Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Escuela de Ciencias Exactas y Naturales
Departamento de Matemática

Algebra y Geometría Analítica II

Año 2023

UNIDAD 2: Matrices y Determinantes. [1]

En esta unidad estudiaremos una herramienta fundamental: las matrices y sus determinantes. Las matrices son una forma de representar datos tabulados con una doble entrada.Una de las aplicaciones que mejor demuestra la potencia de esta herramienta es su aplicación en la resolución de sistemas de ecuaciones lineales (y todo lo que ello implica), tema que abordaremos en la próxima unidad temática de la materia.

En lo que sigue supondremos que $\mathbb{F}$ es el *cuerpo*[2] de los números racionales, reales o complejos. O sea, $\mathbb{F}$ será $\mathbb{Q}$, $\mathbb{R}$ o $\mathbb{C}$.

1 Matrices

Antes de dar la definición precisa de lo que es una matriz, veamos diferentes frases coloquiales que pretenden describir una matriz (pero qué NO lo logran):

- *rectángulo numérico,*
- *un arreglo ordenado de números,*
- *un conjunto bidimensional de números,*
- *un conjunto de números o expresiones dispuesto de forma rectangular,*
- *una tabla bidimensional de números,*
- ... etc.

La definición precisa y formal de una matriz debe hacerse en términos de los conceptos que conocemos: conjuntos, funciones, relaciones, etc.

Ahora si, veamos la definición formal de matriz.

Definición 1.1. Una *matriz* A de tamaño $m \times n$ con coeficientes en $\mathbb{F}$ es una función

$$A : [\![1, m]\!] \times [\![1, n]\!] \to \mathbb{F},$$

en donde $[\![1, m]\!] = \{1, 2, \dots, m\}$ y $[\![1, n]\!] = \{1, 2, \dots, n\}$. El conjunto de todas las matrices $m \times n$ con coeficientes en $\mathbb{F}$ se denotará por $\mathbb{F}^{m \times n}$. En algunos libros se denota el conjunto de matrices $m \times n$ también por $M_{m\times n}(\mathbb{F})$ o $M(m, n, \mathbb{F})$.

[1]Notas de Cátedra basadas en notas de clases de Silvio Reggiani y ediciones anteriores de la cátedra de las docentes Isolda Cardozo y Paola Tolomei. Edición 2023 con modificaciones de Ma. Inés Lopez Pujato y Pablo Fekete.

[2]*Cuerpo* es un concepto que será explorado en otras materias, por ahora sepamos que nos referimos a estos conjuntos numéricos con dos operaciones, una suma y un producto, que verifican las propiedades que todos conocemos.

Notación. Representamos a una matriz $A : [\![1, m]\!] \times [\![1, n]\!] \to \mathbb{F}$ como un arreglo rectangular

$$A = \begin{array}{cl} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} & \begin{array}{ll} \to & \text{fila 1} \\ \to & \text{fila 2} \\ \vdots & \quad \vdots \\ \to & \text{fila m} \end{array} \\ \begin{array}{cccc} \downarrow & \downarrow & \downarrow & \downarrow \\ \text{col } 1 & \text{col } 2 & \cdots & \text{col } n \end{array} & \end{array}$$

en donde $a_{ij} = A(i, j)$ son los llamados *coeficientes* de la matriz A. También se dice que A tiene m *filas* y n *columnas.*

Otras notaciones frecuentes, son las siguientes: Una matriz $A \in \mathbb{F}^{m\times n}$ se denota abreviadamente como $A = (a_{ij})$ o $A = (a_{ij})_{ij}$, o cuando hay dudas sobre el tamaño de la matriz se suele escribir $A = (a_{ij})_{i=1,\ldots,m,\, j=1,\ldots,n}$, en donde los coeficientes de la matriz A son a_{ij}. Otra notación es la que nombra los coeficientes de la matriz A con la misma letra (en mayúscula en este caso) indicando con un subíndice la posición fila-columna, es decir, $A = (A_{ij})$. Por último, a veces para referinos al elemento en la posición ij de una matriz A, escribiremos $[A]_{ij}$.

Definición 1.2. Diremos que dos matrices $A = (a_{ij})$ y $B = (b_{ij})$ son *iguales* si tienen la misma cantidad de filas, la misma cantidad de columnas y $a_{ij} = b_{ij}$ para todos i, j.

Ejemplo 1.3. Observemos que

$$\begin{pmatrix} 1 & 7 & 0 \\ -2 & 1 & 4 \end{pmatrix} \neq \begin{pmatrix} 1 & -2 \\ 7 & 1 \\ 0 & 4 \end{pmatrix}.$$

En efecto, si bien en algún sentido ambas matrices nos dan la "misma información" (pues todo coeficiente de una aparece en alguna posición de la otra), la primera matriz es de tamaño 2×3, en tanto que la segunda es 3×2. Más generalmente, si $A \in \mathbb{F}^{m\times n}$ y $B \in \mathbb{F}^{n\times m}$, con $n \neq m$, entonces $A \neq B$.

Ejemplo 1.4. $\mathbb{F}^{1\times 1} = \mathbb{F}$, o sea, una matriz 1×1 es un escalar. A decir verdad, una matriz 1×1 es una función $A : \{1\} \times \{1\} \to \mathbb{F}$, pero en estas notas identificaremos A con su imagen, o sea $A = A(1, 1) = a_{11}$.

Ejemplo 1.5. Consideremos la matriz $A \in \mathbb{F}^{3\times 3}$ dada por $a_{ij} = i + j$. Entonces

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{pmatrix}.$$

Definición 1.6. Diremos que una matriz $v \in \mathbb{F}^{1\times n}$ es un *vector fila* de tamaño n, o sea

$$v = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix},$$

y lo identificamos con el vector $\overline{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$. Por otra parte, una matriz $v \in \mathbb{F}^{m\times 1}$ se dirá un *vector columna* de tamaño m, o sea

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}.$$

Definición 1.7. Una matriz A se dice *cuadrada* si tiene la misma cantidad de filas y columnas. Es decir, si $A \in \mathbb{F}^{n\times n}$ para algún n. En tal caso, el *vector diagonal* de A es

$$\operatorname{diag}(A) = (a_{11}, a_{22}, \ldots, a_{nn}).$$

Una matriz cuadrada $A = (a_{ij})$ se dice:

- *triangular superior* si $a_{ij} = 0$ para $i > j$;
- *triangular superior estricta* si $a_{ij} = 0$ para $i \geq j$;
- *triangular inferior* si $a_{ij} = 0$ para $i < j$;
- *triangular inferior estricta* si $a_{ij} = 0$ para $i \leq j$;
- *diagonal* si A es triangular superior y triangular inferior.

Las matrices

$$\begin{pmatrix} 1 & 3 & 0 \\ 0 & -1 & 4 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 2 & -4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 4 & -6 & 0 \\ 1 & 2 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -4 \end{pmatrix},$$

son triangular superior, triangular superior estricta, triangular inferior y diagonal, respectivamente.

Definición 1.8. La *matriz nula* $0 = 0_{m\times n} \in \mathbb{F}^{m\times n}$ es la matriz de tamaño $m \times n$ que tiene todas sus entradas iguales a cero, o sea,

$$0_{m\times n} = \underbrace{\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}}_{n \text{ columnas}} \left.\vphantom{\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}\right\} m \text{ filas}$$

Si bien hemos abusado de la notación denotando por 0 a la matriz nula, con el mismo símbolo que usamos para denotar el elemento nulo en $\mathbb{F}$, no debemos confundir estos conceptos. El significado de 0 debería ser claro del contexto. De todas formas, cuando se presente alguna duda, preferiremos la notación $0_{m\times n}$. ¿Es $0_{2\times 3} = 0_{3\times 2}$?

Definición 1.9. La *matriz identidad* de orden n es la matriz $I = I_n \in \mathbb{F}^{n\times n}$ tal que $I_{ij} = \delta_{ij}$, en donde δ_{ij} es la llamada delta de Kronecker

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

Es decir,

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

es la matriz diagonal con $\operatorname{diag}(I) = (1, 1, \ldots, 1)$.

1.1 Operaciones entre matrices

En lo que sigue, veremos diferentes operaciones con matrices. En efecto, definiremos **producto por escalar**, **suma** y **multiplicación** de matrices.

Definición 1.10. La multiplicación de la matriz $A \in \mathbb{F}^{m\times n}$ por el escalar $\alpha \in \mathbb{F}$ se define como la matriz $C = \alpha A \in \mathbb{F}^{m\times n}$ dada por $c_{ij} = \alpha a_{ij}$.

Por ejemplo,

$$2\begin{pmatrix} 1 & 9 \\ -7 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 18 \\ -14 & 6 \end{pmatrix}.$$

Definición 1.11. Si $A, B \in \mathbb{F}^{m\times n}$, la suma de A con B es la matriz $C = A + B \in \mathbb{F}^{m\times n}$ dada por $c_{ij} = a_{ij} + b_{ij}$.

Por ejemplo,

$$\begin{pmatrix} 2 & 1 & 0 \\ -2 & 3 & 15 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 12 \\ -3 & -3 & 7 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 12 \\ -5 & 0 & 22 \end{pmatrix}.$$

Notar que la operación suma para matrices se define para matrices del mismo tamaño. Luego sólo podremos sumar matrices del mismo tamaño.

Teorema 1.12. *Dados $A, B, C \in \mathbb{F}^{m\times n}$ y $\alpha, \beta \in \mathbb{F}$, se verifican las siguientes propiedades.*

1. $0A = 0_{m\times n}$.
2. $\alpha 0_{m\times n} = 0_{m\times n}$.
3. $A + (B + C) = (A + B) + C$, *es decir, la suma de matrices es asociativa.*
4. $\alpha(\beta A) = (\alpha\beta)A$.
5. $A + 0_{m\times n} = 0_{m\times n} + A = A$, *es decir, la matriz nula es el elemento neutro para la suma de matrices.*
6. $(\alpha + \beta)A = \alpha A + \beta A$, *o sea, la suma en $\mathbb{F}$ es distributiva con respecto a la multiplicación de una matriz por un escalar.*
7. $\alpha(A + B) = \alpha A + \alpha B$, *o sea, la multiplicación por un escalar es distributiva con respecto a la suma de matrices.*

Demostración : *Ejercicio.*

A continuación definiremos la multiplicación de matrices. Intuitivamente, uno querría que la multiplicación de matrices respetara las formas que ya tenemos de multiplicar vectores, es decir, el producto escalar y el producto de sus coordenadas. Recordemos que tenemos dos formas matriciales de pensar un vector en $\mathbb{F}^n$: como un vector fila o como un vector columna. La multiplicación de matrices impone ciertos requerimientos al tamaño de las matrices para poder calcular el producto, pero bajo tales condiciones serán validas las siguientes identidades, por ejemplo para $n = 3$, para vectores fila y vectores columna:

$$\begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1b_1 + a_2b_2 + a_3b_3 \qquad \text{(producto escalar)},$$

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix} = \begin{pmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_3 & a_3b_3 \end{pmatrix} \qquad \text{(producto de coordenadas)}.$$

Notar que en el primer caso, el producto de una matriz 1×3 por una matriz 3×1 nos da una matriz 1×1, en tanto que en el segundo caso, el producto de una matriz 3×1 por una matriz 1×3 nos da una matriz 3×3.

