Redes Neuronales

Una Introducción Moderna al Deep Learning

Alejandro Sánchez Yalí

Math & Code Community

15 de noviembre de 2025

Índice general

Índice de figuras

Índice de cuadros

Capítulo 1

Introducción

La inteligencia es la capacidad de adaptarse al cambio.

Albert Einstein

1.1 Preliminares

Definición 1.1 (Grupo simétrico de permutaciones). Sea X un conjunto finito de n elementos. El grupo simétrico de permutaciones, denotado por S_n, es el conjunto de todas las biyecciones de X sobre sí mismo con la operación de composición de funciones.

Definición 1.2 (Signo de una permutación). Sea $\sigma \in S_n$ una permutación. El signo de σ, denotado por $\operatorname{sgn}(\sigma)$, es 1 si σ es par (y se puede expresar como un producto de un número par de transposiciones) y -1 si σ es impar (y se puede expresar como un producto de un número impar de transposiciones).

El signo de una permutación σ se puede calcular contando el número de inversiones en la permutación. Una inversión es un par de índices (i, j) con $i < j$ tal que $\sigma(i) > \sigma(j)$. Si el número total de inversiones es par, entonces $\operatorname{sgn}(\sigma) = 1$; si es impar, entonces $\operatorname{sgn}(\sigma) = -1$.

El proceso paso a paso para determinar el signo de una permutación es el siguiente:

1. Escribe la permutación en forma de lista, por ejemplo, $(\sigma(1), \sigma(2), \ldots, \sigma(n))$.
2. Para cada par de índices $i < j$, compara los valores $\sigma(i)$ y $\sigma(j)$.
3. Cuenta cuántos pares cumplen que $\sigma(i) > \sigma(j)$. Cada uno de estos pares es una inversión.

4. Si el número total de inversiones es par, la permutación es par y $\operatorname{sgn}(\sigma) = 1$.
5. Si el número total de inversiones es impar, la permutación es impar y $\operatorname{sgn}(\sigma) = -1$.

Ejemplo 1.1 (Permutaciones y su signo)*.* Consideremos el conjunto $X = \{1, 2, 3\}$. Las permutaciones de X son:

- $\sigma_1 = (1, 2, 3)$ (identidad)
- $\sigma_2 = (1, 3, 2)$
- $\sigma_3 = (2, 1, 3)$
- $\sigma_4 = (2, 3, 1)$
- $\sigma_5 = (3, 1, 2)$
- $\sigma_6 = (3, 2, 1)$

El signo de cada permutación es:

- $\operatorname{sgn}(\sigma_1) = 1$ (par, 0 inversiones)
- $\operatorname{sgn}(\sigma_2) = -1$ (impar, 1 inversión: $(3, 2)$)
- $\operatorname{sgn}(\sigma_3) = -1$ (impar, 1 inversión: $(2, 1)$)
- $\operatorname{sgn}(\sigma_4) = 1$ (par, 2 inversiones: $(2, 1)$ y $(3, 1)$)
- $\operatorname{sgn}(\sigma_5) = 1$ (par, 2 inversiones: $(3, 1)$ y $(3, 2)$)
- $\operatorname{sgn}(\sigma_6) = -1$ (impar, 3 inversiones: $(3, 2)$, $(3, 1)$ y $(2, 1)$)

Capítulo 2

Álgebra Lineal

La inteligencia es la capacidad de adaptarse al cambio.

Albert Einstein

El álgebra lineal constituye el fundamento matemático de la inteligencia artificial, proporcionando el lenguaje y las herramientas esenciales para representar y manipular datos, modelos y algoritmos. Este capítulo presenta una revisión sistemática de conceptos necesarios para el trabajo con Vectores y Matrizes.

2.1 Matrices

Definición 2.1 (Matriz)**.** Una **Matriz** de orden $m \times n$ sobre un campo $\mathbb{K}$ es un arreglo rectangular de elementos $a_{ij} \in \mathbb{K}$ (llamados entradas o coeficientes), donde $i = 1, \ldots, m$ y $j = 1, \ldots, n$, organizados en m filas y n columnas. Se denota como:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

o de forma abreviada como $A = (a_{ij})_{m \times n}$.

