# Preliminares - Cálculo

## 1 Ecuaciones importantes

En Matemática I y II han estudiado gráficas que consistían en curvas (rectas, parábolas, elipses, senoidales, cosenoidales, exponenciales, logarítimicas, y otras más), y en todos los casos estas representaciones se realizaron a partir de ecuaciones, que en forma sencilla, son igualdades con incógnitas.

Con el fin de refrescar la memoria, a continuación se mencionan algunas formas generales de ecuaciones cuyas representaciones gráficas son curvas en el plano $\mathbb{R}^2$

- Rectas

  $x = a \qquad a \in \mathbb{R}$ (recta paralela al eje Y)

  $y = b \qquad b \in \mathbb{R}$ (recta paralela al eje X)

  $y = mx + b \qquad m, b \in \mathbb{R}$ (recta con pendiente m, $m \neq 0$)

- Prábolas

  $y = a(x-\alpha)^2 + \beta \qquad a, \alpha, \beta \in \mathbb{R}$ (parábola con eje paraleo al eje $Y$ y vértice en V$(\alpha, \beta)$

  $x = b(y-\gamma)^2 + \delta \qquad b, \delta, \gamma \in \mathbb{R}$ (parábola con eje paraleo al eje $X$ y vértice en V$(\gamma, \delta)$

  $y = ax^2 \qquad a \in \mathbb{R}$ (parábola con eje paralelo al eje $Y$ y vértice en el origen de coordenadas)

- Circunferencias

  $(x-\alpha)^2 + (y-\beta)^2 = r^2 \qquad r, \alpha, \beta \in \mathbb{R}$ (circunferencia con centro en el punto C$(\alpha, \beta)$ y radio $r$)

  $x^2 + y^2 = r^2 \qquad r \in \mathbb{R}$ (circunferencia con centro en origen de coordenadas y radio $r$)

- Elipses

$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1 \qquad h,k,a,b \in \mathbb{R}$ (elipse con centro en el punto C$(h,k)$)

$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \qquad a,b \in \mathbb{R}$ (elipse con centro en el origen de coordenadas)

Las siguientes son ecuaciones cuyas gráficas corresponden a superficies en el espacio $\mathbb{R}^3$:

- Planos

$z = 0$ (Plano coordenado XY)

$y = 0$ (Plano coordenado XZ)

$x = 0$ (Plano coordenado YZ)

$\alpha.y + \beta.z + d = 0$ (Plano paralelo al eje coordenado X)

$\gamma.y + \beta.z + d = 0$ (Plano paralelo al eje coordenado Y)

$\gamma.x + \alpha.y + d = 0$ (Plano paralelo al eje coordenado Z)

$\gamma.x + \alpha.y + \beta.z + d = 0$ (Otros planos)

- Esferas

$(x-h)^2 + (y-k)^2 + (z-l)^2 = r^2$ (esfera con centro en $C(h,k,l)$ y radio $r$)

$x^2 + y^2 + z^2 = r^2$ (esfera con centro en en $C(0,0,0)$ y radio $r$)

- Elipsoides

$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} + \frac{(z-l)^2}{c^2} = 1$ (elipsoide con centro en C$(h,k,l)$ y semiejes $a,b,c$)

$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ (elipsoide con centro en $C(0,0,0)$)

- Paraboloides

$z = (x-h)^2 + (y-k)^2 + l$ (paraboloide con vértice en $V(h,k,l)$ y eje de simetría paralelo a $Z$)

$z = x^2 + y^2$ (paraboloide con vértice en $V(0,0,0)$ y eje de simetría $Z$)

## 2 Límites (una variable)

**Definición 2.1.** Definición intuitiva de límite

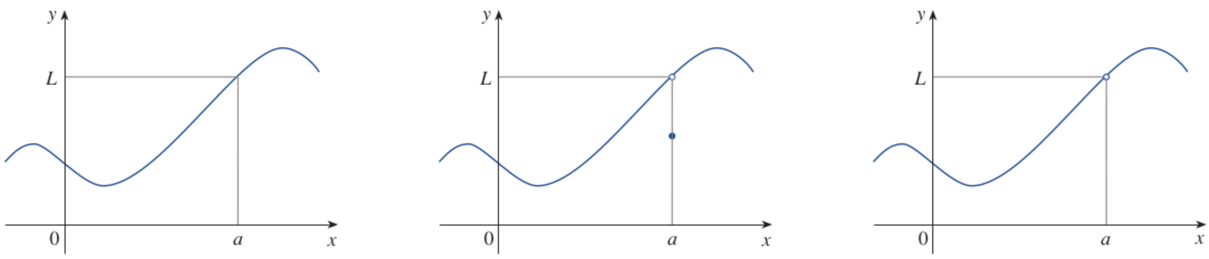
*Dada una función $f(x)$ definida en un intervalo $S$ alrededor del valor $x_0$, exceptuando posiblemente a $x_0$ (definida cuando $x$ está "cerca" de $x_0$). Entonces se escribe*

$$\lim_{x \to x_0} f(x) = L$$

*y se dice que "el límite de $f(x)$, cuando $x$ tiende a $x_0$, es igual a $L$" si se puede hacer que los valores de $f(x)$ estén arbitrariamente cercanos a $L$ (tan cercanos como se quiera), tomando valores de $x$ lo suficientemente cerca de $x_0$ (por ambos lados de $x_0$), pero no iguales a $x_0$*

En términos generales esto quiere decir que los valores de la función $f(x)$ se aproximan al valor $L$ cuando $x$ se acerca cada vez más al número $x_0$, pero $x \neq x_0$. En la siguiente figura se muestran 3 casos que reflejan el significado de la definición intuitiva de límite, en los cuales

$$\lim_{x \to a} f(x) = L$$

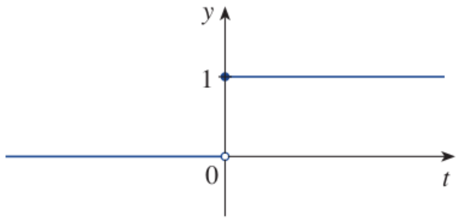



Notemos que no es necesario que la función esté definida en el punto para poder afirmar que el límite existe. Veamos un caso en el que la función está definida en el punto pero donde el límite no existe:

**Ejemplo 2.2.** *La función Heaviside $H$ se define como*

$$H(t) = \begin{cases} 0 & si \quad t < 0 \\ 1 & si \quad t \geq 0 \end{cases}$$

*y cuya gráfica es*




*Veamos que cuando t se aproxima a 0 por la izquierda, $H(t)$ se acerca a 0. En cambio, cuando t se aproxima a 0 por la derecha $H(t)$ toma valores cada vez más cercanos a 1. Al no haber un único número al que se aproxime $H(t)$ cuando $t \to 0$, decimos que el límite de la función cuando t tiende a 0 no existe*

$$\nexists \lim_{t \to 0} H(t)$$

*En este caso notamos que la función se acerca a valores distintos cerca del punto en cuestión dependiendo si se toman valores mayores o menores a 0. Esta situación se indica simbólicamente como*

$$\lim_{t \to 0^-} H(t) = 0 \qquad\qquad \lim_{t \to 0^+} H(t) = 1$$

*El $+$ indica que solo se consideran valores de t mayores a 0 y el $-$ solo valores menores.*

A raíz de este ejemplo definimos el concepto de **límites laterales**:

**Definición 2.3.** Límites laterales

*Dada una función $f(x)$ y una constante $x_0$, cuando se escribe*

$$\lim_{x \to x_0^+} f(x) = L$$

*se expresa que el* ***límite de f(x) cuando*** *$x$* ***tiende a*** *$x_0$* ***por derecha*** *es igual a L si se puede hacer que los valores de $f(x)$ se aproximen arbitrariamente a L, tanto como uno quiera, tomando valores de $x$ lo suficientemente cercanos a $x_0$, pero mayores que $x_0$.*

*De forma análoga se define*

$$\lim_{x \to x_0^-} f(x) = L$$

*y se expresa que el* ***límite de f(x) cuando*** *$x$* ***tiende a*** *$x_0$* ***por izquierda*** *es igual a L si se puede hacer que los valores de $f(x)$ se aproximen arbitrariamente a L, cuando se toman valores cercanos a $x_0$, pero menores que $x_0$.*

De esta forma, uniendo esta definición con la Definición 2.1 se satisface el siguiente teorema:

**Teorema 2.4.** *El límite* $\lim_{x \to x_0} f(x)$ *existe si y sólo si ambos límites laterales* $\lim_{x \to x_0^+} f(x)$ *y* $\lim_{x \to x_0^-} f(x)$ *existen y son iguales*

A continuación presentaremos las denominadas *leyes de los límites* que, a partir del conocimiento de límites de funciones simples, nos permiten calcular los límites de funciones más complejas:

**Propiedades 2.5.** Leyes de los límites

*Suponiendo c una constante y que los límites*

$$\lim_{x \to x_0} f(x) = L_1 \qquad \lim_{x \to x_0} g(x) = L_2$$

*existen. Entonces,*

1. $\lim_{x \to x_0} [f(x) + g(x)] = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x) = L_1 + L_2$
2. $\lim_{x \to x_0} [f(x) - g(x)] = \lim_{x \to x_0} f(x) - \lim_{x \to x_0} g(x) = L_1 - L_2$
3. $\lim_{x \to x_0} [cf(x)] = c. \lim_{x \to x_0} f(x) = c.L_1$
4. $\lim_{x \to x_0} [f(x)g(x)] = \lim_{x \to x_0} f(x). \lim_{x \to x_0} g(x) = L_1.L_2$
5. $\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x \to x_0} f(x)}{\lim_{x \to x_0} g(x)} = \frac{L_1}{L_2} \qquad \lim_{x \to x_0} g(x) \neq 0$

## 2.1 Límites clásicos

Los siguientes son algunos límites importantes para los cuales daremos su resultado:

- $\lim_{x \to 0} \frac{sen(x)}{x} = 1$
- $\lim_{x \to 0} \frac{1 - cos(x)}{x} = 0$
- $\lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^x = e$
- $\lim_{x \to 0} \frac{e^{kx} - 1}{x} = k$
- $\lim_{x \to 0} \frac{a^{kx} - 1}{x} = kln(a)$
- $\lim_{x \to 0} ln \left(\frac{1 + x}{x}\right) = 1$
- $\lim_{t \to 0} (1 + t)^{\frac{1}{t}} = e$

## 3 Continuidad

Como han visto en Matemática I y II, toda función está *definida* para un conjunto de elementos. Si pensamos a la función como una "máquina", la palabra *definida* significa ésta sólo acepta un conjunto (puede ser infinito) de elementos de entrada que podrán serán "transformados". Este conjunto recibe el nombre de **dominio**.

**Definición 3.1.** *Dada $f : D \subset \mathbb{R} \to \mathbb{R}$, se llama* ***dominio de*** *$f$, y se anota $Dom(f)$, al subconjunto $D$ de $\mathbb{R}$ en donde está definida la función $f$.*

Además, cuando observamos la curva que forma la gráfica de una función sobre el plano $\mathbb{R}^2$, buscamos que en cierta forma sea lo más "bonita" posible, es decir, sin interrupciones ni saltos. A partir de esto definimos la siguiente propiedad que pueden cumplir las funciones:

**Definición 3.2.** Continuidad en un punto

*Dada una función $f : D \subset \mathbb{R} \to \mathbb{R}$, decimos que es* ***contínua*** *en un punto $x_0$ si se cumple*

$$\lim_{x \to x_0} f(x) = f(x_0)$$

Observemos que la definición requiere implícitamente tres condiciones para que $f$ sea contínua en el punto $x_0$

1. $f(x_0)$ está definida, es decir, $x_0 \in Dom(f)$
2. $\exists \lim_{x \to x_0} f(x)$
3. $\lim_{x \to x_0} f(x) = f(x_0)$

De esta forma, si una función es continua en todo un intervalo de puntos de su dominio, tiene la propiedad de que un pequeño cambio en $x$ produce solo un pequeño cambio en $f(x)$ en dicho intervalo (no hay saltos bruscos ni interrupciones).

Si alguna de las 3 condiciones anteriores no se cumple se dice que la función presenta una **discontinuidad** en el punto $x_0$, o que la función es **discontinua** en $x_0$. Este es el caso de $H(t)$ vista en el Ejemplo 2.2, donde no se cumple la condición (2) y menos aún la (3).

El siguiente es un resultado que nos facilitará determinar si una función, posiblemente compleja, es continua en un punto.

**Teorema 3.3.** *Si $c$ es un número real y $f$ y $g$ son funciones continuas en $x_0$, entonces las funciones siguientes son continuas en $x_0$:*

1. *Múltiplo escalar:* $cf$
2. *Suma (y diferencia):* $f + g$
3. *Producto:* $fg$
4. *Cociente:* $\frac{f}{g}$ *, si* $g(x_0) \neq 0$.

En general, cuando debemos hallar el intervalo donde una función $f(x)$ es continua, utilizamos tanto el Teorema 3.3 como los siguientes dos resultados:

**Teorema 3.4.** *Los siguientes tipos de funciones son continuas en todo punto de su dominio:*

- *Polinómicas*
- *Trigonométricas*
- *Exponenciales*
- *Racionales*
- *Radicales*
- *Logarítmicas*

**Teorema 3.5.** *Si $g$ es continua en $x_0$ y $f$ es continua en $g(x_0)$, entonces la función compuesta $f \circ g$ dada por $(f \circ g)(x) = f(g(x))$ es continua en $x_0$*

Con ayuda de estos tres teoremas podemos tomar una función relativamente compleja y determinar el intervalo de continuidad con bastante facilidad, pero siempre cuidando y prestando atención al dominio de la función. Usualmente, en los puntos donde la función no esta definida hay que hacer un análisis detenido utilizando la definición de continuidad y las tres condiciones mencionadas.

# 4 Derivación en una variable

El elemento central del Cálulo Diferencial, como el que veremos en esta materia, es la denominada **derivada**. Esta herramiente del cálculo nos permite determinar la pendiente de la recta tangente a una curva en un punto dado.

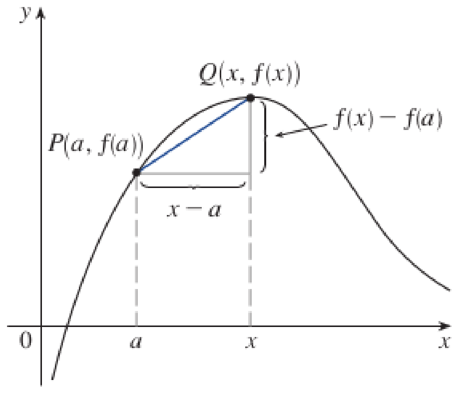
Haciendo memoria, la *pendiente* de una recta es una razón de cambio entre los valores de $x$ e $y$. Determina la cantidad de unidades que aumenta o disminuye el valor de $y$ cuando $x$ cambia en una unidad.

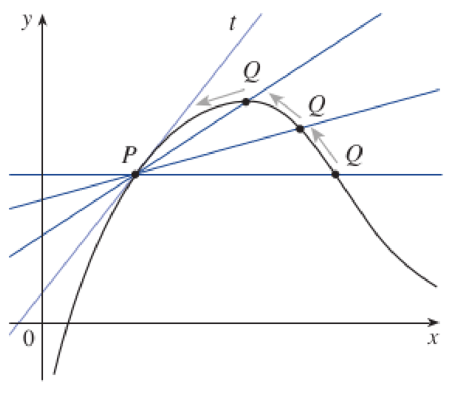
Como sabemos, la pendiente de una recta es única a lo largo de toda su extensión. Este no es el caso cuando tratamos con una curva, la cual puede sufrir cambios constantes en su concavidad y crecimiento. Cuando queremos hallar la razón de cambio instantánea en un punto no resulta tan directo.

Si queremos hallar la recta tangente a la curva $C$ en el punto $P(a, f(x))$ debemos comenzar considerando un punto cercano $Q(x, f(x))$, donde $x \neq a$, y calculamos la pendiente de la recta *secante* (corta en dos puntos a la gráfica) $PQ$:

$$m_{PQ} = \frac{f(x) - f(a)}{x - a}$$

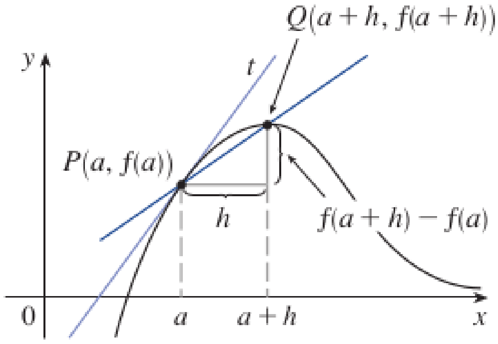
Luego, acercamos $Q$ a $P$ a lo largo de la curva $C$, haciendo que $x$ tienda a $a$. Si $m_{PQ}$ tiende a un número $m$, entonces se define la *tangente* $t$ como la recta que pasa por $P$ con pendiente $m$. Gráficamente




Otra expresión para la pendiente $m_{PQ}$ (que resulta más sencilla de utilizar) se construye utilizando a $h = x - a$. De esta forma $x = a + h$ y la pendiente de la recta secante queda

$$m_{PQ} = \frac{f(a+h) - f(a)}{h}$$




Observemos que conforme $x$ se aproxima a $a$, $h$ se acerca a 0. Es así cómo obtenemos la definición que conocemos de **derivada** o **razón de cambio instantánea**, que representa la pendiente $m$ de la recta tangente $t$.