Antes de pasar a la definición del producto de matrices, pensemos en el siguiente ejemplo:

Ejemplo 1.13. Supongamos que un comerciante tiene tres carribares, uno ubicado en la zona sur, otro en la zona norte y el último en la zona oeste de su ciudad. En cada carribar se venden los siguientes productos: maní, panchos y gaseosas. En las siguientes tablas se resumen en pesos los artículos vendidos y sus precios en el cierre de caja del mediodía:

ARTICULOS VENDIDOS	maní	pancho	gaseosa
local sur	1200	2500	3050
local norte	2070	1400	4200
local oeste	300	1200	1900

PRECIOS	costo	venta
maní	1	2
pancho	2	3
gaseosa	$1,5$	3

Para saber cuánto gastó el local de la zona oeste debemos realizar el siguiente cálculo:

$$300 \times 1 + 1200 \times 2 + 1900 \times 1,5 = 5550,$$

y para saber cuánto vendió el local de zona norte simplemente calculamos

$$2070 \times 2 + 1400 \times 3 + 4200 \times 3 = 62460.$$

De esta forma podemos armarnos otra tabla donde se indiquen los totales (ejercicio: completar la tabla):

TOTALES	costo	venta
local sur	$(1200, 2500, 3050) \times (1, 2, 1.5) =$	
local norte		$(2070, 1400, 4200) \times (2, 3, 3) = 62460$
local oeste	5550	

Vale decir, toda la información está condensada en tres matrices (completar): A representa las cantidades vendidas por artículo y por local, P representa los precios de costo y venta por unidad de cada artículo y T representa los totales de costos y de ingresos por ventas por local.

$$A_{3\times 3} = \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}, \qquad P_{3\times 2} = \begin{pmatrix} & \\ & \\ & \end{pmatrix} \qquad \text{y} \qquad T_{3\times 2} = \begin{pmatrix} & \\ & \\ & \end{pmatrix}.$$

La matriz T es la obtenida *multiplicando* la matriz A por la matriz P. Observemos los tamaños: A es 3×3 referidos a 3 locales y 3 artículos, P es 3×2 referidos a 3 artículos y 2 precios y T es 3×2 referidos a 3 locales y 3 precios totales. Los tamaños no son casuales: la cantidad de columnas de A debe ser igual a la cantidad de filas de P, que en ambos casos refieren a *los artículos*, y es posible entonces efectuar el producto escalar de la fila por la columna correspondiente. En la matriz T la información de los artículos se pierde, y sólo tenemos la información de los locales y los totales de precios.

Teniendo presente este ejemplo, veamos ahora la definición de producto matricial.

Definición 1.14. Si $A \in \mathbb{F}^{m\times n}$ y $B \in \mathbb{F}^{n\times p}$ definimos el *producto* de A con B como la matriz $C = AB \in \mathbb{F}^{m\times p}$ dada por

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

En otras palabras, el lugar $(AB)_{ij}$ está dado por el producto escalar entre la i-ésima fila de A y la j-ésima columna de B.

Ejemplo 1.15. Sean

$$A = \begin{pmatrix} 1 & 2 & 0 \\ -3 & 1 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 & -5 & 3 \\ 2 & 1 & 0 & -2 \\ 7 & 5 & -2 & 10 \end{pmatrix}.$$

Observar que como A es 2×3 y B es 3×4 está definido el producto AB (¡y será una matriz 2×4!):

$$AB = \begin{pmatrix} 4 & 3 & -5 & -1 \\ 30 & 18 & 7 & 29 \end{pmatrix}.$$

Una forma sencilla para recordar la fórmula del producto de matrices es la siguiente: se traza una cruz y se disponen los coeficientes de la matriz A en el cuadrante inferior izquierdo, y los coeficientes de la matriz B en el cuadrante superior derecho. Para calcular el coeficiente $(AB)_{ij}$, se calcula el producto escalar entre la fila i de A y la columna j de B. Observar que si uno traza una línea horizontal según la fila i de A y una línea vertical según la columna j de B, entonces el lugar i, j de la matriz AB se encuentra justamente donde se cortan estas líneas

$$\begin{array}{cccc|cccc}
 & & & & \downarrow & \downarrow & \downarrow & \downarrow \\
 & & & & 0 & 1 & \mathbf{-5} & 3 \\
 & & & & 2 & 1 & \mathbf{0} & -2 \\
 & & & & 7 & 5 & \mathbf{-2} & 10 \\
\cline{2-8}
\to & 1 & 2 & 0 & 4 & 3 & -5 & -1 \\
\to & \mathbf{-3} & \mathbf{1} & \mathbf{4} & 30 & 18 & \mathbf{7} & 29
\end{array}$$

En el ejemplo anterior tenemos resaltado en negrita la fila 2 de A y la columna 3 de B, así como el lugar $(AB)_{23}$. Notemos que efectivamente

$$(AB)_{23} = (-3, 1, 4) \times (-5, 0, -2) = (-3) \cdot (-5) + 2 \cdot 0 + 4 \cdot (-2) = 15 - 8 = 7.$$

Observación 1.16. **El producto de matrices no es conmutativo.** (MUY IMPORTANTE).

En efecto, calculemos el producto de las matrices

$$A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}, \qquad\qquad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Observemos que tanto A como B son matrices 2×2, por lo tanto tiene sentido preguntarse si AB es igual a BA. Por un lado

$$\begin{array}{cc|cc} & & 1 & 0 \\ & & 1 & 1 \\ \hline 1 & 2 & 3 & 2 \\ 0 & -1 & -1 & -1 \end{array} \qquad\qquad AB = \begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix},$$

y por el otro

$$\begin{array}{cc|cc} & & 1 & 2 \\ & & 0 & -1 \\ \hline 1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{array} \qquad\qquad BA = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix},$$

de donde sigue que $AB \neq BA$.

Si bien acabamos de observar que el producto de matrices cuadradas del mismo tamaño no es conmutativo, sí es cierto que este producto es asociativo. Más aún, se tiene el siguiente resultado

Teorema 1.17. *El producto de matrices es asociativo, o sea, si $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{n\times p}$ y $C \in \mathbb{F}^{p\times q}$, entonces*

$$A(BC) = (AB)C.$$

Proof. En primer lugar, observemos que los dos productos mencionados en el teorema están bien definidos. En efecto como A es $m \times n$ y BC es $n \times q$, entonces $A(BC)$ es $m \times q$. Asimismo, como AB es $m \times p$ y C es $p \times q$, $(AB)C$ también es $m \times q$.

La prueba es por cálculo directo, usando la definición del producto entre matrices.

Sean $1 \leq i \leq m$ y $1 \leq j \leq q$. Entonces, por un lado

$$(A(BC))_{ij} = \sum_{k=1}^{n} A_{ik}(BC)_{kj} = \sum_{k=1}^{n} A_{ik} \sum_{\ell=1}^{p} B_{k\ell}C_{\ell j} = \sum_{k=1}^{n}\sum_{\ell=1}^{p} A_{ik}B_{k\ell}C_{\ell j} \tag{1.1}$$

y por el otro

$$((AB)C)_{ij} = \sum_{\ell=1}^{p}(AB)_{i\ell}C_{\ell j} = \sum_{\ell=1}^{p}\sum_{k=1}^{n} A_{ik}B_{k\ell}C_{\ell j} = \sum_{k=1}^{n}\sum_{\ell=1}^{p} A_{ik}B_{k\ell}C_{\ell j}. \tag{1.2}$$

Como (1.1) coincide con (1.2) para todos i, j, concluimos que $A(BC) = (AB)C$. $\square$

El siguiente teorema dice que valen las leyes distributivas para el producto de matrices con respecto a la suma. Observemos que como el producto no es conmutativo, debemos expresar (y probar) ambos enunciados.

Teorema 1.18. *1. Si $A \in \mathbb{F}^{m\times n}$ y $B, C \in \mathbb{F}^{n\times p}$, entonces*

$$A(B + C) = AB + AC.$$

2. *Si* $A, B \in \mathbb{F}^{m\times n}$ *y* $C \in \mathbb{F}^{n\times p}$, *entonces*

$$(A+B)C = AC + BC.$$

Proof. Para probar 1 observamos que para $1 \leq i \leq m$, $1 \leq j \leq p$,

$$\begin{aligned}(A(B+C))_{ij} &= \sum_{k=1}^{n} A_{ik}(B+C)_{kj} = \sum_{k=1}^{n} A_{ik}(B_{kj}+C_{kj}) \\ &= \sum_{k=1}^{n} A_{ik}B_{kj} + \sum_{k=1}^{n} A_{ik}C_{kj} = (AB)_{ij} + (AC)_{ij} = (AB+AC)_{ij}.\end{aligned}$$

La parte 2 queda como ejercicio. □

Teorema 1.19. *Sea* $A \in \mathbb{F}^{m\times n}$, *entonces*

$$I_m A = A I_n = A.$$

En particular, la matriz identidad I_n *es el elemento neutro para el producto en* $\mathbb{F}^{n\times n}$.

Proof. Recordemos que $(I_m)_{ij} = \delta_{ij}$ es la delta de Kronecker. Luego

$$(I_m A)_{ij} = \sum_{k=1}^{m} (I_m)_{ik} A_{kj} = \sum_{k=1}^{m} \delta_{ik} A_{kj} = A_{ij},$$

de donde sigue que $I_m A = A$. Probar como ejercicio que $AI_n = A$. □

Ejemplo 1.20. Consideremos la matriz

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

y calculemos las potencias sucesivas de A. Es decir, $A^2 = AA$, $A^3 = AA^2 = AAA$, $A^4 = AA^3 = A^2A^2 = AAAA$, etc. En efecto,

$$\begin{aligned}A^2 &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \\ A^3 = AA^2 &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} \\ A^4 = AA^3 &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} \\ A^5 = AA^4 &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 5 \\ 5 & 3 \end{pmatrix}\end{aligned}$$

El lector quizás ya pueda intuir que el resultado general es

$$A^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}, \tag{1.3}$$

en donde los números F_n son los números de la sucesión de Fibonacci, dada por $F_0 = 0$, $F_1 = 1$ y $F_{n+1} = F_n + F_{n-1}$ para $n \geq 1$. En efecto, probemos la fórmula (1.3) por inducción. Es claro que (1.3) vale para $n = 1$, pues $A^1 = A$ (de hecho, ya lo probamos para $n = 1, 2, 3, 4, 5$). Supongamos que dicha fórmula vale para un cierto $n \geq 2$ y verifiquemos el resultado para $n + 1$:

$$\begin{aligned}A^{n+1} = AA^n &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} \\ &= \begin{pmatrix} F_{n+1}+F_n & F_n + F_{n-1} \\ F_{n+1} & F_n \end{pmatrix} \\ &= \begin{pmatrix} F_{n+2} & F_{n+1} \\ F_{n+1} & F_n \end{pmatrix},\end{aligned}$$

con lo cual la fórmula (1.3) vale para todo $n \in \mathbb{N}$.

Hay otro tipo de operación que se puede aplicar sobre una matriz, la cual no involucra operaciones algebraicas (como sumar o multiplicar).

Definición 1.21. Dada $A \in \mathbb{F}^{m\times n}$ se define la *matriz transpuesta* de A, como la matriz $A^t \in \mathbb{F}^{n\times m}$ dada por

$$(A^t)_{ij} = A_{ji}.$$

Es decir, A^t es la matriz cuyas columnas son las filas de A.

Ejemplo 1.22. Notemos si una matriz tiene tamaño, digamos, 2×3 entonces su transpuesta tendrá tamaño 3×2. Por ejemplo,

$$\begin{pmatrix} 2 & -7 & 1 \\ 0 & 3 & 6 \end{pmatrix}^t = \begin{pmatrix} 2 & 0 \\ -7 & 3 \\ 1 & 6 \end{pmatrix}.$$

Ejemplo 1.23. Si A es una matriz diagonal, entonces $A^t = A$.

Proposición 1.24. *Sean $A, B \in \mathbb{F}^{m\times n}$, $\alpha \in \mathbb{F}$. Entonces:*

1. $(A^t)^t = A$;
2. $(\alpha A)^t = \alpha A^t$;
3. $(A + B)^t = A^t + B^t$.

Proof. Observar que para todos $1 \leq i \leq m$, $1 \leq j \leq n$ valen

$$((A^t)^t)_{ij} = (A^t)_{ji} = A_{ij},$$

de donde sigue que $(A^t)^t = A$;

$$((\alpha A)^t)_{ij} = (\alpha A)_{ji} = \alpha A_{ji} = \alpha (A^t)_{ij} = (\alpha A^t)_{ij},$$

de donde sigue que $(\alpha A)^t = \alpha A^t$; y

$$((A + B)^t)_{ij} = (A + B)_{ji} = A_{ji} + B_{ji} = (A^t)_{ij} + (B^t)_{ij} = (A^t + B^t)_{ij},$$

de donde sigue que $(A + B)^t = A^t + B^t$. □

Notar que no vale en general que $(AB)^t = A^t B^t$, de hecho, quizás ni siquiera tenga sentido el producto del lado derecho de esta igualdad. Sin embargo, se tiene el siguiente resultado que relaciona la transposición con el producto de matrices.