El conjunto de todas las matrices de orden $m \times n$ con coeficientes en $\mathbb{K}$ se denota por $M_{m\times n}(\mathbb{K})$. Cuando el número de filas coincide con el número de columnas, se dice que la matriz es **Matriz cuadrada**. El conjunto de todas las matrices cuadradas de orden n con coeficientes en $\mathbb{K}$ se denota por $\mathcal{M}_n(\mathbb{K})$, que es equivalente a escribir $\mathcal{M}_{n\times n}(\mathbb{K})$.

Cada posición en una matriz está identificada por dos índices: uno para la fila y otro para la columna. El elemento que se encuentra en la intersección de la fila i-ésima y la columna j-ésima se denomina **entrada** (i, j)**-ésima** de la matriz A. Por convención, este elemento se denota como a_{ij}, lo que permite representar de manera compacta a la matriz A mediante la notación (a_{ij}).

Definición 2.2 (Columna y fila de una matriz)**.** Sea $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$. Dado $j \in \{1, \dots, n\}$ la matriz

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \in \mathcal{M}_{m\times 1}(\mathbb{K})$$

se llama **columna** j**-ésima** de A y se denota como $A_{c(j)}$, y dado $i \in \{1, \dots, m\}$ la matriz

$$\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} \in \mathcal{M}_{1\times n}(\mathbb{K})$$

se denomina **fila** i**-ésima** de A y se denota como $A_{r(i)}$.

A continuación se presentan algunos ejemplos de matrices.

Ejemplo 2.1. La **matriz nula** $0 \in \mathcal{M}_{m\times n}(\mathbb{K})$ es aquella con m filas y n columnas cuyas entradas son todas iguales a 0. En algunas ocasiones escribiremos $0_{m\times n}$ para denotar a la matriz nula de orden $m \times n$.

Ejemplo 2.2. Se dice que una matriz cuadrada $D = (d_{ij}) \in \mathcal{M}_n(\mathbb{K})$ es **diagonal** si $d_{ij} = 0$ para todo $i \neq j$. En otras palabras, una matriz diagonal es aquella que solo tiene elementos no nulos en su diagonal principal. Utilizaremos la notación

$$\operatorname{diag}(\lambda_1, \dots, \lambda_n),$$

con $\lambda_i \in \mathbb{K}$, $i = 1, \dots, n$, para denotar la matriz diagonal D donde $d_{ii} = \lambda_i$ para $i = 1, \dots, n$. Es decir,

$$D = \operatorname{diag}(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Ejemplo 2.3. La **matriz identidad** $I_n \in \mathcal{M}_n(\mathbb{K})$ es aquella con n filas y n columnas cuyas entradas son todas iguales a 0 excepto las de la diagonal principal, que son iguales a 1. Es decir,

$$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

Con la notación habitual de la *delta de Kronecker*

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

se tiene que $I_n = (\delta_{ij}) \in \mathcal{M}_n(\mathbb{K})$.

Definición 2.3 (Igualdad de matrices)**.** Dos matrices $A = (a_{ij})$ y $B = (b_{ij})$ en $\mathcal{M}_{m\times n}(\mathbb{K})$ son **iguales** si y solo si tienen el mismo orden y sus entradas correspondientes son iguales; es decir:

$$A = B \iff a_{ij} = b_{ij} \text{ para todo } i = 1, \dots, m \text{ y } j = 1, \dots, n$$

Definición 2.4 (Submatriz)**.** Sea $A \in \mathcal{M}_{m\times n}(\mathbb{K})$. Llamaremos **submatriz** de A a cualquier matriz obtenida a partir de A suprimiendo algunas de sus filas y/o columnas. Si $I \subseteq \{1, \dots, m\}$ y $J \subseteq \{1, \dots, n\}$ son subconjuntos de índices, denotaremos por A_{IJ} a la submatriz formada por las filas indexadas por I y las columnas indexadas por J.

Ejemplo 2.4. Consideremos la matriz $A \in \mathcal{M}_{5\times 5}(\mathbb{K})$ dada por:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 \\ 21 & 22 & 23 & 24 & 25 \end{bmatrix}.$$

Si tomamos los subconjuntos de índices $I = \{1, 3, 4\}$ y $J = \{2, 4, 5\}$, entonces la submatriz A_{IJ} es:

$$A_{IJ} = \begin{bmatrix} 2 & 4 & 5 \\ 12 & 14 & 15 \\ 17 & 19 & 20 \end{bmatrix}$$

que corresponde al conjunto de subindices $I \times J$ de la matriz A.