**Definición 4.1.** COCIENTE INCREMENTAL DE NEWTON

*Sea $f$ una función, y sea $x$ un número en el dominio de $f$. Llamamos* ***derivada*** *de $f$ en $x$ al límite del cociente incremental de $f$ en $x$ cuando el incremento tiende a* 0*, siempre que ese límite exista. a la derivada la indicamos $f'(x)$. Entonces:*

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Muchas veces deberemos analizar si una función es derivable en un intervalo, es decir, tendremos que ver si existe la derivada de la función para cada punto del intervalo. Es claro que, si tenemos que calcular el cociente incremental para cada punto, es posible que nunca terminemos (si el intervalo es infinito). Por ello presentamos las siguientes *reglas de derivación* que permiten calcular la forma genérica de la derivada para los puntos de la función de una forma más sencilla:

**Propiedades 4.2.** Reglas de derivación para funciones de una variable

*Sean $g(x)$ y $f(x)$ funciones de una variable derivables, entonces tenemos las siguientes propiedades:*

- *Si* $g(x) = c \quad \forall x \in \mathbb{R} \rightarrow g'(x) = 0$.
- *Si* $g(x) = x^n, n \in \mathbb{N} \rightarrow g'(x) = nx^{n-1}$.
- $(f(x) \pm g(x))' = f'(x) \pm g'(x)$.
- $(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x)$.
- $\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$, *si* $g(x) \neq 0$ *en algún intervalo real.*

Por otro lado, cabe mencionar que tanto la continuidad como la derivabilidad son propiedades deseables para una función. El teorema siguiente muestra cómo se relacionan estas propiedades:

**Teorema 4.3.** *Sea $f$ una función, si $f$ es derivable en un punto $x_0$, entonces $f$ es continua en $x_0$*

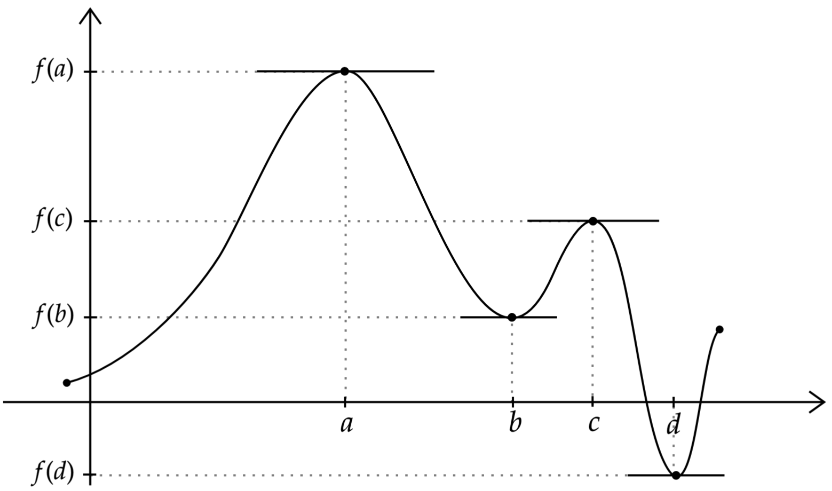
# 5 Máximos y Mínimos

Uno de los usos de la primera derivada de una función es el cálculo de los valores máximos y mínimos absolutos y relativos a lo largo de su dominio. Veamos cómo definimos a estos extremos:

**Definición 5.1.** Máximos y Mínimos

- *Se dice que el punto* $x_0, f(x_0)$ *es un* ***máximo absoluto de f*** *si* $f(x_0) \geq f(x)$ *para todo* $x$ *en el dominio de* $f$
- *Se dice que el punto* $x_0, f(x_0)$ *es un* ***máximo relativo o local de f*** *si* $f(x_0) \geq f(x)$ *para todo* $x$ *en un intervalo alrededor de* $x_0$
- *Se dice que el punto* $x_0, f(x_0)$ *es un* ***mínimo absoluto de f*** *si* $f(x_0) \leq f(x)$ *para todo* $x$ *en el dominio de* $f$
- *Se dice que el punto* $x_0, f(x_0)$ *es un* ***mínimo relativo o local de f*** *si* $f(x_0) \leq f(x)$ *para todo* $x$ *en un intervalo alrededor de* $x_0$

Podemos ver gráficamente estos cuatro casos en la siguiente figura:




La función presenta máximo y mínimo relativo (o local) en los valores $b$ y $c$ respectivamente. Luego sus extremos absolutos se encuentran en $a$ y $d$, dando un valor $f(a)$ máximo y $f(d)$ mínimo.

Teniendo en cuenta la definición de derivada y su significado como la “pendiente” que presenta la curva en un punto (más precisamente la pendiente de la recta tangente a la curva), notemos que en los cuatro valores de $x$ de la figura anterior, la pendiente es 0. Es decir, si trazamos las rectas tangentes a la curva en estos puntos, éstas tienen pendiente 0 y resultan ser rectas horizontales. Esto nos aporta un leve indicio de que la derivada es una herramienta que nos permite detectar máximos y mínimos, determinando los puntos donde ésta toma el valor 0.

**Definición 5.2.** *Se dice que $x_0$ es un **punto crítico** de $f$ si $f'(x_0) = 0$ o si $f'(x_0)$ no existe*

Notamos que $f'(a) = f'(b) = f'(c) = f'(d) = 0$ por lo que, según la definición, son **puntos críticos** de la función graficada. El siguiente teorema muestra que este valor de la derivada no es mera coincidencia.

**Teorema 5.3.** *Si $f$ tiene un máximo local o un mínimo local en un punto c interior de su dominio, y suponiendo que la derivada de $f$ está definida en ese punto, entonces $f'(c) = 0$*

De este resultado obtenemos que los únicos lugares donde la función $f$ posiblemente tiene un valor extremo (local o global) son los puntos críticos de $f$. Por otro lado, si tratamos con una función continua y definida en un intervalo cerrado, los extremos de éste también cuentan como puntos candidatos a tener un máximo o mínimo. Este último tipo de las funciones siempre alcanzarán sus valores máximo y mínimo absolutos en el intervalo.

Esta última afirmación se basa en uno de los teoremas fundamentales del cálculo diferencial en una variable:

**Teorema 5.4.** Teorema de Weierstrass

*Sea $f$ una función continua en un intervalo cerrado $[a, b]$, ésta alcanza un valor máximo absoluto $(M)$ y un valor mínimo absoluto $(m)$ en valores de $x \in [a, b]$.*

*Por definición de máximos y mínimos absolutos en $[a, b]$ debe cumplirse*

$$m \leq f(x) \leq M \qquad \forall\ x \in [a, b]$$

# Preliminares - Estadística

Los siguientes serán conceptos de Matemática 3 utilizados a lo largo del desarrollo teórico del método de *Regresión Lineal*, por lo que es importante que se tomen el tiempo de revisarlos.

## 1 Variables Aleatorias

En cualquier experimento existen múltiples aspectos o características que pueden ser observadas o medidas, pero en la mayoría de los casos un experimentador se enfoca en uno o algunos aspectos específicos de una muestra, como será en nuestro caso.

Por ejemplo, un investigador prueba una muestra de componentes electrónicos para los cuales podía anotar sólo el número de los que han fallado dentro de las 1000 horas, o, en cambio, podría tomar nota de los tiempos de falla de cada uno de ellos.

En general, cada resultado de un experimento puede ser asociado con un número especificando una regla de asociación o función. Tal regla se llama **variable aleatoria**, *variable* porque diferentes valores numéricos son posibles y *aleatoria* porque el valor observado depende de cuál de los posibles resultados experimentales resulte.

**Definición 1.1.** *Para un espacio muestral $\mathcal{S}$ de algún experimento, una* ***variable aleatoria*** *(v.a) es una función que asigna a cada elemento de $\mathcal{S}$ un número real. Es decir, si $X$ es una v.a de $\mathcal{S}$, $X: \mathcal{S} \to \mathbb{R}$*

Por ejemplo, dado que se tira una moneda dos veces y sea $X$ la v.a. tal que $X$: "Número de secas obtenidas luego de los dos tiros". El espacio muestral del experimento será:

$$\mathcal{S} = \{(c,c),(c,s),(s,c),(s,s)\}$$

entonces

$$X(c,c) = 0 \qquad X(c,s) = X(s,c) = 1 \qquad X(s,s) = 2$$

Luego, el rango (o imagen) de $X$ es $R_X = \{0, 1, 2\}$

Las variables aleatorias se clasifican según su rango (o imagen). Sea $X$ una v.a. con rango $R_X$, si $R_X$ es un conjunto *finito* o *infinito numerable*, entonces se dice que $X$ es una **variable aleatoria discreta**. En cambio, si $R_X$ es un conjunto *infinito no numerable* entonces $X$ es una **variable aleatoria continua**

## 2 Esperanza o valor esperado

La **esperanza**, **valor esperado** o **valor medio** de una variable aleatoria $X$ se puede definir informalmente como el promedio ponderado de los valores del rango de la variable, donde los "+pesos" de cada valor $x_i$ es la probabilidad $P(X = x_i)$, es decir, la probabilidad de que $X$ tome el valor $x_i$. El valor esperado de $X$ se anota $E(X)$

Para calcular la esperanza de una variable aleatoria en necesario conocer su distribución de probabilidad, lo que indica cómo está distribuida (asignada) la probabilidad total de 1 entre todos los valores posibles de la variable. Para las variables discretas se define una *función de masa de probabilidad* y para las continuas una *función de densidad de probabilidad.* Sean $X$ una variable aleatoria discreta e $Y$ una variable continua, se define a su valor esperado como:

$$E(X) = \sum_{x_i \in R_X} x_i P(X = x_i) \qquad\qquad E(Y) = \int_{-\infty}^{\infty} y_i P(Y = y_i)\ dx$$

Indistintamente de que la variable aleatoria sea discreta o continua, la esperanza resulta ser una *función lineal*, es decir, dados $a_0, a_1 \in \mathbb{R}$ y una v.a $X$, se cumple:

$$E(a_0 + a_1 X) = a_0 + a_1 E(X)$$

## 3 Varianza

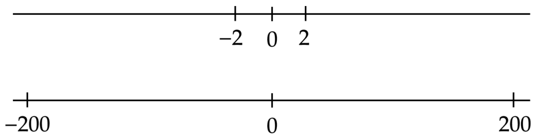
Mientras que la esperanza de una v.a. determina dónde está centrada la distribución de probabilidad, la **varianza** mide cuán dispersos o "alejados" están los valores de la variable con respecto a su esperanza.

Veamos que dadas dos variables aleatorias $X$ e $Y$ discretas, con las siguientes distribuciones de probabilidad

| $x$ | -2 | 2 |
|---|---|---|
| $p(x)$ | 0.5 | 0.5 |

| $y$ | -200 | 200 |
|---|---|---|
| $p(y)$ | 0.5 | 0.5 |

se puede verificar fácilmente, por lo visto en la sección anterior, que $E(X) = E(Y) = 0$, pero notemos que los valores que toma la variable $Y$ están mucho mas "alejados" de su valor medio que los valores de $X$:




El concepto de varianza refleja esta situación, asignando valores más grandes a aquellas variables aleatorias cuyos valores se encuentran más lejos de sus respectivas esperanzas.

Dada una variable aleatoria $X$ cuya esperanza es $E(X) = \mu$ se define a la varianza de $X$ de la siguiente manera:

$$V(X) = E[(X - \mu)^2] = \sigma^2$$

La cantidad $(X - \mu)^2$ es el cuadrado de la desviación de X desde su valor medio, y la varianza resulta la esperanza de dicha desviación al cuadrado. Cuanto más cerca se encuentren los valores de X del valor esperado $\mu$, menor será la varianza. De forma contraria, si hay valores alejados de $\mu$ que tengan alta probabilidad, entonces $\sigma^2$ será grande.

Notemos que la varianza de una v.a. nunca es negativa

Tanto para variables discretas como continuas valen las siguientes propiedades:

Dados $X$ una variable aleatoria, $\mu = E(X)$ y $a$ y $b$ constantes

$$V(X) = E(X^2) - \mu^2$$

$$V(aX + b) = a^2 V(X) \qquad y \qquad \sigma_{aX+b} = |a|\sigma_X$$

## 4 Estadísticos y Estimación puntual

El objetivo de la inferencia estadística casi siempre es obtener algún tipo de conclusión sobre uno o más parámetros (características) de las observaciones realizadas.

A menudo, en los problemas de inferencia estadística es poco práctico o imposible analizar la totalidad de las observaciones, es decir, toda la **población**. En ese caso el investigador debe tomar una parte o subconjunto de la población denominada **muestra**

Para que las inferencias sean válidas, la muestra debe ser representativa de la población. Se selecciona una **muestra aleatoria** como el resultado de un mecanismo aleatorio. En consecuencia, la selección de una muestra es un *experimento aleatorio*, y cada observación de la muestra es el valor observado de una *variable aleatoria*. Las observaciones en la población determinan la distribución de probabilidad de la variable aleatoria.

**Definición 4.1.** *Se dice que las variables aleatorias $X_1, X_2, \ldots, X_n$ forman una* ***muestra aleatoria*** *simple de tamaño $n$ si*

1. *Las $X_i$ son variables aleatorias independientes.*
2. *Cada $X_i$ tiene la misma distribución de probabilidad.*

Esta variación en los valores observados implica, a su vez, que el valor de cualquier función de las observaciones muestrales, tal como la media muestral o la desviación estándar muestral también varía de una muestra a otra. Dicha función o cantidad dependiente de los datos muestrales observados recibe el nombre de **estadístico**.

**Definición 4.2.** *Un* ***estadístico*** *es cualquier cantidad cuyo valor puede ser calculado a partir de datos muestrales (función). Antes de obtener los datos, existe incertidumbre sobre el valor del estadístico particular. En consecuencia, un estadístico es una variable aleatoria.*

El objetivo de la estimación puntual es seleccionar un solo número (*puntual*), con base en los datos muestrales, que represente un valor sensible o aproximado (estimación) de un parámetro

$\theta$ de la población observada.

**Definición 4.3.** *Una* ***estimación puntual*** *de un parámetro $\theta$ es un número único que puede ser considerado como un valor sensible de $\theta$. Se obtiene una estimación puntual seleccionando un estadístico apropiado y calculando su valor con los datos muestrales dados. El estadístico seleccionado se llama* ***estimador puntual*** *de $\theta$ y se lo suele anotar $\hat{\theta}$.*

A continuación veremos estadisticos usualmente utilizados como estimadores puntuales ...

Sean $X_1, X_2, \ldots, X_n$ una muestra aleatoria de la v.a $X$ donde

$$E(X) = \mu \qquad \text{y} \qquad V(X) = \sigma^2$$

En primer lugar, si suponemos que no conocemos el valor de $\mu$ (*media poblacional*), se utiliza como aproximación de éste al estadístico conocido como **media o promedio muestal**:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Es decir, la media muestral es un *estimador puntual* del parámetro $\mu$. Simbólicamente $\hat{\mu} = \overline{X}$.

Por otro lado si desconocemos el valor de $\sigma^2$, un estadístico utilizado como su estimador puantual es la **varianza muestral** ($\hat{\sigma}^2 = S^2$):

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Finalmente, para estimar el valor de la *desviación estándar poblacional* el estadístico que se utiliza es la **desviación estándar muestral** ($\hat{\sigma} = S$):

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

Es claro que, en el mejor de todos los mundos, sería ideal hallar un estimador $\hat{\theta}$ de un parámetro $\theta$ con el cual $\hat{\theta} = \theta$ siempre, es decir, que el estimador siempre de como resultado el valor verdadero del parámtero. Sin embargo, $\hat{\theta}$ es una función de las $X_i$ muestrales, por lo que es una variable aleatoria. Con algunas muestras, $\hat{\theta}$ dará un valor más grande que $\theta$, mientras que con otras muestras $\hat{\theta}$ subestimará $\theta$.

Si pensamos

$$\hat{\theta} = \theta + \text{error de estimación}$$

resulta evidente que un estimador preciso (e ideal) sería uno que genere errores de estimación pequeños, consiguiendo así que los valores estimados se acerquen al valor verdadero.

Ahora pensemos a un estimador puntual como un instrumento de medición que puede: estar calibrado con precisión o dar sistemáticamente valores menores al valor verdadero que se está midiendo. Si el instrumento está calibrado, las mediciones producidas se distribuirán en torno al valor verdadero de tal modo que en promedio este instrumento mide lo que se propone medir, por lo que se conoce como instrumento *insesgado*. En cambio, si siempre resulta dar valores más pequeños, se dice que el instrumento que tiene un *sesgo* sistemático.

**Definición 4.4.** *Un estimador puntual $\hat{\theta}$ es un* ***estimador insesgado*** *de $\theta$ si $E(\hat{\theta}) = \theta$ con todo valor posible de $\theta$. Si $\hat{\theta}$ no es insesgado, la diferencia $E(\hat{\theta}) - \theta$ se conoce como el* ***sesgo*** *de $\hat{\theta}$.*

Es decir, $\hat{\theta}$ es insesgado si su distribución de probabilidad siempre está "centrada" en el valor verdadero del parámetro.