Proposición 1.25. *Si $A \in \mathbb{F}^{m\times n}$ y $B \in \mathbb{F}^{n\times p}$ entonces*

$$(AB)^t = B^t A^t.$$

Proof. Queda como ejercicio comprobar la buena definición, en cuanto a los tamaños de las matrices.

Por un cálculo directo se tiene que

$$((AB)^t)_{ij} = (AB)_{ji} = \sum_{k=1}^{n} A_{jk} B_{ki} = \sum_{k=1}^{n} B_{ki} A_{jk} = \sum_{k=1}^{n} (B^t)_{ik} (A^t)_{kj} = (B^t A^t)_{ij}$$

para todos $1 \leq i \leq p$, $1 \leq j \leq m$. Por tanto, $(AB)^t = B^t A^t$. □

Definición 1.26. Sea $A \in \mathbb{F}^{n\times n}$. Decimos que A es:

- *simétrica* si $A^t = A$;
- *antisimétrica* si $A^t = -A$.

Ejemplo 1.27. La matriz

$$A = \begin{pmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 0 \end{pmatrix}$$

es simétrica, $A^t = A$, en tanto que la matriz

$$B = \begin{pmatrix} 0 & 1 & 5 \\ -1 & 0 & -7 \\ -5 & 7 & 0 \end{pmatrix}$$

es antisimétrica, $B^t = -B$. Observar que "simetría" o "antisimetría" para matrices, significa simetría o antisimetría con respecto a la diagonal. En particular, una matriz simétrica queda determinada por los coeficientes ubicados de la diagonal para arriba.

Nota 1.28. Si A es antisimétrica entonces $\operatorname{diag}(A) = (0, 0, \dots, 0)$.

Proposición 1.29. *Toda matriz $A \in \mathbb{F}^{n\times n}$ puede escribirse de manera única como la suma de una matriz simétrica y una matriz antisimétrica. Es decir,*

$$A = A_{\text{sim}} + A_{\text{anti}},$$

con $(A_{\text{sim}})^t = A_{\text{sim}}$ y $(A_{\text{anti}})^t = -A_{\text{anti}}$.

Proof. Observemos que

$$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t).$$

Luego, si llamamos

$$A_{\text{sim}} = \frac{1}{2}(A + A^t), \qquad A_{\text{anti}} = \frac{1}{2}(A - A^t),$$

entonces por la Proposición 1.24, se tiene que

$$(A_{\text{sim}})^t = \left(\frac{1}{2}(A + A^t)\right)^t = \frac{1}{2}(A^t + A) = \frac{1}{2}(A + A^t) = A_{\text{sim}}$$

y

$$(A_{\text{anti}})^t = \left(\frac{1}{2}(A - A^t)\right)^t = \frac{1}{2}(A^t - A) = -\frac{1}{2}(A - A^t) = -A_{\text{anti}},$$

como se quería demostrar.

Ejercicio : Probar que la descomposición es única. Es decir, si $A = A' + A''$ con $(A')^t = A'$ y $(A'')^t = -A''$, entonces $A' = A_{\text{sim}}$ y $A'' = A_{\text{anti}}$.

□

Definición 1.30. Dada $A \in \mathbb{F}^{n\times n}$, se define la *traza* de A como

$$\operatorname{tr} A = \sum_{i=1}^{n} A_{ii}.$$

Ejemplo 1.31. Para la matriz

$$A = \begin{pmatrix} 2 & -5 & 7 \\ 0 & 7 & 2 \\ 0 & 1 & -9 \end{pmatrix}$$

tenemos que

$$\operatorname{tr} A = 2 + 7 - 9 = 0.$$

Teorema 1.32. *Dados $A, B \in \mathbb{F}^{n\times n}$, $\alpha \in \mathbb{F}$, se verifica:*

1. $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$;
2. $\operatorname{tr}(\alpha A) = \alpha \operatorname{tr} A$;
3. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Demostración : *Ejercicio.*

Proposición 1.33. *Si $A \in \mathbb{F}^{n\times n}$, entonces*

$$\operatorname{tr}(AA^t) = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}^2.$$

Proof. Observemos que

$$(AA^t)_{ii} = \sum_{j=1}^{n} A_{ij}(A^t)_{ji} = \sum_{j=1}^{n} A_{ij}A_{ij} = \sum_{j=1}^{n} A_{ij}^2,$$

de donde se obtiene que

$$\operatorname{tr}(AA^t) = \sum_{i=1}^{n}(AA^t)_{ii} = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}^2.$$

□

Comentario. 1. Si $\mathbb{F} = \mathbb{R}$, puede pensarse en el número $\operatorname{tr}(AA^t)$ como el "módulo" al cuadrado de la matriz $A \in \mathbb{R}^{n\times n}$, pues este número se calcula sumando los cuadrados de todos los coeficientes de la matriz A. Esto introduce una noción geométrica de distancia en el conjunto $\mathbb{R}^{n\times n}$, es decir, la distancia entre dos matrices $A, B \in \mathbb{F}^{n\times n}$ se define como el módulo de la diferencia $B - A$,

$$\operatorname{dist}(A,B) = \operatorname{tr}\left((B-A)(B-A)^t\right)^{\frac{1}{2}}.$$

A partir de esta distancia, es posible definir muchas otras nociones geométricas.

2. El conjunto de todas las matrices $A \in \mathbb{R}^{n\times n}$ tales que

$$\operatorname{tr} A = \text{const.}$$

puede pensarse como un "hiperplano" en $\mathbb{R}^{n\times n}$, pues dichas matrices satisfacen una especie de "ecuación general del plano": una ecuación lineal generalizada a n^2 variables.

Ejemplo 1.34. El hiperplano de todas las matrices 2×2 de traza cero, con coeficientes reales, está dado por

$$\left\{\begin{pmatrix} x & y \\ z & w \end{pmatrix} \in \mathbb{R}^{2\times 2} : x + w = 0\right\}.$$

1.2 Ejemplo de aplicación.

Así como guardamos información en vectores (por ejemplo velocidades, posición, color RGB de un pixel, etc.) también en las matrices podemos almacenar datos. Vimos en un ejemplo cómo un comerciante almacenaba datos de las ventas de los artículos por local, y los costos. Para ubicar un pixel en la pantalla también se utilizan matrices. Al complejizar la información, el ente matemático que precisamos también se vuelve más complejo. Incluso para ciertos datos se utilizan *arreglos* de mayor *dimensión*, donde estas nociones tienen una definición precisa y unas ciertas reglas aritméticas. En este curso no trabajaremos con éstas, pues debemos para ello primero comprender bien el concepto de matriz.

En esta sección queremos dar una idea muy simplificada de aplicación de matrices a la Ciencia de Datos. Si bien aún no contamos con los temas propios del álgebra lineal, sino tan solo con la definición de matriz, intentaremos describir al menos el modelo de un sistema de recomendación. Netflix, Google, Spotify, YouTube y miles de apps más utilizan algoritmos de recomendación, y los principios de los mismos están dados en términos de álgebra lineal, y consiguientemente, en términos de vectores y matrices.

Consideremos un servicio de streaming que tiene guardados los siguientes datos de 6 clientes: 1: Alex, 2: Betty, 3: Carmen, 4: Dani, 5: Enrique, 6: Fausto que han indicado cuáles películas han visto y les han gustado de entre las 7 películas: 1: Godzilla, 2: Hamlet, 3:Ishtar, 4: JFK, 5: King Kong, 6:Lincoln, 7: Macbeth.

Las películas 1 y 5 son ambas películas de monstruos, las 2 y 7 están basadas en obras de Shakespeare, las 4 y 6 están basadas en vidas de presidentes norteamericanos y la película 3 es una comedia. Esta información no la podemos tabular, sin embargo buscamos una forma de que el algoritmo *infiera* similaridades y diferencias entre las películas a partir de los datos. Esto es lo que se busca en la ciencia de datos.

Al usuario 1, Alex, le han gustado Godzilla y King Kong. A Betty le ha gustado Hamlet, JFK, Lincoln y Macbeth. A Carmen le gustó Ishtar y King Kong. Esta información, y la de los demás usuarios, la han tabulado en una matriz A 7×6 como sigue:

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

En esta matriz claramente las 7 filas se corresponden a las 7 películas mientras que las 6 columnas se corresponden a los 6 usuarios. En la primera columna está la información de Alex, que es el cliente 1, a quien le gustaron las películas 1 y 5 (Godzilla y King Kong) cuya información la encontramos en las filas 1 y 5 respectivamente. Podemos observar que en las posiciones 11 y 15 hay números 1 y en las demás posiciones de la primera columna hay 0. Así se tabula la información. Si pasamos a la segunda columna, vemos que Betty tiene 1 en la segunda, cuarta, sexta y séptima columnas, como era esperable. Idéntico análisis para Carmen. Para Dani no dijimos qué información teníamos pero podemos recuperarla mirando la columna 4: a Dani le han gustado las películas 1, 2 y 5, o sea, Godzilla, Hamlet y King Kong. ¿Qué películas les han gustado a Enrique y a Fausto? ¿A quiénes les ha gustado la película Macbeth? Toda esta información se recupera fácilmente con nuestra matriz. Una matriz que encierra este tipo de información se conoce como *term-by-document matrix.*

Se une al servicio de streaming una nueva usuaria, Greta. Greta indica que ha visto y le ha gustado Godzilla (1), JFK (4) y Macbeth (7). ¿Qué otras películas le podrían gustar a Greta? Hay muchísimas formas de dar una respuesta a esta pregunta. Vamos a intentar describir una solución al problema de la recomendación, muy sencilla.

La información que Greta proveyó puede, en el espíritu del razonamiento previo, guardarse en un vector columna $v = (1, 0, 0, 1, 0, 0, 1)^T$. Para poder recomendar una película, podríamos pensar en las columnas $A_1, \ldots, A_6$ de nuestra matriz A, y tratar de considerar la columna que más se aproxime a v, en algún sentido. Esto nos estaría indicando cuál usuario del servicio tiene gustos más parecidos a los de Greta, para luego sugerirle alguna película que ella no haya visto y que dicho usuario haya disfrutado. El sentido de *cercanía* que le vamos a dar en esta resolución es el del ángulo que forma la columna correspondiente al usuario con el vector v: buscaremos el menor ángulo entre todos los vectores columna de A y v.

Cada vector columna $\overline{u}$, identificado como el vector libre A_j^T de $\mathbb{R}^7$, podemos proyectarlo sobre el vector libre $\overline{v}$ que también lo interpretamos como un vector de $\mathbb{R}^7$, como sigue:

$$\text{proy}_{\overline{v}}\overline{u} = \frac{\overline{u} \times \overline{v}}{|\overline{v}|},$$

de donde se deduce que el ángulo es

$$\cos(\alpha_j) = \frac{\overline{u} \times \overline{v}}{|\overline{u}||\overline{v}|} = \frac{A_j^T v}{||A_j||||v||},$$

donde hemos denotado con $||v||$ a la norma del vector $\overline{v}$, es decir, $||v|| = |v|$ (la diferencia está en que v es una matriz columna y $\overline{v}$ es un vector, o sea, es sólo de notación).

Resolviendo estos cálculos (queda como ejercicio para el lector), vemos que los usuarios 1 (Alex) y 4 (Dani) producen valores del coseno cercanos a $0,577$. En general, un valor mayor a $0,5$ se considera bastante bueno. Si miramos los gustos de Alex y Dani, vemos que ambos comparten Godzilla y King Kong. Además, a Dani le gustó Hamlet. Se le podría sugerir a Greta que mire King Kong, o Hamlet.

Con más datos, seguramente nuestro recomendador podría hacer un mejor trabajo. Efectivamente, cuanto más sepamos de matrices y álgebra lineal, mejor podremos aprovechar las herramientas para armar un modelo más apropiado. Pero, como primera aproximación, este modelo es bastante interesante.

2 Determinantes

2.1 Permutaciones

Recordemos que para $n \in \mathbb{N}$, se denota por $[\![1, n]\!] = \{1, 2, \ldots, n\}$ el intervalo de todos los enteros comprendidos entre 1 y n.

Definición 2.1. Dado $n \in \mathbb{N}$, denotaremos por

$$S_n = \{\sigma : [\![1, n]\!] \to [\![1, n]\!] : \sigma \text{ es biyectiva}\}$$

el conjunto de permutaciones de n elementos.