Para visualizar mejor este proceso de selección de submatriz, consideremos el siguiente diagrama:

$$\begin{array}{c|ccccc|} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 1 & 2 & 3 & 4 & 5 \\ 2 & 6 & 7 & 8 & 9 & 10 \\ 3 & 11 & 12 & 13 & 14 & 15 \\ 4 & 16 & 17 & 18 & 19 & 20 \\ 5 & 21 & 22 & 23 & 24 & 25 \\ \hline \end{array} \longrightarrow \begin{array}{c|ccc|} & 1 & 2 & 3 \\ \hline 1 & 2 & 4 & 5 \\ 2 & 12 & 14 & 15 \\ 3 & 17 & 19 & 20 \\ \hline \end{array}$$

donde las casillas sombreadas corresponden a los elementos seleccionados para formar la submatriz A_{IJ} con $I = \{1,3,4\}$ y $J = \{2,4,5\}$.

2.2 Operaciones con matrices

Definición 2.5 (Suma de matrices)**.** La suma de dos matrices $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, $B \in \mathcal{M}_{m\times n}(\mathbb{K})$ se define como la suma elemento a elemento, es decir,

$$A + B := \begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \in \mathcal{M}_{m\times n}(\mathbb{K}).$$

La suma de matrices cumple las siguientes propiedades fundamentales:

1. **Asociatividad:** Para cualesquiera matrices A, B, $C \in \mathcal{M}_{m\times n}(k)$, se cumple que $(A+B)+C = A+(B+C)$.
2. **Conmutatividad:** Para cualesquiera matrices A, $B \in \mathcal{M}_{m\times n}(k)$, se cumple que $A+B = B+A$.
3. **Elemento neutro:** La matriz nula $0 \in \mathcal{M}_{m\times n}(k)$ actúa como elemento neutro de la suma, es decir, para toda matriz $A \in \mathcal{M}_{m\times n}(k)$ se cumple que $A+0 = 0+A = A$.
4. **Elemento opuesto:** Para toda matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}(k)$, existe una matriz opuesta $-A = (-a_{ij})$ tal que $A+(-A) = (-A)+A = 0$. Esta matriz opuesta se obtiene multiplicando cada elemento de A por -1.

Estas propiedades dotan al conjunto $\mathcal{M}_{m\times n}(k)$ de una estructura de grupo abeliano respecto a la operación suma.

Definición 2.6 (Producto de un escalar por una matriz)**.** Sea $A = (a_{ij}) \in \mathcal{M}_{m\times n}(k)$ y $\lambda \in k$. Se define el producto de λ por A como la matriz

$$\lambda \cdot A := \begin{bmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{bmatrix} \in \mathcal{M}_{m\times n}(k),$$

esto es, el producto de un escalar por una matriz es la matriz que resulta al multiplicar cada una de las entradas de la matriz por el escalar.

Para $\lambda, \mu \in \mathbb{K}$ y $A, B \in \mathcal{M}_{m\times n}(\mathbb{K})$, el producto de un escalar por una matriz cumple las siguientes propiedades:

1. **Distributividad respecto a la suma de escalares:** $\lambda(A+B) = \lambda A + \lambda B$.

2. **Distributividad respecto a la suma de matrices:** $(\lambda + \mu)A = \lambda A + \mu A$.
3. **Compatibilidad con la multiplicación de escalares:** $(\lambda\mu)A = \lambda(\mu A)$.
4. **Asociatividad del producto por escalares:** $\lambda(\mu A) = (\lambda\mu)A$.

Definición 2.7 (Producto de matrices)**.** Sean $A = (a_{ij}) \in \mathcal{M}_{m\times p}(\mathbb{K})$ y $B = (b_{ij}) \in \mathcal{M}_{p\times n}(\mathbb{K})$ dos matrices. El producto $C = AB = (c_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{K})$ es una matriz cuyos elementos c_{ij} se calculan como:

$$c_{ij} = \sum_{l=1}^{p} a_{il}b_{lj}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n. \tag{2.13}$$

Para visualizar mejor cómo se realiza este producto, consideremos la siguiente representación gráfica:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix} \times \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ b_{p1} & \cdots & b_{pj} & \cdots & b_{pn} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mj} & \cdots & c_{mn} \end{bmatrix}$$

En el diagrama anterior, se muestra cómo cada elemento c_{ij} del producto (resaltado en azul) se obtiene del producto escalar entre la fila i-ésima de A (resaltada en verde) y la columna j-ésima de B (resaltada en verde). Utilizamos la notación $A_{r(i)}$ para referirnos a la fila i-ésima de A y $B_{c(j)}$ para la columna j-ésima de B.