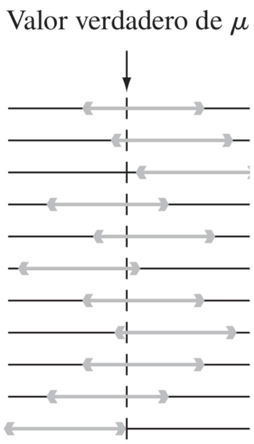
Notar que los estadísticos mencionados anteriormente resultan ser estimadores puntuales insesgados de la media, varianza y desviación poblacional, respectivamente.

# 5 Intervalos de confianza

Como vimos en la sección anterior, la estimación puntual de un parámetro da como resultado un solo número. Éste no proporciona información sobre la precisión y confiabilidad de esta estimación. Consideremos, por ejemplo, utilizar el estadístico $\overline{X}$ para calcular una estimación puntual de la media poblacional ($\mu$) en un experimento. Debido a la variabilidad del muestreo, casi nunca es el caso de que el valor calculado $\overline{x}$ cumpla $\overline{x} = \mu$. La estimación puntual no dice nada sobre qué tan *cerca* pudiera estar a $\mu$.

Una variante a dar un solo número puntual es calcular un intervalo completo de los valores posibles del parámetro estudiado, un **intervalo de confianza** (IC). Para construir el intervalo de confianza primero debe determinarse un **nivel de confianza**, es decir, el grado de confiabilidad de que el verdadero valor del parámetro se encuentre en dicho intervalo ¿Pero qué indica este "nivel de confianza" en verdad?

Pensemos que deseamos hallar un intervalo de confianza del 95% para la media $\mu$ de una población normalmente distribuída cuando se conoce el valor de $\sigma$. Entonces si tomáramos muestras aleatorias una y otra vez, y calculáramos el intervalo de confianza, a la larga, el 95% de los intervalos resultantes contendrán a $\mu$.




La variación en los intervalos es originada por la participación del estadístico $\overline{X}$ en su cálculo. En particular, los intervalos son de la forma:

$$\left(\overline{X} - 1,96\ \frac{\sigma}{\sqrt{n}} \quad ; \quad \overline{X} + 1.96\ \frac{\sigma}{\sqrt{n}}\right)$$

El intervalo es aleatorio porque los dos puntos extremos del intervalo implican a la variable aleatoria $\overline{X}$. Esto significa que, en cada una de las muestras aleatorias mencionadas, el valor de $\overline{X}$ va cambiando y por ende el de los extremos del intervalo.

Por otro lado el **ancho** del intervalo da información sobre la precisión de la estimación de intervalo. Si el nivel de confianza es alto y el intervalo resultante es angosto, el conocimiento del valor del parámetro es razonablemente preciso. Un muy amplio intervalo de confianza, sin embargo, indica que existe gran cantidad de incertidumbre sobre el valor verdadero de lo que se está estimando.

Entonces, si podemos elegir el nivel de confianza, ¿por qué no elegimos un nivel mayor al 95%, como 99% o 100%? Esto es debido a que cuanto más alto es el grado de confianza, más ancho es el intervalo resultante. En particular, el único intervalo de 100% para $\mu$ es $(-\infty, \infty)$, que no nos aporta mucha información sobre el valor verdadero de este parámetro. Por lo tanto notemos que al elegir un alto nivel de confianza estamos "sacrificando" precisión en la estimación.

## 6 Tests de Hipótesis

Como hemos visto, el valor de un parámetro puede ser estimado a partir de datos muestrales mediante estimadores puntuales o intervalos de confianza. Con frecuencia, sin embargo, el objetivo de una investigación no es estimar un parámetro sino decidir cuál de dos presunciones contradictorias sobre el parámetro es la correcta. Los métodos para conseguir esto forman parte de la inferencia estadística llamada *prueba de hipótesis*.

Una **hipótesis estadística** es una aseveración sobre el valor de uno o varios parámetros de una población. Un ejemplo de una hipótesis es la pretensión de que $\mu = 1000$ es el número de horas promedio de funcionamiento de las componentes electrónicas fabricadas en una empresa.

En cualquier problema de prueba de hipótesis siempre existen dos hipótesis contradictorias que se consideran. Una podría ser la presunción $\mu = 1000$ y la otra $\mu \neq 1000$. O bien, si se toma como aseveración que la proporción de componentes electrónicas defectuosas es $p < 0.1$, la hipótesis contraria es que $p \geq 0.1$. El objetivo de estas pruebas es determinar, en base a los datos muestrales, qué afirmación es la correcta.

Así como dice la frase conocida "serás inocente hasta que se demuestre lo contrario", en el caso de las pruebas de hipótesis, siempre una de las ellas será inicialmente favorecida, es decir, se tomará como cierta. Luego, esta pretensión inicial no será rechazada a menos que se presente evidencia muestral suficiente que la contradiga y apoye fuertemente la hipótesis alternativa.

**Definición 6.1.** *La* ***hipótesis nula*** *denotada por $H_0$, es la pretensión que inicialmente se supone cierta. La* ***hipótesis alternativa*** *denotada por $H_a$, es la aseveración contradictoria a $H_0$.*

La hipótesis nula será rechazada en favor de la hipótesis alternativa sólo si la evidencia muestral sugiere que $H_0$ es falsa. Si la muestra no contradice fuertemente a $H_0$, se continuará creyendo en la verdad de la hipótesis nula. Las dos posibles conclusiones derivadas de un análisis de prueba de hipótesis son entonces *rechazar* $H_0$ o *no rechazar* $H_0$.

Una **prueba de hipótesis** es el método de utilizar datos muestrales para decidir si la hipótesis nula debe ser rechazada. Por consiguiente se podría probar $H_0 : \mu = 1000$ contra $H_a : \mu \neq 1000$. Sólo si los datos muestrales sugieren fuertemente que $\mu$ es diferente de 1000 deberá ser rechazada la hipótesis nula. En el tratamiento de la prueba de hipótesis, $H_0$ siempre será formulada como una afirmación de igualdad. Si $\theta$ denota el parámetro de interés, la hipótesis nula tendrá la forma $H_0 : \theta = \theta_0$ donde $\theta_0$ es un número específico llamado **valor nulo** del parámetro.

Una vez recordados estos conceptos, pasemos a especificar cómo es el proceso de prueba de una hipótesis. Como mencionamos, deseamos determinar si existe información en los datos muestrales que funcionen como contradicción a nuestra hipótesis nula. De esta forma, el procedimiento de prueba consta de:

1. Un **estadístico de prueba**, una función de los datos muestrales en la cual se basará la decisión.

2. Una **región de rechazo**, el conjunto de todos los valores del estadístico de prueba para los cuales $H_0$ será rechazada.

Por lo tanto, la hipótesis nula será rechazada si y sólo si el valor estadístico de prueba calculado queda en la región de rechazo. La base de elección del rango de valores de rechazo radica en la consideración de los errores que podrían cometerse al sacar una conclusión. En una prueba de hipótesis siempre existen dos tipos de errores:

Un **error de tipo I** consiste en rechazar la hipótesis nula cuando ésta es verdadera.
Un **error de tipo II** consiste en no rechazar la hipótesis nula cuando ésta es falsa.

La dificultad con la utilización de un procedimiento basado en datos muestrales es que, debido a la *variabilidad* del muestreo, el resultado podría ser una muestra no representativa, resultando en errores a la hora de tomar una decisión.

Como no podemos considerar un procedimiento que no tengan errores (son aquellos en los que se prueba con toda la población), habrá que buscar procedimientos con los cuales sea *improbable* que ocurra cualquier tipo de error. Es decir, un buen procedimiento es uno con el cual la probabilidad de cometer cualquier tipo de error es pequeña. La selección de un valor de corte en una región de rechazo particular fija las probabilidades de errores de tipo I y tipo II. Estas probabilidades de error son usualmente denotadas por $\alpha$ y $\beta$ respectivamente.

Con respecto a estas dos probabilidades de error debemos aclarar lo siguiente: Suponiendo que un tenemos un experimento y un tamaño de muestra fijos, dado un estadístico de prueba, entonces si se reduce el tamaño de la región de rechazo para obtener un valor más pequeño de $\alpha$ se obtiene un valor más grande de $\beta$. De esta forma, no existe una región de rechazo que haga que al mismo tiempo $\alpha$ y $\beta$ sean pequeños. Se debe seleccionar una región para establecer un "compromiso" entre $\alpha$ y $\beta$. El método más utilizado para esto es especificar el valor más grande de $\alpha$ que pueda ser tolerado y encontrar una región de rechazo que tenga valor de $\alpha$. Esto hace a $\beta$ tan pequeño como sea posible dependiendo del límite en $\alpha$.

Al valor resultante de $\alpha$ se lo conoce como **nivel de significación** de la prueba y el procedimiento de prueba correspondiente se llama **prueba de nivel** $\alpha$. Usualmente el nivel de significación será pequeño, del orden de 0,1, 0,01 o 0,05, como veremos en las pruebas de hipótesis sobre parámteros de regresión lineal.

Finalmente, cabe notar que la probabilidad de estos dos errores $\alpha$ y $\beta$ esta atada a la distribución de probabilidad del estadístico que se elija para la prueba.

# 7 Aclaraciones

Este apunte preliminar de conceptos de estadística se desarrolló en su mayoría sin ejemplos ya que, por un lado, fueron vistos previamente en Matemática 3, y por otro, el uso que les daremos a los conceptos será con un enfoque hacia el método de regresión lineal.

A pesar de esto, no dejamos de destacar la importancia de la probabilidad y la estadística para el desarrollo de modelos matemáticos e informáticos, así como para el análisis de datos.

Aquellas personas que deseen profundizar, tanto en ejemplos como en conceptos, pueden dirigirse a los apuntes provistos por la profesora María Beatriz Pintarelli en el sitio del departamento de matemática de la Facultad de Ciencias Exactas: Apuntes Matemática 3

# Preliminares - Lógica y Conjuntos

## 1 Lógica proposicional

**Definición 1.1.** *Se llama **proposición** a toda oración que tiene un valor de verdad. Es decir, una **proposición** es una oración que puede ser verdadera o falsa pero no ambas a la vez.*

Sea p una proposición; se define *valor de verdad de p* y se escribe V(p) = V, si p es una proposición verdadera y V(p) = F, si p es una proposición falsa. Recordemos que a cada proposición le corresponde un único valor de verdad de los dos posibles.

Por lo general, a las proposiciones se las representa por las letras del alfabeto desde la letra p, es decir, p, q, r, s, t,... etc.

Así, por ejemplo, podemos citar las siguientes proposiciones y su valor de verdad:
p: $15 + 5 = 21$ (F)
q: Santa Fe es una provincia Argentina. (V)
r: El número 15 es divisible por 3. (V)
s: El perro es un ave. (F)

### Expresiones No Proposicionales

Son aquellos enunciados a los que no se les puede asignar un valor de verdad. Entre ellos tenemos a los exclamativos, interrogativos o imperativos, por ejemplo:

- ¿Cómo te llamas?
- Prohibido pasar.

- Haz lo tuyo, muchacho

No son proposiciones porque no se les puede asignar un valor de verdad.

#### Clasificación de las Proposiciones

Aquellas proposiciones que no se pueden descomponer se llaman **simples o atómicas**. Por ejemplo, sea la proposición $p : 3 + 6 = 9$ es una proposición simple o atómica.

Cuando una proposición consta de dos o más enunciados simples, se le llama **proposición compuesta o molecular**.
Así, por ejemplo, la proposición: "*Pitágoras era griego y geómetra*" es una proposición compuesta por las proposiciones simples,*p : Pitágoras era griego* y *q : Pitágoras era geómetra*.

No es necesario conocer si una afirmación es verdadera o falsa (es decir, su valor de verdad) para saber que es una proposición. Por ejemplo: "Hay vida extraterrestre" es una proposición, independientemente de que algunos crean que es verdadera y otros que es falsa, puesto que claramente o bien existe vida extraterrestre o bien no existe.

## 1.2 Conectivos Lógicos

Como dijimos a partir de proposiciones simples es posible generar otras, simples o compuestas. La propiedad fundamental de una proposición compuesta es que su valor de verdad está completamente determinado por los valores de verdad de las proposiciones que la componen.
Estudiaremos las distintas formas de conectar proposiciones entre sí, es decir como se puede operar con proposiciones, y para ello utilizaremos los llamados **conectivos lógicos**.

### 1.2.1 Operaciones Proposicionales

Definiremos las operaciones entre proposiciones, de las que se conoce su valor de verdad, y veremos como asignar a la proposición resultante el suyo.

**Negación**

Dada una proposición $p$, se denomina la negación de $p$ a otra proposición denotada por $\neg p$ (se lee no $p$) que le asigna el valor de verdad opuesto al de $p$.

Por ejemplo:

$p$: Diego estudia matemática.

$\neg p$: Diego no estudia matemática.

Por lo que nos resulta sencillo construir su tabla de verdad:

| $p$ | $\neg p$ |
|---|---|
| V | F |
| F | V |

Se trata de una operación unaria, pues a partir de una sola proposición se obtiene otra, que es su negación.

Otro ejemplo: La negación de $p$: Santa Fe es una provincia argentina, es:

$\neg p$: Santa Fe no es una provincia argentina.

**Conjunción**

Dadas dos proposiciones $p$ y $q$, se denomina conjunción de estas proposiciones a la proposición $p \wedge q$ (se lee "$p$ y $q$"), cuya tabla de verdad es:

| $p$ | $q$ | $p \wedge q$ |
|---|---|---|
| V | V | V |
| V | F | F |
| F | V | F |
| F | F | F |

La tabla que define esta operación, establece que la conjunción es verdadera sólo si lo son las dos proposiciones componentes. En todo otro caso es falsa.

Ejemplos: Sea la declaración:

$$5 \textit{ es un número impar } \text{ y } 6 \textit{ es un número par}$$

Vemos que está compuesta de dos proposiciones a las que simbolizaremos por:

$p$:*5 es un número impar*

$q$: 6 es un número par

Por ser ambas verdaderas, la conjunción es verdadera.

Ahora bien, sea la declaración:

*5 es un número impar y 7 es un número par*

Esta conjunción es falsa, ya que no son simultáneamente verdaderas las proposiciones que la componen.

**Disyunción**

Dadas dos proposiciones $p$ y $q$, la disyunción de las proposiciones $p$ y $q$ es la proposición compuesta $p \vee q$ (se lee "$p$ o $q$"), cuya tabla de verdad es:

| $p$ | $q$ | $p \vee q$ |
|---|---|---|
| V | V | V |
| V | F | V |
| F | V | V |
| F | F | F |

Ejemplos:

*Córdoba es una provincia argentina o Uruguay es un país latinoamericano.*

*El 3 es par o el 8 es primo.*

Como vemos en la tabla y se desprende de los ejemplos para que la disyunción sea verdadera *alcanza* con que al menos una de las proposiciones que la componen lo sea, o dicho de otra forma, una disyunción será falsa sólo si todas las proposiciones que la componen lo son.

**Disyunción Exclusiva**

Dadas dos proposiciones $p$ y $q$, la disyunción exclusiva de las proposiciones $p$ y $q$ es la proposición compuesta $p \veebar q$ (también $p \oplus q$), se lee "$p$ o $q$ pero no ambas" (a veces $p$ XOR $q$), cuya tabla de verdad es:

| $p$ | $q$ | $p \veebar q$ |
|---|---|---|
| V | V | F |
| V | F | V |
| F | V | V |
| F | F | F |

Ejemplos:

*El televisor está prendido o está apagado*

*La proposición es verdadera o es falsa*

**Condicional (o implicación)**

Consideremos el enunciado: *Si paga en efectivo obtiene un descuento*

Este enunciado está formado por dos proposiciones atómicas:

$p$: *Paga en efectivo*

$q$: *Obtiene un descuento*

Lo que nuestro enunciado original afirma es esto: si "pasa" $p$ entonces "pasa" $q$, si $p$ es verdad, entonces $q$ también es verdad, o, dicho de modo más sencillo, si $p$ entonces $q$.

En el enunciado $p \to q$, se dice que $p$ es el antecedente y $q$ el consecuente.

El condicional $p \to q$ se lee "$p$ condicional $q$", "$p$ implica $q$" o bien "si $p$ entonces $q$".

Un condicional siempre es verdadero, excepto cuando el antecedente es verdadero y el consecuente falso.

Por lo tanto, su valor de verdad queda definido por la siguiente tabla de verdad.

| $p$ | $q$ | $p \to q$ |
|---|---|---|
| V | V | V |
| V | F | F |
| F | V | V |
| F | F | V |

Otras expresiones que representan también la proposición "*si p entonces q*" y que se simbolizan por $p \to q$:

- p sólo si q
- q si p
- p es condición suficiente para q
- q es condición necesaria para p

**Ejemplos**:
- Obtiene el descuento si paga en efectivo.
- Es suficiente pagar en efectivo para obtener el descuento.
- Ser tucumano es condición sufiente para ser argentino, pero no necesaria (podría ser de otra provincia).
- Es necesario ser argentino para ser cordobés, pero no suficiente (no alcanza con ser argentino, es necesario pero además hay que nacer en algún lugar de Córdoba)

**Condicionales asociados**

Se puede ver por medio de las tablas de verdad, que tanto la conjunción como la disyunción tienen la **propiedad conmutativa**, es decir el orden de las componentes de una conjunción o de una disyunción no altera su valor de verdad: es lo mismo $p \wedge q$ que $q \wedge p$, y también es lo mismo $p \vee q$ que $q \vee p$. Pero, ¿ocurre lo mismo con el condicional? ¿Es lo mismo $p \to q$ que $q \to p$? La respuesta es que no.

