \to {}_{i|}\mathsf{A}_{|j}$ (Página 20)
(c) La transposición de una matriz: $\mathsf{A} \to \mathsf{A}^{\mathsf{T}}$ (Página 20)
(d) La selección de la fila $i$ésima de una matriz: $\mathsf{A} \to {}_{i|}\mathsf{A}$ (Página 20)
(e) El producto de una matriz por el vector $\boldsymbol{b}$ por su derecha: $\mathsf{A} \to \mathsf{A}\boldsymbol{b}$ (Página 26)
(f) El producto de un vector por una matriz A a su izquierda: $\boldsymbol{b} \to \mathsf{A}\boldsymbol{b}$ (Página 24)
(g) El producto de una matriz por el vector $\boldsymbol{a}$ por su izquierda: $\mathsf{B} \to \boldsymbol{a}\mathsf{B}$ (Página 27)
(h) El producto de un vector por una matriz B a su derecha: $\boldsymbol{a} \to \boldsymbol{a}\mathsf{B}$ (Página 28)
(i) El producto de una matriz por una matriz B a su derecha: $\mathsf{A} \to \mathsf{AB}$ (Página 31)
(j) El producto de una matriz por una matriz A a su izquierda: $\mathsf{B} \to \mathsf{AB}$ (Página 31)
(k) Una transformación elemental $\boldsymbol{\tau}$ de las columnas de una matriz: $\mathsf{A} \to \mathsf{A}_{\boldsymbol{\tau}}$ (Página 46)
(l) Una transformación elemental $\boldsymbol{\tau}$ de las filas de una matriz: $\mathsf{A} \to {}_{\boldsymbol{\tau}}\mathsf{A}$ (Página 54)

☞ ¡Fíjese que podríamos haber titulado *“Ejemplos de funciones lineales”* a las cuatro primeras lecciones!

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 34. Demuestre las siguientes proposiciones (ambas demostraciones son muy similares):

(a) **Proposición 6.2.1.** *La composición* $g \circ f$ *de dos funciones lineales* $f\colon \mathcal{D} \to \mathcal{V}$ *y* $g\colon \mathcal{V} \to \mathcal{W}$ *es una función lineal.*
(b) **Proposición 6.2.2.** *Si* $f\colon \mathcal{D} \to \mathcal{V}$ *es lineal e invertible, su inversa* $f^{-1}\colon \mathcal{V} \to \mathcal{D}$ *también es lineal.*

¡En las primeras lecciones también 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 $f(\mathsf{A}) = {}_{i|}\mathsf{A}_{|j}$, es decir

  $${}_{i|}\_|_j : \mathbb{R}^{m\times n} \to \mathbb{R}$$

  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.

  $${}_{i|}\_|_j : \mathbb{R}^{m\times n} \xrightarrow{{}_{i|}\_} \mathbb{R}^n \xrightarrow{\_|_j} \mathbb{R}$$

  (fíjese que del mismo modo 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 la 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

$$\left[f_{\mathbf{A}} \circ f_{\mathbf{B}}\right](\boldsymbol{x}) \;=\; f_{\mathbf{A}}\big(f_{\mathbf{B}}(\boldsymbol{x})\big) \;=\; f_{\mathbf{A}}\big(\mathbf{B}\boldsymbol{x}\big) \;=\; \mathbf{AB}\boldsymbol{x} \;=\; f_{\mathbf{AB}}(\boldsymbol{x}).$$

  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 se multiplica por la izquierda: si $g_{\mathbf{A}}(\boldsymbol{x}) = \boldsymbol{x}\mathbf{A}$ y $g_{\mathbf{B}}(\boldsymbol{x}) = \boldsymbol{x}\mathbf{B}$, entonces

$$\left[g_{\mathbf{A}} \circ g_{\mathbf{B}}\right](\boldsymbol{x}) \;=\; g_{\mathbf{A}}\big(g_{\mathbf{B}}(\boldsymbol{x})\big) \;=\; g_{\mathbf{A}}\big(\boldsymbol{x}\mathbf{B}\big) \;=\; \boldsymbol{x}\mathbf{BA} \;=\; g_{\mathbf{BA}}(\boldsymbol{x}).$$

  Otro ejemplo es el producto de varias matrices: $\mathbf{ABX}$.

- 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 una sucesión de transformaciones de las filas... o también 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).

Como ejemplos de funciones lineales invertibles podemos indicar

- Las transformaciones elementales.
- Si $\mathbf{A}$ es invertible, entonces $f_{\mathbf{A}}$ es invertible; en particular $\big(f_{\mathbf{A}}\big)^{-1} = f_{\mathbf{A}^{-1}}$.

  Explicación: por una parte $f_{\mathbf{A}}\big(f_{\mathbf{A}^{-1}}(\boldsymbol{x})\big) = \mathbf{AA}^{-1}\boldsymbol{x} = \boldsymbol{x}$ y por otra $f_{\mathbf{A}^{-1}}\big(f_{\mathbf{A}}(\boldsymbol{y})\big) = \mathbf{A}^{-1}\mathbf{A}\boldsymbol{y} = \boldsymbol{y}$. (Proposición 6.A.1 en la página 83 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 $\mathcal{V}$ (y con las mismas operaciones de $\mathcal{V}$) son espacios vectoriales por si 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 de por escalares. Es decir, si para todo $\vec{x}$ e $\vec{x}$ en $\mathcal{W}$ y todo escalar a,*

1. $\vec{x} + \vec{y}$ esta en el subespacio $\mathcal{W}$

2. $a\vec{x}$ esta 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 $\vec{x}$ en $\mathcal{W}$ y para todo $a, b \in \mathbb{R}$

$$a\vec{x} + b\vec{y} \quad \text{esta en el subespacio } \mathcal{W}$$

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 de cualquier para de vectores* es equivalente a decir que el conjunto es cerrado para la suma y el producto por escalares.

EJERCICIO 35. ¿Es subespacio el conjunto de combinaciones lineales de n vectores de un espacio vectorial?

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 está en W, o si está $\vec{x}$ 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.

- Polinomios ($\mathbb{R}[x]$).
- Variables aleatorias con esperanza cero: su suma tiene esperanza cero y cualquier múltiplo también.
- Variables aleatorias con distribución normal (gaussiana).

☞ Nótese que como todo conjunto está contenido en si mismo, todo espacio vectorial es un *subespacio* de si 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 $\mathcal{V}_2$.

- 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 son cero ni la primera, ni la última componentes (al sumar 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:

$$(\vec{a}, \vec{b}) + (\vec{b}, \vec{c}) = (\vec{a} + \vec{c},\ \vec{b} + \vec{d}); \tag{6.1}$$

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

$$(\boldsymbol{x}, \mathbf{A}) + (\boldsymbol{y}, \mathbf{B}) = (\boldsymbol{x} + \boldsymbol{y},\ \mathbf{A} + \mathbf{B})$$

evidentemenete los tres tipos de sumas, lo son de objetos muy distintos.

Y definimos el producto de un escalar λ por un par ordenado como el par ordenado que resulta de multiplicar cada componente por el escalar λ

$$\lambda(\vec{a}, \vec{b}) = (\lambda\vec{a}, \lambda\vec{b}); \tag{6.2}$$

siguiendo con los productos el mismo criterio que con las sumas en la definición anterior. Por ejemplo

$$\lambda(\boldsymbol{x}, \mathbf{A}) = (\lambda\boldsymbol{x},\ \lambda\mathbf{A}).$$

EJERCICIO 36. 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.*

Generalizando, el *producto Cartesiano* de n conjuntos $A, B, \dots M$, que denotamos con $A \times B \times \cdots M$, e consiste en el conjunto de n-tuplas ordenadas $(a, b, \dots, m)$ donde $a \in A,\ b \in B$,... y $m \in M$:

$$A \times B \times \cdots M = \{\, (a, b, \dots m) \mid a \in A,\ b \in B, \dots \ \text{y} \ \ m \in M \,\}.$$

Pero alternativamente podemos considerar esta generalización como el producto cartesiano del primer conjunto con el producto cartesiano del los $n-1$ restantes:

$$A \times B \times \cdots M = A \times (B \times \cdots M).$$

Ejemplo 8.

$$A \times B \times C \times D = A \times \big(B \times (C \times D)\big).$$

De este modo, aprovechamos el resultado anterior para saber que el producto Cartesiano $\mathcal{A} \times \mathcal{B} \times \cdots \mathcal{X}$ de los espacios vectoriales $\mathcal{A}$, $\mathcal{B}$,...$\mathcal{X}$, 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

☞ Con la *suma de funciones* y el *producto de un escalar por una función* lograremos una potente herramienta para crear nuevos espacios vectoriales... y con algunos de ellos se encontrará 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*

$$\begin{aligned}[f+g]\colon X &\to \mathcal{V} \\ \vec{x} &\to f(\vec{x}) + g(\vec{x})\end{aligned}\ . \tag{6.3}$$

Definición 6.6. *Se define producto de un escalar α por una función* $f\colon X \to \mathcal{V}$ *como la función*

$$\begin{aligned}[\alpha\cdot f]\colon X &\to \mathcal{V} \\ \vec{x} &\to \alpha\ f(\vec{x})\end{aligned}\ . \tag{6.4}$$

Con $\mathcal{V}^D$ denotaremos al conjunto de funciones cuyo dominio es D y cuya imagen está contenida en $\mathcal{V}$:

$$\mathcal{V}^D = \{f \mid f\colon D \to \mathcal{V}\}\,.$$

Ejercicio 37. 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 de números reales: $S\colon \mathbb{N} \to \mathbb{R}$.
- $\mathbb{R}^\mathbb{Z}$: sucesiones de números reales: $S\colon \mathbb{Z} \to \mathbb{R}$.
- $\mathbb{R}^\Omega$: variables aleatorias (donde Ω es el conjunto de sucesos elementales): $S\colon \Omega \to \mathbb{R}$.
- $\left(\mathbb{R}^\Omega\right)^\mathbb{Z}$: series temporales: $S\colon \mathbb{Z} \to \mathbb{R}^\Omega$.
- Etcetera.

6.3.4. Subespacio de funciones lineales

Ejercicio 38. 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.*

Por tanto, concluimos que 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 $\mathcal{V}$).

Por tanto, cerramos la lección con 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)

Apendices 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 los que no existen dos pares distintos que tengan su primeras componentes iguales.*

Dada una función f, se llama *dominio* de f, al conjunto de primeras componentes de los pares de f.

$$\text{dom}(f) = \{x \mid \text{existe } y \text{ tal que } (x, y) \in f\}\,.$$

Dada una función f, se llama *imagen* de f, al conjunto de segundas componentes de f.

$$\mathit{imagen}(f) = \{y \mid \text{existe } x \text{ tal que } (x, y) \in f\}\,.$$

Ejemplos.

Ejemplo 9. El siguiente conjunto (en el que las segundas componentes son el cuadrado de las primeras) es una función

$$f = \{(x, x^2) \mid x \in \mathbb{Z}\} = \{(0,0),\ (1,1),\ (-1,1),\ (2,4),\ (-2,4),\ (3,9),\ (-3,9),\ \dots\}$$

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 función:

$$\left\{(x^2, x) \mid x \in \mathbb{Z}\right\} = \{(0,0),\ (1,1),\ (1,-1),\ (4,2),\ (4,-2),\ (9,3),\ (9,-3),\ \dots\}$$

pues hay pares distintos que tienen la misma primera componente; por ejemplo $(1,1)$ y $(1,-1)$.

Ejemplo 10 (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 11 (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

$$f\colon X \to Y$$

es una forma abreviada de expresar lo siguiente:

1. f es una función
2. $\mathrm{dom}(f) = X$
3. $imagen(f) \subset Y$

Y el esquema

$$\begin{aligned} f\colon X &\to Y \\ x &\to \textit{expresión de } x \end{aligned}$$

es una forma abreviada de expresar el siguiente conjunto de pares:

$$f = \{(x, \textit{expresión de } x) \mid x \in X\}\,.$$

Así, el Ejemplo 9 se puede expresar como $\begin{aligned} f\colon \mathbb{Z} &\to \mathbb{N}\,. \\ x &\to x^2 \end{aligned}$

Y la función *nula* que asigna a todo número real el vector nulo de $\mathbb{R}^2$ como

$$\begin{aligned} n\colon \mathbb{R} &\to \mathbb{R}^2 \quad, \\ x &\to \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

donde aquí el elemento *nulo* de $\mathbb{R}^2$ es $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Y podemos expresar la función *identidad* en $\mathbb{R}^3$ como $\begin{aligned} id\colon \mathbb{R}^3 &\to \mathbb{R}^3. \\ \boldsymbol{x} &\to \boldsymbol{x} \end{aligned}$

Uso de la notación funcional. Posiblemente esté familiarizado con la siguiente notación funcional

$$y = f(x)$$

que equivale a escribir $(x, y) \in f$.

Así, el Ejemplo 9 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) asociados a esta función, y sabrá que describen una parábola.

De igual modo, expresamos la función *nula* que asigna a todo número real la matriz nula de $\mathbb{R}^{2\times 2}$:

$$f(x) = \underset{2\times 2}{\mathbf{0}}, \qquad \text{donde } x \in \mathbb{R} \quad \text{y} \quad \underset{2\times 2}{\mathbf{0}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Y podemos expresar la función *identidad* en $\mathbb{R}^3$ como:

$$f(\boldsymbol{x}) = \boldsymbol{x}, \qquad \text{donde } \boldsymbol{x} \in \mathbb{R}^3.$$

Por tanto, con esta notación podemos expresar la imagen de f como:

$$\mathit{imagen}(f) = \{y \mid \text{existe } x \in \operatorname{dom}(f) \text{ tal que } y = f(x)\}. \tag{6.5}$$

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 pueden obtener pares distintos que tienen la misma primera componente (si invertimos el orden de los pares de la función nula, todos tendrán como primera componente la matriz *nula* de $\mathbb{R}^{2\times 2}$; y más arriba ya vimos qué pasaba al invertir los pares ordenados del primer ejemplo).

Evidentemente

$$\left(f^{-1}\right)^{-1} = f;$$

pues si intercambiamos el orden de los pares dos veces, recuperamos los pares originales.

También es evidente que $\operatorname{dom}(f^{-1}) = \mathit{imagen}(f)$, y que $\mathit{imagen}(f^{-1}) = \operatorname{dom}(f)$.

Proposición 6.A.1. *Si f y g son dos funciones tales que*

$$g(f(x)) = x \;\; \textit{para todo} \;\; x \in \operatorname{dom}(f) \qquad \textit{y} \qquad f(g(y)) = y \;\; \textit{para todo} \;\; y \in \operatorname{dom}(g)$$

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

$$g \circ f = \big\{ ({\color{blue}x}, {\color{blue}z}) \;\big|\; \text{ existe } \; z \; \text{ de modo que } \; ({\color{blue}x}, y) \in f \; \text{ y } \; (y, {\color{blue}z}) \in g \big\}\,.$$

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 $z = g(y)$.

Así, resulta evidente que si f es invertible entonces

$$[f^{-1} \circ f](x) = f^{-1}\big(f(x)\big) = x, \qquad \text{para todo } \; x \in \operatorname{dom}(f).$$

Igualmente evidente es que

$$[f \circ f^{-1}](y) = f\big(f^{-1}(y)\big) = y, \qquad \text{para todo } \; y \in \operatorname{dom}(f^{-1}).$$

6.B. Subespacios y funciones lineales

☞ Hay una estrecha relación entre subespacios y funciones lineales que no se suele contar. La incluyo en este apéndice a la lección.

Proposición 6.B.1. *Si $\mathcal{V}$ y $\mathcal{W}$ son dos espacios vectoriales y $f \subset \mathcal{V} \times \mathcal{W}$ 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*[6] *de $\{\vec{0}\} \times \mathcal{W}$.*

Demostración. 1. $\Rightarrow$ 2. Veamos en primer lugar que f es un subespacio vectorial de $\mathcal{V} \times \mathcal{W}$.

1. Por ser f 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 $\mathcal{V}$ es un espacio vectorial $\mathcal{V} \neq \emptyset$. Luego, $\operatorname{dom}(f) = \mathcal{V} \neq \emptyset$ y por consiguiente $f \neq \emptyset$.
3. Si $(\vec{v}_1, \vec{w}_1)$ y $(\vec{v}_2, \vec{w}_2)$ pertenecen a f y $\alpha_1, \alpha_2 \in \mathbb{R}$ entonces, como $f(\vec{v}_1) = \vec{w}_1$ y $f(\vec{v}_2) = \vec{w}_2$ y f es lineal, tendremos que

$$f(\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2) = \alpha_1 f(\vec{v}_1) + \alpha_2 f(\vec{v}_2) = \alpha_1 \vec{w}_1 + \alpha_2 \vec{w}_2$$

Luego $(\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2, \alpha_1 \vec{w}_1 + \alpha_2 \vec{w}_2) \in f$ y por consiguiente

$$\alpha(\vec{u}_1, \vec{w}_1) + \beta(\vec{u}_2, \vec{w}_2) = (\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2, \alpha_1 \vec{w}_1 + \alpha_2 \vec{w}_2) \in f$$

Veamos en segundo lugar que f y $\{\vec{0}\} \times \mathcal{W}$ son subespacios suplementarios:

1. Si $(\vec{v}, \vec{w}) \in f \cap (\{\vec{0}\} \times \mathcal{W})$ tendremos que $\vec{w} = f(\vec{v})$ y que $\vec{v} = \vec{0}$. Luego, por ser f lineal, necesariamente $\vec{w} = \vec{0}$. Por tanto

$$f \cap (\{\vec{0}\} \times \mathcal{W}) = (\vec{0}, \vec{0})$$

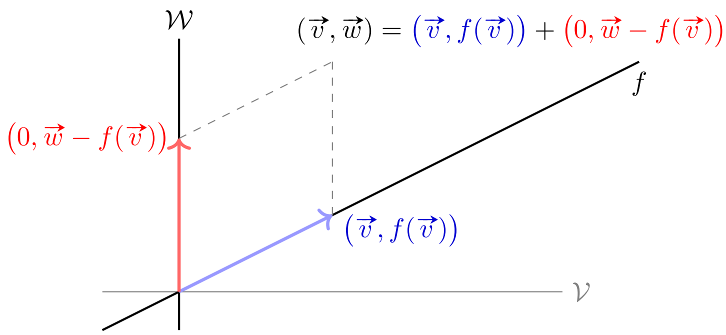
[6] es decir $\mathcal{V} \times \mathcal{W} = f \oplus (\{\vec{0}\} \times \mathcal{W})$. (véase la Definición 10.2 en la página 119)

2. Dado $(\vec{v},\vec{w}) \in \mathcal{V}\times\mathcal{W}$, como $\mathcal{V} = \operatorname{dom}(f)$, tendremos 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 $\mathcal{V}\times\mathcal{W} = f + (\{\vec{0}\}\times\mathcal{W})$.

La siguiente figura es una representación esquemática de que $\mathcal{V}\times\mathcal{W} = f + (\{\vec{0}\}\times\mathcal{W})$:



1. $\Leftarrow$ 2. Veamos 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 de primeras componentes en $\mathcal{V}$ y de 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 dicha diferencia pertenece al subespacio $\{\vec{0}\}\times\mathcal{W}$ que es suplementario de f, necesariamente[7] $\vec{w}_1-\vec{w}_2 = \vec{0}$. Por tanto $\vec{w}_1 = \vec{w}_2$, con lo que es imposible encontrar dos pares distintos de f cuyas primeras componentes sean coincidentes.

   3. De momento ya sabemos que f es una función cuyo dominio está contenido en $\mathcal{V}$ y cuya imagen está contenida en $\mathcal{W}$. Veamos que $\mathcal{V} = \operatorname{dom}(f)$.

      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 dicha descomposición necesariamente $\vec{x}=\vec{v}$, concluimos que $\vec{v}\in\operatorname{dom}(f)$.

   Ya 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, tendremos 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 $f(\alpha_1\vec{v}_1+\alpha_2\vec{v}_2) = \alpha_1 f(\vec{v}_1)+\alpha_2 f(\vec{v}_2)$.

□

[7] $f\cap(\{\vec{0}\}\times\mathcal{W}) = (\vec{0},\vec{0})$

Resolviendo Ax = 0

7.1. Ecuaciones lineales y 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 alguna frecuencia no son capaces de ejecutar el método correctamente.

El método que veremos es deliberadamente distinto del que se cuenta habitualmente. Con ello pretendo, entre otras cosas, que el estudiante no aplique ciegamente una batería de recetas (mal aprendidas en muchos casos). Así pues, *por el momento olvide lo que sabe y trate de entender el método que expondré aquí.*

Por suerte el método es muy sencillo, 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.

7.1.1. Sistemas de ecuaciones lineales

Se llama sistema de ecuaciones lineales a una colección de ecuaciones en la que cada variable es reemplazada por idéntico valor en todas las ecuaciones. 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 anterior sistema también lo podemos escribir como

$$\begin{cases} z + y - 2x = 3 \\ y + x - 2z = 1 \\ x + z - 2y = 2 \end{cases},$$

pero esta libertad en la notación tradicional no ayuda a trabajar con ellos.

7.1.2. Notación matricial

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 con el anterior ejemplo de sistema de ecuaciones, requiere tomar una decisión inicial y arbitraria. Hemos de establecer quién es la primera variable, quién la segunda y quién la tercera; así como también qué ecuación será la primera, cuál la segunda y cuál la tercera. La decisión que tomemos no es importante, pero una vez tomada, debe ser mantenida.

Por ejemplo, podemos decidir que el orden de las variables es, primero x, luego y y por último z; definiendo de este modo 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 con los números del lado derecho de cada ecuación: $\boldsymbol{b} = (1, 2, 3)$. A este vector lo denominamos *vector del lado derecho*.

Así, con la siguiente ecuación matricial denotamos de manera muy compacta un sistema de ecuaciones:

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b},$$

donde ahora la variable $\boldsymbol{x}$ es una etiqueta para ser reemplazada por un vector. Así, con

$$\begin{bmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 2 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{denotamos} \quad \begin{cases} x + y - 2z = 1 \\ x - 2y + z = 2 \\ -2x + y + z = 3 \end{cases}.$$

Llamaremos *solución* al conjunto de vectores que, reemplazando a $\boldsymbol{x} = (x, y, z)$, hacen cierta la igualdad.

☞ La ecuación matricial, $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ define un *sistema de ecuaciones lineales*.

7.2. Sistemas de ecuaciones homogéneos

Cuando el vector del lado derecho es un vector nulo el sistema de ecuaciones de denomina *homogéneo*. En esta Lección 7 solo nos ocuparemos del los sistemas de ecuaciones que son homogéneos:

$$\mathbf{A}\boldsymbol{x} = \mathbf{0}.$$

Fíjese que el lado izquierdo de un sistema de ecuaciones esta 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 el vector de incógnitas $\boldsymbol{x}$ es igual a $\mathbf{0}$, tenemos que

$$\mathbf{A0} = \mathbf{0}.$$

Pero aparte de la solución trivial $\boldsymbol{x} = \mathbf{0}$, ¿existen otros vectores $\boldsymbol{x} \neq 0$ 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

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

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 $\mathbf{I0} = \mathbf{0}$).

Sin embargo, es posible encontrar soluciones no triviales para el sistema

$$\begin{bmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

por ejemplo: $\boldsymbol{x} = (2, 2, 2)$... piense un poco y seguro que es capaz de encontrar muchísimas más.

7.2.1. Espacio nulo de una matriz $\mathcal{N}(\mathbf{A})$

Los ejemplos anteriores muestran que algunos sistemas homogéneos tienen únicamente la *solución trivial*, $\boldsymbol{x} = \mathbf{0}$; pero que otros tienen más soluciones (además de la trivial). Para algunos sistemas de ecuaciones homogéneos el conjunto de soluciones contiene únicamente el vector nulo $\mathbf{0}$, pero para otros el conjunto de soluciones contiene otros vectores además de $\mathbf{0}$.

Vamos a dar un nombre más corto al "*conjunto de soluciones del 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 del $\mathbf{0}$). Pues bien, resulta que el *espacio nulo* de cualquier matriz de n columnas es un subespacio vectorial de $\mathbb{R}^n$.

Proposición 7.2.1. *Para cualquier* $\mathbf{A}$ *de orden* $m \times n$*, el espacio nulo* $\mathcal{N}(\mathbf{A})$ *es un subespacio de* $\mathbb{R}^n$*.*

EJERCICIO 39. Demuestre la Proposición 7.2.1.

Pista. Basta comprobar que el conjunto es cerrado para las combinaciones lineales.

Es muy importante subrayar que el conjunto de soluciones es subespacio solo para los sistemas *homogéneos*... pero nunca para sistemas NO homogéneos $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ (con $\boldsymbol{b} \neq \mathbf{0}$); pues evidentemente el vector nulo $\mathbf{0}$ no pertenece al conjunto de soluciones.

7.2.2. Resolución de un Sistema Homogéneo por eliminación

Puesto que $\mathcal{N}(\mathbf{A})$ 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, ¿como calcular $\mathcal{N}(\mathbf{A})$?

[1] Esta forma de caracterizar los elementos de un conjunto como soluciones de un sistema de ecuaciones se denomina *ecuación cartesiana*.

Revisitando el método de eliminación. Vimos en la Sección 5.2 que si aplicamos una secuencia de transformaciones elementales $\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k$ a las columnas de la matriz particionada que resulta de concatenar $\mathbf{A}$ con $\mathbf{I}$, tenemos:

$$\left[\begin{array}{c}\mathbf{A}\\ \hline \mathbf{I}\end{array}\right]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}=\left[\begin{array}{c}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\end{array}\right]=\left[\begin{array}{c}\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)\\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.\end{array}\right]=\left[\begin{array}{c}\mathbf{AE}\\ \hline \mathbf{E}.\end{array}\right],\qquad\text{donde}\quad \mathbf{E}=\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

Ahora fijémonos en la columna j-ésima de la matriz particionada resultante

$$\left[\begin{array}{c}\mathbf{AE}\\ \hline \mathbf{E}\end{array}\right]_{|j}=\left(\begin{array}{c}\mathbf{A}\left(\mathbf{E}_{|j}\right)\\ \hline \mathbf{E}_{|j}\end{array}\right);$$

para cada columna $j$ésima tenemos que el vector de la parte inferior, $\mathbf{E}_{|j}$, contiene los coeficientes ("la receta") de la combinación lineal de las columnas de $\mathbf{A}$ que hemos usado para calcular ("cocinar") la parte superior de la columna, $\mathbf{AE}_{|j}$.

Ejemplo 12. Apliquemos la eliminación (de izquierda a derecha) para la concatenación de 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. Y fíjese 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:

$$\left[\begin{array}{c}\mathbf{A}\\ \hline \mathbf{I}\end{array}\right]=\left[\begin{array}{cccc}1&2&0&1\\0&1&1&0\\1&2&0&1\\ \hline 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right]\xrightarrow[{[(-1)\mathbf{1}+\mathbf{4}]}]{\substack{\boldsymbol{\tau}\\ [(-2)\mathbf{1}+\mathbf{2}]}}\left[\begin{array}{cccc}1&0&0&0\\0&1&1&0\\1&0&0&0\\ \hline 1&-2&0&-1\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right]\xrightarrow[{[(-1)\mathbf{2}+\mathbf{3}]}]{\boldsymbol{\tau}}\left[\begin{array}{cccc}1&0&0&0\\0&1&0&0\\1&0&0&0\\ \hline 1&-2&2&-1\\0&1&-1&0\\0&0&1&0\\0&0&0&1\end{array}\right]=\left[\begin{array}{c}\mathbf{A}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2\boldsymbol{\tau}_3}\\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2\boldsymbol{\tau}_3}\end{array}\right]$$

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, pues ya sabemos 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 todos y cada uno de los pasos de eliminación.

☞ En todo momento a lo largo de la secuencia de transformaciones

$$\left[\begin{array}{c}\mathbf{A}\\ \hline \mathbf{I}\end{array}\right]\to\left[\begin{array}{c}\mathbf{A}\\ \hline \mathbf{I}\end{array}\right]_{\boldsymbol{\tau}_1}\to\ldots\to\left[\begin{array}{c}\mathbf{A}\\ \hline \mathbf{I}\end{array}\right]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}=\left[\begin{array}{c}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\end{array}\right].$$

la parte inferior de cada columna indica qué combinación lineal de las columnas de $\mathbf{A}$ hemos usado para calcular la parte superior de la columna: $\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_h}\right)_{|j}=\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_h}\right)_{|j};\quad h=1:k.$

Por el Teorema 4.4.2 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

$$\left[\frac{\mathbf{A}}{\mathbf{I}}\right]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \left[\frac{\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}{\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}\right] = \left[\frac{\mathbf{A}\big(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\big)}{\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}}\right] = \left[\frac{\mathbf{AE}}{\mathbf{E}}\right] = \left[\frac{\mathbf{K}}{\mathbf{E}}\right]; \quad \text{donde} \quad \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

Y ahora 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

$$\mathbf{K}_{|j} = \mathbf{A}\big(\mathbf{E}_{|j}\big) = \mathbf{0} \qquad \text{y por lo tanto} \qquad \mathbf{E}_{|j} \in \mathcal{N}\left(\mathbf{A}\right).$$

Es decir, que si $\mathbf{K}$ tiene columnas nulas, los vectores que aparecen por debajo de las columnas nulas son *soluciones* al sistema $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ (y como veremos, también todas sus combinaciones lineales).

Por tanto, en este caso encontrar soluciones al sistema $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ es tan sencillo como alargar $\mathbf{A}$ poniendo 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 de $\mathbf{A}\boldsymbol{x} = \mathbf{0}$. Con respecto al Ejemplo 12 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 al sistema $\mathbf{A}\boldsymbol{x} = \mathbf{0}$. Pero también 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 pues tenemos que

$$\mathbf{A}\boldsymbol{x} = \mathbf{0} \quad\Rightarrow\quad \mathbf{AEE}^{-1}\boldsymbol{x} = \mathbf{0} \quad\Rightarrow\quad \mathbf{R}\big(\mathbf{E}^{-1}\boldsymbol{x}\big) = \mathbf{0} \quad\Rightarrow\quad \mathbf{E}^{-1}\boldsymbol{x} = \mathbf{0} \quad\Rightarrow\quad \mathbf{EE}^{-1}\boldsymbol{x} = \mathbf{E0} \quad\Rightarrow\quad \boldsymbol{x} = \mathbf{0}.$$

Así que la única solución es $\boldsymbol{x} = \mathbf{0}$ (la solución trivial).

Y ahora nos puede surgir una duda, ¿existen otras soluciones que no sean combinación lineal de las encontradas mediante eliminación?

La Proposición 7.2.4 en la página siguiente demuestra que no; es decir, *todas las soluciones son combinación lineal de las soluciones que hemos encontrado aplicando el método de Gauss*.

Pero antes necesitamos demostrar un resultado previo:

EJERCICIO 40. Demuestre que

Lema 7.2.2. *Si $\mathbf{K}$ es pre-escalonada, 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 7.2.3. *Si $\mathbf{A}$ tiene soluciones especiales, todo vector de $\mathcal{N}\left(\mathbf{A}\right)$ 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 pre-escalonada. Puesto que $\mathbf{E}$ es invertible, para todo vector $\boldsymbol{x}$ de $\mathbb{R}^n$ existe un *único* vector tal que

$$\mathbf{E}\boldsymbol{y}_x = \boldsymbol{x};$$

concretamente $\boldsymbol{y}_x = \left(\mathbf{E}^{-1}\right)\boldsymbol{x}$.

Si $\mathbf{K}_{|p_1}, \dots \mathbf{K}_{|p_r}$ son las r columnas no nulas de $\mathbf{K}$ (es decir, las columnas con pivote), por el Lemma 7.2.2 sabemos que la condición necesaria y suficiente para que $\mathbf{K}\boldsymbol{y}_x = \mathbf{0}$ es que las componentes del vector $\boldsymbol{y}_x$ que multiplican a las columnas con pivote sean nulas, es decir, $\left(\boldsymbol{y}_x\right)_{|p_1} = \dots = \left(\boldsymbol{y}_x\right)_{|p_r} = 0$.

Así, si $\{e_1, ..., e_{n-r}\}$ es el conjunto de índices correspondientes a las columnas nulas de $\mathbf{K}$, es decir, si $\{p_1, ..p_r\} \cup \{e_1, ..., e_{n-r}\} = \{1, .., n\}$; como

$$\boldsymbol{x} \;=\; \mathbf{E}\boldsymbol{y}_x \;=\; \left(\mathbf{E}_{|p_1}\right)\left(\boldsymbol{y}_x\right)_{|p_1} + \dots + \left(\mathbf{E}_{|p_r}\right)\left(\boldsymbol{y}_x\right)_{|p_r} \;+\; \left(\mathbf{E}_{|e_1}\right)\left(\boldsymbol{y}_x\right)_{|e_1} + \dots + \left(\mathbf{E}_{|e_{(n-r)}}\right)\left(\boldsymbol{y}_x\right)_{|e_{(n-r)}},$$

tendremos que $\mathbf{A}\boldsymbol{x} = \mathbf{0}$, es decir, que $\mathbf{AE}\boldsymbol{y}_x = \mathbf{K}\boldsymbol{y}_x = \mathbf{0}$ y como sabemos que las componentes $\left(\boldsymbol{y}_x\right)_{|p_1}$, $\dots$, $\left(\boldsymbol{y}_x\right)_{|p_r}$ de $\boldsymbol{y}_x$ son nulas, llegamos a que $\boldsymbol{x}$ es combinación lineal de $\left(\mathbf{E}_{|e_1}\right), \dots, \left(\mathbf{E}_{|e_{(n-r)}}\right)$, pues

$$\boldsymbol{x} = \left(\mathbf{E}_{|e_1}\right)\left(\boldsymbol{y}_x\right)_{|e_1} + \dots + \left(\mathbf{E}_{|e_{(n-r)}}\right)\left(\boldsymbol{y}_x\right)_{|e_{(n-r)}}.$$

□

Por tanto tenemos el siguiente corolario:

Corolario 7.2.4. *O todas las soluciones de* $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ *son combinaciones lineales de las soluciones encontradas por el método de eliminación, o bien, la forma pre-escalonada de* $\mathbf{A}$ *no tiene columnas nulas y la única solución es* $\mathbf{0}$*, es decir:*

$$\text{Ó}\quad \mathcal{N}\left(\mathbf{A}\right) = \{\textit{combinaciones lineales de las soluciones encontradas por eliminación}\} \quad \text{ó} \quad \mathcal{N}\left(\mathbf{A}\right) = \{\mathbf{0}\}.$$

Ahora ya podemos terminar el Ejemplo 12 en la página 90 concluyendo que el espacio nulo de $\mathbf{A}$ (es decir, el conjunto de soluciones del sistema $\mathbf{A}\boldsymbol{x} = \mathbf{0}$) es:

$$\mathcal{N}\left(\mathbf{A}\right) = \left\{ \boldsymbol{x} \in \mathbb{R}^4 \;\middle|\; \text{existen } a, b \in \mathbb{R} \text{ tales que } \boldsymbol{x} = a\begin{pmatrix}2\\-1\\1\\0\end{pmatrix} + b\begin{pmatrix}-1\\0\\0\\1\end{pmatrix} \right\}.$$

Es decir, $\mathcal{N}\left(\mathbf{A}\right)$ *es el subconjunto de vectores* $\boldsymbol{x}$ *de* $\mathbb{R}^4$ *que verifican que, para cada* $\boldsymbol{x}$*, es posible encontrar dos números reales* a *y* b *tales que* a *veces la primera solución especial más* b *veces la segunda es igual a* $\boldsymbol{x}$*.*[2]

Nomenclatura usada por G. Strang Siguiendo la nomenclatura de G. Strang, a las columnas de $\mathbf{A}$ que acaban siendo nulas en la forma pre-escalonada $\mathbf{K}$ se denominarán *columnas libres* y el resto (i.e. las que mantienen un pivote) se llamarán *columnas pivote.*

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).

[2] Esta forma de identificar los elementos $\boldsymbol{x}$ de un subconjunto como aquellos que para los que es posible encontrar parámetros que permitan expresar cada $\boldsymbol{x}$ como una combinación lineal se denomina *ecuación paramétrica*.

Fíjese que con las *ecuaciones paramétricas* es muy fácil obtener elementos del conjunto, basta emplear elegir dos valores cualesquiera para a y b 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 ecuación sirve para una cosa distinta, unas para generar ejemplos, y las otras para verificar la pertenencia!

☞ **Algoritmo para resolver $\mathbf{A}\boldsymbol{x} = \mathbf{0}$**

1. Aplicamos la eliminación: $\begin{bmatrix}\mathbf{A}\\ \hline \mathbf{I}\end{bmatrix} \xrightarrow{\boldsymbol{\tau}_1,\ldots,\boldsymbol{\tau}_k} \begin{bmatrix}\mathbf{K}\\ \hline \mathbf{E}\end{bmatrix}$, donde $\mathbf{K} = \mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ y $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$.
2. Si hay *soluciones especiales* (columnas de $\mathbf{E}$ bajo las columnas nulas de $\mathbf{K}$)
   - Solución completa: $\mathcal{N}(\mathbf{A}) = \{$combinaciones lineales de las *soluciones especiales*$\}$
3. Si no hay *soluciones especiales* (si $\mathbf{K}$ no tiene columnas nulas)
   - Solución completa: $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$

Librería NAcAL para Python

```python
A = Matrix( [ [1,2,0,1],[0,1,1,0],[1,2,0,1] ] )
Homogenea(A, 1)                              # Resuelve la ecuación homogenea
```

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

$$\text{Conjunto de combinaciones lineales de:} \quad \begin{pmatrix}2\\-1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\0\\1\end{pmatrix}.$$

7.3. Generalizaciones o visiones alternativas de resultados previos

Podemos generalizar la Proposición 7.2.1:

EJERCICIO 41. Demuestre la siguiente

Proposición 7.3.1. *Para toda función lineal $f\colon \mathcal{V} \to \mathcal{W}$, el conjunto*

$$\mathcal{N}(f) = \left\{\vec{v} \,\middle|\, f(\vec{v}) = \vec{0}\right\}$$

es un subespacio de $\mathcal{V}$.

Fíjese que la Proposición 7.2.1 es un caso particular en el que la función es $\begin{array}{rl} f_{\mathbf{A}}\colon \mathbb{R}^n & \to \mathbb{R}^m. \\ \boldsymbol{x} & \to \mathbf{A}\boldsymbol{x}\end{array}$

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 $\boldsymbol{b} \neq \mathbf{0}$). Es posible re-escribir el sistema para que tenga una apariencia de *sistema homogéneo* ya que:

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b} \qquad \iff \qquad \mathbf{A}\boldsymbol{x} - \boldsymbol{b} = \mathbf{0} \qquad \iff \qquad \left[\ \mathbf{A} \mid -\boldsymbol{b}\ \right]\begin{pmatrix}\boldsymbol{x}\\1\end{pmatrix} = \mathbf{0}; \tag{8.1}$$

donde el vector de incógnitas, $\begin{pmatrix}\boldsymbol{x}\\1\end{pmatrix}$, tiene una restricción: *su última componente debe ser un 1.*

La matriz $\left[\ \mathbf{A} \mid -\boldsymbol{b}\ \right]$ se denomina *matriz de coeficientes "ampliada"*, y marcamos 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 *vector del lado derecho* (**con el signo cambiado**).

Pues bien, el método de resolución que exponemos aquí intenta resolver este "pseudo-*sistema homogéneo*" tratando de encontrar soluciones que tengan un 1 en su última componente (es decir, trataremos de *encontrar algún vector de* $\mathcal{N}\left(\left[\ \mathbf{A} \mid -\boldsymbol{b}\ \right]\right)$ *que tenga un 1 en la última componente*).

☞ Las implicaciones "si y solo si" de más arriba indican que encontrar soluciones al sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ 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 sistema tiene solución. ¡Si además logramos mantener el 1 de la solución especial correspondiente a la columna $n+1$, entonces tendremos a la vista una solución a $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$!

Veamos primero un ejemplo de un sistema sin solución:

Ejemplo 13. 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 $\boldsymbol{d} = \begin{pmatrix}5\\-1\\1\\-1\end{pmatrix}$.

Para resolver el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{d}$ (con $\mathbf{A}$ de orden m por n) por el método de eliminación, "alargamos" la matriz ampliada $\left[\ \mathbf{A}\ \middle|\ -\boldsymbol{d}\ \right]$ (de orden m por $(n+1)$) con una matriz identidad de orden $(n+1)$:

$$\left[\begin{array}{c|c} \underset{m\times n}{\mathbf{A}} & -\boldsymbol{d} \\ \hline \underset{n\times n}{\mathbf{I}} & \\ \hline & 1 \end{array}\right] \qquad \text{donde} \qquad \underset{(n+1)\times(n+1)}{\mathbf{I}} = \left[\begin{array}{c|c} \underset{n\times n}{\mathbf{I}} & \\ \hline & 1 \end{array}\right];$$

donde distinguimos los bloques correspondientes a las distintas partes con una partición (mediante líneas horizontales y verticales).

Y ahora pre-escalonamos la matriz de coeficientes ampliada (fila superior de bloques de la matriz particionada)... pero aplicando las transformaciones a las columnas completas, es decir, a las columnas "alargadas" (con un total de $m+n+1$ componentes):

$$\left[\begin{array}{c|c} \mathbf{A} & -\boldsymbol{d} \\ \hline \mathbf{I} & \\ \hline & 1 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 1 & 2 & -5 \\ 2 & 2 & 4 & 1 \\ 3 & 1 & 4 & -1 \\ 4 & 1 & 5 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[{[(5)\mathbf{1}+\mathbf{4}]}]{\substack{\boldsymbol{\tau} \\ [(-1)\mathbf{1}+\mathbf{2}] \\ [(-2)\mathbf{1}+\mathbf{3}]}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 11 \\ 3 & -2 & -2 & 14 \\ 4 & -3 & -3 & 21 \\ \hline 1 & -1 & -2 & 5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[{[(7)\mathbf{2}+\mathbf{4}]}]{\substack{\boldsymbol{\tau} \\ [(-1)\mathbf{2}+\mathbf{3}]}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 11 \\ 3 & -2 & 0 & 0 \\ 4 & -3 & 0 & 0 \\ \hline 1 & -1 & -1 & -2 \\ 0 & 1 & -1 & 7 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

La parte superior de la matriz está pre-escalonada, y como no hay ningún pivote en la segunda fila de l bloque de la izquierda (correspondiente a la matriz de coeficientes), no es posible eliminar el pivote (el 11) de la última columna de la matriz de coeficientes ampliada. Por tanto, $\boldsymbol{b}$ no es combinación lineal de las columnas de $\mathbf{A}$, es decir, el sistema no tiene solución, pues no es posible encontrar una combinación de columnas $\mathbf{A}\boldsymbol{x}$ que sea igual a $\boldsymbol{b}$.

O visto de otro modo, no existe un vector de $\mathcal{N}\left(\left[\ \mathbf{A}\ \middle|\ -\boldsymbol{d}\ \right]\right)$ cuya última componente sea un uno, así que el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{d}$ NO tiene solución.

Ejemplo 14. 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

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & -5 \\ 2 & 2 & 4 & -10 \\ 3 & 1 & 4 & -1 \\ 4 & 1 & 5 & 1 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[{[(5)\mathbf{1}+\mathbf{4}]}]{\substack{\boldsymbol{\tau} \\ [(-1)\mathbf{1}+\mathbf{2}] \\ [(-2)\mathbf{1}+\mathbf{3}]}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 3 & -2 & -2 & 14 \\ 4 & -3 & -3 & 21 \\ \hline 1 & -1 & -2 & 5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[{[(7)\mathbf{2}+\mathbf{4}]}]{\substack{\boldsymbol{\tau} \\ [(-1)\mathbf{2}+\mathbf{3}]}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 \\ 4 & -3 & 0 & 0 \\ \hline 1 & -1 & -1 & -2 \\ 0 & 1 & -1 & 7 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

Hemos encontrado dos soluciones especiales (los vectores por debajo de las dos columnas nulas), de las cuales la segunda tiene un 1 como ultima 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 $\boldsymbol{x} = (-2,\ \ 7,\ \ 0)$ es *una* solución (compruébelo).

¿Una solución o infinitas soluciones?

Fíjese que si sumamos cualquier múltiplo de la tercera columna a la cuarta, obtenemos un nuevo vector de $\mathcal{N}\left(\left[\ \mathbf{A}\ \middle|\ -\boldsymbol{b}\ \right]\right)$ con un 1 como última componente. De hecho, el conjunto de vectores de $\mathcal{N}\left(\left[\ \mathbf{A}\ \middle|\ -\boldsymbol{b}\ \right]\right)$

con un 1 como última componente son los vectores de la forma

$$\left\{ \boldsymbol{y} \in \mathbb{R}^4 \;\middle|\; \text{existe } a \in \mathbb{R} \text{ tal que } \boldsymbol{y} = \begin{pmatrix} -2 \\ 7 \\ 0 \\ \hline 1 \end{pmatrix} + a \begin{pmatrix} -1 \\ -1 \\ 1 \\ \hline 0 \end{pmatrix} \right\}$$

Consecuentemente (Ecuación 8.1) todas las soluciones del sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ son de la forma

$$\left\{ \boldsymbol{x} \in \mathbb{R}^3 \;\middle|\; \text{existe } a \in \mathbb{R} \text{ tal que } \boldsymbol{x} = \begin{pmatrix} -2 \\ 7 \\ 0 \end{pmatrix} + a \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\}$$

Estas ecuaciones paramétricas nos permiten obtener tantos ejemplos de soluciones como queramos, basta con dar valores arbitrarios al parámetro a (obtenga alguna otra solución $\boldsymbol{y}$ y compruebe que $\mathbf{A}\boldsymbol{y} = \boldsymbol{b}$).

En general, considere un sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ que tiene solución y donde el rango de la matriz de coeficientes $\underset{m\times n}{\mathbf{A}}$ es r; Y sea $\left[\begin{array}{c|c} \mathbf{A} & -\boldsymbol{b} \end{array}\right]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ una forma pre-escalonada de la matriz de coeficientes ampliada y $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$. Sabemos que el espacio nulo $\mathcal{N}\left(\left[\begin{array}{c|c} \mathbf{A} & -\boldsymbol{b} \end{array}\right]\right)$ son las combinaciones lineales de las soluciones especiales. Como en la demostración de la Proposición 7.2.4 en la página 92, sea $\{e_1, ..., e_{n-r}, (n+1)\}$ el conjunto de índices correspondientes a las columnas nulas de la forma pre-escalonada, y donde sabemos que el último índice es $n+1$ puesto que $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ tiene solución; entonces:

$$\mathcal{N}\left(\left[\begin{array}{c|c} \mathbf{A} & -\boldsymbol{b} \end{array}\right]\right) = \left\{ \boldsymbol{x} \in \mathbb{R}^{(n+1)} \text{ tales que } \boldsymbol{x} = \left(\mathbf{E}_{|e_1}\right)\lambda_{e_1} + \ldots + \left(\mathbf{E}_{|e_{(n-r)}}\right)\lambda_{e_{(n-r)}} + \left(\mathbf{E}_{|(n+1)}\right)\lambda_{(n+1)} \right\};$$

y como únicamente la última columna $\mathbf{E}_{|n+1}$ tiene un 1 en su última componente[1], el conjunto de vectores de $\mathcal{N}\left(\left[\begin{array}{c|c} \mathbf{A} & -\boldsymbol{b} \end{array}\right]\right)$ con un 1 en su última componente es

$$\left\{ \boldsymbol{x} \in \mathbb{R}^{n+1} \text{ tales que } \boldsymbol{x} = \left(\mathbf{E}_{|e_1}\right)\lambda_{e_1} + \ldots + \left(\mathbf{E}_{|e_{(n-r)}}\right)\lambda_{e_{(n-r)}} + \mathbf{E}_{|n+1} \right\}; \tag{8.2}$$

es decir, necesariamente $\lambda_{(n+1)} = 1$. 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 a los vectores del conjunto de la Ecuación (8.2).
2. si la forma pre-escalonada de $\mathbf{A}$ tiene columnas nulas (si rango$(\mathbf{A}) < n$) entonces el sistema tiene infinitas soluciones. Y si por el contrario $\mathcal{N}(\mathbf{A}) = \{\mathbf{0}\}$, entonces existe un único vector solución... (el vector cuyas n componentes son las n primeras componentes de $\mathbf{E}_{|n+1}$).

Algoritmo de resolución de un sistema no homogéneo

Aplicando el método de eliminación de izquierda a derecha para pre-escalonar, logramos:

$$\left[\begin{array}{c|c} \underset{m\times n}{\mathbf{A}} & -\boldsymbol{b} \\ \hline \underset{n\times n}{\mathbf{I}} & \\ \hline & 1 \end{array}\right]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_p} = \left[\begin{array}{c|c} \mathbf{K} & \boldsymbol{c} \\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \boldsymbol{v} \\ \hline & d \end{array}\right], \tag{8.3}$$

donde $\left[\begin{array}{c|c} \mathbf{K} & \boldsymbol{c} \end{array}\right]$ es una matriz pre-escalonada de la matriz de coeficientes ampliada $\left[\begin{array}{c|c} \mathbf{A} & -\boldsymbol{b} \end{array}\right]$; y donde la única componente distinta de cero, de la última fila, es d en la última columna. Si sigue el algoritmo de la demostración del Teorema 4.4.1 en la página 51, entonces d siempre será igual a 1, pues no emplea

[1] recuerde que la última componente del resto de vectores $\mathbf{E}_{|j}$ con $j \neq (n+1)$ es un 0

transformaciones de *Tipo II*. Pero si usted opera como lo hace la librería NAcAL (que multiplica algunas columnas por números no nulos para evitar así operar con fracciones cuando ello es posible), entonces d podrá tomar cualquier valor no nulo.

Tras la eliminación en (8.3), solo caben dos posibilidades:

1. Si $\boldsymbol{c} = \boldsymbol{0}$, entonces el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ tiene solución, pues debajo de $\boldsymbol{c}$ hay un vector del espacio nulo de la matriz ampliada cuya última componente d es distinta de cero (es lo que pasó en el segundo de los dos ejemplos anteriores). Si $d \neq 1$, basta dividir la última columna por d para obtener un vector del espacio nulo de la matriz ampliada.

$$\left[\begin{array}{c|c} \underset{m\times n}{\mathbf{K}} & \boldsymbol{0} \\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \boldsymbol{v} \\ \hline & d \end{array}\right] \xrightarrow[{[(\frac{1}{d})\boldsymbol{(n+1)}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccccc|c} * & 0 & \cdots & 0 & 0 & 0 \\ : & * & \cdots & 0 & 0 & 0 \\ : & : & & 0 & 0 & 0 \\ : & : & \cdots & * & 0 & 0 \\ : & : & \cdots & : & 0 & 0 \\ \hline : & : & \cdots & : & : & \boldsymbol{s}_p \\ \hline 0 & 0 & \cdots & 0 & 0 & 1 \end{array}\right] \quad \text{donde “}*\text{” son pivotes,}$$

   y donde $\boldsymbol{s}_p = \frac{1}{d}\boldsymbol{v}$ es un vector solución, es decir $\mathbf{A}(\boldsymbol{s}_p) = \boldsymbol{b}$.

2. Si $\boldsymbol{c} \neq \boldsymbol{0}$, entonces el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ NO tiene solución.

$$\left[\begin{array}{c|c} \mathbf{K} & \boldsymbol{c} \\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \boldsymbol{v} \\ \hline & d \end{array}\right] = \left[\begin{array}{ccccc|c} * & 0 & \cdots & 0 & 0 & 0 \\ : & * & \cdots & 0 & 0 & 0 \\ : & : & & 0 & 0 & * \\ : & : & \cdots & * & 0 & : \\ : & : & \cdots & : & 0 & : \\ \hline : & : & \cdots & : & : & \boldsymbol{v} \\ \hline 0 & 0 & \cdots & 0 & 0 & d \end{array}\right] \quad \text{donde “}*\text{” son pivotes,}$$

Por tanto, *el sistema tiene solución si y solo si logramos transformar $-\boldsymbol{b}$ en un $\boldsymbol{0}$*.