Observación. Las matrices solo pueden multiplicarse si sus dimensiones «vecinas» coinciden. Por ejemplo, una matriz A de dimensión $n \times p$ puede multiplicarse con una matriz B de dimensión $p \times m$, pero solo por la izquierda:

$$\underset{n\times p}{A} \cdot \underset{p\times m}{B} = \underset{n\times m}{C}$$

El producto BA no está definido si $m \neq n$ ya que las dimensiones vecinas no coinciden.

Observación. La multiplicación de matrices *no* está definida como una operación elemento a elemento, es decir, $c_{ij} \neq a_{ij}b_{ij}$ (incluso si el tamaño de A y B fuera apropiado). Este tipo de multiplicación elemento a elemento aparece frecuentemente en lenguajes de programación cuando multiplicamos arreglos (multidimensionales) entre sí, y se denomina *producto de Hadamard*.

Ahora que hemos definido la multiplicación de matrices, la suma de matrices y la matriz identidad, veamos algunas propiedades de las operaciones matriciales:

1. **Propiedad asociativa de la multiplicación:** Sea $A \in \mathcal{M}_{m\times n}$, $B \in \mathcal{M}_{n\times p}$, $C \in \mathcal{M}_{p\times q}$, entonces
$$(AB)C = A(BC)$$
Esta propiedad nos permite calcular productos de tres o más matrices en cualquier orden de agrupación.

2. **Propiedades distributivas:** Sea $A, B \in \mathcal{M}_{m\times n}$, $C, D \in \mathcal{M}_{n\times p}$, entonces
$$(A + B)C = AC + BC$$
$$A(C + D) = AC + AD$$
La multiplicación matricial distribuye tanto por la izquierda como por la derecha respecto a la suma.

3. **Propiedades de la matriz identidad:** Sea $A \in \mathcal{M}_{m\times n}$, entonces
$$I_m A = AI_n = A$$
La matriz identidad actúa como elemento neutro de la multiplicación. Es importante observar que $I_m \neq I_n$ cuando $m \neq n$, y que necesitamos matrices identidad de diferentes dimensiones para multiplicar por la izquierda (I_m) y por la derecha (I_n).

Con las propiedades anteriores, el conjunto $\mathcal{M}_{m\times n}(\mathbb{K})$ con las operaciones de suma y producto por escalares, y con la matriz identidad, cumple las propiedades de un anillo conmutativo con unidad.

Definición 2.8 (Matriz traspuesta)**.** Dada una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, definimos su **Matriz transpuesta** $A^\top \in \mathcal{M}_{n\times m}(\mathbb{K})$ como aquella que se obtiene al intercambiar filas por columnas, es decir, $(A^\top)_{ij} = A_{ji}$.

Ejemplo 2.5. Si tenemos la matriz
$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathcal{M}_{2\times 3}(\mathbb{R})$$
su matriz traspuesta es
$$A^\top = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in \mathcal{M}_{3\times 2}(\mathbb{R})$$
Observe que las columnas de $A^\top$ son las filas de A y viceversa.

Definición 2.9 (Matriz simétrica y antisimétrica)**.** Una matriz cuadrada $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ es:

1. **Simétrica** si $A = A^\top$, es decir, si $a_{ij} = a_{ji}$ para todo i, j.
2. **Antisimétrica** si $A = -A^\top$, es decir, si $a_{ij} = -a_{ji}$ para todo i, j.

Ejemplo 2.6. Consideremos las siguientes matrices:

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \quad \text{y} \quad C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

La matriz B es Matriz simétrica ya que $B = B^\top$, mientras que C es Matriz antisimétrica pues $C = -C^\top$. Puede verificarse esto observando que en B, los elementos simétricos respecto a la diagonal principal son iguales ($b_{12} = b_{21} = 2$), mientras que en C, los elementos simétricos son opuestos ($c_{12} = -c_{21} = 1$).