**El recíproco del condicional**

Se dice que $q \to p$ es el **recíproco** de $p \to q$. El implicador no tiene la propiedad conmutativa y esto se aprecia en la comparación de las tablas de verdad de $p \to q$ y de su recíproco $q \to p$:

| $p$ | $q$ | $p \to q$ | $q \to p$ |
|---|---|---|---|
| V | V | V | V |
| V | F | F | V |
| F | V | V | F |
| F | F | V | V |

Veámoslo con un ejemplo:

Sean $p$ : Ahora llueve y $q$ : El suelo está mojado, siendo, por consiguiente $p \to q$ : Si ahora llueve, entonces el suelo está mojado. Veamos el recíproco de este enunciado: $q \to p$: Si el suelo está mojado entonces ahora llueve. Supongamos que $p$ es falso, y $q$ verdadero, lo que se corresponde con la tercera fila de la tabla anterior.

- $p \to q$ (Si ahora llueve, entonces el suelo está mojado) es verdadero.
- $q \to p$ (Si el suelo está mojado, entonces ahora llueve) es falso.

**El contrarrecíproco del condicional**

Aunque un enunciado condicional y su recíproco no tienen los mismos valores de verdad, si los tienen el condicional y su contrarrecíproco.

El contrarrecíproco del enunciado $p \to q$ es $\neg q \to \neg p$. Veámoslo comparando tablas de verdad:

| $p$ | $q$ | $\neg q$ | $\neg p$ | $p \to q$ | $\neg q \to \neg p$ |
|---|---|---|---|---|---|
| V | V | F | F | V | V |
| V | F | V | F | F | F |
| F | V | F | V | V | V |
| F | F | V | V | V | V |

Comparemos el mismo ejemplo: En el ejemplo anterior donde $p$: Ahora llueve, $q$: El suelo está mojado, $p \to q$ : Si ahora llueve entonces el suelo está mojado.

El contrarrecíproco es $\neg q \to \neg p$ : Si el suelo no está mojado entonces ahora no llueve, que es lógicamente equivalente al enunciado primitivo $p \to q$.

**El contrario del condicional**

Se dice que $\neg p \rightarrow \neg q$ es el **contrario** de $p \rightarrow q$.

| $p$ | $q$ | $\neg p$ | $\neg q$ | $\neg p \rightarrow \neg q$ |
|---|---|---|---|---|
| V | V | F | F | V |
| V | F | F | V | V |
| F | V | V | F | F |
| F | F | V | V | V |

**Bicondicional**

Ya hemos comprobado que $p \rightarrow q$ no es lo mismo que $q \rightarrow p$. Puede ocurrir, sin embargo, que tanto $p \rightarrow q$ como $q \rightarrow p$ sean verdaderos. Por ejemplo, si $p$: La Tierra es plana, y $q$: El Sol es un planeta,

entonces tanto $p \rightarrow q$ como $q \rightarrow p$ son verdaderos, porque tanto $p$ como $q$ son falsos. Es necesario tener esto en cuenta para entender bien el concepto de bicondicional. Mediante el bicondicional ($\leftrightarrow$) lo que queremos decir es que un enunciado es a la vez condición necesaria y suficiente para otro. El **bicondicional** $p \leftrightarrow q$, que se lee "$p$ si y sólo si $q$".

Así, si digo que $p$: apruebo Filosofía y $q$: saco un 5 o más en el examen de Lógica la fórmula $p \leftrightarrow q$ significa apruebo Filosofía si y sólo si saco un 5 o más en el examen de Lógica.

La proposición compuesta: apruebo Filosofía si y sólo si saco un 5 o más en el exámen de Lógica, se puede formalizar de dos formas equivalentes: $(p \rightarrow q) \wedge (q \rightarrow p)$, o bien $p \leftrightarrow q$. En consecuencia, el enunciado $p \leftrightarrow q$ queda definido por el enunciado $(p \rightarrow q) \wedge (q \rightarrow p)$. Por esta razón, el símbolo $\leftrightarrow$ se llama bicondicional, y la tabla de verdad para $p \leftrightarrow q$ es la misma que la de $(p \rightarrow q) \wedge (q \rightarrow p)$.

| $p$ | $q$ | $p \leftrightarrow q$ |
|---|---|---|
| V | V | V |
| V | F | F |
| F | V | F |
| F | F | V |

La doble flecha horizontal $\leftrightarrow$ es el operador bicondicional.

Mirando con atención la tabla de verdad deducimos que para que $p \leftrightarrow q$ sea verdadera, tanto $p$

como $q$ han de tener los mismos valores de verdad, y en caso contrario es falsa.

También de la observación de la tabla notamos que es igual a la correspondiente a $\neg(p \veebar q)$

**Formalización del bicondicional**

El bicondicional puede tener varias expresiones equivalentes en lenguaje natural. Así $p \leftrightarrow q$ es la formalización de las siguientes expresiones de lenguaje natural:

- $p$ si y sólo si $q$
- $p$ es necesario y suficiente para $q$

Notar que $p \leftrightarrow q$ y $q \leftrightarrow p$ tendrían totalmente los mismos valores de verdad, puesto que ambas son coimplicaciones y por lo tanto si sus valores de verdad son los mismos, son verdaderas, y son falsas en los demás casos. En consecuencia, podemos reformular los enunciados anteriores intercambiando $p$ y $q$.

**Ejemplos del bicondicional**

Ejemplos de bicondicionales verdaderos:

a. La Tierra es plana si y sólo si el Sol es un planeta; si llamamos: $p$:La Tierra es plana, sabemos que es Falsa y sea $q$: El Sol es un planeta, también es Falsa.

b. La Tierra es esférica si y sólo si el Sol es una estrella, sea $p$: La Tierra es esférica es Verdadera, y sea $q$: El Sol es una estrella, también es Verdadera.

Ejemplos de bicondicionales falsos:

a. La Tierra es plana si y sólo si el Sol es una estrella; si llamamos: $p$: La Tierra es plana, sabemos que es Falsa y sea $q$: El Sol es una estrella, es Verdadera.

b. La Tierra es esférica si y sólo si el Sol es un planeta, sea $p$: La Tierra es esférica, es Verdadera, y sea $q$: El Sol es un planeta, en cambio, es Falsa.

**Equivalencia Lógica**

**Definición 1.2.** *Decimos que dos proposiciones $P$ y $Q$ formadas ambas por las mismas letras proposicionales, son **lógicamente equivalentes**, o simplemente **equivalentes**, si coinciden sus valores de verdad para los mismos valores de verdad de las componentes. Se nota como $P \Leftrightarrow Q$, siendo $P$ y $Q$ formas proposicionales no necesariamente atómicas.*

*Importante: usamos la $\Leftrightarrow$ para indicar la equivalencia, mientras que $\leftrightarrow$ simboliza al bicondicional.*

Decimos que dos proposiciones $P$ y $Q$ formadas ambas por las mismas letras proposicionales, son lógicamente equivalentes, o simplemente equivalentes, si coinciden sus valores de verdad para los mismos valores de verdad de las componentes.
Se nota como $P \Leftrightarrow Q$, siendo $P$ y $Q$ formas proposicionales no necesariamente atómicas.
Importante: usamos la $\Leftrightarrow$ para indicar la equivalencia, mientras que $\leftrightarrow$ simboliza al bicondicional.

**Propiedades**

1.- Doble negación: $\neg\neg p \Leftrightarrow p$

2.- Leyes conmutativas: a) $(p \wedge q) \Leftrightarrow (q \wedge p)$ b) $(p \vee q) \Leftrightarrow (q \vee p)$

3.- Leyes asociativas: a) $[(p \wedge q) \wedge r] \Leftrightarrow [p \wedge (q \wedge r)]$ b) $[(p \vee q) \vee r] \Leftrightarrow [p \vee (q \vee r)]$

4.- Leyes distributivas: a) $[p \vee (q \wedge r)] \Leftrightarrow [(p \vee q) \wedge (p \vee r)]$ b) $[p \wedge (q \vee r)] \Leftrightarrow [(p \wedge q) \vee (p \wedge r)]$

5.- Leyes de De Morgan: a) $\neg(p \vee q) \Leftrightarrow (\neg p \wedge \neg q)$ b) $\neg(p \wedge q) \Leftrightarrow (\neg p \vee \neg q)$

6.- Implicación: $(p \rightarrow q) \Leftrightarrow (\neg p \vee q)$

| $p$ | $q$ | $r$ | $p \wedge q$ | $q \wedge r$ | $(p \wedge q) \wedge r$ | $p \wedge (q \wedge r)$ |
|---|---|---|---|---|---|---|
| V | V | V | V | V | V | V |
| V | V | F | V | F | F | F |
| V | F | V | F | F | F | F |
| V | F | F | F | F | F | F |
| F | V | V | F | V | F | F |
| F | V | F | F | F | F | F |
| F | F | V | F | F | F | F |
| F | F | F | F | F | F | F |

Notar que las últimas dos columnas coinciden los valores de verdad para los mismos valores de verdad de las componentes $p$, $q$ y $r$. Entonces, si miramos la tabla del bicondicional estaremos en la primera o cuarta fila, por lo que $P \leftrightarrow Q$ resultará verdadera.

### Esquemas proposicionales en una indeterminada

En Álgebra y Aritmética suele decirse que la siguiente expresión: $x + 2 = 5$ es una ecuación. Tal expresión no es una proposición, pues no tiene sentido afirmar que sea verdadera o falsa, pero existe algún reemplazo de $x$ por un número de modo tal que se transforma en una proposición.

Por ejemplo, si reemplazamos $x$ por 7 queda la expresión $7 + 2 = 5$, es una proposición, la cual en este caso es Falsa. Si reemplazamos $x$ por 3 queda la expresión $3 + 2 = 5$, es una proposición, la cual en este caso es Verdadera.

**Definición 1.3.** $p(x)$ *es un* ***esquema proposicional en la variable o indeterminada*** $x$ *si y sólo si existe, al menos, una sustitución de* $x$ *por una constante que la transforma en proposición.*
*Esto es, "Existe por lo menos un nombre tal que la expresión obtenida sustituyendo la indeterminada por dicho nombre, es una proposición".*

Convención: Llamaremos simplemente esquema en lugar de esquema proposicional. Las indeterminadas suelen llamarse variables o incógnitas.

Ejemplos

1. "La x es blanca" es esquema pues existe una constante "flor" que si ocupa el lugar de la variable $x$ produce la siguiente proposición: "La flor es blanca".

Que esta proposición sea Verdadera o Falsa dependerá de cual sea la flor particular que se está eligiendo.

2. ¿Qué es x? NO es un esquema, pues no hay constante que sustituida en la variable produzca una proposición.

**Definición**: Si $p(x)$ es un esquema en $x$ y $a$ es una constante, se llama **valor de** $p(x)$ **en la constante** $a$ a la expresión obtenida de $p(x)$ sustituyendo $x$ por $a$. El valor de $p(x)$ para $a$ se designa $p(a)$.

Se llama *conjunto Universal* $U$, a un conjunto del cual se extraen las constantes para reemplazar a la indeterminada.

Para vincular los esquemas proposicionales utilizamos los mismos conectivos ya estudiados y con los mismos significados;

Vamos a definir al conjunto de valores de verdad de $p$, lo simbolizamos con $\mathrm{V}(p)$, al conjunto formado por todas las constantes $a$ que hacen verdadera la proposición $p(a)$.

## Cuantificadores: Universal y Existencial

Hasta ahora se ha visto un método para obtener proposiciones a partir de esquemas $p(x)$ consiste en sustituir la variable $x$ por una constante adecuada $a$ de tal forma que $p(a)$ sea una proposición.

Hay otro método distinto que transforma un esquema en proposición a partir del esquema $p(x)$, es el método de los **operadores** o **cuantificadores**.

Como vimos en los ejemplos, uno trata de reemplazar la incógnita por valores que tenga cierto sentido como para obtener una proposición, por ejemplo, si el esquema es $p(x) : x > 5$, pensamos que $x$ puede ser un número, y dependiendo de si $x$ es 8 ó $x$ es 2, será verdadera o falsa; pero no pensamos en reemplazar $x$ por algún color del arco iris. Esto nos conduce a la siguiente definición:

> **Definición 1.4.** *Llamaremos **conjunto universal** al conjunto de variables que al reemplazar la $x$ por un elemento de ese conjunto se obtenga una proposición. Lo notaremos por U y lo nombraremos por conjunto universal o, simplemente, universo. Debe contener, al menos, un elemento.*

Por medio de los cuantificadores podemos convertir en proposiciones a los esquemas de la siguiente manera:

**El cuantificador existencial**,

"Para algún $x$ se verifica $p(x)$"
"Existe $x$ tal que se cumple $p(x)$"
"Para al menos un $x$ se satisface $p(x)$"
son proposiciones que se escriben como $(\exists x)(p(x))$

*Ejemplo: Hay flores rojas.*

**El cuantificador universal**

"Para todo $x$ se verifica $p(x)$"
"Para cualquier $x$ tal que se cumple $p(x)$"
"Para cada $x$ se satisface $p(x)$"
son proposiciones que se escriben como $(\forall x)(p(x))$

*Ejemplo: Todas las flores son rojas.*

**Ejemplo**

Vamos a escribir en forma simbólica las siguientes proposiciones, (Necesitamos dar un conjunto universal y luego los esquemas, además de identificar el cuantificador):

- *Hay números pares*

  Como la propiedad de ser *par* es una propiedad de los números enteros podemos tranquilamente usar como conjunto universal al conjuntos de los números enteros.

  Luego, si $p(x)$ : x es par nuestra proposición quedaría simbolizada por: $(\exists x)p(x)$

- *Todos los números racionales son números enteros*

  Podemos usar como conjunto universal al conjunto de los números racionales, luego nuestra proposición se puede simbolizar como:

  $(\forall x)p(x)$ siendo $p(x)$ : $x$ es un número entero

  Ahora, observemos lo siguiente: Si en lugar de elegir de universo al conjunto de los números racionales elegimos al conjunto de los reales tendremos que tomar otras proposiciones para simbolizar la misma frase (ya que hay que dejar claro que sólo hablamos de los racionales y no de todos los elementos del conjunto, los reales). Vamos a necesitar usar más proposiciones y conectarlas.

  Sean $q(x)$ : $x$ es un número racional, y $p(x)$ : $x$ es un número entero, entonces *Todos los números racionales son números enteros* se simboliza por: $(\forall x)(q(x) \rightarrow p(x)$

**Alcance de un operador**

Pensemos en el siguiente ejemplo:

$$(\exists x)(v(x)) \wedge r(x) \; (*)$$

con $v(x) : x \; es \; verde$ y $r(x) : x \; es \; rojo$

Vemos que el operador existencial se refiere únicamente al esquema x es verde y NO a $x$ es rojo, o sea que el alcance del operador llega únicamente al primer esquema, si quisiéramos que alcance a los dos esquemas, tendríamos que poner

$$(\exists x)(v(x) \wedge r(x))$$

o sea, usaríamos paréntesis.

Del ejemplo precedente podemos deducir que: La expresión "x es verde" es el esquema más simple que aparece en $(*)$ inmediatamente después del operador.

La expresión "$x$ *es verde* $\wedge$ $x$ *es rojo*", también es un esquema pero no es el más simple. La expresión "$x$ *es rojo*" es un esquema también simple pero no aparece después del operador.

**Definición 1.5.** *Se llama* ***alcance de un operador en*** $x$ *al esquema más simple que aparece inmediatamente después del operador, salvo que se presenten paréntesis, en cuyo caso deben aplicarse las reglas habituales referentes al uso de paréntesis.*

**Negación de operadores**

Sea la siguiente proposición: $(\forall n)(p(n))$ con $p(n)$ : $n$ es un número primo, que en lenguaje coloquial dice: *Todos los números con primos* (la cuál sabemos es Falsa en el universo de los números naturales).

Vamos ahora a negarla: *No es cierto que todos los números sean primos*, $\neg(\forall n)(\mathrm{p(n)})$

En lenguaje corriente esto nos dice que no todos los números son primos que es lo mismo que si dijéramos: algunos números no son primos, y simbólicamente

$$(\exists n)(n \text{ no es un número primo})$$

De lo anterior se puede deducir, y vale de manera general y no sólo el ejemplo que:

$$\neg(\forall x)(p(x)) \Leftrightarrow (\exists x)(\neg p(x))$$

De manera análoga se obtiene:

$$\neg(\exists x)(p(x)) \Leftrightarrow (\forall x)(\neg p(x))$$

Por lo tanto, en palabras decimos que:

*La negación de un cuantificador universal (existencial, respectivamente) es equivalente a la afirmación de un cuantificador existencial (universal) cuyo alcance es la negación del alcance del primero.*

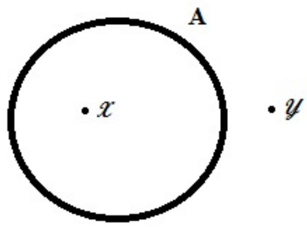
## Teoría de Conjuntos

Un **conjunto** es la reunión de entes u objetos bien definidos y diferenciables entre sí, que en general tienen características similares. Estos entes u objetos se llaman **elementos del conjunto**. Si $a$ es un elemento del conjunto $A$ se denota con la **relación de pertenencia** $a \in A$. En caso contrario, si $a$ no es un elemento de $A$ se denota $a \notin A$.