Resolviendo varios sistemas a la vez Este método, con una mínima variación, permite resolver simultáneamente sistemas que compartan la matriz de coeficientes. Por ejemplo, podemos resolver simultáneamente los dos ejemplos anteriores:[2]

$$\left[\begin{array}{c||c|c} \mathbf{A} & -\boldsymbol{b} & -\boldsymbol{d} \\ \hline \mathbf{I} & & \\ \hline & 1 & 1 \end{array}\right] = \left[\begin{array}{ccc||c|c} 1 & 1 & 2 & -5 & -5 \\ 2 & 2 & 4 & -10 & 1 \\ 3 & 1 & 4 & -1 & -1 \\ 4 & 1 & 5 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 \end{array}\right] \to \left[\begin{array}{ccc||c|c} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 11 \\ 3 & -2 & -2 & 14 & 14 \\ 4 & -3 & -3 & 21 & 21 \\ \hline 1 & -1 & -2 & 5 & 5 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 \end{array}\right] \to \left[\begin{array}{ccc||c|c} 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 11 \\ 3 & -2 & 0 & 0 & 0 \\ 4 & -3 & 0 & 0 & 0 \\ \hline 1 & -1 & -1 & -2 & -2 \\ 0 & 1 & -1 & 7 & 7 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

Es decir, $\mathbf{A}\boldsymbol{x} = \boldsymbol{d}$ NO tiene solución pero $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ si es resoluble, y el conjunto de todas las soluciones es

$$\left\{ \boldsymbol{x} \in \mathbb{R}^3 \;\middle|\; \text{existe } a \in \mathbb{R} \text{ tal que } \boldsymbol{x} = \begin{pmatrix} -2 \\ 7 \\ 0 \end{pmatrix} + a \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\}.$$

[2]¡pero tenga en cuenta que aquí no se puede usar la penúltima columna para eliminar componentes de la última!

☞ 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]:

$$\left[\begin{array}{c|c}\mathbf{A} & -\boldsymbol{b}\\ \hline \mathbf{I} & \\ \hline & 1\end{array}\right]\xrightarrow{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_p}\left[\begin{array}{c|c}\mathbf{K} & \boldsymbol{f}\\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \boldsymbol{v}\\ \hline & d\end{array}\right]\xrightarrow[{[(\frac{1}{d})\boldsymbol{(n+1)}]}]{\boldsymbol{\tau}}\left[\begin{array}{c|c}\mathbf{K} & \boldsymbol{c}\\ \hline \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \boldsymbol{s}\\ \hline & 1\end{array}\right]$$

donde $\mathbf{K}=\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ está pre-escalonada; y por tanto: $\mathbf{K}_{|j}=\mathbf{0}\implies\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)_{|j}\in\mathcal{N}(\mathbf{A})$.

- 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

$$\{\boldsymbol{x}\in\mathbb{R}^n \mid \text{existe } \boldsymbol{y}\in\mathcal{N}(\mathbf{A}) \text{ tal que } \boldsymbol{x}=\boldsymbol{s}+\boldsymbol{y}\}.$$

Por tanto, cuando $\mathbf{K}$ no tiene columnas nulas ($\mathcal{N}(\mathbf{A})=\{\mathbf{0}\}$) 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 4.4.1 en la página 51, pero si también se emplean transformaciones *Tipo II*, entonces d podría tomar cualquier valor no nulo.

Librería 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
```

$$\left[\begin{array}{ccc|c}7&-2&0&-4\\6&4&2&-12\\ \hline 1&0&0&0\\0&1&0&0\\0&0&1&0\\ \hline 0&0&0&1\end{array}\right]\xrightarrow{\substack{\boldsymbol{\tau}\\ [(7)\mathbf{2}]\\ [(2)\mathbf{1+2}]\\ [(7)\mathbf{4}]\\ [(4)\mathbf{1+4}]}}\left[\begin{array}{ccc|c}7&0&0&0\\6&40&2&-60\\ \hline 1&2&0&4\\0&7&0&0\\0&0&1&0\\ \hline 0&0&0&7\end{array}\right]\xrightarrow{\substack{\boldsymbol{\tau}\\ [(20)\mathbf{3}]\\ [(-1)\mathbf{2+3}]\\ [(2)\mathbf{4}]\\ [(3)\mathbf{2+4}]}}\left[\begin{array}{ccc|c}7&0&0&0\\6&40&0&0\\ \hline 1&2&-2&14\\0&7&-7&21\\0&0&20&0\\ \hline 0&0&0&14\end{array}\right]\xrightarrow{\substack{\boldsymbol{\tau}\\ [(\frac{1}{14})\mathbf{4}]}}\left[\begin{array}{ccc|c}7&0&0&0\\6&40&0&0\\ \hline 1&2&-2&1\\0&7&-7&\frac{3}{2}\\0&0&20&0\\ \hline 0&0&0&1\end{array}\right]$$

$$\text{Conjunto de vectores: }\left\{\boldsymbol{x}\in\mathbb{R}^3\;\middle|\;\exists\boldsymbol{p}\in\mathbb{R}^1\text{ tal que }\boldsymbol{x}=\begin{pmatrix}1\\\frac{3}{2}\\0\end{pmatrix}+\begin{bmatrix}-2\\-7\\20\end{bmatrix}\boldsymbol{p}\right\}.$$

(Como se ha multiplicado la cuarta columna por 7 y luego por 2, al final hay que dividir por 14.)

8.1.1. El Teorema de de Rouché-Frobenius

☞ Del procedimiento anterior se deduce que los rangos de la matriz de coeficientes $\mathbf{A}$ y de la matriz ampliada $\left[\mathbf{A}\middle|-\boldsymbol{b}\right]$ nos 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

$$\left[\mathbf{A}\,\middle|\,-\boldsymbol{b}\right]\xrightarrow{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\left[\mathbf{K}\,\middle|\,\boldsymbol{c}\right]$$

solo se pueden dar dos casos: que la última columna sea nula ($\boldsymbol{c}=\mathbf{0}$), o que no lo sea ($\boldsymbol{c}\neq\mathbf{0}$):

1. Si $\boldsymbol{c} \neq \boldsymbol{0}$, entonces la matriz ampliada tiene una columna pivote adicional, es decir,

$$\text{rango}\left(\left[\mathbf{A} \mid -\boldsymbol{b}\right]\right) > \text{rango}\left(\mathbf{A}\right) \qquad \Longleftrightarrow \qquad \text{el sistema NO tiene solución.}$$

2. Si $\boldsymbol{c} = \boldsymbol{0}$, entonces

$$\text{rango}\left(\left[\mathbf{A} \mid -\boldsymbol{b}\right]\right) = \text{rango}\left(\mathbf{A}\right) \qquad \Longleftrightarrow \qquad \text{el sistema es resoluble.}$$

   Si además $\mathbf{A}$ es de rango completo por columnas (si todas sus columnas son columnas pivote), entonces la forma pre-escalonada $\mathbf{K}$ no tiene columnas nulas (es decir, $\mathcal{N}\left(\mathbf{A}\right) = \{\mathbf{0}\}$), por lo que el sistema tiene una única solución.

   Pero 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 $\mathcal{N}\left(\mathbf{A}\right)$ para obtener otra solución.

El anterior resultado se resume en el siguiente teorema:

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* $\text{rango}\left(\mathbf{A}\right) = n$*, 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 las columnas pivote de $\mathbf{A}$ y haber un pivote en cada fila).

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 (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. $\mathbf{A}$ *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

$$\left\{\boldsymbol{b} \in \mathbb{R}^n \;\middle|\; \text{existen } x_1, \ldots, x_n \in \mathbb{R} \text{ tales que } \boldsymbol{b} = x_1\mathbf{A}_{|1} + \cdots + x_n\mathbf{A}_{|n}\right\}$$

o bien, usando la notación matricial,

$$\{\boldsymbol{b} \in \mathbb{R}^n \mid \text{existe } \boldsymbol{x} \in \mathbb{R}^n \text{ tal que } \boldsymbol{b} = \mathbf{A}\boldsymbol{x}\}$$

Puesto que este subconjunto de $\mathbb{R}^n$ *contiene todas* las combinaciones lineales de las columnas, obviamente *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 $\mathcal{C}\left(\mathbf{A}\right)$. Por ejemplo

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}; \qquad \Rightarrow \qquad \mathcal{C}\left(\mathbf{A}\right) = \left\{\boldsymbol{b} \in \mathbb{R}^n \;\middle|\; \text{existen } x, y, z \in \mathbb{R} \text{ tales que } \boldsymbol{b} = x\begin{pmatrix}1\\1\end{pmatrix} + y\begin{pmatrix}1\\0\end{pmatrix} + z\begin{pmatrix}0\\1\end{pmatrix}\right\}.$$

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 $\mathbf{A}$:

$$\mathcal{C}\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) \subset \mathcal{C}(\mathbf{A});$$

pero como las transformaciones elementales son "reversibles", mediante una sucesión de transformaciones elementales se puede retornar de $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ 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

$$\mathcal{C}(\mathbf{A}) \subset \mathcal{C}\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right).$$

Por tanto

$$\boxed{\mathcal{C}(\mathbf{A}) = \mathcal{C}\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)}.$$

¡Al aplicar la eliminación, todas las matrices que aparecen en el proceso tienen el mismo espacio columna!

Así pues, 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 de resultados previos

Ejercicio 42. Considere $\underset{m\times n}{\mathbf{A}}$, y compruebe que $\mathcal{C}(\mathbf{A}) = imagen(f_{\mathbf{A}})$, donde $\begin{array}{rl} f_{\mathbf{A}}: \mathbb{R}^n & \to \mathbb{R}^n \\ \boldsymbol{x} & \to \mathbf{A}\boldsymbol{x} \end{array}$.

Ejercicio 43. Demuestre la siguiente

Proposición 8.3.1. *Si $f: \mathcal{D} \to \mathcal{W}$ es una función lineal, entonces $imagen(f)$ es un subespacio de $\mathcal{W}$.*

Independencia, base y dimensión

9.1. Combinaciones lineales de vectores de $\mathcal{V}$

[Recuerde las definiciones de *Espacio Vectorial* (Pag. 74) y de *Subespacio* (Pag. 78)]

En esta lección vamos a tratar con *sistemas* de vectores de un subespacio genérico $\mathcal{V}$:

$$\mathsf{Z} = \left[\vec{z}_1;\ \vec{z}_2; \ldots; \vec{z}_n\right], \quad \text{con} \quad \vec{z}_i \in \mathcal{V};$$

y dicho sistemas los denotaremos con caracteres del tipo: A, B,… X, Y, Z. Una vez más podemos emplear el operador selector para denotar el $j$ésimo vector de un sistema $\mathsf{Z} = \left[\vec{z}_1; \ldots; \vec{z}_n\right]$:

$$\mathsf{Z}_{|j} = \vec{z}_j; \quad \text{y por tanto} \quad \mathsf{Z} = \left[\mathsf{Z}_{|1};\ \mathsf{Z}_{|2}; \ldots; \mathsf{Z}_{|n}\right];$$

Fíjese que las matrices son un caso particular, pues $\mathbf{A}$ de orden m por n es un sistema de n vectores de $\mathbb{R}^m$:

$$\mathbf{A} = \left[\mathbf{A}_{|1};\ \mathbf{A}_{|2}; \ldots; \mathbf{A}_{|n}\right], \quad \text{con} \quad \mathbf{A}_{|j} \in \mathbb{R}^m.$$

Librería NAcAL para Python

```
A = Matrix( [ [1,0,0], [-1,0,1] ] )
A.sis()              # Devuelve el Sistema correspondiente a las columnas de A
```

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} = \left[\begin{pmatrix} 1 \\ -1 \end{pmatrix}; \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 0 \\ 1 \end{pmatrix};\right].$$

9.1.1. Extendiendo la notación matricial a los espacios vectoriales

Aunque no es habitual, extenderé la notación empleada con las matrices a espacios vectoriales genéricos [1] empezando por 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 Z, cuyos coeficientes son las componentes de $\boldsymbol{a}$:

$$\mathsf{Z}\boldsymbol{a} = (\mathsf{Z}_{|1})a_1 + \cdots + (\mathsf{Z}_{|n})a_n;$$

donde $\boldsymbol{a} \in \mathbb{R}^n$ contiene los coeficientes de la combinación lineal.

Por tanto $\mathsf{Z}\boldsymbol{a}$ es un vector de $\mathcal{V}$. Así, en el caso particular del producto $\mathbf{A}\boldsymbol{x}$ obtenemos un vector de $\mathbb{R}^m$. Pero en el caso general $\mathsf{Z}\boldsymbol{a}$, y dependiendo de la naturaleza de $\mathcal{V}$, el vector resultante pudiera ser una función, o matriz, o polinomio, o variable aleatoria, o etc.… todo depende de la naturaleza de $\mathcal{V}$.

[1] espacios vectoriales genéricos *de dimensión finita*.

Ejemplo:

Librería NAcAL para Python

```
A = Matrix([ [ 1, 1, 1], [1, 0, 0] ])
B = Matrix([ [-1, 1, 1], [0, 1, 0] ])
C = Matrix([ [ 2, 0, 1], [0, 0, 1] ])
Z = Sistema( [ A, B, C] )          # ¡Sistema de Matrices!
A = Vector ( [ 1, 2, 3] )          # Sistema por Vector (comb. lineal de matrices)
Z*a
```

$$\mathsf{Z}\boldsymbol{a} \;=\; \left[\begin{bmatrix}1&1&1\\1&0&0\end{bmatrix};\begin{bmatrix}-1&1&1\\0&1&0\end{bmatrix};\begin{bmatrix}2&0&1\\0&0&1\end{bmatrix};\right]\begin{pmatrix}1\\2\\3\end{pmatrix} \;=\; \begin{bmatrix}5&3&6\\1&2&3\end{bmatrix}.$$

El siguiente ejercicio muestra que el producto $\mathsf{Z}\boldsymbol{a}$ verifica las *propiedades de linealidad*: la primera respecto a la suma, $\mathsf{Z}\big(\boldsymbol{b}+\boldsymbol{c}\big)=\mathsf{Z}\boldsymbol{b}+\mathsf{Z}\boldsymbol{c}$, 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.** *Sean* Z*, un sistema de* n *vectores de* $\mathcal{V}$*, y sean* $\boldsymbol{b}$ *y* $\boldsymbol{c}$ *vectores de* $\mathbb{R}^n$*; entonces*

$$\mathsf{Z}\big(\boldsymbol{b}+\boldsymbol{c}\big)=\mathsf{Z}\boldsymbol{b}+\mathsf{Z}\boldsymbol{c}.$$

(b) **Proposición 9.1.2.** *Sean* Z*, un sistema de* n *vectores de* $\mathcal{V}$*, el escalar* λ *y* $\boldsymbol{b}$ *un vector de* $\mathbb{R}^n$*; entonces*

$$\mathsf{Z}\big(\lambda\boldsymbol{b}\big)=\lambda\big(\mathsf{Z}\boldsymbol{b}\big).$$

Podemos extender aún más la notación matricial.

Sistema de n vectores de $\mathcal{V}$ por una matriz de $\mathbb{R}^{n\times p}$

Si Z es un sistema de n vectores y $\mathbf{A}$ una matriz de n filas y p columnas; entonces denotamos por $\mathsf{Z}\mathbf{A}$ al sistema de vectores

$$\mathsf{Z}\mathbf{A}=\Big[\mathsf{Z}\big(\mathbf{A}_{|1}\big);\ldots;\mathsf{Z}\big(\mathbf{A}_{|p}\big)\Big];$$

es decir, el vector $j$ésimo del sistema $\mathsf{Z}\mathbf{A}$ es

$$\big(\mathsf{Z}\mathbf{A}\big)_{|j}=\mathsf{Z}\big(\mathbf{A}_{|j}\big);$$

(compare esta expresión con la Definición 3.1). Así que podemos escribir sin ambigüedad el vector $\mathsf{Z}\mathbf{A}_{|j}$.

Continuando con el ejemplo de más arriba

Librería NAcAL para Python

```
D = Matrix([ Vector([1,0,0]), Vector([-1,0,1]) ])
Z*D                                   # Sistema por Matriz (Sistema de Matrices)
```

$$\mathsf{Z}\mathbf{D} \;=\; \left[\begin{bmatrix}1&1&1\\1&0&0\end{bmatrix};\begin{bmatrix}-1&1&1\\0&1&0\end{bmatrix};\begin{bmatrix}2&0&1\\0&0&1\end{bmatrix};\right]\begin{bmatrix}1&-1\\0&0\\0&1\end{bmatrix} \;=\; \left[\begin{bmatrix}1&1&1\\1&0&0\end{bmatrix};\begin{bmatrix}1&-1&0\\-1&0&1\end{bmatrix};\right].$$

Además de las propiedades de linealidad de más arriba, también se verifica la asociatividad con el producto de matrices: $\mathsf{Z}(\mathbf{B}\boldsymbol{c})=(\mathsf{Z}\mathbf{B})\boldsymbol{c}$ 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.** *Sean* 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.** *Sean* Z *un sistema de* p *vectores de* $\mathcal{V}$*, y* $\underset{p\times q}{\mathbf{B}}$ *y* $\underset{q\times n}{\mathbf{C}}$*, entonces* $\quad \mathsf{Z}\big(\mathbf{B}\mathbf{C}\big) = \big(\mathsf{Z}\mathbf{B}\big)\mathbf{C}.$

Por lo que 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 logrado algunos resultados sobre combinaciones lineales de sistemas de vectores de $\mathcal{V}$ que nos 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 $\mathbf{A}$ como el conjunto de todas las combinaciones lineales de sus columnas:

$$\mathcal{C}\left(\mathbf{A}\right) = \left\{ \boldsymbol{y}\in\mathbb{R}^n \;\middle|\; \text{existen } x_1,\ldots,x_n\in\mathbb{R} \text{ tales que } \boldsymbol{y} = \big(\mathbf{A}_{|1}\big)x_1 + \cdots + \big(\mathbf{A}_{|n}\big)x_n \right\}.$$

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 9.1 ($\mathcal{L}(\mathsf{Z})$: subespacio engendrado por un sistema Z)**.** *Dado un sistema* $\mathsf{Z} = \big[\vec{z}_1;\ldots;\vec{z}_n\big]$*, 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 9.2 (Sistema generador de $\mathcal{V}$)**.** *Decimos que* Z *es un* sistema generador *de* $\mathcal{V}$ *si* $\mathcal{V} = \mathcal{L}(\mathsf{Z})$.

Ejemplo 15.

- Las columnas de $\mathbf{A}$ son un *sistema generador* de $\mathcal{C}\left(\mathbf{A}\right)$.

Librería 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
```

$$\left\{ \boldsymbol{v}\in\mathbb{R}^3 \;\middle|\; \exists\boldsymbol{p}\in\mathbb{R}^2 \text{ tal que } \boldsymbol{v} = \begin{bmatrix}1&0\\0&0\\0&1\end{bmatrix}\boldsymbol{p} \right\} \;=\; \left\{\boldsymbol{v}\in\mathbb{R}^3 \;\middle|\; \begin{bmatrix}0&1&0\end{bmatrix}\boldsymbol{v} = \boldsymbol{0}\right\}$$

(*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 $\mathcal{N}(\mathbf{A})$.

Librería NAcAL para Python

```
C = Matrix([ [1,0,0], [1,1,1] ])
Homogenea(C,1).sgen    # Sistema generador del espacio nulo de C
```

$$\left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 1 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow[{[(-1)\mathbf{2}+\mathbf{3}]}]{\boldsymbol{\tau}} \left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \hline 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}; \right]$$

Librería NAcAL para Python

```
Homogenea(C).enulo   # Espacio nulo de C
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^1 \text{ tal que } \boldsymbol{v} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \boldsymbol{v} = \mathbf{0} \right\}$$

☞ Un sistema de vectores Z es un sistema generador de $\mathcal{V}$ si: $\mathcal{V} = \mathcal{L}(\mathsf{Z})$

$\mathcal{L}(\mathsf{Z})$ es el conjunto de combinaciones lineales de los vectores de Z; y puesto que

- la suma de dos combinaciones lineales $\mathsf{Z}\boldsymbol{b} + \mathsf{Z}\boldsymbol{c}$ es la combinación lineal $\mathsf{Z}(\boldsymbol{b} + \boldsymbol{c})$
- el producto de una combinación lineal por un escalar $(\mathsf{Z}\boldsymbol{b})\lambda$ es la combinación lineal $\mathsf{Z}(\boldsymbol{b}\lambda)$

$\mathcal{L}(\mathsf{Z})$ *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 $\mathsf{Y} = [\vec{y}_1; \ldots; \vec{y}_n]$, 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 $\vec{y}_1, \ldots, \vec{y}_n$.

Ahora ya podemos establecer un criterio para saber si dos sistemas Y y Z generan el mismo subespacio:

Sea 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* Y *con* n *vectores y* Z *con* k *vectores; entonces* $\mathcal{L}(\mathsf{Y}) = \mathcal{L}(\mathsf{Z})$ *si y sólo si*

$$\mathsf{Y}_{|i} \in \mathcal{L}(\mathsf{Z}) \textit{ para todo } i = 1:n \qquad y \qquad \mathsf{Z}_{|j} \in \mathcal{L}(\mathsf{Y}) \textit{ para todo } j = 1:k.$$

Definición 9.3. *Diremos que dos sistemas generadores son* equivalentes *si generan el mismo espacio.*

Librería 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

$$\mathsf{Z} \;=\; \left[\vec{z}_1;\ldots;\vec{z}_i;\ldots;\vec{z}_j;\ldots;\vec{z}_n\right] \xrightarrow[{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}]{\boldsymbol{\tau}} \left[\vec{z}_1;\ldots;(\alpha\vec{z}_i+\vec{z}_j);\ldots;\vec{z}_j;\ldots;\vec{z}_n\right] \;=\; \mathsf{Z}_{\underset{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}} \;=\; \mathsf{Z}\Big(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\Big)$$

- *(Transformación* Tipo II *sobre el sistema de vectores)* Si $\alpha \neq 0$ entonces

$$\mathsf{Z} \;=\; \left[\vec{z}_1;\ldots;\vec{z}_i;\ldots;\vec{z}_n\right] \xrightarrow[{[(\alpha)\boldsymbol{i}]}]{\boldsymbol{\tau}} \left[\vec{z}_1;\ldots;(\alpha\vec{z}_i);\ldots;\vec{z}_n\right] \;=\; \mathsf{Z}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}} \;=\; \mathsf{Z}\Big(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}\Big)$$

Sistemas equivalentes

Así, si mediante una sucesión de transformaciones elementales $\boldsymbol{\tau}_1,\ldots,\boldsymbol{\tau}_k$ transformamos Z en un nuevo sistema $\mathsf{Y} = \mathsf{Z}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$; por el Corolario 9.2.1 concluimos que ambos sistemas son equivalentes:

$$\mathcal{L}(\mathsf{Z}) \;=\; \mathcal{L}(\mathsf{Z}\mathbf{E}) \;=\; \mathcal{L}(\mathsf{Y}), \quad \text{donde } \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}; \tag{9.2}$$

puesto que, por una parte, $\mathsf{Y}_{|j} = \mathsf{Z}\mathbf{E}_{|j}$ y por otra $\mathsf{Z}_{|j} = \mathsf{Y}(\mathbf{E}^{-1})_{|j}$.

Consideremos ahora una transformación de naturaleza completamente distinta: *quitar vectores nulos de un sistema*:

- *(Quitando vectores nulos)* $\left[\vec{z}_1;\ \vec{0};\ldots;\vec{z}_i;\ \vec{0};\ldots;\vec{z}_n;\ \vec{0}\,\right] \xrightarrow{\text{Quitando vectores nulos}} \left[\vec{z}_1;\ldots;\vec{z}_i;\ldots;\vec{z}_n\right].$

Por el Corolario 9.2.1

$$\mathcal{L}\Big(\left[\vec{z}_1;\ \vec{0};\ldots;\vec{z}_i;\ \vec{0};\ldots;\vec{z}_n;\ \vec{0}\,\right]\Big) = \mathcal{L}\Big(\left[\vec{z}_1;\ldots;\vec{z}_i;\ldots;\vec{z}_n\right]\Big).$$

Es mas, podemos generalizar el resultado con la siguiente

Proposición 9.3.1. *Si* $\mathsf{Z} = \left[\vec{z}_1;\ldots;\vec{z}_n\right]$ *es un sistema de vectores entonces*

$$\mathcal{L}\Big(\left[\vec{z}_1;\ldots;\vec{z}_{i-1};\ \vec{z}_i;\ \vec{z}_{i+1},\ldots\vec{z}_n\right]\Big) = \mathcal{L}\Big(\left[\vec{z}_1;\ldots;\vec{z}_{i-1};\ \vec{z}_{i+1};\ldots;\vec{z}_{n-1}\right]\Big).$$

si, y solo si, $\vec{z}_i$ *es combinación lineal del resto de vectores de* 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 $\mathbf{L}$ es una forma escalonada de $\mathbf{A}$ con r columnas no nulas (rango $(\mathbf{A}) = r$), entonces

$$\mathcal{C}\left(\mathbf{A}\right) = \mathcal{C}\left(\mathbf{L}\right) = \mathcal{C}\left(\mathbf{L}_{|(1:r)}\right),$$

donde $\mathbf{L}_{|(1:r)}$ es la submatriz correspondiente a las r primeras columnas (véase Apéndice 3.A en la página 33).

Por tanto, las columnas no nulas de cualquier forma pre-escalonada de $\mathbf{A}$ generan $\mathcal{C}\left(\mathbf{A}\right)$.

Observación. Como $\vec{0} \in \mathcal{V}$, y puesto que $\mathcal{L}\Big(\left[\vec{0};\ \vec{0};\ \vec{0}\right]\Big) = \mathcal{L}\Big(\left[\vec{0};\ \vec{0}\right]\Big) = \mathcal{L}\Big(\left[\vec{0}\right]\Big) =$ ¿...?, convenimos que $\mathcal{L}\Big(\left[\;\right]\Big) = \{\vec{0}\}$; es decir, admitimos el sistema vacío como sistema generador del subespacio $\{\vec{0}\}$.

9.4. Sistemas linealmente dependientes y sistemas linealmente independientes

EJERCICIO 46. Demuestre la siguiente...

Proposición 9.4.1. *Dado un sistema de vectores* $\mathsf{Z} = \left[\vec{z}_1; \ldots; \vec{z}_n\right]$*, existe* $i \in \{1, \ldots, n\}$ *tal que* z_i *es combinación lineal del resto de vectores de* Z *si, y solo si, existe* $\boldsymbol{a} \neq \boldsymbol{0}$ *en* $\mathbb{R}^n$ *tal que* $\mathsf{Z}\boldsymbol{a} = \vec{0}$.

La anterior proposición da lugar a la definición de *sistema linealmente dependiente* y de *sistema linealmente independiente*:

Definición 9.4. *Diremos que el sistema de* n *vectores,* Z

- *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, es decir,* $\mathsf{Z}\boldsymbol{a} = \vec{0}$ *si y solo si* $\boldsymbol{a} = \boldsymbol{0}$.

Definición 9.5. *Sea* Z *de* n *vectores de* $\mathcal{V}$*. Denominamos espacio nulo de* Z (que denotamos con $\mathcal{N}(\mathsf{Z})$) *al subconjunto de vectores* $\boldsymbol{x}$ *de* $\mathbb{R}^n$ *tales que* $\mathsf{Z}\boldsymbol{x} = \vec{0}$*, es decir*

$$\mathcal{N}(\mathsf{Z}) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \mathsf{Z}\boldsymbol{x} = \vec{0}\}.$$

Por tanto Z es linealmente independiente si y solo si $\mathcal{N}(\mathsf{Z}) = \{\boldsymbol{0}\}$.

Ejemplos de sistemas linealmente independientes

- Las columnas de una matriz $\mathbf{A}$ *invertible* constituyen un sistema linealmente independiente, puesto que
$$\mathbf{A}\boldsymbol{x} = \boldsymbol{0} \quad \Leftrightarrow \quad \mathbf{A}^{-1}\mathbf{A}\boldsymbol{x} = \mathbf{A}^{-1}\boldsymbol{0} \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{0}.$$
- Si, de un sistema Z de vectores linealmente independiente, quitamos 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) 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:

Proposición 9.4.2. *Para un sistema* $\mathbf{A} = \left[\boldsymbol{v}_1; \ldots; \boldsymbol{v}_n\right]$ *de* n *vectores de* $\mathbb{R}^m$*, las siguientes propiedades son equivalentes:*

- *El rango de* $\mathbf{A} = \left[\boldsymbol{v}_1; \ldots; \boldsymbol{v}_n\right]$ *es* n
- $\mathcal{N}(\mathbf{A}) = \{\boldsymbol{0}\}$
- *La combinación lineal* $\mathbf{A}\boldsymbol{x}$ *es* $\boldsymbol{0}$ *si y solo si* $\boldsymbol{x} = \boldsymbol{0}$
- *Las columnas de* $\mathbf{A}$ *son linealmente independientes*

Demostración. ¡Son propiedades que ya hemos demostrado anteriormente! □

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, al retirar cualquiera de los vectores se reduce el espacio engendrado por ellos. Un sistema de vectores que tenga el tamaño justo para generar el subespacio $\mathcal{V}$ sin que sobre ningún vector tiene un nombre especial:

Definición 9.6. *Diremos que el sistema* B *es* una base *de un subespacio* $\mathcal{V}$ *si simultáneamente*

1. B *genera el subespacio, es decir, si* $\mathcal{L}(\mathsf{B}) = \mathcal{V}$.
2. B *es linealmente independiente, es decir,* $\mathsf{B}\boldsymbol{x} = \vec{0}$ *si y solo si* $\boldsymbol{x} = \mathbf{0}$.

Es decir, una base de $\mathcal{V}$ es un sistema de vectores B con un número suficiente elevado de vectores como para generar el subespacio $\mathcal{V}$ pero, simultáneamente, suficientemente reducido como para que el sistema sea linealmente independiente.

Ejemplo 16. Las soluciones *especiales* encontradas al resolver $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ constituyen una base de $\mathcal{N}(\mathbf{A})$.

Justificación: Por la Proposición 7.2.3 en la página 91 sabemos que las soluciones especiales son un sistema generador de $\mathcal{N}(\mathbf{A})$ y, puesto que son una selección de columnas una matriz invertible, son linealemente independientes.

Un resultado importantísimo 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* X *y* Z *son dos sistemas con* p *y* q *vectores de* $\mathcal{V}$ *respectivamente, donde* $p > q$*; tales que* $\mathcal{L}(\mathsf{X}) \subset \mathcal{L}(\mathsf{Z})$ *entonces* X *es linealmente dependiente.*

Demostración. Como cada $\mathsf{X}_{|i}$ es combinación lineal de los vectores del sistema 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 tuviera más que la otra, por la Proposición 9.5.1, sería un sistema dependiente, y por tanto ya no sería una base.

Definición 9.7. *Decimos que un subespacio* $\mathcal{V}$ *es de dimensión finita si 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* $[\vec{z}_1;\ldots;\vec{z}_n]$ *es un sistema linealmente independiente y* $\vec{z}_{n+1} \notin \mathcal{L}\Big([\vec{z}_1;\ldots;\vec{z}_n]\Big)$ *entonces* $[\vec{z}_1;\ldots;\vec{z}_n;\ \vec{z}_{n+1}]$ *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 $\vec{z}_{n+1} \notin \mathcal{L}\Big([\vec{z}_1;\ldots;\vec{z}_n]\Big)$, necesariamente $a_{n+1} = 0$. Y como $[\vec{z}_1;\ldots;\vec{z}_n]$ es un sistema linealmente independiente, entonces también $a_1,\ldots,a_n = 0$. □

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 $\mathcal{V}$, entonces sabemos que existen sistemas linealmente independientes formados por vectores de $\mathcal{W}$. Por ejemplo el vacío. Por otra parte, cualquier sistema de vectores linealmente independientes tiene como mucho tantos vectores como la dimensión del espacio. Tomemos de todos los posibles sistemas linealmente independientes de $\mathcal{W}$, un sistema M con el mayor número posible de vectores. Entonces $\mathcal{L}(\mathsf{M}) \subset \mathcal{W}$. Veamos que si suponemos que existe $\vec{v} \in \mathcal{W}$ tal que $\vec{v} \in \mathcal{L}(\mathsf{M})$, necesariamente llegamos a una contradicción. Basta añadir $\vec{v}$ al sistema M para obtener un sistema linealmente independiente formado por vectores de $\mathcal{W}$. Pero esto contradice que 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 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 B y obtendríamos un sistema linealmente independiente con más elementos que la dimensión de $\mathcal{V}$. □

Corolario 9.5.5. *Cualquier sistema B de vectores de $\mathcal{W}$ linealmente independiente cuyo número de vectores coincida con la dimensión de $\mathcal{W}$ es necesariamente es una base.*

☞ En la siguiente sección diseñaremos un algoritmo para encontrar una base de $\mathcal{L}(\mathsf{Z})$. La mecánica es sencilla y se basa en que $\mathcal{L}(\mathsf{Z}) = \mathcal{L}(\mathsf{Z}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k})$ (Ecuación 9.2).

El método consiste en aplicar la eliminación de "izquierda a derecha" sobre Z, y observar qué vectores de Z no se anulan. Aquellos vectores de Z que no se anulen constituyen una base de $\mathcal{L}(\mathsf{Z})$.

Aunque el procedimiento es muy sencillo, demostrar por qué funciona no será inmediato. Aprovecharemos para introducir notación y operaciones nuevas, pero a cambio lograremos una comprensión más profunda de las propiedades de los sistemas obtenidos en el proceso de eliminación ("de izquierda a derecha").

9.5.1. Eliminación "de izquierda a derecha" y sistemas "acoplados" de vectores

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

$$\mathsf{Y}_{|(1:s)} = \Big[\vec{y}_1; \dots; \vec{y}_s\Big] = \Big[\vec{y}_j \mid j < s\Big] \quad \text{con} \quad s \leq n,$$

y asumiremos que $\mathcal{L}(\mathsf{Z}_{|(1:0)}) = \mathcal{L}([\ ]) = \{\vec{0}\}$; donde $[\ ]$ es la lista vacía.

Definición 9.8 (Sistemas acoplados)**.** *Los sistemas Y y Z, de n vectores cada uno, están* acoplados *si*

$$\mathcal{L}\Big([\vec{y}_1; \dots; \vec{y}_i]\Big) = \mathcal{L}\Big([\vec{z}_1; \dots; \vec{z}_i]\Big) \quad \textit{para } i = 1:n,$$

es decir, si

$$\mathcal{L}\Big(\mathsf{Y}_{|(1:i)}\Big) = \mathcal{L}\Big(\mathsf{Z}_{|(1:i)}\Big) \quad \textit{para } i = 1:n.$$

Proposición 9.5.6 (Eliminación "de izquierda a derecha" en Z)**.** *Si transformamos un sistema de vectores $\mathsf{Z} = [\vec{z}_1; \dots; \vec{z}_n]$ mediante una secuencia $\boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_k$ de transformaciones elementales* "de izquierda a derecha"[2]

$$\mathsf{Z}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathsf{Z}\mathbf{E} = \mathsf{Y} \quad \textit{donde} \quad \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k},$$

de tal forma que cada $\mathsf{Y}_{|j}$ es de la forma $\mathsf{Z}\mathbf{E}_{|j}$ con $e_{jj} = 1$, es decir,

$$\mathsf{Y} = [\vec{y}_1; \dots; \vec{y}_n] = [\vec{z}_1; \dots; \vec{z}_n] \begin{bmatrix} e_{11} & e_{12} & e_{13} & \cdots & e_{1n} \\ & e_{22} & e_{23} & \cdots & e_{2n} \\ & & e_{33} & \cdots & e_{3n} \\ & & & \ddots & \vdots \\ & & & & e_{nn} \end{bmatrix} = \mathsf{Z}\mathbf{E},$$

entonces Y y Z están acoplados.

[2] del mismo modo que en Teorema 4.4.1 en la página 51

Demostración. Por una parte, como $\mathbf{E}$ es triangular superior sin elementos nulos en la diagonal principal, cada vector $\mathsf{Y}_{|j} = \mathsf{Z}\mathbf{E}_{|j}$ es un múltiplo no nulo ($e_{jj} \neq 0$) de $\mathsf{Z}_{|j}$ más una combinación lineal de los vectores $\mathsf{Z}_{|k}$ con $k = 1 : (j-1)$, y por tanto $\mathcal{L}(\mathsf{Y}_{|(1:j)}) \subset \mathcal{L}(\mathsf{Z}_{|(1:j)})$ para $j = 1 : n$.

Por otra parte, ya que la inversa de una matriz triangular superior es triangular superior (Sección 5.3), sabemos que $\mathbf{E}^{-1}$ es triangular superior sin elementos nulos en la diagonal, y como $\mathsf{Z} = \mathsf{Y}\mathbf{E}^{-1}$, usando el mismo argumento del párrafo anterior concluimos que $\mathcal{L}(\mathsf{Z}_{|(1:j)}) = \mathcal{L}(\mathsf{Y}_{|(1:j)})$ para $j = 1 : n$. $\square$

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* Y (de n vectores) y Z (de k vectores)

$$\mathsf{Y} \# \mathsf{Z} \;=\; [\vec{y}_1; \dots; \vec{y}_n;\; \vec{z}_1; \dots; \vec{z}_k].$$

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 $\{\vec{z}_1, \dots, \vec{z}_k\} \cup \{\vec{x}_1, \dots, \vec{x}_r\} \subset \mathcal{L}(\mathsf{Y} \# \mathsf{X})$. $\square$

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* 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 késimo 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 Z y 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 $k < i$

$$\begin{aligned}
\mathcal{L}\Big([\vec{y}_j \mid j \neq k]\Big) \;=\; \mathcal{L}\Big(\mathsf{Y}_{|(1:k-1)} \# \mathsf{Y}_{|(k+1:i)}\Big) \;&\overset{*}{=}\; \mathcal{L}\Big(\mathsf{Y}_{|(1:k)} \# \mathsf{Y}_{|(k+1:i)}\Big) = \mathcal{L}\Big(\mathsf{Y}_{|(1:i)}\Big) = \mathcal{L}\Big(\mathsf{Z}_{|(1:i)}\Big) \\
&=\; \mathcal{L}\Big(\mathsf{Z}_{|(1:k)} \# \mathsf{Z}_{|(k+1:i)}\Big) \\
&\overset{*}{=}\; \mathcal{L}\Big(\mathsf{Z}_{|(1:k-1)} \# \mathsf{Z}_{|(k+1:i)}\Big) \;=\; \mathcal{L}\Big([\vec{z}_j \mid j \neq k]\Big).
\end{aligned}$$

$\square$

Corolario 9.5.9. *Si* Y *y* 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 9.9. *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* $[\vec{y}_1;\ldots;\vec{y}_n]$ *y* $[\vec{z}_1;\ldots;\vec{z}_n]$ *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 quitar los pares superfluos correspondientes a los ceros en el sistema $[\vec{y}_1;\ldots;\vec{y}_n]$. □

☞ El anterior corolario nos muestra una forma de encontrar una base de $\mathcal{L}(\mathsf{Z})$ entre los vectores de 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

$$\left[\mathsf{Z}_{|j} \mid \mathsf{Z}\mathbf{E}_{|j} \neq \vec{0}\right], \quad \text{donde} \quad \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}.$$

Así pues, el siguiente es otro ejemplo de base de un subespacio:

Ejemplo 17. Las columnas no nulas de una forma escalonada de $\mathbf{A}$ constituyen una base de $\mathcal{C}(\mathbf{A})$.

Justificación: Sea $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ una forma escalonada de $\mathbf{A}$, por (9.2) sabemos que $\mathcal{L}(\mathbf{A}) = \mathcal{L}(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k})$; así, al selecionar las columnas no nulas de $\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ creamos un sistema linealmente independiente.

Encontrando dos bases de $\mathcal{C}(\mathbf{A})$ y una base de $\mathcal{N}(\mathbf{A})$

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 $\mathcal{C}(\mathbf{A})$ y una base de $\mathcal{N}(\mathbf{A})$:

$$\left[\frac{\mathbf{A}}{\mathbf{I}}\right] = \left[\begin{array}{cccc} 0&0&1&2\\ 1&2&1&3\\ 2&4&3&8\\ \hline 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0\\ 0&0&0&1 \end{array}\right] \xrightarrow[{[(-2)\mathbf{3}+\mathbf{4}]}]{\boldsymbol{\tau}} \left[\begin{array}{cccc} 0&0&1&0\\ 1&2&1&1\\ 2&4&3&2\\ \hline 1&0&0&0\\ 0&1&0&0\\ 0&0&1&-2\\ 0&0&0&1 \end{array}\right] \xrightarrow[{\substack{[(-2)\mathbf{1}+\mathbf{2}]\\ [(-1)\mathbf{1}+\mathbf{3}]\\ [(-1)\mathbf{1}+\mathbf{4}]}}]{\boldsymbol{\tau}} \left[\begin{array}{cccc} 0&0&1&0\\ 1&0&0&0\\ 2&0&1&0\\ \hline 1&-2&-1&-1\\ 0&1&0&0\\ 0&0&1&-2\\ 0&0&0&1 \end{array}\right] = \left[\frac{\mathbf{K}}{\mathbf{E}}\right];$$

por una parte, $\mathcal{C}(\mathbf{A}) = \mathcal{L}\Big(\big[\mathbf{K}_{|1};\ \mathbf{K}_{|3}\big]\Big)$ es de dimensión 2 por ser $\mathbf{K}_{|1}$ y $\mathbf{K}_{|3}$ linealmente independientes. Y por el Corolario 9.5.10 también sabemos que $\mathcal{C}(\mathbf{A}) = \mathcal{L}\Big(\big[\mathbf{A}_{|1};\ \mathbf{A}_{|3}\big]\Big)$. Como $\big[\mathbf{A}_{|1};\ \mathbf{A}_{|3}\big]$ tiene dos vectores, dicho sistema es otra base de $\mathcal{C}(\mathbf{A})$. Pero además ya vimos (Página 109) que las soluciones especiales $\mathbf{E}_{|2}$ y $\mathbf{E}_{|4}$ constituyen una base de $\mathcal{N}(\mathbf{A})$.

Así pues, es evidente que en este prodecimiento, el número de columnas que se anulan (que es la dimensión de $\mathcal{N}(\mathbf{A})$) más el número de columnas que no se anulan (que es la dimensión de $\mathcal{C}(\mathbf{A})$) es igual a número total de columnas n de $\mathbf{A}$. Por tanto podemos enunciar el siguiente

Corolario 9.5.11. $\dim\mathcal{C}(\mathbf{A}) + \dim\mathcal{N}(\mathbf{A})$ *es igual al número de columnas de* $\mathbf{A}$.

Librería 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
```

$$\left[\begin{pmatrix}0\\1\\2\end{pmatrix};\ \begin{pmatrix}1\\0\\1\end{pmatrix};\right]$$

Librería NAcAL para Python

```
SubEspacio( A.sis() ).dim    # dimensión del SubEspacio
```

```
2
```

Librería NAcAL para Python

```
Homogenea(A).enulo.dim     # dimensión del Espacio Nulo
```

```
2
```

Pero este resultado no se limita a sistemas de vectores de $\mathbb{R}^m$. De manera análoga, si mediante eliminación anulamos todos los vectores del sistema Z que son combinación lineal de los vectores que están a su izquierda: $\mathsf{Y} = \mathsf{Z}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$, donde $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$

$$\left[\begin{array}{c}\mathsf{Z}\\ \hline \mathbf{I}\end{array}\right] \xrightarrow{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \left[\begin{array}{c}\mathsf{Y}\\ \hline \mathbf{E}\end{array}\right] = \left[\begin{array}{cccccc} \vec{y}_1; & \vec{0}; & \vec{y}_3; & \vec{0}; & \cdots & \vec{y}_n;\\ \hline 1 & e_{12} & e_{13} & e_{14} & \cdots & e_{1n}\\ & 1 & e_{23} & e_{24} & \cdots & e_{2n}\\ & & 1 & e_{34} & \cdots & e_{3n}\\ & & & 1 & \cdots & e_{3n}\\ & & & & \cdots & \vdots\\ & & & & & 1\end{array}\right]; \tag{9.4}$$

el número total de vectores no anulados más el número total de soluciones especiales es igual al número de vectores del sistema Z. Por tanto, también es cierto el siguiente

Corolario 9.5.12. $\dim \mathcal{L}\big(\mathsf{Z}\big) + \dim \mathcal{N}\left(\mathsf{Z}\right)$ *es igual al número de vectores de* Z.

☞ Hemos visto que 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 Z, y la sucesión de transformaciones elementales $\boldsymbol{\tau}_1,\ldots,\boldsymbol{\tau}_k$ tal que $\mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ es triangular superior

$$\text{Si}\quad \mathsf{Y} = \mathsf{Z}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathsf{Z}\mathbf{E}, \quad \text{entonces } \mathsf{Y} \text{ y } \mathsf{Z} \text{ están acoplados.}$$

es decir,

$$\mathcal{L}\Big(\big[\mathsf{Z}_{|1};\ldots;\mathsf{Z}_{|j}\big]\Big) \;=\; \mathcal{L}\Big(\big[\mathsf{Y}_{|1};\ldots;\mathsf{Y}_{|j}\big]\Big) \quad \text{para } j = 1:n.$$

Pero si además algunos vectores de Y acaban siendo nulos, los correspondientes vectores de Z se pueden quitar y el sistema resultante continúa generando el mismo subespacio (Corolario 9.5.10).

Hemos usado esta idea para encontrar una base de $\mathcal{C}\left(\mathbf{A}\right)$ entre las columnas de $\mathbf{A}$ pre-escalonando la matriz mediante eliminación (de *"izquierda a derecha"*).

Recordando la Proposición 5.4.5, 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 de resultados previos

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 Z un sistema de n vectores de $\mathcal{V}$, denotaremos con f_{Z} a la función lineal:

$$\begin{aligned} f_{\mathsf{Z}}\colon \mathbb{R}^n &\to \mathcal{V} \\ \boldsymbol{x} &\to \mathsf{Z}\boldsymbol{x} \end{aligned}.$$

Pues para cualquier función *lineal* $f\colon \mathbb{R}^n \to \mathcal{V}$, existe 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.

$$\begin{aligned} f(\boldsymbol{x}) =& f(\mathbf{I}\boldsymbol{x}) & \\ =& f\big(x_1\mathbf{I}_{|1} + \cdots + x_n\mathbf{I}_{|n}\big) & \text{por la definición de matriz por vector de } \mathbb{R}^n \\ =& x_1 f(\mathbf{I}_{|1}) + \cdots + x_n f(\mathbf{I}_{|n}) & \text{por ser } f \text{ una función lineal} \\ =& \big[f(\mathbf{I}_{|1});\ \dots f(\mathbf{I}_{|n})\big]\boldsymbol{x} & \text{por la definición de sistema por vector de } \mathbb{R}^n \\ =& \mathsf{Z}\boldsymbol{x} & \text{donde } \mathsf{Z} = \big[f(\mathbf{I}_{|1});\ \dots f(\mathbf{I}_{|n})\big]. \end{aligned}$$

□

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:

$$[f_{\mathsf{Z}} \circ f_{\mathbf{B}}](\boldsymbol{c}) \;=\; f_{\mathsf{Z}}\big(f_{\mathbf{B}}(\boldsymbol{c})\big) \;=\; f_{\mathsf{Z}}\big(\mathbf{B}\boldsymbol{c}\big) \;=\; \mathsf{Z}\mathbf{B}\boldsymbol{c} \;=\; f_{\mathsf{Z}\mathbf{B}}(\boldsymbol{c});$$

es decir

$$f_{\mathsf{Z}\mathbf{B}}(\boldsymbol{c})\colon \mathbb{R}^n \xrightarrow{f_{\mathbf{B}}} \mathbb{R}^n \xrightarrow{f_{\mathsf{Z}}} \mathcal{V}$$

y por otra, si $\underset{n\times p}{\mathbf{B}}$ y $\underset{p\times q}{\mathbf{C}}$

$$[f_{\mathsf{Z}} \circ f_{\mathbf{B}}](\mathbf{C}) \;=\; f_{\mathsf{Z}}\big(f_{\mathbf{B}}(\mathbf{C})\big) \;=\; f_{\mathsf{Z}}\big(\mathbf{B}\mathbf{C}\big) \;=\; \mathsf{Z}\mathbf{B}\mathbf{C} \;=\; f_{\mathsf{Z}\mathbf{B}}(\mathbf{C});$$

es decir

$$f_{\mathsf{Z}\mathbf{B}}(\mathbf{C})\colon \mathbb{R}^{p\times q} \xrightarrow{f_{\mathbf{B}}} \mathbb{R}^{n\times q} \xrightarrow{f_{\mathsf{Z}}} \mathcal{V}^n$$

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 Z es la imagen de f_{Z}.

Ejercicio 48. Demuestre la siguiente

Proposición 9.6.1. *Sean las funciones* $f\colon \mathcal{V} \to \mathcal{W}$ *y* $g\colon \mathcal{W} \to \mathcal{Y}$, *entonces la imagen de* $g \circ f$ *está contenida en la imagen de* g.

Ejercicio 49. Demuestre el siguiente

Corolario 9.6.2. *Sean las funciones* $f\colon \mathcal{V} \to \mathcal{W}$ *y* $g\colon \mathcal{W} \to \mathcal{Y}$, *y donde* f *es invertible, entonces* *imagen*$(g \circ f) =$ *imagen*(g).

Sistemas linealmente independientes y coordenadas

Proposición 9.6.3. *Un sistema* Z *de* n *vectores es independiente si y solo si* f_{Z} *es invertible.*

Demostración. Si 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 Z es independiente, necesariamente $(\boldsymbol{x}-\boldsymbol{y}) = \mathbf{0}$; es decir $\boldsymbol{x} = \boldsymbol{y}$.

Supongamos 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 $\mathbf{0}$. $\square$

Cuando Z es independiente, la función inversa $(f_{\mathsf{Z}})^{-1}\colon \mathcal{L}\big(\mathsf{Z}\big) \to \mathbb{R}^n$ se conoce como *función de coordenadas respecto a* Z; que denotaremos como

$$\vec{v}_{/\mathsf{Z}} = (f_{\mathsf{Z}})^{-1}(\vec{v}) \qquad \Rightarrow \qquad -_{/\mathsf{Z}}\colon \mathcal{V} \to \mathbb{R}^n;$$

por tanto, $\vec{v}_{/\mathsf{Z}}$, es decir, las coordenadas de $\vec{v}$ respecto de Z, es el vector de $\mathbb{R}^n$ que nos indica la única combinación lineal de los vectores el sistema 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

$$\mathsf{Z}\big(\vec{v}_{/\mathsf{Z}}\big) = \vec{v}.$$

Como la función *función de coordenadas respecto a* 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 50. Demuestre la siguiente

Proposición 9.6.4. *Si* B *es una base de* $\mathcal{V}$*, entonces para cualquier función lineal* $f\colon \mathcal{V} \to \mathcal{W}$ *existe un sistema* Z *de vectores de* $\mathcal{W}$ *tal que* $f(\vec{v}) = \mathsf{Z}\big(\vec{v}_{/\mathsf{B}}\big)$

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. $\mathcal{A}+\mathcal{B}$ es no 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 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 $\Big(x(\vec{b}_1)+y(\vec{b}_2)\Big)\in\mathcal{B}$. □

Librería NAcAL para Python

```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)
```

$$\left\{\boldsymbol{v}\in\mathbb{R}^4\;\middle|\;\exists\boldsymbol{p}\in\mathbb{R}^3\text{ tal que }\boldsymbol{v}=\begin{bmatrix}1&0&0\\0&-1&0\\1&0&-1\\0&-1&0\end{bmatrix}\boldsymbol{p}\right\}\;=\;\left\{\boldsymbol{v}\in\mathbb{R}^4\;\middle|\;\begin{bmatrix}0&-1&0&1\end{bmatrix}\boldsymbol{v}=\boldsymbol{0}\right\}$$

Proposición 10.1.2. *Sean $\mathsf{A}=\left[\vec{a}_1;\ldots;\vec{a}_p;\right]$ y $\mathsf{B}=\left[\vec{b}_1;\ldots;\vec{b}_q;\right]$ dos sistemas de vectores de $\mathcal{V}$. Y sean un sistema $\left[\boldsymbol{x}_1;\ldots;\boldsymbol{x}_s;\right]$ de vectores de $\mathbb{R}^p$ y un sistema $\left[\boldsymbol{y}_1;\ldots;\boldsymbol{y}_s;\right]$ de vectores de $\mathbb{R}^q$; tales que*

$$\mathcal{L}\Big(\big[(\boldsymbol{x}_1\mathbin{\#}\boldsymbol{y}_1);\ldots(\boldsymbol{x}_s\mathbin{\#}\boldsymbol{y}_s);\big]\Big)=\mathcal{N}\,(\mathsf{A}\mathbin{\#}\mathsf{B})$$

donde $(\boldsymbol{x}_i\mathbin{\#}\boldsymbol{y}_i)$ es el vector de $\mathbb{R}^{p+q}$ resultante de concatenar los vectores $\boldsymbol{x}_i$ e $\boldsymbol{y}_i$, y donde $\mathsf{A}\mathbin{\#}\mathsf{B}$ es el sistema de $p+q$ vectores resultante de concatenar A y B. Entonces $\mathcal{L}\Big(\big[\mathsf{A}\boldsymbol{x}_1;\ \ldots\mathsf{A}\boldsymbol{x}_s\big]\Big)=\mathcal{L}\big(\mathsf{A}\big)\cap\mathcal{L}\big(\mathsf{B}\big)$.