Definición 2.10 (Matriz invertible)**.** Una matriz $A \in \mathcal{M}_n(\mathbb{K})$ es **invertible** (o **no singular**) si existe una matriz $B \in \mathcal{M}_n(\mathbb{K})$ tal que $AB = BA = I_n$. En tal caso, la matriz B es única y se denomina **Matriz inversa de** A, denotándose por A^{-1}.

Ejemplo 2.7. Consideremos la matriz

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

Su matriz inversa es

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

En efecto, podemos verificar que $AA^{-1} = A^{-1}A = I_2$ realizando el producto:

$$AA^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

En este caso particular, A es su propia inversa, es decir, $A^{-1} = A$. Las matrices con esta propiedad se denominan **involutivas**.

Más adelante daremos un criterio para saber si una matriz es invertible y, en este caso, una fórmula para calcular la matriz inversa.

Definición 2.11 (Matriz ortogonal)**.** Una matriz $A \in \mathcal{M}_n(\mathbb{R})$ es **ortogonal** si su transpuesta coincide con su inversa, es decir, si $A^\top = A^{-1}$. De forma equivalente, A

es ortogonal si y solo si $AA^\top = A^\top A = I_n$.

Ejemplo 2.8. Consideremos la matriz

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Podemos verificar que A es ortogonal realizando el producto:

$$A^\top A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$$

Por lo tanto, A es ortogonal.

Definición 2.12 (Matriz traspuesta conjugada)**.** Sea $A = (a_{ij}) \in \mathcal{M}_{m\times n}(\mathbb{C})$. La matriz $A^* = (\overline{a_{ji}}) \in \mathcal{M}_{n\times m}(\mathbb{C})$ se denomina **matriz traspuesta conjugada** o **matriz adjunta**, siendo $\overline{a_{ji}}$ el conjugado complejo de a_{ji}, para $i = 1, \ldots, m$ y $j = 1, \ldots, n$. Esta operación combina la transposición de la matriz con la conjugación compleja de sus elementos.

La matriz traspuesta conjugada tiene varias propiedades importantes:

1. $(A^*)^* = A$ (involución)
2. Para matrices reales, $A^* = A^\top$, ya que los números reales son iguales a sus conjugados
3. $(AB)^* = B^*A^*$ para matrices compatibles
4. Para escalares $\lambda \in \mathbb{C}$, $(\lambda A)^* = \overline{\lambda} A^*$

Para un vector columna

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{C}^n,$$

su traspuesta conjugada es el vector fila $\mathbf{v}^* = (\overline{v_1}, \ldots, \overline{v_n})$.

Definición 2.13 (Matrices especiales complejas)**.** Sea $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{C})$. Se dice que A es:

1. **Hermítica** (o auto-adjunta) si $A = A^*$, es decir, $a_{ij} = \overline{a_{ji}}$ para todo $i, j = 1, 2, \ldots, n$. Los elementos de la diagonal principal de una matriz hermítica son necesariamente números reales.
2. **Unitaria** si $A^* = A^{-1}$, es decir, $AA^* = A^*A = I_n$. Las matrices unitarias son la generalización compleja de las matrices ortogonales reales.

3. **Normal** si $AA^* = A^*A$, es decir, si la matriz conmuta con su traspuesta conjugada.

Teorema 2.1. *Las matrices hermíticas y unitarias tienen propiedades especiales que las relacionan con las matrices normales. En particular:*

1. *Toda matriz hermítica o unitaria es también normal.*
2. *Si A es hermítica, entonces $A = A^*$. Por consiguiente, $AA^* = AA = A^*A$.*
3. *Si A es unitaria, entonces $AA^* = I_n$ y $A^*A = I_n$. Por consiguiente, $AA^* = A^*A$.*

2.2.1. La traza y el determinante de una matriz

Definición 2.14 (Traza de una matriz)**.** Sea $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ una matriz cuadrada de orden n. La **traza** de A, denotada por $\operatorname{tr}(A)$, es la suma de los elementos de su diagonal principal:

$$\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}.$$

Ejemplo 2.9. Si tenemos la matriz

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

su traza es

$$\operatorname{tr}(A) = 1 + 5 + 9 = 15.$$

Definición 2.15 (Determinante de una matriz)**.** Sea $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ una matriz cuadrada de orden n. El **determinante** de A, denotado por $\det(A)$ o $|A|$, es un escalar que se define por la expresión:

$$|A| = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

donde S_n es el conjunto de todas las permutaciones de los índices $\{1, 2, \ldots, n\}$ y $\operatorname{sign}(\sigma)$ es la función signo que asigna $+1$ a las permutaciones pares y -1 a las impares.