Observación 2.2. Ya hemos probado los siguientes hechos:

i) Hay $n!$ permutaciones de n elementos, o sea $|S_n| = n!$.

ii) Si $\sigma, \tau \in S_n$ entonces $\sigma \circ \tau \in S_n$, en donde $\sigma \circ \tau$ es la *composición* de σ con τ y está definida por $(\sigma \circ \tau)(i) = \sigma(\tau(i))$ para todo $i \in [\![1, n]\!]$.

iii) La función identidad $\mathrm{id} : S_n \to S_n$ se comporta como el "elemento neutro" de S_n con respecto a la operación composición, es decir,

$$\mathrm{id} \circ \sigma = \sigma \circ \mathrm{id} = \sigma$$

para toda $\sigma \in S_n$.

iv) Toda $\sigma \in S_n$ tiene una inversa, es decir, existe una permutación $\sigma^{-1} \in S_n$ tal que

$$\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \mathrm{id}\,.$$

v) Se puede identificar una permutación $\sigma \in S_n$ con una n-upla de números entre 1 y n, todos distintos,

$$\sigma \longleftrightarrow (\sigma(1), \sigma(2), \ldots, \sigma(n)).$$

Más aún, en la presente unidad abusaremos de la notación escribiendo

$$\sigma = (\sigma(1), \sigma(2), \ldots, \sigma(n)).$$

Por ejemplo, la n-upla $(1, 2, \ldots, n)$ representa la permutación identidad, en tanto que $(2, 1, 3, 4, \ldots, n)$ representa la permutación que intercambia el 1 con el 2 y deja fijos los demás elementos.

Las permutaciones que intercambian solamente dos elementos forman una familia muy importante y llevan nombre propio.

Definición 2.3. Una *trasposición* es una permutación que intercambia sólo dos elementos. O sea, $\tau \in S_n$ es una trasposición si existen $i, j \in [\![1, n]\!]$, $i \neq j$, tales que

$$\tau(k) = \begin{cases} j & \text{si } k = i, \\ i & \text{si } k = j, \\ k & \text{si } k \neq i, j. \end{cases}$$

Notación. Denotaremos por $\tau_{i,j}$ la transposición que intercambia i con j. Observemos que esta notación es ambigua pues, por ejemplo, $\tau_{1,2}$ representa permutaciones distintas según consideremos $\tau_{1,2} \in S_2$ o $\tau_{1,2} \in S_3$. En el primer caso $\tau_{1,2} = (2,1)$ y en el segundo $\tau_{1,2} = (2,1,3)$. Por ende, para indicar una trasposición $\tau_{i,j}$ también deberíamos indicar el n tal que $\tau_{i,j} \in S_n$. Sin embargo, en lo que sigue siempre podremos deducir quién es n a partir del contexto. Observemos que para $i < j$, la transposición que intercambia i con j se representa mediante la siguiente n-upla, sabiendo así que $\tau_{i,j} \in S_n$.

$$\tau_{i,j} = (1, \dots, i-1, j, i+1, \dots, j-1, i, j+1, \dots, n).$$

Observación 2.4. Es importante notar que si τ es una trasposición, entonces $\tau \circ \tau = \mathrm{id}$, o sea, τ es su propia inversa, $\tau = \tau^{-1}$.

Usando trasposiciones, uno puede calcular fácilmente la inversa de una permutación arbitraria $\sigma \in S_n$. En efecto, simplemente debemos reordenar de menor a mayor la n-upla $(\sigma(1), \sigma(2), \dots, \sigma(n))$ intercambiando sólo dos elementos en cada paso.

Ejemplo 2.5. Sea $\sigma \in S_6$ la permutación representada por la 6-upla $\sigma = (2,4,3,1,6,5)$. En un primer paso podemos intercambiar 1 con 2 para obtener la 6-upla $(1,4,3,2,6,5)$, o más formalmente

$$\tau_{1,2} \circ \sigma = (1,4,3,2,6,5).$$

Luego podemos intercambiar 2 con 4 para obtener

$$\tau_{2,4} \circ \tau_{1,2} \circ \sigma = (1,2,3,4,6,5).$$

Finalmente, para llegar a la identidad, debemos intercambiar 5 con 6,

$$\tau_{5,6} \circ \tau_{2,4} \circ \tau_{1,2} \circ \sigma = (1,2,3,4,5,6) = \mathrm{id}\,.$$

Por lo tanto,

$$\sigma^{-1} = \tau_{5,6} \circ \tau_{2,4} \circ \tau_{1,2},$$

de donde se desprende que $\sigma = (\tau_{1,2})^{-1} \circ (\tau_{2,4})^{-1} \circ (\tau_{5,6})^{-1}$ y como cada trasposición es su propia inversa, se obtiene que

$$\sigma = \tau_{1,2} \circ \tau_{2,4} \circ \tau_{5,6},$$

es decir, logramos escribir σ como una composición de trasposiciones. Este es un hecho general, como lo refleja el siguiente teorema.

Teorema 2.6.

1. *Toda $\sigma \in S_n$ se puede escribir como una composición de trasposiciones, es decir:*

$$\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k,$$

   con τ_i trasposición.

2. *Dada $\sigma \in S_n$, la cantidad k de trasposiciones τ_i usadas en una descomposición σ como la dada en el ítem anterior es siempre par o siempre impar.*

Proof. Daremos un esquema de la prueba. *Ejercicio : completar la demostración.*

1. Dada $\sigma \in S_n$, definimos $\sigma_0 = \sigma$ y luego recursivamente $\sigma_j \in S_n$, para $j = 1, \dots, n$ como sigue:

- si $\sigma_{j-1}(j) \neq j$, $\sigma_j := \tau_{j\sigma_{j-1}(j)} \circ \sigma_{j-1}$;

- si $\sigma_{j-1}(j) = j$, $\sigma_j := \sigma_{j-1}$.

Así definida sigue que $\sigma_j(k) = k$ para todo $k = 1, \dots, j$. Finalmente, $\sigma_n = id$, de donde vemos que una composición de transposiciones debe ser la inversa de σ, luego σ no es otra cosa que la inversa de esa composición de transposiciones, vale decir, una composición de transposiciones en si misma.

Sugerencia: para comprender el enunciado correctamente y luego completar la demostración: armar $\sigma_1, \sigma_2, \sigma_3$ y chequear en cada caso las afirmaciones hechas.

2. Dejamos como ejercicio probar que para describir la identidad como composición de transposiciones *siempre* se requiere de una cantidad par. Luego, si $\sigma \in S_n$ puede escribirse de dos maneras como composición de transposiciones, por ejemplo: $\sigma = \tau_{i_1 j_1} \circ \cdots \circ \tau_{i_k j_k}$ y $\sigma = \tau_{i'_1 j'_1} \circ \cdots \circ \tau_{i'_l j'_l}$, con k y l las cantidades respectivas, podemos considerar $\sigma \circ \sigma^{-1} = id$, de donde podemos obtener la identidad componiendo $k + l$ transposiciones. Como $k + l$ debe ser par (por el ejercicio anterior), se tiene que o bien k y l son ambos pares o bien k y l son ambos impares.

□

Ejemplo 2.7. Consideremos $\sigma = (2, 4, 3, 1) \in S_4$. Razonando como en el Ejemplo 2.5 obtenemos que

$$\begin{aligned} \tau_{1,2} \circ \sigma &= (1, 4, 3, 2) \\ \tau_{2,4} \circ \tau_{1,2} \circ \sigma &= (1, 2, 3, 4), \end{aligned}$$

de donde $\sigma = \tau_{1,2} \circ \tau_{2,4}$ y logramos escribir σ como una composición de $k = 2$ trasposiciones. Obviamente esta no es la única manera de reordenar de menor a mayor los números $(2, 4, 3, 1)$, uno podría hacer también

$$\begin{aligned} \tau_{1,3} \circ \sigma &= (2, 4, 1, 3), \\ \tau_{1,4} \circ \tau_{1,3} \circ \sigma &= (2, 1, 4, 3), \\ \tau_{1,2} \circ \tau_{1,4} \circ \tau_{1,3} \circ \sigma &= (1, 2, 4, 3), \\ \tau_{3,4} \circ \tau_{1,2} \circ \tau_{1,4} \circ \tau_{1,3} \circ \sigma &= (1, 2, 3, 4), \end{aligned}$$

de donde sigue que σ también se puede escribir como la composición $\sigma = \tau_{1,3} \circ \tau_{1,4} \circ \tau_{1,2} \circ \tau_{3,4}$ usando $k = 4$ trasposiciones. Notemos que en ambos casos necesitamos una cantidad par de trasposiciones para descomponer σ.

Definición 2.8. El *signo* de una permutación $\sigma \in S_n$ se define por

$$\operatorname{sg}(\sigma) = (-1)^k,$$

en donde $\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_k$ con τ_i trasposición (notar que el signo está bien definido por el Teorema 2.6). Diremos que σ es

- una permutación *par* si $\operatorname{sg}(\sigma) = 1$,
- o una permutación *impar* si $\operatorname{sg}(\sigma) = -1$.

Teorema 2.9. *1.* $\operatorname{sg}(\mathrm{id}) = 1$.

2. $\operatorname{sg}(\sigma \circ \tau) = \operatorname{sg}(\sigma) \operatorname{sg}(\tau)$ *para todas* $\sigma, \tau \in S_n$.

3. $\operatorname{sg}(\sigma^{-1}) = \operatorname{sg}(\sigma)^{-1} = \operatorname{sg}(\sigma)$ *para toda* $\sigma \in S_n$.

Proof. La primera parte sigue inmediatamente del Teorema 2.6 y la definición de signo de una permutación. Para la segunda, usando el Teorema 2.6, podemos escribir $\sigma = \sigma_1 \circ \cdots \circ \sigma_k$ y $\tau = \tau_1 \circ \cdots \circ \tau_\ell$ con σ_i, τ_j trasposiciones. Luego

$$\sigma \circ \tau = \sigma_1 \circ \cdots \circ \sigma_k \circ \tau_1 \circ \cdots \circ \tau_\ell$$

es composición de $k + \ell$ trasposiciones y por consiguiente

$$\operatorname{sg}(\sigma \circ \tau) = (-1)^{k+\ell} = (-1)^k (-1)^\ell = \operatorname{sg}(\sigma) \operatorname{sg}(\tau).$$

Para probar la tercera parte podemos usar lo que acabamos de demostrar. En efecto, como $\sigma \circ \sigma^{-1} = \mathrm{id}$, sigue que

$$1 = \operatorname{sg}(\mathrm{id}) = \operatorname{sg}(\sigma \circ \sigma^{-1}) = \operatorname{sg}(\sigma) \operatorname{sg}(\sigma^{-1}),$$

con lo cual

$$\operatorname{sg}(\sigma^{-1}) = \frac{1}{\operatorname{sg}(\sigma)} = \operatorname{sg}(\sigma)^{-1}.$$

Observar que si $\operatorname{sg}(\sigma) = 1$, entonces $\operatorname{sg}(\sigma^{-1}) = 1/1 = 1$ y si $\operatorname{sg}(\sigma) = -1$, entonces $\operatorname{sg}(\sigma^{-1}) = 1/(-1) = -1$. En cualquiera de los dos casos vale $\operatorname{sg}(\sigma) = \operatorname{sg}(\sigma^{-1})$. □

Con estos preliminares sobre permutaciones estamos ya en condiciones de definir el determinante de una matriz cuadrada.

2.2 Determinantes

Definición 2.10. Sea $A \in \mathbb{F}^{n\times n}$, el *determinante* de A se define como

$$\det A = \sum_{\sigma\in S_n} \operatorname{sg}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)}. \tag{2.1}$$

En otras palabras, para calcular el determinante de A hay que sumar todos los posibles factores que se pueden armar con coeficientes de A, tomando un elemento en cada fila y cada columna, en donde el factor $A_{1\sigma(1)}A_{2\sigma(2)}\cdots A_{n\sigma(n)}$ debe ir multiplicado por el signo de la permutación σ.

Es común la notación que indica el determinante de la matriz $A = (a_{ij})$ utilizando barras verticales:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

A continuación estudiamos algunos casos particulares que ayudarán a clarificar la definición del determinante.