Si además, A, B y $\big[(\boldsymbol{x}_1\mathbin{\#}\boldsymbol{y}_1);\ldots(\boldsymbol{x}_s\mathbin{\#}\boldsymbol{y}_s);\big]$ son linealmente independientes, también lo es $\big[\mathsf{A}\boldsymbol{x}_1;\ \ldots\mathsf{A}\boldsymbol{x}_s\big]$.

Demostración.

1. *Demostremos que $\mathsf{A}\boldsymbol{x}_i\in\mathcal{L}\big(\mathsf{A}\big)\cap\mathcal{L}\big(\mathsf{B}\big)$.*

Puesto que los vectores $(\boldsymbol{x}_i \# \boldsymbol{y}_i)$ pertenecen a $\mathcal{N}(\mathsf{A} \# \mathsf{B})$, es inmediato ver que

$$\left[\mathsf{A} \# \mathsf{B}\right]\begin{pmatrix}\boldsymbol{x}_i\\ \boldsymbol{y}_i\end{pmatrix} = \vec{0} \quad \Rightarrow \quad \mathsf{A}\boldsymbol{x}_i + \mathsf{B}\boldsymbol{y}_i = \vec{0}; \quad \Rightarrow \quad \mathsf{A}\boldsymbol{x}_i = -\mathsf{B}\boldsymbol{y}_i \quad \Rightarrow \quad \mathsf{A}\boldsymbol{x}_i \in \mathcal{L}(\mathsf{A}) \cap \mathcal{L}(\mathsf{B}),$$

pues $\mathsf{A}\boldsymbol{x}_i$ es también una combinación lineal de los vectores de B.

2. *Demostremos que si $\vec{v} \in \mathcal{L}(\mathsf{A}) \cap \mathcal{L}(\mathsf{B})$ entonces $\vec{v}$ es una combinación lineal de los vectores $\mathsf{A}\boldsymbol{x}_i$, es decir, $\vec{v} = \lambda_1\mathsf{A}\boldsymbol{x}_1 + \cdots + \lambda_s\mathsf{A}\boldsymbol{x}_s$.*

   Sea $\vec{v} \in \mathcal{L}(\mathsf{A}) \cap \mathcal{L}(\mathsf{B})$, 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 $(\boldsymbol{x} \# \boldsymbol{y}) \in \mathcal{N}(\mathsf{A} \# \mathsf{B})$ y por tanto $(\boldsymbol{x} \# \boldsymbol{y})$ es combinación lineal del sistema generador de $\mathcal{N}(\mathsf{A} \# \mathsf{B})$:

   $$(\boldsymbol{x} \# \boldsymbol{y}) = \lambda_1(\boldsymbol{x}_1 \# \boldsymbol{y}_1) + \cdots + \lambda_s(\boldsymbol{x}_s \# \boldsymbol{y}_s)$$

   consecuentemente $\boldsymbol{x} = \lambda_1\boldsymbol{x}_1 + \cdots + \lambda_s\boldsymbol{x}_s$; y por tanto $\vec{v} = \mathsf{A}\boldsymbol{x} = \lambda_1\mathsf{A}\boldsymbol{x}_1 + \cdots + \lambda_s\mathsf{A}\boldsymbol{x}_s$.

3. *Demostremos que si A, B y $\left[(\boldsymbol{x}_1 \# \boldsymbol{y}_1);\ldots(\boldsymbol{x}_s \# \boldsymbol{y}_s);\right]$ son linealmente independientes, también lo es $[\mathsf{A}\boldsymbol{x}_1;\ldots\ \mathsf{A}\boldsymbol{x}_s]$.*

   Supongamos que $\lambda_1\mathsf{A}\boldsymbol{x}_1 + \cdots + \lambda_s\mathsf{A}\boldsymbol{x}_s = \vec{0}$. En tal caso, por linealidad, $\mathsf{A}(\lambda_1\boldsymbol{x}_1 + \cdots + \lambda_s\boldsymbol{x}_s) = \vec{0}$, con lo que al ser A linealmente independiente $\lambda_1\boldsymbol{x}_1 + \cdots + \lambda_s\boldsymbol{x}_s = \mathbf{0}$. Por otra parte, como $\lambda_1(\boldsymbol{x}_1 \# \boldsymbol{y}_1) + \cdots + \lambda_s(\boldsymbol{x}_s \# \boldsymbol{y}_s)$ pertenece al conjunto $\mathcal{N}(\mathsf{A} \# \mathsf{B})$, tendremos que

   $$\vec{0} = \mathsf{A}(\lambda_1\boldsymbol{x}_1 + \cdots + \lambda_s\boldsymbol{x}_s) + \mathsf{B}(\lambda_1\boldsymbol{y}_1 + \cdots + \lambda_s\boldsymbol{y}_s) = \mathsf{B}(\lambda_1\boldsymbol{y}_1 + \cdots + \lambda_s\boldsymbol{y}_s)$$

   Por lo que al ser B linealmente independiente $\lambda_1\boldsymbol{y}_1 + \cdots + \lambda_s\boldsymbol{y}_s = \mathbf{0}$. Pero entonces tenemos que $\lambda_1(\boldsymbol{x}_1 \# \vec{y_1}) + \cdots + \lambda_s(\boldsymbol{x}_s;\boldsymbol{y}_s) = (\mathbf{0} \# \mathbf{0})$, con lo que al ser $\left[(\boldsymbol{x}_1 \# \boldsymbol{y}_1);\ldots(\boldsymbol{x}_s \# \boldsymbol{y}_s);\right]$ linealmente independiente $\lambda_1 = \cdots = \lambda_s = 0$.

□

Consecuentemente, si A y B son linealmente independientes y $\left[(\boldsymbol{x}_1 \# \boldsymbol{y}_1);\ldots(\boldsymbol{x}_s \# \boldsymbol{y}_s);\right]$ es una base de $\mathcal{N}(\mathsf{A} \# \mathsf{B})$, entonces $\left[\mathsf{A}\boldsymbol{x}_1;\ \ldots\mathsf{A}\boldsymbol{x}_s\right]$ es una base de $\mathcal{L}(\mathsf{A}) \cap \mathcal{L}(\mathsf{B})$.

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 de arriba sabemos que si tomamos una base A de $\mathcal{A}$ y una base B de $\mathcal{B}$, la dimensión de $\mathcal{L}(\mathsf{A}) \cap \mathcal{L}(\mathsf{B})$ es igual a la dimensión de $\mathcal{N}(\mathsf{A} \# \mathsf{B})$. Por tanto, todo se reduce a contar el número de "soluciones especiales" del sistema $(\mathsf{A} \# \mathsf{B})\boldsymbol{z} = \vec{0}$. De propina veremos que empleando ciertos "sub-vectores" de las soluciones "soluciones especiales" podemos construir una base de $\dim(\mathcal{A}\cap\mathcal{B})$:

Si aplicamos la eliminación sobre el sistema $\left(\mathsf{A} \# \mathsf{B}\right)$ mediante una secuencia $\boldsymbol{\tau}_1,\ldots,\boldsymbol{\tau}_m$ de transformaciones elementales *"de izquierda a derecha"* (como en la Proposición 9.5.6 en la página 110 o la Ecuación 9.4 en la página 113):

$$\left[\mathsf{A} \# \mathsf{B}\right]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_m} = \left[\mathsf{A} \# \mathsf{B}\right]\mathbf{E} \quad \text{donde} \quad \mathbf{E} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_m} \quad \text{es triangular superior con unos en la diagonal,}$$

observaremos el siguiente esquema

$$\left[\begin{array}{c}\left[\mathsf{A} \# \mathsf{B}\right]\\ \hline \mathbf{I}\end{array}\right]\mathbf{E} = \left[\begin{array}{c}\left[\mathsf{A} \# \mathsf{B}\right]\mathbf{E}\\ \hline \mathbf{E}\end{array}\right] = \left[\begin{array}{ccc}\cdots & \vec{0} & \cdots\\ \hline \cdots & \boldsymbol{x}_j & \cdots\\ \cdots & \boldsymbol{y}_j & \cdots\end{array}\right] \quad \text{con} \quad \mathbf{E}_{|j} = \begin{pmatrix}\boldsymbol{x}_j\\ \boldsymbol{y}_j\end{pmatrix};\ \Rightarrow\ \mathsf{A}\boldsymbol{x}_j + \mathsf{B}\boldsymbol{y}_j = \vec{0},$$

donde $\boldsymbol{x}_j$ contiene las p primeras componentes de la solución especial, e $\boldsymbol{y}_j$ contiene las q últimas.

Así, tras aplicar la eliminación al sistema $\mathsf{A} \# \mathsf{B}$, que engendra el subespacio suma $\mathcal{A}+\mathcal{B}$, habremos anulado todos los vectores que eran combinación lineal de los que les anteceden. Por tanto, los vectores no nulos de $\left[\mathsf{A} \mathbin{+\!\!+} \mathsf{B}\right]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_m}$ constituyen una base de $\mathcal{A}+\mathcal{B}$ (Corolario 9.5.10). Si tras la eliminación se han anulado s vectores de $\mathsf{A} \mathbin{+\!\!+} \mathsf{B}$, entonces $\boxed{\dim(\mathcal{A}+\mathcal{B}) = p+q-s}$.

Y como hemos encontrado s "soluciones especiales" sabemos que $\boxed{\dim \mathcal{A}\cap\mathcal{B} = s}$, y que $\left[\mathsf{A}\boldsymbol{x}_1;\ \dots \mathsf{A}\boldsymbol{x}_s\right]$ es una base de $\mathcal{L}\big(\mathsf{A}\big)\cap\mathcal{L}\big(\mathsf{B}\big)$ (Proposición 10.1.2).

□

Ejercicio 51. 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 $\dim(\mathcal{A}+\mathcal{B}) = \dim(\mathcal{A}) + \dim(\mathcal{B}) - \dim(\mathcal{A}\cap\mathcal{B})$:

Librería NAcAL para Python

```
print( (SubEspacio(A) + SubEspacio(B)).dim ) # dimensión SubEspacio suma
print(  SubEspacio(A).dim )                  # dimensión SubEspacio generado por A
print(  SubEspacio(B).dim )                  # dimensión SubEspacio generado por B
print( (SubEspacio(A) & SubEspacio(B)).dim ) # dimensión SubEspacio intersección
```

```
3
2
2
1
```

Suma directa y subespacios complementarios

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.

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

$$\vec{v}-\vec{v} = \big(\vec{a}+\vec{b}\big)-\big(\vec{a}_*+\vec{b}_*\big) = \underbrace{\big(\vec{a}-\vec{a}_*\big)}_{\in\mathcal{A}} - \underbrace{\big(\vec{b}+\vec{b}_*\big)}_{\in\mathcal{B}} = \vec{0} \qquad\Longrightarrow\qquad \big(\vec{a}-\vec{a}_*\big) = -\big(\vec{b}+\vec{b}_*\big)\in\mathcal{B}.$$

Es decir, $\big(\vec{a}-\vec{a}_*\big)\in\mathcal{A}\cap\mathcal{B} = \{\vec{0}\}$, y por tanto $\big(\vec{a}-\vec{a}_*\big) = \vec{0} = \big(\vec{b}-\vec{b}_*\big)$, es decir, $\vec{a} = \vec{a}_*$ y $\vec{b} = \vec{b}_*$. □

10.2. Los cuatro subespacios fundamentales de una matriz A

En esta sección vamos a verificar que toda matriz $\mathbf{A}$, de orden m por n, define una la descomposición en espacios suplementarios tanto de $\mathbb{R}^m$ como de $\mathbb{R}^n$. En particular

- $\mathbb{R}^n$ queda descompuesto en el *espacio nulo* $\mathcal{N}\left(\mathbf{A}\right)$ y $\mathcal{C}\left(\mathbf{A}^\intercal\right)$ (que denominamos *espacio fila*):

$$\boxed{\mathbb{R}^n = \mathcal{C}\left(\boldsymbol{\mathcal{N}}\left(\mathbf{A}\right)\right)\oplus\mathbf{A}^\intercal}.$$

- $\mathbb{R}^m$ queda descompuesto en el *espacio columna* $\mathcal{C}\left(\mathbf{A}\right)$ y $\mathcal{N}\left(\mathbf{A}^\intercal\right)$ (que denominamos *espacio nulo por la izquierda*).

$$\boxed{\mathbb{R}^m = \mathcal{C}(\mathbf{A}) \oplus \mathcal{N}\left(\mathbf{A}^\intercal\right)}.$$

Además, la dimensión de los espacios columna $\mathcal{C}(\mathbf{A})$ y fila $\mathcal{C}\left(\mathbf{A}^\intercal\right)$ es igual al rango de la matriz. 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:

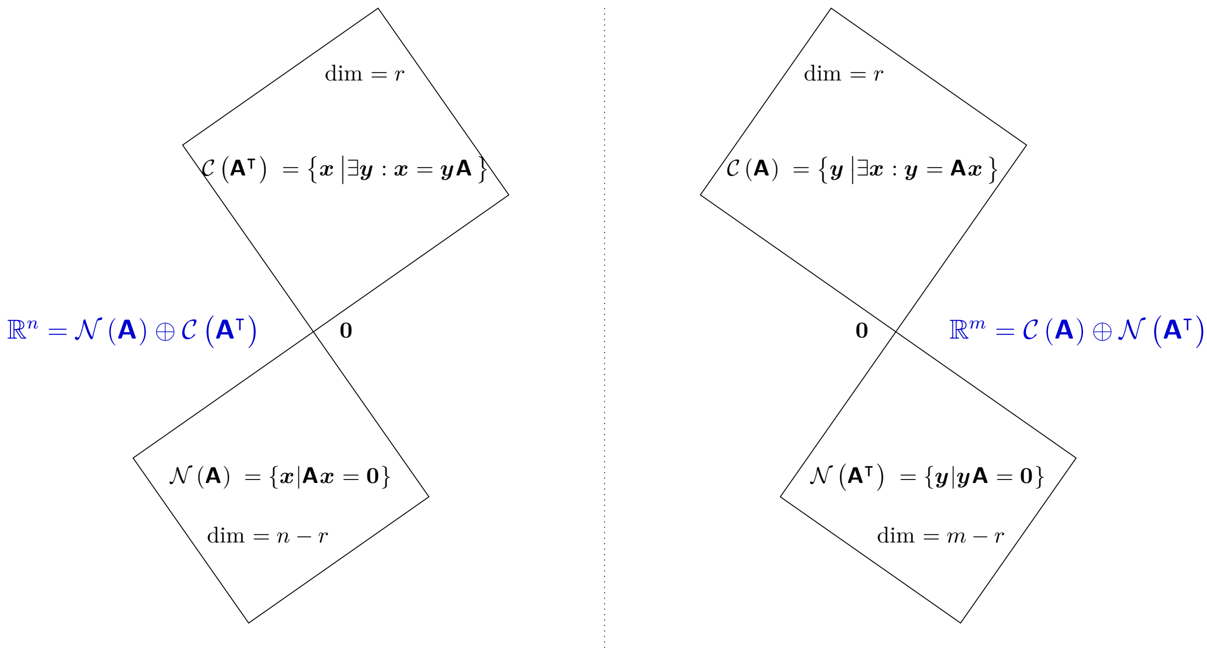


Figura 10.1: Esquema de descomposición de los cuatro subespacios fundamentales de una matriz (parte 1)

El profesor Strang (2007) se refiere a estos cuatro subespacios: $\mathcal{N}(\mathbf{A})$, $\mathcal{C}(\mathbf{A})$, $\mathcal{N}\left(\mathbf{A}^\intercal\right)$ y $\mathcal{C}\left(\mathbf{A}^\intercal\right)$, con el nombre "los cuatro subespacios fundamentales de una matriz" $\mathbf{A}$.

Fíjese que la notación indica que los nuevos espacios son los espacios columna y nulo de la matriz transpuesta. Veamos los dos nuevos espacios en detalle.

10.2.1. El espacio fila

Definimos el espacio fila de $\underset{m\times n}{\mathbf{A}}$ como 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 $\mathbb{R}^n$, pues 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

$$\begin{aligned}\mathcal{C}\left(\mathbf{A}^\intercal\right) =& \{\boldsymbol{x} \mid \text{existe } \boldsymbol{y} \in \mathbb{R}^m : \boldsymbol{x} = \big(\mathbf{A}^\intercal\big)\boldsymbol{y}\}\\ =& \{\boldsymbol{x} \mid \text{existe } \boldsymbol{y} \in \mathbb{R}^m : \boldsymbol{x} = \boldsymbol{y}\mathbf{A}\} \qquad\qquad \text{ya que } \big(\mathbf{A}^\intercal\big)\boldsymbol{y} = \boldsymbol{y}\mathbf{A};\end{aligned}$$

es decir, *el espacio fila* está formado por los vectores que se pueden describir como combinación lineal de las filas. Además, como el rango de una matriz no cambia al transponerla (Proposición 5.4.5), tenemos que

$$\dim \mathcal{C}(\mathbf{A}) \;=\; \dim \mathcal{C}\left(\mathbf{A}^\intercal\right) \;=\; \operatorname{rango}(\mathbf{A}).$$

Es decir, para una matriz $\mathbf{A}$ de orden $m\times n$ y rango r, tenemos que $\dim \mathcal{C}\left(\mathbf{A}^\intercal\right) = r$ y $\dim \mathcal{N}(\mathbf{A}) = n - r$.

Fíjese que $\dim \mathcal{C}\left(\mathbf{A}^{\intercal}\right) + \dim \mathcal{N}(\mathbf{A}) = n$, es decir, es igual a la dimensión de todo el espacio $\mathbb{R}^n$. Así, si probamos que $\mathcal{N}(\mathbf{A}) \cap \mathcal{C}\left(\mathbf{A}^{\intercal}\right) = \{\mathbf{0}\}$ (i.e., que la dimensión de la intersección es cero), por la Proposición 10.1.4 habremos demostrado que $\mathbb{R}^n = \mathcal{C}\left(\mathbf{A}^{\intercal}\right) \oplus \mathcal{N}(\mathbf{A})$:

Proposición 10.2.1. *Para toda* $\mathbf{A}$ *se verifica que* $\mathcal{N}(\mathbf{A}) \cap \mathcal{C}\left(\mathbf{A}^{\intercal}\right) = \{\mathbf{0}\}$.

Demostración. Para todo vector $\boldsymbol{x} \in \mathcal{C}\left(\mathbf{A}^{\intercal}\right)$ (es decir, de la forma $\boldsymbol{x} = \boldsymbol{z}\mathbf{A}$) y para todo vector $\boldsymbol{y} \in \mathcal{N}(\mathbf{A})$ (es decir, tal que $\mathbf{A}\boldsymbol{y} = \mathbf{0}$) el producto punto $\boldsymbol{x} \cdot \boldsymbol{y}$ es nulo:

$$\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{z}\mathbf{A}\boldsymbol{y} = \boldsymbol{z} \cdot \mathbf{0} = 0.$$

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 $\boldsymbol{v} \in \mathcal{N}(\mathbf{A}) \cap \mathcal{C}\left(\mathbf{A}^{\intercal}\right)$ entonces necesariamente $\boldsymbol{v} = \mathbf{0}$. □

Por tanto $\boxed{\mathbb{R}^n = \mathcal{C}\left(\mathbf{A}^{\intercal}\right) \oplus \mathcal{N}(\mathbf{A})}$.

10.2.2. El nulo por la izquierda

El subespacio $\mathcal{N}\left(\mathbf{A}^{\intercal}\right)$ es el conjunto de soluciones al sistema $(\mathbf{A}^{\intercal})\boldsymbol{x} = \mathbf{0}$, que nos pregunta ¿qué combinaciones lineales de las columnas de $\mathbf{A}^{\intercal}$ son iguales a cero? Pero esto es lo mismo que preguntar ¿qué combinaciones lineales de las *filas* de $\mathbf{A}$ son iguales a cero? Es decir, $\mathcal{N}\left(\mathbf{A}^{\intercal}\right)$ son las soluciones del siguiente *sistema de ecuaciones homogéneo*

$$\boldsymbol{x}\mathbf{A} = \mathbf{0}.$$

lo que justifica el nombre de *espacio nulo por la izquierda* (pues el vector de incógnitas $\boldsymbol{x}$ multiplica por la izquierda). Consecuentemente definimos el *espacio nulo por la izquierda* de $\mathbf{A}$ (de orden m por n) como

$$\mathcal{N}\left(\mathbf{A}^{\intercal}\right) \;=\; \{\boldsymbol{x} \mid \boldsymbol{x}\mathbf{A} = \mathbf{0}\}\,.$$

Si llamamos $\underset{n\times m}{\mathbf{B}}$ a $\mathbf{A}^{\intercal}$, por la Proposición 10.2.1 sabemos que $\mathbb{R}^m = \mathcal{C}\left(\mathbf{B}^{\intercal}\right) \oplus \mathcal{N}(\mathbf{B})$.

Por tanto $\boxed{\mathbb{R}^m = \mathcal{C}(\mathbf{A}) \oplus \mathcal{N}\left(\mathbf{A}^{\intercal}\right)}$.

Ejemplo 18. Por ejemplo, para la matriz $\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & -1 & -1 & 0 \end{bmatrix}$ tenemos que:

Librería NAcAL para Python

```
A = Matrix([ [0,1,1,0], [0,2,2,0],[0,-1,-1,0] ])
print( A.rg() )            # Rango de A
```

```
1
```

Librería NAcAL para Python

```
ColA = SubEspacio(A.sis()); # Espacio Columna de A (en R3)
print(ColA.dim);  ColA
```

```
1
```

$$\left\{\boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^1 \text{ tal que } \boldsymbol{v} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \boldsymbol{p} \right\} \;=\; \left\{\boldsymbol{v} \in \mathbb{R}^3 \;\middle|\; \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \boldsymbol{v} = \mathbf{0} \right\}$$

Librería NAcAL para Python

```
NulAT = SubEspacio( ~A );   # Espacio Nulo Izda de A  (en R3)
print(NulAT.dim);  NulAT
```

2

$$\left\{ \boldsymbol{v}\in\mathbb{R}^3 \;\middle|\; \exists\boldsymbol{p}\in\mathbb{R}^2 \text{ tal que } \boldsymbol{v}=\begin{bmatrix}-2&1\\1&0\\0&1\end{bmatrix}\boldsymbol{p} \right\} = \left\{ \boldsymbol{v}\in\mathbb{R}^3 \;\middle|\; \begin{bmatrix}-1&-2&1\end{bmatrix}\boldsymbol{v}=\boldsymbol{0} \right\}$$

Librería NAcAL para Python

```
FilA = SubEspacio( (~A).sis() ); # Espacio Fila de A (en R4)
print(FilA.dim);  FilA
```

1

$$\left\{ \boldsymbol{v}\in\mathbb{R}^4 \;\middle|\; \exists\boldsymbol{p}\in\mathbb{R}^1 \text{ tal que } \boldsymbol{v}=\begin{bmatrix}0\\1\\1\\0\end{bmatrix}\boldsymbol{p} \right\} = \left\{ \boldsymbol{v}\in\mathbb{R}^4 \;\middle|\; \begin{bmatrix}1&0&0&0\\0&-1&1&0\\0&0&0&1\end{bmatrix}\boldsymbol{v}=\boldsymbol{0} \right\}$$

Librería NAcAL para Python

```
NulA = SubEspacio( A );     # Espacio Nulo de A (en R4)
print(NulA.dim);  NulA
```

3

$$\left\{ \boldsymbol{v}\in\mathbb{R}^4 \;\middle|\; \exists\boldsymbol{p}\in\mathbb{R}^3 \text{ tal que } \boldsymbol{v}=\begin{bmatrix}1&0&0\\0&-1&0\\0&1&0\\0&0&1\end{bmatrix}\boldsymbol{p} \right\} = \left\{ \boldsymbol{v}\in\mathbb{R}^4 \;\middle|\; \begin{bmatrix}0&1&1&0\end{bmatrix}\boldsymbol{v}=\boldsymbol{0} \right\}$$

Librería 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*)

Librería 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* $\mathcal{C}(\mathbf{A})$ y *nulo* $\mathcal{N}(\mathbf{A})$. Por tanto, 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 esto supone 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 que aplicando la eliminación tan solo sobre las *columnas* de $\mathbf{A}$ encontremos bases para cada uno de los cuatro subespacios?… ¡Afortunadamente si!

La siguiente proposición nos da la pista necesaria para hacerlo.

Proposición 10.3.1. *Sean* $\mathbf{A}$ *de orden* m *por* n *y* $\mathbf{E}$ *invertible y de orden* n*, entonces* $\mathcal{N}\left(\mathbf{A}^\intercal\right) = \mathcal{N}\left((\mathbf{AE})^\intercal\right)$*, es decir, dado* $\boldsymbol{x} \in \mathbb{R}^m$ *se verifica que*

$$\boldsymbol{x}\mathbf{A} = \mathbf{0} \quad \Longleftrightarrow \quad \boldsymbol{x}\mathbf{AE} = \mathbf{0}.$$

Demostración. Veamos las implicaciones en uno y otro sentido.

- Si $\boldsymbol{x} \in \mathcal{N}\left(\mathbf{A}^\intercal\right) \implies \boldsymbol{x}\mathbf{A} = \mathbf{0} \implies \boldsymbol{x}\mathbf{AE} = \mathbf{0E} \implies \boldsymbol{x}(\mathbf{AE}) = \mathbf{0} \implies \boldsymbol{x} \in \mathcal{N}\left((\mathbf{AE})^\intercal\right)$
- Si $\boldsymbol{x} \in \mathcal{N}\left((\mathbf{AE})^\intercal\right) \implies \boldsymbol{x}(\mathbf{AE}) = \mathbf{0} \implies \boldsymbol{x}\mathbf{AEE}^{-1} = \mathbf{0E}^{-1} \implies \boldsymbol{x}\mathbf{A} = \mathbf{0} \implies \boldsymbol{x} \in \mathcal{N}\left(\mathbf{A}^\intercal\right).$

□

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 $\left[{}_{i_1|}\mathbf{A};\ \dots {}_{i_r|}\mathbf{A};\right]$ es linealmente independiente si y sólo si el sistema de vectores $\left[{}_{i_1|}\mathbf{AE};\ \dots {}_{i_r|}\mathbf{AE};\right]$ es linealmente independiente.
- la dimensión del espacio fila de $\mathbf{A}$ es igual a la dimensión del espacio fila de $\mathbf{AE}$.

Así pues, podemos bien pre-escalonar, bien escalonar o incluso escalonar y reducir una matriz para obtener otra matriz mucho más simple en la que buscar filas linealmente independientes.

El siguiente lema demuestra que las filas con pivote de cualquier forma escalonada $\mathbf{L}$ de $\mathbf{A}$ son linealmente independientes.

EJERCICIO 52. Demuestre que

Lema 10.3.2. *Si* $\mathbf{L}$ *está pre-escalonada, entonces 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, las filas con pivote de la forma pre-escalonada $\mathbf{K}$ forman una base del espacio fila de la matriz $\mathbf{K}$, pero no necesariamente forman una base del espacio fila de $\mathbf{A}$. Veamos cómo asegurarnos de escoger correctamente las filas 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 $\mathbf{S}$ una matriz que selecciona las filas de $\mathbf{AE}$ que forman una base del espacio fila de $\mathbf{AE}$; por ejemplo, las filas pivote:

$$\mathbf{S}\big(\mathbf{AE}\big) = \begin{bmatrix} 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&0&1&0 \end{bmatrix} \begin{bmatrix} 1&0&0&0\\ -2&-4&0&0\\ 4&4&0&0\\ -1&-1&1&0\\ -6&-7&3&0 \end{bmatrix} = \begin{bmatrix} 1&0&0&0\\ -2&-4&0&0\\ -1&-1&1&0 \end{bmatrix}.$$

Como $\mathbf{S}$ selecciona una base del espacio fila, entonces para cualquier vector $\boldsymbol{u}$ existe un único vector $\boldsymbol{x}$ (una única combinación lineal de las filas) tal que

$$\boldsymbol{x}\mathbf{SAE} = \boldsymbol{u}\mathbf{AE}.$$

Pero entonces también la ecuación $\boldsymbol{x}\mathbf{SA} = \boldsymbol{u}\mathbf{A}$ tiene una única solución ya que por ser $\mathbf{E}$ invertible

$$\boldsymbol{x}\mathbf{SA} = \boldsymbol{u}\mathbf{A} \Longleftrightarrow \boldsymbol{x}\mathbf{SAE} = \boldsymbol{u}\mathbf{AE}.$$

Luego $\mathbf{SA}$ (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 $\mathcal{C}\left(\mathbf{A}^{\intercal}\right)$**

- Pre-escalonamos $\mathbf{A}$
- Identificamos las filas con pivote de la forma escalonada
- Las correspondientes filas de $\mathbf{A}$ forman una base $\mathcal{C}\left(\mathbf{A}^{\intercal}\right)$

Ejemplo 19.

$$\left[\frac{\mathbf{A}}{\mathbf{I}}\right]=\left[\begin{array}{cccc}1&-2&-1&-1\\-2&0&1&2\\4&-4&-3&-4\\-1&1&1&1\\-6&5&5&6\\\hline 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right]\xrightarrow{\substack{\boldsymbol{\tau}\\ [(2)\mathbf{1}+\mathbf{2}]\\ [(1)\mathbf{1}+\mathbf{3}]\\ [(1)\mathbf{1}+\mathbf{4}]}}\left[\begin{array}{cccc}1&0&0&0\\-2&-4&-1&0\\4&4&1&0\\-1&-1&0&0\\-6&-7&-1&0\\\hline 1&2&1&1\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right]\xrightarrow{\substack{\boldsymbol{\tau}\\ [(4)\mathbf{3}]\\ [(-1)\mathbf{2}+\mathbf{3}]}}\left[\begin{array}{cccc}1&0&0&0\\-2&-4&0&0\\4&4&0&0\\-1&-1&1&0\\-6&-7&3&0\\\hline 1&2&2&1\\0&1&-1&0\\0&0&4&0\\0&0&0&1\end{array}\right]=\left[\frac{\mathbf{K}}{\mathbf{E}}\right]$$

De este proceso de eliminación, deducimos que:

$$\mathcal{C}(\mathbf{A})=\mathcal{L}\Big(\left[\mathbf{K}_{|1};\ \mathbf{K}_{|2};\ \mathbf{K}_{|3}\right]\Big),\qquad \mathcal{C}\left(\mathbf{A}^{\intercal}\right)=\mathcal{L}\Big(\left[{}_{1|}\mathbf{A};\ {}_{2|}\mathbf{A};\ {}_{4|}\mathbf{A}\right]\Big)\quad \text{y}\quad \mathcal{N}(\mathbf{A})=\mathcal{L}\Big(\left[\mathbf{E}_{|4}\right]\Big).$$

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 mirar ahí qué combinaciones de las filas son nulas. Por tanto aplicaremos la eliminación Gauss-Jordan hasta llegar a $\mathbf{R}$ (es decir, la forma escalonada *reducida*, donde los pivotes son unos con ceros a derecha e izquierda). Continuando el ejemplo anterior:

$$\mathbf{A}=\begin{bmatrix}1&-2&-1&-1\\-2&0&1&2\\4&-4&-3&-4\\-1&1&1&1\\-6&5&5&6\end{bmatrix}\xrightarrow{\substack{\boldsymbol{\tau}\\ [(2)\mathbf{1}+\mathbf{2}]\\ [(1)\mathbf{1}+\mathbf{3}]\\ [(1)\mathbf{1}+\mathbf{4}]\\ [(4)\mathbf{3}]\\ [(-1)\mathbf{2}+\mathbf{3}]}}\begin{bmatrix}1&0&0&0\\-2&-4&0&0\\4&4&0&0\\-1&-1&1&0\\-6&-7&3&0\end{bmatrix}\xrightarrow{\substack{\boldsymbol{\tau}\\ [(2)\mathbf{1}]\\ [(-1)\mathbf{2}+\mathbf{1}]}}\begin{bmatrix}2&0&0&0\\0&-4&0&0\\4&4&0&0\\-1&-1&1&0\\-5&-7&3&0\end{bmatrix}$$

$$\xrightarrow{\substack{\boldsymbol{\tau}\\ [(1)\mathbf{3}+\mathbf{1}]\\ [(1)\mathbf{3}+\mathbf{2}]}}\begin{bmatrix}2&0&0&0\\0&-4&0&0\\4&4&0&0\\0&0&1&0\\-2&-4&3&0\end{bmatrix}\xrightarrow{\substack{\boldsymbol{\tau}\\ [(\frac{1}{2})\mathbf{1}]\\ \left[\left(\frac{-1}{4}\right)\mathbf{2}\right]}}\begin{bmatrix}1&0&0&0\\0&1&0&0\\2&-1&0&0\\0&0&1&0\\-1&1&3&0\end{bmatrix}=\mathbf{R}.$$

Ahora es inmediato ver en $\mathbf{R}$ que cada fila sin pivote es combinación de las filas con pivote que la anteceden:

$$\begin{cases}{}_{3|}\mathbf{R}=2\big({}_{1|}\mathbf{R}\big)-1\big({}_{2|}\mathbf{R}\big) & \Rightarrow\quad (-2,\ \ 1,\ \ 1,\ \ 0,\ \ 0)\mathbf{R}=\mathbf{0}\\ \qquad\text{y} & \\ {}_{5|}\mathbf{R}=-1\big({}_{1|}\mathbf{R}\big)+1\big({}_{2|}\mathbf{R}\big)+3\big({}_{3|}\mathbf{R}\big) & \Rightarrow\quad (1,\ \ -1,\ \ 0,\ \ -3,\ \ 1)\mathbf{R}=\mathbf{0}\end{cases}$$

Y las mismas relaciones se verifican en la matriz original $\mathbf{A}$;

$$\begin{cases} {}_{3|}\mathbf{A} = 2\big({}_{1|}\mathbf{A}\big) - 1\big({}_{2|}\mathbf{A}\big) & \Rightarrow \quad (-2, \quad 1, \quad 1, \quad 0, \quad 0)\mathbf{A} = \mathbf{0} \\ \quad\text{y} & \\ {}_{5|}\mathbf{A} = -1\big({}_{1|}\mathbf{A}\big) + 1\big({}_{2|}\mathbf{A}\big) + 3\big({}_{3|}\mathbf{A}\big) & \Rightarrow \quad (1, \quad -1, \quad 0, \quad -3, \quad 1)\mathbf{A} = \mathbf{0} \end{cases}$$

☞ En el proceso de eliminación $\mathbf{A} \to \mathbf{R}$ “se ven” bases para los cuatro subespacios fundamentales de $\mathbf{A}$

Parte IV

Ortogonalidad (Espacio Euclideo)

Vectores ortogonales. Subespacios ortogonales

Una de las aplicaciones más importantes de las matemáticas es la medición tanto de distancias como de ángulos. El espacio euclideo nos da las herramientas necesarias para medir.

11.1. 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.

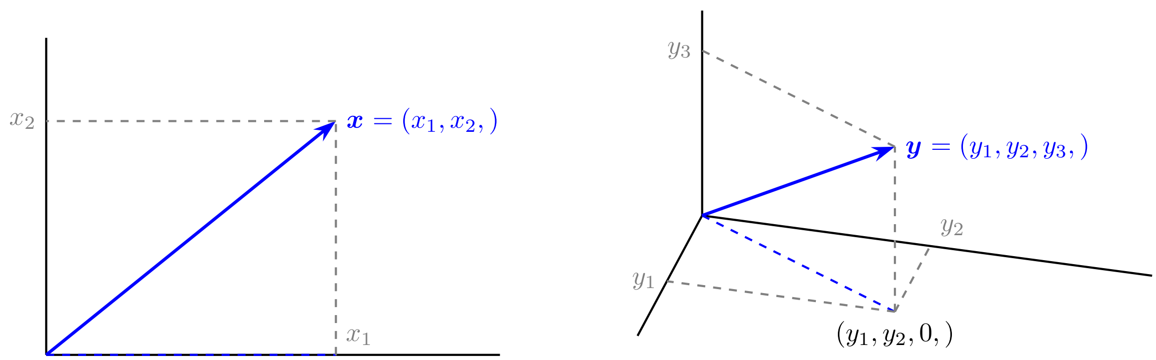


Figura 11.1: Longitud de un vector en $\mathbb{R}^2$ y en $\mathbb{R}^3$

Así, el cuadrado de la longitud de la hipotenusa (el cuadrado de la longitud del vector) es igual a la suma del cuadrado de las longitudes de los catetos.

En el caso de $\mathbb{R}^2$, la longitud de cada cateto es el valor de cada una de las dos componentes. Por tanto

$$\|\boldsymbol{x}\|^2 = x_1^2 + x_2^2 \qquad \Rightarrow \qquad \|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2};$$

donde $\|\boldsymbol{x}\|$ denota la longitud o *norma* de un vector $\boldsymbol{x}$.

En el caso de $\mathbb{R}^3$, la longitud del cateto vertical es el valor de la tercera componente del vector, pero la longitd del cateto horizontal ha de ser calculada. Como el cateto horizontal es la hipotenusa de de un triángulo rectángulo semejante al de la figura en $\mathbb{R}^2$, tenemos que el cuadrado de su longitud es $(y_1^2 + y_2^2)$; así

$$\|\boldsymbol{y}\|^2 = (y_1^2 + y_2^2) + y_3^2 \qquad \Rightarrow \qquad \|\boldsymbol{y}\| = \sqrt{y_1^2 + y_2^2 + y_3^2}.$$

Es decir, tanto en $\mathbb{R}^2$ como en $\mathbb{R}^3$, la longitud de un vector es la raiz cuadrada de la suma de los cuadrados de los componentes del vector.

11.2. Ángulo formado por dos vectores tanto en $\mathbb{R}^2$ como en $\mathbb{R}^3$

Un modo de indicar el angulo del vértice formado por dos segmentos es mediante la función *coseno*. Dicha función toma valores entre 1 y -1: cuando el ángulo θ es cero el coseno toma el valor 1 (marcado con un pequeño rombo sobre el eje horizontal de la primera figura de la izquierda). A medida que el ángulo se va abriendo el valor del coseno va disminuyendo. Al alcanzar un ángulo recto (90 grados) el coseno toma el valor 0 (figura central), y conforme el ángulo sigue abriéndose el coseno continúa disminuyendo, hasta que el ángulo alcanza la apertura máxima de 180 grados, en cuyo caso la función coseno toma el mínimo valor posible, es decir el valor -1 (última figura de la derecha).

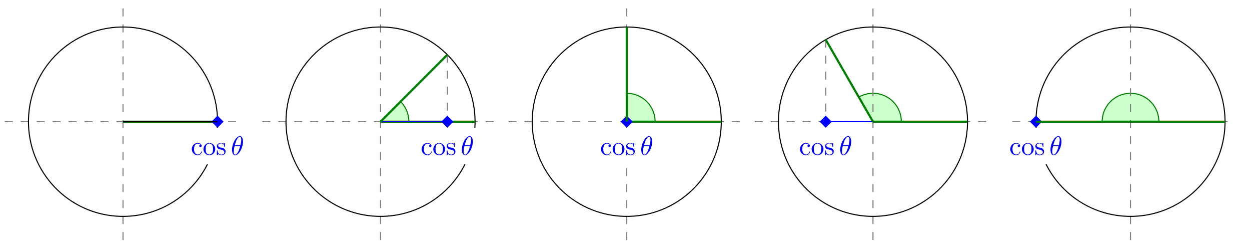


Figura 11.2: Relación entre el ángulo y su coseno (el círculo exterior es de radio uno)

Ahora piense que tiene tres listones de madera y que ninguno de ellos tiene una longitud mayor que la suma de los otros dos. Entonces con ellos se puede formar un triángulo único; es decir, los ángulos del triángulo quedan univocamente determinados por las longitudes de los lados (las longitudes de los listones de madera). Esto quiere decir que conociendo los lados de un triángulo podemos calcular sus ángulos. Para ello, de nuevo echaremos mano del teorema de Pitágoras.

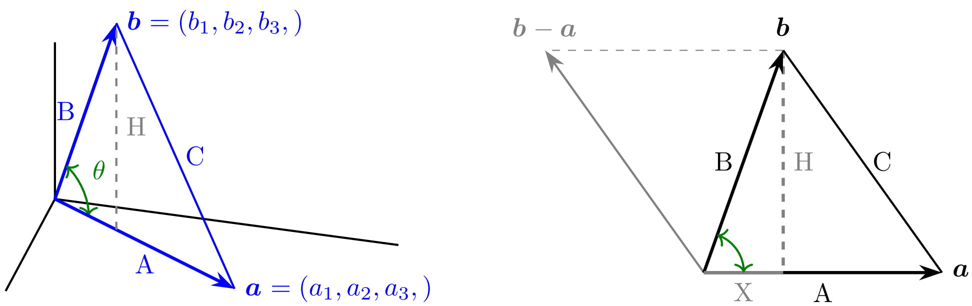


Figura 11.3: Triángulo formado por $\mathbf{0}$ y dos vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ no nulos.

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 indica el cateto X de uno de los triángulos rectángulos resultantes de dividir el triángulo ABC con el segmento vertical H.

En todo *triángulo rectángulo*[1], el coseno del ángulo θ formado por cada uno de los catetos con la hipotenusa

[1]Como en un triángulo rectángulo ningún angulo tiene más de 90 grados, el coseno del ángulo de cualquiera de sus vértices siempre es mayor o igual a cero

es el ratio entre las longitudes de dicho cateto y la hipotenusa. Por tanto,

$$\cos\theta = \frac{X}{B} \tag{11.1}$$

donde con X denotamos la longitud del cateto X y con B denotamos la longitud de la hipotenusa B.

Puesto que el ángulo formado por los vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ es el mismo que el formado por los segmentos X y B, solo necesitamos calcular el ratio X/B. Por el teorema de Pitágoras, sabemos que

$$X^2 + H^2 = B^2. \tag{11.2}$$

Y si nos fijamos en el segundo triángulo rectángulo, también sabemos que

$$\begin{aligned} (A-X)^2 + H^2 &= C^2 \\ A^2 - 2AX + X^2 + H^2 &= C^2 \qquad \text{desarrollando el cuadrado.} \end{aligned} \tag{11.3}$$

Restando (11.2) de (11.3) tenemos

$$\begin{aligned} A^2 - 2AX &= C^2 - B^2 \\ -2AX &= C^2 - B^2 - A^2 \\ X &= (A^2 + B^2 - C^2)/2A \end{aligned}$$

Y sustituyendo en (11.1) llegamos a que

$$\cos\theta = \frac{A^2 + B^2 - C^2}{2AB}$$

Así, como en la figura asociamos el segmento A con el vector $\boldsymbol{a}$ y el segmento B con $\boldsymbol{b}$, podemos sustituir A, B y C por las longitudes de los correspondientes vectores asociados (y C por la longitud de $\boldsymbol{b} - \boldsymbol{a}$).

$$\begin{aligned} A = \|\boldsymbol{a}\| &= \sqrt{a_1^2 + a_2^2 + a_3^2} \\ B = \|\boldsymbol{b}\| &= \sqrt{b_1^2 + b_2^2 + b_3^2} \\ C = \|\boldsymbol{b} - \boldsymbol{a}\| &= \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2} \end{aligned}$$

Así

$$\cos\theta \;=\; \frac{2(a_1b_1 + a_2b_2 + a_3b_3)}{2\sqrt{a_1^2 + a_2^2 + a_3^2}\sqrt{b_1^2 + b_2^2 + b_3^2}} \;=\; \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|}$$

donde hemos sustituido la suma de productos de las componentes por el producto punto en $\mathbb{R}^3$.

11.3. Desigualdad de Cauchy-Schwarz

Veamos que $\left|\langle \vec{v}\,,\, \vec{w}\rangle\right| \leq \|\vec{v}\|\|\vec{w}\|$

Caso trivial: $\vec{v}$ *y* $\vec{w}$ *son linealmente dependientes*

En tal caso, o bien $\vec{v} = \vec{0}$ o bien existe $\alpha \in \mathbb{R}$ tal que $\alpha\vec{v} = \vec{w}$. Si ocurre lo primero $\langle \vec{v}\,,\, \vec{w}\rangle = 0$ y $\|\vec{v}\|\|\vec{w}\| = 0$. Si ocurre lo segundo

$$\left|\langle \vec{v}\,,\, \vec{w}\rangle\right| = \left|\langle \vec{v}\,,\, \alpha\vec{v}\rangle\right| = \left|\alpha\|\vec{v}\|^2\right| = \|\vec{v}\|\,|\alpha|\,\|\vec{v}\| = \|\vec{v}\|\|\vec{w}\|$$

Caso no trivial: $\vec{v}$ *y* $\vec{w}$ *son linealmente independientes*

En tal caso existe un vector $\vec{h}$ ortogonal a $\vec{v}$ tal que $L(\vec{v}, \vec{h}) = L(\vec{v}, \vec{w})$. Dicho vector lo construimos de la forma $\vec{h} = \vec{w} + \alpha\vec{v}$, con lo que garantizamos que $L(\vec{v}, \vec{h}) = L(\vec{v}, \vec{w})$ e imponemos que

$$0 = \left\langle \vec{v}\,,\, \vec{h} \right\rangle = \left\langle \vec{v}\,,\, \vec{w} + \alpha\vec{v} \right\rangle = \left\langle \vec{v}\,,\, \vec{w} \right\rangle + \alpha \left\langle \vec{v}\,,\, \vec{v} \right\rangle$$

con lo que $\alpha = -\frac{\left\langle \vec{v}\,,\, \vec{w} \right\rangle}{\left\langle \vec{v}\,,\, \vec{v} \right\rangle}$. Por tanto tendremos que

$$\begin{aligned} \|\vec{w}\|^2 &= \left\langle \vec{w}\,,\, \vec{w} \right\rangle = \left\langle \vec{h} - \alpha\vec{v}\,,\, \vec{h} - \alpha\vec{v} \right\rangle = \left\langle \vec{h}\,,\, \vec{h} \right\rangle - 2\alpha \left\langle \vec{h}\,,\, \vec{v} \right\rangle + \alpha^2 \left\langle \vec{v}\,,\, \vec{v} \right\rangle \\ &= \|\vec{h}\|^2 + \alpha^2\|\vec{v}\|^2 \quad > \quad \alpha^2\|\vec{v}\|^2 = \frac{\left\langle \vec{v}\,,\, \vec{w} \right\rangle^2}{\left\langle \vec{v}\,,\, \vec{v} \right\rangle^2}\|\vec{v}\|^2 = \frac{\left\langle \vec{v}\,,\, \vec{w} \right\rangle^2}{\|\vec{v}\|^2} \end{aligned}$$

de donde $\left\langle \vec{v}\,,\, \vec{w} \right\rangle^2 < \|\vec{v}\|^2\|\vec{w}\|^2$ y por consiguiente $\left|\left\langle \vec{v}\,,\, \vec{w} \right\rangle\right| < \|\vec{v}\|\|\vec{w}\|$.

11.4. Generalizando longitudes y ángulos a $\mathbb{R}^n$

Generalicemos los conceptos (vistos para $\mathbb{R}^2$ y $\mathbb{R}^3$) de longitud y coseno del ángulo entre vectores al espacio genérico $\mathbb{R}^n$:

Definición 11.1. *La longitud (o norma) de un vector $\boldsymbol{a}$ de $\mathbb{R}^n$ es la raíz cuadrada de $\boldsymbol{a}\cdot\boldsymbol{a}$:*

$$\textit{longitud de } \boldsymbol{a} \;=\; \|\boldsymbol{a}\| \;=\; \sqrt{\boldsymbol{a}\cdot\boldsymbol{a}}.$$

Por ejemplo, la longitud de $\boldsymbol{x} = \begin{pmatrix}6\\0\\-2\\3\end{pmatrix}$ es $\sqrt{\boldsymbol{x}\cdot\boldsymbol{x}} = \sqrt{\begin{pmatrix}6\\0\\-2\\3\end{pmatrix}\cdot\begin{pmatrix}6\\0\\-2\\3\end{pmatrix}} = \sqrt{6^2+0^2+(-2)^2+3^2} = \sqrt{49} = 7.$

Nótese que $\quad \|\boldsymbol{x}\| = 0 \;\Leftrightarrow\; \boldsymbol{x} = \boldsymbol{0}$.

Definición 11.2. *El coseno del ángulo θ formado por los vectores no nulos $\boldsymbol{a}$ y $\boldsymbol{b}$ de $\mathbb{R}^n$ es*

$$\cos\theta \;=\; \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|}; \qquad -1 \leq \cos\theta \leq 1.$$

Con esta generalización, la parte derecha de la Figura 11.3 se puede interpretar como una representación esquemática de un triángulo formado por vectores de $\mathbb{R}^n$.

Vectores alineados. Vectores unitarios.

Si el ángulo formado por dos vectores $\boldsymbol{a}$ y $\boldsymbol{b}$ es cero, $(\theta = 0)$, entonces $\cos\theta = \cos 0 = 1$; y por tanto

$$\boldsymbol{a}\cdot\boldsymbol{b} \;=\; \|\boldsymbol{a}\|\cdot\|\boldsymbol{b}\|.$$

Así, se llega al conocido resultado de que

$$\boldsymbol{a}\cdot\boldsymbol{a} \;=\; \|\boldsymbol{a}\|^2.$$

Nótese que el efecto de multiplicar un vector $\boldsymbol{x}$ por α es obtener un múltiplo de $\boldsymbol{x}$ cuya longitud es $|\alpha|$ veces la longitud original.

$$\|\alpha\boldsymbol{x}\| = \sqrt{\alpha^2\big(\boldsymbol{x}\cdot\boldsymbol{x}\big)} = |\alpha|\,\|\boldsymbol{x}\|$$

Cuando un vector $\boldsymbol{y}$ es un múltiplo de $\boldsymbol{x}$ el producto escalar de ambos es

$$\boldsymbol{y}\cdot\boldsymbol{x} = \alpha\big(\boldsymbol{x}\cdot\boldsymbol{x}\big) = \alpha\|\boldsymbol{x}\|^2$$

En tal caso (y si $\alpha \neq 0$) el ángulo que forman es cero (o 180) ya que:

$$\cos\theta = \frac{\boldsymbol{x}\cdot\boldsymbol{y}}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} = \frac{\alpha\|\boldsymbol{x}\|^2}{|\alpha|\,\|\boldsymbol{x}\|\|\boldsymbol{x}\|} = \frac{\alpha}{|\alpha|} = \pm 1.$$

Definición 11.3 (Vector unitario)**.** *Se dice que un vector es unitario si su norma es uno.*

Partiendo de un vector no nulo $\boldsymbol{x}$, es sencillo obtener un múltiplo unitario (i.e., de longitud uno); basta dividirlo por su longitud. Por ejemplo si $\boldsymbol{x} = \begin{pmatrix}1, & 1, & 1, & 1,\end{pmatrix}$, entonces $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}\cdot\boldsymbol{x}} = \sqrt{4} = 2$. Así,

$$\boldsymbol{y} = \frac{1}{\|\boldsymbol{x}\|}(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{x}) = \begin{pmatrix}\frac{1}{2}, & \frac{1}{2}, & \frac{1}{2}, & \frac{1}{2},\end{pmatrix} \quad \text{es unitario. ¡Compruébelo!}$$

Vectores perpendiculares.

Definición 11.4. *Decimos que* $\boldsymbol{a}$ *y* $\boldsymbol{b}$ *son ortogonales* o perpendiculares ($\boldsymbol{a} \perp \boldsymbol{b}$) *cuando* $\boldsymbol{a}\cdot\boldsymbol{b} = 0$.

Por tanto $\mathbf{0}$ es ortogonal a todos los vectores (incluido él mismo).