### Formas de Expresar un Conjunto

- **Diagrama de Venn**

  Seguramente esté familiarizado con gráficos como el siguiente, donde el elemento $x \in A$ y el elemento $y \notin A$.




- **Expresado por Extensión**

  Cuando damos de manera explícita los elementos, es decir, se especifica cuales son cada uno de los elementos de $A$. Por ejemplo: $A = \{1, 3, 5, 7\}$

- **Expresado por Comprensión**

  Cuando expresamos la propiedad que define el conjunto. En el ejemplo anterior, $A = \{x : x \text{ } es \text{ } un \text{ } nro. \text{ } impar \text{ } menor \text{ } o \text{ } igual \text{ } a \text{ } 7\} = \{x : x \text{ } es \text{ } impar \ \wedge x \leq 7\}$

  Se lee $x$ tal que $x$ es impar y $x$ es menor o igual que 7.

Considere $\{x : x > 1 \wedge 2.x = 1\}$ ¿Es un conjunto, qué elementos tiene?

`Ejemplos de Conjuntos muy utilizados en matemática.`

- Ya vimos en lógica, al conjunto universal o universo, que contiene a todos los elementos en un determinado contexto. Notado por la letra $\mathcal{U}$.

- El conjunto vacío, es aquél que no contiene elementos. Se nota $\emptyset$ ó sólo con $\{\}$.
- Conjuntos numéricos: números naturales, enteros, racionales, irracionales, reales, que desarrollaremos en el siguiente módulo.

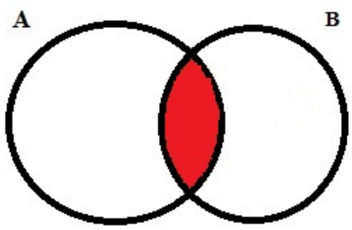
### Operaciones con Conjuntos

Sean A y B dos conjuntos. Diremos que:

- $A$ está contenido en $B$ ó $A$ es un subconjunto de $B$, si todo elemento de $A$ es también un elemento de $B$. En símbolos $A \subseteq B$ si y sólo si se verifica el condicional $(\forall x)(x \in A \rightarrow x \in B)$
- Los conjuntos $A$ y $B$ son iguales, si y sólo si tienen los mismos elementos. En símbolos $A = B$ si y sólo si se verifica el bicondicional $(\forall x)(x \in A \leftrightarrow x \in B)$

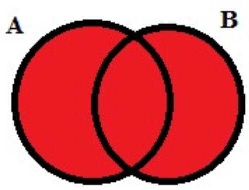
**Intersección**

Dados los conjuntos $A$ y $B$, se define el conjunto $A \cap B$ que llamaremos **intersección de $A$ y $B$**, como el conjunto formado por los elementos que son comunes a ambos conjuntos. $A \cap B = \{x : x \in A \wedge x \in B\}$



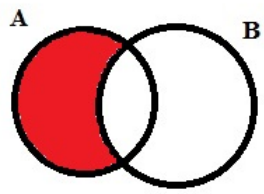

**Unión**

Dados los conjuntos $A$ y $B$, se define el conjunto $A \cup B$ que llamaremos **unión de $A$ y $B$**, como el conjunto formado por los elementos que pertenecen a uno u otro conjunto. $A \cup B = \{x : x \in A \vee x \in B\}$



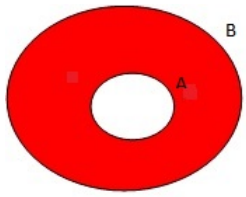
**Diferencia**

Dados los conjuntos $A$ y $B$, se define el conjunto $A - B$ que llamaremos **diferencia entre** $A$ **y** $B$ (en ese orden), como el conjunto formado por los elementos de $A$ que no pertenecen a $B$. $A - B = \{x : x \in A \wedge x \notin B\}$




**Complemento**

Si $A \subseteq B$, se define el **complemento de** $A$ **con respecto a** $B$ como el conjunto formado por los elementos de $B$ que no pertenecen a $A$. $C_B A = \{x : x \in B \wedge x \notin A\}$




En particular, si $B = \mathcal{U}$, decimos directamente el complemento de $A$, sin necesidad de aclarar respecto a quién. En general, usaremos $B = \mathcal{U}$ y simplificamos la notación usando: $A^C$.

**Propiedades:**

1. Idempotencia: $A \cup A = A$ ; $A \cap A = A$
2. Absorción: $A \cup (A \cap B) = A$ ; $A \cap (A \cup B) = A$
3. Distributiva: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ ; $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
4. Complementariedad $A \cup C_{\mathcal{U}} A = U$ ; $A \cap C_{\mathcal{U}} A = \emptyset$
5. $C_{\mathcal{U}}(C_{\mathcal{U}} A) = A$

## Producto Cartesiano

Sean $A_1, A_2, .....A_n$ conjuntos, el *producto cartesiano*[1] de $A_1, .....A_n$ denotado $A_1 \times ..... \times A_n$ es el conjunto de las n-uplas $(a_1, ...., a_n)$ tales que $a_i \in A_i$ para todo $i$ con $i = 1, ...n$.

El elemento $a_i \in A_i$ se denomina *i-ésima coordenada* de $(a_1, ...., a_n)$. Cuando $A_i = A$ para todo $i$, entonces $A_1 \times ..... \times A_n$ se denota $A^n$.

**Ejemplo**

Sean $A = \{a, b, c\}$ y $B = \{0, 1\}$ el producto cartesiano entre los conjuntos $A$ y $B$ es

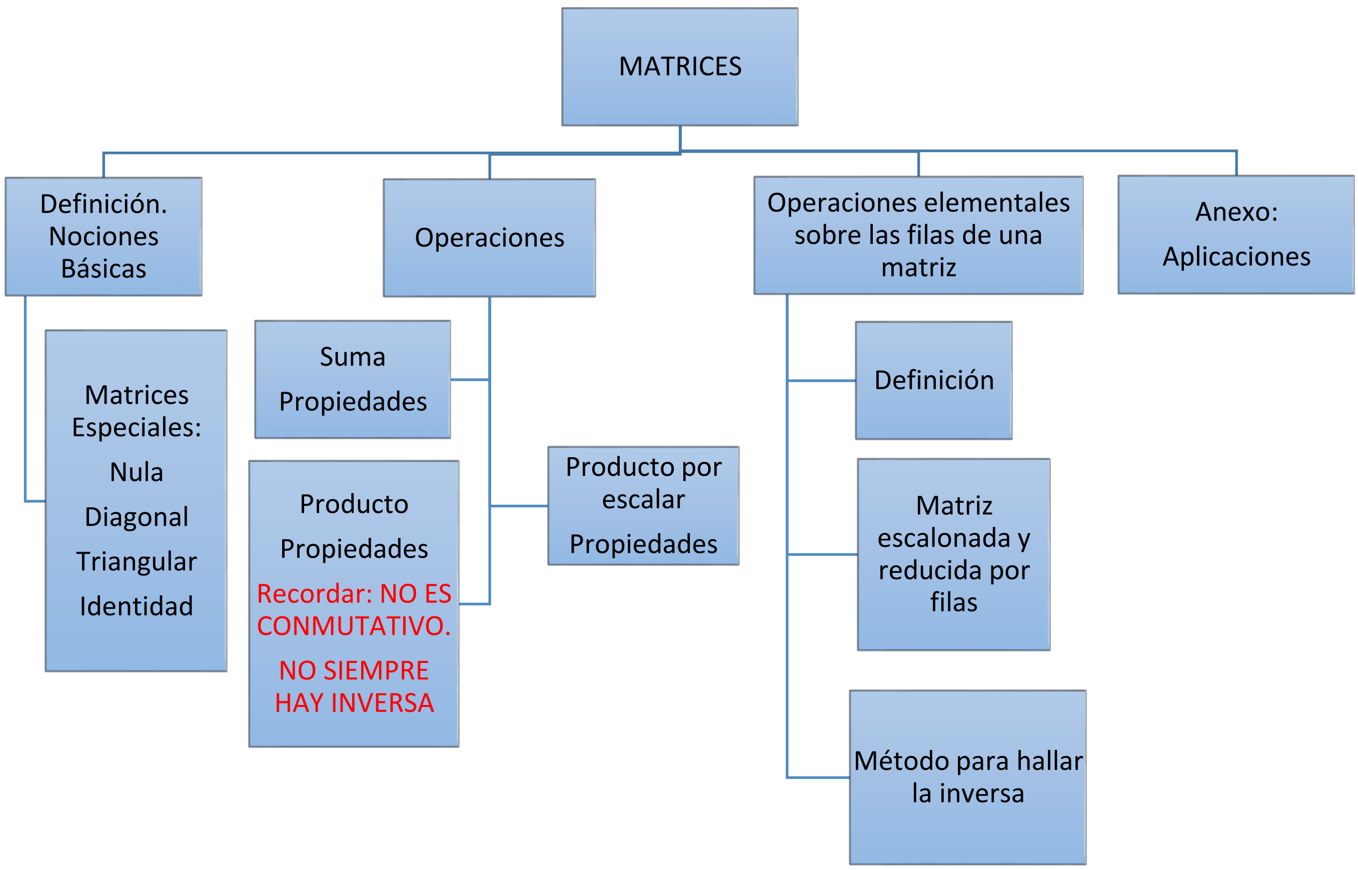
$$A \times B = \{(a, 0); (a, 1); (b, 0); (b, 1); (c, 0); (c, 1)\}$$

[1] El nombre producto cartesiano fue puesto en honor al matemático, físico y filósofo francés René Descartes, 1596-1650.

Capítulo 6

# MATRICES

CONTENIDOS:




## Matemáticos invitados: Arthur Cayley y James Joseph Sylvester

Las matrices como cuadros numéricos han sido usadas desde la antigüedad, fundamentalmente para resolver sistemas de ecuaciones lineales.

El libro chino *Nueve capítulos sobre el Arte de las matemáticas* (*Jiu Zhang Suan Shu*) que proviene del año 300 a 200 a.c., es el primer ejemplo conocido de uso del método de matrices para resolver sistemas de ecuaciones.

Después del desarrollo de la teoría de determinantes por Seki Kowa y Leibniz para facilitar la resolución de ecuaciones lineales, Carl Friederich Gauss y Wilhelm Jordan desarrollaron la eliminación de Gauss-Jordan, basada en las operaciones elementales, en el siglo XIX.

Fue James Joseph Sylvester quien utilizó por primera vez el término MATRIZ alrededor de 1848.

Cayley introdujo en 1858 la notación matricial, como forma abreviada de escribir un sistema de *m* ecuaciones lineales con *n* incógnitas.

Ambos nacieron en Inglaterra, Cayley nació en 1821 y fue abogado de profesión, mientras trabajaba en matemáticas en su tiempo libre, más adelante, conoce a Sylvester, que había nacido unos años antes, en 1814, y abandona la abogacía para dedicarse por completo a la matemática. Ambos trabajaron muchos años juntos e hicieron grandes aportes a la teoría de invariantes, campo relacionado con el álgebra lineal.

## 1. Introducción. Nociones básicas.

Una ecuación lineal con coeficientes reales $a_1, a_2, ..., a_n$, término independiente $b$ e incógnitas $x_1, x_2, ..., x_n$ es una expresión de la forma $a_1.x_1 + a_2.x_2 + ... + a_n.x_n = b$, por ejemplo $\frac{2}{5}.x_1 + 5.x_2 - 8.x_3 = -1$.

Si se tienen dos o más ecuaciones lineales en las mismas incógnitas se tiene un sistema de *m* ecuaciones con *n* incógnitas. El siguiente es un sistema de 2 ecuaciones con 3 incógnitas (sistema 2x3):

$$S_1 : \begin{cases} \frac{2}{5}.x_1 + 5.x_2 - 8.x_3 = -1 \\ x_1 + 6.x_2 - 4.x_3 = 9 \end{cases}.$$

Una ***solución*** del mismo, cuando existe, será una terna ordenada de números $(x_1^0, x_2^0, x_3^0)$ que satisfaga simultáneamente todas las ecuaciones del sistema.

Si en $S_1$ se cambian $x_1, x_2, x_3$ por, respectivamente, *x, y, z* o bien por $u_1, u_2, u_3$ el sistema es el mismo. De modo que toda la información del sistema se encuentra en los coeficientes y términos independientes *en el orden* que aparecen dispuestos, es decir en los siguientes "cuadros" de números $A = \begin{pmatrix} \frac{2}{5} & 5 & -8 \\ 1 & 6 & -4 \end{pmatrix}$, $b = \begin{pmatrix} -1 \\ 9 \end{pmatrix}$ que llamaremos ***matrices.***

**A** es una matriz **2**x**3** (2 filas por 3 columnas), **b** es una matriz **2**x**1** (2 filas por 1 columna).

Este es uno de los problemas más importantes en los que se aplican las matrices: la resolución de sistemas lineales de ecuaciones.

En general una ***matriz m****x****n*** tendrá la forma:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \ldots.. & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \ldots.. & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \ldots.. & a_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & a_{m3} & a_{m4} & \ldots.. & a_{mn} \end{pmatrix}$$

Usamos en general letras mayúsculas de imprenta para nombrar a las matrices y la misma letra en minúscula con dos subíndices para nombrar a sus elementos, así

Para la matriz A, el elemento que está en la fila $i$ columna $j$ se nota $a_{ij}$ .

El primer subíndice $i$ de cada coeficiente indica la fila donde se encuentra dicho coeficiente, el segundo subíndice ***j*** indica en qué columna está, para ***i*** = *1,2,…,m*, ***j*** **=** *1, 2, …,n*.

El conjunto de todas las matrices ***m*x*n*** con coeficientes $a_{ij} \in \mathbb{R}$, se indica $\mathbb{R}^{mxn}$, decimos entonces que si:

A es una matriz de m filas y n columnas con números reales $A \in \mathbb{R}^{mxn}$

En el caso particular en que ***m=n*** decimos que las matrices son **cuadradas**, de lo contrario son **rectangulares**.

En una **matriz cuadrada** la ***diagonal principal*** está dada por los elementos que tienen igual número de fila que de columna, es decir: $(a_{11}, a_{22}, a_{33}, ..., a_{nn})$.

Por ejemplo $(5, \sqrt{3}, 9)$ es la diagonal principal de la siguiente matriz 3x3

$$B = \begin{pmatrix} 5 & 0 & -3 \\ \frac{1}{8} & \sqrt{3} & 0 \\ -1 & 35 & 9 \end{pmatrix}$$

**Algunas clases especiales de matrices**

**Matriz nula**:

Es una matriz cuadrada o rectangular en la que todos sus coeficientes son ceros.

Si llamamos $b_{ij}$ a los elementos de la matriz nula nxm, definimos:

$\mathbf{O}_{nxm}$ es la matriz tal que $b_{ij} = 0 \;\; \forall i, 1 \le i \le n \;\; y \, \forall j, 1 \le j \le m$

Observar que hay una matriz nula para cada dimensión.

Por ejemplo la matriz nula 2x2 es $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, se escribe con un subíndice para indicar la dimensión: $O_{2x2}$

La matriz nula 3x2 es $\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$, se escribe: $O_{3x2}$

**Matriz *triangular superior*** (respectivamente ***inferior***):

Es una matriz cuadrada en la que son ceros todos los coeficientes debajo (respectivamente arriba) de la diagonal principal.

Si $A \in \mathbb{R}^{nxn}$ es triangular superior entonces $\forall i, \forall j,\ si\ \ i > j\ \ entonces\ a_{ij} = 0$

Si $A \in \mathbb{R}^{nxn}$ es triangular inferior entonces $\forall i, \forall j,\ si\ \ i < j\ \ entonces\ a_{ij} = 0$

La matriz A es una matriz 4x4, triangular superior y la matriz B es una matriz 4x4 triangular inferior:

$$A=\begin{pmatrix} 5 & 6 & 1 & 0 \\ 0 & 1 & -8 & 5 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 5 \end{pmatrix} \qquad B=\begin{pmatrix} 3 & 0 & 0 & 0 \\ 1 & 6 & 0 & 0 \\ 0 & 3 & 4 & 0 \\ 36 & -5 & 0 & -3 \end{pmatrix}$$

Los elementos que están debajo de la diagonal principal cumplen que la fila es mayor o igual que la columna y los que están arriba cumplen que la fila es menor o igual que la columna.

Notar que en la definición no se dice nada del resto de los elementos, es decir que también puede haber 0 en otros lugares como en la matriz B del ejemplo.

**Matriz diagonal**:

Es una matriz ***n*x*n*** en la que son **0** todos los coeficientes que no están en la diagonal principal.

Si $A \in \mathbb{R}^{nxn}$ es diagonal entonces $\forall i, \forall j\ \ si\ \ i \neq j\ \ entonces\ \ a_{ij} = 0$

Las siguientes son matrices diagonales:

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$$

**Matriz identidad**:

Es una matriz diagonal, en la que los coeficientes de la diagonal principal son todos 1.