Para el caso de una matriz $A = (a) \in \mathbb{F}^{1\times 1}$ tenemos que $S_1 = \{\mathrm{id}\}$ y por ende

$$\det A = a.$$

Cuando $n = 2$ tenemos que $S_2 = \{(1,2),(2,1)\}$, en donde $\operatorname{sg}((1,2)) = \operatorname{sg}(\mathrm{id}) = 1$ y $\operatorname{sg}((2,1)) = -1$, pues $(2,1) = \tau_{1,2}$ es una trasposición. Luego, si

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

tendremos que

$$\begin{aligned} \det A &= \sum_{\sigma\in S_2} \operatorname{sg}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \\ &= \operatorname{sg}((1,2) a_{11}a_{22} + \operatorname{sg}((2,1)) a_{12}a_{21} \\ &= a_{11}a_{22} - a_{12}a_{21}. \end{aligned}$$

Esta fórmula a veces se escribe como

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Estudiemos ahora el caso de matrices 3×3. Notemos que para $n = 3$, la suma (2.1) tiene 6 sumandos pues $|S_3| = 3! = 6$. Más precisamente,

$$S_3 = \{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\}.$$

Pero para calcular el determinante de una matriz 3×3 no solo debemos conocer todas las permutaciones de 3 elementos, sino también sus signos. Para hacer esto notemos que:

- $(1,2,3) = \mathrm{id} \Rightarrow \operatorname{sg}((1,2,3)) = 1$,
- $(1,3,2) = \tau_{2,3} \Rightarrow \operatorname{sg}((1,3,2)) = -1$,
- $(2,1,3) = \tau_{1,2} \Rightarrow \operatorname{sg}((2,1,3)) - 1$,

- $(2,3,1) = \tau_{2,3} \circ \tau_{1,3} \Rightarrow \operatorname{sg}((2,3,1)) = 1$,
- $(3,1,2) = \tau_{1,2} \circ \tau_{1,3} \Rightarrow \operatorname{sg}((3,1,2)) = 1$,
- $(3,2,1) = \tau_{1,3} \Rightarrow \operatorname{sg}((3,2,1)) = -1$.

Luego, si

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

entonces

$$\begin{aligned} \det A &= \sum_{\sigma \in S_3} \operatorname{sg}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)} \\ &= \operatorname{sg}((1,2,3)) a_{11}a_{22}a_{33} + \operatorname{sg}((1,3,2)) a_{11}a_{23}a_{32} + \operatorname{sg}((2,1,3)) a_{12}a_{21}a_{33} \\ &\qquad + \operatorname{sg}((2,3,1)) a_{12}a_{23}a_{31} + \operatorname{sg}((3,1,2)) a_{13}a_{21}a_{32} + \operatorname{sg}((3,2,1)) a_{13}a_{22}a_{31}. \end{aligned}$$

Reemplazando los signos de las permutaciones por los calculados más arriba y reordenando los sumandos, se obtiene la fórmula

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \tag{2.2}$$

Nota 2.11. El determinante de una matriz 3×3 puede calcularse mediante una regla nemotécnica llama *regla de Sarrus*. El procedimiento es el siguiente: primero hay que disponer un arreglo de cinco filas que consiste de las tres filas de la matriz A, agregando como cuarta y quinta filas la primera y la segunda fila de A respectivamente. O sea,

$$\begin{array}{ccccc} & a_{11} & a_{12} & a_{13} & \\ & a_{21} & a_{22} & a_{23} & \\ & a_{31} & a_{32} & a_{33} & \\ -\swarrow & a_{11} & a_{12} & a_{13} & \searrow + \\ -\swarrow & a_{21} & a_{22} & a_{23} & \searrow + \\ -\swarrow & & & & \searrow + \end{array}$$

y luego se suman los productos de los elementos en las diagonales (de izquierda a derecha en dirección sudeste) y se restan los productos de los elementos en las antidiagonales (de derecha a izquierda en dirección sudoeste). Así se obtiene

$$\det A = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{13}a_{22}a_{31} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21}.$$

Observar que esta fórmula coincide, reordenando sumandos e intercambiando factores, con la que dedujimos en (2.2).

A continuación derivamos propiedades importantes de la función determinante.

2.3 Propiedades

Proposición 2.12. *Si* $I \in \mathbb{F}^{n \times n}$ *es la matriz identidad, entonces* $\det I = 1$.

Proof. Por definición tenemos que

$$\det I = \sum_{\sigma \in S_n} \operatorname{sg}(\sigma) \prod_{i=1}^{n} I_{i\sigma(i)},$$

pero $I_{i\sigma(i)} \neq 0$ si y sólo si $\sigma(i) = i$, y en tal caso se tiene $I_{ii} = 1$. Luego, la suma anterior tiene un único sumando no nulo, el correspondiente a $\sigma = \mathrm{id}$, por tanto

$$\det I = \operatorname{sg}(\mathrm{id}) \prod_{i=1}^{n} I_{ii} = 1. \qquad \square$$

Usando la misma idea que en la demostración de la Proposición 2.12, se puede probar el siguiente resultado.

Proposición 2.13. *Si $A \in \mathbb{F}^{n\times n}$ es una matriz diagonal entonces* $\det A = A_{11}A_{22}\cdots A_{nn}$.

Demostración : *Ejercicio.*

Proposición 2.14. *Si $A \in \mathbb{F}^{n\times n}$ entonces,*

$$\det A = \det A^t.$$

Proof. Si $\sigma \in S_n$, para $i \in [\![1,n]\!]$ será $\sigma(i) = j \in [\![1,n]\!]$. Luego, $i = \sigma^{-1}(j)$. Entonces

$$\prod_{i=1}^{n} A_{i\sigma(i)} = \prod_{j=1}^{n} A_{\sigma^{-1}(j)j},$$

de modo que

$$\begin{aligned} \operatorname{sg}(\sigma)A_{1\sigma(1)}A_{2\sigma(2)}\cdots A_{n\sigma(n)} &= \operatorname{sg}(\sigma^{-1})A_{\sigma^{-1}(1)1}A_{\sigma^{-1}(2)2}\cdots A_{\sigma^{-1}(n)n} \\ &= \operatorname{sg}(\sigma^{-1})(A^t)_{1\sigma^{-1}(1)}(A^t)_{2\sigma^{-1}(2)}\cdots(A^t)_{n\sigma^{-1}(n)}. \end{aligned}$$

Luego

$$\det A = \sum_{\sigma\in S_n} \operatorname{sg}(\sigma)\prod_{i=1}^{n} A_{i\sigma(i)} = \sum_{\sigma^{-1}\in S_n} \operatorname{sg}(\sigma^{-1})\prod_{i=1}^{n} (A^t)_{i\sigma^{-1}(i)} = \det A^t.$$

□

Proposición 2.15. *Si $A \in \mathbb{F}^{n\times n}$ tiene dos columnas (o filas) iguales, entonces*

$$\det A = 0.$$

Proof. Supongamos primero que A tiene dos columnas iguales, digamos la columna k es igual a la columna j con $k < j$. Esto quiere decir que $a_{ik} = a_{ij}$ para todo $i = 1,\ldots,n$. Dada $\sigma \in S_n$, sea $\tilde{\sigma} = \tau_{k,j}\circ\sigma$. Se tiene que

$$A_{1\sigma(1)}A_{2\sigma(2)}\cdots A_{n\sigma(n)} = A_{1\tilde{\sigma}(1)}A_{2\tilde{\sigma}(2)}\cdots A_{n\tilde{\sigma}(n)},$$

pues $\sigma(i) = k$ implica $\tilde{\sigma}(i) = j$ y por tanto $A_{i\sigma(i)} = A_{i\tilde{\sigma}(i)}$. Análogamente $A_{j\sigma(j)} = A_{j\tilde{\sigma}(j)}$ y $A_{k\sigma(k)} = A_{k\tilde{\sigma}(k)}$ para $k \neq i, j$. Además

$$\operatorname{sg}(\tilde{\sigma}) = \operatorname{sg}(\tau_{k,j}\circ\sigma) = \operatorname{sg}(\tau_{k,j})\operatorname{sg}(\sigma) = -\operatorname{sg}(\sigma).$$

Por ende

$$\operatorname{sg}(\sigma)\prod_{i=1}^{n} A_{i\sigma(i)} + \operatorname{sg}(\tilde{\sigma})\prod_{i=1}^{n} A_{i\tilde{\sigma}(i)} = 0.$$

Así, los sumandos en la fórmula

$$\det A = \sum_{\sigma\in S_n} \operatorname{sg}(\sigma)\prod_{i=1}^{n} A_{i\sigma(i)}$$

se van cancelando de a pares, de donde se concluye que $\det A = 0$. Para justificar este último paso observemos que S_n se descompone como la unión disjunta de las permutaciones pares y las permutaciones impares: $S_n = S_n^{\text{par}} \cup S_n^{\text{impar}}$. Más aún, dada la transposición $\tau_{k,j}$, los subconjuntos de transposiciones pares e impares se relacionan de la siguiente forma: $\sigma \in S_n^{\text{par}} \iff \tau_{k,j}\circ\sigma = \tilde{\sigma} \in S_n^{\text{impar}}$.

Además, es claro que $\tau_{k,j} \circ \sigma_1 = \tau_{k,j} \circ \sigma_2$ si y sólo si $\sigma_1 = \sigma_2$. Luego,

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \mathrm{sg}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} \\
&= \sum_{\sigma \in S_n^{\text{par}}} \mathrm{sg}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} + \sum_{\sigma \in S_n^{\text{impar}}} \mathrm{sg}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} \\
&= \sum_{\sigma \in S_n^{\text{par}}} \mathrm{sg}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} + \sum_{\sigma \in S_n^{\text{par}}} \mathrm{sg}(\tilde{\sigma}) \prod_{i=1}^{n} A_{i\tilde{\sigma}(i)} \\
&= \sum_{\sigma \in S_n^{\text{par}}} \left[\mathrm{sg}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} + \mathrm{sg}(\tilde{\sigma}) \prod_{i=1}^{n} A_{i\tilde{\sigma}(i)} \right] = 0.
\end{aligned}$$

Finalmente, si A tiene dos filas iguales, entonces A^t tiene dos columnas iguales. Por lo que acabamos de ver, y usando la Proposición 2.14, tenemos que $\det A = \det A^t = 0$. □

Notación. En lo que sigue usaremos la siguiente convención para describir una matriz en términos de sus columnas o filas. En primer lugar si $C_1, C_2, \ldots, C_n \in \mathbb{F}^{n\times 1}$ son vectores columna, denotaremos por

$$A = (C_1\, C_2\, \cdots\, C_n)$$

la matriz cuyas columnas son $C_1, C_2, \ldots, C_n$. En tanto que si $F_1, F_2, \ldots, F_n \in \mathbb{F}^{1\times n}$ son vectores fila, denotaremos por

$$A = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{pmatrix}$$

la matriz cuyas filas son $F_1, F_2, \ldots, F_n$.

Proposición 2.16. *Sean $C_1, C_2, \ldots, C_n \in \mathbb{F}^{n\times 1}$ vectores columna y sea $\alpha \in \mathbb{F}$. Entonces*

$$\det(C_1\, \cdots\, \alpha C_k\, \cdots\, C_n) = \alpha \det(C_1\, \cdots\, C_k\, \cdots\, C_n).$$

En otras palabras, si $A \in \mathbb{F}^{n\times n}$ es la matriz cuyas columnas son $C_1, C_2, \ldots, C_n$ y A' es la matriz que se obtiene de A multiplicando la k-ésima columna por α, entonces

$$\det A' = \alpha \det A.$$

Proof. Observemos que

$$A'_{ij} = \begin{cases} A_{ij} & \text{si } j \neq k \\ \alpha A_{ik} & \text{si } j = k. \end{cases}$$

Luego, dada $\sigma \in S_n$, se tiene

$$A'_{1\sigma(1)} A'_{2\sigma(2)} \cdots A'_{n\sigma(n)} = \alpha A_{1\sigma(1)} A_{2\sigma(2)} \cdots A_{n\sigma(n)}$$

y por tanto

$$\det A = \sum_{\sigma \in S_n} \mathrm{sg}(\sigma) \prod_{i=1}^{n} A'_{i\sigma(i)} = \alpha \sum_{\sigma \in S_n} \mathrm{sg}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} = \alpha \det A.$$ □

Usando la proposición anterior, pueden probarse los siguientes corolarios:

Corolario 2.17. *Sea $A \in \mathbb{F}^{n\times n}$ y $\alpha \in \mathbb{F}$, se verifica:*

1. $\det(\alpha A) = \alpha^n \det A$;

2. *Si $A \in \mathbb{F}^{n\times n}$ tiene una columna (o fila) nula, entonces* $\det A = 0$.