La propiedad más importante que verifican los vectores perpendiculares es el **Teorema de Pitágoras** que también vamos a generalizar a $\mathbb{R}^n$ (recuerde las propiedades del *producto punto* en la Página 23):

Teorema 11.4.1 (Teorema de Pitágoras)**.** *Sean* $\boldsymbol{x}$ *e* $\boldsymbol{y}$ *son dos vectores de* $\mathbb{R}^n$*; entonces* $\boldsymbol{x}\cdot\boldsymbol{y} = 0$ (son perpendiculares) *si y solo si*

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2.$$

Demostración. Si $\boldsymbol{x}$ e $\boldsymbol{y}$ son *perpendiculares*, entonces $\boldsymbol{x}\cdot\boldsymbol{y} = 0$ y

$$\begin{aligned}
\|\boldsymbol{x}+\boldsymbol{y}\|^2 = (\boldsymbol{x}+\boldsymbol{y})\cdot(\boldsymbol{x}+\boldsymbol{y}) &= \boldsymbol{x}\cdot(\boldsymbol{x}+\boldsymbol{y}) + \boldsymbol{y}\cdot(\boldsymbol{x}+\boldsymbol{y}) \\
&= (\boldsymbol{x}+\boldsymbol{y})\cdot\boldsymbol{x} + (\boldsymbol{x}+\boldsymbol{y})\cdot\boldsymbol{y} \\
&= \boldsymbol{x}\cdot\boldsymbol{x} + \boldsymbol{y}\cdot\boldsymbol{x} + \boldsymbol{x}\cdot\boldsymbol{y} + \boldsymbol{y}\cdot\boldsymbol{y} \\
&= \boldsymbol{x}\cdot\boldsymbol{x} + \quad 2(\boldsymbol{x}\cdot\boldsymbol{y}) \quad + \boldsymbol{y}\cdot\boldsymbol{y} \\
&= \boldsymbol{x}\cdot\boldsymbol{x} + \qquad 0 \qquad + \boldsymbol{y}\cdot\boldsymbol{y} \qquad = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2.
\end{aligned}$$

□

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.

Un sistema de vectores perpendiculares y no nulos es un sistema linealmente independiente. Para demostrarlo, veamos antes el producto $\mathbf{A}^{\intercal}\mathbf{A}$ cuando $\mathbf{A}$ tiene columnas perpendiculares entre si.

Proposición 11.4.2. *Sean las columnas de* $\mathbf{A}$ *perpendiculares entre si, entonces* $\mathbf{A}^{\intercal}\mathbf{A}$ *es diagonal, y las componentes de la diagonal son el cuadrado de las normas de cada una de las columnas de* $\mathbf{A}$*.*

Demostración. $${}_{i|}\big(\mathbf{A}^{\intercal}\mathbf{A}\big)_{|j} = \big(\mathbf{A}_{|i}\big)\cdot\big(\mathbf{A}_{|j}\big) = \begin{cases} 0 & \text{cuando } i \neq j \\ \|\mathbf{A}_{|i}\|^2 & \text{cuando } i = j \end{cases}$$ □

Puesto que $\mathbf{A}^{\intercal}\mathbf{A}$ es diagonal, su fila $i$ésima es un múltiplo de la fila $i$ésima de la matriz identidad:

$${}_{i|}\big(\mathbf{A}^{\intercal}\mathbf{A}\big) = \big(\|\mathbf{A}_{|i}\|^2\big)\,{}_{i|}\mathbf{I};$$

multiplicando por $\boldsymbol{x}$ deducimos que la componente $i$ésima de $\mathbf{A}^{\intercal}\mathbf{A}\boldsymbol{x}$ es:

$${}_{i|}\mathbf{A}^{\intercal}\mathbf{A}\boldsymbol{x} = \big(\|\mathbf{A}_{|i}\|^2\big)\,{}_{i|}\mathbf{I}\boldsymbol{x} = \|\mathbf{A}_{|i}\|^2 x_i,$$

ya que $\big({}_{i|}\mathbf{I}\boldsymbol{x}\big)$ selecciona el elemento $i$ésimo de $\boldsymbol{x}$. Usemos este resultado para demostrar que un sistema de vectores no nulos y perpendiculares entre si, es linealmente independiente:

Proposición 11.4.3. *Sea un sistema de vectores* $\left[\boldsymbol{a}_1;\ldots;\boldsymbol{a}_n;\right]$ *distintos de* $\mathbf{0}$ *y perpendiculares entre sí, entonces el sistema es linealmente independiente.*

Demostración. Consideremos la matriz $\mathbf{A} = \left[\boldsymbol{a}_1;\ldots;\boldsymbol{a}_n;\right]$ y demostremos que $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ si y solo si $\boldsymbol{x} = \mathbf{0}$. Multiplicando por $\mathbf{A}^\intercal$ a ambos lados del sistema $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ tenemos que $\mathbf{A}^\intercal\mathbf{A}\boldsymbol{x} = \mathbf{0}$, y como

$$_{i|}\mathbf{A}^\intercal\mathbf{A}\boldsymbol{x} = \|\mathbf{A}_{|i}\|^2(x_i); \qquad \text{resulta que} \qquad _{i|}\mathbf{A}^\intercal\mathbf{A}\boldsymbol{x} = {}_{i|}\mathbf{0} \iff \|\mathbf{A}_{|i}\|^2(x_i) = 0.$$

Puesto que los vectores $\boldsymbol{a}_i$ (que son las columnas de $\mathbf{A}$) son distintos de cero, sus normas también lo son. Así que necesariamente $x_i = 0$ para $i = 1,\ldots,n$, es decir, $\boldsymbol{x} = \mathbf{0}$. □

11.5. Ortogonalidad de los 4 subespacios fundamentales de A

Definición 11.5 (Subespacios ortogonales)**.** *Dos* subespacios $\mathcal{A}$ *y* $\mathcal{B}$ *de* $\mathbb{R}^n$ *son ortogonales* ($\mathcal{A} \perp \mathcal{B}$)*, si todo vector de* $\mathcal{A}$ *es perpendicular a todo vector de* $\mathcal{B}$*; es decir, si* $\boldsymbol{a}\cdot\boldsymbol{b} = 0$ *para cualesquiera* $\boldsymbol{a} \in \mathcal{A}$ *y* $\boldsymbol{b} \in \mathcal{B}$*.*

La siguiente proposición muestra que los subespacios fundamentales de $\mathbf{A}$ son ortogonales dos a dos.

Proposición 11.5.1. *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:

$$\boldsymbol{y}\cdot\boldsymbol{b} = \boldsymbol{y}\mathbf{A}\boldsymbol{x} = \mathbf{0}\cdot\boldsymbol{x} = 0.$$

□

Definición 11.6 (Complementos ortogonales)**.** *Dos* subespacios $\mathcal{A}$ *y* $\mathcal{B}$ *de* $\mathbb{R}^n$ *se dice que son uno el* complemento ortogonal *del otro si son son suplementarios y ortogonales entre sí.*

Librería 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 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 Z.

Ejemplo 20. 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:

$$\left[\begin{array}{cccc}1&-3&0&-1\\0&-1&1&1\\1&-4&1&0\\\hline 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right]\xrightarrow[\substack{[(3)\mathbf{1}+\mathbf{2}]\\[(1)\mathbf{1}+\mathbf{4}]\\[(1)\mathbf{2}+\mathbf{3}]\\[(1)\mathbf{2}+\mathbf{4}]}]{\boldsymbol{\tau}}\left[\begin{array}{cccc}1&0&0&0\\0&-1&0&0\\1&-1&0&0\\\hline 1&3&3&4\\0&1&1&1\\0&0&1&0\\0&0&0&1\end{array}\right]\implies \text{Base del complemento ortogonal: }\left[\begin{pmatrix}3\\1\\1\\0\end{pmatrix};\begin{pmatrix}4\\1\\0\\1\end{pmatrix};\right].$$

Librería 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
```

11.5.1. Ecuaciones cartesianas (o implícitas) y ecuaciones paramétricas

En la Sección 7.2.2 de la Lección 7 aparecieron por primera vez las ecuaciones paramétricas y las ecuaciones cartesianas (véase las notas a pie de página números 2 en la página 92 y 1 en la página 89). Veamos ahora cómo pasar de unas a las otras; así dispondremos en todo momento de dos modos de describir el conjunto de soluciones. 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 de las ecuaciones cartesianas).

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, expresamos el conjunto de soluciones con la siguiente *ecuación cartesiana*:

$$\{\boldsymbol{x}\in\mathbb{R}^n\,|\mathbf{A}\boldsymbol{x}=\boldsymbol{b}\}$$

Cuando resolvemos el sistema encontrando una solución particular $\boldsymbol{s}$ y un sistema generador $\big[\boldsymbol{n}_1;\;\dots\;\boldsymbol{n}_k\big]$ de $\mathcal{N}(\mathbf{A})$, re-expresamos el conjunto de soluciones con la *ecuación paramétrica*:

$$\left\{\boldsymbol{x}\in\mathbb{R}^n\;\middle|\;\exists\boldsymbol{p}\in\mathbb{R}^k \text{ tal que } \boldsymbol{x}=\boldsymbol{s}+\big[\boldsymbol{n}_1;\;\dots\;\boldsymbol{n}_k\big]\boldsymbol{p}\right\}.$$

Hasta ahora siempre hemos pasado de las ecuaciones cartesianas (o implícitas) a las ecuaciones paramétricas; veamos como recorrer el camino inverso.

Recorriendo el camino inverso: de las ecuaciones paramétricas a las cartesianas

En el camino inverso, dado el conjunto de soluciones de "un sistema desconocido, $\mathbf{D}\boldsymbol{x} = \boldsymbol{f}$*", encontraremos las filas de una matriz de coeficientes* $\mathbf{A}$ *(ortogonales a los vectores que generan el espacio nulo) y un vector del lado derecho* $\boldsymbol{b}$ *que definan un sistema de ecuaciones* $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ *cuyo conjunto de soluciones coincide con el del sistema desconocido,* $\mathbf{D}\boldsymbol{x} = \boldsymbol{f}$*.*

Procedimiento. Cuando disponemos de unas ecuaciones paramétricas

$$\left\{\boldsymbol{x} \in \mathbb{R}^n \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^k \text{ tal que } \boldsymbol{x} = \boldsymbol{s} + \left[\boldsymbol{n}_1;\; \dots\; \boldsymbol{n}_k\right]\boldsymbol{p}\right\},$$

disponemos tanto de una solución $\boldsymbol{s}$ (un vector del conjunto de soluciones) como de un sistema $\left[\boldsymbol{n}_1;\; \dots\; \boldsymbol{n}_k\right]$ generador del espacio nulo de la matriz $\mathbf{A}$ que buscamos, es decir, $\mathcal{N}(\mathbf{A}) = \mathcal{L}\left(\left[\boldsymbol{n}_1;\; \dots\; \boldsymbol{n}_k\right]\right)$.

Por eliminación podemos encontrar una base del complemento ortogonal de $\mathcal{N}(\mathbf{A})$, es decir una base de $\mathcal{C}\left(\mathbf{A}^{\intercal}\right)$. Usaremos los vectores de dicha base como *filas* de la matriz $\mathbf{A}$ que buscamos. Los pasos a seguir son:

Primero encontramos una base del complemento ortogonal de $\mathcal{N}(\mathbf{A}) = \mathcal{L}\left(\left[\boldsymbol{n}_1;\; \dots\; \boldsymbol{n}_k\right]\right)$. Para ello usamos la matriz $\mathbf{Z}$ cuyas k filas son los vectores $\boldsymbol{n}_1, \dots, \boldsymbol{n}_k$ perpendiculares a $\mathcal{N}(\mathbf{A})$ y aplicamos la eliminación "de izquierda a derecha" siguiendo el esquema:

$$\left[\frac{\mathbf{Z}}{\mathbf{I}}\right] \to \left[\frac{\mathbf{K}}{\mathbf{E}}\right]; \qquad \text{donde } \mathbf{Z} = \left[\boldsymbol{n}_1\; \dots\; \boldsymbol{n}_k\right]^{\intercal};$$

Después usamos las "soluciones especiales" encontradas (las columnas de $\mathbf{E}$ bajo vectores nulos de $\mathbf{K}$) como *filas* para construir la matriz $\mathbf{A}$; de manera que dichas soluciones son una base de $\mathcal{C}\left(\mathbf{A}^{\intercal}\right)$.

Comprobación. Si multiplicamos por $\mathbf{A}$ la ecuación paramétrica $\boldsymbol{x} = \boldsymbol{s} + \left[\boldsymbol{n}_1;\; \dots\; \boldsymbol{n}_k\right]\boldsymbol{p}$ tenemos que

$$\mathbf{A}\boldsymbol{x} \;=\; \mathbf{A}\boldsymbol{s} \;+\; \mathbf{A}\left[\boldsymbol{n}_1;\; \dots\; \boldsymbol{n}_k\right]\boldsymbol{p} \quad\Rightarrow\quad \mathbf{A}\boldsymbol{x} \;=\; \mathbf{A}\boldsymbol{s},$$

donde $\boldsymbol{b} = \mathbf{A}\boldsymbol{s}$ es el vector del lado derecho. Este sistema es, por construcción, un sistema cuyas soluciones son las ecuaciones paramétricas de las que hemos partido, ya que

$$\begin{cases} \mathbf{A}\boldsymbol{x} = \boldsymbol{b} & \text{cuando} \quad \boldsymbol{x} = \boldsymbol{s} \\ \mathbf{A}\boldsymbol{x} = \boldsymbol{0} & \text{cuando} \quad \boldsymbol{x} \in \mathcal{L}\left(\left[\boldsymbol{n}_1;\; \dots\; \boldsymbol{n}_k\right]\right) = \mathcal{N}(\mathbf{A}) \end{cases}.$$

Procedimiento abreviado para usar con lápiz y papel. No es necesario encontrar primero los vectores perpendiculares a $\boldsymbol{n}_1, \dots, \boldsymbol{n}_k$ para luego formar con ellos las filas de $\mathbf{A}$ y después multiplicar la ecuación paramétrica inicial para obtener el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$... En realidad podemos hacer todo esto a la vez con nuestra "herramienta multiusos', es decir, con *la eliminación de izquierda a derecha.*

Si $\mathbf{M}$ es la matriz

$$\mathbf{M} = \left[\boldsymbol{n}_1, \dots, \boldsymbol{n}_k \,||\, \boldsymbol{x} \,|\, \boldsymbol{s}\right]^{\intercal} \qquad \text{(nótese el símbolo de transposición)},$$

donde $\boldsymbol{x}$ es el vector de incógnitas, $\boldsymbol{s}$ es una solución particular y $\left[\boldsymbol{n}_1;\; \dots\; \boldsymbol{n}_k\right]$ es un sistema generador de $\mathcal{N}(\mathbf{A})$; podemos tratar de hacer columnas de ceros en la submatriz con las k primeras filas (correspondientes a los vectores $\boldsymbol{n}_1, \dots, \boldsymbol{n}_k$). Una vez finalizado el proceso de eliminación, "aparecerá" bajo cada columna nula una ecuación del sistema de ecuaciones deseado. Veámoslo.

Ejemplo 21. Sea el conjunto $\left\{\boldsymbol{x} \in \mathbb{R}^4 \;\middle|\; \exists a, b \in \mathbb{R} \text{ tales que } \boldsymbol{x} = \begin{pmatrix}0\\2\\0\\5\end{pmatrix} + a\begin{pmatrix}1\\2\\0\\-2\end{pmatrix} + b\begin{pmatrix}0\\0\\1\\3\end{pmatrix}\right\}.$

Apliquemos la eliminación sobre la siguiente *matriz particionada* M cuya tercera fila es $\boldsymbol{x} = (x, y, z, w)$ y cuya cuarta y última fila es la solución particular $\boldsymbol{s} = (0, 2, 0, 5)$:

$$\left[\begin{array}{c}\mathsf{M}\\ \hline \hline \mathsf{I}\end{array}\right]=\left[\begin{array}{cccc}1&2&0&-2\\0&0&1&3\\ \hline\hline x&y&z&w\\ \hline 0&2&0&5\\ \hline\hline 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right]\xrightarrow[{[(2)\mathbf{1}+\mathbf{4}]}]{\substack{\boldsymbol{\tau}\\ {[(-2)\mathbf{1}+\mathbf{2}]}}}\left[\begin{array}{cccc}1&0&0&0\\0&0&1&3\\ \hline\hline x&-2x+y&z&w+2x\\ \hline 0&5&0&5\\ \hline\hline 1&-2&0&2\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{array}\right]\xrightarrow{\substack{\boldsymbol{\tau}\\ {[(-3)\mathbf{3}+\mathbf{4}]}}}\left[\begin{array}{cccc}1&0&0&0\\0&0&1&0\\ \hline\hline x&-2x+y&z&w+2x-3z\\ \hline 0&2&0&5\\ \hline\hline 1&-2&0&2\\0&1&0&0\\0&0&1&-3\\0&0&0&1\end{array}\right].$$

Entonces $\mathsf{A} = \begin{bmatrix}-2&1&0&0\\2&0&-3&1\end{bmatrix}$; y consecuentemente $\mathsf{A}\boldsymbol{x} = \begin{pmatrix}-2x+y\\2x+w-3z\end{pmatrix}$ y $\mathsf{A}\boldsymbol{s} = \boldsymbol{b} = \begin{pmatrix}2\\5\end{pmatrix}$. Es decir

$$\mathsf{A}\boldsymbol{x} = \boldsymbol{b} \quad\longrightarrow\quad \begin{cases}-2x+y & =2\\ 2x+w-3w & =5\end{cases}.$$

Así pues, la *ecuación cartesiana* correspondiente a la ecuación paramétrica de este ejemplo es

$$\left\{\boldsymbol{x}\in\mathbb{R}^4 \;\middle|\; \begin{bmatrix}-2&1&0&0\\2&0&-3&1\end{bmatrix}\boldsymbol{x} = \begin{pmatrix}2\\5\end{pmatrix}\right\}.$$

Pero si nos fijamos bien, es evidente que para obtener el sistema de ecuaciones

$$\begin{cases}-2x+y & =2\\ 2x+w-3w & =5\end{cases}$$

no es necesario añadir la matriz identidad por debajo de la matriz particionada M, pues en el proceso de eliminación ya hemos calculado $\mathsf{A}\boldsymbol{x}$ y $\mathsf{A}\boldsymbol{s}$. Así pues, para encontrar las ecuaciones paramétricas con lápiz y papel basta aplicar la eliminación:

$$\left[\begin{array}{cccc}1&2&0&-2\\0&0&1&3\\ \hline\hline x&y&z&w\\ \hline 0&2&0&5\end{array}\right]\xrightarrow[{[(2)\mathbf{1}+\mathbf{4}]}]{\substack{\boldsymbol{\tau}\\ {[(-2)\mathbf{1}+\mathbf{2}]}}}\left[\begin{array}{cccc}1&0&0&0\\0&0&1&3\\ \hline\hline x&-2x+y&z&w+2x\\ \hline 0&5&0&5\end{array}\right]\xrightarrow{\substack{\boldsymbol{\tau}\\ {[(-3)\mathbf{3}+\mathbf{4}]}}}\left[\begin{array}{cccc}1&0&0&0\\0&0&1&0\\ \hline\hline x&-2x+y&z&w+2x-3z\\ \hline 0&2&0&5\end{array}\right],$$

mirar qué ecuaciones aparecen bajo las columnas de ceros y deducir cuál es la correspondiente matriz de coeficientes A del sistema formado por dichas ecuaciones.

Puntos, rectas, planos e hiper-planos de $\mathbb{R}^n$. Espacios afines

Dado un sistema de ecuaciones $\mathsf{A}\boldsymbol{x} = \boldsymbol{b}$, el conjunto de soluciones posee una forma geométrica especial, ya que es "plano", es decir, no es curvado. En particular, si $\dim\mathcal{N}(\mathsf{A}) = 0$ el conjunto es *un punto*; si $\dim\mathcal{N}(\mathsf{A}) = 1$ el conjunto es *una recta*; si $\dim\mathcal{N}(\mathsf{A}) = 2$ el conjunto es *un plano*, y si $\dim\mathcal{N}(\mathsf{A}) = d > 2$ el conjunto es *un hiper-plano 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. tanto a las ecuaciones paramétricas como a las cartesianas que describen el conjunto de soluciones de un sistema $\mathsf{A}\boldsymbol{x} = \boldsymbol{b}$. Por tanto, podemos preguntarnos por las ecuaciones cartesianas (o las ecuaciones paramétricas) de, por ejemplo, una recta en $\mathbb{R}^n$ (y al decir *recta* implícitamente estamos indicando que el espacio nulo de la matriz de coeficientes tiene dimensión uno).

De hecho, estos subconjuntos de $\mathbb{R}^n$ descritos como la suma de un vector particular $\boldsymbol{s}$ más las combinaciones lineales de un sistema de vectores $\left[\boldsymbol{n}_1; \ \ldots \ \boldsymbol{n}_k\right]$ se denominan *espacios afines.* La librería de Python guarda el conjunto de soluciones de sistema de ecuaciones en forma de espacio afín en el atributo `eafin`; y su representación en Jupyter muestra tanto las ecuaciones cartesianas como las paramétricas de dichos espacios.

Librería NAcAL para Python

```
A = Matrix( [ [-1,1,0,0], [-1,0,-1,1] ] )
b = Vector([0,1])
SEL(A,b).eafin
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^2 \text{ tal que } \boldsymbol{v} = \begin{pmatrix}0\\0\\0\\1\end{pmatrix} + \begin{bmatrix}-1&0\\-1&0\\1&1\\0&1\end{bmatrix}\boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix}-1&1&0&0\\-1&0&-1&1\end{bmatrix}\boldsymbol{v} = \begin{pmatrix}0\\1\end{pmatrix} \right\}$$

Podemos generar un espacio afín con una matriz y un vector

Librería NAcAL para Python

```
A = Matrix( [ [-1,1,0,0], [-1,0,-1,1] ] )
v = Vector([0,0,0,1])
EAfin(A,v)
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^2 \text{ tal que } \boldsymbol{v} = \begin{pmatrix}0\\0\\0\\1\end{pmatrix} + \begin{bmatrix}-1&1\\-1&1\\1&0\\0&1\end{bmatrix}\boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix}-1&1&0&0\\-1&0&-1&1\end{bmatrix}\boldsymbol{v} = \begin{pmatrix}0\\1\end{pmatrix} \right\}$$

o como con un sistema de vectores y un vector

Librería NAcAL para Python

```
S = Sistema([ Vector([-1,-1,1,0]), Vector([0,0,1,1]) ])
v = Vector([0,0,0,1])
EAfin(S,v)
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^2 \text{ tal que } \boldsymbol{v} = \begin{pmatrix}0\\0\\0\\1\end{pmatrix} + \begin{bmatrix}-1&0\\-1&0\\1&1\\0&1\end{bmatrix}\boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix}-1&1&0&0\\-1&0&-1&1\end{bmatrix}\boldsymbol{v} = \begin{pmatrix}0\\1\end{pmatrix} \right\}$$

Proyecciones sobre subespacios

12.1. Proyecciones sobre subespacios

Puesto que dada $\mathbf{A}$ es de orden m por n se verifica que $\mathbb{R}^n = \mathcal{C}\left(\mathbf{A}^\intercal\right) \oplus \mathcal{N}\left(\mathbf{A}\right)$ y que $\mathbb{R}^m = \mathcal{C}\left(\mathbf{A}\right) \oplus \mathcal{N}\left(\mathbf{A}^\intercal\right)$, sabemos que todo vector de $\mathbb{R}^n$ se puede descomponer como suma de un vector de $\mathcal{C}\left(\mathbf{A}\right)$ más otro de $\mathcal{N}\left(\mathbf{A}^\intercal\right)$. Además, como $\mathcal{C}\left(\mathbf{A}^\intercal\right) \cap \mathcal{N}\left(\mathbf{A}\right) = \{\mathbf{0}\}$, dicha descomposición es única. E igualmente, todo vector de $\mathbb{R}^m$ se puede descomponer de manera úncia como suma de un vector de $\mathcal{C}\left(\mathbf{A}^\intercal\right)$ más otro de $\mathcal{N}\left(\mathbf{A}\right)$.

Proposición 12.1.1. $\mathcal{N}\left(\mathbf{A}\right) = \mathcal{N}\left(\mathbf{A}^\intercal\mathbf{A}\right)$.

Demostración. Para demostrar que dos conjuntos son iguales basta mostrar que todo elemento del primer conjunto está en el segundo, y viceversa. Veamos que así ocurre con $\mathcal{N}\left(\mathbf{A}\right)$ y $\mathcal{N}\left(\mathbf{A}^\intercal\mathbf{A}\right)$.

Primero comprobemos que todo vector de $\mathcal{N}\left(\mathbf{A}\right)$ es también un vector de $\mathcal{N}\left(\mathbf{A}^\intercal\mathbf{A}\right)$:

> 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 $\mathbf{A}^\intercal\mathbf{A}\boldsymbol{x} = \left(\mathbf{A}^\intercal\right)\mathbf{0} = \mathbf{0}$; y por tanto $\boldsymbol{x} \in \mathcal{N}\left(\mathbf{A}^\intercal\mathbf{A}\right)$.

Ahora toca la segunda parte:

> 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 $\boldsymbol{x}\mathbf{A}^\intercal\mathbf{A}\boldsymbol{x} = \boldsymbol{x} \cdot \mathbf{0} = 0$. Pero esto quiere decir que el vector $\mathbf{A}\boldsymbol{x}$ tiene norma cero, ya que $\boldsymbol{x}\left(\mathbf{A}^\intercal\right)\mathbf{A}\boldsymbol{x} = (\mathbf{A}\boldsymbol{x}) \cdot (\mathbf{A}\boldsymbol{x})$, es el cuadrado de la norma de $\mathbf{A}\boldsymbol{x}$. Como el único vector de norma cero es $\mathbf{0}$ (el vector nulo) tenemos que $\mathbf{A}\boldsymbol{x} = \mathbf{0}$, y por tanto $\boldsymbol{x} \in \mathcal{N}\left(\mathbf{A}\right)$.

□

La proposición anterior tiene consecuencias sobre el rango de las matrices $\mathbf{A}$ y $\mathbf{A}^\intercal\mathbf{A}$:

Corolario 12.1.2. *El rango de* $\mathbf{A}^\intercal\mathbf{A}$ *es igual al rango de* $\mathbf{A}$.

Demostración. El rango de una matriz es igual al número de columnas con pivote, es decir, al número total de columnas menos las columnas sin pivote. El número de columnas sin pivote es la dimensión del espacio nulo. Como tanto $\mathbf{A}^\intercal\mathbf{A}$ como $\mathbf{A}$ tienen n columnas e idéntico espacio nulo ($\mathcal{N}\left(\mathbf{A}\right) = \mathcal{N}\left(\mathbf{A}^\intercal\mathbf{A}\right)$), necesariamente tienen el mismo rango: $k = n - \dim\mathcal{N}\left(\mathbf{A}\right)$. □

Mínimos cuadrados

Aunque el sistema $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ no tiene solución cuando $\boldsymbol{b} \notin \mathcal{C}(\mathbf{A})$, ¡El sistema $(\mathbf{A}^\intercal\mathbf{A})\boldsymbol{x} = \big(\mathbf{A}^\intercal\big)\boldsymbol{b}$ siempre se puede resolver! Antes de demostrarlo hay que recordar que $\mathbf{A}$ y $\mathbf{A}^\intercal$ tienen el mismo rango; y que rango$(\mathbf{AB}) \leq$ rango$(\mathbf{A})$, pues las columnas de $\mathbf{AB}$ son combinaciones de las columnas de $\mathbf{A}$ y por tanto $\mathcal{C}(\mathbf{AB}) \subset \mathcal{C}(\mathbf{A})$.

Proposición 13.0.1. *Sea* $\mathbf{A}$ *de orden* m *por* n *y* $\boldsymbol{b} \in \mathbb{R}^m$*, entonces el sistema* $(\mathbf{A}^\intercal\mathbf{A})\boldsymbol{x} = \big(\mathbf{A}^\intercal\big)\boldsymbol{b}$ *siempre tiene solución.*

Demostración. El sistema $(\mathbf{A}^\intercal\mathbf{A})\boldsymbol{x} = \big(\mathbf{A}^\intercal\big)\boldsymbol{b}$ tiene solución si el rango de la matriz de coeficientes $\mathbf{A}^\intercal\mathbf{A}$ es igual que el rango de la matriz ampliada $[\mathbf{A}^\intercal\mathbf{A}\,|\big(\mathbf{A}^\intercal\big)\boldsymbol{b}]$; es decir, si $\big(\mathbf{A}^\intercal\big)\boldsymbol{b}$ 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 rango$\big([\mathbf{A}^\intercal\mathbf{A}\,|\big(\mathbf{A}^\intercal\big)\boldsymbol{b}]\big) \geq$ rango$(\mathbf{A}^\intercal\mathbf{A}) = k$.

Por otra parte, $[\mathbf{A}^\intercal\mathbf{A}\,|\big(\mathbf{A}^\intercal\big)\boldsymbol{b}] = \mathbf{A}^\intercal[\mathbf{A}\,|\boldsymbol{b}]$ que es el producto de la matriz $\mathbf{A}^\intercal$ por la matriz ampliada $[\mathbf{A}\,|\boldsymbol{b}]$, y como recordábamos más arriba rango$\big([\mathbf{A}^\intercal\mathbf{A}\,|\big(\mathbf{A}^\intercal\big)\boldsymbol{b}]\big) =$ rango$(\mathbf{A}^\intercal[\mathbf{A}\,|\boldsymbol{b}]) \leq$ rango$(\mathbf{A}^\intercal) =$ rango$(\mathbf{A}) = k$.

Así que, como por una parte rango$\big([\mathbf{A}^\intercal\mathbf{A}\,|\big(\mathbf{A}^\intercal\big)\boldsymbol{b}]\big) \geq k$ y por la otra rango$\big([\mathbf{A}^\intercal\mathbf{A}\,|\big(\mathbf{A}^\intercal\big)\boldsymbol{b}]\big) \leq k$, se deduce que rango$\big([\mathbf{A}^\intercal\mathbf{A}\,|\big(\mathbf{A}^\intercal\big)\boldsymbol{b}]\big) = k$; que es el rango de la matriz de coeficientes $\mathbf{A}^\intercal\mathbf{A}$ del sistema $\mathbf{A}^\intercal\mathbf{A}\boldsymbol{x} = \big(\mathbf{A}^\intercal\big)\boldsymbol{b}$, y por tanto el sistema siempre tiene solución ($\big(\mathbf{A}^\intercal\big)\boldsymbol{b} \in \mathcal{C}\left(\mathbf{A}^\intercal\mathbf{A}\right)$). $\square$

Se puede obtener el mismo resultado empleando la interpretación geométrica de la proyección ortogonal: cuando proyectamos el vector $\boldsymbol{b}$ sobre el espacio columna de $\mathbf{A}$, descomponemos $\boldsymbol{b}$ en un vector $\boldsymbol{p} = \mathbf{A}\widehat{\boldsymbol{x}}$ (en el espacio columna de $\mathbf{A}$) y un vector $\boldsymbol{e}$ (ortogonal a dicho espacio columna): $\mathbf{A}\widehat{\boldsymbol{x}} + \boldsymbol{e} = \boldsymbol{b}$. Entonces $\mathbf{A}^\intercal\mathbf{A}\widehat{x} + \big(\mathbf{A}^\intercal\big)\boldsymbol{e} = \big(\mathbf{A}^\intercal\big)\boldsymbol{b}$. Pero $\boldsymbol{e}$ es ortogonal al espacio fila de $\mathbf{A}^\intercal$, así que $\big(\mathbf{A}^\intercal\big)\boldsymbol{e} = \mathbf{0}$, y por tanto $\mathbf{A}^\intercal\mathbf{A}\widehat{\boldsymbol{x}} = \big(\mathbf{A}^\intercal\big)\boldsymbol{b}$

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] ($base \times altura$) de un paralelepípedo de dimensión n, cuyas aristas son 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 del hipercubo de dimensión n descrito por las columnas de $\mathbf{I}$ es siempre uno:

$$\text{Volumen}\big(\underset{n\times n}{\mathbf{I}}\big) = 1.$$

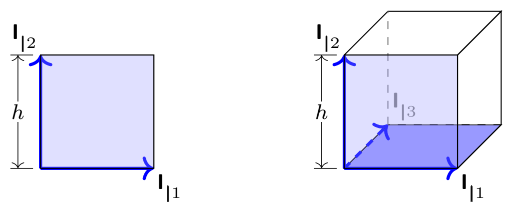


Figura 14.1: $\text{Area}\big[\mathbf{I}_{|1};\ \mathbf{I}_{|2}\big] = 1$ (izquierda) y $\text{Volumen}\big[\mathbf{I}_{|1};\ \mathbf{I}_{|2};\ \mathbf{I}_{|3}\big] = 1$ (derecha).

- Al sumar a una de las arístas (a una de las columnas) un múltiplo de otra arista (columna), el volumen del paralelogramo resultante no cambia; pues se mantiene la misma base y la misma altura h. Por tanto, aplicar una transformación elemental de Tipo I no modifica el volumen:

$$\text{Vol}\big(\mathbf{A}\big) \;=\; \text{Vol}\Big(\mathbf{A}_{\underset{[(\alpha)\boldsymbol{k}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\Big)\text{para } i \neq k.$$

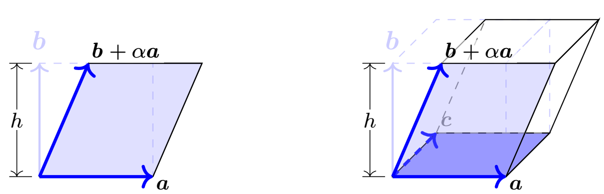


Figura 14.2: $\text{Area}\big[\boldsymbol{a};\ \boldsymbol{b}\big] = \text{Area}\big[\boldsymbol{a};\ (\alpha\boldsymbol{a}+\boldsymbol{b})\big]$ (izquierda) y $\text{Vol}\big[\boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c}\big] = \text{Vol}\big[\boldsymbol{a};\ (\alpha\boldsymbol{a}+\boldsymbol{b});\ \boldsymbol{c}\big]$ (derecha).

[1] en dimensión 2 se denomina área

- Al multiplicar uno de los lados (uno de los vectores) por un escalar, el volumen (o área) queda multiplicado por el valor absoluto de dicho escalar (no hay áreas o volúmenes negativos).

$$|\alpha| \cdot \mathrm{Vol}\big(\mathbf{A}\big) \;=\; |\alpha| \cdot \mathrm{Vol}\big[\dots; \mathbf{A}_{|k}; \dots\big] \;=\; \mathrm{Vol}\big[\dots; \alpha\mathbf{A}_{|k}; \dots\big]$$

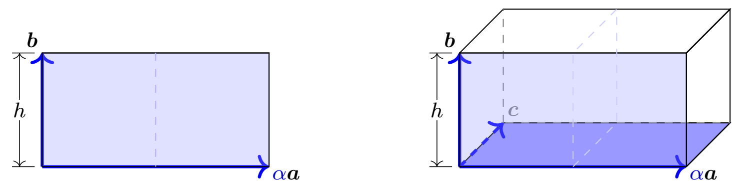


Figura 14.3: $|\alpha| \cdot \mathrm{Area}\big[\boldsymbol{a};\ \boldsymbol{b}\big] = \mathrm{Area}\big[\alpha\boldsymbol{a};\ \boldsymbol{b}\big]$ (izquierda) y $|\alpha| \cdot \mathrm{Vol}\big[\boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c}\big] = \mathrm{Vol}\big[\alpha \cdot \boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c}\big]$ (derecha).

Por tanto, y como caso particular, aplicar la transformación elemental de Tipo II, $\underset{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}$, donde $\alpha \neq 0$, multiplica el volumen por $|\alpha|$.

14.1.2. Las tres primeras propiedades (P-1 to P-3)

Por analogía con la función volumen, definimos el determinante del siguiente modo

Definición 14.1. *Denominamos* función determinante *a toda función que asigna a cada sistema de n vectores de $\mathbb{R}^n$ (o a cada matriz cuadrada de orden n) un número real*

$$\det: \quad \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$$

y que verifica las siguientes tres propiedades:

P-1 P-1 **Determinante de las matrices identidad:**

$$\det \mathbf{I} = 1$$

P-2 P-2 **Sumar una columna a otra no altera el determinante***:*

$$\det \mathbf{A} = \det\left(\mathbf{A}\underset{[(\alpha)\boldsymbol{k}+\boldsymbol{j}]}{\boldsymbol{\tau}}\right) \qquad \textit{para } i \neq k.$$

P-3 P-3 **Multiplicar una columna multiplica el determinante**

$$\alpha \cdot \det \mathbf{A} \;=\; \det\Big[\dots; \alpha\mathbf{A}_{|k}; \dots\Big] \; \textit{para cualquier } k \in \{1:n\} \textit{ y } \alpha \in \mathbb{R}$$

Nótese que como caso particular, concluimos que aplicar una transformación elemental de Tipo II, $\underset{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}$, multiplica el determinante por $\alpha \neq 0$. Consecuentemente, los determinantes pueden tomar valores negativos.

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, por ejemplo, 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: $\begin{vmatrix} a_1 & (\beta b_1) & c_1 \\ a_2 & (\beta b_2) & c_2 \\ a_3 & (\beta b_3) & c_3 \end{vmatrix} = \beta \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix};$ $\qquad \det\left[\boldsymbol{a};\ \beta\boldsymbol{b};\ \boldsymbol{c}\right] = \beta \det\left[\boldsymbol{a};\ \boldsymbol{b};\ \boldsymbol{c}\right].$

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 $\mathbf{A}$:

$$\boxed{\text{Vol}\,(\mathbf{A}) = \text{Valor absoluto de } |\mathbf{A}|}$$

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 viene dado por el objeto encerrado entre ellas: si es un número es el *valor absoluto*, y si es una matriz es el *determinante*. Jugando con esto, podemos decir que

$$\text{Vol}\,(\mathbf{A}) = \text{Valor absoluto de } \det(\mathbf{A}) = |\det \mathbf{A}| = \big|\,|\mathbf{A}|\,\big|.$$

☞ Las propiedades ***P-1***, ***P-2*** y ***P-3*** definen la función determinante. Ahora queda deducir todo lo demás…

14.2. Resto de propiedades (P-4 a P-9)

14.2.1. Determinante de una matriz con una columna de ceros

Ejercicio 53. Demuestre la siguiente propiedad:

$\boxed{\text{P-4}}$ **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

$$\det\left(\mathbf{A}_{\underset{[(\alpha)\boldsymbol{k}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)=|\mathbf{A}|\,;\qquad \det\left(\mathbf{A}_{\underset{[(\alpha)\boldsymbol{k}]}{\boldsymbol{\tau}}}\right)=\alpha\,|\mathbf{A}|\,;\tag{14.1}$$

así que en particular, y puesto que $\det(\mathbf{I})=1$:

$$\det\left(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{k}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)=1\qquad\text{y}\qquad \det\left(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{j}]}{\boldsymbol{\tau}}}\right)=\alpha.$$

Ahora recordando que $\mathbf{A}_{\boldsymbol{\tau}}=\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}}\right)$, y analizando tanto la igualdad de la izquierda como la de la derecha de (14.1), concluimos que

$$\left|\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}}\right)\right|=|\mathbf{A}|\cdot\left|\mathbf{I}_{\boldsymbol{\tau}}\right|\tag{14.2}$$

donde $\mathbf{I}_{\boldsymbol{\tau}}$ es una matriz elemental (nos da igual de qué tipo sea).

14.2.3. Sucesión de transformaciones elementales por columnas

Veamos más resultados relacionados con las matrices elementales.

Ejercicio 54. Demuestre las siguientes proposiciones.

(a) $\det\left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)=|\mathbf{A}|\cdot\left|\mathbf{I}_{\boldsymbol{\tau}_1}\right|\cdots\left|\mathbf{I}_{\boldsymbol{\tau}_k}\right|$.

(b) Si $\mathbf{B}$ es de rango completo, es decir, si $\mathbf{B}=\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$, entonces $|\mathbf{B}|=\left|\mathbf{I}_{\boldsymbol{\tau}_1}\right|\cdots\left|\mathbf{I}_{\boldsymbol{\tau}_k}\right|$, 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 22. Una sucesión $\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k$ de transformaciones *Tipo I* de la matriz $\mathbf{A}$ no altera el determinante.

$$\left|\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right|=\left|\mathbf{A}\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right|=|\mathbf{A}|\cdot\left|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right|=|\mathbf{A}|\cdot 1=|\mathbf{A}|$$

Ejemplo 23. Pero una sucesión de transformaciones *Tipo II* si puede modificar el valor del determinante:

$$\begin{vmatrix}2a&3c\\2b&3d\end{vmatrix}=6\begin{vmatrix}a&c\\b&d\end{vmatrix}$$

14.2.4. Propiedad antisimétrica. Determinante de una matriz singular, de la inversa y del producto de matrices

Permutación (o intercambio) de columnas. Propiedad antisimétrica

P-5 $\boxed{\text{P-5}}$ **[Propiedad antisimétrica]** *Intercambiar dos columnas cambia el signo del determinante.*

Demostración. Un intercambio de columnas es una sucesión de transformaciones elementales *Tipo I* y una única de *Tipo II* que multiplica por -1 una columna (Ejercicio 28 en la página 49). □

Así

$$\begin{vmatrix}c_1&b_1&a_1\\c_2&b_2&a_2\\c_3&b_3&a_3\end{vmatrix}=(-1)\begin{vmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\\a_3&b_3&c_3\end{vmatrix}$$

Matrices singulares. Matrices inversas

Vamos a demostrar conjuntamente dos propiedades más de la función determinante:

P-6 **Si $\mathbf{A}$ es singular entonces $|\mathbf{A}| = 0$.** **P-6**

P-7 $\det\left(\mathbf{A}^{-1}\right) = \left(\det \mathbf{A}\right)^{-1}$. **P-7**

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 $|\mathbf{A}| \cdot \left|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right| = |\mathbf{R}|$ y solo hay dos casos posibles:

$$\begin{cases} \mathbf{A} \text{ singular } (\mathbf{R}_{|n} = \mathbf{0}) : & |\mathbf{R}| = |\mathbf{A}| \cdot \left|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right| = 0 \quad \Rightarrow \quad |\mathbf{A}| = 0 \\ \mathbf{A} \text{ invertible } (\mathbf{R} = \mathbf{I}) : & |\mathbf{R}| = |\mathbf{A}| \cdot \left|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right| = 1 \quad \Rightarrow \quad \left|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right| = \left|\mathbf{A}^{-1}\right| = \frac{1}{|\mathbf{A}|} \end{cases}.$$

□

Así pues, 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 24. Sea $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 2 \end{bmatrix}$, entonces

$$\left[\begin{array}{cc} 1 & 2 \\ 2 & 2 \\ \hline 1 & 0 \\ 0 & 1 \end{array}\right] \xrightarrow[(Tipo I)]{\underset{[(-2)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}} \left[\begin{array}{cc} 1 & 0 \\ 2 & -2 \\ \hline 1 & -2 \\ 0 & 1 \end{array}\right] \xrightarrow[(Tipo II)]{\underset{[(-1/2)\mathbf{2}]}{\boldsymbol{\tau}}} \left[\begin{array}{cc} 1 & 0 \\ 2 & 1 \\ \hline 1 & 1 \\ 0 & -1/2 \end{array}\right] \xrightarrow[(Tipo I)]{\underset{[(-2)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}} \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline -1 & 1 \\ 1 & -1/2 \end{array}\right]$$

Por tanto, $\left|\mathbf{A}^{-1}\right| = \left|\mathbf{I}_{\underset{[(-2)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}\right| \cdot \left|\mathbf{I}_{\underset{[(-1/2)\mathbf{2}]}{\boldsymbol{\tau}}}\right| \cdot \left|\mathbf{I}_{\underset{[(-2)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}}\right| = 1 \cdot \frac{-1}{2} \cdot 1 = \frac{-1}{2}$; 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 $\mathbf{B}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \mathbf{L}$ 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

$$\left(\mathbf{AB}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)\right)_{|n} = \left(\mathbf{AL}\right)_{|n} = \mathbf{0};$$

y como $\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right)_{|n} \neq \mathbf{0}$, necesariamente $(\mathbf{AB})$ es singular (véase Corolario 8.1.2 en la página 100).

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. $\begin{cases} \text{Si } \mathbf{B} \text{ es singular, también lo es} \mathbf{AB} \Rightarrow & \det(\mathbf{AB}) = 0 = \det(\mathbf{A}) \cdot \det(\mathbf{B}) \\ \\ \text{Si } \mathbf{B} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k} \Rightarrow & \det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B}) \qquad \text{(Ecuación (14.3))} \end{cases}$ □

14.2.5. Determinante de la matriz transpuesta.

EJERCICIO 55.

(a) ¿Qué relación hay entre el determinante de $\mathbf{I}_{\boldsymbol{\tau}}$ y el determinante de su transpuesta ${}_{\boldsymbol{\tau}}\mathbf{I}$?
(b) Sea $\mathbf{B}$ de rango completo, demuestre que $|\mathbf{B}| = |\mathbf{B}^{\intercal}|$.

P-9 **P-9** **Determinante de la transpuesta**

$$|\mathbf{A}| = |\mathbf{A}^{\intercal}|.$$

Demostración. Por una parte, si $\mathbf{A}$ es singular, $\mathbf{A}^{\intercal}$ también es singular (véase Proposición 9.5.13 en la página 113); 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. □

Librería NAcAL para Python

```
A = Matrix([ [1,-1,0], [1,1,1], [2,0,3] ])
A.determinante()
( Matrix(A) & T({ 1,3}) ).determinante()
(~Matrix(A) & T((10,2)) ).determinante()
```

Cofactores y fórmulas para el cálculo del determinante

Si la matriz elemental $\mathbf{I}_{\boldsymbol{\tau}}$ es de orden n, sabemos que necesariamente la transformación $\boldsymbol{\tau}$ actúa únicamente sobre las n primeras columnas; 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 $\boldsymbol{p}, \boldsymbol{s}$ son necesariamente menores o iguales a n.

Por tanto, para la siguiente matriz por bloques de orden $n+1$, e idéntica transformación $\boldsymbol{\tau}$ se verifica que

$$\begin{bmatrix} \mathbf{B}_{\boldsymbol{\tau}} & \\ & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{B} & \\ & 1 \end{bmatrix}_{\boldsymbol{\tau}}. \tag{15.1}$$

Y si $\mathbf{B} = \mathbf{I}$ es la matriz identidad, y puesto que $\begin{bmatrix} \mathbf{I} & \\ & 1 \end{bmatrix}_{\boldsymbol{\tau}}$ y $\mathbf{I}_{\boldsymbol{\tau}}$ son matrices elementales del mismo tipo, ambas tienen idéntico determinante. Repitiendo k veces el paso dado en (15.1) tenemos que

$$\begin{vmatrix} \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \\ & 1 \end{vmatrix} = \det\left(\begin{bmatrix} \mathbf{I} & \\ & 1 \end{bmatrix}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} \right) = \det\left(\begin{bmatrix} \mathbf{I} & \\ & 1 \end{bmatrix}_{\boldsymbol{\tau}_1} \cdots \begin{bmatrix} \mathbf{I} & \\ & 1 \end{bmatrix}_{\boldsymbol{\tau}_k} \right) = \left|\mathbf{I}_{\boldsymbol{\tau}_1}\right| \cdots \left|\mathbf{I}_{\boldsymbol{\tau}_k}\right| = \left|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right|. \tag{15.2}$$

Vamos a emplear este resultado para relacionar el determinante de dos matrices que tienen distinto orden:

Proposición 15.0.1. *Si* $\underset{n\times n}{\mathbf{A}}$ *tiene la forma* $\mathbf{A} = \left[\begin{array}{ccc|c} & & & 0 \\ & \underset{n\times n}{\mathbf{B}} & & \vdots \\ & & & 0 \\ \hline 0 & \cdots & 0 & 1 \end{array}\right]$, *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

$$|\mathbf{A}| = \begin{vmatrix} \mathbf{B} & \\ & 1 \end{vmatrix} = \begin{vmatrix} \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} & \\ & 1 \end{vmatrix} = \left|\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right| = |\mathbf{B}|.$$

Y si $\mathbf{B}$ es singular, es decir si 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 matrices triangulares

EJERCICIO 56.

(a) ¿Cómo es el determinante de una matriz triangular inferior $\mathbf{L}$ de rango completo?
(b) ¿Cuál es el determinante de una matriz cuadrada y triangular con algún elemento nulo en su diagonal
(c) ¿Cómo es el determinante de una matriz triangular superior $\mathbf{U}$?

15.2. Cálculo del determinante por eliminación Gaussiana

Por tanto, el determinante de cualquier matriz cuadrada y triangular es igual al producto de los elementos de la diagonal principal. Este descubrimiento nos permite calcular el determinante con menos operaciones, pues basta con encontrar una forma escalonada de la matriz extendida, compensando en su última columna las transformaciones elementales Tipo II aplicadas sobre las n primeras columnas.

Ejemplo 25. Considere de nuevo la matriz $\begin{bmatrix}1&5\\2&3\end{bmatrix}$ entonces

$$\left[\begin{array}{cc|c}1&5&\\2&3&\\\hline &&1\end{array}\right]\xrightarrow[{[(-5)\mathbf{1}+\mathbf{2}]}]{\boldsymbol{\tau}}\left[\begin{array}{cc|c}1&0&\\2&-7&\\\hline &&1\end{array}\right]$$

Multiplicando los elementos de la diagonal obtenemos el valor del determiante de $\mathbf{A}$. En este caso $\det\mathbf{A}=-7$.

Ejemplo 26. Para la matriz $\begin{bmatrix}0&2&1\\9&6&3\\0&1&1\end{bmatrix}$ tenemos

$$\left[\begin{array}{ccc|c}0&2&1&\\9&6&3&\\0&1&1&\\\hline &&&1\end{array}\right]\xrightarrow[\substack{[(2)\mathbf{3}]\\ [(-1)\mathbf{2}+\mathbf{3}]\\ [(\frac{1}{2})\mathbf{4}]}]{\boldsymbol{\tau}}\left[\begin{array}{ccc|c}0&2&0&\\9&6&0&\\0&1&1&\\\hline &&&\frac{1}{2}\end{array}\right]\xrightarrow[\substack{[(3)\mathbf{2}]\\ [(-2)\mathbf{1}+\mathbf{2}]\\ [(\frac{1}{3})\mathbf{4}]}]{\boldsymbol{\tau}}\left[\begin{array}{ccc|c}0&6&0&\\9&0&0&\\0&3&1&\\\hline &&&\frac{1}{6}\end{array}\right]\xrightarrow[\substack{[\mathbf{1}\rightleftharpoons\mathbf{2}]\\ [(-1)\mathbf{4}]}]{\boldsymbol{\tau}}\left[\begin{array}{ccc|c}6&0&0&\\0&9&0&\\3&0&1&\\\hline &&&\frac{-1}{6}\end{array}\right]$$

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 y $\mathbf{I}$ de orden n, entonces, aplicando la anterior igualdad repetidamente n veces, deducimos que

$$\begin{vmatrix}\mathbf{A}&\underset{m\times n}{\mathbf{0}}\\\underset{n\times m}{\mathbf{0}}&\underset{n\times n}{\mathbf{I}}\end{vmatrix}=\begin{vmatrix}\mathbf{A}&\underset{m\times(n-1)}{\mathbf{0}}\\\underset{(n-1)\times m}{\mathbf{0}}&\underset{(n-1)\times(n-1)}{\mathbf{I}}\end{vmatrix}=\begin{vmatrix}\mathbf{A}&\underset{m\times(n-2)}{\mathbf{0}}\\\underset{(n-2)\times m}{\mathbf{0}}&\underset{(n-2)\times(n-2)}{\mathbf{I}}\end{vmatrix}=\dots=\begin{vmatrix}\mathbf{A}&0\\0&1\end{vmatrix}=|\mathbf{A}|.$$

Pero todo lo realizado en las anteriores secciones se podía 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 particionadas por bloques. En la Sección 3.B.2 se mostró cómo calcular el producto de matrices particionadas por bloques siguiendo la fórmula de la Ecuación 3.2 en la página 36.

EJERCICIO 57. Demuestre las siguientes propiedades:

(a) Sean $\mathbf{A}$ de orden m y $\mathbf{B}$ de orden n, entonces $\begin{vmatrix}\mathbf{A}&\underset{m\times n}{\mathbf{0}}\\\underset{n\times m}{\mathbf{0}}&\mathbf{B}\end{vmatrix}=|\mathbf{A}|\cdot|\mathbf{B}|$.

(b) Sean $\mathbf{A}$ de orden m, $\mathbf{B}$ de orden n y $\mathbf{C}$ de orden n por m, entonces

$$\begin{vmatrix}\mathbf{A}&\underset{m\times n}{\mathbf{0}}\\\mathbf{C}&\mathbf{B}\end{vmatrix}=|\mathbf{A}|\cdot|\mathbf{B}|.$$

15.4. Fórmulas sencillas para matrices de orden menor que 4

Para matrices de orden 1, 2 o 3, hay formulas sencillas. ¡Ojo, para matrices de orden 4 o más NO HAY REGLAS SENCILLAS! No extrapole que lo que funciona para una matriz 3 por 3 se puede hacer de manera similar para una matriz de orden 4 por 4.