Si llamamos $e_{ij}$ a los elementos de la matriz identidad nxn, definimos:

$I_n$ es la matriz tal que $e_{ij} = \begin{cases} 1 & si\ i = j \\ 0 & si\ \ i \neq j \end{cases}$

Observar que al igual que la matriz nula, hay una matriz identidad para cada dimensión.

Por ejemplo la matriz identidad 2x2 es: $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ , se escribe con un subíndice para indicar la dimensión

La matriz identidad 3x3 es: $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

---

**2. Operaciones**

**Definimos en principio la igualdad de matrices.**

Dadas dos matrices de la misma dimensión, $A \in \mathbb{R}^{mxn}$ y $B \in \mathbb{R}^{mxn}$ decimos que:

$A = B$ si $a_{ij} = b_{ij} \quad \forall i, 1 \le i \le m, \forall j, 1 \le j \le n$

**2.1 Suma**

Dadas dos matrices de la misma dimensión, $A \in \mathbb{R}^{mxn}$ y $B \in \mathbb{R}^{mxn}$ la suma de las matrices se define como una nueva matriz $C \in \mathbb{R}^{mxn}$ donde sus elementos son la suma de los elementos de A y de B en la misma posición:

**Ejemplo 2.1:**

Dadas $A\ y\ B \in \mathbb{R}^{2x3}$ $A = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 5 & 4 \end{pmatrix}$ $B = \begin{pmatrix} 3 & -1 & 1 \\ -3 & 3 & 4 \end{pmatrix}$, $A + B = \begin{pmatrix} 3 & 0 & 3 \\ 0 & 8 & 8 \end{pmatrix}$

La matriz $C = A + B$ es el resultado de: $\begin{pmatrix} 0+3 & 1-1 & 2+1 \\ 3-3 & 5+3 & 4+4 \end{pmatrix}$

**Suma de matrices**

Sean las matrices, $A \in \mathbb{R}^{mxn}$ y $B \in \mathbb{R}^{mxn}$ se define la matriz suma $C = A + B$

Donde $C \in \mathbb{R}^{mxn}$ y sus elementos $c_{ij} = a_{ij} + b_{ij} \quad \forall i, 1 \le i \le m, \forall j, 1 \le j \le n$

**Propiedades**

**1) Asociatividad:**

Para toda terna de matrices $A, B\ y\ C \in \mathbb{R}^{mxn}$ se cumple que

$$A + (B + C) = (A + B) + C$$

**2) Existencia de elemento neutro:**

Existe la matriz $0_{mxn} \in \mathbb{R}^{mxn}$, tal que para toda matriz $A \in \mathbb{R}^{mxn}$ se cumple que:

$$A + 0_{mxn} = 0_{mxn} + A = A \quad .$$

**3) Existencia de opuesto:**

Para toda matriz $A \in \mathbb{R}^{mxn}$ , existe una matriz $B \in \mathbb{R}^{mxn}$ tal que

$$A + B = B + A = 0_{mxn} \quad .$$

$B$ es la opuesta de $A$ y se indica $-A$

**Ejemplo 2.2**: dada la matriz $A = \begin{pmatrix} 2 & -30 \\ \sqrt{3} & 4 \\ -1 & \frac{2}{5} \end{pmatrix}$, su opuesta es $-A = \begin{pmatrix} -2 & 30 \\ -\sqrt{3} & -4 \\ 1 & -\frac{2}{5} \end{pmatrix}$

Notar que gracias a la existencia del opuesto se define la **resta de matrices**, como la suma de una matriz y la opuesta de otra:

| |
|---|
| Dadas $C\ y\ D \in \mathbb{R}^{mxn}$ definimos $C - D = C + (-D)$ |

**4) Conmutatividad:**

Para todo par de matrices $A\ y\ B \in \mathbb{R}^{mxn}$ se cumple que

$$A + B = B + A$$

**Demostración de la propiedad conmutativa**:

Queremos ver que: $A + B = B + A$ .

Llamemos $C = A + B \quad y \quad D = B + A$ , entonces por la igualdad de matrices basta con probar que $c_{ij} = d_{ij} \quad \forall i, 1 \le i \le m, \forall j, 1 \le j \le n$

En efecto: $c_{ij} = a_{ij} + b_{ij} = b_{ij} + a_{ij} = d_{ij} \qquad \forall i, 1 \leq i \leq m, \forall j, 1 \leq j \leq n$

Por definición de Suma de matrices

Por propiedad conmutativa de Los números reales

Por definición de suma de matrices

Hemos probado que C=D y por lo tanto $A + B = B + A$

**Observación:**

**Como hemos visto que la operación $+$ es cerrada, asociativa, tiene neutro, tiene opuesto y es conmutativa, tenemos que $(\mathbb{R}^{mxn}, +)$ es un Grupo Conmutativo.**

### 2.2 Producto de un escalar (número real) por una matriz

Dadas una matriz $A \in \mathbb{R}^{mxn}$ y un número real $\alpha$ la matriz $C = \alpha.A,$ $C \in \mathbb{R}^{mxn}$ es la matriz que resulta de multiplicar todos los elementos de $A$ por el número real $\alpha$ .

**Ejemplo 2.3:**

Dados $A \in \mathbb{R}^{2x3}$ $y$ $\alpha \in \mathbb{R}$ , $\alpha = 3$ , $A = \begin{pmatrix} 1 & 2 & 5 \\ -2 & 1 & 0 \end{pmatrix}$, $\alpha.A = \begin{pmatrix} 3 & 6 & 15 \\ -6 & 3 & 0 \end{pmatrix}$

La matriz $C = \alpha A$ es el resultado de: $\begin{pmatrix} 3.1 & 3.2 & 3.5 \\ 3.(-2) & 3.1 & 3.0 \end{pmatrix}$

**Producto por un escalar**

Sea la matriz, $A \in \mathbb{R}^{mxn}$ y $\alpha \in \mathbb{R}$ se define la matriz producto por escalar $C = \alpha A$

Donde $C \in \mathbb{R}^{mxn}$ y sus elementos $c_{ij} = \alpha . a_{ij}$ $\forall i, 1 \leq i \leq m, \forall j, 1 \leq j \leq n$

Observación: escribimos indistintamente $\alpha A$ o $A\alpha$

**Propiedades**

**1)** Para todo par de matrices $A$ $y$ $B$ $\in \mathbb{R}^{mxn}$ $y$ $\alpha \in \mathbb{R}$ se cumple que:

$$\alpha.(A + B) = \alpha.A + \alpha.B \quad \text{y} \quad (A + B).\alpha = A.\alpha + B.\alpha$$

**2)** Para toda matriz $A$ $\in \mathbb{R}^{mxn}$ $y$ $\alpha$ $y$ $\beta \in \mathbb{R}$ se cumple que:

$$(\alpha + \beta).A = \alpha.A + \beta.A \quad \text{y} \quad A.(\alpha + \beta) = A.\alpha + A.\beta$$

**3)** Para toda matriz $A$ $\in \mathbb{R}^{mxn}$ $y$ $\alpha$ $y$ $\beta \in \mathbb{R}$ se cumple que:

$$(\alpha.\beta).A = \alpha.(\beta.A)$$

**4)** Para toda matriz $A \in \mathbb{R}^{mxn}$ se cumple que:

$$1.A = A$$

La definición de suma y producto por escalar con sus respectivas propiedades le dan a las matrices con esas operaciones la estructura de **Espacio vectorial.** No trabajaremos en este curso con espacios vectoriales pero se verá más adelante como una importante estructura con muchas aplicaciones, entre ellas la detección de errores en la transmisión de datos.

Las propiedades vistas para la suma no difieren de las propiedades de la suma de números reales. Son importantes porque nos permiten operar con matrices en ecuaciones.

**Ejemplo 2.4:**

Hallar la matriz $X \in \mathbb{R}^{3x2}$ que cumpla:

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

Tenemos que despejar $X$ utilizando las propiedades vistas:

Por la propiedad 1 de producto por escalar y por definición de suma de matrices:

$$3.\begin{pmatrix} 1 & 0 \\ -2 & 4 \\ 1 & 1 \end{pmatrix} + 3.\begin{pmatrix} 3 & 1 \\ 5 & 1 \\ 0 & 1 \end{pmatrix} = 3.\left(\begin{pmatrix} 1 & 0 \\ -2 & 4 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 3 & 1 \\ 5 & 1 \\ 0 & 1 \end{pmatrix}\right) = 3.\begin{pmatrix} 4 & 1 \\ 3 & 5 \\ 1 & 2 \end{pmatrix}$$

Por definición de producto por escalar tenemos entonces que:

$$\begin{pmatrix} 2 & -3 \\ 5 & 4 \\ 0 & 1 \end{pmatrix} + X = \begin{pmatrix} 12 & 3 \\ 9 & 15 \\ 3 & 6 \end{pmatrix}$$

Entonces por la existencia del opuesto podemos sumar a ambos miembros el opuesto de $A$:

$$\begin{pmatrix} 2 & -3 \\ 5 & 4 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} -2 & 3 \\ -5 & -4 \\ 0 & -1 \end{pmatrix} + X = \begin{pmatrix} 12 & 3 \\ 9 & 15 \\ 3 & 6 \end{pmatrix} + \begin{pmatrix} -2 & 3 \\ -5 & -4 \\ 0 & -1 \end{pmatrix}$$

Por definición de suma de matrices y por la existencia de elemento neutro tenemos:

$$X = \begin{pmatrix} 10 & 6 \\ 4 & 11 \\ 3 & 5 \end{pmatrix}$$

### 2.3 Producto de matrices

Dadas dos matrices, $A \in \mathbb{R}^{mxn}$ y $B \in \mathbb{R}^{nxt}$ el producto de las matrices se define como una nueva matriz $C \in \mathbb{R}^{mxt}$ donde sus elementos son el resultado sumar las multiplicaciones de los elementos de cada fila de $A$ con cada columna de $B$:

**Ejemplo 2.5:**

Sean $A = \begin{pmatrix} 4 & 2 \\ 0 & 0 \\ 1 & 6 \end{pmatrix} \in \mathbb{R}^{3x2}$ $y$ $B = \begin{pmatrix} 3 & 5 \\ 1 & 3 \end{pmatrix} \in \mathbb{R}^{2x2}$, $A.B = \begin{pmatrix} 4.3 + 2.1 & 4.5 + 2.3 \\ 0.3 + 0.1 & 0.5 + 0.3 \\ 1.3 + 6.1 & 1.5 + 6.3 \end{pmatrix} = \begin{pmatrix} 14 & 26 \\ 0 & 0 \\ 9 & 23 \end{pmatrix}$

Una forma práctica de hacer el producto es ubicando las matrices de la siguiente manera:

$$\begin{array}{c|c} & \begin{pmatrix} \boxed{\begin{matrix} 3 \\ 1 \end{matrix}} & \begin{matrix} 5 \\ 3 \end{matrix} \end{pmatrix} = B \\ \hline A = \begin{pmatrix} \boxed{4 \quad 2} \\ 0 \quad 0 \\ 1 \quad 6 \end{pmatrix} & \begin{pmatrix} \boxed{4.3 + 2.1} & 4.5 + 2.3 \\ 0.3 + 0.1 & 0.5 + 0.3 \\ 1.3 + 6.1 & 1.5 + 6.3 \end{pmatrix} = C \end{array}$$

El elemento $c_{11} = a_{11}.b_{11} + a_{12}.b_{21}$, se recorre la fila 1 de $A$ y la columna 1 de $B$

El elemento $c_{12} = a_{11}.b_{12} + a_{12}.b_{22}$, se recorre la fila 1 de $A$ y la columna 2 de $B$

…

El elemento $c_{23} = a_{21}.b_{13} + a_{22}.b_{23}$, se recorre la fila 2 de $A$ y la columna 3 de $B$

Notar que para construir un elemento de $C$ en la posición $ij$ se recorre la fila $i$ de $A$ y la columna $j$ de $B$, lo que da lugar a la siguiente definición:

**Producto de matrices**

Sean las matrices, $A \in \mathbb{R}^{mxn}$ y $B \in \mathbb{R}^{nxt}$ se define la matriz producto $C = A.B$

Donde $C \in \mathbb{R}^{mxt}$ y sus elementos son:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \,.\, b_{kj} \qquad \forall i, 1 \leq i \leq m, \quad \forall j, 1 \leq j \leq t$$

**Observación:**

Las matrices que multiplicamos en el ejemplo NO TIENEN LA MISMA DIMENSION, al multiplicar cada elemento de una fila de $A$ por una columna de $B$, **deben coincidir el número de columnas de la primer matriz con el número de filas de la segunda.**

Con las matrices del ejemplo, el producto $B.A$ no se puede hacer. Esto dice que el producto no está definido para cualquier par de matrices, solo para las matrices que cumplen esa relación entre columnas y filas.

Esta observación sugiere que EL PRODUCTO NO ES CONMUTATIVO, como veremos a continuación.

**Propiedades**

**1) Asociatividad:**

Para toda terna de matrices $A \in \mathbb{R}^{mxn}, B \in \mathbb{R}^{nxt}\ y\ C \in \mathbb{R}^{txq}$ se cumple que:

$$A.(B.C) = (A.B).C$$

Observar que todos los productos se pueden hacer.

**2) Existencia de elemento neutro:**

Existe la matriz $I_n$, tal que para toda matriz $A \in \mathbb{R}^{nxn}$ :

$$A.\mathrm{I}_n = \mathrm{I}_n.A = A$$

La matriz Identidad es el elemento neutro en el producto de matrices cuadradas. Observar que al ser cuadradas pueden realizarse los productos a derecha y a izquierda.

Si $A$ no es cuadrada y $A \in \mathbb{R}^{mxn}$ decimos que tiene un neutro a derecha y un neutro a izquierda, esto es: $A.\mathrm{I}_n = \mathrm{I}_m.A = A$

A derecha multiplicamos por la Identidad nxn y a izquierda por la Identidad mxm.

**3) Distributividad del producto en la suma:**

Para toda terna de matrices $A, B\ y\ C \in \mathbb{R}^{nxn}$ se cumple que:

$$A.(B + C) = A.B + A.C \quad \text{y} \quad (B + C).A = B.A + C.A$$

Observar que las matrices al distribuir se mantienen en EL MISMO ORDEN.

**4) NO SE CUMPLE la propiedad Conmutativa:**

**Ejemplo 2.6:**

Sean $A=\begin{pmatrix}1 & 3\\ 2 & 5\end{pmatrix}$, $B=\begin{pmatrix}-1 & 0\\ 2 & 3\end{pmatrix}$, resulta $A.B=\begin{pmatrix}5 & 9\\ 8 & 15\end{pmatrix}$ y $B.A=\begin{pmatrix}-1 & -3\\ 8 & 21\end{pmatrix}$

**Observación:** esto no quiere decir que no haya matrices para las que sí se cumple, pero para hablar de una propiedad debe cumplirse **para todas las matrices** y esto no es cierto.

Por ejemplo:

$$\begin{array}{c|c} & \begin{pmatrix}3 & 0\\ 0 & 5\end{pmatrix}=B \\ \hline A=\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix} & \underbrace{\begin{pmatrix}1.3+0.0 & 1.0+0.5\\ 0.3+2.0 & 0.0+2.5\end{pmatrix}}_{\begin{pmatrix}3 & 0\\ 0 & 10\end{pmatrix}}=C \end{array} \qquad \begin{array}{c|c} & \begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}=A \\ \hline B=\begin{pmatrix}3 & 0\\ 0 & 5\end{pmatrix} & \underbrace{\begin{pmatrix}3.1+0.0 & 3.0+0.2\\ 0.1+5.0 & 0.0+5.2\end{pmatrix}}_{\begin{pmatrix}3 & 0\\ 0 & 10\end{pmatrix}}=C \end{array}$$

**5)** En los números naturales, enteros, reales se tiene la propiedad: $a.b=0 \Rightarrow a=0 \vee b=0$.

En el producto de matrices tampoco es válida esa propiedad:

$A.B=O$ <u>no</u> implica $A=O$ o $B=O$

**Ejemplo 2.7:**

Sean $A=\begin{pmatrix}1 & 3\\ 2 & 6\end{pmatrix}$ , $B=\begin{pmatrix}-3 & 3\\ 1 & -1\end{pmatrix}$, el producto $A.B=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$ da la matriz nula, mientras que ni $A$ ni $B$ son nulas

**6) La existencia de inversa NO ES una propiedad general de las matrices.**

En los números reales, todo número tiene inverso multiplicativo, esto es, dado un número real $a$ se busca otro que multiplicado por él de como resultado el neutro del producto, es decir : $a.a^{-1}=1$, por ejemplo: $2.\frac{1}{2}=1$, $5.\frac{1}{5}=1$, $34.\frac{1}{34}=1$ , etc.

En las matrices deberíamos buscar, dada una matriz $A$ otra matriz que multiplicada por ella de como resultado el neutro del producto, es decir $I$, teniendo en cuenta además que como el producto no es conmutativo, debe poder multiplicarse a derecha y a izquierda y ambos resultados deben dar la identidad. Para esto la única opción es que la matriz sea cuadrada, pero veremos que esto no es suficiente:

**Ejemplo 2.8:**

Sea $A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$ , tenemos que buscar una matriz $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ que multiplicada por $A$ de como resultado la identidad:

$$\begin{array}{r|l} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} = B \\ \hline A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}$$

Entonces $1.a + 2.c = 1$ y $1.b + 2.d = 0$

$0.a + 0.c = 0$ y $0.b + 0.d = 1$

La última ecuación nos dice $0 = 1$ que es absurdo, es decir que no existen valores de $a, b, c$ y $d$ que cumplan las ecuaciones, por lo tanto la matriz $A$ no tiene inversa.