Demostración : *Ejercicio.*

Corolario 2.18. *Sean $A_1, \dots, A_n \in \mathbb{F}^{1\times n}$ vectores fila y sea $\alpha \in \mathbb{F}$. Entonces*

$$\det \begin{pmatrix} A_1 \\ \vdots \\ \alpha A_k \\ \vdots \\ A_n \end{pmatrix} = \alpha \det \begin{pmatrix} A_1 \\ \vdots \\ A_k \\ \vdots \\ A_n \end{pmatrix}.$$

En otras palabras, si $A \in \mathbb{F}^{n\times n}$ es la matriz cuyas filas son $A_1, A_2, \dots, A_n$ y A' es la matriz que se obtiene de A multiplicando la k-ésima fila por α, entonces

$$\det A' = \alpha \det A.$$

Demostración : *Ejercicio (aplicar la Proposición*2.16 *a la matriz A^t).*

Ejemplo 2.19. Se pueden usar los resultados anterior para calcular el determinante de una matriz diagonal sin usar la definición:

$$\det \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} = a \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} = ab \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{pmatrix} \qquad = abc \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = abc.$$

Proposición 2.20. *Sea $A \in \mathbb{F}^{n\times n}$ y sean $A_1, \dots, A_n \in \mathbb{F}^{n\times 1}$ las columnas de A. Supongamos que $A_k = B_k + C_k$ y sean las matrices*

$$B = (A_1 \cdots A_{k-1}\, B_k\, A_{k+1} \cdots A_n), \qquad C = (A_1 \cdots A_{k-1}\, C_k\, A_{k+1} \cdots A_n).$$

Entonces

$$\det A = \det B + \det C.$$

Proof. Denotemos

$$B_k = \begin{pmatrix} B_{1k} \\ B_{2k} \\ \vdots \\ B_{nk} \end{pmatrix}, \qquad C_k = \begin{pmatrix} C_{1k} \\ C_{2k} \\ \vdots \\ C_{nk} \end{pmatrix},$$

de donde sigue que

$$B_{ij} = \begin{cases} A_{ij} & \text{si } j \neq k \\ B_{ik} & \text{si } j = k, \end{cases} \qquad C_{ij} = \begin{cases} A_{ij} & \text{si } j \neq k \\ C_{ik} & \text{si } j = k. \end{cases}$$

Luego, si $\sigma \in S_n$, tenemos que

$$A_{1\sigma(1)}A_{2\sigma(2)}\cdots A_{n\sigma(n)} = B_{1\sigma(1)}B_{2\sigma(2)}\cdots B_{n\sigma(n)} + C_{1\sigma(1)}C_{2\sigma(2)}\cdots C_{n\sigma(n)}$$

pues siempre existe un $\ell \in [\![1, n]\!]$ tal que $\sigma(\ell) = k$, y así $A_{\ell\sigma(\ell)} = B_{\ell\sigma(\ell)} + C_{\ell\sigma(\ell)}$. Finalmente

$$\begin{aligned} \det A &= \sum_{\sigma\in S_n} \operatorname{sg}(\sigma) \prod_{i=1}^n A_{i\sigma(i)} = \sum_{\sigma\in S_n} \operatorname{sg}(\sigma) \left[\prod_{i=1}^n B_{i\sigma(i)} + \prod_{i=1}^n C_{i\sigma(i)} \right] \\ &= \sum_{\sigma\in S_n} \operatorname{sg}(\sigma) \prod_{i=1}^n B_{i\sigma(i)} + \sum_{\sigma\in S_n} \operatorname{sg}(\sigma) \prod_{i=1}^n C_{i\sigma(i)} = \det B + \det C. \end{aligned}$$

□

A partir del resultado anterior pueden probarse los siguientes corolarios:

Corolario 2.21. *Sea* $A \in \mathbb{F}^{n\times n}$ *y sean* $A_1, \ldots, A_n \in \mathbb{F}^{1\times n}$ *las filas* A. *Supongamos que* $A_k = B_k + C_k$ *y sean las matrices* B *cuyas filas son* $A_1, \ldots, A_{k-1}, B_k, A_{k+1}, \ldots A_n$, *y* C *cuyas filas son* $A_1, \ldots, A_{k-1}, C_k, A_{k+1}, \ldots A_n$. *Entonces*

$$\det A = \det B + \det C.$$

Demostración : Ejercicio.

Corolario 2.22. *Sean* $A_1, \ldots, A_n \in \mathbb{F}^{n\times 1}$ *las columnas* A, *entonces*

$$\det(A_1 \cdots \underbrace{A_j}_{columna\ i} \cdots \underbrace{A_i}_{columna\ j} \cdots A_n) = -\det(A_1 \cdots \underbrace{A_i}_{columna\ i} \cdots \underbrace{A_j}_{columna\ j} \cdots A_n).$$

En otras palabras, si $A \in \mathbb{F}^{n\times n}$ *y* A' *es la matriz que se obtiene de* A *intercambiando la* i*-ésima columna con la* j*-ésima columna, con* $i \neq j$, *entonces*

$$\det A' = -\det A.$$

Proof. Consideremos la matriz

$$\tilde{A} = (A_1 \cdots (A_i + A_j) \cdots (A_i + A_j) \cdots A_n),$$

o sea, $\tilde{A}$ es la matriz cuyas columnas i y j son iguales a $A_i + A_j$ y por lo tanto tiene determinante nulo. Además, usando las Proposiciones 2.20 y 2.15 obtenemos que

$$\begin{aligned}
0 = \det \tilde{A} &= \det(A_1 \cdots (A_i + A_j) \cdots (A_i + A_j) \cdots A_n) \\
&= \det(A_1 \cdots A_i \cdots (A_i + A_j) \cdots A_n) + \det(A_1 \cdots A_j \cdots (A_i + A_j) \cdots A_n) \\
&= \det(A_1 \cdots A_i \cdots A_i \cdots A_n) + \det(A_1 \cdots A_i \cdots A_j \cdots A_n) \\
&\qquad + \det(A_1 \cdots A_j \cdots A_i \cdots A_n) + \det(A_1 \cdots A_j \cdots A_j \cdots A_n) \\
&= \det(A_1 \cdots A_i \cdots A_j \cdots A_n) + \det(A_1 \cdots A_j \cdots A_i \cdots A_n),
\end{aligned}$$

de donde se concluye que

$$\det(A_1 \cdots A_j \cdots A_i \cdots A_n) = -\det(A_1 \cdots A_i \cdots A_j \cdots A_n). \qquad \square$$

Ejercicio 2.23. Enunciar y demostrar el resultado análogo para el intercambio de filas.

Proposición 2.24. *Sea* $A \in \mathbb{F}^{n\times n}$ *y sea* A' *la matriz que se obtiene de* A *sumando a la* j*-ésima columna un múltiplo de la* k*-ésima columna, con* $j \neq k$. *Entonces*

$$\det A' = \det A.$$

Proof. Supongamos sin perder generalidad que $j < k$. Denotemos por $A_1, \ldots, A_n$ las columnas de la matriz A. Entonces

$$A' = (A_1 \cdots A_{j-1}\,(A_j + \alpha A_k)\,A_{j+1} \cdots A_k \cdots A_n)$$

para cierto $\alpha \in \mathbb{F}$. Luego, usando las Proposiciones 2.15, 2.16 y 2.20 tenemos que

$$\begin{aligned}
\det A' &= \det(A_1 \cdots A_{j-1}\,(A_j + \alpha A_k)\,A_{j+1} \cdots A_k \cdots A_n) \\
&= \det(A_1 \cdots A_{j-1}\,A_j\,A_{j+1} \cdots A_k \cdots A_n) \\
&\qquad + \det(A_1 \cdots A_{j-1}\,\alpha A_k\,A_{j+1} \cdots A_k \cdots A_n) \\
&= \det A + \alpha \det(A_1 \cdots A_{j-1}\,A_k\,A_{j+1} \cdots A_k \cdots A_n) \\
&= \det A. \qquad \square
\end{aligned}$$

Ejercicio 2.25. Enunciar y demostrar el resultado análogo para operaciones sobre las filas de la matriz.

Con los resultados anteriores el calculo del determinante se simplifica. En efecto, veamos un ejemplo de esto.

Ejemplo 2.26. Calculemos el determinante de la matriz

$$A = \begin{pmatrix} 0 & 1 & -2 & 0 \\ 1 & 4 & 0 & 3 \\ 2 & 1 & -2 & 5 \\ 1 & 0 & 2 & 3 \end{pmatrix}.$$

La idea es "mejorar" la matriz A aplicando las operaciones de intercambio de filas o columnas y sumando o restando a una fila/columna un múltiplo de otra fila/columna. Como ya demostramos, al hacer intercambios de filas o columnas, el determinante cambia de signo, en tanto que si a una fila o columna le sumamos un múltiplo de otra, el determinante no cambia. Para justificar las operaciones que usamos en cada paso usaremos la siguiente notación. Para el intercambio de filas o columnas escribiremos, por ejemplo, $\mathtt{f1} \leftrightarrow \mathtt{f2}$ para indicar que intercambiamos la fila 1 con la fila 2, en tanto que $\mathtt{c2} \leftrightarrow \mathtt{c4}$ indica que intercambiamos la columna 2 con la columna 4. Por otro lado, la notación $\mathtt{f3} \to \mathtt{f3} - 2\mathtt{f1}$ significa que reemplazamos la fila 3 por la fila 3 menos 2 veces la fila 1. Aclarado esto, tenemos que

$$\begin{aligned}
\det A &\overset{\mathtt{f1}\leftrightarrow\mathtt{f2}}{=} -\det \begin{pmatrix} 1 & 4 & 0 & 3 \\ 0 & 1 & -2 & 0 \\ 2 & 1 & -2 & 5 \\ 1 & 0 & 2 & 3 \end{pmatrix} \overset{\mathtt{f3}\to\mathtt{f3}-2\mathtt{f1}}{=} -\det \begin{pmatrix} 1 & 4 & 0 & 3 \\ 0 & 1 & -2 & 0 \\ 0 & -7 & -2 & -1 \\ 1 & 0 & 2 & 3 \end{pmatrix} \\
&\overset{\mathtt{f4}\to\mathtt{f4}-\mathtt{f1}}{=} -\det \begin{pmatrix} 1 & 4 & 0 & 3 \\ 0 & 1 & -2 & 0 \\ 0 & -7 & -2 & -1 \\ 0 & -4 & 2 & 0 \end{pmatrix} \overset{\mathtt{c2}\leftrightarrow\mathtt{c4}}{=} \det \begin{pmatrix} 1 & 3 & 0 & 4 \\ 0 & 0 & -2 & 1 \\ 0 & -1 & -2 & -7 \\ 0 & 0 & 2 & -4 \end{pmatrix} \\
&\overset{\mathtt{f2}\leftrightarrow\mathtt{f3}}{=} -\det \begin{pmatrix} 1 & 3 & 0 & 4 \\ 0 & -1 & -2 & -7 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 2 & -4 \end{pmatrix} \overset{\mathtt{f4}\leftrightarrow\mathtt{f4}+\mathtt{f3}}{=} -\det \begin{pmatrix} 1 & 3 & 0 & 4 \\ 0 & -1 & -2 & -7 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -3 \end{pmatrix}
\end{aligned}$$

Luego hemos mostrado que el determinante de A es igual al determinante de una matriz triangular superior, los cuales son muy fáciles de calcular usando la siguiente proposición.