Para matrices de orden 1: Para $\mathbf{A} = \begin{bmatrix} a \end{bmatrix}$ tenemos

$$\left[\begin{array}{c|c} a & \\ \hline & 1 \end{array}\right] \qquad \Rightarrow \qquad \boxed{\det \mathbf{A} = \begin{vmatrix} a \end{vmatrix} = a.}$$

Para matrices de orden 2. Deduzcamos la formula general para una matriz de la forma $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

- Caso $a \neq 0$:

$$\left[\begin{array}{cc|c} a & b & \\ c & d & \\ \hline & & 1 \end{array}\right] \xrightarrow[\substack{[(\frac{1}{a})\mathbf{3}]}]{\substack{\boldsymbol{\tau} \\ [(a)\mathbf{2}] \\ [(-b)\mathbf{1}+\mathbf{2}] \\ [(\frac{1}{a})\mathbf{3}]}} \left[\begin{array}{cc|c} a & 0 & 0 \\ c & ad-cb & 0 \\ \hline 0 & 0 & \frac{1}{a} \end{array}\right] \qquad \Rightarrow \qquad \det \mathbf{A} = ad - bc.$$

- Caso $a = 0$:

$$\left[\begin{array}{cc|c} a & b & \\ c & d & \\ \hline & & 1 \end{array}\right] \xrightarrow{\substack{\boldsymbol{\tau} \\ [\mathbf{1}\rightleftharpoons\mathbf{2}] \\ [(-1)\mathbf{3}]}} \left[\begin{array}{cc|c} b & a & \\ d & c & \\ \hline & & -1 \end{array}\right] \text{(que es como el caso anterior)} \qquad \Rightarrow \qquad \det \mathbf{A} = (bc-ad)(-1) = ad-bc.$$

Otra vez la misma regla. Por tanto, siempre

$$\boxed{\det \mathbf{A} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.}$$

Para matrices de orden 3. Deduzcamos la formula general para una matriz de la forma $\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$.

- Caso $a \neq 0$:

$$\left[\begin{array}{ccc|c} a & b & c & \\ d & e & f & \\ g & h & i & \\ \hline & & & 1 \end{array}\right] \xrightarrow{\substack{\boldsymbol{\tau} \\ [(a)\mathbf{2}] \\ [(a)\mathbf{3}] \\ [(-b)\mathbf{1}+\mathbf{2}] \\ [(-c)\mathbf{1}+\mathbf{3}] \\ [(\frac{1}{a})\mathbf{4}] \\ [(\frac{1}{a})\mathbf{4}]}} \left[\begin{array}{ccc|c} a & 0 & 0 & \\ d & ae-bd & af-cd & \\ g & ah-bg & ai-cg & \\ \hline & & & \frac{1}{a^2} \end{array}\right]$$

donde ahora tenemos una matriz diagonal por bloques y por tanto

$$\begin{aligned}
\det \mathbf{A} = |a| \cdot \begin{vmatrix} ae-bd & af-cd \\ ah-bg & ai-cg \end{vmatrix} \cdot \frac{1}{a^2} =& \frac{1}{a}\Big((ae-bd)(ai-cg) - (af-cd)(ah-bg)\Big) \\
=& \frac{1}{a}\Big((a^2ei - aecg - bdai + bdcg) - (a^2fh - afbg - cdah + cdbg)\Big) \\
=& (aei - ecg - bdi + \frac{bdcg}{a}) - (afh - fbg - cdh + \frac{cdbg}{a}) \\
=& aei + fbg + cdh - ecg - bdi - afh + \frac{bdcg}{a} - \frac{cdbg}{a} \\
=& aei + fbg + cdh - ecg - bdi - afh.
\end{aligned}$$

- Caso $a = 0$ y $b \neq 0$:

$$\left[\begin{array}{ccc|c} a & b & c & \\ d & e & f & \\ g & h & i & \\ \hline & & & 1 \end{array}\right] \xrightarrow[\quad]{\substack{\boldsymbol{\tau} \\ [\mathbf{1}\rightleftharpoons\mathbf{2}] \\ [(-1)\mathbf{4}]}} \left[\begin{array}{ccc|c} b & a & c & \\ e & d & f & \\ h & g & i & \\ \hline & & & 1 \end{array}\right] \text{(como el caso anterior)} \quad \Rightarrow \det \mathbf{A} = aei{+}fbg{+}cdh{-}ecg{-}bdi{-}afh.$$

- ...y de modo similar para el caso $a = 0$, $b = 0$ y $c \neq 0$.

Por tanto, siempre

$$\boxed{\det \mathbf{A} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + fbg + cdh - ecg - bdi - afh.}$$

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 (para abreviar, asumiremos que las componentes de la matriz son tales que todas las fraciones que aparecen en la derivación fracción están definidas, es decir, que todos los denominadores son distintos de cero).

$$\left[\begin{array}{cccc|c} a & b & c & d & 0 \\ e & f & h & i & 0 \\ k & l & m & n & 0 \\ o & p & q & r & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{\boldsymbol{\tau} \\ [(1)\mathbf{5}]}]{\substack{\boldsymbol{\tau} \\ [(-\frac{b}{a})\mathbf{1}+\mathbf{2}] \\ [(-\frac{c}{a})\mathbf{1}+\mathbf{3}] \\ [(-\frac{d}{a})\mathbf{1}+\mathbf{4}]}} \left[\begin{array}{cccc|c} a & 0 & 0 & 0 & 0 \\ e & f-\frac{be}{a} & h-\frac{ce}{a} & i-\frac{de}{a} & 0 \\ k & l-\frac{bk}{a} & m-\frac{ck}{a} & n-\frac{dk}{a} & 0 \\ o & p-\frac{bo}{a} & q-\frac{co}{a} & r-\frac{do}{a} & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{\boldsymbol{\tau} \\ [(1)\mathbf{5}]}]{\substack{\boldsymbol{\tau} \\ \left[\left(\frac{-ah+ce}{af-be}\right)\mathbf{2}+\mathbf{3}\right] \\ \left[\left(\frac{-ai+de}{af-be}\right)\mathbf{2}+\mathbf{4}\right]}}$$

$$\left[\begin{array}{cccc|c} a & 0 & 0 & 0 & 0 \\ e & f-\frac{be}{a} & 0 & 0 & 0 \\ k & l-\frac{bk}{a} & \frac{afm-ahl-bem+bhk+cel-cfk}{af-be} & \frac{afn-ail-ben+bik+del-dfk}{af-be} & 0 \\ o & p-\frac{bo}{a} & \frac{afq-ahp-beq+bho+cep-cfo}{af-be} & \frac{afr-aip-ber+bio+dep-dfo}{af-be} & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow[\substack{\boldsymbol{\tau} \\ [(1)\mathbf{5}]}]{\substack{\boldsymbol{\tau} \\ \left[\left(\frac{-afn+ail+ben-bik-del+dfk}{afm-ahl-bem+bhk+cel-cfk}\right)\mathbf{3}+\mathbf{4}\right]}}$$

$$\left[\begin{array}{cccc|c} a & 0 & 0 & 0 & 0 \\ e & f-\frac{be}{a} & 0 & 0 & 0 \\ k & l-\frac{bk}{a} & \frac{afm-ahl-bem+bhk+cel-cfk}{af-be} & 0 & 0 \\ o & p-\frac{bo}{a} & \frac{afq-ahp-beq+bho+cep-cfo}{af-be} & r+\frac{\left(-n+\frac{\left(i-\frac{de}{a}\right)\left(l-\frac{bk}{a}\right)}{f-\frac{be}{a}}+\frac{dk}{a}\right)\left(-q+\frac{\left(h-\frac{ce}{a}\right)\left(p-\frac{bo}{a}\right)}{f-\frac{be}{a}}+\frac{co}{a}\right)}{-m+\frac{\left(h-\frac{ce}{a}\right)\left(l-\frac{bk}{a}\right)}{f-\frac{be}{a}}+\frac{ck}{a}}-\frac{\left(i-\frac{de}{a}\right)\left(p-\frac{bo}{a}\right)}{f-\frac{be}{a}}-\frac{do}{a} & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

Ahora multiplicando los elementos de la diagonal y simplificando las expresiones se llega a la expresión final:

$$\begin{aligned} \det \mathbf{A} = \begin{vmatrix} a & b & c & d \\ e & f & h & i \\ k & l & m & n \\ o & p & q & r \end{vmatrix} \\ = afmr - afnq - ahlr + ahnp + ailq - aimp - bemr + benq + bhkr - bhno - bikq + bimo \\ + celr - cenp - cfkr + cfno + cikp - cilo - delq + demp + dfkq - dfmo - dhkp + dhlo. \end{aligned}$$

Está claro que la esta fórmula no es ni sencilla ni fácil de recordar... *¡No puedo ni 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

Pero afortunadamente hay otra forma de calcular determinantes de matrices de orden n empleando fórmulas para calcular el determinante de orden $n-1$. Esta descripción recursiva de la función determinante se llama expansión de Laplace. No obstante, con cualquier procedimiento el cálculo del determinante de una matriz genérica de orden elevado es computacionalmente intenso.

Para deducir la citada Expansión de Laplace, antes necesitamos enunciar una última propiedad de la función determinate, y definir los *menores* y los *cofactores*.

15.6. Expansión de Laplace (desarrollo por cofactores).

15.6.1. Propiedad multilineal

P-10 **Propiedad multilineal** **P-10**

$$\det\Big[\ldots;(\beta\boldsymbol{b}+\psi\boldsymbol{c});\ldots\Big]=\beta\det\Big[\ldots;\boldsymbol{b};\ldots\Big]+\psi\det\Big[\ldots;\boldsymbol{c};\ldots\Big]$$

Demostración. Puesto que se verifica la propiedad del producto **P-3**, nos basta con demostrar que $\det\Big[\ldots;\boldsymbol{b}+\boldsymbol{c};\ldots\Big]=\det\Big[\ldots;\boldsymbol{b};\ldots\Big]+\det\Big[\ldots;\boldsymbol{c};\ldots\Big]$. Haremos la demostración para la primera columna. La demostración para el resto de columnas es similar.

Consideremos dos casos. Si $\boldsymbol{y}_1,\ldots,\boldsymbol{y}_{n-1}$ son linealmente dependientes, entonces la demostración es inmediata:

$$\det\Big[\boldsymbol{a}+\boldsymbol{b};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big]=0=0+0=\det\Big[\boldsymbol{a};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big]+\det\Big[\boldsymbol{b};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big].$$

En caso contrario, existe $\boldsymbol{x}$ tal que $\Big[\boldsymbol{x};\boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big]$ es una base de $\mathbb{R}^n$. Así,

$$\det\Big[\alpha\boldsymbol{x}+\psi_1\boldsymbol{y}_1+\cdots+\psi_{n-1}\boldsymbol{y}_{n-1};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big]\overset{*}{=}\det\Big[\alpha\boldsymbol{x};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big]\tag{15.3}$$
$$=\alpha\det\Big[\boldsymbol{x};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big]\tag{15.4}$$

(*) ya que con una secuencia de transformaciones elementales se puede "simplificar" la primera columna. Ahora, si

$$\begin{aligned}\boldsymbol{a}&=\alpha\boldsymbol{x}+\psi_1\boldsymbol{y}_1+\cdots+\psi_{n-1}\boldsymbol{y}_{n-1}\\ \boldsymbol{b}&=\beta\boldsymbol{x}+\gamma_1\boldsymbol{y}_1+\cdots+\gamma_{n-1}\boldsymbol{y}_{n-1}\end{aligned}$$

tendremos que

$$\boldsymbol{a}+\boldsymbol{b}=(\alpha+\beta)\boldsymbol{x}+(\psi_1+\gamma_1)\boldsymbol{y}_1+\cdots+(\psi_{n-1}+\gamma_{n-1})\boldsymbol{y}_{n-1}$$

y consecuentente

$$\begin{aligned}\det\Big[\boldsymbol{a}+\boldsymbol{b};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big]&=(\alpha+\beta)\det\Big[\boldsymbol{x};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big] && \text{por (15.4)}\\ &=\alpha\det\Big[\boldsymbol{x};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big]+\beta\det\Big[\boldsymbol{x};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big] && \text{pues }\det\mathsf{A}\text{ es un número}\\ &=\det\Big[\alpha\boldsymbol{x};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big]+\det\Big[\beta\boldsymbol{x};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big] && \text{por } \textbf{P-3}\\ &=\det\Big[\boldsymbol{a};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big]+\det\Big[\boldsymbol{b};\ \boldsymbol{y}_1;\ldots;\boldsymbol{y}_{n-1}\Big] && \text{por (15.3)}\end{aligned}$$

□

Ejemplo 27. Para $\mathbb{R}^2$

$$\begin{vmatrix} a+\alpha & c \\ b+\beta & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} + \begin{vmatrix} \alpha & c \\ \beta & d \end{vmatrix};$$

es decir

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix} = \begin{vmatrix} a & c \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & c \\ b & d \end{vmatrix}.$$

15.6.2. Menores y cofactores

☞ En esta sección veremos como calcular los determinantes mediante formulas más sencillas. Fórmulas que expresan determinantes de matrices de orden n como sumas de determinantes de matrices de un orden más pequeño $(n-1)$.

Proposición 15.6.1. *Si* $\underset{n\times n}{\mathbf{A}}$ *es de la forma* $\left[\begin{array}{ccc|c} & & & 0 \\ & \underset{(n-1)\times(n-1)}{\mathbf{B}} & & \vdots \\ & & & 0 \\ \hline a_{11} & \cdots & a_{1n-1} & 1 \end{array}\right]$ *entonces* $\boxed{\det(\mathbf{A}) = \det(\mathbf{B}).}$

Demostración. Mediante operaciones elementales por columnas de *Tipo I*, podemos reducir $\mathbf{A}$ a

$$\mathbf{A} = \left[\begin{array}{ccc|c} & & & 0 \\ & \mathbf{B} & & \vdots \\ & & & 0 \\ \hline a_{11} & \cdots & a_{1n-1} & 1 \end{array}\right] \xrightarrow{\text{transformaciones } \textit{Tipo I}} \left[\begin{array}{ccc|c} & & & 0 \\ & \mathbf{B} & & \vdots \\ & & & 0 \\ \hline 0 & \cdots & 0 & 1 \end{array}\right] \equiv \begin{bmatrix} \mathbf{B} & \\ & 1 \end{bmatrix}$$

Donde solo hemos empleado transformaciones de *Tipo I*, así, por la Proposición 15.0.1 sabemos que

$$|\mathbf{A}| = \begin{vmatrix} \mathbf{B} & \\ & 1 \end{vmatrix} = |\mathbf{B}|. \tag{15.5}$$

□

Nueva notación y definición de menores y cofactores

Sea $\boldsymbol{q} = (q_1, \ldots, q_n)$ de $\mathbb{R}^n$, si quitamos la $j$ésima componente, q_j, denotamos al nuevo vector de $\mathbb{R}^{(n-1)}$ como

$$\boldsymbol{q}^{\Rsh j} = \left(q_1, \ldots,\ q_{j-1},\quad q_{j+1},\ \ldots, q_n\right) \in \mathbb{R}^{n-1}.$$

Análogamente, si $\mathbf{A}$ es de orden m por n, denotamos a la submatriz que resulta de quitar la $i$ésima fila con ${}^{i\Lsh}\mathbf{A}$, y a la submatriz que resulta de quitar la $j$ésima columna con $\mathbf{A}^{\Rsh j}$. Así, ${}^{i\Lsh}\mathbf{A}^{\Rsh j}$, es la submatriz de orden $m-1$ por $n-1$ que resulta de quitar la fila $i$ésima y la columna $j$ésima.

Definición 15.1 (menores y cofactores)**.** *Denotamos a la submatriz resultante de eliminar la fila i y la columna j con*

$${}^{i\Lsh}\mathbf{A}^{\Rsh j};$$

su determinante $\det\left({}^{i\Lsh}\mathbf{A}^{\Rsh j}\right)$, *se denomina* menor *de a_{ij}.*

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,

$$\text{cof}_{ij}\left(\mathbf{A}\right) = (-1)^{i+j} \det\left({}^{i\Lsh}\mathbf{A}^{\Rsh j}\right)$$

es el cofactor *de a_{ij}.*

Ejemplo 28. Para $\quad \mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, tenemos

$$ {}^{1\Lsh}\mathbf{A}^{\Rsh 2} = \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix}, \qquad {}^{3\Lsh}\mathbf{A}^{\Rsh 3} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} $$

así

$$ \mathrm{cof}_{12}\left(\mathbf{A}\right) = (-1)^{1+2}\det\left({}^{1\Lsh}\mathbf{A}^{\Rsh 2}\right) = (-1)\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix}. $$

y

$$ \mathrm{cof}_{33}\left(\mathbf{A}\right) = (-1)^{3+3}\det\left({}^{3\Lsh}\mathbf{A}^{\Rsh 3}\right) = \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}. $$

15.6.3. Desarrollo del determinante por cofactores (Expansión de Laplace).

Teorema 15.6.2 ([**Expansión de Laplace**])**.** *Para cualquier matriz de orden n,* $\det(\mathbf{A})$ *se puede expresar como suma de los productos de los elementos de cualquier columna (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

$$ \det(\mathbf{A}) = \sum_{j=1}^{n} a_{ij}\,\mathrm{cof}_{ij}\left(\mathbf{A}\right), \qquad \textit{expansión por la fila } i\textit{ésima} $$

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:

$$ (\mathbf{A})_{\underset{[j\rightleftharpoons(j+1)]}{\boldsymbol{\tau}},\ \underset{[(j+1)\rightleftharpoons(j+2)]}{\boldsymbol{\tau}},\ \cdots,\ \underset{[(n-1)\rightleftharpoons n]}{\boldsymbol{\tau}}} = \left[\mathbf{A}_{|1},\ldots,\ \mathbf{A}_{|j-1},\ \mathbf{A}_{|j+1},\ \ldots,\mathbf{A}_{|n}\middle|\mathbf{A}_{|j}\right]; = \left[\ \mathbf{A}^{\Rsh j}\ \middle|\ \mathbf{A}_{|j}\right]. $$

Puesto que hay $n-j$ permutaciones en la sucesión, tenemos que

$$ \det\mathbf{A} \quad = \quad \det\left[\mathbf{A}_{|1},\ldots,\mathbf{A}_{|n}\right] \quad = \quad (-1)^{n-j}\cdot\det\left[\ \mathbf{A}^{\Rsh j}\ \middle|\ \mathbf{A}_{|j}\right]. $$

Escribiendo $\mathbf{A}_{|j}$ como $\quad \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} = \begin{pmatrix} a_{1j} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_{2j} \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{nj} \end{pmatrix}$, y empleando la propiedad multilineal (***P-12***):

$$ \det\mathbf{A} = (-1)^{n-j}\left(\det\left[\ \mathbf{A}^{\Rsh j}\ \middle|\ \begin{matrix} a_{1j} \\ 0 \\ \vdots \\ 0 \end{matrix}\right] + \det\left[\ \mathbf{A}^{\Rsh j}\ \middle|\ \begin{matrix} 0 \\ a_{2j} \\ \vdots \\ 0 \end{matrix}\right] + \cdots + \det\left[\ \mathbf{A}^{\Rsh j}\ \middle|\ \begin{matrix} 0 \\ \vdots \\ 0 \\ a_{nj} \end{matrix}\right]\right). $$

Si "hundimos" la primera fila de la primera matriz hasta la última posición mediante una sucesión de $(n-1)$ intercambios: $\underset{[1\rightleftharpoons 2]}{\boldsymbol{\tau}},\ \underset{[2\rightleftharpoons 3]}{\boldsymbol{\tau}},\ \cdots,\ \underset{[(n-1)\rightleftharpoons n]}{\boldsymbol{\tau}}$; y la segunda fila de la segunda matriz mediante $(n-2)$

intercambios: $\underset{[\boldsymbol{2}\rightleftharpoons\boldsymbol{3}]}{\boldsymbol{\tau}}, \underset{[\boldsymbol{3}\rightleftharpoons\boldsymbol{4}]}{\boldsymbol{\tau}}, \cdots, \underset{[(\boldsymbol{n-1})\rightleftharpoons\boldsymbol{n}]}{\boldsymbol{\tau}}$; —y en general, “hundimos” la fila $i$ésima de la matriz $i$ésima con los $(n-i)$ intercambios:

$$\underset{[\boldsymbol{i}\rightleftharpoons(\boldsymbol{i+1})]}{\boldsymbol{\tau}}, \underset{[(\boldsymbol{i+1})\rightleftharpoons(\boldsymbol{i+2})]}{\boldsymbol{\tau}}, \cdots, \underset{[(\boldsymbol{n-1})\rightleftharpoons\boldsymbol{n}]}{\boldsymbol{\tau}}$$

obtenemos:

$$\det \mathbf{A} = (-1)^{n-j}\left((-1)^{n-1}\cdot\det\left[\begin{array}{c|c} {}^{1\urcorner}\mathbf{A}^{\urcorner j} & \mathbf{0} \\ \hline {}_{1|}\mathbf{A}^{\urcorner j} & a_{1j}\end{array}\right] + (-1)^{n-2}\cdot\det\left[\begin{array}{c|c} {}^{2\urcorner}\mathbf{A}^{\urcorner j} & \mathbf{0} \\ \hline {}_{2|}\mathbf{A}^{\urcorner j} & a_{2j}\end{array}\right] + \cdots\right.$$

$$\left.\cdots + (-1)\cdot\det\left[\begin{array}{c|c} {}^{(n-1)\urcorner}\mathbf{A}^{\urcorner j} & \mathbf{0} \\ \hline {}_{(n-1)|}\mathbf{A}^{\urcorner j} & a_{(n-1)j}\end{array}\right] + \det\left[\begin{array}{c|c} {}^{n\urcorner}\mathbf{A}^{\urcorner j} & \mathbf{0} \\ \hline {}_{n|}\mathbf{A}^{\urcorner j} & a_{nj}\end{array}\right]\right),$$

donde ${}_{i|}\mathbf{A}^{\urcorner j}$ es la fila $i$ésima de la submatriz $\mathbf{A}^{\urcorner j}$. Así, por la Proposición 15.6.1

$$\det(\mathbf{A}) = (-1)^{n-j}\sum_{i=1}^{n} a_{ij}(-1)^{n-i}\det\left({}^{i\urcorner}\mathbf{A}^{\urcorner j}\right);$$

y puesto que $(-1)^{n-j}\cdot(-1)^{n-i} = (-1)^{2n-(i+j)} = (-1)^{-(i+j)} = (-1)^{i+j}$, sustituyendo llegamos a

$$\boxed{\det(\mathbf{A}) = \textstyle\sum_{i=1}^{n} a_{ij}\operatorname{cof}_{ij}\left(\mathbf{A}\right).}$$

$\square$

Ejercicio 58. 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 59. Calcule el determinante de la matriz genérica de orden 3, desarrollado por los cofactores de la segunda columna.

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$$

15.7. Aplicación de los determinantes

15.7.1. Regla de Cramer para la resolución de un sistema de ecuaciones

Suponga el sistema de ecuaciones $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$, donde $|\mathbf{A}| \neq 0$. Sabemos que dicho sistema siempre tiene solución única, sea cual sea el vector $\boldsymbol{b}$. Dicha solución es

$$\boldsymbol{x} = \mathbf{A}^{-1}\boldsymbol{b};$$

que verifica:

$$\boldsymbol{b} = (\mathbf{A}_{|1})x_1 + \cdots + (\mathbf{A}_{|j})x_j + \cdots + (\mathbf{A}_{|n})x_n, \tag{15.6}$$

donde los coeficiente x_i son los componentes[1] del vector solución $\boldsymbol{x}$.

Calculemos el determinante de una nueva matriz, idéntica a $\mathbf{A}$ excepto por que su $j$ésima columna $\mathbf{A}_{|j}$ 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 (15.6) 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 encontramos la regla de Cramer para la resolución de sistemas de ecuaciones determinados: *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 15.2. *Para $\underset{n\times n}{\mathbf{A}}$ la matriz* **Adj($\mathbf{A}$)** *(la* matriz adjunta *de $\mathbf{A}$) se define como la transpuesta de la matriz resultante de sustituir cada elemento a_{ij} de $\mathbf{A}$ por su correspondiente cofactor $\text{cof}_{ij}(\mathbf{A})$. Es decir,*

$$\mathbf{Adj}(\mathbf{A}) = \begin{bmatrix} \text{cof}_{11}(\mathbf{A}) & \text{cof}_{12}(\mathbf{A}) & \cdots & \text{cof}_{1n}(\mathbf{A}) \\ \text{cof}_{21}(\mathbf{A}) & \text{cof}_{22}(\mathbf{A}) & \cdots & \text{cof}_{2n}(\mathbf{A}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cof}_{n1}(\mathbf{A}) & \text{cof}_{n2}(\mathbf{A}) & \cdots & \text{cof}_{nn}(\mathbf{A}) \end{bmatrix}^{\mathsf{T}}$$

¿Qué obtenemos si multiplicamos la matriz adjunta de $\mathbf{A}$ por la matriz $\mathbf{A}$?

$$[\mathbf{Adj}(\mathbf{A})] \cdot \mathbf{A} = \begin{bmatrix} \text{cof}_{11}(\mathbf{A}) & \text{cof}_{21}(\mathbf{A}) & \cdots & \text{cof}_{n1}(\mathbf{A}) \\ \text{cof}_{12}(\mathbf{A}) & \text{cof}_{22}(\mathbf{A}) & \cdots & \text{cof}_{n2}(\mathbf{A}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cof}_{1n}(\mathbf{A}) & \text{cof}_{2n}(\mathbf{A}) & \cdots & \text{cof}_{nn}(\mathbf{A}) \end{bmatrix} \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}}_{\mathbf{A}}$$

$$= \begin{bmatrix} |\mathbf{A}| & 0 & \cdots & 0 \\ 0 & |\mathbf{A}| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |\mathbf{A}| \end{bmatrix} = |\mathbf{A}| \cdot \mathbf{I}$$

[1] las coordenadas de $\boldsymbol{b}$ en la base formada por las columnas de $\mathbf{A}$

$$\left(\frac{1}{|\mathbf{A}|} \cdot \mathbf{Adj(A)}\right) \cdot \mathbf{A} = \mathbf{I}$$

El primer elemento de la diagonal de la matriz resultante es el desarrollo de $|\mathbf{A}|$ por la 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:

$$a_{1j} \operatorname{cof}_{1j}(\mathbf{A}) + a_{2j} \operatorname{cof}_{2j}(\mathbf{A}) + a_{3j} \operatorname{cof}_{3j}(\mathbf{A}) + \cdots + a_{nj} \operatorname{cof}_{nj}(\mathbf{A}) = \sum_{i=1}^{n} a_{ij} \operatorname{cof}_{ij}(\mathbf{A}) = \det(\mathbf{A})$$

Así pues, la diagonal de la matriz resultante está compuesta por el valor del determinante de $\mathbf{A}$.

Los elementos fuera de la diagonal son determinantes de matrices con dos columnas iguales. Por ejemplo, el segundo elemento de la primera fila de $[\mathbf{Adj(A)}] \cdot \mathbf{A}$ es el desarrollo por la primera columna de

$$\begin{vmatrix} a_{12} & a_{12} & \dots & a_{1n} \\ a_{22} & a_{22} & \dots & a_{2n} \\ a_{32} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & & \vdots \\ a_{n2} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{12} \operatorname{cof}_{11}(\mathbf{A}) + a_{22} \operatorname{cof}_{21}(\mathbf{A}) + a_{32} \operatorname{cof}_{31}(\mathbf{A}) + \cdots + a_{n2} \operatorname{cof}_{n1}(\mathbf{A}) = \sum_{i=1}^{n} a_{i2} \operatorname{cof}_{i1}(\mathbf{A}) = 0,$$

donde la columna 2 aparece repetida en la primera posición, y por lo tanto el determinante es igual a cero.

Y el elemento $k$ésimo ($k \neq 1$) de la primera fila es

$$\begin{vmatrix} a_{1k} & a_{12} & \dots & a_{1n} \\ a_{2k} & a_{22} & \dots & a_{2n} \\ a_{3k} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & & \vdots \\ a_{nk} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{1k} \operatorname{cof}_{11}(\mathbf{A}) + a_{2k} \operatorname{cof}_{21}(\mathbf{A}) + a_{3k} \operatorname{cof}_{31}(\mathbf{A}) + \cdots + a_{nk} \operatorname{cof}_{n1}(\mathbf{A}) = \sum_{i=1}^{n} a_{ik} \operatorname{cof}_{i1}(\mathbf{A}) = 0,$$

donde la columna $k$ésima aparece repetida en primera posición, y por lo tanto el determinante es igual a cero.

Se deduce entonces que $[\mathbf{Adj(A)}] \cdot \mathbf{A} = |\mathbf{A}| \cdot \mathbf{I}$; y despejando $\mathbf{I}$ tenemos que:

$$\left[\frac{\mathbf{Adj(A)}}{|\mathbf{A}|}\right] \cdot \mathbf{A} = \mathbf{I};$$

donde la primera matriz es necesariamente la matriz inversa de $\mathbf{A}$.

Parte VI

Autovalores y autovectores. Diagonalización y formas cuadráticas

En este Capítulo VI todas las matrices son cuadradas.

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 λ tal que (16.1) tiene soluciones no nulas $\boldsymbol{x} \neq \mathbf{0}$. En tal caso, a los vectores $\boldsymbol{x}$ se les llama autovectores de $\mathbf{A}$ correspondientes al autovalor λ. Al conjunto de autovalores de $\mathbf{A}$ se le denomina espectro de $\mathbf{A}$.*

Por ejemplo, es fácil ver 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 de $\mathbf{A}$, $\lambda_1 = 6$ y $\lambda_2 = 1$ respectivamente (basta multiplicar $\mathbf{A}\boldsymbol{x}_i$ para comprobarlo).

Ejercicio 60. Demuestre la siguiente proposición:

Proposición 16.0.1. *Una combinación lineal de autovectores correspondientes al autovalor λ es otro autovector correspondiente al mismo autovalor λ.*

El anterior resultado justifica la siguiente definición:

Definición 16.2 (Autoespacio)**.** *Sea $\mathbf{A}$ de orden n. Se denomina autoespacio correspondiente al autovalor λ de $\mathbf{A}$ al subespacio formado por los autovectores correspondientes al autovalor λ de $\mathbf{A}$ junto con el vector nulo.*

$$\mathcal{E}_\lambda(\mathbf{A}) = \left\{ \boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x} \right\}.$$

16.0.1. Cálculo de los autovalores y los autovectores

¿Cómo podemos encontrar los autovalores de una matriz? ¿Y qué podemos decir acerca de la existencia de autovalores de una matriz en general? Para responder, escribamos la Ecuación (16.1) de manera diferente

$$\begin{aligned} \mathbf{A}\boldsymbol{x} &= \lambda\boldsymbol{x} \\ \mathbf{A}\boldsymbol{x} &= \lambda\mathbf{I}\boldsymbol{x} \\ \mathbf{A}\boldsymbol{x} - \lambda\mathbf{I}\boldsymbol{x} &= \mathbf{0} \\ \big(\mathbf{A} - \lambda\mathbf{I}\big)\boldsymbol{x} &= \mathbf{0}, \end{aligned} \tag{16.2}$$

es decir, el autoespacio correspondiente a λ 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 $\big(\mathbf{A}-\lambda\mathbf{I}\big)$ es singular, y por tanto:

$$\det\big(\mathbf{A}-\lambda\mathbf{I}\big)=0. \tag{16.3}$$

Así pues, el problema reside en encontrar los valores de λ para los que $(\mathbf{A}-\lambda\mathbf{I})$ es singular (y por tanto $|\mathbf{A}-\lambda\mathbf{I}|=0$). Como veremos a continuación, al desarrollar el determinante se comprueba que existen $n+1$ coeficientes $p_0,\dots,p_n$ de $\mathbb{R}$, tales que para todo λ

$$\det\big(\mathbf{A}-\lambda\mathbf{I}\big)= \ p_0+p_1\lambda+\cdots+p_{n-1}\lambda^{n-1}+p_n\lambda^n.$$

$|\mathbf{A}-\lambda\mathbf{I}|$ es un polinomio en λ que denominamos *polinomio característico* de $\mathbf{A}$ y que denotamos con $P_{\mathbf{A}}(\lambda)$.

En el ejemplo de más arriba:

$$P_{\mathbf{A}}(\lambda)=\det\big(\mathbf{A}-\lambda\mathbf{I}\big)=\begin{vmatrix}5-\lambda & 4\\ 1 & 2-\lambda\end{vmatrix}=6-7\lambda+\lambda^2=0.$$

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: $\det\big(a_{11}-\lambda\big)=a_{11}-\lambda$. Supuesto que es cierto para cualquier matriz de orden $(n-1)$, veamos que también es cierto para cualquier matriz de orden n:

$$\det\big(\mathbf{A}-\lambda\mathbf{I}\big)=\det\left[\begin{array}{ccc|c} & & & a_{1n}\\ & {}^{n\urcorner}\mathbf{A}^{\ulcorner n}-\lambda\mathbf{I} & & \vdots\\ & & & a_{(n-1)n}\\ \hline * & \dots & * & a_{nn}-\lambda\end{array}\right]$$

$$=\det\left[\begin{array}{ccc|c} & & & a_{1n}\\ & {}^{n\urcorner}\mathbf{A}^{\ulcorner n}-\lambda\mathbf{I} & & \vdots\\ & & & a_{(n-1)n}\\ \hline * & \dots & * & a_{nn}\end{array}\right]-\lambda\det\left[\begin{array}{ccc|c} & & & 0\\ & {}^{n\urcorner}\mathbf{A}^{\ulcorner n}-\lambda\mathbf{I} & & \vdots\\ & & & 0\\ \hline * & \dots & * & 1\end{array}\right]$$

$$=\det\left[\begin{array}{ccc|c} & & & a_{1n}\\ & {}^{n\urcorner}\mathbf{A}^{\ulcorner n}-\lambda\mathbf{I} & & \vdots\\ & & & a_{(n-1)n}\\ \hline * & \dots & * & a_{nn}\end{array}\right]-\lambda\det\Big({}^{n\urcorner}\mathbf{A}^{\ulcorner n}-\lambda\mathbf{I}\Big), \tag{16.4}$$

donde ${}^{n\urcorner}\mathbf{A}^{\ulcorner n}$ es la submatriz de $\mathbf{A}$ que resulta tras quitar la fila y columna $n$ésimas. Por hipótesis de inducción, $\det\big({}^{n\urcorner}\mathbf{A}^{\ulcorner n}-\lambda\mathbf{I}\big)$ 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

$$\det\left[\begin{array}{ccc|c} & & & a_{1n}\\ & {}^{n\urcorner}\mathbf{A}^{\ulcorner n}-\lambda\mathbf{I} & & \vdots\\ & & & a_{(n-1)n}\\ \hline * & \dots & * & a_{nn}\end{array}\right]=\det\left[\begin{array}{ccc|c} & & & a_{1n}\\ & \mathbf{B}-\lambda\mathbf{I} & & \vdots\\ & & & a_{(n-1)n}\\ \hline 0 & \dots & 0 & a_{nn}\end{array}\right]=a_{nn}\cdot\det\big(\mathbf{B}-\lambda\mathbf{I}\big),$$

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

$$\det\left[\begin{array}{ccc|c} & {}_{n|}\mathbf{A}_{|n}-\lambda\mathbf{I} & & \begin{matrix}a_{1n}\\ \vdots\\ a_{(n-1)n}\end{matrix}\\ \hline * & \dots & * & 0\end{array}\right] = \det\left[\begin{array}{ccc|c} & {}_{n|}\mathbf{A}_{|n}-\lambda\mathbf{I} & & \begin{matrix}a_{1n}\\ \vdots\\ a_{(n-1)n}\end{matrix}\\ \hline * & \dots & * & -1\end{array}\right] + \det\left[\begin{array}{ccc|c} & {}_{n|}\mathbf{A}_{|n}-\lambda\mathbf{I} & & \begin{matrix}0\\ \vdots\\ 0\end{matrix}\\ \hline * & \dots & * & 1\end{array}\right].$$

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 16.0.2. *Todo autovalor de* $\mathbf{A}$ *es raíz del polinomio característico* $P_{\mathbf{A}}(\lambda)$.

Para el ejemplo de más arriba, puesto que $\det\big(\mathbf{A}-\lambda\mathbf{I}\big)=\lambda^2-7\lambda+6=0$; tenemos que

$$\lambda=\frac{7\pm\sqrt{49-24}}{2}=\frac{7\pm5}{2}=\begin{cases}6\\1\end{cases}\qquad\Longrightarrow\quad \text{El } \textit{espectro} \text{ de } \mathbf{A} \text{ es } \{6,1\}.$$

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

$$P_{\mathbf{A}}(\lambda)\;=\;p_0+p_1\lambda+\cdots+p_{n-1}\lambda^{n-1}+p_n\lambda^n\;=\;\alpha(\lambda_1-\lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda),\quad \text{donde } \alpha,\lambda_1,\dots,\lambda_n\in\mathbb{C}.$$

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.*

Como consecuencia tenemos el siguiente resultado:

Teorema 16.0.3. *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 mucho* 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* λ *es una raíz de* $P_{\mathbf{A}}$ *de multiplicidad* k *diremos que* λ *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 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, respectivamente, las soluciones no nulas de los 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 λ*; 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 $P_{\mathbf{A}}(\lambda) = \det\big(\mathbf{A} - \lambda\mathbf{I}\big)$.
2. Encontrar las raíces λ_i de la ecuación característica $P_{\mathbf{A}}(\lambda) = 0$.
3. Resolver los sistemas homogéneos $\big(\mathbf{A} - \lambda_i\mathbf{I}\big)\boldsymbol{x} = \mathbf{0}$ para encontrar los autovectores.

La multiplicidad *algebraica* de λ es el número de veces que se repite la raíz λ en el polinomio característico $P_{\mathbf{A}}(\lambda)$.

La multiplicidad *geométrica* de λ es la dimensión del correspondiente autoespacio $\mathcal{E}_\lambda(\mathbf{A})$.

Matrices semejantes y diagonalización por bloques triangulares

En esta lección vamos a ver que toda matriz cuadrada se puede transformar en una matriz diagonal **por bloques triangulares**, y que cada submatriz en la diagonal (cada bloque) es triangular y con un autovalor en su diagonal. Además, veremos cuál es la condición necesaria para que los bloques sean de tamaño 1, es decir, para que la matriz sea *diagonalizable*.

Primero introduzcamos nueva notación relacionada con las matrices elementales y las operaciones con filas y columnas.

17.1. Transformación elemental "*espejo*" de otra transformación

Resulta que podemos obtener una misma matriz elemental tanto operando sobre las filas como sobre las columnas de la matriz identidad. En el caso de las matrices elementales *Tipo I* es muy sencillo. Por ejemplo, para $\mathbf{I}_{\underset{[(\alpha)\mathbf{2}+\mathbf{3}]}{\boldsymbol{\tau}}}$

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

En general, para una misma matriz elemental $\mathbf{I}_{\underset{[(\alpha)\boldsymbol{j}+\boldsymbol{k}]}{\boldsymbol{\tau}}}$, la transformación elemental necesaria para crearla es distinta si actuamos sobre las filas, o sobre sus columnas, pues es necesario el *intercambio entre los índices* $\boldsymbol{j}$ y $\boldsymbol{k}$.

$$_{\underset{[(\alpha)\boldsymbol{k}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\mathbf{I} = \mathbf{I}_{\underset{[(\alpha)\boldsymbol{j}+\boldsymbol{k}]}{\boldsymbol{\tau}}}$$

A falta de un mejor nombre, llamaremos a cada una el "*espejo*" de la otra:

Definición 17.1. *Llamaremos* "espejo" *de una transformación elemental* $\boldsymbol{\tau}$ *(que denotaremos con* $esp(\boldsymbol{\tau})$*) a aquella transformación elemental que actuando por el otro lado de la matriz identidad arroja el mismo resultado; es decir*

$$_{esp(\boldsymbol{\tau})}\mathbf{I} = \mathbf{I}_{\boldsymbol{\tau}} \qquad ó \qquad _{\boldsymbol{\tau}}\mathbf{I} = \mathbf{I}_{esp(\boldsymbol{\tau})}$$

El *espejo* de la transformación elemental $\underset{[(\alpha)\boldsymbol{j}+\boldsymbol{k}]}{\boldsymbol{\tau}}$ es a la transformación que resulta de intercambiar los índices j y k, es decir,

$$esp\Big(\underset{[(\alpha)\boldsymbol{j}+\boldsymbol{k}]}{\boldsymbol{\tau}}\Big) = \underset{[(\alpha)\boldsymbol{k}+\boldsymbol{j}]}{\boldsymbol{\tau}}.$$

Para las matrices elementales *Tipo II* (y para las matrices intercambio) el caso es mucho más sencillo. Como las matrices elementales *Tipo II* (y las matrices intercambio) son simétricas, $_{\boldsymbol{\tau}}\mathbf{I} = \mathbf{I}_{\boldsymbol{\tau}}$, resulta que tanto las transformaciones elementales *Tipo II* (y los intercambios), son su propio *espejo*. Por ejemplo

$$\underset{[(\alpha)\mathbf{2}]}{\boldsymbol{\tau}}\begin{bmatrix}1&&\\&1&\\&&1\end{bmatrix} = \begin{bmatrix}1&&\\&\alpha&\\&&1\end{bmatrix} = \begin{bmatrix}1&&\\&1&\\&&1\end{bmatrix}\underset{[(\alpha)\mathbf{2}]}{\boldsymbol{\tau}}.$$

Por otra parte, dada una sucesión de transformaciones elementales, su *espejo* es la sucesión de transformaciones espejo:

$$esp\big(\boldsymbol{\tau}_1\dots\boldsymbol{\tau}_k\big) \;=\; esp(\boldsymbol{\tau}_1)\dots esp(\boldsymbol{\tau}_k);$$

es decir

$${}_{esp(\boldsymbol{\tau}_1\dots\boldsymbol{\tau}_k)}\mathbf{I} \;=\; \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$$

Librería NAcAL para Python

```
Tr = T((3,2,1)) & T((-2,3,2)) & T((7,1,3)) & T({2,3}) & T((20,1)) # Secuencia de 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
```

Así, como la inversa de $\underset{[(\alpha)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}$ es $\underset{[(-\alpha)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}$; la matriz $\Big(\underset{[(-\alpha)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}\mathbf{I}\Big)$ es la inversa de $\Big(\mathbf{I}\underset{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}\Big)$; y por otra parte, la matriz $\Big(\underset{[(\frac{1}{\alpha})\boldsymbol{i}]}{\boldsymbol{\tau}}\mathbf{I}\Big)$ es la inversa de $\Big(\mathbf{I}\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}\Big)$.

Por tanto,

$$\underset{[(-\alpha)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}\mathbf{I}\underset{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}} = \mathbf{I}; \qquad \text{y} \qquad \underset{[(\frac{1}{\alpha})\boldsymbol{i}]}{\boldsymbol{\tau}}\mathbf{I}\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}} = \mathbf{I}.$$

Y en general

$${}_{esp(\boldsymbol{\tau}^{-1})}\mathbf{I}_{\boldsymbol{\tau}} = \mathbf{I} \qquad \text{y} \qquad {}_{esp(\boldsymbol{\tau}_1^{-1}\dots\boldsymbol{\tau}_k^{-1})}\mathbf{I}_{(\boldsymbol{\tau}_1\dots\boldsymbol{\tau}_k)} = \mathbf{I}.$$

Librería NAcAL para Python

```
Tr = T((3,2,1)) & T((-2,3,2)) & T((7,1,3)) & T({2,3}) & T((20,1))  # Sec. de transf.
(Tr**-1).espejo() & I(3) & Tr
```

En lo que queda de lección aplicaremos a matrices cuadradas transformaciones de la forma

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} \;=\; \big(\mathbf{I}_{(\boldsymbol{\tau}_1\dots\boldsymbol{\tau}_k)}\big)^{-1}\mathbf{A}\big(\mathbf{I}_{(\boldsymbol{\tau}_1\dots\boldsymbol{\tau}_k)}\big) \;=\; \big({}_{esp(\boldsymbol{\tau}_1^{-1}\dots\boldsymbol{\tau}_k^{-1})}\mathbf{I}\big)\mathbf{A}\big(\mathbf{I}_{(\boldsymbol{\tau}_1\dots\boldsymbol{\tau}_k)}\big) \;=\; {}_{esp(\boldsymbol{\tau}_1^{-1}\dots\boldsymbol{\tau}_k^{-1})}\mathbf{A}_{(\boldsymbol{\tau}_1\dots\boldsymbol{\tau}_k)},$$

donde $\mathbf{S} = \mathbf{I}_{(\boldsymbol{\tau}_1\dots\boldsymbol{\tau}_k)}$.

17.2. Diagonalización por bloques triangulares

17.2.1. Matrices semejantes

Definición 17.2. *Decimos que* $\mathbf{A}$ *y* $\mathbf{C}$ *(del mismo orden) son* semejantes *si existe una matriz invertible* $\mathbf{S}$ *tal que*

$$\mathbf{A} = \mathbf{S}^{-1}\mathbf{C}\mathbf{S}.$$

Propiedades compartidas por dos matrices semejantes

Las *matrices semejantes* comparten muchas propiedades; por ejemplo, *tienen el mismo determinante*:

$$\det \mathbf{A} \;=\; \det \mathbf{S}^{-1} \cdot \det \mathbf{C} \cdot \det \mathbf{S} \;=\; \det \mathbf{C} \cdot \det \mathbf{S} \cdot \det \mathbf{S}^{-1} \;=\; \det \mathbf{C}.$$

Las matrices semejantes también tienen idéntico polinomio característico:

Proposición 17.2.1. *Si* $\mathbf{A}$ *y* $\mathbf{C}$ *son semejantes, entonces tienen el mismo polinomio característico.*

Demostración. Puesto que son similares, existe una matriz invertible $\mathbf{S}$ tal que $\mathbf{C} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$, entonces

$$\begin{aligned}
P_{\mathbf{C}}(\lambda) = \det\left(\mathbf{C} - \lambda\mathbf{I}\right) &= \det\left(\mathbf{S}^{-1}\mathbf{A}\mathbf{S} - \lambda\mathbf{S}^{-1}\mathbf{S}\right) && \text{pues } \mathbf{C} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S} \text{ y } \mathbf{S}^{-1}\mathbf{S} = \mathbf{I}\\
&= \det\left(\mathbf{S}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{S}\right) && \text{sacando factor común}\\
&= \det\left(\mathbf{S}^{-1}\right) \cdot \det\left(\mathbf{A} - \lambda\mathbf{I}\right) \cdot \det\left(\mathbf{S}\right) && \text{determinante de un producto de matrices}\\
&= \det\left(\mathbf{A} - \lambda\mathbf{I}\right) = P_{\mathbf{A}}(\lambda) && \text{ya que } \det\left(\mathbf{S}^{-1}\right) = \frac{1}{\det\left(\mathbf{S}\right)}.
\end{aligned}$$

□

Y puesto que cada raíz del polinomio característico es un autovalor de la matrices, y 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.*

Pero 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 n columnas y el mismo número de pivotes, por tanto $\dim\mathcal{N}\left(\mathbf{A}\mathbf{S}\right) = \dim\mathcal{N}\left(\mathbf{A}\right)$. Por otra parte sabemos que si $\mathbf{S}$ es invertible, $\mathcal{N}\left(\mathbf{S}\mathbf{A}\right) = \mathcal{N}\left(\mathbf{A}\right)$. Combinando ambos resultados y teniendo en cuenta que

$$\mathbf{S}^{-1}\mathbf{A}\mathbf{S} - \lambda_i\mathbf{I} \quad = \quad \mathbf{S}^{-1}\mathbf{A}\mathbf{S} - \lambda_i\mathbf{S}^{-1}\mathbf{S} \quad = \quad \mathbf{S}^{-1}\Big(\mathbf{A} - \lambda_i\mathbf{I}\Big)\mathbf{S},$$

tenemos que para cada autovalor λ_i de $\mathbf{A}$,

$$\dim\mathcal{N}\left(\mathbf{S}^{-1}\mathbf{A}\mathbf{S} - \lambda_i\mathbf{I}\right) \quad = \quad \dim\mathcal{N}\left(\mathbf{A} - \lambda_i\mathbf{I}\right).$$

Definición 17.3. *La* traza *de* $\underset{n\times n}{\mathbf{A}}$ *es la suma de los elementos de su diagonal principal, es decir,*

$$\operatorname{tr}\left(\mathbf{A}\right) = a_{11} + a_{22} + \cdots + a_{nn}$$

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, $\left(\mathbf{I}_{\boldsymbol{\tau}}\right)^{-1}$, por la izquierda. Evidentemente la matriz resultante es similar a $\mathbf{A}$. Veamos que la traza no cambia…

Proposición 17.2.2. *Si* $\mathbf{A} = \left(\mathbf{I}_{\boldsymbol{\tau}}\right)^{-1}\mathbf{B}\left(\mathbf{I}_{\boldsymbol{\tau}}\right)$, *es decir, si* $\mathbf{A} = {}_{esp(\boldsymbol{\tau}^{-1})}\mathbf{B}_{\boldsymbol{\tau}}$, *entonces* $\operatorname{tr}\left(\mathbf{A}\right) = \operatorname{tr}\left(\mathbf{B}\right).$

Demostración. Veamos que la traza no cambia ni con transformaciones *Tipo I* ni con transformaciones *Tipo II*:

Si $\boldsymbol{\tau}$ es de Tipo II y $\boldsymbol{\tau}$ multiplica por α a la columna $\boldsymbol{i}$ésima, entonces $esp(\boldsymbol{\tau}^{-1})$ dividirá por α la fila $\boldsymbol{i}$ésima. Por tanto, el $\boldsymbol{i}$ésimo componente de la diagonal no cambiará. Veámoslo:

$$\mathbf{A}\Big(\mathbf{I}_{\underset{[(\alpha)\boldsymbol{j}]}{\boldsymbol{\tau}}}\Big)_{|j} = \Big(\mathbf{A}_{\underset{[(\alpha)\boldsymbol{j}]}{\boldsymbol{\tau}}}\Big)_{|j} = \alpha\big(\mathbf{A}_{|j}\big);$$

pero al aplicar la inversa también tenemos

$$_{j|}\Big({}_{\underset{[(\frac{1}{\alpha})\boldsymbol{j}]}{\boldsymbol{\tau}}}\mathbf{I}\Big)\mathbf{A} = {}_{j|}\Big({}_{\underset{[(\frac{1}{\alpha})\boldsymbol{j}]}{\boldsymbol{\tau}}}\mathbf{A}\Big) = \frac{1}{\alpha}\big({}_{j|}\mathbf{A}\big);$$

Como la primera operación multiplica la componente ${}_{j|}\mathbf{A}_{|j}$ por α y la segunda la divide por α, la diagonal no cambia; y por tanto tampoco cambia la traza.

Si $\boldsymbol{\tau}$ es de Tipo I y $\boldsymbol{\tau}$ suma α veces la columna $\boldsymbol{i}$ésima a la $\boldsymbol{j}$ésima, y a la correspondiente matriz elemental la denominamos $\mathbf{E} = \mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}$ entonces su inversa es $\mathbf{E}^{-1} = {}_{\underset{[(-\alpha)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}}\mathbf{I}$ y

$$\big(\mathbf{AE}\big)_{|j} = \alpha\big(\mathbf{A}_{|i}\big) + \mathbf{A}_{|j} \quad \Rightarrow \quad {}_{j|}\big(\mathbf{AE}\big)_{|j} = \alpha\big({}_{j|}\mathbf{A}_{|i}\big) + {}_{j|}\mathbf{A}_{|j}$$

pero

$${}_{i|}\big(\mathbf{E}^{-1}\mathbf{A}\big) = -\alpha\big({}_{j|}\mathbf{A}\big) + {}_{i|}\mathbf{A} \quad \Rightarrow \quad {}_{i|}\big(\mathbf{E}^{-1}\mathbf{A}\big)_{|i} = -\alpha\big({}_{j|}\mathbf{A}_{|i}\big) + {}_{i|}\mathbf{A}_{|i}$$

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:

Corolario 17.2.3. *Si* $\mathbf{A}$ *y* $\mathbf{B}$ *son semejantes, entonces tienen la misma traza.*

Librería 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 $\mathbf{A}$ 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 su producto es el determinante.