Esta definición se puede expresar de manera compacta como:

$$|A| = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}.$$

Veamos algunos ejemplos de determinantes:

Ejemplo 2.10. Consideremos la matriz $A = (a_{ij}) \in \mathcal{M}_2(\mathbb{K})$, entonces su determinante se calcula como:

$$|A| = a_{11}a_{22} - a_{12}a_{21}.$$

En donde $S_2 = \{(1,2),(2,1)\}$ y $\operatorname{sign}((1,2)) = 1$, $\operatorname{sign}((2,1)) = -1$.

Ejemplo 2.11. Consideremos la matriz $B = (b_{ij}) \in \mathcal{M}_3(\mathbb{K})$, entonces su determinante se calcula como:

$$|B| = b_{11}b_{22}b_{33} + b_{12}b_{23}b_{31} + b_{13}b_{21}b_{32} - b_{13}b_{22}b_{31} - b_{12}b_{21}b_{33} - b_{11}b_{23}b_{32}.$$

En donde $S_3 = \{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\}$ y las funciones signo son $\operatorname{sign}((1,2,3)) = 1$, $\operatorname{sign}((1,3,2)) = -1$, $\operatorname{sign}((2,1,3)) = -1$, $\operatorname{sign}((2,3,1)) = 1$, $\operatorname{sign}((3,1,2)) = 1$, y $\operatorname{sign}((3,2,1)) = -1$.

Definición 2.16 (Matrices menores)**.** Sea $A \in \mathcal{M}_{m\times n}(k)$. Para un entero positivo p tal que $p \leq \text{mín}(m,n)$, se denomina *menor de orden* p de A a cualquier determinante de una submatriz cuadrada de A de tamaño $p \times p$.

En el caso en que A es cuadrada ($m = n$), el **menor principal de orden** p es el determinante de la submatriz obtenida al conservar las primeras p filas y columnas de A (es decir, eliminando las últimas $n - p$ filas y columnas).

Definición 2.17 (Matriz de cofactores)**.** Sea $A \in \mathcal{M}_{n\times n}(k)$. La **matriz de cofactores** de A, denotada por $C = (c_{ij})$, es la matriz obtenida al calcular el cofactor c_{ij} de cada elemento a_{ij} de A. El cofactor se define como:

$$c_{ij} = (-1)^{i+j}|A_{ij}|,$$

donde $|A_{ij}|$ es el determinante del menor obtenido al eliminar la fila i y la columna j de A.

Teorema 2.2. *Sea $A \in \mathcal{M}_{n\times n}(k)$ una matriz cuadrada de orden n. Entonces el determinante de A es igual a la suma de los productos de los elementos de cualquier fila (o columna) de A por sus cofactores correspondientes, es decir, si elegimos la fila*

i, entonces:

$$|A| = \sum_{j=1}^{n} a_{ij}c_{ij} = a_{i1}c_{i1} + a_{i2}c_{i2} + \cdots + a_{in}c_{in}.$$

o si elegimos la columna j, entonces:

$$|A| = \sum_{i=1}^{n} a_{ij}c_{ij} = a_{1j}c_{1j} + a_{2j}c_{2j} + \cdots + a_{nj}c_{nj}.$$

Observación. Es imporante resaltar que el determinante de la suma de matrices no es igual a la suma de los determinantes, es decir, en general $|A+B| \neq |A| + |B|$ para matrices $A, B \in \mathcal{M}_{n\times n}(k)$, y por lo tanto, tampoco se $|\lambda A| = \lambda|A|$ para un escalar $\lambda \in k$ y una matriz $A \in \mathcal{M}_{n\times n}(k)$.

2.2.2. Matriz inversa

Apéndice: Material Complementario

.1 Demostraciones Adicionales

Aquí puedes incluir demostraciones más detalladas o extensas que no encajan en el cuerpo principal del libro.

.2 Tablas de Datos

Aquí puedes incluir tablas de datos extensas o información de referencia.