Proposición 2.27. *Si $A \in \mathbb{F}^{n\times n}$ es triangular superior, entonces*

$$\det A = A_{11}A_{22}\cdots A_{nn}.$$

Proof. Una matriz triangular superior tiene la forma

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ 0 & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn} \end{pmatrix},$$

es decir $A_{ij} = 0$ si $i > j$. Ahora bien, para $\sigma \in S_n$, $\sigma \neq \mathrm{id}$, siempre existe un i tal que $i > \sigma(i)$. En efecto, si esto no sucediera, se tendría $n \leq \sigma(n) \leq n$, de donde sigue $\sigma(n) = n$. Luego $n-1 \leq \sigma(n-1) \leq n-1$, lo cual implica $\sigma(n-1) = n-1$. Así siguiendo, se obtiene que $\sigma(i) = i$ para todo $i = 1, \ldots, n$. Absurdo, pues supusimos $\sigma \neq \mathrm{id}$. Esto dice que en la definición de $\det A$ todos los sumandos correspondientes a $\sigma \neq \mathrm{id}$ son nulos, por tanto

$$\det A = \sum_{\sigma\in S_n} \mathrm{sg}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} = \mathrm{sg}(\mathrm{id}) \prod_{i=1}^{n} A_{ii} = \prod_{i=1}^{n} A_{ii}$$

como queríamos probar. □

Ejemplo 2.26 (continuación)**.** Usando el resultado anterior concluimos que

$$\det A = -\det \begin{pmatrix} 1 & 3 & 0 & 4 \\ 0 & -1 & -2 & -7 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -3 \end{pmatrix} = -(1 \cdot (-1) \cdot (-2) \cdot (-3)) = 6.$$

2.4 Matrices inversibles

Una de las propiedades más importantes de la función determinante es la siguiente.

Teorema 2.28. *Si $A, B \in \mathbb{F}^{n\times n}$, entonces*

$$\det(AB) = (\det A)(\det B).$$

Proof. La veremos más adelante. □

Definición 2.29. Una matriz $A \in \mathbb{F}^{n\times n}$ se dice *invertible* si existe $B \in \mathbb{F}^{n\times n}$ tal que $AB = BA = I$. En caso de que exista una tal B, ésta se llama la *matriz inversa* de A y se denota por $B = A^{-1}$.

Observación 2.30. Si $A \in \mathbb{F}^{n\times n}$ es invertible, entonces la inversa es única. En efecto, supongamos que existen $B, C \in \mathbb{F}^{n\times n}$ tales que $BA = AB = I = CA = AC$. Sigue que

$$B = BI = BAC = IC = C.$$

Ejemplo 2.31. 1. La matriz identidad es invertible. ¿Por qué?

2. La matriz nula no es invertible. ¿Por qué?

Ejemplo 2.32. ¿Cuándo es invertible una matriz 2×2? Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es invertible, entonces existe una matriz $B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$ tal que

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

lo cual es equivalente al siguiente sistema de cuatro ecuaciones con cuatro incógnitas (las incógnitas son e, f, g, h):

$$ae + bg = 1 \tag{2.3}$$
$$af + bh = 0 \tag{2.4}$$
$$ce + dg = 0 \tag{2.5}$$
$$cf + dh = 1 \tag{2.6}$$

Si multiplicamos la ecuación (2.3) por d y le restamos b veces la ecuación (2.5) obtenemos

$$d(ae + bg) - b(ce + dg) = (ad - bc)e = d.$$

Trabajando análogamente con las ecuaciones (2.4) y (2.6) obtenemos que

$$\begin{aligned} d(af + bh) - b(cf + dh) &= (ad - bc)f = -b \\ a(ce + dg) - c(ae + bg) &= (ad - bc)g = -c \\ a(cf + dh) - c(af + bh) &= (ad - bc)h = a, \end{aligned}$$

de donde sigue que si $\det A = ad - bc \neq 0$ entonces el sistema tiene solución y la matriz inversa es

$$B = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Corolario 2.33. *Una matriz* $A \in \mathbb{F}^{n\times n}$ *es invertible si y sólo si* $\det A \neq 0$. *Más aún, si* A *es invertible, entonces* $\det A^{-1} = (\det A)^{-1}$.

Proof. Si A es invertible entonces existe la inversa A^{-1} de A y vale $AA^{-1} = I$. Luego por el Teorema 2.28 se tiene $1 = \det I = \det A \det A^{-1}$, por ende, tanto $\det A$ como $\det A^{-1}$ deben ser no nulos. Notar que esta ecuación también implica que $\det A^{-1} = (\det A)^{-1}$.

La recíproca la probaremos en el próximo apartado, cuando exhibamos un método para calcular la matriz inversa de una matriz A tal que $\det A \neq 0$. □

2.5 Desarrollo del determinante por filas o columnas

En este apartado presentamos un método alternativo para el cálculo del determinante. Dicho método no presenta ninguna ventaja sobre la Definición 2.10 en lo que respecta a la complejidad de cálculo, pero sí tiene importancia teórica como veremos en breve.

Definición 2.34. Sea $A \in \mathbb{F}^{n\times n}$ y sean $i, j \in [\![1, n]\!]$. Se define $A(i|j)$ como la matriz $(n-1)\times(n-1)$ que se obtiene suprimiendo la i-ésima fila y la j-ésima columna de la matriz A.

Ejemplo 2.35. Si

$$A = \begin{pmatrix} 0 & 1 & -2 & 0 \\ 1 & 4 & 0 & 3 \\ 2 & 1 & -2 & 5 \\ 1 & 0 & 2 & 6 \end{pmatrix},$$

entonces

$$A(1|1) = \begin{pmatrix} 4 & 0 & 3 \\ 1 & -2 & 5 \\ 0 & 2 & 6 \end{pmatrix}, \quad A(1|3) = \begin{pmatrix} 1 & 4 & 3 \\ 2 & 1 & 5 \\ 1 & 0 & 6 \end{pmatrix}, \quad A(3|2) = \begin{pmatrix} 0 & -2 & 0 \\ 1 & 0 & 3 \\ 1 & 2 & 6 \end{pmatrix}.$$

Más aún, el proceso se puede repetir,

$$A(1|1)(2|3) = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}, \qquad A(1|1)(2|3)(2|2) = 4.$$

A continuación presentamos una fórmula recursiva para el cálculo de la función determinante, para lo cual necesitamos hacer unas observaciones.

Observemos que el conjunto de todas las permutaciones de n elementos, puede describirse como la unión disjunta de las permutaciones que mandan el 1 a un elemento especificado $j \in [\![1, n]\!]$. Más precisamente, si llamamos

$$S_n^j = \{\sigma \in S_n : \sigma(1) = j\},$$

entonces se tiene

$$S_n = S_n^1 \cup S_n^2 \cup \cdots \cup S_n^n$$

y esta unión es disjunta, es decir, $S_n^j \cap S_n^k = \varnothing$, si $j \neq k$. Esto es, si $\sigma \in S_n$ entonces $\sigma \in S_n^{\sigma(1)}$.

Ahora bien, un elemento $\sigma \in S_n^j$ puede pensarse como una permutación de $n-1$ elementos, o sea, como un elemento de S_{n-1}. En efecto, sabemos que todos los elementos de S_n^j mandan 1 en j, luego, podemos pensar que los elementos de S_n^j permutan los números $\{1, 2, \ldots, j-1, j+1, \ldots, n\}$.

Esquemáticamente:

$$\begin{array}{ccc}
\sigma \in S_n^j & \leftrightsquigarrow & \sigma^j \in S_{n-1} \\
\sigma : [\![1, n]\!] \to [\![1, n]\!] & \leftrightsquigarrow & \sigma^j : [\![1, n-1]\!] \to [\![1, n-1]\!] \\
1 & \leftrightsquigarrow & 1 \\
2 & \leftrightsquigarrow & 2 \\
 & \vdots & \\
j-1 & \leftrightsquigarrow & j-1 \\
j+1 & \leftrightsquigarrow & j \\
 & \vdots & \\
n-1 & \leftrightsquigarrow & n-2 \\
n & \leftrightsquigarrow & n-1.
\end{array}$$

Esto puede formalizarse diciendo que a cada elemento $\sigma \in S_n^j$ le corresponde un único elemento $\sigma^j \in S_{n-1}$. Esta identificación puede resultar un poco difícil de entender, pues $\sigma \in S_n^j$ permuta los elementos $\{1, 2, \ldots, j-1, j+1, \ldots, n\}$, pero σ^j permuta los elementos $\{1, 2, \ldots, n-1\}$.

Veamos unos ejemplos para clarificar esta noción.

Ejemplo 2.36. Consideremos la permutación $\sigma \in S_7^3$ dada por $\sigma = (3, 2, 1, 6, 5, 4, 7)$. Luego $\sigma(1) = 3$ y σ reordena la 6-upla $(1, 2, 4, 5, 6, 7)$ en $(2, 1, 6, 5, 4, 7)$. O sea, el 1er elemento va al 2do lugar, el 2do elemento va al 1er lugar, el 3er elemento va al 5to lugar, el 4to elemento va al 4to lugar, el 5to elemento va al 3er lugar y el 6to elemento va al 6to lugar. Luego, la identificación en este caso sería

$$\sigma = (3, 2, 1, 6, 5, 4, 7) \longleftrightarrow \sigma^3 = (2, 1, 5, 4, 3, 6).$$

Ejemplo 2.37. Otro ejemplo: ¿qué permutación $\sigma \in S_7^3$ corresponde a la permutación $\mathrm{id} \in S_6$? En este caso debería ser $\sigma(1) = 3$ pero σ tiene que mantener el orden de los restantes elementos $(1, 2, 4, 5, 6, 7)$. Luego

$$\sigma = (3, 1, 2, 4, 5, 6, 7) \longleftrightarrow \sigma^3 = (1, 2, 3, 4, 5, 6) = \mathrm{id}\,.$$

La pregunta clave en la demostración del teorema siguiente será: dada $\sigma \in S_n^j$, ¿cuál es el signo de la permutación $\sigma^j \in S_{n-1}$? Para responder esta pregunta, observemos que para calcular $\mathrm{sg}(\sigma)$ uno debe expresar σ como una composición de trasposiciones, o dicho de otra manera, necesitamos contar la cantidad de intercambios de dos elementos necesarios para llegar de $(1, 2, \ldots, n)$ a $(\sigma(1), \sigma(2), \ldots, \sigma(n))$. En tanto que, al ser $\sigma(1) = j$, para calcular el signo de σ^j uno tiene contar la cantidad de intercambios de dos elementos necesarios para pasar de $(1, 2, \ldots, j-1, j+1, \ldots, n)$ a $(\sigma(2), \sigma(3), \ldots, \sigma(n))$. Ahora bien, la relación entre $\mathrm{sg}(\sigma)$ y $\mathrm{sg}(\sigma^j)$ viene dada como sigue: para pasar de $(1, 2, \ldots, n)$ a

$$(\sigma(1), \sigma(2), \ldots, \sigma(n)) = (j, \sigma(2), \ldots, \sigma(n))$$

uno puede hacer primero $j-1$ intercambios para pasar de $(1, 2, \ldots, n)$ a $(j, 2, 3, \ldots, n)$ y luego hacer los intercambios necesarios sobre los últimos elementos para transformar $(j, 2, 3, \ldots, n)$ en $(j, \sigma(2), \sigma(3), \ldots, \sigma(n))$. Por tanto

$$\mathrm{sg}(\sigma) = (-1)^{j-1}\,\mathrm{sg}(\sigma^j) = (-1)^{1+j}\,\mathrm{sg}(\sigma^j).$$

Teorema 2.38 (Desarrollo del determinante por la primera fila)**.** *Si $A \in \mathbb{F}^{n\times n}$, entonces*

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \det A(1|j).$$

Proof. Para probar el teorema usaremos a lo largo de los cálculos las observaciones que hemos hecho anteriormente.

$$
\begin{aligned}
\det A = \sum_{\sigma \in S_n} \operatorname{sg}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} &= \sum_{j=1}^{n} \sum_{\sigma \in S_n^j} \operatorname{sg}(\sigma) \prod_{i=1}^{n} A_{i\sigma(i)} \\
&= \sum_{j=1}^{n} \sum_{\sigma \in S_n^j} \operatorname{sg}(\sigma) A_{1j} \prod_{i=2}^{n} A_{i\sigma(i)} \\
&= \sum_{j=1}^{n} \sum_{\sigma^j \in S_{n-1}} (-1)^{1+j} \operatorname{sg}(\sigma^j) A_{1j} \prod_{i=1}^{n-1} A(1|j)_{i\sigma^j(i)} \\
&= \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \sum_{\sigma^j \in S_{n-1}} \operatorname{sg}(\sigma^j) \prod_{i=1}^{n-1} A(1|j)_{i\sigma^j(i)} \\
&= \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \det A(1|j).
\end{aligned}
$$
□

Veamos algunos casos particulares de aplicación del teorema anterior.