Nota 3. *¡Recuerde que para obtener matrices semejantes hay que operar tanto con las filas como con las columnas! (...hay que multiplicar por la derecha con una matriz y por la izquierda con la inversa de dicha matriz).*

17.2.2. Diagonalización por bloques triangulares

Teorema 17.2.4. *Para toda matriz cuadrada* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *existe una matriz invertible* $\mathbf{S}$ *tal que*

$$\mathbf{S}^{-1}\mathbf{AS} = \begin{bmatrix} \boxed{\mathbf{T}_{\lambda_k}} & & & \\ & \ddots & & \\ & & \boxed{\mathbf{T}_{\lambda_2}} & \\ & & & \boxed{\mathbf{T}_{\lambda_1}} \end{bmatrix} \quad \textit{donde } \mathbf{T}_{\lambda_i} = \begin{bmatrix} \lambda_i & & & \\ * & \lambda_i & & \\ * & * & \ddots & \\ * & * & \cdots & \lambda_i \end{bmatrix} \quad \textit{es triangular y de orden igual a la multiplicidad}$$

algebraica de λ_i*, y donde* $\{\lambda_1, \dots, \lambda_k\}$ *es el conjunto*[1] *de autovalores de* $\mathbf{A}$*.*

☞ Antes de demostrar este importante teorema, vamos a ver algunos resultados previos que nos ayudarán a entender los pasos del algoritmo de diagonalización por bloques.

Lema 17.2.5 (De paso inicial)**.** *Si* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *es de la forma*

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{C} & \\ \hline \ast & \mathbf{L} \end{array}\right]$$

donde $\mathbf{C}$ *(de orden* m*) es singular y* $\mathbf{L}$ *es una triangular inferior sin ceros en la diagonal principal, entonces existe una matriz invertible* $\mathbf{S}$ *tal que*

$$\mathbf{S}^{-1}\mathbf{AS} = \left[\begin{array}{c|c|c} \mathbf{C'} & & \\ \hline & 0 & \\ \cline{2-3} \ast & & \mathbf{L} \end{array}\right], \quad \textit{donde } \mathbf{C'} \textit{ es de orden } (m-1).$$

Y si $\mathbf{A}$ *es simplemente singular, entonces* $\mathbf{S}^{-1}\mathbf{AS} = \left[\begin{array}{c|c} \mathbf{C'} & \\ \hline \ast & 0 \end{array}\right]$, *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 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, ${}_{esp(\boldsymbol{\tau}_1^{-1}\dots\boldsymbol{\tau}_k^{-1})}\mathbf{I} = \mathbf{R}^{-1}$, 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

$${}_{esp(\boldsymbol{\tau}_1^{-1}\dots\boldsymbol{\tau}_k^{-1})}\mathbf{A}_{(\boldsymbol{\tau}_1\dots\boldsymbol{\tau}_k)} = \mathbf{R}^{-1}\mathbf{AR} = \left[\begin{array}{c|c|ccccc} & 0 & & & & & \\ \underset{m\times(m-1)}{\ast} & \vdots & & & & & \\ & 0 & & & & & \\ \hline & d_{m+1} & \beta_{m+1} & & & & \\ & d_{m+2} & \ast & \beta_{m+2} & & & \\ \ast & \vdots & \ast & \ast & \ddots & & \\ & d_n & \ast & \ast & \cdots & \beta_n & \end{array}\right]$$

Etapa 2 [*Anulando el resto de coeficientes de la columna* m*ésima de* $\mathbf{A}$ *(los que quedan a la izquierda de* $\mathbf{L}$*)*]. Gracias a que las componentes β_j de la diagonal principal de $\mathbf{L}$ son pivotes, mediante una sucesión de transformaciones elementales $\mathbf{I}_{\boldsymbol{\tau}_{(k+1)}\cdots\boldsymbol{\tau}_p}$, $\left(\text{del tipo } \mathbf{I}_{\underset{[(\alpha_j)\boldsymbol{j}+\boldsymbol{m}]}{\boldsymbol{\tau}}} = \mathbf{P}, \text{ con } j > m\right)$, se pueden anular las componentes $d_{m+1}, \dots, d_n$ de la columna $m$ésima. Al aplicar la sucesión de las correspondientes transformaciones inversas "*espejo*", ${}_{esp(\boldsymbol{\tau}_{k+1}^{-1}\dots\boldsymbol{\tau}_p^{-1})}\mathbf{I} = \mathbf{P}^{-1}$, $\left(\text{del tipo } {}_{\underset{[(-\alpha_j)\boldsymbol{m}+\boldsymbol{j}]}{\boldsymbol{\tau}}}\mathbf{I}, \text{ con } j > m\right)$, solo varían las columnas correspondientes a los asteriscos "$\ast$", ya que la fila $m$ésima contiene únicamente ceros a partir de la posición

[1] Asumimos que en $\{\lambda_1, \dots, \lambda_k\}$ no hay elementos repetidos.

m; resultando la matriz

$$\,_{esp(\boldsymbol{\tau}_1^{-1}\dots\boldsymbol{\tau}_k^{-1},\boldsymbol{\tau}_{(k+1)}^{-1}\dots\boldsymbol{\tau}_p^{-1})}\mathbf{A}_{(\boldsymbol{\tau}_1\dots\boldsymbol{\tau}_k,\boldsymbol{\tau}_{k+1}\dots\boldsymbol{\tau}_p)} = \big(\mathbf{P}^{-1}\mathbf{R}^{-1}\big)\mathbf{A}\big(\mathbf{RP}\big) = \left[\begin{array}{c|c|c} \mathbf{C'} & & \\ \hline & 0 & \\ \hline \boldsymbol{*} & & \mathbf{L} \end{array}\right], \quad \text{donde } \mathbf{P} = \mathbf{I}_{\boldsymbol{\tau}_{(k+1)}\cdots\boldsymbol{\tau}_p}.$$

Llamando $\mathbf{S}$ a la matriz $\mathbf{RP} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p}$ 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 17.2.6 (Paso de continuación)**.** *Si* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *es de la forma*

$$\mathbf{A} = \left[\begin{array}{c|c|c} \mathbf{C} & & \\ \hline \boldsymbol{*} & \mathbf{T} & \\ \hline \boldsymbol{*} & & \mathbf{L} \end{array}\right].$$

donde $\mathbf{C}$ *(de orden* k*) es singular, donde* $\mathbf{L}$ *(de orden* $n-m$*) es una 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*

$$\mathbf{S}^{-1}\mathbf{AS} = \left[\begin{array}{c|c|c} \mathbf{C'} & & \\ \hline \boldsymbol{*} & \mathbf{T'} & \\ \hline \boldsymbol{*} & & \mathbf{L} \end{array}\right],$$

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* $\left[\begin{array}{c|c} \mathbf{C} & \\ \hline \boldsymbol{*} & \mathbf{T} \end{array}\right]$, *entonces* $\mathbf{S}^{-1}\mathbf{AS} = \left[\begin{array}{c|c} \mathbf{C'} & \\ \hline \boldsymbol{*} & \mathbf{T'} \end{array}\right]$, *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

$$\left[\begin{array}{c|c|c|cccc} \underset{k\times(k-1)}{\boldsymbol{*}} & \begin{matrix}0\\ \vdots\\ 0\end{matrix} & & & & & \\ \hline \boldsymbol{*} & * & \mathbf{T} & & & & \\ \hline \boldsymbol{*} & \begin{matrix} d_{m+1}\\ d_{m+2}\\ \vdots\\ d_n\end{matrix} & & \begin{matrix}\beta_{m+1} & & & \\ * & \beta_{m+2} & & \\ * & * & \ddots & \\ * & * & \cdots & \beta_n\end{matrix} \end{array}\right], \quad \text{donde } \mathbf{T'} = \left[\begin{array}{c|c} 0 & \\ \hline * & \mathbf{T} \end{array}\right].$$

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

$$\left[\begin{array}{c|c|c|c} \mathbf{C'} & & & \\ \hline & 0 & & \\ \boldsymbol{*} & * & \mathbf{T} & \\ & \mathbf{0} & & \mathbf{L} \end{array}\right], \quad \text{donde } \mathbf{T'} = \left[\begin{array}{c|c} 0 & \\ \hline * & \mathbf{T} \end{array}\right].$$

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 el *Etapa 1.* $\square$

Antes de pasar al siguiente corolario, recuérdese que para una *matriz triangular por bloques* (véase el EJERCICIO 57 en la página 152)

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{B} & \mathbf{0} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right] \qquad \text{se verifica que} \qquad |\mathbf{A}| = |\mathbf{B}| \cdot |\mathbf{D}|\,;$$

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:

$$\det\Big(\mathbf{A} - \lambda\mathbf{I}\Big) = \det\Big(\mathbf{B} - \lambda\mathbf{I}\Big) \cdot \det\Big(\mathbf{D} - \lambda\mathbf{I}\Big).$$

☞ Usaremos este resultado sobre polinomios característicos en la última parte de la demostración del siguiente corolario, que nos indica como iniciar el algoritmo de diagonalización para generar un primer bloque:

Corolario 17.2.7. *Si* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *y* λ_1 *es un autovalor de* $\mathbf{A}$*, entonces existe* $\mathbf{S}$ *invertible tal que*

$$\textit{o bien} \qquad \mathbf{S}^{-1}\mathbf{AS} = \left[\begin{array}{c|c} \mathbf{C'} & \\ \hline \ast & \mathbf{T}_{\lambda_1} \end{array}\right]; \qquad \textit{o bien} \qquad \mathbf{S}^{-1}\mathbf{AS} = \mathbf{T}_{\lambda_1};$$

donde $\quad \mathbf{T}_{\lambda_1} = \begin{bmatrix} \lambda_1 & & & \\ \ast & \lambda_1 & & \\ \ast & \ast & \ddots & \\ \ast & \ast & \cdots & \lambda_1 \end{bmatrix}$ *es de orden igual a la multiplicidad* algebraica *de* λ_1.

Demostración. Como $\big(\mathbf{A} - \lambda_1\mathbf{I}\big)$ 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

$$\big(\mathbf{S}_k^{-1}\cdots\mathbf{S}_1^{-1}\big)\cdot\big(\mathbf{A}-\lambda_1\mathbf{I}\big)\cdot\big(\mathbf{S}_1\cdots\mathbf{S}_k\big) = \begin{cases} \big[\mathbf{T}\big] & \text{si hemos podido continuar hasta el final} \\[2ex] \left[\begin{array}{c|c} \mathbf{C'} & \\ \hline \ast & \mathbf{T} \end{array}\right] & \text{si hemos topado un una submatriz } \mathbf{C'} \text{ no singular,} \end{cases} \tag{17.1}$$

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 $\mathbf{S}^{-1}\big(\mathbf{A}-\lambda_1\mathbf{I}\big)\mathbf{S} = \mathbf{S}^{-1}\mathbf{AS}-\lambda_1\mathbf{I}$, sumando $\big(\lambda_1\mathbf{I}\big)$ en (17.1) tenemos

$$\big(\mathbf{S}_k^{-1}\cdots\mathbf{S}_1^{-1}\big)\cdot\mathbf{A}\cdot\big(\mathbf{S}_1\cdots\mathbf{S}_k\big) = \begin{cases} \big[\mathbf{T}+\lambda_1\mathbf{I}\big] \\[2ex] \left[\begin{array}{c|c} \mathbf{C'}+\lambda_1\mathbf{I} & \\ \hline \ast & \mathbf{T}+\lambda_1\mathbf{I} \end{array}\right] \end{cases} ; \quad \text{donde } \mathbf{T}_{\lambda_1} = \Big[\mathbf{T}+\lambda_1\mathbf{I}\Big].$$

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, $\left(P_{\left(\mathbf{C'}+\lambda_1\mathbf{I}\right)}(\lambda)\right) \cdot (\lambda_1 - \lambda)^k$. Ahora bien, como λ_1 no es un autovalor[2] de $\left(\mathbf{C'} + \lambda_1\mathbf{I}\right)$, 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 (nótese que es casi idéntico al anterior) …

Corolario 17.2.8. *Si* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *de la forma* $\left[\begin{array}{c|ccc} \mathbf{C} & & & \\ \hline & \boxed{\mathbf{T}_{\lambda_r}} & & \\ * & & \ddots & \\ & & & \boxed{\mathbf{T}_{\lambda_1}} \end{array}\right]$, *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* λ_i *y donde* λ_{r+1} *es un autovalor*[3] *de* $\mathbf{C}$*, entonces existe* $\mathbf{S}$ *invertible tal que*

$$\text{o bien} \quad \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \left[\begin{array}{c|ccc} \mathbf{C'} & & & \\ \hline & \boxed{\mathbf{T}_{\lambda_{r+1}}} & & \\ * & & \ddots & \\ & & & \boxed{\mathbf{T}_{\lambda_1}} \end{array}\right], \qquad \text{o bien} \quad \mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \left[\begin{array}{c|ccc} \mathbf{T}_{\lambda_{r+1}} & & & \\ \hline & \boxed{\mathbf{T}_{\lambda_r}} & & \\ * & & \ddots & \\ & & & \boxed{\mathbf{T}_{\lambda_1}} \end{array}\right],$$

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 *del autovalor* λ_{r+1}.

Demostración. Como $\left(\mathbf{C} - \lambda_{r+1}\mathbf{I}\right)$ es singular, repitiendo los mismos pasos de la anterior demostración llegamos a

$$\left(\mathbf{S}_k^{-1}\cdots\mathbf{S}_1^{-1}\right)\cdot\mathbf{A}\cdot\left(\mathbf{S}_1\cdots\mathbf{S}_k\right) = \begin{cases} \left[\begin{array}{c|ccc} \mathbf{T}_{\lambda_{r+1}} & & & \\ \hline & \boxed{\mathbf{T}_{\lambda_r}} & & \\ * & & \ddots & \\ & & & \boxed{\mathbf{T}_{\lambda_1}} \end{array}\right] & \text{si hemos podido continuar hasta el final} \\[2em] \left[\begin{array}{c|ccc} \mathbf{C'}+\lambda_{r+1}\mathbf{I} & & & \\ \hline & \boxed{\mathbf{T}_{\lambda_{r+1}}} & & \\ * & & \ddots & \\ & & & \boxed{\mathbf{T}_{\lambda_1}} \end{array}\right] & \text{si hemos topado un una submatriz } \mathbf{C'} \text{ no singular,} \end{cases}$$

[2] si fuera autovalor, entonces $\mathbf{C'}$ sería singular.
[3] por tanto también es un autovalor de $\mathbf{A}$ distinto de $\lambda_1, \ldots, \lambda_r$

donde en ambos casos $\mathbf{T}_{\lambda_{r+1}}$ es triangular inferior de orden k, con λ_{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 λ_i.

En el segundo caso, el polinomio característico de $\mathbf{A}$ es

$$P_{\mathbf{A}}(\lambda) \;=\; \Big(P_{\left(\mathbf{C'}+\lambda_{r+1}\mathbf{I}\right)}(\lambda)\Big)\cdot(\lambda_1-\lambda)^{\mu(\lambda_1)}\cdots(\lambda_r-\lambda)^{\mu(\lambda_r)}\cdot(\lambda_{r+1}-\lambda)^k.$$

Ahora bien, como λ_{r+1} no es un autovalor de $\left(\mathbf{C'}+\lambda_{r+1}\mathbf{I}\right)$, entonces $\boxed{k \text{ es la multiplicidad del autovalor } \lambda_{r+1}}$. □

Ya solo resta demostrar el Teorema 17.2.4, pero…

Demostración del Teorema 17.2.4. Ya no hay nada que demostrar. Basta aplicar el primer corolario e iterar el segundo hasta finalizar la generación de bloques. □

Ejemplo 29. Sea la matriz $\mathbf{A}=\begin{bmatrix}1&-1&0\\0&0&0\\0&-2&1\end{bmatrix}$ con autovalores 1, 1 y 0. Vamos a diagonalizar por bloques triangulares.

$$\begin{bmatrix}\mathbf{A}\\\hline\mathbf{I}\end{bmatrix}\xrightarrow[0\mathbf{I}]{(-)}\left[\begin{array}{ccc}1&-1&0\\0&0&0\\0&-2&1\\\hline 1&0&0\\0&1&0\\0&0&1\end{array}\right]\xrightarrow{\substack{\boldsymbol{\tau}\\ [(1)\mathbf{1}+\mathbf{2}]\\ [(2)\mathbf{3}+\mathbf{2}]\\ [\mathbf{2}\rightleftharpoons\mathbf{3}]}}\left[\begin{array}{ccc}1&0&0\\0&0&0\\0&1&0\\\hline 1&0&1\\0&0&1\\0&1&2\end{array}\right]\xrightarrow[\substack{\boldsymbol{\tau}\\ [(-1)\mathbf{2}+\mathbf{1}]\\ [(-2)\mathbf{2}+\mathbf{3}]\\ [\mathbf{2}\rightleftharpoons\mathbf{3}]}]{}\left[\begin{array}{ccc}1&0&0\\0&1&0\\0&0&0\\\hline 1&0&1\\0&0&1\\0&1&2\end{array}\right]\xrightarrow[0\mathbf{I}]{(+)}\left[\begin{array}{ccc}1&0&0\\0&1&0\\0&0&0\\\hline 1&0&1\\0&0&1\\0&1&2\end{array}\right]=\begin{bmatrix}\mathbf{D}\\\hline\mathbf{S}\end{bmatrix}$$

Así pues,

$$\mathbf{D}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}=\begin{bmatrix}1&0&1\\0&0&1\\0&1&2\end{bmatrix}^{-1}\begin{bmatrix}1&-1&0\\0&0&0\\0&-2&1\end{bmatrix}\begin{bmatrix}1&0&1\\0&0&1\\0&1&2\end{bmatrix}=\mathbf{S}^{-1}\mathbf{A}\mathbf{S};$$

Por lo que hemos llegado a una matriz diagonal.

Librería NAcAL para Python

```
A = Matrix([[1,-1,0],[0,0,0],[0,-2,1]]); L=[1,1,0];
Diagonaliza(A, L , 1)
```

Ejemplo 30. 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.

$$\left[\begin{array}{c}\mathbf{A}\\\hline\mathbf{I}\end{array}\right]\xrightarrow[1\mathbf{I}]{(-)}\left[\begin{array}{ccc}-3&0&3\\3&-3&-9\\-1&2&5\\\hline 1&0&0\\0&1&0\\0&0&1\end{array}\right]\xrightarrow{\substack{\boldsymbol{\tau}\\[(1)\mathbf{1}+\mathbf{3}]\\[(-2)\mathbf{2}+\mathbf{3}]}}\left[\begin{array}{ccc}-3&0&0\\3&-3&0\\-1&2&0\\\hline 1&0&1\\0&1&-2\\0&0&1\end{array}\right]\xrightarrow[\substack{\boldsymbol{\tau}\\[(-1)\mathbf{3}+\mathbf{1}]\\[(2)\mathbf{3}+\mathbf{2}]}]{}\left[\begin{array}{ccc}-2&-2&0\\1&1&0\\-1&2&0\\\hline 1&0&1\\0&1&-2\\0&0&1\end{array}\right]\xrightarrow[1\mathbf{I}]{(+)}\left[\begin{array}{ccc}-1&-2&0\\1&2&0\\-1&2&1\\\hline 1&0&1\\0&1&-2\\0&0&1\end{array}\right]$$

$$\xrightarrow[1\mathbf{I}]{(-)}\left[\begin{array}{ccc}-2&-2&0\\1&1&0\\\hline -1&2&0\\\hline 1&0&1\\0&1&-2\\0&0&1\end{array}\right]\xrightarrow{\substack{\boldsymbol{\tau}\\[(-1)\mathbf{1}+\mathbf{2}]}}\left[\begin{array}{ccc}-2&0&0\\1&0&0\\\hline -1&3&0\\\hline 1&-1&1\\0&1&-2\\0&0&1\end{array}\right]\xrightarrow[\substack{\boldsymbol{\tau}\\[(1)\mathbf{2}+\mathbf{1}]}]{}\left[\begin{array}{ccc}-1&0&0\\1&0&0\\\hline -1&3&0\\\hline 1&-1&1\\0&1&-2\\0&0&1\end{array}\right]\xrightarrow[1\mathbf{I}]{(+)}\left[\begin{array}{ccc}0&0&0\\1&1&0\\-1&3&1\\\hline 1&-1&1\\0&1&-2\\0&0&1\end{array}\right]$$

$$\xrightarrow[0\mathbf{I}]{(-)}\left[\begin{array}{ccc}0&0&0\\\hline 1&1&0\\-1&3&1\\\hline 1&-1&1\\0&1&-2\\0&0&1\end{array}\right]\xrightarrow{\substack{\boldsymbol{\tau}\\[(-1)\mathbf{2}+\mathbf{1}]\\[(4)\mathbf{3}+\mathbf{1}]}}\left[\begin{array}{ccc}0&0&0\\\hline 0&1&0\\0&3&1\\\hline 6&-1&1\\-9&1&-2\\4&0&1\end{array}\right]\xrightarrow[\substack{\boldsymbol{\tau}\\[(1)\mathbf{1}+\mathbf{2}]\\[(-4)\mathbf{1}+\mathbf{3}]}]{}\left[\begin{array}{ccc}0&0&0\\\hline 0&1&0\\0&3&1\\\hline 6&-1&1\\-9&1&-2\\4&0&1\end{array}\right]\xrightarrow[0\mathbf{I}]{(+)}\left[\begin{array}{ccc}0&0&0\\0&1&0\\0&3&1\\\hline 6&-1&1\\-9&1&-2\\4&0&1\end{array}\right]=\left[\begin{array}{c}\mathbf{C}\\\hline\mathbf{S}\end{array}\right]$$

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.

Librería NAcAL para Python

```
A = Matrix([[-2,0,3],[3,-2,-9],[-1,2,6]]); L=[1,1,0];
C = Diagonaliza (A, L, 1) # Matriz diagonal por bloques triangulares
S = C.S                   # C = (S**-1) * A * S  (La matriz S se guarda como un atributo de C)
(S**-1) * A * S           # Comprobación
```

Como se ha visto, el algoritmo es un poco pesado para ser calculado con “papel y lápiz”…¡pero para eso se inventaron los ordenadores!

Ejemplo 31. $\mathbf{A}=\begin{bmatrix}3&0&-1&1\\3&2&-2&2\\1&-2&2&0\\-3&-2&3&-1\end{bmatrix}$ con autovalores 2 (doble) y 1 (doble). Diagonalicemos por bloques:

Librería 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 = Diagonaliza (A, L)      # Añada como tercer argumento un 1 si quiere ver los pasos
C.S                         # La matriz S se guarda como atributo S
```

Así pues, $\mathbf{C} = \begin{bmatrix}2&0&0&0\\2&2&0&0\\0&0&1&0\\0&0&1&1\end{bmatrix} = \begin{bmatrix}0&0&0&-2\\1&0&2&-2\\1&-1&2&-2\\1&-1&0&2\end{bmatrix}^{-1} \begin{bmatrix}3&0&-1&1\\3&2&-2&2\\1&-2&2&0\\-3&-2&3&-1\end{bmatrix} \begin{bmatrix}0&0&0&-2\\1&0&2&-2\\1&-1&2&-2\\1&-1&0&2\end{bmatrix} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}.$

17.2.3. Autovalores, determinante y traza

Como matrices semejantes tienen el mismo determinante, traza y autovalores (con la misma multiplicidad aritmética y geométrica); y puesto que por el Teorema 17.2.4 sabemos que para toda matriz cuadrada podemos encontrar otra similar a ella, que es diagonal por bloques *con sus autovalores en la diagonal principal*, se deduce que:

Corolario 17.2.9. *La suma de los autovalores de* $\mathbf{A}$ *es igual a la traza de* $\mathbf{A}$*.*

Corolario 17.2.10. *El producto de los autovalores de* $\mathbf{A}$ *es igual al determinante de* $\mathbf{A}$*.*

17.2.4. 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 λ_i en su diagonal principal. Fijémonos ahora en aquellas columnas de $\mathbf{C}$ en la que aparezca un autovalor λ_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 17.2.11. *Si* $\mathbf{C} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ *es diagonal por bloques triangulares y* $\mathbf{C}_{|j}$ *tan solo tiene ceros por debajo del autovalor* λ_i*, entonces* $\mathbf{S}_{|j}$ *es un autovector asociado al autovalor* λ_i*.*

Así, usando nuestra librería de Python con el último ejemplo, llegamos a $\left[\begin{array}{c}\mathbf{C}\\ \hline \mathbf{S}\end{array}\right] = \left[\begin{array}{cccc}2&0&0&0\\2&2&0&0\\0&0&1&0\\0&0&1&1\\ \hline 0&0&0&-2\\1&0&2&-2\\1&-1&2&-2\\1&-1&0&2\end{array}\right],$

y deducimos que $\begin{pmatrix}0, & 0, & -1, & -1\end{pmatrix}$ es un autovector para $\lambda = 2$; y $\begin{pmatrix}-2, & -2, & -2, & 2\end{pmatrix}$ lo es para $\lambda = 1$.

Nótese que tras diagonalizar por bloques triangulares

$$\left[\begin{array}{c}\mathbf{A}\\ \hline \mathbf{I}\end{array}\right] \xrightarrow[esp(\boldsymbol{\tau}_1^{-1}\cdots\boldsymbol{\tau}_p^{-1})]{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p} \left[\begin{array}{c}\mathbf{C}\\ \hline \mathbf{S}\end{array}\right] \qquad \text{donde} \quad \mathbf{S} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p},$$

puesto que $\mathbf{S}$ es invertible, necesariamente *los autovectores correspondientes a autovalores distintos son linealmente independientes*.

17.3. Matrices diagonalizables

Definición 17.4 (Matriz diagonalizable). *Se dice que* $\mathbf{A}$ *(de orden* n*) es diagonalizable si existe una matriz* $\mathbf{S}$ *tal que* $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ *es diagonal.*

Ejemplo 32.

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

$$\xrightarrow[1\mathbf{I}]{(-)} \left[\begin{array}{ccc}2&4&0\\0&0&0\\ \hline 0&0&-1\\ \hline 1&0&2\\0&1&-2\\0&0&1\end{array}\right] \xrightarrow{\substack{\boldsymbol{\tau}\\ [(-2)\mathbf{1}+\mathbf{2}]}} \left[\begin{array}{ccc}2&0&0\\0&0&0\\ \hline 0&0&-1\\ \hline 1&-2&2\\0&1&-2\\0&0&1\end{array}\right] \xrightarrow[\substack{\boldsymbol{\tau}\\ [(2)\mathbf{2}+\mathbf{1}]}]{} \left[\begin{array}{ccc}2&0&0\\0&0&0\\ \hline 0&0&-1\\ \hline 1&-2&2\\0&1&-2\\0&0&1\end{array}\right] \xrightarrow[1\mathbf{I}]{(+)} \left[\begin{array}{ccc}3&0&0\\0&1&0\\0&0&0\\ \hline 1&-2&2\\0&1&-2\\0&0&1\end{array}\right] = \left[\begin{array}{c}\mathbf{D}\\ \hline \mathbf{S}\end{array}\right]$$

$$\mathbf{AS}_{|1} = \begin{bmatrix}3&4&2\\0&1&2\\0&0&0\end{bmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix} = 3\cdot\mathbf{A}_{|1} \qquad \mathbf{AS}_{|2} = \begin{bmatrix}3&4&2\\0&1&2\\0&0&0\end{bmatrix}\begin{pmatrix}-2\\1\\0\end{pmatrix} = 1\cdot\mathbf{A}_{|2}$$

$$\mathbf{AS}_{|3} = \begin{bmatrix}3&4&2\\0&1&2\\0&0&0\end{bmatrix}\begin{pmatrix}2\\-2\\1\end{pmatrix} = 0\cdot\mathbf{A}_{|3}.$$

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 λ_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* λ_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).*

Por tanto:

Corolario 17.3.2. *Si todos los autovalores* $\mathbf{A}$ *son distintos entre si, 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

$$\mathbf{A}^{-1} = \left(\mathbf{SDS}^{-1}\right)^{-1} = \mathbf{SD}^{-1}\mathbf{S}^{-1}.$$

Además, si seguimos el siguiente convenio

$$\mathbf{A}^0 = \mathbf{A}^{(1-1)} = \mathbf{A}^1\left(\mathbf{A}^{-1}\right) = \mathbf{I};$$

entonces, para toda $\mathbf{A}$ diagonalizable e invertible y todo numero entero n: $\mathbf{A}^n = \mathbf{SD}^n\mathbf{S}^{-1}$.

Diagonalización de matrices simétricas

18.1. Método de Gram-Schmidt

Definición 18.1. *Decimos que 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. Si $\mathbf{0} = a_1\boldsymbol{z}_1 + \cdots + a_n\boldsymbol{z}_n$ entonces para cada j

$$0 \;=\; \mathbf{0}\cdot\boldsymbol{z}_j \;=\; \Big(a_1\boldsymbol{z}_1 + \cdots + a_n\boldsymbol{z}_n\Big)\cdot\boldsymbol{z}_j \;=\; a_j\boldsymbol{z}_i\cdot\boldsymbol{z}_j$$

como $\boldsymbol{z}_j\cdot\boldsymbol{z}_j \neq 0$, entonces, necesariamente $a_j = 0$ para $j = 1,\ldots,n$. □

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 $\big[\boldsymbol{y}_1;\;\ldots\;\boldsymbol{y}_n;\;\boldsymbol{y}_{(n+1)}\big]$, aplicando la hipótesis de inducción, existe una sucesión, $\boldsymbol{\tau}_1,\ldots,\boldsymbol{\tau}_k$, de transformaciones elementales de "izquierda a derecha" que transforman el sub-sistema formado por los n primeros vectores en otro sistema $\big[\boldsymbol{z}_1;\;\ldots\;\boldsymbol{z}_n\big]$ que es equivalente, pero además es ortogonal; por tanto, aplicando las transformaciones elementales sobre el sistema completo tenemos

$$\big[\boldsymbol{y}_1;\;\ldots\;\boldsymbol{y}_n;\;\boldsymbol{y}_{(n+1)}\big]_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \big[\boldsymbol{z}_1;\;\ldots\;\boldsymbol{z}_n;\;\boldsymbol{y}_{(n+1)}\big].$$

Y por último, mediante una segunda sucesión de transformaciones elementales "de izquierda a derecha", $\boldsymbol{\tau}_{(k+1)},\ldots,\boldsymbol{\tau}_p$, 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

$$\boldsymbol{z}_{(n+1)}\cdot\boldsymbol{z}_j \;=\; 0,\qquad \text{para } j = 1,\ldots,n.$$

Para cada $j = 1,\ldots,n$ 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

Por tanto, la segunda sucesión de transformaciones elementales realiza la siguiente transformación.

$$\boldsymbol{z}_{(n+1)} = \boldsymbol{y}_{(n+1)} - \sum_{\boldsymbol{z}_j \neq \boldsymbol{0}} \widehat{a_j}\, \boldsymbol{z}_j, \qquad \text{donde} \quad \widehat{a_j} \;=\; \frac{\boldsymbol{y}_{(n+1)} \cdot \boldsymbol{z}_j}{\boldsymbol{z}_j \cdot \boldsymbol{z}_j},$$

de manera que tras las dos sucesiones de transformaciones tenemos que

$$\begin{bmatrix}\boldsymbol{y}_1; & \dots & \boldsymbol{y}_n; & \boldsymbol{y}_{(n+1)}\end{bmatrix} \xrightarrow{\boldsymbol{\tau}_1,\dots\boldsymbol{\tau}_k} \begin{bmatrix}\boldsymbol{z}_1; & \dots & \boldsymbol{z}_n; & \boldsymbol{y}_{(n+1)}\end{bmatrix} \xrightarrow{\boldsymbol{\tau}_{(k+1)},\dots\boldsymbol{\tau}_p} \begin{bmatrix}\boldsymbol{z}_1; & \dots & \boldsymbol{z}_n; & \boldsymbol{z}_{(n+1)}\end{bmatrix}$$

es un sistema equivalente y ortogonal. □

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 $\begin{bmatrix}\boldsymbol{z}_1; & \dots & \boldsymbol{z}_r\end{bmatrix}$ un sistema ortogonal de vectores de $\mathbb{R}^n$ y sea $\begin{bmatrix}\boldsymbol{y}_1; & \dots & \boldsymbol{y}_m\end{bmatrix}$ un sistema generador de $\mathbb{R}^n$. Entonces aplicando Gram-Schmidt sobre el sistema ampliado

$$\begin{bmatrix}\boldsymbol{z}_1; & \dots & \boldsymbol{z}_r; & \boldsymbol{y}_1; & \dots & \boldsymbol{y}_m\end{bmatrix}$$

comenzando sobre el vector $(r+1)$ésimo, obtenemos un nuevo sistema equivalente al ampliado y ortogonal[2]

$$\begin{bmatrix}\boldsymbol{z}_1; & \dots & \boldsymbol{z}_r; & \boldsymbol{z}_{r+1}; & \dots & \boldsymbol{z}_{r+m}\end{bmatrix}.$$

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 un sistema 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 61. Demuestre el siguiente corolario:

Corolario 18.2.1. *Dado un vector* $\boldsymbol{q}$ *de* $\mathbb{R}^n$ *de norma uno, existe una base ortonormal de* $\mathbb{R}^n$ *cuyo último vector es* $\boldsymbol{q}$*.*

Definición 18.3 (Vectores ortonormales)**.** *Los vectores* $\boldsymbol{q}_1,\dots,\boldsymbol{q}_k$ *de* $\mathbb{R}^n$ *son* ortonormales *si son perpendiculares entre si y de norma uno, es decir, si*

$$(\boldsymbol{q}_i)\cdot(\boldsymbol{q}_j) = \begin{cases} 0 & \textit{cuando } i \neq j \quad (\textit{vectores } \text{ortogonales}) \\ 1 & \textit{cuando } i = j \quad (\textit{vectores } \text{unitarios}: \|\boldsymbol{q}_1\| = 1) \end{cases}$$

Habitualmente denotamos con $\mathbf{Q}$ las matrices cuyas columnas son *orto-normales*.

EJERCICIO 62. Demuestre la siguiente proposición:

Proposición 18.2.2. *Las columnas de* $\underset{m\times n}{\mathbf{Q}}$ *son orto-normales si y solo si* $\mathbf{Q}^\intercal\mathbf{Q} = \underset{n\times n}{\mathbf{I}}$

[2]En realidad se puede empezar desde la posición 1, pues los vectores $\boldsymbol{y}_i$ ya son ortogonales entre si, y por tanto el método no los modifica.

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

$$\mathbf{Q}^\intercal\mathbf{Q} = \begin{bmatrix}1&0&0\\0&1&0\end{bmatrix}\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix} = \begin{bmatrix}1&0\\0&1\end{bmatrix} = \underset{2\times2}{\mathbf{I}}; \quad \text{pero} \quad \mathbf{Q}(\mathbf{Q}^\intercal) = \begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix} \neq \underset{3\times3}{\mathbf{I}}.$$

Definición 18.4. *Decimos que* $\mathbf{Q}$ *es* ortogonal *si es cuadrada y sus columnas son ortonormales.*

Corolario 18.2.3. *Cuando* $\mathbf{Q}$ *es ortogonal,* $\mathbf{Q}^\intercal = \mathbf{Q}^{-1}$.

Nótese por tanto que las columnas de una matriz ortogonal $\mathbf{Q}$ de orden n forman una base de $\mathbb{R}^n$.

EJERCICIO 63. 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* λ *de una matriz real y simétrica son reales.*

Demostración. Suponga $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ (con $\boldsymbol{x} \neq \mathbf{0}$); y donde tanto λ como $\boldsymbol{x}$ son complejos. Entonces multiplicando $\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ por el conjugado $\overline{\boldsymbol{x}}$, tenemos

$$\overline{\boldsymbol{x}}\,\mathbf{A}\boldsymbol{x} = \overline{\boldsymbol{x}}\cdot\lambda\boldsymbol{x} = \lambda(\overline{\boldsymbol{x}}\cdot\boldsymbol{x}),$$

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

$$\overline{\boldsymbol{x}}\,\mathbf{A}\boldsymbol{x} = \overline{\lambda}\,\overline{\boldsymbol{x}}\cdot\boldsymbol{x} = \overline{\lambda}\,(\overline{\boldsymbol{x}}\cdot\boldsymbol{x}).$$

Como en las dos ecuaciones de más arriba 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. $\square$

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 una matriz ortogonal $\mathbf{Q}$ tal que $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ es diagonal.*

Demostración. Vamos a demostrarlo describiendo el algoritmo que construye dicha base, y que consta de dos etapas.

Paso inicial. Sea $\boldsymbol{q}_n$ un autovector de $\mathbf{A}$ correspondiente al autovalor λ_n y de norma 1. Aplicando el Corolario 18.2.1 extendemos el sistema formado por dicho autovector hasta obtener una base ortonormal de $\mathbb{R}^n$

$$\mathbf{Q}_n=\begin{bmatrix}\boldsymbol{q}_1 & \cdots & \boldsymbol{q}_n\end{bmatrix}.$$

Por tanto $\mathbf{Q}_n$ es ortogonal, y $\mathbf{Q}_n^{-1}\mathbf{A}\mathbf{Q}_n$ es simétrica y de la forma $\mathbf{Q}_n^{-1}\mathbf{A}\mathbf{Q}_n=\begin{bmatrix} & & & 0\\ & \underset{(n-1)\times(n-1)}{\mathbf{A}'} & & 0\\ & & & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}$, ya que

$$\left(\mathbf{Q}_n^{-1}\mathbf{A}\mathbf{Q}_n\right)_{|n}=\left(\mathbf{Q}_n^{\mathsf{T}}\mathbf{A}\mathbf{Q}_n\right)_{|n}=\mathbf{Q}_n^{\mathsf{T}}\mathbf{A}\boldsymbol{q}_n=\lambda_n\cdot\mathbf{Q}_n^{\mathsf{T}}\boldsymbol{q}_n=\lambda_n\begin{pmatrix}\boldsymbol{q}_1\cdot\boldsymbol{q}_n\\ \boldsymbol{q}_2\cdot\boldsymbol{q}_n\\ \vdots\\ \boldsymbol{q}_n\cdot\boldsymbol{q}_n\end{pmatrix}=\begin{pmatrix}0\\0\\ \vdots\\ \lambda_n\end{pmatrix}=\lambda_n\cdot\mathbf{I}_{|n}.$$

Paso de continuación. Supongamos que tras $n-k$ pasos tenemos

$$\mathbf{Q}_{k+1}^{-1}\cdots\mathbf{Q}_n^{-1}\cdot\mathbf{A}\cdot\mathbf{Q}_n\cdots\mathbf{Q}_{k+1}=\left[\begin{array}{c|ccc}\underset{k\times k}{\mathbf{A}'} & & & \\ \hline & \lambda_{k+1} & & \\ & & \ddots & \\ & & & \lambda_n\end{array}\right],\tag{18.1}$$

como $\mathbf{A}'$ es simétrica, tomamos un autovector $\boldsymbol{x}_k\in\mathbb{R}^k$ de $\mathbf{A}'$ correspondiente al autovalor λ_k y de norma 1. Aplicando el Corolario 18.2.1 formamos una base ortonormal $\mathbf{R}=\begin{bmatrix}\boldsymbol{x}_1, & \cdots, & \boldsymbol{x}_k\end{bmatrix}$ de $\mathbb{R}^k$. Ahora, si consideramos la matriz por bloques

$$\mathbf{Q}_k=\left[\begin{array}{c|c}\mathbf{R} & \mathbf{0}\\ \hline \mathbf{0} & \underset{(n-k)\times(n-k)}{\mathbf{I}}\end{array}\right]=\begin{bmatrix}\boldsymbol{y}_1, & \ldots, & \boldsymbol{y}_n\end{bmatrix},$$

entonces $\mathbf{Q}_k$ es ortogonal y sus últimas columnas $\boldsymbol{y}_k,\ldots,\boldsymbol{y}_n$ son autovectores de la matriz de la Ecuación (18.1); es decir, llamando $\mathbf{B}$ a la matriz de dicha ecuación tenemos que

$$\mathbf{B}\boldsymbol{y}_j=\lambda_j\boldsymbol{y}_j;\quad\text{para}\quad j=k,\ldots,n.$$

Consecuentemente, multiplicando (18.1) por $\mathbf{Q}_k^{-1}$ y por $\mathbf{Q}_k$ tenemos

$$\mathbf{Q}_k^{-1}\cdot\left(\mathbf{Q}_{k+1}^{-1}\cdots\mathbf{Q}_n^{-1}\cdot\mathbf{A}\cdot\mathbf{Q}_n\cdots\mathbf{Q}_{k+1}\right)\cdot\mathbf{Q}_k\;=\;\mathbf{Q}_k^{-1}\mathbf{B}\mathbf{Q}_k\;=\left[\begin{array}{c|ccc}\underset{k-1\times k-1}{\mathbf{A}''} & & & \\ \hline & \lambda_k & & \\ & & \ddots & \\ & & & \lambda_n\end{array}\right],$$

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 $\lambda_1,\ldots,\lambda_n$ de $\mathbf{A}$ correspondientes a las columnas de $\mathbf{Q}$ en su diagonal principal. □

Definición 18.5. $\underset{n\times n}{\mathbf{A}}$ *es diagonalizable ortogonalmente si existe una matriz ortogonal* $\mathbf{Q}$ *tal que* $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ *es diagonal.*

Como consecuencia del teorema anterior tenemos el siguiente e importante corolario que se conoce por 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 si*

Demostración. Considere dos autovectores, $\boldsymbol{x}$ e $\boldsymbol{y}$, correspondientes a autovalores distintos: $\mathbf{A}\boldsymbol{x} = \lambda_1\boldsymbol{x}$ y $\mathbf{A}\boldsymbol{y} = \lambda_2\boldsymbol{y}$. Entonces
$$\lambda_1\boldsymbol{x}\cdot\boldsymbol{y} \;=\; \mathbf{A}\boldsymbol{x}\cdot\boldsymbol{y} \;=\; \boldsymbol{x}\left(\mathbf{A}^{\intercal}\right)\boldsymbol{y} \;=\; \boldsymbol{x}\mathbf{A}\boldsymbol{y} \;=\; \lambda_2\left(\boldsymbol{x}\cdot\boldsymbol{y}\right).$$
Puesto que $\lambda_1 \neq \lambda_2$ necesariamente:
$$\lambda_1(\boldsymbol{x}\cdot\boldsymbol{y}) - \lambda_2(\boldsymbol{x}\cdot\boldsymbol{y}) = 0 \quad\Longrightarrow\quad \left(\lambda_1-\lambda_2\right)\boldsymbol{x}\cdot\boldsymbol{y} = 0 \quad\Longrightarrow\quad \boldsymbol{x}\cdot\boldsymbol{y} = 0.$$
□

Ejercicio 64. Demuestre la siguiente proposición:

Proposición 18.4.6. *Si una matriz es diagonalizable ortogonalmente, entonces es simétrica.*

Para finalizar, la demostración del Teorema 18.4.3 nos describe el algoritmo para diagonalizar $\mathbf{A}$ ortogonalmente y obtener una matriz ortogonal $\mathbf{Q}$ de autovectores; pero, como toda matriz simétrica es diagonalizable, también podemos obtener una matriz $\mathbf{Q}$ aplicando el algoritmo descrito en la demostración del Teorema 17.2.4 de diagonalización por bloques triangulares: primero aplicamos la diagonalización por semejanza
$$\left[\frac{\mathbf{A}}{\mathbf{I}}\right] \xrightarrow[esp(\boldsymbol{\tau}_1^{-1}\cdots\boldsymbol{\tau}_p^{-1})]{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p} \left[\frac{\mathbf{C}}{\mathbf{S}}\right] \qquad \text{donde} \quad \mathbf{S} = \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p},$$
y a continuación aplicamos Gram-Schmidt sobre las columnas de $\mathbf{S}$ para asegurarnos que las columnas de $\mathbf{S}$ que corresponden a autovalores repetidos son perpendiculares; finalmente normalizamos todas las columnas para obtener una matriz $\mathbf{Q}$.

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, producto y exponentes no negativos de las variables. Un ejemplo de polinomio en x es $x^2 - 4x + 7$. Y un ejemplo de polinomio en tres variables es $x^3 + 2xyz^2 - yz + 1$.

El grado de un polinomio es el mayor de los grados de sus monomios (o términos individuales) con coeficientes no nulos. El grado de cada término es la suma de los exponentes de las variables que aparecen en él, y por tanto nunca es negativo. 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 $(1 + 1 + 2)$.

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 A simétrica y de orden n define una forma cuadrática q_{A} en n variables mediante la siguiente formula:

$$q_{\mathsf{A}}(\boldsymbol{x}) = \boldsymbol{x}\mathsf{A}\boldsymbol{x}.$$

Recíprocamente, dada una forma cuadrática en n variables, sus coeficientes se pueden arreglar 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 $\begin{bmatrix} 1 & -2 & \\ -2 & 4 & \\ & & 5 \end{bmatrix}$ 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* $\mathsf{A} \in \mathbb{R}^{n\times n}$ *es* definida positiva *si para todo vector* $\boldsymbol{x} \neq \boldsymbol{0}$ *de* $\mathbb{R}^n$, $\boldsymbol{x}\mathsf{A}\boldsymbol{x} > 0$.

Como consecuencia de la definición tenemos que

Proposición 19.1.1. *Si* $\underset{m\times n}{\mathsf{A}}$ *tiene rango* n, *entonces* $\mathsf{A}^\intercal\mathsf{A}$ *es definida positiva.*

Demostración. Por una parte,

$$\boldsymbol{x}\big(\mathbf{A}^{\intercal}\big)\mathbf{A}\boldsymbol{x} = \big(\mathbf{A}\boldsymbol{x}\big)\cdot\big(\mathbf{A}\boldsymbol{x}\big) = \|\mathbf{A}\boldsymbol{x}\|^2 \geq 0.$$

Pero 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 $\boldsymbol{x}\mathbf{A}\boldsymbol{x} = \|\mathbf{A}\boldsymbol{x}\|^2 > 0$. □

Hay una estrecha relación entre los signos de los autovalores de la matriz y el signo de la forma cuadrática asociada.

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 la misma idea de la última demostración junto al hecho de que toda matriz real y simétrica es diagonalizable ortogonalmente de la forma $\mathbf{D} = \mathbf{Q}^{-1}\,\mathbf{A}\,\mathbf{Q}$, donde $\mathbf{D}$ tiene los autovalores de $\mathbf{A}$ en la diagonal principal, y donde las columnas de $\mathbf{Q}$ son autovectores correspondientes a dichos autovalores...

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

$$\boldsymbol{x}\mathbf{A}\boldsymbol{x} = \boldsymbol{x}\mathbf{Q}\mathbf{D}\big(\mathbf{Q}^{\intercal}\big)\boldsymbol{x} = \big(\big(\mathbf{Q}^{\intercal}\big)\boldsymbol{x}\big)\mathbf{D}\big(\big(\mathbf{Q}^{\intercal}\big)\boldsymbol{x}\big) = \boldsymbol{y}\mathbf{D}\boldsymbol{y}, \qquad (\text{donde} \quad \boldsymbol{y} = \big(\mathbf{Q}^{\intercal}\big)\boldsymbol{x}).$$

De esta expresión es evidente que $\boldsymbol{x}\mathbf{A}\boldsymbol{x}$ es 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 λ_j son los autovalores de $\mathbf{A}$. Como $\mathbf{Q}$ es invertible, $\boldsymbol{y} = \mathbf{Q}^{\intercal}\boldsymbol{x}$ es distinto de cero siempre que $\boldsymbol{x}$ también lo sea. Por tanto, si los autovalores son positivos entonces la suma $\sum_{j=1}^{n} \lambda_j \big(y_j^2\big)$ es positiva.

Por otra parte, la forma cuadrática nunca puede ser positiva si algún autovalor λ_j es negativo o cero; para verlo basta elegir $\boldsymbol{y} = \mathbf{I}_{|j}$ para ver que $\boldsymbol{y}\mathbf{D}\boldsymbol{y} = {}_{j|}\mathbf{I}\mathbf{D}\mathbf{I}_{|j} = {}_{j|}\mathbf{D}_{|j} = \lambda_j$. □

Ejercicio 65. Demuestre las siguientes proposiciones.

(a) Si $\mathbf{A}$ y $\mathbf{B}$ son definidas positivas, entonces la suma $(\mathbf{A} + \mathbf{B})$ también es definida positiva.
(b) Si $\mathbf{A}$ es simétrica y definida positiva, entonces $\mathbf{A}^{-1}$ también es definida positiva.

19.2. Diagonalización de matrices simétricas por congruencia

☞ 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 comprobar que la expresión solo contiene sumas de cuadrados).

Este resultado tiene importancia práctica. No es posible encontrar las raíces de un polinomio cualquiera (tan solo está asegurado para polinomios de grado menor o igual a 4). Pero al menos, la diagonalización por congruencia nos revelará *los signos* de los autovalores de *cualquier matriz simétrica*. Este resultado se llama *Ley de inercia*[a].

[a] Aunque este año no incluiré su demostración, pues nos hemos dejado algunas cosas importantes por el camino que necesito para la demostración.

To-Do: Incluir demostración *ley de inercia* cuando haya incluido aplicaciones lineales en la primera parte del curso

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 lección vamos a ver que, dada una matriz simétrica $\mathbf{A}$ siempre es posible encontrar una matriz diagonal que es congruente con $\mathbf{A}$ mediante transformaciones elementales de sus filas y columnas. La demostración que veremos más abajo describe los pasos a seguir para diagonalizar por congruencia (paso de inicio y paso de continuación). Pero antes de exponer la demostración, veamos un ejemplo de diagonalización por congruencia. El método consiste en tratar de escalonar una matriz, pero aplicaremos a las filas todas las operaciones que hayamos aplicado a las columnas.

Ejemplo 33. 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:

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \xrightarrow[{[(1)\mathbf{1}+\mathbf{2}]}]{\substack{\boldsymbol{\tau} \\ [(2)\mathbf{2}]}} \begin{bmatrix} 2 & 0 & 0 \\ -1 & 3 & -1 \\ 0 & -2 & 2 \end{bmatrix} \xrightarrow[\substack{[(1)\mathbf{1}+\mathbf{2}] \\ [(2)\mathbf{2}]}]{\boldsymbol{\tau}} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & -2 \\ 0 & -2 & 2 \end{bmatrix} \xrightarrow[{[(1)\mathbf{2}+\mathbf{3}]}]{\substack{\boldsymbol{\tau} \\ [(3)\mathbf{3}]}} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & -2 & 4 \end{bmatrix} \xrightarrow[\substack{[(1)\mathbf{2}+\mathbf{3}] \\ [(3)\mathbf{3}]}]{\boldsymbol{\tau}} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{bmatrix}.$$

Fíjese que aunque la matriz obtenida es diagonal, ni 6 ni 12 son autovalores de la matriz original (este método NO nos dice quienes son los autovalores). Pero como los tres componentes de la diagonal son positivos, sabemos que $\mathbf{A}$ es definida positiva.

Y si en la diagonal tenemos un cero ¿qué podemos hacer?...Podemos sumar a la primera columna otra columna (aunque en un ejemplo posterior veremos que hay que hacerlo con cuidado, puesto no siempre funciona).

Ejemplo 34. 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 si 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*):

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \xrightarrow{\substack{\boldsymbol{\tau} \\ [(1)\mathbf{2}+\mathbf{1}]}} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \xrightarrow[\substack{\boldsymbol{\tau} \\ [(1)\mathbf{2}+\mathbf{1}]}]{} \begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \xrightarrow{\substack{\boldsymbol{\tau} \\ [(2)\mathbf{2}] \\ [(-1)\mathbf{1}+\mathbf{2}] \\ [(-1)\mathbf{1}+\mathbf{3}]}} \begin{bmatrix} 2 & 0 & 0 \\ 1 & -1 & 0 \\ 2 & 0 & -2 \end{bmatrix} \xrightarrow[\substack{\boldsymbol{\tau} \\ [(-1)\mathbf{1}+\mathbf{3}] \\ [(-1)\mathbf{1}+\mathbf{2}] \\ [(2)\mathbf{2}]}]{} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Un tercer ejemplo nos indica que debemos andar con cuidado cuando encontramos un cero en la diagonal.

Ejemplo 35. Considere la matriz, $\begin{bmatrix} 0 & 1 \\ 1 & a \end{bmatrix}$ que tiene un cero en donde querríamos tener un pivote con el que eliminar todo lo que queda a su derecha (y por debajo).

Primer intento: podemos intentar la estrategia del ejemplo anterior, y sumar la segunda columna a la primera (y repetir la operación con las filas):

$$\begin{bmatrix} 0 & 1 \\ 1 & a \end{bmatrix} \xrightarrow{\substack{\boldsymbol{\tau} \\ [(1)\mathbf{2}+\mathbf{1}]}} \begin{bmatrix} 1 & 1+a \\ 1 & a \end{bmatrix} \xrightarrow[\substack{\boldsymbol{\tau} \\ [(1)\mathbf{2}+\mathbf{1}]}]{} \begin{bmatrix} 2+a & 1+a \\ 1+a & a \end{bmatrix}$$

y ahora con el pivote $(2+a)$ podemos anular todo lo que está a su derecha (y por debajo), pero…¿cómo sabemos que es un pivote? si $a=-2$, estamos como al principio…(en el anterior ejemplo la estrategia funcionó porque sumamos un vector que tenía un cero en la diagonal,…este tenía el número a que puede hacer fallar el algoritmo)

Segundo intento: para que la estrategia anterior no falle, debemos asegurarnos de que el vector que sumamos a la primera columna tiene un cero en la componente situada en la diagonal. Como en este caso no hay un vector así, sencillamente intercambiamos filas y columnas para colocar el cero de la primera posición de la diagonal en otro lugar.

$$\begin{bmatrix}0&1\\1&-2\end{bmatrix}\xrightarrow[]{\substack{\boldsymbol{\tau}\\ [\mathbf{1}\rightleftharpoons\mathbf{2}]}}\begin{bmatrix}1&0\\-2&1\end{bmatrix}\xrightarrow[\substack{\boldsymbol{\tau}\\ [\mathbf{1}\rightleftharpoons\mathbf{2}]}]{}\begin{bmatrix}-2&1\\1&0\end{bmatrix}$$

Ahora ya podemos eliminar todo lo que queda a la derecha y por debajo del *pivote* sin problemas.

$$\begin{bmatrix}0&1\\1&-2\end{bmatrix}\xrightarrow[]{\substack{\boldsymbol{\tau}\\ [\mathbf{1}\rightleftharpoons\mathbf{2}]}}\begin{bmatrix}1&0\\-2&1\end{bmatrix}\xrightarrow[\substack{\boldsymbol{\tau}\\ [\mathbf{1}\rightleftharpoons\mathbf{2}]}]{}\begin{bmatrix}-2&1\\1&0\end{bmatrix}\xrightarrow[]{\substack{\boldsymbol{\tau}\\ [(2)\mathbf{2}]\\ [(1)\mathbf{1}+\mathbf{2}]}}\begin{bmatrix}-2&0\\1&1\end{bmatrix}\xrightarrow[\substack{\boldsymbol{\tau}\\ [(1)\mathbf{1}+\mathbf{2}]\\ [(2)\mathbf{2}]}]{}\begin{bmatrix}-2&0\\0&2\end{bmatrix}$$

Vayamos con la demostración de los pasos para 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*

$$\left[\begin{array}{c|c} * & \\ \hline & \underset{(n-1)\times(n-1)}{\mathbf{A}'}\end{array}\right].$$

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, y aplicando las mismas operaciones sobre las filas llegamos a la siguiente matriz congruente con $\mathbf{A}$:

$${}_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}=\mathbf{B}^{\intercal}\mathbf{A}\mathbf{B}=\left[\begin{array}{c|c} a_{11} & \\ \hline & \underset{(n-1)\times(n-1)}{\mathbf{A}'}\end{array}\right];\quad\text{donde}\quad \mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}=\mathbf{B}.$$

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 finalmente termina por situarse en la posición $(1,1)$. Con lo que hemos llegado al 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 $\mathbf{A}_{|j}$ a la primera columna, y por

tanto, también sumamos ${}_{j|}\mathbf{A}$ a la primera fila. Entonces tendremos que

$$\underset{\boldsymbol{\tau}_{[(1)\boldsymbol{j}+\mathbf{1}]}}{}\mathbf{A}\underset{\boldsymbol{\tau}_{[(1)\boldsymbol{j}+\mathbf{1}]}}{} = \left[\begin{array}{c|c} 2a_{1j} & * \quad \dots \quad * \\ \hline * & \\ \vdots & \underset{(n-1)\times(n-1)}{\mathbf{A}'} \\ * & \end{array}\right];$$

así que de nuevo hemos llegado al Caso 1.

□

Proposición 19.2.2 (Paso de continuación)**.** *Dada una matriz simétrica de orden* n *de la forma*

$$\left[\begin{array}{ccc|c} * & & & \\ & \ddots & & \\ & & * & \\ \hline & & & \underset{k\times k}{\mathbf{A}'} \end{array}\right],$$

existe una matriz no singular $\mathbf{B}$ *tal que* $\mathbf{B}^{\mathsf{T}}\mathbf{A}\mathbf{B}$ *es una matriz simétrica de la forma*

$$\left[\begin{array}{ccc|c} * & & & \\ & \ddots & & \\ & & * & \\ \hline & & & \underset{(k-1)\times(k-1)}{\mathbf{A}''} \end{array}\right].$$

Demostración. La demostración es como la del *Paso de Inicio*, pero trabajando con las filas y columnas de la submatriz $\mathbf{A}'$. □

De la combinación de las dos proposiciones anteriores llegamos al siguiente corolario:

Corolario 19.2.3. *Para toda matriz simétrica de orden* n*, existe una matriz invertible* $\mathbf{B}$ *tal que* $\mathbf{B}^{\mathsf{T}}\mathbf{A}\mathbf{B}$ *es diagonal.*

Ejemplo 36*.* Diagonalice por congruencia la matriz

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

Piense en qué pasos debe sumar una columna (y una fila) para generar un pivote en la diagonal, y cuando debe intercambiar columnas (y filas).

Librería NAcAL para Python