Si una matriz no es cuadrada no posee inversa.

Algunas matrices cuadradas tienen inversa y otras no.

Como existen infinitas matrices cuadradas que tienen inversa e interesa conocerlas, será un objetivo identificar aquéllas que tengan inversa y en tal caso calcularla.

**Observación:**

Vimos que $(\mathbb{R}^{nxn}, +)$ es un Grupo Conmutativo y como la multiplicación en $\mathbb{R}^{nxn}$ es cerrada, asociativa y distributiva con respecto a la suma, decimos que $(\mathbb{R}^{nxn}, +, .)$ **es un Anillo** .

En el parágrafo que sigue se dará la definición de matriz inversa y algunas propiedades, más adelante se tratará el problema de hallar la inversa cuando ésta exista.

**Matriz Inversa**

Sea $A \in \mathbb{R}^{nxn}$ , $A$ tiene inversa (se dice también que es invertible o no singular) si existe una matriz $B \in \mathbb{R}^{nxn}$ tal que: $A.B = I$ y $B.A = I$

Recordar que como la conmutatividad no es una propiedad general del producto de matrices, se piden las dos condiciones.

**En caso de que la matriz $B$ de la definición exista, esta matriz se nota $A^{-1}$ .**

**Observación:** en los números reales notamos a la inversa de un número $a$ como $a^{-1}$ y también puede escribirse $\frac{1}{a}$ . En el caso de las matrices esta última notación no es válida ya que supondría dividir por una matriz, operación que no está definida en este conjunto.

**Proposición 2.1:**

Si existe ***B*** en las condiciones de la definición de Matriz Inversa, ésta es única.

**Demostración**

Lo probaremos por el método del absurdo:

Existe $B$ tal que $A.B = I$ y $B.A = I$ y no es única, es decir que existe otra matriz $C$ tal que $A.C = I$ y $C.A = I$

Si esto pasara entonces multiplicando la igualdad $B.A = I$ por $C$ a ambos miembros a derecha tenemos:

$(B.A).C = I.C = C$ ***(1)***

Y multiplicando la igualdad $A.C = I$ por $B$ a ambos miembros a izquierda tenemos $B.(A.C) = B.I = B$ ***(2)***

Como el producto de matrices es asociativo de ***(1)*** $(B.A).C = I.C = C$ tenemos que $B.(A.C) = C$ y en ***(2)*** teníamos que $B.(A.C) = B$ Absurdo

Es decir que ***C=B***, por lo tanto la matriz ***B*** que cumple la Definición es única.

**Proposición 2.2:**

Sean $A, B$ matrices *nxn* inversibles entonces $(A.B)^{-1} = B^{-1}.A^{-1}$.

**Demostración**

Lo demostraremos por el método directo:

Por hipótesis existen $A^{-1}$ y $B^{-1}$.

Notar que como las matrices son nxn sus inversas y el producto entre ellas también lo son.

De acuerdo a la definición de matriz inversa debe probarse que multiplicando $B^{-1}.A^{-1}$ a izquierda y a derecha de $A.B$, en ambos casos se obtiene la identidad.

$$(A.B).(B^{-1}.A^{-1}) \underset{\substack{\text{Por propiedad}\\ \text{Asociativa}}}{=} A.(B.B^{-1}).A^{-1} \underset{\substack{\text{Por definición}\\ \text{de inversa}}}{=} A.I.A^{-1} \underset{\substack{\text{Por ser } I \text{ neutro}\\ \text{en las matrices nxn}}}{=} A.A^{-1} = I$$

Además:

$$(B^{-1}.A^{-1}).(A.B) \underset{\substack{\text{Por propiedad}\\ \text{Asociativa}}}{=} B^{-1}.(A^{-1}.A).B \underset{\substack{\text{Por definición}\\ \text{de inversa}}}{=} B^{-1}.I.B \underset{\substack{\text{Por ser } I \text{ neutro}\\ \text{en las matrices nxn}}}{=} B^{-1}.B = I$$

Hemos encontrado entonces una matriz $B^{-1}.A^{-1}$ que multiplicada a izquierda y a derecha por la matriz $A.B$ da la matriz Identidad, entonces $B^{-1}.A^{-1}$ y es la inversa de $A.B$

Esto se nota como $(A.B)^{-1} = B^{-1}.A^{-1}$

**Corolario**

El resultado anterior se generaliza a cualquier número finito de matrices inversibles de un mismo orden *nxn*, esto es: $(A_1.A_2 \dots .A_n)^{-1} = (A_n)^{-1} \dots .(A_2)^{-1}.(A_1)^{-1}$

**Las operaciones vistas con sus propiedades nos permiten resolver ecuaciones con matrices.**

**Ejemplo 2.9:**

Hallar la matriz $X \in \mathbb{R}^{2x3}$ que cumpla $\begin{pmatrix} 3 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} - X = \begin{pmatrix} 0 & -2 & 1 \\ 1 & 4 & 0 \end{pmatrix} . \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & -1 \\ -1 & 1 & 1 \end{pmatrix}$

Resolvemos primero el lado derecho de la ecuación realizando el producto:

$$\begin{array}{c|c} & \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & -1 \\ -1 & 1 & 1 \end{pmatrix} \\ \hline \begin{pmatrix} 0 & -2 & 1 \\ 1 & 4 & 0 \end{pmatrix} & \begin{pmatrix} -1 & -3 & 3 \\ 1 & 9 & -4 \end{pmatrix} \end{array}$$

Tenemos entonces que: $\begin{pmatrix} 3 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} - X = \begin{pmatrix} -1 & -3 & 3 \\ 1 & 9 & -4 \end{pmatrix}$

Sumamos el opuesto de $\begin{pmatrix} 3 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$ a ambos miembros:

$$\begin{pmatrix} 3 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix} + \begin{pmatrix} -3 & -1 & 0 \\ 1 & -2 & -1 \end{pmatrix} - X = \begin{pmatrix} -1 & -3 & 3 \\ 1 & 9 & -4 \end{pmatrix} + \begin{pmatrix} -3 & -1 & 0 \\ 1 & -2 & -1 \end{pmatrix}$$

Como sumar una matriz con su opuesta nos da la matriz nula y ésta es el elemento neutro en la suma de matrices:

$$-X = \begin{pmatrix} -4 & -4 & 3 \\ 2 & 7 & -5 \end{pmatrix}$$

Ahora aplicamos la operación de producto por escalar, en este caso -1:

$$X = \begin{pmatrix} 4 & 4 & -3 \\ -2 & -7 & 5 \end{pmatrix}$$

**Ejemplo 2.10:**

Dos comercios de venta de tecnología, T1 y T2, reciben mensualmente tablets y celulares, de dos empresas de fabricación distintas (f y g). Se registran esas cantidades y sus precios (expresados en precio en pesos/mil pesos) en las siguientes matrices:

$$\begin{array}{ll} & \begin{array}{cc} \text{Tablets} & \text{Celulares} \end{array} \\ \begin{array}{c} \text{f} \\ \text{g} \end{array} & A = \begin{pmatrix} 20 & 30 \\ 50 & 40 \end{pmatrix} \end{array} \qquad \begin{array}{ll} \begin{array}{cc} \text{Precio T1} & \text{Precio T2} \end{array} & \\ B = \begin{pmatrix} 20 & 21 \\ 14 & 13 \end{pmatrix} & \begin{array}{c} \text{Tablets} \\ \text{Celulares} \end{array} \end{array}$$

Si queremos obtener cuál será la ganancia total de cada comercio vendiendo la totalidad de los productos que recibieron, por cada fábrica, realizamos el producto A.B:

$$\begin{pmatrix} 20 & 30 \\ 50 & 40 \end{pmatrix} \cdot \begin{pmatrix} 20 & 21 \\ 14 & 13 \end{pmatrix} =$$

Ganancia de T1 vendiendo los productos de f — Ganancia de T2 vendiendo los productos de f

$$= \begin{pmatrix} 20.20 + 30.14 & 20.21 + 30.13 \\ 50.20 + 40.14 & 50.21 + 40.13 \end{pmatrix} = \begin{pmatrix} 820 & 810 \\ 1560 & 1570 \end{pmatrix}$$

Ganancia de T1 vendiendo los productos de g — Ganancia de T2 vendiendo los productos de g

**Ejemplo 2.11:**

Si $A$ es una matriz cuadrada que cumple que: $A^3 = -4A - I$, hallar $A^{-1}$.

Tenemos que encontrar una matriz $B$ que cumpla que: $AB = BA = I$

Observemos que no conocemos a la matriz $A$, solo conocemos que cumple:

$$A^3 = -4A - I$$

Entonces: $-A^3 - 4A = I$

Podemos sacar factor común $A$ : $A(-A^2 - 4I) = I$

O también: $(-A^2 - 4I)A = I$

Observemos que cuando sacamos como factor común una matriz que está multiplicada por un número, el término dentro del paréntesis es el número multiplicado por la matriz identidad, que es el neutro del producto de matrices

$A(-A^2) = -A^3$ y $(-A^2)A = -A^3$
$A(-4I) = -4A$ y $(-4I)A = -4A$

Esto nos dice que la matriz $(-A^2 - 4I)$ multiplicada a derecha y a izquierda por la matriz $A$ da como resultado la matriz Identidad, por lo tanto $(-A^2 - 4I)$ es la inversa de $A$.

$$\boxed{A^{-1} = (-A^2 - 4I)}$$

---

**Ejercicios:**

**1)** Dadas las matrices $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ , $B = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$, C= $\begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix}$

Calcular: a) 3A-2B+C

b) A- 3(B-C)

c) Hallar una matriz D de 2x2 que cumpla que: $A - D = B$

d) Hallar una matriz E de 2x2 que cumpla que: $A + B + H$ sea una matriz triangular superior.

e) Hallar una matriz F de 2x2 que sea el opuesto de $C - B + A$

f) Hallar una matriz G de 2x2 que cumpla que: $B - 3G + 4(A + B) = 0_{2x2}$

**2)** Sean A $\in \mathbb{R}^{4x5}$, B$\in \mathbb{R}^{5x7}$, C $\in \mathbb{R}^{4x5}$, D $\in \mathbb{R}^{7x5}$. Indicar cuáles de las siguientes operaciones son posibles y, en caso afirmativo, cuál es la cantidad de filas y de columnas de la matriz resultado.

a) A .B  b) B.A  c) A.C  d) C.B  e) A.B.D

**3)** En los casos que sea posible calcular A.B y B.A, ¿Es A.B = B.A ?

a) $A = \begin{pmatrix} 2 & 3 \\ 1 & -4 \end{pmatrix}$  B= $\begin{pmatrix} 3 & -2 & 2 \\ 1 & 0 & -1 \end{pmatrix}$

b) $A = (1 \quad 2 \quad 3)$  $B = \begin{pmatrix} -2 \\ 4 \\ 1 \end{pmatrix}$

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

**4)** Al igual que en los números reales definimos recursivamente la potencia natural de una matriz: $A^n = \begin{cases} I & si\ n = 0 \\ A.A^{n-1} & si\ n \geq 1 \end{cases}$ , es decir que $A^n$ no es otra cosa que multiplicar n veces $A$ por sí misma.

Dadas $A = \begin{pmatrix} 3 & 4 \\ -1 & 0 \end{pmatrix}$ , $B = \begin{pmatrix} 6 & -1 \\ 0 & 4 \end{pmatrix}$  calcular: a) $A^2$  b) $B^3$  c) A.B  d) $2A^2$+ B.A

**5)** Dadas  $A = \begin{pmatrix} 1 & 3 \\ -2 & 0 \end{pmatrix}$  $B = \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$  establecer si es cierto que:

a) (A+B)(A - B) = $A^2$ - $B^2$

b) $(A+B)^2$ = $A^2$ + 2 AB + $B^2$

c) ¿Por qué valen (o no) las igualdades?

**6)** Dadas las matrices  $A = \begin{pmatrix} 6 & 9 \\ 4 & 6 \end{pmatrix}$  $B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$  $C = \begin{pmatrix} -2 & 1 \\ 3 & 2 \end{pmatrix}$

a) Calcular *A.B* y *A.C*.

b) ¿Es válida la propiedad cancelativa en el producto de matrices? (Recordar que la propiedad cancelativa de los números reales dice que: si $a.b = a.c$ entonces $b = c$ )

**7)** Hallar los valores de $a$ y de $b$ para que se cumpla la siguiente igualdad:

$$\begin{pmatrix}2 & 3\\ 1 & 4\end{pmatrix} + \begin{pmatrix}1 & 0\\ 2 & 2\end{pmatrix} . \begin{pmatrix}a & 5\\ 0 & b\end{pmatrix} = \begin{pmatrix}3 & 8\\ 3 & 21\end{pmatrix}$$

**8)** Hallar la matriz $X \in \mathbb{R}^{2x3}$ que cumpla $\begin{pmatrix}3 & 1 & 0\\ -1 & 0 & 1\end{pmatrix} + 3X = \begin{pmatrix}2 & 0 & 0\\ 0 & 3 & 1\end{pmatrix} . \begin{pmatrix}1 & 1 & 0\\ 0 & 2 & -1\\ -1 & 1 & 1\end{pmatrix}$

**9)** Sea *A* una matriz cuadrada 2x2. Probar que si *A* tiene una fila (o una columna) nula, entonces *A* no tiene inversa.

**10)** Si $A, B, C, D \in R^{nxn}$ tienen inversa, deducir cuál es $(A.B.C.D)^{-1}$

**11)** Una empresa que fabrica autos tiene 4 sucursales y fabrica 3 modelos distintos. Tiene guardada en dos matrices, la siguiente información de las ventas del último mes:

$$A = \begin{array}{rc} & \begin{array}{ccc}\text{Mod 1} & \text{Mod 2} & \text{Mod 3}\end{array} \\ \begin{array}{r}\text{Suc. 1}\\ \text{Suc. 2}\\ \text{Suc. 3}\\ \text{Suc. 4}\end{array} & \begin{pmatrix}30 & 25 & 5\\ 23 & 20 & 6\\ 20 & 22 & 4\\ 15 & 10 & 1\end{pmatrix}\end{array} \qquad B = \begin{array}{cl} \begin{array}{cc}\text{Precio/1000} & \text{Ganancia/1000}\end{array} & \\ \begin{pmatrix}700 & 100\\ 900 & 150\\ 980 & 200\end{pmatrix} & \begin{array}{l}\text{Mod. 1}\\ \text{Mod. 2}\\ \text{Mod. 3}\end{array}\end{array}$$

a) ¿Qué dimensión tiene el producto A.B? ¿Qué información obtiene al realizarlo?

b) Indique qué representa cada fila y cada columna de la matriz producto.

**12)** Sea $A$, una matriz cuadrada, que cumple que: $A^2 = 2A - I$ . Hallar $A^{-1}$

**13)** Sea $A$, una matriz cuadrada, que cumple que: $A^3 = 0$ (0 es la matriz nula de la misma dimensión que $A$) . Demostrar que $I + A + A^2$ es la matriz inversa de $I - A$.

---

## 3. Operaciones elementales sobre las filas de una matriz

Definimos en primer lugar qué es una matriz escalonada y reducida por filas, ya que las operaciones elementales las haremos sobre las filas de una matriz para transformarla en una escalonada y reducida por filas.

### 3.1 Matriz escalonada y reducida por filas

Sea $A \in \mathbb{R}^{mxn}$, $A$ está en forma *escalonada y reducida por filas si:*

1) el primer número distinto de cero de cada fila es 1, ese coeficiente se llama **coeficiente principal**.

2) si el número de ceros que precede al primer coeficiente no nulo va aumentando en las filas sucesivas.

3) en el resto de la columna donde hay coeficiente principal los elementos son 0.

4) si hay filas de 0 están al final.

**Ejemplos 3.1:**

$A$ y $B$ no son escalonadas y reducidas por filas, sólo $C$ es reducida y escalonada por filas.

$$A = \begin{pmatrix} 1 & 3 & 5 & 0 & 5 & 0 & 3 \\ 0 & 0 & 2 & 8 & 0 & 3 & 9 \\ 0 & 0 & 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 5 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 3 & 0 & 5 \\ 0 & 0 & 0 & 1 & 8 & 5 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 5 & 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

### 3.2 Operaciones elementales sobre las filas de una matriz.

Existen 3 tipos de ***operaciones elementales*** sobre las filas de una matriz que permiten llevar cualquier matriz $A \in \mathbb{R}^{mxn}$ a una matriz equivalente reducida y escalonada por filas:

1) Multiplicar una fila $F_k$ de $A$ por un escalar $\alpha \in \mathbb{R}$ no nulo, $F_k \leftarrow \alpha F_k$

2) Sumar a la fila $F_h$ la fila $F_k$ multiplicada por $\alpha \in \mathbb{R}$ no nulo, $F_h \leftarrow F_h + \alpha F_k$

3) Permutar dos filas $F_h \longleftrightarrow F_k$

Cada operación elemental **e** tiene su inversa **e**$^{-1}$ que es también elemental y del mismo tipo que **e.** Es decir que si realizamos:

1) la operación de permutar la fila i con la fila k, la operación inversa es permutar la fila i con la fila k para volver a la misma matriz.