Ejemplo 2.39. Calculemos, usando el Teorema 2.38, el determinante de

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

En efecto,

$$\det A = a_{11} \det A(1|1) - a_{12} \det A(1|2) = a_{11}a_{22} - a_{12}a_{21}.$$

También podemos calcular el determinante de una matriz 3×3 usando determinantes de matrices 2×2. Si

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

entonces

$$
\begin{aligned}
\det A &= a_{11} \det A(1|1) - a_{12} \det A(1|2) + a_{13} \det A(1|3) \\
&= a_{11} \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \\
&= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\
&= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.
\end{aligned}
$$

Observación 2.40. En algunos libros de texto el determinante se define recursivamente usando la fórmula del Teorema 2.38. Es decir, para matrices $A \in \mathbb{F}^{1\times 1} = \mathbb{F}$, se define $\det A = A$ y luego, dado $n \in \mathbb{N}$, se define el determinante de una matriz $A \in \mathbb{F}^{(n+1)\times(n+1)}$ como

$$\det A = \sum_{j=1}^{n+1} (-1)^{1+j} A_{1j} \det A(1|j).$$

Como el determinante de una matriz es igual al determinante de su transpuesta y el determinante cambia de signo si en una matriz intercambiamos filas o columnas, se pueden deducir fórmulas para el desarrollo del determinante por cualquier fila o columna de la matriz. Queda como ejercicio hacer las demostraciones de los siguientes resultados.

Corolario 2.41 (Desarrollo del determinante por la i-ésima fila)**.** *Sean $A \in \mathbb{F}^{n\times n}$, e $i \in [\![1, n]\!]$. Entonces*

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det A(i|j).$$

Corolario 2.42 (Desarrolo del determinante por la j-ésima columna)**.** *Sean* $A \in \mathbb{F}^{n\times n}$*, y* $j \in [\![1, n]\!]$*. Entonces*

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det A(i|j).$$

A continuación mostramos un método para calcular la inversa de una matriz A (si es que ésta existe).

Definición 2.43. Sea $A \in \mathbb{F}^{n\times n}$, el escalar $C_{ij} = (-1)^{i+j} \det A(i|j)$ se llama el *cofactor* i, j de A. La matriz $C = (C_{ij})$ se llama *matriz de los cofactores* de A.

Definición 2.44. La *matriz adjunta* de $A \in \mathbb{F}^{n\times n}$, denotada por $\operatorname{adj} A$, es la matriz transpuesta de la matriz de los cofactores de A. Es decir, el lugar i, j de $\operatorname{adj} A$ está dado por el cofactor j, i de A:

$$(\operatorname{adj} A)_{ij} = (-1)^{i+j} \det A(j|i).$$

Teorema 2.45. *Si* $A \in \mathbb{F}^{n\times n}$*, entonces*

$$A \operatorname{adj} A = (\operatorname{adj} A)A = (\det A)I$$

Proof. Debemos probar que $(A \operatorname{adj} A)_{ij} = ((\operatorname{adj} A)A)_{ij} = (\det A)\delta_{ij}$, en donde δ_{ij} es la delta de Kronecker (vale 1 si $i = j$ y 0 si $i \neq j$). Probaremos que $(A \operatorname{adj} A)_{ij} = \delta_{ij}$, dejando como ejercicio el comprobar que $((\operatorname{adj} A)A)_{ij} = \delta_{ij}$. Para ello, calculamos la matriz producto $A \operatorname{adj} A$ por definición. En primer lugar, si $i \neq j$ tenemos que

$$(A \operatorname{adj} A)_{ij} = \sum_{k=1}^{n} A_{ik}(\operatorname{adj} A)_{kj} = \sum_{k=1}^{n} (-1)^{k+j} A_{ik} \det A(j|k) = \det A',$$

usando el Corolario 2.41, en donde A' es la matriz tal que $(A)_{ik} = (A')_{jk}$ para todo $k = 1, \ldots, n$ y tiene todas sus otras entradas iguales a las entradas de A. Pero en A', la fila i es igual a la fila j, luego, por la Proposición 2.15, se tiene que $0 = \det A' = (A \operatorname{adj} A)_{ij}$.

Finalmente,

$$(A \operatorname{adj} A)_{ii} = \sum_{k=1}^{n} A_{ik}(\operatorname{adj} A)_{ki} = \sum_{k=1}^{n} (-1)^{i+k} A_{ik} \det A(i|k) = \det A,$$

nuevamente por el Corolario 2.41, lo cual concluye la prueba del teorema. □

Usando el teorema anterior podemos completar la prueba del Corolario 2.33. De hecho, podemos encontrar una fórmula para la matriz inversa de una matriz invertible.

Corolario 2.46. *Sea* $A \in \mathbb{F}^{n\times n}$ *tal que* $\det A \neq 0$*. Entonces* A *es invertible y vale*

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A.$$

Proof. Es inmediato del teorema anterior que si $\det A \neq 0$, entonces

$$\left(\frac{1}{\det A} \operatorname{adj} A\right) A = A \left(\frac{1}{\det A} \operatorname{adj} A\right) = I.$$ □

Se dice que una matriz $A \in \mathbb{F}^{n\times n}$ tiene una *inversa a izquierda* (resp. *a derecha*) si existe $B \in \mathbb{F}^{n\times n}$ tal que $BA = I$ (resp. $AB = I$).

Corolario 2.47. *Si* $A \in \mathbb{F}^{n\times n}$ *tiene una inversa a izquierda (resp. a derecha) entonces* A *es invertible.*

Proof. Supongamos que existe $B \in \mathbb{F}^{n\times n}$ tal que $BA = I$. Sigue que

$$1 = \det(BA) = (\det B)(\det A)$$

y en particular $\det A \neq 0$. Luego A es invertible por el corolario anterior. El caso en el que A admite una inversa a derecha es análogo. □

Ejemplo 2.48. Consideremos la matriz

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

La matriz de los cofactores de A es

$$\begin{pmatrix} (-1)^{1+1} \det A(1|1) & (-1)^{1+2} \det A(1|2) \\ (-1)^{2+1} \det A(2|1) & (-1)^{2+2} \det A(2|2) \end{pmatrix} = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

Luego, si $ad - bc \neq 0$, entonces A es invertible y su inversa viene dada por

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Ejemplo 2.49. Decidir si la matriz

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 2 & 5 \\ 2 & -1 & 4 \end{pmatrix}$$

es invertible y calcular su inversa en caso afirmativo. Calculamos primero la matriz de los cofactores de A, para esto, necesitamos calcular $\det A(i|j)$ para todos $i, j = 1, 2, 3$.

$$\det A(1|1) = \begin{vmatrix} 2 & 5 \\ -1 & 4 \end{vmatrix} = 13, \qquad \det A(1|2) = \begin{vmatrix} 0 & 5 \\ 2 & 4 \end{vmatrix} = -10, \qquad \det A(1|3) = \begin{vmatrix} 0 & 2 \\ 2 & -1 \end{vmatrix} = -4,$$

$$\det A(2|1) = \begin{vmatrix} 2 & -1 \\ -1 & 4 \end{vmatrix} = 7, \qquad \det A(2|2) = \begin{vmatrix} 1 & -1 \\ 2 & 4 \end{vmatrix} = 6, \qquad \det A(2|3) = \begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix} = -5,$$

$$\det A(3|1) = \begin{vmatrix} 2 & -1 \\ 2 & 5 \end{vmatrix} = 12, \qquad \det A(3|2) = \begin{vmatrix} 1 & -1 \\ 0 & 5 \end{vmatrix} = 5, \qquad \det A(3|3) = \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} = 2.$$

Luego la matriz de los cofactores de A y la matriz adjunta están dadas por

$$C = \begin{pmatrix} 13 & 10 & -4 \\ -7 & 6 & 5 \\ 12 & -5 & 2 \end{pmatrix}, \qquad \operatorname{adj} A = \begin{pmatrix} 13 & -7 & 12 \\ 10 & 6 & -5 \\ -4 & 5 & 2 \end{pmatrix}.$$

Para chequear si A es invertible deberíamos calcular $\det A$, pero observemos que con lo ya obtenido, esta cuenta puede hacerse como

$$A(\operatorname{adj} A) = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 2 & 5 \\ 2 & -1 & 4 \end{pmatrix} \begin{pmatrix} 13 & -7 & 12 \\ 10 & 6 & -5 \\ -4 & 5 & 2 \end{pmatrix} = \begin{pmatrix} 37 & 0 & 0 \\ 0 & 37 & 0 \\ 0 & 0 & 37 \end{pmatrix}.$$

Luego $\det A = 37 \neq 0$ y por lo tanto A es invertible con inversa

$$A^{-1} = \frac{1}{37} \begin{pmatrix} 13 & -7 & 12 \\ 10 & 6 & -5 \\ -4 & 5 & 2 \end{pmatrix}.$$

El método que desarrollamos para el cálculo de la inversa de una matriz resulta muy tedioso de aplicar y nada eficiente para matrices de tamaño grande. En la próxima unidad desarrollaremos un métodos más eficientes, tanto para el calculo del determinante como para encontrar la inversa de una matriz invertible.

Comentario. El *permanente* de una matriz $A \in \mathbb{F}^{n \times n}$ se define de manera análoga al determinante, sumando todos los factores que se pueden armar eligiendo un elemento en cada fila recorriendo todas las columnas, pero sin tener en cuenta el signo de la permutación de las columnas, es decir,

$$\operatorname{perm} A = \sum_{\sigma \in S_n} \prod_{i=1}^{n} A_{i\sigma(i)}.$$

Por ejemplo,

$$\operatorname{perm}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad + bc,$$

$$\operatorname{perm}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + afh + bdi + bfg + cdh + ceg.$$

Con las mismas técnicas que aprendimos en esta unidad, uno puede probar que el permanente comparte algunas propiedades con el determinante, por ejemplo $\operatorname{perm} A = \operatorname{perm} A^t$, o que si uno multiplica la columna k de A por el escalar α y llama A' a esta nueva matriz, entonces $\operatorname{perm} A' = \alpha \operatorname{perm} A$. Pero otras propiedades ya no son válidas, por ejemplo, si una matriz A tiene dos columnas iguales, no necesariamente vale $\operatorname{perm} A = 0$. Tampoco es cierto que $\operatorname{perm}(AB) = (\operatorname{perm} A)(\operatorname{perm} B)$ para todas $A, B \in \mathbb{F}^{n\times n}$. En efecto,

$$4 = \operatorname{perm}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \operatorname{perm}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \neq \operatorname{perm}\left(\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right) = \operatorname{perm}\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} = 8.$$

Es la falta de estas propiedades la que hace que el permanente, si bien tiene una definición similar al determinante, sea mucho más difícil de calcular. En efecto, si uno utiliza la definición, tanto para el permanente como para el determinante deben realizarse $n!n$ operaciones sobre los coeficientes de la matriz (porque tenemos $n!$ sumandos de productos de n elementos). Sin embargo, para calcular el determinante de una matriz uno puede realizar operaciones por fila y columna para transformarla en una matriz triangular superior. Puede verse (como lo probaremos en la próxima unidad y quizás ya podamos intuirlo de los ejemplos en esta unidad) que la cantidad necesaria de operaciones para pasar de una matriz arbitraria a una matriz triangular superior es del orden de n^3. Es por esto que se dice que el cálculo del determinante de una matriz es un problema que puede resolverse en tiempo polinomial, o que tiene complejidad $O(n^3)$, pues n^3 es un polinomio de grado 3 en el número de operaciones necesarias para calcular el determinante. Estas consideraciones no son ciertas para el cálculo del permanente, de hecho se cree que el problema de calcular el permanente de una matriz no puede resolverse en tiempo polinomial.

Observar que $n!n$ es mucho más grande que n^3 cuando n es suficientemente grande.

n	n^3	$n!n$
1	1	1
2	8	4
3	27	18
4	64	96
5	125	600
6	216	4320
7	343	35280
8	512	322560
9	729	3265920
10	1000	36288000
15	3375	19615115520000
20	8000	48658040163532800000