```
A = Matrix([[0,1,0,0],[1,0,1,1],[0,1,0,1],[0,1,1,1]])
D = DiagonalizaC(A,1)
```

19.3. Algunos tipos de 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.

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 66. Demuestre que si $\mathbf{A}$ es definida positiva, entonces todos los elementos de la diagonal deben ser positivos (y si es semidefinida positiva todos los elementos de la diagonal deben ser no negativos).

To-Do: Me falta la otra regla: $|a_{ij}| < |a_{ii}|$

19.4. Completar el cuadrado para clasificar 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 muchas formas distintas 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, tenemos que

$$\mathbf{D} = \mathbf{B}^{\intercal}\mathbf{A}\mathbf{B} \qquad \Rightarrow \qquad \mathbf{A} = \left(\mathbf{B}^{-1}\right)^{\intercal} \cdot \mathbf{D} \cdot \mathbf{B}^{-1}$$

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 del vector $\boldsymbol{y} = \mathbf{B}^{-1}\boldsymbol{x}$ 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 19.4.1. *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.*

Es lo que ocurría con la matriz del Ejemplo 33 en la página 189. Vamos repetir su diagonalización, pero ahora realizando solo transformaciones *Tipo I* (así minimizamos el número de operaciones, con lo que facilitamos pensar quien es la inversa de $\mathbf{B}$; 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):

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

Así, la forma cuadrática $\boldsymbol{x}\mathbf{A}\boldsymbol{x}$ se puede re-escribir como suma de cuadrados: [1]

$$\begin{aligned}
\boldsymbol{x}\mathbf{A}\boldsymbol{x} &= \begin{pmatrix}x & y & z\end{pmatrix}\begin{bmatrix}1&0&0\\-1/2&1&0\\0&-2/3&1\end{bmatrix}\begin{bmatrix}2&&\\&3/2&\\&&4/3\end{bmatrix}\begin{bmatrix}1&-1/2&0\\0&1&-2/3\\0&0&1\end{bmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix} && = \boldsymbol{x}\left(\mathbf{B}^{-1}\right)^{\intercal}\mathbf{D}\mathbf{B}^{-1}\boldsymbol{x}\\
&= \begin{pmatrix}(x-1/2y) & (y-2/3z) & z\end{pmatrix}\begin{bmatrix}2&&\\&3/2&\\&&4/3\end{bmatrix}\begin{pmatrix}(x-1/2y)\\(y-2/3z)\\z\end{pmatrix} && = \left(\mathbf{B}^{-1}\boldsymbol{x}\right)\cdot\mathbf{D}\cdot\left(\mathbf{B}^{-1}\boldsymbol{x}\right)\\
&= 2\cdot(x-1/2y)^2 \;+\; 3/2\cdot(y-2/3z)^2 \;+\; 4/3\cdot(z)^2 && = \sum{}_{j|}\mathbf{D}_{|j}\cdot\left(y_j\right)^2 \quad (\text{con } \boldsymbol{y} = \mathbf{B}^{-1}\boldsymbol{x}).
\end{aligned}$$

[1] Esto está relacionado con la factorización LDU, Y ahora que he separado la eliminación (de izquierda a derecha) de la eliminación Gaussiana (intercambio), el resultado de la factorización $\mathbf{AP} = \mathbf{LU}$ es casi inmediato. Quizá debería volver a contar la factorización LU tras la Lección 5.

$\boldsymbol{x}\mathbf{A}\boldsymbol{x}$ es una suma de tres términos al cuadrado, cada uno de ellos multiplicado por un pivote de $\mathbf{D}$; como todos los pivotes son positivos, la forma cuadrática es evidentemente definida positiva. Algo que también podemos verificar en este caso si calculamos los tres autovalores, pues la ecuación característica de $\mathbf{A}$ es

$$\begin{vmatrix} 2-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 2-\lambda \end{vmatrix} = (2-\lambda)^3 - 2(2-\lambda) = 0 \quad \Longrightarrow \quad \begin{cases} \lambda = 2 \\ (2-\lambda)^2 - 2 = \lambda^2 - 4\lambda + 2 \begin{cases} \lambda = 2+\sqrt{2} \\ \lambda = 2-\sqrt{2} \end{cases} \end{cases} .$$

De manera similar, como la matriz diagonal del Ejemplo 34 en la página 189 tiene dos pivotes negativos y uno positivo, deducimos que la correspondiente forma cuadrática es indefinida (y de hecho, los correspondientes autovalores son -1 (doble) y 2).

Librería NAcAL para Python

```
A = Matrix([[0,1,0,0],[1,0,1,1],[0,1,0,1],[0,1,1,1]])
D = DiagonalizaC(A,1)
B = D.B
~B*A*B                    # D es congruente con A
```

☞ **Nota 4.** *El Teorema Espectral muestra una forma muy especial de completar el cuadrado de una forma cuadrática; pues la diagonalización*

$$\mathbf{D} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{Q}^{\intercal}\mathbf{A}\mathbf{Q},$$

lo es simultáneamente por semejanza y por congruencia.

Corolario 19.4.2. *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.*

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

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

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

$$\boxed{(\mathbf{A}+\mathbf{B})_{|j} = \mathbf{A}_{|j} + \mathbf{B}_{|j}} \qquad \text{y} \qquad \boxed{\big(\lambda \mathbf{B}\big)_{|j} = \lambda\big(\mathbf{B}_{|j}\big)}$$

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

$$\boxed{\big(\mathbf{A}^{\intercal}\big)_{|j} = {}_{j|}\mathbf{A}}$$

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

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

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

Reglas asociativas (desplazando el paréntesis)

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

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

[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

$$(\boldsymbol{b}\lambda)_{|i} = (\boldsymbol{b}_{|i})\lambda \qquad\qquad {}_{i|}(\lambda\boldsymbol{b}) = \lambda({}_{i|}\boldsymbol{b})$$
$$(\mathbf{A}\lambda)_{|j} = (\mathbf{A}_{|j})\lambda \qquad\qquad {}_{i|}(\lambda\mathbf{A}) = \lambda({}_{i|}\mathbf{A})$$

Lección 2

Comienza con el *producto punto* (o producto escalar usual) de dos vectores de $\mathbb{R}^n$ y sus propiedades.[2]

$$\boxed{\boldsymbol{a}\cdot\boldsymbol{b} = \sum_i a_i b_i}$$

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 $\mathbf{B}$).

$$\boxed{\mathbf{A}\boldsymbol{b} = \sum_j (\mathbf{A}_{|j})b_j} \qquad \text{y} \qquad \boxed{\boldsymbol{a}\mathbf{B} = (\mathbf{B}^{\intercal})\boldsymbol{a}}$$

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

- $\mathbf{A}(\lambda\boldsymbol{b}) = (\lambda\mathbf{A})\boldsymbol{b}$
- $(\mathbf{A}+\mathbf{B})\boldsymbol{c} = \mathbf{A}\boldsymbol{c} + \mathbf{B}\boldsymbol{c}$
- $\mathbf{A}(\mathbf{B}\boldsymbol{c}) = \Big[\mathbf{A}(\mathbf{B}_{|1}), \quad \dots, \quad \mathbf{A}(\mathbf{B}_{|n})\Big]\boldsymbol{c}$

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*.

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

(*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

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

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 $\mathbf{A}\big(\mathbf{B}_{|j}\big)$, como el elemento $j$ésimo de una combinación de las filas $\big(_{i|}\mathbf{A}\big)\mathbf{B}$, 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 $(\mathbf{AB})^{\intercal} = \mathbf{B}^{\intercal}\big(\mathbf{A}^{\intercal}\big)$ 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:

- $\boxed{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}} \text{ suma } \lambda \text{ veces el vector } i\text{ésimo al } j\text{ésimo.}}$
- $\boxed{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}} \text{ multiplica por } \alpha \text{ el vector } i\text{ésimo.}}$

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

$$\mathbf{A}_{\boldsymbol{\tau}} = \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}}\right) \qquad \text{y} \qquad _{\boldsymbol{\tau}}\mathbf{A} = \left(_{\boldsymbol{\tau}}\mathbf{I}\right)\mathbf{A}.$$

Para describir la aplicación de la secuencia $\boldsymbol{\tau}_1 \dots \boldsymbol{\tau}_k$ de k trasformaciones elementales de las *columnas* de $\mathbf{A}$ usamos el esquema. [3]

$$\mathbf{A} \xrightarrow{\boldsymbol{\tau}_1} \left(\mathbf{A}_{\boldsymbol{\tau}_1}\right) \xrightarrow{\boldsymbol{\tau}_2} \left(\mathbf{A}_{\boldsymbol{\tau}_1\boldsymbol{\tau}_2}\right) \cdots \xrightarrow{\boldsymbol{\tau}_k} \left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) \quad \text{o agregando varios pasos} \quad \mathbf{A} \xrightarrow[]{\substack{\boldsymbol{\tau}_1\\ \vdots\\ \boldsymbol{\tau}_p}} \left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p}\right) \xrightarrow[]{\substack{\boldsymbol{\tau}_{p+1}\\ \vdots\\ \boldsymbol{\tau}_k}} \left(\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right).$$

Cuando se aplica la sucesión $\boldsymbol{\tau}_1 \dots \boldsymbol{\tau}_k$ de k transformaciones elementales sobre las *columnas*, o bien la sucesión $\boldsymbol{\tau}_k \dots \boldsymbol{\tau}_1$ 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:

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

[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

$$\mathbf{A} \xrightarrow[\boldsymbol{\tau}_1]{} \left(_{\boldsymbol{\tau}_1}\mathbf{A}\right) \xrightarrow[\boldsymbol{\tau}_2]{} \left(_{\boldsymbol{\tau}_2\boldsymbol{\tau}_1}\mathbf{A}\right) \cdots \xrightarrow[\boldsymbol{\tau}_k]{} \left(_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{A}\right) \quad \text{o agregando varios pasos} \quad \mathbf{A} \xrightarrow[\substack{\boldsymbol{\tau}_p\\ \vdots\\ \boldsymbol{\tau}_1}]{} \left(_{\boldsymbol{\tau}_p\cdots\boldsymbol{\tau}_1}\mathbf{A}\right) \xrightarrow[\substack{\boldsymbol{\tau}_k\\ \vdots\\ \boldsymbol{\tau}_{p+1}}]{} \left(_{\boldsymbol{\tau}_k\cdots\boldsymbol{\tau}_1}\mathbf{A}\right);$$

donde la secuencia $\boldsymbol{\tau}_k \dots \boldsymbol{\tau}_1$ es la transpuesta de la secuencia $\boldsymbol{\tau}_1 \dots \boldsymbol{\tau}_k$, es decir, $\big(\boldsymbol{\tau}_1 \dots \boldsymbol{\tau}_k\big)^{\intercal} = \boldsymbol{\tau}_k \dots \boldsymbol{\tau}_1$, 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:

- $\boxed{\underset{[i\rightleftharpoons j]}{\boldsymbol{\tau}} \text{ intercambia de posición los vectores } i\text{ésimo y } j\text{ésimo.}}$

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

$$\mathbf{A}_{\underset{[\circlearrowright]}{\boldsymbol{\tau}}} = \mathbf{A}\left(\mathbf{I}_{\underset{[\circlearrowright]}{\boldsymbol{\tau}}}\right) \qquad \text{y} \qquad {}_{\underset{[\circlearrowright]}{\boldsymbol{\tau}}}\mathbf{A} = \left({}_{\underset{[\circlearrowright]}{\boldsymbol{\tau}}}\mathbf{I}\right)\mathbf{A}$$

(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 $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\left(\mathbf{A}^{-1}\right)$.
- 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 $\left(\mathbf{A}^{\intercal}\right)^{-1} = \left(\mathbf{A}^{-1}\right)^{\intercal}$.
- 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 4.4.2 en la página 53)

$$\mathbf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_p} = \mathbf{K},$$

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 60)

$$\mathbf{K}_{\boldsymbol{\tau}_{(p+1)}\cdots\boldsymbol{\tau}_k} = \mathbf{I}.$$

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,

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

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. (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. 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 incorrecta, pues entendemos $\boldsymbol{b} + \boldsymbol{b}_{|2}$ como 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*.

El modo correcto de escribir la operación sería $\big(\boldsymbol{b} + \boldsymbol{b}\big)_{|2}$, así que, por claridad, en Python también deberíamos escribir `(b+b)|2`.

□

Ejercicio 3. Mostraremos ambas estrategias en cada caso. En el caso de la segunda, en azul aparecen las operaciones entre números reales.

1. Estrategia 1:

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

   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:

$$\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c}) = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} + \left(\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} + \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix}\right) = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} + \begin{pmatrix} b_1+c_1 \\ \vdots \\ b_m+c_m \end{pmatrix} = \begin{pmatrix} a_1+b_1+c_1 \\ \vdots \\ a_m+b_m+c_m \end{pmatrix} =$$
$$= \begin{pmatrix} a_1+b_1 \\ \vdots \\ a_m+b_m \end{pmatrix} + \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} = \left(\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}\right) + \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} = (\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}.$$

   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:

$$\boldsymbol{0}+\boldsymbol{a} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} 0+a_1 \\ \vdots \\ 0+a_m \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = \boldsymbol{a}.$$

   Estrategia 2: $\big(\boldsymbol{0}+\boldsymbol{a}\big)_{|i} = \boldsymbol{0}_{|i} + \boldsymbol{a}_{|i} = 0 + \boldsymbol{a}_{|i} = \boldsymbol{a}_{|i}.$

4. Estrategia 1:

$$\boldsymbol{a}+(-\boldsymbol{a})=\begin{pmatrix}a_1\\ \vdots\\ a_m\end{pmatrix}+\begin{pmatrix}-a_1\\ \vdots\\ -a_m\end{pmatrix}=\begin{pmatrix}a_1-a_1\\ \vdots\\ a_m-a_m\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}=\boldsymbol{0}.$$

Estrategia 2: $\Big(\boldsymbol{a}+(-\boldsymbol{a})\Big)_{|i}=\boldsymbol{a}_{|i}+(-\boldsymbol{a})_{|i}=0=\boldsymbol{0}_{|i}.$

5. Estrategia 1:

$$\lambda(\boldsymbol{a}+\boldsymbol{b})=\begin{pmatrix}\lambda(a_1+b_1)\\ \vdots\\ \lambda(a_m+b_m)\end{pmatrix}=\begin{pmatrix}\lambda a_1+\lambda b_1\\ \vdots\\ \lambda a_m+\lambda b_m\end{pmatrix}=\begin{pmatrix}\lambda a_1\\ \vdots\\ \lambda a_m\end{pmatrix}+\begin{pmatrix}\lambda b_1\\ \vdots\\ \lambda b_m\end{pmatrix}=\lambda\boldsymbol{a}+\lambda\boldsymbol{b}.$$

Estrategia 2: $\Big(\lambda(\boldsymbol{a}+\boldsymbol{b})\Big)_{|i}=\lambda\Big((\boldsymbol{a}+\boldsymbol{b})_{|i}\Big)=\lambda\Big(\boldsymbol{a}_{|i}+\boldsymbol{b}_{|i}\Big)=\lambda(\boldsymbol{a}_{|i})+\lambda(\boldsymbol{b}_{|i})=(\lambda\boldsymbol{a})_{|i}+(\lambda\boldsymbol{b})_{|i}=$ $(\lambda\boldsymbol{a}+\lambda\boldsymbol{b})_{|i}.$

6. Estrategia 1:

$$(\lambda+\eta)\boldsymbol{a}=\begin{pmatrix}(\lambda+\eta)a_1\\ \vdots\\ (\lambda+\eta)a_m\end{pmatrix}=\begin{pmatrix}\lambda a_1+\eta a_1\\ \vdots\\ \lambda a_m+\eta a_m\end{pmatrix}=\begin{pmatrix}\lambda a_1\\ \vdots\\ \lambda a_m\end{pmatrix}+\begin{pmatrix}\eta b_1\\ \vdots\\ \eta b_m\end{pmatrix}=\lambda\boldsymbol{a}+\eta\boldsymbol{a}.$$

Estrategia 2: $\Big((\lambda+\eta)\boldsymbol{a}\Big)_{|i}=(\lambda+\eta)(\boldsymbol{a}_{|i})=\lambda(\boldsymbol{a}_{|i})+\eta(\boldsymbol{a}_{|i})=(\lambda\boldsymbol{a})_{|i}+(\eta\boldsymbol{a})_{|i}=(\lambda\boldsymbol{a}+\eta\boldsymbol{a})_{|i}.$

7. Estrategia 1:

$$\lambda(\eta\boldsymbol{a})=\lambda\begin{pmatrix}\eta a_1\\ \vdots\\ \eta a_m\end{pmatrix}=\begin{pmatrix}\lambda\eta a_1\\ \vdots\\ \lambda\eta a_m\end{pmatrix}=\lambda\eta\begin{pmatrix}a_1\\ \vdots\\ a_m\end{pmatrix}=(\lambda\eta)\boldsymbol{a}.$$

Estrategia 2: $\Big(\lambda(\eta\boldsymbol{a})\Big)_{|i}=\lambda\Big((\eta\boldsymbol{a})_{|i}\Big)=\lambda\Big(\eta(\boldsymbol{a}_{|i})\Big)=(\lambda\eta)(\boldsymbol{a}_{|i})=((\lambda\eta)\boldsymbol{a})_{|i}.$

8. Estrategia 1:

$$1\boldsymbol{a}=\begin{pmatrix}1\cdot a_1\\ \vdots\\ 1\cdot a_m\end{pmatrix}=\begin{pmatrix}a_1\\ \vdots\\ a_m\end{pmatrix}=\boldsymbol{a}.$$

Estrategia 2: $(1\boldsymbol{a})_{|i}=1(\boldsymbol{a}_{|i})=\boldsymbol{a}_{|i}.$

□

Ejercicio 4(a) Primero multiplica el vector por 3 y luego selecciona la primera componente.

□

Ejercicio 4(b) ¡Selecciona la tercera componente del vector!

□

Ejercicio 4(c) ¡También selecciona la tercera componente del vector!

□

Ejercicio 4(d) Como en la primera expresión, primero multiplica el vector por 3 y luego selecciona la primera componente.

$\square$

Ejercicio 5(a) $\quad {}_{k|}(\mathbf{A}^\intercal)_{|j} = {}_{k|}\Big((\mathbf{A}^\intercal)_{|j}\Big) = {}_{k|}({}_{j|}\mathbf{A}) = ({}_{j|}\mathbf{A})_{|k} = {}_{j|}\mathbf{A}_{|k}.$

$\square$

Ejercicio 5(b) $\quad {}_{j|}\Big((\mathbf{A}^\intercal)^\intercal\Big)_{|k} = {}_{k|}(\mathbf{A}^\intercal)_{|j} = {}_{j|}\mathbf{A}_{|k}.$

$\square$

Ejercicio 5(c) $\quad \mathbf{A}_{|i} = \Big((\mathbf{A}^\intercal)^\intercal\Big)_{|i} = {}_{i|}(\mathbf{A}^\intercal).$

$\square$

Ejercicio 6. Las demostraciones son prácticamente idénticas a las vistas en el caso de los vectores. En azul aparecen las operaciones entre vectores.

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

$\square$

Ejercicio 7. Comenzamos por la suma.

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

Y ahora el producto.

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

$\square$

Ejercicio 8(a) $\quad \Big((\lambda\mathbf{A})^\intercal\Big)_{|j} = {}_{j|}(\lambda\mathbf{A}) = \lambda\Big({}_{j|}\mathbf{A}\Big) = \lambda\Big((\mathbf{A}^\intercal)_{|j}\Big) = \Big(\lambda(\mathbf{A}^\intercal)\Big)_{|j}.$

$\square$

Ejercicio 8(b) $\big((\mathbf{A}+\mathbf{B})^\intercal\big)_{|j} = {}_{j|}(\mathbf{A}+\mathbf{B}) = {}_{j|}\mathbf{A} + {}_{j|}\mathbf{B} = \big(\mathbf{A}^\intercal\big)_{|j} + \big(\mathbf{B}^\intercal\big)_{|j} = \big(\mathbf{A}^\intercal + \mathbf{B}^\intercal\big)_{|j}.$

$\square$

Ejercicio 9. Comenzamos con la suma

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

Y continuamos con el producto:

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

$\square$

Ejercicio 10(a) $\boldsymbol{x}\cdot\boldsymbol{y} \;=\; x_1y_1+\cdots+x_ny_n \;=\; y_1x_1+\cdots+y_nx_n \;=\; \boldsymbol{y}\cdot\boldsymbol{x}$

$\square$

Ejercicio 10(b)

- $(a\boldsymbol{x})\cdot\boldsymbol{y} \;=\; ax_1y_1+\cdots+ax_ny_n \;=\; a\big(x_1y_1+\cdots+x_ny_n\big) \;=\; a(\boldsymbol{x}\cdot\boldsymbol{y}).$
- $(\boldsymbol{x}+\boldsymbol{y})\cdot\boldsymbol{z} \;=\; (x_1+y_1)z_1+\cdots+(x_n+y_n)z_n \;=\; x_1z_1+\cdots+x_nz_n \;+\; y_1z_1+\cdots+y_nz_n \;=\; \boldsymbol{x}\cdot\boldsymbol{z}+\boldsymbol{y}\cdot\boldsymbol{z}.$

$\square$

Ejercicio 10(c) $\boldsymbol{x}\cdot\boldsymbol{x} \;=\; x_1^2+\cdots+x_n^2 \;\geq 0.$

$\square$

Ejercicio 10(d) $\boldsymbol{x}\cdot\boldsymbol{x} \;=\; x_1^2+\cdots+x_n^2 \;=0 \iff x_i=0$ para $i=1:n$.

$\square$

Ejercicio 11. $2\begin{pmatrix}1\\1\\-2\end{pmatrix}+2\begin{pmatrix}1\\-2\\1\end{pmatrix}+2\begin{pmatrix}-2\\1\\1\end{pmatrix} = \begin{pmatrix}2\\2\\-4\end{pmatrix}+\begin{pmatrix}2\\-4\\2\end{pmatrix}+\begin{pmatrix}-4\\2\\2\end{pmatrix} = \begin{pmatrix}2+2-4\\2-4+2\\-4+2+2\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix} = \mathbf{0}.$

$\square$

Ejercicio 12.

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

$\square$

Ejercicio 13. $\mathbf{A}\big(\mathbf{I}_{|j}\big) = \mathbf{A}\begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix} = \big(\mathbf{A}_{|1}\big)0+\cdots+\big(\mathbf{A}_{|j}\big)1+\cdots+\big(\mathbf{A}_{|n}\big)0 = \mathbf{A}_{|j}.$

$\square$

Ejercicio 14. Sea $\mathbf{A}$, de n columnas y $\boldsymbol{b}$ de $\mathbb{R}^n$, entonces

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

□

Ejercicio 15(a)

$${}_{i|}\big(\mathbf{A}(\boldsymbol{b}+\boldsymbol{c})\big) = \big({}_{i|}\mathbf{A}\big)\cdot(\boldsymbol{b}+\boldsymbol{c}) = \big({}_{i|}\mathbf{A}\big)\cdot\boldsymbol{b} + \big({}_{i|}\mathbf{A}\big)\cdot\boldsymbol{c} = {}_{i|}\big(\mathbf{A}\boldsymbol{b}\big) + {}_{i|}\big(\mathbf{A}\boldsymbol{c}\big) = {}_{i|}\big(\mathbf{A}\boldsymbol{b}+\mathbf{A}\boldsymbol{c}\big).$$

□

Ejercicio 15(b)

$${}_{i|}\big(\mathbf{A}(\lambda\boldsymbol{b})\big) = \big({}_{i|}\mathbf{A}\big)\cdot(\lambda\boldsymbol{b}) = \lambda\Big(\big({}_{i|}\mathbf{A}\big)\cdot\boldsymbol{b}\Big) = \lambda\Big({}_{i|}\big(\mathbf{A}\boldsymbol{b}\big)\Big) = {}_{i|}\big(\lambda(\mathbf{A}\boldsymbol{b})\big).$$

□

Ejercicio 16.

$$\begin{aligned}
\mathbf{A}(\mathbf{B}\boldsymbol{c}) &= \mathbf{A}\Big(\big(\mathbf{B}_{|1}\big)c_1 + \cdots + \big(\mathbf{B}_{|n}\big)c_n\Big) && \text{por la Definición 2.3 de producto} \\
&= \mathbf{A}\big(\mathbf{B}_{|1}\big)c_1 + \cdots + \mathbf{A}\big(\mathbf{B}_{|n}\big)c_n && \text{proposiciones 2.2.1 y 2.2.2} \\
&= \Big[\mathbf{A}\big(\mathbf{B}_{|1}\big), \quad \ldots, \quad \mathbf{A}\big(\mathbf{B}_{|n}\big)\Big]\,\boldsymbol{c} && \text{por la Definición 2.3 de producto.}
\end{aligned}$$

□

Ejercicio 17(a)

$${}_{i|}\big(\mathbf{A}(\lambda\boldsymbol{b})\big) \;=\; \big({}_{i|}\mathbf{A}\big)\cdot(\lambda\boldsymbol{b}) \;=\; \lambda\Big(\big({}_{i|}\mathbf{A}\big)\cdot\boldsymbol{b}\Big) \;=\; \Big(\lambda\big({}_{i|}\mathbf{A}\big)\Big)\cdot\boldsymbol{b} \;=\; \Big({}_{i|}\big(\lambda\mathbf{A}\big)\Big)\cdot\boldsymbol{b} \;=\; {}_{i|}\big((\lambda\mathbf{A})\boldsymbol{b}\big)$$

□

Ejercicio 17(b)

$${}_{i|}\big((\mathbf{A}+\mathbf{B})\boldsymbol{c}\big) = \big({}_{i|}(\mathbf{A}+\mathbf{B})\big)\cdot\boldsymbol{c} = \big({}_{i|}\mathbf{A} + {}_{i|}\mathbf{B}\big)\cdot\boldsymbol{c} = \big({}_{i|}\mathbf{A}\big)\cdot\boldsymbol{c} + \big({}_{i|}\mathbf{B}\big)\cdot\boldsymbol{c} = {}_{i|}\big(\mathbf{A}\boldsymbol{c}\big) + {}_{i|}\big(\mathbf{B}\boldsymbol{c}\big) = {}_{i|}\big(\mathbf{A}\boldsymbol{c}+\mathbf{B}\boldsymbol{c}\big).$$

□

Ejercicio 18(a) $\quad \boldsymbol{a}\mathbf{I} = \big(\mathbf{I}^{\intercal}\big)\boldsymbol{a} = \mathbf{I}\boldsymbol{a} = \boldsymbol{a}.$

□

Ejercicio 18(b) $\quad \big({}_{i|}\mathbf{I}\big)\mathbf{A} = \big(\mathbf{A}^{\intercal}\big)\big({}_{i|}\mathbf{I}\big) = \big(\mathbf{A}^{\intercal}\big)\big(\mathbf{I}_{|i}\big) = \big(\mathbf{A}^{\intercal}\big)_{|i} = {}_{i|}\mathbf{A}.$

□

Ejercicio 18(c) $(\boldsymbol{a}+\boldsymbol{b})\mathbf{C} = \big(\mathbf{C}^{\intercal}\big)(\boldsymbol{a}+\boldsymbol{b}) \underset{(*)}{=} \big(\mathbf{C}^{\intercal}\big)\boldsymbol{a} + \big(\mathbf{C}^{\intercal}\big)\boldsymbol{b} = \boldsymbol{a}\mathbf{C} + \boldsymbol{b}\mathbf{C}.$

□

Ejercicio 18(d) $(\lambda\boldsymbol{a})\mathbf{B} = \big(\mathbf{B}^{\intercal}\big)(\lambda\boldsymbol{a}) \underset{(*)}{=} \lambda\Big(\big(\mathbf{B}^{\intercal}\big)\boldsymbol{a}\Big) = \lambda(\boldsymbol{a}\mathbf{B})$

□

Ejercicio 18(e) $\boldsymbol{a}(\lambda\mathbf{B}) = \big((\lambda\mathbf{B})^{\intercal}\big)\boldsymbol{a} = \Big(\lambda\big(\mathbf{B}^{\intercal}\big)\Big)\boldsymbol{a} \underset{(*)}{=} \big(\mathbf{B}^{\intercal}\big)(\lambda\boldsymbol{a}) = (\lambda\boldsymbol{a})\mathbf{B}.$

□

Ejercicio 18(f) $\boldsymbol{a}(\mathbf{B}+\mathbf{C}) = (\mathbf{B}+\mathbf{C})^{\intercal}\boldsymbol{a} = \Big(\big(\mathbf{B}^{\intercal}\big) + \big(\mathbf{C}^{\intercal}\big)\Big)\boldsymbol{a} \underset{(*)}{=} \big(\mathbf{B}^{\intercal}\big)\boldsymbol{a} + \big(\mathbf{C}^{\intercal}\big)\boldsymbol{a} = \boldsymbol{a}\mathbf{B} + \boldsymbol{a}\mathbf{C}.$

□

Ejercicio 19(a) Recordando la Proposición 2.2.3 en la página 26 (*) y usando la Definición 3.1 de producto:

$$\mathbf{A}(\mathbf{B}\boldsymbol{c}) \underset{(*)}{=} \Big[\mathbf{A}\big(\mathbf{B}_{|1}\big), \quad \dots, \quad \mathbf{A}\big(\mathbf{B}_{|n}\big)\Big]\,\boldsymbol{c} = (\mathbf{A}\mathbf{B})\boldsymbol{c}.$$

□

Ejercicio 19(b) Aplicando repetidamente la definición de producto y una vez $(*)$ la proposición anterior tenemos:

$$\Big(\mathbf{A}\big(\mathbf{B}\mathbf{C}\big)\Big)_{|j} = \mathbf{A}\Big(\big(\mathbf{B}\mathbf{C}\big)_{|j}\Big) = \mathbf{A}\Big(\mathbf{B}\big(\mathbf{C}_{|j}\big)\Big) \underset{(*)}{=} \Big(\mathbf{A}\mathbf{B}\Big)\big(\mathbf{C}_{|j}\big) = \Big(\big(\mathbf{A}\mathbf{B}\big)\mathbf{C}\Big)_{|j}, \quad j=1:n.$$

□

Ejercicio 20(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

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

□

Ejercicio 20(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

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

□

Ejercicio 20(c)

$$\Big(\mathbf{A}(\lambda\mathbf{B})\Big)_{|j} = \mathbf{A}\Big((\lambda\mathbf{B})_{|j}\Big) = \mathbf{A}\Big(\lambda\big(\mathbf{B}_{|j}\big)\Big)\begin{cases} = (\lambda\mathbf{A})\big(\mathbf{B}_{|j}\big) = \Big((\lambda\mathbf{A})\mathbf{B}\Big)_{|j} \\ \\ = \lambda\Big(\mathbf{A}\big(\mathbf{B}_{|j}\big)\Big) = \lambda\Big(\big(\mathbf{A}\mathbf{B}\big)_{|j}\Big) = \Big(\lambda\big(\mathbf{A}\mathbf{B}\big)\Big)_{|j} \end{cases}, \quad j=1:n.$$

$\square$

Ejercicio 20(d) $\big(\mathsf{IA}\big)_{|j} = \mathsf{I}\big(\mathsf{A}_{|j}\big) = \mathsf{A}_{|j},\quad j = 1:n.$

$\square$

Ejercicio 20(e) $\big(\mathsf{AI}\big)_{|j} = \mathsf{A}\big(\mathsf{I}_{|j}\big) = \mathsf{A}_{|j},\quad j = 1:n.$

$\square$

Ejercicio 21.

$${}_{i|}\big(\mathsf{AB}\big)_{|j} = {}_{i|}\Big(\big(\mathsf{AB}\big)_{|j}\Big) = {}_{i|}\Big(\mathsf{A}\big(\mathsf{B}_{|j}\big)\Big) = \big({}_{i|}\mathsf{A}\big)\cdot\big(\mathsf{B}_{|j}\big)$$

$\square$

Ejercicio 22.

Veamos que las componentes $j$ésimas son iguales:

$$\Big({}_{i|}\big(\mathsf{AB}\big)\Big)_{|j} = {}_{i|}\big(\mathsf{AB}\big)_{|j} = \big({}_{i|}\mathsf{A}\big)\cdot\big(\mathsf{B}_{|j}\big) = \Big(\big({}_{i|}\mathsf{A}\big)\mathsf{B}\Big)_{|j};$$

por tanto: ${}_{i|}\big(\mathsf{AB}\big) = \big({}_{i|}\mathsf{A}\big)\mathsf{B}$.

$\square$

Ejercicio 23. ${}_{i|}(\mathsf{AB})^{\intercal}{}_{|j} \;=\; {}_{j|}\mathsf{AB}_{|i} \;=\; {}_{j|}\mathsf{A}\cdot\mathsf{B}_{|i} \;=\; \mathsf{B}_{|i}\cdot{}_{j|}\mathsf{A} \;=\; {}_{i|}\big(\mathsf{B}^{\intercal}\big)\big(\mathsf{A}^{\intercal}\big)_{|j}$

$\square$

Ejercicio 24(a) $\begin{bmatrix}1&0&-7&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}.$

$\square$

Ejercicio 24(b) $\begin{bmatrix}1&0&0&1\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}.$

$\square$

Ejercicio 24(c) $\begin{bmatrix}1&0&0&0\\3&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}.$

$\square$

Ejercicio 24(d) $\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&-10&0\\0&0&0&1\end{bmatrix}.$

$\square$

Ejercicio 25. a) $\underset{[(-7)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}$ transforma $\mathbf{B}_{|3}$, restándole siete veces $\mathbf{B}_{|1}$, es decir: $\Big(\mathbf{B}_{\underset{[(-7)\mathbf{1}+\mathbf{3}]}{\boldsymbol{\tau}}}\Big)_{|3} = -7\mathbf{B}_{|1} + \mathbf{B}_{|3}$.

b) $\underset{[(1)\mathbf{1}+\mathbf{4}]}{\boldsymbol{\tau}}$ transforma la cuarta columna de $\mathbf{B}$, sumándole la primera columna: $\Big(\mathbf{B}_{\underset{[(1)\mathbf{1}+\mathbf{4}]}{\boldsymbol{\tau}}}\Big)_{|4} = \mathbf{B}_{|1} + \mathbf{B}_{|4}$.

c) $\underset{[(3)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}$ transforma la primera columna de $\mathbf{B}$, sumándole tres veces la segunda: $\Big(\mathbf{B}_{\underset{[(3)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}}\Big)_{|1} = 3\mathbf{B}_{|2}+\mathbf{B}_{|1}$.

d) $\underset{[(-10)\mathbf{3}]}{\boldsymbol{\tau}}$ transforma la tercera columna de $\mathbf{B}$, multiplicándola por -10, es decir: $\Big(\mathbf{B}_{\underset{[(-10)\mathbf{3}]}{\boldsymbol{\tau}}}\Big)_{|3} = -10\mathbf{B}_{|3}$.

$\square$

Ejercicio 26(a) $\mathbf{A}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}} = \mathbf{A}\Big(\mathbf{I}_{\underset{[(3)\mathbf{1}]}{\boldsymbol{\tau}}}\Big) = \begin{bmatrix}1&1&1&1\\0&1&1&1\\1&2&3&4\end{bmatrix}\begin{bmatrix}3&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} = \begin{bmatrix}3&1&1&1\\0&1&1&1\\3&2&3&4\end{bmatrix}$.

$\square$

Ejercicio 26(b) $\mathbf{A}_{\underset{[(-1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}} = \mathbf{A}\Big(\mathbf{I}_{\underset{[(-1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}}\Big) = \begin{bmatrix}1&1&1&1\\0&1&1&1\\1&2&3&4\end{bmatrix}\begin{bmatrix}1&-1&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} = \begin{bmatrix}1&0&1&1\\0&1&1&1\\1&1&3&4\end{bmatrix}$.

$\square$

Ejercicio 27. Por una parte,

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

y por otra

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

$\square$

Ejercicio 28(a) Una posible sucesión de transformaciones elementales por columnas es:

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} \xrightarrow[\text{Tipo I}]{\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}} \begin{bmatrix}1&1&0\\0&1&0\\0&0&1\end{bmatrix} \xrightarrow[\text{Tipo I}]{\underset{[(-1)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}} \begin{bmatrix}0&1&0\\-1&1&0\\0&0&1\end{bmatrix} \xrightarrow[\text{Tipo I}]{\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}} \begin{bmatrix}0&1&0\\-1&0&0\\0&0&1\end{bmatrix} \xrightarrow[\text{Tipo II}]{\underset{[(-1)\mathbf{1}]}{\boldsymbol{\tau}}} \begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},$$

es decir, $\mathbf{I}_{\underset{[\mathbf{1}\rightleftharpoons\mathbf{2}]}{\boldsymbol{\tau}}} = \mathbf{I}_{\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-1)\mathbf{2}+\mathbf{1}]}{\boldsymbol{\tau}}\underset{[(1)\mathbf{1}+\mathbf{2}]}{\boldsymbol{\tau}}\underset{[(-1)\mathbf{1}]}{\boldsymbol{\tau}}}$.

$\square$

Ejercicio 28(b) Hay muchas combinaciones posibles. Por ejemplo: $\mathbf{I}_{\underset{[\boldsymbol{i}\rightleftharpoons\boldsymbol{j}]}{\boldsymbol{\tau}}} = \mathbf{I}_{\underset{[(-1)\boldsymbol{j}]}{\boldsymbol{\tau}}\underset{[(-1)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}\underset{[(1)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}\underset{[(-1)\boldsymbol{j}+\boldsymbol{i}]}{\boldsymbol{\tau}}}$.

$\square$

Ejercicio 29(a) Por una parte $\quad \mathbf{A}\mathbf{B}\mathbf{B}^{-1}\mathbf{A}^{-1} = \mathbf{A}\mathbf{I}\mathbf{A}^{-1} = \mathbf{I}, \quad$ y por otra $\quad \mathbf{B}^{-1}\mathbf{A}^{-1}\mathbf{A}\mathbf{B} = \mathbf{B}^{-1}\mathbf{I}\mathbf{B} = \mathbf{I}.$

$\square$

Ejercicio 29(b) Basta aplicar iterativamente la proposición anterior:

$$(\mathbf{A}_1 \cdots \mathbf{A}_k)^{-1} = (\mathbf{A}_2 \cdots \mathbf{A}_k)^{-1}\mathbf{A}_1^{-1} = (\mathbf{A}_3 \cdots \mathbf{A}_k)^{-1}\mathbf{A}_2^{-1}\mathbf{A}_1^{-1} = \cdots = (\mathbf{A}_{k-1}\mathbf{A}_k)^{-1}\mathbf{A}_{k-2}^{-1} \cdots \mathbf{A}_1^{-1} = \mathbf{A}_k^{-1} \cdots \mathbf{A}_1^{-1}.$$

$\square$

Ejercicio 29(c) Si $\mathbf{A}$ es invertible, entonces $\mathbf{A}\mathbf{B}$ es invertible (Proposición 5.1.2 en la página 57).

Supongamos que $\mathbf{A}\mathbf{B}$ es invertible. Entonces sabemos que existe $\mathbf{C}$ tal que

$$\big(\mathbf{A}\mathbf{B}\big)\mathbf{C} = \mathbf{I}, \qquad \text{y} \qquad \mathbf{C}\big(\mathbf{A}\mathbf{B}\big) = \mathbf{I}.$$

De la primera igualdad tenemos que $\mathbf{A}$ es invertible por la derecha, pues

$$\mathbf{A}\big(\mathbf{B}\mathbf{C}\big) = \mathbf{I}.$$

Solo nos falta comprobar que también es invertible por la izquierda, es decir, que $\big(\mathbf{B}\mathbf{C}\big)\mathbf{A} = \mathbf{I}$. 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

$$\mathbf{B}\mathbf{C}\big(\mathbf{A}\mathbf{B}\big)\mathbf{B}^{-1} = \mathbf{B}\mathbf{I}\mathbf{B}^{-1} \qquad \Rightarrow \qquad \big(\mathbf{B}\mathbf{C}\big)\mathbf{A} = \mathbf{I}.$$

$\square$

Ejercicio 29(d) Si $\mathbf{B}$ es invertible, entonces $\mathbf{A}\mathbf{B}$ es invertible (Proposición 5.1.2 en la página 57).

Supongamos que $\mathbf{A}\mathbf{B}$ es invertible. Entonces sabemos que existe $\mathbf{C}$ tal que

$$\big(\mathbf{A}\mathbf{B}\big)\mathbf{C} = \mathbf{I}, \qquad \text{y} \qquad \mathbf{C}\big(\mathbf{A}\mathbf{B}\big) = \mathbf{I}.$$

De la segunda igualdad tenemos que $\mathbf{B}$ es invertible por la izquierda, pues

$$\big(\mathbf{C}\mathbf{A}\big)\mathbf{B} = \mathbf{I}.$$

Solo nos falta comprobar que también es invertible por la derecha, es decir, que $\mathbf{B}\big(\mathbf{C}\mathbf{A}\big) = \mathbf{I}$. 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

$$\mathbf{A}^{-1}\big(\mathbf{A}\mathbf{B}\big)\mathbf{C}\mathbf{A} = \mathbf{A}^{-1}\mathbf{I}\mathbf{A}, \qquad \Rightarrow \qquad \big(\mathbf{B}\mathbf{C}\big)\mathbf{A} = \mathbf{I}.$$

$\square$

Ejercicio 29(e) Aplicando la pista tenemos por un lado: $\mathbf{A}^{\intercal}\big(\mathbf{A}^{-1}\big)^{\intercal} = \big(\mathbf{A}^{-1}\mathbf{A}\big)^{\intercal} = (\mathbf{I})^{\intercal} = \mathbf{I}$ y por el otro $\big(\mathbf{A}^{-1}\big)^{\intercal}\mathbf{A}^{\intercal} = \big(\mathbf{A}\mathbf{A}^{-1}\big)^{\intercal} = (\mathbf{I})^{\intercal} = \mathbf{I}$. Es decir, $\big(\mathbf{A}^{-1}\big)^{\intercal} = \big(\mathbf{A}^{\intercal}\big)^{-1}$.

$\square$

Ejercicio 29(f) Sea $\mathbf{A}$ de orden n una matriz cuya $i$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 $\mathbf{A}^{-1}$:

$$\mathbf{A}^{-1}\big(\mathbf{A}_{|j}\big) = \mathbf{A}^{-1}(\mathbf{0}) \qquad \Rightarrow \qquad \Big(\mathbf{A}^{-1}\mathbf{A}\Big)_{|j} = \mathbf{0} \qquad \Rightarrow \qquad \mathbf{I}_{|j} = \mathbf{0}.$$

Pero sabemos que $\mathbf{I}_{|i}$ (la columna *i*ésima de $\mathbf{I}$) no es $\mathbf{0}$, así que hemos llegado a una contradicción. Por 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 ${}_{i|}\mathbf{A} = \left(\mathbf{A}^\intercal\right)_{|i} = \mathbf{0}$, llegamos a la misma contradicción.

□

Ejercicio 30. Si para alguna columna $\mathbf{A}_{|j}$ se verifica que $\mathbf{A}_{|j} = \mathbf{A}\boldsymbol{x}$ (con $\boldsymbol{x}_{|j}$ =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 $\mathbf{A}\boldsymbol{x}$:

$$\mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_{(n-1)}}\right) = \mathbf{A}_{\substack{\boldsymbol{\tau}\\ [(-x_1)\mathbf{1}+\mathbf{j}]\\ [(-x_2)\mathbf{2}+\mathbf{j}]\\ \vdots\\ [(-x_n)\mathbf{n}+\mathbf{j}]}} = \mathbf{A}\begin{bmatrix} 1 & & & -x_1 & & \\ & 1 & & -x_2 & & \\ & & \ddots & \vdots & & \\ & & & 1 & & \\ & & & \vdots & \ddots & \\ & & & -x_n & & 1 \end{bmatrix} = \left[\mathbf{A}_{|1};\ \mathbf{A}_{|2}; \cdots \mathbf{0}; \cdots \mathbf{A}_{|n}\right],$$

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 57).

□

Ejercicio 31(a)

$$\begin{aligned} \vec{0} &= \vec{z} + \vec{0} && \text{por la propiedad que hemos pedido a } \vec{z} \\ &= \vec{z} && \text{por la propiedad que tiene } \vec{0}. \end{aligned}$$

□

Ejercicio 31(b)

$$\vec{z} = \vec{z} + \vec{0} = \vec{z} + \left(\vec{x} + (-\vec{x})\right) = (\vec{z} + \vec{x}) + (-\vec{x}) = \vec{0} + (-\vec{x}) = -\vec{x}.$$

□

Ejercicio 32.

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 $\vec{x} + \vec{0} = x1 = x = \vec{x}$; por tanto $\vec{0} = 1$.

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. $(ab)\overrightarrow{x} = x^{(ab)} = (x^b)^a = (b\overrightarrow{x})^a = a(b\overrightarrow{x})$.

7. $a(\overrightarrow{x} + \overrightarrow{y}) = (xy)^a = x^a \cdot y^a = a\overrightarrow{x} + a\overrightarrow{y}$.

8. $(a+b)\overrightarrow{x} = (x)^{a+b} = x^a \cdot x^b = a\overrightarrow{x} + b\overrightarrow{x}$.

□

Ejercicio 33(a) Esquema de la función lineal: $\begin{array}{rl} \_\_{}_{|j} : \mathbb{R}^{m\times n} & \to \mathbb{R}^m \\ \mathbf{A} & \to (a_{1j}, \dots a_{mj}) \end{array}.$

Es lineal puesto que

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

□

Ejercicio 33(b) Esquema de la función lineal: $\begin{array}{rl} {}_{i|}\_\_{}_{|j} : \mathbb{R}^{m\times n} & \to \mathbb{R} \\ \mathbf{A} & \to a_{ij} \end{array}.$

Es lineal puesto que

$${}_{i|}(\mathbf{A}+\mathbf{B})_{|j} = {}_{i|}\mathbf{A}_{|j} + {}_{i|}\mathbf{B}_{|j}$$
$${}_{i|}(\lambda\mathbf{A})_{|j} = \lambda\Big({}_{i|}\mathbf{A}_{|j}\Big).$$

□

Ejercicio 33(c) Esquema de la función lineal: $\begin{array}{rl} \_\_^{\intercal} : \mathbb{R}^{m\times n} & \to \mathbb{R}^{n\times m} \\ \mathbf{A} & \to [{}_{1|}\mathbf{A};\ \dots {}_{m|}\mathbf{A}] \end{array}.$

Es lineal puesto que

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

□

Ejercicio 33(d) Esquema de la función lineal: $\begin{array}{rl} {}_{i|}\_\_ : \mathbb{R}^{m\times n} & \to \mathbb{R}^n \\ \mathbf{A} & \to (a_{m1}, \dots a_{mn}) \end{array}.$

Es lineal puesto que

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

□

Ejercicio 33(e) Esquema de la función lineal: $\begin{array}{rl} \_\_\boldsymbol{b} : \mathbb{R}^{m\times n} & \to \mathbb{R}^n \\ \mathbf{A} & \to \big(\mathbf{A}_{|1}\big)b_1 + \big(\mathbf{A}_{|2}\big)b_2 + \cdots + \big(\mathbf{A}_{|n}\big)b_n \end{array}.$

Es lineal puesto que

$$(\mathbf{A}+\mathbf{C})\boldsymbol{b} = \mathbf{A}\boldsymbol{b} + \mathbf{C}\boldsymbol{b}$$
$$(\lambda\mathbf{A})\boldsymbol{b} = \lambda(\mathbf{A}\boldsymbol{b}).$$

$\square$

Ejercicio 33(f) Esquema de la función lineal: $\begin{array}{rl} \mathbf{A}\_ : \mathbb{R}^n & \to \mathbb{R}^n \\ \boldsymbol{b} & \to \big(\mathbf{A}_{|1}\big)b_1 + \big(\mathbf{A}_{|2}\big)b_2 + \cdots + \big(\mathbf{A}_{|n}\big)b_n \end{array}.$

Es lineal puesto que

$$\mathbf{A}(\boldsymbol{b}+\boldsymbol{c}) = \mathbf{A}\boldsymbol{b} + \mathbf{A}\boldsymbol{c}$$
$$\mathbf{A}(\lambda\boldsymbol{b}) = \lambda(\mathbf{A}\boldsymbol{b}).$$

$\square$

Ejercicio 33(g) Esquema de la función lineal: $\begin{array}{rl} \boldsymbol{a}\_ : \mathbb{R}^{m\times n} & \to \mathbb{R}^m \\ \mathbf{B} & \to a_1\big({}_{1|}\mathbf{B}\big) + \cdots + a_m\big({}_{m|}\mathbf{B}\big) \end{array}.$

Es lineal puesto que

$$\boldsymbol{a}(\mathbf{B}+\mathbf{C}) = \boldsymbol{a}\mathbf{B} + \boldsymbol{a}\mathbf{C}$$
$$\boldsymbol{a}(\lambda\mathbf{B}) = \lambda(\boldsymbol{a}\mathbf{B}).$$

$\square$

Ejercicio 33(h) Esquema de la función lineal: $\begin{array}{rl} \_\mathbf{B} : \mathbb{R}^m & \to \mathbb{R}^m \\ \boldsymbol{a} & \to a_1\big({}_{1|}\mathbf{B}\big) + \cdots + a_m\big({}_{m|}\mathbf{B}\big) \end{array}.$

Es lineal puesto que

$$\big(\boldsymbol{a}+\boldsymbol{c}\big)\mathbf{B} = \boldsymbol{a}\mathbf{B} + \boldsymbol{c}\mathbf{B}$$
$$(\lambda\boldsymbol{a})\mathbf{B} = \lambda(\boldsymbol{a}\mathbf{B}).$$

$\square$

Ejercicio 33(i) Esquema de la función lineal: $\begin{array}{rl} \mathbf{A}\_ : \mathbb{R}^{m\times n} & \to \mathbb{R}^{m\times p} \\ \mathbf{A} & \to [\mathbf{A}\mathbf{B}_{|1};\ \ldots \mathbf{A}\mathbf{B}_{|p}] \end{array}.$

Es lineal puesto que

$$(\mathbf{A}+\mathbf{C})\mathbf{B} = \mathbf{A}\mathbf{B} + \mathbf{C}\mathbf{B}$$
$$(\lambda\mathbf{A})\mathbf{B} = \lambda(\mathbf{A}\mathbf{B}).$$

$\square$

Ejercicio 33(j) Esquema de la función lineal: $\begin{array}{rl} \_\mathbf{B} : \mathbb{R}^{m\times n} & \to \mathbb{R}^{k\times n} \\ \mathbf{B} & \to [{}_{1|}\mathbf{A}\mathbf{B};\ \ldots {}_{k|}\mathbf{A}\mathbf{B}]^{\intercal} \end{array}.$

Es lineal puesto que

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{A}\mathbf{B} + \mathbf{A}\mathbf{C}$$
$$\mathbf{A}(\lambda\mathbf{B}) = \lambda(\mathbf{A}\mathbf{B}).$$

□

Ejercicio 33(k) Esquema de la función lineal: $\begin{array}{rcl} \_{}_{\boldsymbol{\tau}}: \mathbb{R}^{m\times n} & \to & \mathbb{R}^{m\times n} \\ \mathbf{A} & \to & \mathbf{A}\left(\mathbf{I}_{\boldsymbol{\tau}}\right) \end{array}$ donde $\mathbf{I}_{\boldsymbol{\tau}}$ es una matriz elemental del orden n. Es lineal puesto que

$$\begin{aligned} (\mathbf{A}+\mathbf{B})_{\boldsymbol{\tau}} &= \mathbf{A}_{\boldsymbol{\tau}} + \mathbf{B}_{\boldsymbol{\tau}} \\ (\lambda\mathbf{A})_{\boldsymbol{\tau}} &= \lambda(\mathbf{A}_{\boldsymbol{\tau}}). \end{aligned}$$

□

Ejercicio 33(l) Esquema de la función lineal: $\begin{array}{rcl} {}_{\boldsymbol{\tau}}\_: \mathbb{R}^{m\times n} & \to & \mathbb{R}^{m\times n} \\ \mathbf{A} & \to & \left({}_{\boldsymbol{\tau}}\mathbf{I}\right)\mathbf{A} \end{array}$ donde ${}_{\boldsymbol{\tau}}\mathbf{I}$ es una matriz elemental del orden m. Es lineal puesto que

$$\begin{aligned} {}_{\boldsymbol{\tau}}(\mathbf{A}+\mathbf{B}) &= {}_{\boldsymbol{\tau}}\mathbf{A} + {}_{\boldsymbol{\tau}}\mathbf{B} \\ {}_{\boldsymbol{\tau}}(\lambda\mathbf{A}) &= \lambda({}_{\boldsymbol{\tau}}\mathbf{A}). \end{aligned}$$

□

Ejercicio 34(a) Tenemos que 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 34(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 35. Si, lo es; 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 es cerrado para la suma y el producto por escalares.

$\square$

Ejercicio 36. 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 λ y μ. Entonces

1.

$$\begin{aligned}
(\vec{a},\vec{x})+(\vec{b},\vec{y}) &=(\vec{a}+\vec{b},\vec{x}+\vec{y}) && \text{aplicado la definición de suma de pares}\\
&=(\vec{b}+\vec{a},\vec{y}+\vec{x}) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}\\
&=(\vec{b},\vec{y})+(\vec{a},\vec{x}) && \text{aplicado la definición de suma de pares.}
\end{aligned}$$

2.

$$\begin{aligned}
(\vec{a},\vec{x})+\Big((\vec{b},\vec{y})+(\vec{c},\vec{z})\Big) &=(\vec{a},\vec{x})+\Big((\vec{b}+\vec{c}),\ (\vec{y}+\vec{z})\Big) && \text{aplicado la definición de suma de pares}\\
&=\Big(\vec{a}+(\vec{b}+\vec{c}),\ \vec{x}+(\vec{y}+\vec{z})\Big)) && \text{aplicado la definición de suma de pares}\\
&=\Big((\vec{a}+\vec{b})+\vec{c},\ (\vec{x}+\vec{y})+\vec{z}\Big) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}\\
&=\Big((\vec{a}+\vec{b}),\ (\vec{x}+\vec{y})\Big)+(\vec{c},\vec{z}) && \text{aplicado la definición de suma de pares}\\
&=\Big((\vec{a},\vec{x})+(\vec{b},\vec{y})\Big)+(\vec{c},\vec{z}).
\end{aligned}$$

3.

$$\begin{aligned}
(\vec{a},\vec{x})+(\vec{0}_{\mathcal{A}},\vec{0}_{\mathcal{X}}) &=(\vec{a}+\vec{0}_{\mathcal{A}},\ \vec{x},\vec{0}_{\mathcal{X}}) && \text{aplicado la definición de suma de funciones}\\
&=(\vec{a},\vec{x}) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}\\
& && \text{por tanto } \ \vec{0}=(\vec{0}_{\mathcal{A}},\vec{0}_{\mathcal{X}}).
\end{aligned}$$

4.

$$\begin{aligned}
(\vec{a},\vec{x})+(-\vec{a},-\vec{x}) &=(\vec{a}-\vec{a},\ \vec{x}-\vec{x}) && \text{aplicado la definición de suma de funciones}\\
&=(\vec{0}_{\mathcal{A}},\vec{0}_{\mathcal{X}}) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}\\
& && \text{por tanto } \ (-\vec{a},-\vec{x})=-(\vec{a},\vec{x}).
\end{aligned}$$

5.

$$\begin{aligned}
\lambda\big((\vec{a},\vec{x})+(\vec{b},\vec{y})\big) &=\lambda\big((\vec{a}+\vec{b}),\ (\vec{x}+\vec{y})\big) && \text{aplicado la definición de suma de pares}\\
&=\big(\lambda(\vec{a}+\vec{b}),\ \lambda(\vec{x}+\vec{y})\big) && \text{aplicado la definición de escalar por par}\\
&=(\lambda\vec{a}+\lambda\vec{b},\ \lambda\vec{x}+\lambda\vec{y}) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}\\
&=(\lambda\vec{a},\ \lambda\vec{x})+(\lambda\vec{b},\ \lambda\vec{y}) && \text{aplicado la definición de suma de pares}\\
&=\lambda(\vec{a},\vec{x})+\lambda(\vec{b},\vec{y}) && \text{aplicado la definición de escalar por par}
\end{aligned}$$

6.

$$
\begin{aligned}
(\lambda+\mu)(\vec{a},\vec{x}) =& \big((\lambda+\mu)\vec{a},\ (\lambda+\mu)\vec{x}\big) && \text{aplicado la definición de escalar por par}\\
=& \big((\lambda\vec{a}+\mu\vec{a}),\ (\lambda\vec{x}+\mu\vec{x})\big) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}\\
=& (\lambda\vec{a},\ \lambda\vec{x}) + (\mu\vec{a},\ \mu\vec{x}) && \text{aplicado la definición de suma de pares}\\
=& \lambda(\vec{a},\vec{x}) + \mu(\vec{a},\vec{x}) && \text{aplicado la definición de escalar por par}
\end{aligned}
$$

7.

$$
\begin{aligned}
(\lambda\mu)(\vec{a},\vec{x}) =& \big((\lambda\mu)\vec{a},\ (\lambda\mu)\vec{x}\big) && \text{aplicado la definición de escalar por par}\\
=& \big(\lambda(\mu\vec{a}),\ \lambda(\mu\vec{x})\big) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}\\
=& \lambda(\mu\vec{a},\ \mu\vec{x}) && \text{aplicado la definición de escalar por par}
\end{aligned}
$$

8.

$$
\begin{aligned}
1(\vec{a},\vec{x}) =& (1\vec{a}, 1\vec{x}) && \text{aplicado la definición de escalar por par}\\
=& (\vec{a},\vec{x}) && \text{puesto que } \mathcal{A} \text{ y } \mathcal{X} \text{ son espacios vectoriales}
\end{aligned}
$$

□

Ejercicio 37. 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 λ y μ. Entonces

1.

$$
\begin{aligned}
[f+g](\vec{x}) =& f(\vec{x}) + g(\vec{x}) && \text{aplicado la definición de suma de funciones}\\
=& g(\vec{x}) + f(\vec{x}) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial}\\
=& [g+f](\vec{x}) && \text{aplicado la definición de suma de funciones}
\end{aligned}
$$

2.

$$
\begin{aligned}
[f+(g+h)](\vec{x}) =& f(\vec{x}) + [g+h](\vec{x}) && \text{aplicado la definición de suma de funciones}\\
=& f(\vec{x}) + \big(g(\vec{x}) + h(\vec{x})\big) && \text{aplicado la definición de suma de funciones}\\
=& \big(f(\vec{x}) + g(\vec{x})\big) + h(\vec{x}) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial}\\
=& [f+g](\vec{x}) + h(\vec{x}) && \text{aplicado la definición de suma de funciones}\\
=& [(f+g)+h](\vec{x}) && \text{aplicado la definición de suma de funciones}
\end{aligned}
$$

3. Sea $\begin{aligned} 0\colon X &\to \mathcal{V} \\ \vec{x} &\to \vec{0} \end{aligned}$, entonces

$$
\begin{aligned}
[f+0](\vec{x}) =& f(\vec{x}) + 0(\vec{x}) && \text{aplicado la definición de suma de funciones}\\
=& f(\vec{x}) + \vec{0} && \text{pues } 0(\vec{x}) = \vec{0}\\
=& f(\vec{x})
\end{aligned}
$$

4. Sea $\begin{array}{l} -f\colon X \to \mathcal{V} \\ \quad \vec{x} \to -f(\vec{x}) \end{array}$, entonces

$$\begin{aligned}
[f+(-f)](\vec{x}) &= f(\vec{x}) + \big(-f(\vec{x})\big) && \text{aplicado la definición de suma de funciones} \\
&= \vec{0} && \text{puesto que en } \mathcal{V} \text{ se verifica que } f(\vec{x}) + \big(-f(\vec{x})\big) = \vec{0} \\
&= 0(\vec{x}) && \text{pues } 0(\vec{x}) = \vec{0}
\end{aligned}$$

5.

$$\begin{aligned}
[\lambda(f+g)](\vec{x}) &= \lambda\big([f+g](\vec{x})\big) && \text{aplicado la definición de función por escalar} \\
&= \lambda\big(f(\vec{x}) + g(\vec{x})\big) && \text{aplicado la definición de suma de funciones} \\
&= \lambda f(\vec{x}) + \lambda g(\vec{x}) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial} \\
&= [\lambda f](\vec{x}) + [\lambda g](\vec{x}) && \text{aplicado la definición de función por escalar} \\
&= [\lambda f + \lambda g](\vec{x}) && \text{aplicado la definición de suma de funciones}
\end{aligned}$$

6.

$$\begin{aligned}
[(\lambda+\mu)f](\vec{x}) &= (\lambda+\mu)f(\vec{x}) && \text{aplicado la definición de función por escalar} \\
&= \lambda f(\vec{x}) + \mu f(\vec{x}) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial} \\
&= [\lambda f](\vec{x}) + [\mu f](\vec{x}) && \text{aplicado la definición de función por escalar} \\
&= [\lambda f + \mu f](\vec{x}) && \text{aplicado la definición de suma de funciones}
\end{aligned}$$

7.

$$\begin{aligned}
[(\lambda\cdot\mu)f](\vec{x}) &= (\lambda\cdot\mu)f(\vec{x}) && \text{aplicado la definición de función por escalar} \\
&= \lambda\big(\mu f(\vec{x})\big) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial} \\
&= \lambda\big([\mu f](\vec{x})\big) && \text{aplicado la definición de función por escalar} \\
&= [\lambda(\mu f)](\vec{x}) && \text{aplicado la definición de función por escalar}
\end{aligned}$$

8.

$$\begin{aligned}
[1\cdot f](\vec{x}) &= 1\cdot f(\vec{x}) && \text{aplicado la definición de función por escalar} \\
&= f(\vec{x}) && \text{puesto que } \mathcal{V} \text{ es un espacio vectorial}
\end{aligned}$$

$\square$

Ejercicio 38(a) Tenemos que comprobar que la función $[g+f]\colon \mathcal{D} \to \mathcal{W}$ 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

$$[g+f](\alpha\vec{x}) = g(\alpha\vec{x}) + f(\alpha\vec{x}) = \alpha g(\vec{x}) + \alpha f(\vec{x}) = \alpha\big(g(\vec{x}) + f(\vec{x})\big) = \alpha[g+f](\vec{x}).$$

$\square$

Ejercicio 38(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

$$[\alpha \cdot f](\gamma\vec{x}) = \alpha f(\gamma\vec{x}) = \alpha\big(\gamma f(\vec{x})\big) = \gamma\big(\alpha f(\vec{x})\big) = \gamma[\alpha \cdot f](\vec{x}).$$

$\square$

Ejercicio 39. 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 $\mathcal{N}(\mathbf{A})$, es decir, $\mathbf{A}\boldsymbol{x} = \mathbf{0}$ y $\mathbf{A}\boldsymbol{y} = \mathbf{0}$, y sean b y c dos números reales; entonces

$$\mathbf{A}\big(b\boldsymbol{x} + c\boldsymbol{x}\big) = \mathbf{A}b\boldsymbol{x} + \mathbf{A}c\boldsymbol{y} = b\mathbf{A}\boldsymbol{x} + c\mathbf{A}\boldsymbol{y} = b\mathbf{0} + c\mathbf{0} = \mathbf{0} \quad \Rightarrow \quad \big(b\boldsymbol{x} + c\boldsymbol{x}\big) \in \mathcal{N}(\mathbf{A}).$$

$\square$

Ejercicio 40. Si son nulos los coeficientes correspondientes a las columnas no nulas, entonces

$$\mathbf{K}\boldsymbol{x} \;=\; \big(\mathbf{K}_{|1}\big)x_1 + \big(\mathbf{K}_{|2}\big)x_2 + \cdots + \big(\mathbf{K}_{|n}\big)x_n \;=\; \mathbf{0},$$

pues en cada producto $\big(\mathbf{K}_{|j}\big)x_j$ es $\mathbf{0}$, ya que o bien $x_j = 0$ o bien $\big(\mathbf{K}_{|j}\big) = \mathbf{0}$.

Recíprocamente, supongamos que algún coeficiente x_k es distinto de cero para una columna no nula, y tomemos $j = \min\Big\{k \mid x_k \neq 0 \text{ y } \big(\mathbf{K}_{|j}\big) \neq \mathbf{0}\Big\}$. Por la elección de j, si $h < j$ entonces el producto $\big(\mathbf{K}_{|h}\big)x_h = \mathbf{0}$ (ya que alguno de los dos es cero). Por tanto tendremos que $\big(\mathbf{K}_{|1}\big)x_1 + \cdots + \big(\mathbf{K}_{|n}\big)x_n = \big(\mathbf{K}_{|j}\big)x_j + \cdots + \big(\mathbf{K}_{|n}\big)x_n$. Entonces si i es la posición de pivote de $\mathbf{K}_{|j}$ tendremos que ${}_{i|}\Big(\big(\mathbf{K}_{|j}\big)x_j + \cdots + \big(\mathbf{K}_{|n}\big)x_n\Big) = \big({}_{i|}\mathbf{K}_{|j}\big)x_j \neq 0$; pues todas las componentes a la derecha del pivote son nulas por estar $\mathbf{K}$ pre-escalonada. Y por tanto $\mathbf{K}\boldsymbol{x} \neq \mathbf{0}$.

$\square$

Ejercicio 41. Basta comprobar que cualquier combinación de vectores de $\mathcal{N}(f) = \big\{\vec{v} \,\big|\, f(\vec{v}) = \vec{0}\big\}$ pertenece a $\mathcal{N}(f)$.

Sean $\vec{x}$ y $\vec{y}$ dos vectores de $\mathcal{N}(f)$, es decir, $f(\vec{x}) = \vec{0}$ y $f(\vec{y}) = \vec{0}$, y sean b y c dos escalares; entonces

$$f\big(b\vec{x} + c\vec{x}\big) = f(\vec{x}) + f(c\vec{y}) = b \cdot f(\vec{x}) + c \cdot f(\vec{x}) = b\vec{0} + c\vec{0} = \vec{0} \quad \Rightarrow \quad \big(b\vec{x} + c\vec{x}\big) \in \mathcal{N}(f).$$

(Compare esta demostración con la de la Proposición 7.2.1).

$\square$

Ejercicio 42. 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 83). Así pues, la imagen de $f_{\mathbf{A}}$ es

$$\begin{aligned} \mathit{imagen}(f_{\mathbf{A}}) &= \{\boldsymbol{b} \mid \text{existe } \boldsymbol{x} \text{ tal que } (\boldsymbol{x}, \boldsymbol{b}) \in f_{\mathbf{A}}\} \\ &= \{\boldsymbol{b} \mid \text{existe } \boldsymbol{x} \in \mathbb{R}^n \text{ tal que } \boldsymbol{b} = f_{\mathbf{A}}(\boldsymbol{x}) = \mathbf{A}\boldsymbol{x}\} \\ &= \mathcal{C}(\mathbf{A}). \end{aligned}$$

$\square$

Ejercicio 43. Basta demostrar que el conjunto

$$imagen(f) = \{\vec{y} \mid \text{existe } \vec{x} \in \mathcal{D} \text{ tal que } \vec{y} = f(\vec{x})\}$$

es cerrado para la suma y el producto por escalares.

Sean $\vec{x}, \vec{y} \in imagen(f)$, es decir, vectores de $\mathcal{W}$ tales que existen $\vec{z}$ y $\vec{w}$ de $\mathcal{D}$ de manera que $\vec{x} = f(\vec{z})$ e $\vec{y} = f(\vec{w})$, y sean los escalares a y b. Entonces para el vector $(a\vec{x} + b\vec{y}) \in \mathcal{W}$ se verifica que

$$a\vec{x} + b\vec{y} \;=\; af(\vec{z}) + bf(\vec{w}) \;=\; f(a\vec{z} + b\vec{w});$$

pues f es lineal. Es decir, $(a\vec{x} + b\vec{y}) \in imagen(f)$.

$\square$

Ejercicio 44(a)

$$\mathsf{Z}(\boldsymbol{b}+\boldsymbol{c}) = \sum_{k=1}^{n}\left(\mathsf{Z}_{|k}(b_k+c_k)\right) = \sum_{k=1}^{n}\left((\mathsf{Z}_{|k})b_k \;+\; (\mathsf{Z}_{|k})c_k\right) = \sum_{k=1}^{n}(\mathsf{Z}_{|k})b_k \;+\; \sum_{k=1}^{n}(\mathsf{Z}_{|k})c_k = \mathsf{Z}\boldsymbol{b} + \mathsf{Z}\boldsymbol{c}.$$

$\square$

Ejercicio 44(b)

$$\mathsf{Z}(\lambda\boldsymbol{b}) = \sum_{k=1}^{n}(\mathsf{Z}_{|k})\lambda b_k = \lambda\sum_{k=1}^{n}(\mathsf{Z}_{|k})b_k = \lambda(\mathsf{Z}\boldsymbol{b}).$$

$\square$

Ejercicio 45(a)

$$\mathsf{Z}(\mathbf{B}\boldsymbol{c}) = \mathsf{Z}\left(\sum_{k=1}^{n}(\mathbf{B}_{|k})c_k\right) = \sum_{k=1}^{n}\left(\mathsf{Z}(\mathbf{B}_{|k})c_k\right) = \sum_{k=1}^{n}\left((\mathsf{Z}\mathbf{B})_{|k}\right)c_k = \left[\mathsf{Z}\mathbf{B}_{|1};\ldots;\mathsf{Z}\mathbf{B}_{|n}\right]\boldsymbol{c} = (\mathsf{Z}\mathbf{B})\boldsymbol{c}.$$

$\square$

Ejercicio 45(b) Puesto 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, tenemos:

$$\left(\mathsf{Z}(\mathbf{BC})\right)_{|j} = \mathsf{Z}(\mathbf{BC})_{|j} = \mathsf{Z}\left(\mathbf{B}(\mathbf{C}_{|j})\right) \underset{\text{(Prop. 9.1.3)}}{=} \left(\mathsf{Z}\mathbf{B}\right)(\mathbf{C}_{|j}) = \left((\mathsf{Z}\mathbf{B})\mathbf{C}\right)_{|j}, \qquad \forall j = \{1,\ldots,n\}.$$

$\square$

Ejercicio 46. Por una parte, si z_i es combinación lineal del resto de vectores de Z, entonces

$$\begin{aligned}
\vec{z}_i &= a_1\vec{z}_1 + \cdots + a_{i-1}\vec{z}_{i-1} \phantom{\;+(-1)\vec{z}_i} + a_{i+1}\vec{z}_{i+1} + \cdots + a_n\vec{z}_n \qquad \Rightarrow \\
\vec{0} &= a_1\vec{z}_1 + \cdots + a_{i-1}\vec{z}_{i-1} + (-1)\vec{z}_i + a_{i+1}\vec{z}_{i+1} + \cdots + a_n\vec{z}_n \qquad \Rightarrow \\
\vec{0} &= \mathsf{Z}\boldsymbol{a} \quad (\text{con } a_i = -1).
\end{aligned}$$

Por otra parte, si $\vec{0} = \mathsf{Z}\boldsymbol{a}$, con $\boldsymbol{a} \neq \mathbf{0}$, entonces existe un $i \in 1,\ldots,n$ tal que $a_i \neq 0$, y por lo tanto

$$\begin{aligned}
\vec{0} &= a_1\vec{z}_1 + \cdots + a_{i-1}\vec{z}_{i-1} + a_i\vec{z}_i + a_{i+1}\vec{z}_{i+1} + \cdots + a_n\vec{z}_n \qquad \Rightarrow \\
-a_i\vec{z}_i &= a_1\vec{z}_1 + \cdots + a_{i-1}\vec{z}_{i-1} \phantom{\;+a_i\vec{z}_i} + a_{i+1}\vec{z}_{i+1} + \cdots + a_n\vec{z}_n \qquad \Rightarrow \\
\vec{z}_i &= \frac{-1}{a_i}\left(a_1\vec{z}_1 + \cdots + a_{i-1}\vec{z}_{i-1} + a_{i+1}\vec{z}_{i+1} + \cdots + a_n\vec{z}_n\right).
\end{aligned}$$

$\square$

Ejercicio 47. Basta expresar la imagen de f_{Z} con la notación funcional (6.5) y comprobar que coincide con la notación matricial de la Definición 9.1 de espacio engendrado por Z (9.1):

$$imagen(f_{\mathsf{Z}}) = \left\{ \vec{v} \in \mathcal{V} \;\middle|\; \text{existe } \boldsymbol{a} \in \mathbb{R}^n \text{ tal que } \vec{v} = \mathsf{Z}\boldsymbol{a} \right\} = \mathcal{L}(\mathsf{Z})$$

$\square$

Ejercicio 48. 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 $\vec{z} \in imagen(g)$.

$\square$

Ejercicio 49. Por una parte sabemos que $imagen(g \circ f) \subset imagen(g)$.

Por otra, puesto que $imagen(g) = imagen(g \circ f \circ f^{-1})$, aplicando la proposición anterior tenemos que

$$imagen(g) = imagen\big([g \circ f] \circ f^{-1}\big) \subset imagen(g \circ f).$$

$\square$

Ejercicio 50. La función $f \circ f_{\mathsf{B}}$ es lineal y va de $\mathbb{R}^n$ a $\mathcal{W}$. Por tanto existe un sistema Z de vectores de $\mathcal{W}$ tal que $f \circ f_{\mathsf{B}} = f_{\mathsf{Z}}$. Por tanto

$$f(\vec{v}) = f\Big(\mathsf{B}\big(\vec{v}_{/\mathsf{B}}\big)\Big) = f\Big(f_{\mathsf{B}}\big(\vec{v}_{/\mathsf{B}}\big)\Big) = f \circ f_{\mathsf{B}}\big(\vec{v}_{/\mathsf{B}}\big) = f_{\mathsf{Z}}\big(\vec{v}_{/\mathsf{B}}\big) = \mathsf{Z}\big(\vec{v}_{/\mathsf{B}}\big).$$

$\square$

Ejercicio 51.

Librería 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)
```

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

De donde deducimos que las últimas dos componentes de la única solución especial nos dicen como combinar los vectores del sistema `B` (es decir, que el primer vector de dicho sistema es una base de la intersección. Comprobación:

Librería NAcAL para Python

```
SubEspacio(A) & SubEspacio(B) # SubEspacio intersección
```

$$\left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \exists \boldsymbol{p} \in \mathbb{R}^1 \text{ tal que } \boldsymbol{v} = \begin{bmatrix} 1\\1\\1\\1 \end{bmatrix} \boldsymbol{p} \right\} = \left\{ \boldsymbol{v} \in \mathbb{R}^4 \;\middle|\; \begin{bmatrix} -1 & 1 & 0 & 0\\ -1 & 0 & 1 & 0\\ -1 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{v} = \boldsymbol{0} \right\}$$

□

Ejercicio 52. La matriz L es de la forma

$$\mathsf{L} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0\\ *_1 & 0 & \cdots & 0 & 0 & \cdots & 0\\ : & 0 & \cdots & 0 & 0 & \cdots & 0\\ : & *_2 & \cdots & 0 & 0 & \cdots & 0\\ : & : & & 0 & 0 & \cdots & 0\\ : & : & \cdots & *_r & 0 & \cdots & 0\\ : & : & \cdots & : & 0 & \cdots & 0 \end{bmatrix}$$

y puesto que el último pivote $*_r \neq 0$, necesariamente la última fila con pivote tiene que estar multiplicada por cero ($x_r = 0$); pero entonces, como el penúltimo pivote $*_{r-1} \neq 0$, necesariamente la penúltima fila con pivote debe estar multiplicada por cero ($x_{r-1} = 0$); pero entonces, como $*_{r-2} \neq 0$, necesariamente ...y razonando sucesivamente, es evidente que $x_i = 0$ para $i = 1, \ldots, r$.

Para ver que cada fila sin pivote es combinación lineal de las filas con pivote que la preceden, basta aplicar la eliminación sobre las filas de L para anular todas las filas sin pivote[4].

□

Ejercicio 53. 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 A es cero:

$$\det\left[\mathsf{A}_{|1}\ldots;\mathbf{0};\ldots;\mathsf{A}_{|n}\right] = \det\left[\mathsf{A}_{|1}\ldots;0\boldsymbol{x};\ldots;\mathsf{A}_{|n}\right] = 0 \cdot \det\left[\mathsf{A}_{|1}\ldots;\boldsymbol{x};\ldots;\mathsf{A}_{|n}\right] = 0$$

□

Ejercicio 54(a) Puesto que $\mathsf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k} = \mathsf{A}\left(\mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \mathsf{A}\left(\mathsf{I}_{\boldsymbol{\tau}_1}\right)\cdots\left(\mathsf{I}_{\boldsymbol{\tau}_k}\right)$, aplicando repetidamente (14.2) tenemos

$$\begin{aligned} \det\left(\mathsf{A}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) &= \left|\mathsf{A}\left(\mathsf{I}_{\boldsymbol{\tau}_1}\right)\cdots\left(\mathsf{I}_{\boldsymbol{\tau}_k}\right)\right| \\ &= \left|\mathsf{A}\left(\mathsf{I}_{\boldsymbol{\tau}_1}\right)\cdots\left(\mathsf{I}_{\boldsymbol{\tau}_{(k-1)}}\right)\right| \cdot \left|\mathsf{I}_{\boldsymbol{\tau}_k}\right| \\ &= \left|\mathsf{A}\left(\mathsf{I}_{\boldsymbol{\tau}_1}\right)\cdots\left(\mathsf{I}_{\boldsymbol{\tau}_{(k-2)}}\right)\right| \cdot \left|\mathsf{I}_{\boldsymbol{\tau}_{(k-1)}}\right| \cdot \left|\mathsf{I}_{\boldsymbol{\tau}_k}\right| \\ &\vdots \\ &= \left|\mathsf{A}\right| \cdot \left|\mathsf{I}_{\boldsymbol{\tau}_1}\right|\cdots\left|\mathsf{I}_{\boldsymbol{\tau}_k}\right|. \end{aligned}$$

□

Ejercicio 54(b)

$$\left|\mathsf{B}\right| = \det\left(\mathsf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}\right) = \left|\mathsf{I}_{\boldsymbol{\tau}_1}\right|\cdots\left|\mathsf{I}_{\boldsymbol{\tau}_k}\right|.$$

[4] aplicar la eliminación con las filas es equivalente a aplicar transformaciones elementales *"de izquierda a derecha"* sobre el sistema de vectores $\left[{}_{1|}\mathsf{L}, \ldots, {}_{m|}\mathsf{L}\right]$.

y como los determinantes de las matrices elementales son distintos de cero, necesariamente $|\mathbf{B}| \neq 0$.

□

Ejercicio 54(c) Si $\mathbf{B}$ es de rango completo entonces es producto de k matrices elementales; por tanto $\mathbf{B} = \mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}$. Así

$$\det(\mathbf{AB}) = \det\left(\mathbf{A}\mathbf{I}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}\right) = \det\left(\mathbf{A}_{\boldsymbol{\tau}_1 \cdots \boldsymbol{\tau}_k}\right) = |\mathbf{A}| \cdot \left(\left|\mathbf{I}_{\boldsymbol{\tau}_1}\right| \cdots \left|\mathbf{I}_{\boldsymbol{\tau}_k}\right| \right) = |\mathbf{A}| \cdot |\mathbf{B}|\,.$$

□

Ejercicio 55(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 55(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

$$\left|\mathbf{B}^{\intercal}\right| \;=\; \prod_{i=1}^{k} \det\left({}_{\boldsymbol{\tau}_i}\mathbf{I}\right) \;=\; \prod_{i=1}^{k} \det\left(\mathbf{I}_{\boldsymbol{\tau}_i}\right) \;=\; |\mathbf{B}|\,.$$

□

Ejercicio 56(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).

$$\mathbf{L} = \begin{bmatrix} *_1 & & & \\ : & *_2 & & \\ : & : & \ddots & \\ : & : & : & *_n \end{bmatrix} \qquad \text{donde } *_j \text{ son números distintos de cero.}$$

Dividiendo cada columna $j$ésima por su pivote $*_j$ para normalizar los pivotes (y compensando dichas transformaciones multiplicado la última fila por cada pivote); y aplicando, en una segunda fase, la eliminación de izquierda a derecha con transformaciones de Tipo I para anular todo lo que queda a la izquierda de los pivotes (ahora basta multiplicar la última fila por 1), llegamos a:

$$\left[\begin{array}{cccc|c} *_1 & & & & \\ : & *_2 & & & \\ : & : & \ddots & & \\ : & : & : & *_n & \\ \hline & & & & 1 \end{array}\right] \xrightarrow[\substack{\boldsymbol{\tau} \\ [(*_1)(\mathbf{n+1})] \\ \vdots \\ [(*_n)(\mathbf{n+1})]}]{\substack{\boldsymbol{\tau} \\ \left[\left(\frac{1}{*_1}\right)\mathbf{1}\right] \\ \vdots \\ \left[\left(\frac{1}{*_n}\right)\mathbf{n}\right]}} \left[\begin{array}{cccc|c} 1 & & & & \\ : & 1 & & & \\ : & : & \ddots & & \\ : & : & : & 1 & \\ \hline & & & & *_1 \cdot *_2 \cdots *_n \end{array}\right] \xrightarrow[\substack{\boldsymbol{\tau} \\ [(1)(\mathbf{n+1})]}]{\substack{\boldsymbol{\tau}_1 \ldots \boldsymbol{\tau}_q \\ \text{(de Tipo I)}}} \left[\begin{array}{cccc|c} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ \hline & & & & *_1 \cdot *_2 \cdots *_n \end{array}\right]$$

por tanto, si la matriz es triangular inferior es de rango completo, su determinante es igual al producto de sus pivotes; es decir, al producto de los elementos de la diagonal.

$$\det(\mathbf{L}) = \text{ producto de los elementos de la diagonal}$$

$\square$

Ejercicio 56(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. Por tanto, si la matriz tringular es singular, necesariamente tiene algún cero en su diagonal principal. Como su determinante es cero, por ser singular, su determinante es igual al producto de los elementos de la diagonal (donde uno de ellos es cero).

$\square$

Ejercicio 56(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.

$\square$

Ejercicio 57(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 $\begin{vmatrix}\mathbf{A} & \\ & \mathbf{B}\end{vmatrix} = \begin{vmatrix}\mathbf{A} & \\ & \mathbf{I}\end{vmatrix} \cdot \begin{vmatrix}\mathbf{I} & \\ & \mathbf{B}\end{vmatrix} = \left|\mathbf{A}\right| \cdot \left|\mathbf{B}\right|.$

$\square$

Ejercicio 57(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 $\boldsymbol{\tau}_1, \dots \boldsymbol{\tau}_k$ de transformaciones elementales *Tipo I* es posible la transformación $\begin{bmatrix}\mathbf{A} & \\ \mathbf{C} & \mathbf{I}\end{bmatrix} \xrightarrow{\boldsymbol{\tau}_1,\dots\boldsymbol{\tau}_k} \begin{bmatrix}\mathbf{A} & \\ & \mathbf{I}\end{bmatrix}$ tenemos que $\begin{vmatrix}\mathbf{A} & \\ \mathbf{C} & \mathbf{I}\end{vmatrix} = \begin{vmatrix}\mathbf{A} & \\ & \mathbf{I}\end{vmatrix}$; y por tanto

$$\begin{vmatrix}\mathbf{A} & \\ \mathbf{C} & \mathbf{B}\end{vmatrix} = \begin{vmatrix}\mathbf{A} & \\ \mathbf{C} & \mathbf{I}\end{vmatrix}\begin{vmatrix}\mathbf{I} & \\ & \mathbf{B}\end{vmatrix} = \begin{vmatrix}\mathbf{A} & \\ & \mathbf{I}\end{vmatrix}\begin{vmatrix}\mathbf{I} & \\ & \mathbf{B}\end{vmatrix} = \left|\mathbf{A}\right| \cdot \left|\mathbf{B}\right|.$$

$\square$

Ejercicio 58. Desarrollando por la segunda columna tenemos

$$\det \mathbf{A} \;=\; -0\begin{vmatrix}5&2&4\\3&1&2\\5&2&1\end{vmatrix} + \begin{vmatrix}2&3&2\\3&1&2\\5&2&1\end{vmatrix} - 0\begin{vmatrix}2&3&2\\5&2&4\\5&2&1\end{vmatrix} + 3\begin{vmatrix}2&3&2\\5&2&4\\3&1&2\end{vmatrix} \;=\; 0 + 17 - 0 + 3 \times 4 \;=\; 29$$

$\square$

Ejercicio 59.

$$\begin{aligned}
\begin{vmatrix}a&b&c\\d&e&f\\g&h&i\end{vmatrix} &= -\,b\begin{vmatrix}d&f\\g&i\end{vmatrix} + e\begin{vmatrix}a&c\\g&i\end{vmatrix} - h\begin{vmatrix}a&c\\d&f\end{vmatrix}\\
&= -\,b(-fg+id) + e(-cg+ia) - h(-cd+fa)\\
&= bfg - bdi - ecg + eia + hcd - hfa && \text{desarrollando términos}\\
&= a(ei-hf) - d(bi-hc) + g(bf-ec) && \text{reordenando}\\
&= aei + bfg + dhc - gec - dbi - hfa && \text{reordenando.}
\end{aligned}$$

(esta fórmula se denomina regla de Sarrus)

$\square$

Ejercicio 60. Sea λ un autovalor de A y sean $\boldsymbol{x}$ y $\boldsymbol{y}$ dos autovectores correspondientes al autovalor λ, es decir, tales que $\mathsf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$ y $\mathsf{A}\boldsymbol{y} = \lambda\boldsymbol{y}$; entonces

$$\mathsf{A}\big(a\boldsymbol{x} + b\boldsymbol{y}\big) = a\mathsf{A}\boldsymbol{x} + b\mathsf{A}\boldsymbol{y} = \lambda a\boldsymbol{x} + \lambda b\boldsymbol{y} = \lambda\big(a\boldsymbol{x} + b\boldsymbol{y}\big).$$

□

Ejercicio 61. 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 62. $${}_{i|}\mathsf{Q}^{\intercal}\mathsf{Q}_{|j} = \big(\mathsf{Q}_{|i}\big)\cdot\big(\mathsf{Q}_{|j}\big) = \begin{cases}1 & \text{cuando } i=j\\ 0 & \text{cuando } i\neq j\end{cases}, \quad \text{es decir} \quad \mathsf{Q}^{\intercal}\mathsf{Q} = \begin{bmatrix}1 & 0 & \dots & 0\\ 0 & 1 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 1\end{bmatrix} = \underset{n\times n}{\mathsf{I}}.$$

□

Ejercicio 63(a) $$\Big(\mathsf{QR}\Big)^{\intercal} = \mathsf{R}^{\intercal}\big(\mathsf{Q}^{\intercal}\big) = \mathsf{R}^{-1}\left(\mathsf{Q}^{-1}\right) = (\mathsf{QR})^{-1}.$$

□

Ejercicio 63(b) $$\Big(\mathsf{Q}^{-1}\mathsf{A}\mathsf{Q}\Big)^{\intercal} = \mathsf{Q}^{\intercal}\mathsf{A}\left(\mathsf{Q}^{-1}\right)^{\intercal} = \mathsf{Q}^{\intercal}\mathsf{A}\big(\mathsf{Q}^{\intercal}\big)^{\intercal} = \mathsf{Q}^{-1}\mathsf{A}\mathsf{Q}.$$

□

Ejercicio 64. Si $\mathsf{Q}^{-1}\mathsf{A}\mathsf{Q} = \mathsf{D}$ es diagonal, entonces $\mathsf{A} = \mathsf{Q}\mathsf{D}\mathsf{Q}^{\intercal}$. Por tanto, $\mathsf{A}^{\intercal} = \Big(\mathsf{Q}\mathsf{D}\mathsf{Q}^{\intercal}\Big)^{\intercal} = \mathsf{Q}\mathsf{D}\mathsf{Q}^{\intercal} = \mathsf{A}$.

□

Ejercicio 65(a) $\boldsymbol{x}(\mathsf{A}+\mathsf{B})\boldsymbol{x} = (\boldsymbol{x}\mathsf{A}+\boldsymbol{x}\mathsf{B})\boldsymbol{x} = \boldsymbol{x}\mathsf{A}\boldsymbol{x}+\boldsymbol{x}\mathsf{B}\boldsymbol{x} > 0$; por ser la suma de dos números positivos.

□

Ejercicio 65(b) A es invertible y, como es simétrica, es ortogonalmente diagonalizable $\mathsf{A} = \mathsf{Q}\mathsf{D}\mathsf{Q}^{-1} = \mathsf{Q}\mathsf{D}\mathsf{Q}^{\intercal}$; así

$$\mathsf{A}^{-1} = \big(\mathsf{Q}\mathsf{D}\mathsf{Q}^{-1}\big)^{-1} = \Big(\mathsf{Q}^{-1}\Big)^{-1}\mathsf{D}^{-1}\mathsf{Q}^{-1} = \mathsf{Q}\mathsf{D}^{-1}\mathsf{Q}^{\intercal};$$

y como los autovalores de A son positivos, sus inversos (los autovalores de A^{-1} que se encuentran en la diagonal de D^{-1}) también lo son.

□

Ejercicio 66. Si A es definida positiva, entonces para cualquier $\boldsymbol{x}\in\mathbb{R}^n$ no nulo se verifica que $\boldsymbol{x}\mathsf{A}\boldsymbol{x} > 0$. En particular, si $\boldsymbol{x}$ es la columna $j$ésima de la matriz identidad, tenemos

$${}_{j|}\mathsf{IAI}_{|j} = {}_{j|}\mathsf{A}_{|j} > 0.$$

La demostración es análoga para el caso semidefinido.

□

Bibliografía

Arvesú Carballo, J., Marcellán Español, F., and Sánchez Ruiz, J. (2005). *Problemas resueltos de Álgebra Lineal*. Thomson Learning, Madrid. España. ISBN 84-9732-284-3.

Cullen, C. G. (1972). *Matrices and Linear Transformations*. Dover publications, Inc., New York, USA., second ed.

Larson, R., Edwars, B. H., and Falvo, D. C. (2004). *Álgebra lineal*. Ediciones Pirámide, Madrid. España, fifth 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., third 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, third ed. ISBN 0-9614088-9-8.

Strang, G. (2007). *Álgebra Lineal y sus Aplicaciones*. Thomsom Learning, Inc, Santa Fe, México, D. F., fourth ed. ISBN 970686609-4.

Glosario

C

combinación lineal de los vectores de Z ($\mathsf{Z}\boldsymbol{a}$).
concatenación de los sistemas Y **y** Z ($\mathsf{Y} \# \mathsf{Z}$).
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 de n **números reales** $\mathbb{R}^n$.

E

espacio columna de $\mathbf{A}$ ($\mathcal{C}(\mathbf{A})$).
espacio nulo de $\mathbf{A}$ ($\mathcal{N}(\mathbf{A})$).
espacios vectoriales ($\mathcal{C}$, $\mathcal{N}$, $\mathcal{X}$, $\mathcal{V}$, $\mathcal{W}$).

M

matrices particionadas ($\boldsymbol{\mathcal{A}}$, $\boldsymbol{\mathcal{B}}$, $\boldsymbol{\mathcal{C}}$).
matrices ($\mathbb{R}^{m \times n}$) ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$).
- **elemental** ($\mathbf{I}_{\boldsymbol{\tau}}$).
  - **Tipo I** ($\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}$).
  - **Tipo II** ($\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}$).
- **escalonada reducida (por columnas)** ($\mathbf{R}$).
- **identidad** ($\mathbf{I}$).
- **intercambio** ($\mathbf{I}_{\underset{[\boldsymbol{p}\rightleftharpoons\boldsymbol{q}]}{\boldsymbol{\tau}}}$).
- **invertible, de rango completo, o no singular** ($\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$).
- **inversa (de** $\mathbf{A}$**)** ($\mathbf{A}^{-1}$).
- **nula** ($\mathbf{0}$).
- **permutación** ($\mathbf{I}_{\underset{[\circlearrowleft]}{\boldsymbol{\tau}}}$).
- **pre-escalonada** ($\mathbf{K}$).
- **rango de** $\mathbf{A}$ ($\text{rango}(\mathbf{A})$).
- **transpuesta** ($\mathbf{A}^{\intercal}$; transpuesta de $\mathbf{A}$).
- **triangular inferior** ($\mathbf{L}$).
- **triangular superior** ($\mathbf{U}$).

O

operador selector de componentes ($|$).
- **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 transposición ($\intercal$).

P

producto punto (producto escalar usual en $\mathbb{R}^n$) ($\boldsymbol{a} \cdot \boldsymbol{b}$).

S

Sistema de ecuaciones lineales ($\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$).
sistemas de vectores de un subespacio genérico $\mathcal{V}$ (A, B,... X, Y, Z).
subespacio engendrado por un sistema Z ($\mathcal{L}(\mathsf{Z})$).
suma directa de los subespacios $\mathcal{A}$ y $\mathcal{B}$ ($\mathcal{A} \oplus \mathcal{B}$).

T

Transformación ($\boldsymbol{\tau}$).
  de las columnas de $\mathbf{A}$ ($\mathbf{A}_{\boldsymbol{\tau}}$).
  elemental Tipo I; suma λ veces el vector $\boldsymbol{i}$ésimo al vector $\boldsymbol{j}$ésimo $\left(\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}\right)$.
  elemental Tipo II; multiplica por α el vector $\boldsymbol{i}$ésimo vector $\left(\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}\right)$.
  intercambio entre los vectores $\boldsymbol{i}$ésimo y $\boldsymbol{j}$ésimo $\left(\underset{[\boldsymbol{i} \rightleftharpoons \boldsymbol{j}]}{\boldsymbol{\tau}}\right)$.
  inversa de $\boldsymbol{\tau}$ ($\boldsymbol{\tau}^{-1}$).
  permutación de vectores $\left(\underset{[\circlearrowleft]}{\boldsymbol{\tau}}\right)$.

V

vectores de $\mathbb{R}^n$ ($\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$).
  columna $j$ésima de $\mathbf{A}$ ($\mathbf{A}_{|j}$).
  combinación lineal de las columas de $\mathbf{A}$ ($\mathbf{A}\boldsymbol{b}$).
  combinación lineal de las filas de $\mathbf{A}$ ($\boldsymbol{a}\mathbf{B}$).
  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

conjuntos

$\{X,\ Y,\ X,\dots\}$ conjunto.
$\mathbb{C}$ conjunto de numeros complejos.
$\mathbb{R}$ conjunto de numeros reales.
$\mathbb{R}^{m\times n}$ conjunto de matrices de orden $m\times n$.
$\mathbb{R}^{n}$ conjunto de sistemas de n números reales.

espacios y subespacios vectoriales

$\mathcal{C}\left(\mathbf{A}\right)$ espacio columna de $\mathbf{A}$.
$\mathcal{C}$, $\mathcal{N}$, $\mathcal{X}$, $\mathcal{V}$, $\mathcal{W}$ espacios vectoriales.
$\mathcal{N}\left(\mathbf{A}\right)$ espacio nulo de $\mathbf{A}$.
$\mathcal{L}\big(\mathsf{Z}\big)$ subespacio engendrado por un sistema Z.
$\mathcal{A}\oplus\mathcal{B}$ suma directa de los subespacios $\mathcal{A}$ y $\mathcal{B}$.

matrices

$\begin{bmatrix}\boldsymbol{a} & \boldsymbol{b} & \boldsymbol{c}\dots\end{bmatrix}$ con $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}\in\mathbb{R}^m$; matriz de $\mathbb{R}^{m\times n}$.
$\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ matrices ($\mathbb{R}^{m\times n}$).
$\mathbf{R}$ matriz escalonada reducida (por columnas).
$\mathbf{A}^{-1}$ matriz inversa (de $\mathbf{A}$).
$\mathbf{I}_{\boldsymbol{\tau}}$ matriz elemental.
$\mathbf{I}_{\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}}$ matriz elemental Tipo I.
$\mathbf{I}_{\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}}$ matriz elemental Tipo II.
$\mathbf{I}$ matriz identidad.
$\mathbf{I}_{\underset{[\boldsymbol{p}\rightleftharpoons\boldsymbol{q}]}{\boldsymbol{\tau}}}$ matriz intercambio.
$\mathbf{I}_{\boldsymbol{\tau}_1\cdots\boldsymbol{\tau}_k}$ matriz invertible, de rango completo, o no singular.
$\mathbf{0}$ matriz nula.
$\mathbf{I}_{\underset{[\circlearrowleft]}{\boldsymbol{\tau}}}$ matriz permutación.
$\mathbf{A}^{\mathsf{T}}$ transpuesta de $\mathbf{A}$.
$\mathbf{K}$ matriz pre-escalonada.
$\text{rango}\left(\mathbf{A}\right)$ rango de $\mathbf{A}$.
$\mathbf{L}$ matriz triangular inferior.
$\mathbf{U}$ matriz triangular superior.

matrices particionadas

$\boldsymbol{\mathcal{A}}$, $\boldsymbol{\mathcal{B}}$, $\boldsymbol{\mathcal{C}}$ matrices particionadas.

operaciones

$\boldsymbol{a} \cdot \boldsymbol{b}$ producto punto (producto escalar usual en $\mathbb{R}^n$).

operadores

T operador transposición.
$|$ operador selector de componentes.
$i|$ por la izquierda.
$|i$ por la derecha.

Sistema de ecuaciones lineales

$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$ Sistema de ecuaciones lineales.

sistema de vectores de un espacio vectorial generico

$\mathsf{Y} +\!\!\!+ \mathsf{Z}$ concatenación de los sistemas Y y Z.
A, B,... X, Y, Z sistemas de vectores de un subespacio genérico $\mathcal{V}$.

sistemas genéricos

$\left[X;\ Y;\ Z;\ldots\right]$ sistema.

trasnformaciones elementales (o sucesiones de transformaciones elementales)

$\boldsymbol{\tau}$ Transformación.
$\underset{[\boldsymbol{i}\rightleftharpoons\boldsymbol{j}]}{\boldsymbol{\tau}}$ intercambio entre los vectores $\boldsymbol{i}$ésimo y $\boldsymbol{j}$ésimo.
$\underset{[\circlearrowleft]}{\boldsymbol{\tau}}$ permutación de vectores.
$\mathbf{A}_{\boldsymbol{\tau}}$ de las columnas de $\mathbf{A}$.
$\underset{[(\alpha)\boldsymbol{i}]}{\boldsymbol{\tau}}$ elemental Tipo II; multiplica por α el vector $\boldsymbol{i}$ésimo vector.
$\underset{[(\lambda)\boldsymbol{i}+\boldsymbol{j}]}{\boldsymbol{\tau}}$ elemental Tipo I; suma λ veces el vector $\boldsymbol{i}$ésimo al vector $\boldsymbol{j}$ésimo.
$\boldsymbol{\tau}^{-1}$ transformación inversa de $\boldsymbol{\tau}$.

vectores

$\mathsf{Z}\boldsymbol{a}$ combinación lineal de los vectores de Z.
$\vec{x}$, $\vec{y}$, $\vec{z}$ vectores de un espacio vectorial genérico $\mathcal{V}$.
$\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$ vectores de $\mathbb{R}^n$.
$\mathbf{A}_{|j}$ columna $j$ésima de $\mathbf{A}$.
$\mathbf{A}\boldsymbol{b}$ combinación lineal de las columas de $\mathbf{A}$.
$\boldsymbol{a}\mathbf{B}$ combinación lineal de las filas de $\mathbf{A}$.
${}_{i|}\mathbf{A}$ fila $i$ésima de $\mathbf{A}$.
$(a,\ b,\ c,\ldots)$ vector de $\mathbb{R}^n$; con $a, b, c, \ldots \in \mathbb{R}$.
$\mathbf{0}$ nulo.

Índice alfabético