2) la operación de sumarle a la fila i la fila k multiplicada por $\alpha$, la operación inversa es sumarle a la fila i la fila k multilicada por ($-\alpha$)

3) la operación de multiplicar la fila j por $\alpha$, la operación inversa es multiplicar la fila j por $1/\alpha$.

**Definición:**

Dos matrices $A$, $B$ se llaman **equivalentes por filas** si de una de ellas se pasa a la otra aplicando un número finito de operaciones elementales de fila.

$$A \approx_f B \quad \text{si} \quad e_k \circ e_{k-1} \circ \ldots \circ e_2 \circ e_1(A) = B$$

(Como cada $e_i$ posee su inversa que también es una operación elemental, de $B$ se puede llegar a $A$ aplicando las inversas respectivas).

**Ejemplo 3.2:**

Hallar la matriz escalonada y reducida por filas equivalente con $A = \begin{pmatrix} 2 & 4 & 6 \\ 2 & 0 & 1 \end{pmatrix}$

**Importante:** Realizaremos operaciones elementales sobre la matriz en un orden determinado, esto es por columnas, primero pondremos el 1 en la posición fila 1 columna 1, luego pondremos 0 en el resto de la columna, después pondremos el 1 en la posición fila 2 columna 2 y buscamos 0 en el resto de la columna. Este orden no es caprichoso, nos garantiza que los 1 y 0 que vayamos obteniendo no cambiarán con ninguna operación.

$$A = \begin{pmatrix} 2 & 4 & 6 \\ 2 & 0 & 1 \end{pmatrix} \quad \left(F_1 \leftarrow \frac{1}{2}.F_1\right) \sim \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 1 \end{pmatrix} \quad (F_2 \leftarrow F_2 - 2.F_1) \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & -4 & -5 \end{pmatrix}$$

$$\left(F_2 \leftarrow -\frac{1}{4}.F_2\right) \sim \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & \frac{5}{4} \end{pmatrix} \quad (F_1 \leftarrow F_1 - 2.F_2)) \sim \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{5}{4} \end{pmatrix} = A_R$$

**A esta última matriz la llamamos $A_R$ que es la matriz escalonada y reducida por filas equivalente con $A$, esto se escribe $A \approx_f A_R$ o simplemente $A \approx A_R$**

Notar que cada vez que realizamos una operación elemental no ponemos el signo =, sino $\sim$ , también puede ponerse $\approx o \rightarrow$, pero no =, ya que las matrices son distintas, veremos que mantienen propiedades importantes.

**Ejemplo 3.3:**

Hallar la matriz escalonada y reducida por filas equivalente con $B = \begin{pmatrix} 3 & 6 \\ 1 & 2 \\ 9 & 18 \end{pmatrix}$

Nuevamente realizaremos operaciones elementales sobre la matriz en el mismo orden.

$$B = \begin{pmatrix} 3 & 6 \\ 1 & 2 \\ 9 & 18 \end{pmatrix} \quad \left(F_1 \leftarrow \frac{1}{3}.F_1\right) \sim \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 9 & 18 \end{pmatrix} \quad (F_2 \leftarrow F_2 - 1.F_1) \sim \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 9 & 18 \end{pmatrix}$$

$$(F_3 \leftarrow F_3 - 9.F_1) \sim \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = B_R$$

### 3.3 Rango

**Definición**

Se llama **rango** de una matriz *A*, indicado con $r(A)$, al número de filas no nulas de la escalonada y reducida por filas $A_R$ equivalente con *A.*

En los ejemplos de arriba, la matriz $\mathrm{A}$ tiene rango 2 mientras que la matriz $\mathrm{B}$ tiene rango 1.

Existen dos aplicaciones importantes de las operaciones elementales: el cálculo de la inversa de una matriz y la resolución de sistemas de ecuaciones lineales.

En el caso del cálculo de la inversa recuerde que la matriz debe ser cuadrada, si es cuadrada puede tener o no inversa, pero si no lo es, seguro que no tiene inversa.

### 3.4 Cálculo de la inversa de una matriz

El procedimiento que seguiremos es: dada una matriz $A \in \mathbb{R}^{nxn}$, construimos un cuadro con la matriz dada y la matriz Identidad a la derecha, luego realizaremos operaciones elementales para hallar la escalonada y reducida por filas, $A_R$ , equivalente con A. Las operaciones elementales se hacen sobre $A$ y sobre la identidad, como si fuera una matriz de nx2n. Si al llegar a $A_R$ nos quedó la Identidad entonces la matriz identidad que transformamos es la inversa de $A$, sino, no tiene inversa.

**Ejemplo 3.4:**

Sea $A=\begin{pmatrix} 2 & 6 & -4 \\ 1 & 4 & 1 \\ 2 & 6 & 1 \end{pmatrix}$

$$\begin{array}{c} \begin{array}{cc} \qquad\quad A \qquad\quad & \qquad\quad I_3 \qquad\quad \end{array} \\ \left(\begin{array}{ccc|ccc} 2 & 6 & -4 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 & 1 & 0 \\ 2 & 6 & 1 & 0 & 0 & 1 \end{array}\right) \end{array}$$

$$F_1.\frac{1}{2} \rightarrow \left(\begin{array}{ccc|ccc} 1 & 3 & -2 & \frac{1}{2} & 0 & 0 \\ 1 & 4 & 1 & 0 & 1 & 0 \\ 2 & 6 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$F_2 - F_1 \rightarrow \left(\begin{array}{ccc|ccc} 1 & 3 & -2 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 3 & -\frac{1}{2} & 1 & 0 \\ 2 & 6 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$F_3 - 2.F_1 \rightarrow \left(\begin{array}{ccc|ccc} 1 & 3 & -2 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 3 & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 5 & -1 & 0 & 1 \end{array}\right)$$

$$F_1 - 3.F_2 \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -11 & 2 & -3 & 0 \\ 0 & 1 & 3 & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 5 & -1 & 0 & 1 \end{array}\right)$$

$$F_3.\frac{1}{5} \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & -11 & 2 & -3 & 0 \\ 0 & 1 & 3 & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 & -\frac{1}{5} & 0 & \frac{1}{5} \end{array}\right)$$

$$F_1 + 11F_3 \quad y \quad F_2 - 3F_3 \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{5} & -3 & \frac{11}{5} \\ 0 & 1 & 0 & \frac{1}{10} & 1 & -\frac{3}{5} \\ 0 & 0 & 1 & -\frac{1}{5} & 0 & \frac{1}{5} \end{array}\right)$$

La matriz B= $\begin{pmatrix} -\frac{1}{5} & -3 & \frac{11}{5} \\ \frac{1}{10} & 1 & -\frac{3}{5} \\ -\frac{1}{5} & 0 & \frac{1}{5} \end{pmatrix}$ es la inversa de A

Por la definición de equivalencia por filas resulta que**:**

$A \approx_f I$ ya que la Identidad es el resultado de haberle aplicado a A 7 operaciones elementales: $e_7$ **o** $e_6$ **o** ... **o** $e_2$ **o** $e_1(A)$ = I , y además

$I \approx_f B$ ya que B es el resultado de haberle aplicado a la Identidad 7 operaciones elementales: $e_7$ **o** $e_6$ **o** ... **o** $e_2$ **o** $e_1$(I) = B , y como

$e_7$ **o** $e_6$ **o** ... **o** $e_2$ **o** $e_1(A)$ = $e_7$ **o** $e_6$ **o** ... **o** $e_2$ **o** $e_1(I.A)$ = $\underbrace{e_7 \text{ o } e_6 \text{ o } \ldots \text{ o } e_2 \text{ o } e_1(I)}_{B}\, A$ = $\boxed{B.A = I}$

Además, como se mencionó, las operaciones elementales tienen inversa, en el sentido de que pueden aplicarse para volver a la matriz anterior, por eso A es el resultado de haberle aplicado a la identidad las inversas de las operaciones elementales:

$A$=$e_1^{-1}$ **o** $e_2^{-1}$ **o** ... **o** $e_6^{-1}$ **o** $e_7^{-1}$ (I) entonces

$e_1^{-1}$ **o** $e_2^{-1}$ **o** ... **o** $e_6^{-1}$ **o** $e_7^{-1}$ $(B)$ = $e_1^{-1}$ **o** $e_2^{-1}$ **o** ... **o** $e_6^{-1}$ **o** $e_7^{-1}$ $(I.B)$ =

= $e_1^{-1}$ **o** $e_2^{-1}$ **o** ... **o** $e_6^{-1}$ **o** $e_7^{-1}$ $(I)B = A.B = I$

$A$

De lo anterior resulta que $\boxed{B = A^{-1}}$

**Importante:**

Este procedimiento nos dice que:

$A \in \mathbb{R}^{nxn}$ y su rango es n si y sólo si $A$ tiene inversa

Que es equivalente a decir:

$A \in \mathbb{R}^{nxn}$ y $A \approx I$ si y sólo si $A$ tiene inversa

**Ejemplo 3.4:**

En el ejemplo 2.7) se mostró que si $A \in \mathbb{R}^{nxn}\ y\ B \in \mathbb{R}^{nxn}$ y $A.B = 0_{nxn}$ eso **no implica** que $A = 0_{nxn}$ o que $B = 0_{nxn}$.

Veremos que si $A$ tiene inversa entonces $B = 0_{nxn}$:

Sea $A.B = 0_{nxn}$

Como $A$ tiene inversa, podemos **multiplicar a ambos miembros, a izquierda**, por $A^{-1}$:

$$A^{-1}.(A.B) = A^{-1}.0_{nxn}$$

Y por la propiedad asociativa del producto de matrices, podemos escribir:

$$(A^{-1}.A).B = A^{-1}.0_{nxn}$$

Y como $A^{-1}.A = I$, tenemos: $I.B = A^{-1}.0_{nxn}$

Y como $I.B = B$ y $A^{-1}.0_{nxn} = 0_{nxn}$: $\boxed{B = 0_{nxn}}$

---

**Ejercicios:**

**14)** Llevar las siguientes matrices a la forma escalonada y reducida, indicando las operaciones elementales, indicar el rango de cada una.

$$A = \begin{pmatrix} 1 & -1 & 5 & 3 & 1 \\ 2 & -2 & 10 & 8 & 3 \\ 6 & -6 & 30 & 22 & 8 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 & 1 & 5 \\ 2 & 3 & -1 & 4 \\ -3 & -3 & 0 & -9 \end{pmatrix} \qquad C = \begin{pmatrix} -3 & 4 \\ 6 & -8 \end{pmatrix}$$

**15)** Hallar la inversa, si existe, de las siguientes matrices, mediante operaciones elementales e indicar el rango de cada una:

$$A=\begin{pmatrix} 1 & 4 \\ -8 & 2 \end{pmatrix} \qquad B=\begin{pmatrix} 1 & 0 & 3 \\ 2 & 3 & 4 \\ 0 & 2 & 1 \end{pmatrix} \qquad C=\begin{pmatrix} 2 & 1 & 0 \\ 5 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix} \qquad D=\begin{pmatrix} -2 & 5 \\ 6 & -15 \end{pmatrix}$$

**16)** Dada $A=\begin{pmatrix} 2 & -4 \\ 1 & k \end{pmatrix}$

**a)** Hallar $k$ para que sea $A^2 = \mathbf{0}_{2x2}$

**b)** Hallado $k$, encontrar el rango de A.

**17)** Sea A una matriz nxn:

a) Indicar, justificando su respuesta, si es verdadero o falso que: $(A - I_n)(A + I_n) = A^2 - I_n$

b) Si $A^2 = \mathbf{0}_{nxn}$ , ¿Cuál es la inversa de $(A + I_n)$ ?

**18)** Dadas las matrices: $A = \begin{pmatrix} -3 & -9 & 0 \\ 0 & 2 & 0 \\ 6 & 18 & 1 \end{pmatrix}$, $B = \begin{pmatrix} a & 0 \\ 1 & -1 \\ 5 & b \end{pmatrix}$, $C = \begin{pmatrix} -12 & 9 \\ 2 & -2 \\ 29 & -8 \end{pmatrix}$

a) Encontrar los números $a$ y $b$ tales que se cumpla $A.B = C$

b) Encontrar si es que existe, la inversa de $A$.

**19) a)** Dada $D=\begin{pmatrix} 1 & k \\ 5 & -1 \end{pmatrix}$ encontrar el valor de *k* para que sea $D^2 = \mathrm{O}$ (matriz nula)

**b)** Con el valor ***k*** encontrado, calcular el rango de $D$.

**20) a)** Encontrar los números $a,\ b$ tales que $(A+B).C = D$, siendo:

$$A=\begin{pmatrix} 3 & 2 & 5 \\ 3 & 1 & 2 \end{pmatrix}, \quad B=\begin{pmatrix} -2 & 1 & -4 \\ -2 & 0 & -1 \end{pmatrix}, \quad C=\begin{pmatrix} a & 0 \\ 2 & b \\ 0 & 3 \end{pmatrix}, \quad D=\begin{pmatrix} 9 & 6 \\ 5 & 4 \end{pmatrix}$$

c) Hallar (si existe) $D^{-1}$.

**21)** Demostrar que si $A \in \mathbb{R}^{nxn}\ y\ B \in \mathbb{R}^{nxn}$ y $A.B = 0_{nxn}$ y $B$ tiene inversa entonces $A = 0_{nxn}$

**22)** Sean las matrices A, B y C de nxn, tales que B tiene inversa. Indique si la siguiente afirmación es verdadera o falsa, justificando su respuesta:

“Si $C.B = A$ , entonces $C = B^{-1}.A$ “

**23)** Demuestre que si una matriz de 2x2 tiene dos filas iguales entonces no tiene inversa.

**24)** Sean A, B y C matrices cuadradas de la misma dimensión. Demuestre que si A tiene inversa y se cumple que A.B=A.C entonces B=C.

---

## 4. Anexo: Aplicaciones

Una de las aplicaciones importantes que tienen las matrices es la **criptografía**. La criptografía es el estudio de las formas de transmitir mensajes en forma segura. En la década del 40 cuando se establece el inicio de la criptografía, estaba restringida prácticamente al campo de la estrategia militar. Hoy, la criptografía es de gran utilidad en muchísimas cosas que hacemos a diario: cuando utilizamos la tarjeta de crédito en un cajero automático para realizar una operación bancaria, necesitamos identificarnos con una clave, cuando accedemos a nuestra cuenta de correo electrónico, se nos pide una contraseña, etc.

Si numeramos las letras del abecedario de 1 a 27 y quisiéramos enviar como mensaje la letra J, puede elegirse como código el número 7 y multiplicar 7.10 ya que 10 es el número que le corresponde a la J. El mensaje enviado seria 70.

Si el receptor no conoce que el código es 7 podría pensar que se envió 5.14=70 o 10.7=70 o 7.10=70, y no puede decidir si la letra enviada es N (que correspondería al número 14 y el código sería 5) o G(que correspondería al número 7 y el código sería 10) o J(que corresponde al número 10 y el código sería 7).

El receptor entonces debe conocer el código y multiplicar el mensaje recibido por el inverso de ese número, en el caso de haber elegido código 7 sería: $70.\frac{1}{7} = 10$. Claro que este método para encriptar es muy vulnerable ya que hay un número muy pequeño de posibilidades y del contexto del mensaje podría deducirse la letra. También es cierto que si en lugar de elegir el 7 eligiéramos el 345.678, aumenta significativamente la cantidad de posibilidades.

Este mismo proceso se utiliza para la encriptación con matrices, el código ya no es un número sino una matriz y el mensaje que se envía también se envía en una matriz, de modo que habrá que conocer la inversa de la matriz de código para recuperar el mensaje.

**Ejemplo:**

Para codificar un mensaje los elementos que se requieren son: un emisor, un receptor, un mensaje y un código.

Cuando hablamos de código, estamos hablando de un método de codificación, es decir algún algoritmo biunívoco (una *función biyectiva*), que asigne a cada carácter del mensaje otro carácter.

Este método hace que el mensaje enviado por el emisor se transforme en una cadena de símbolos ilegibles para el resto de los receptores que no sean legales. Dependiendo de la calidad del método de codificación, el mensaje será más o menos difícil de descifrar si es capturado por receptores ilegales.

**Método de encriptación:**

1. A las letras del alfabeto se le asignan los números del 1 al 27. Al espacio en blanco se le asigna el número 28, para poder separar palabras. Esta es una posibilidad, también podrían asignarse las letras en orden decreciente o comenzando por el número 3, etc.
2. Cualquier matriz cuyos elementos sean enteros positivos y sea invertible se puede usar como matriz de codificación.
3. Si la matriz de código es de nxn se construye con el mensaje una matriz de n filas y tantas columnas como sean necesarias, escribiendo los números por columna y rellenando al final con espacios en blanco si fuera necesario.
4. Luego se multiplica a izquierda por la matriz de código y el resultado es el mensaje codificado.
5. Para recuperar el mensaje se multiplica la matriz anterior a izquierda por la inversa de la matriz de código.

Mensaje a codificar: **“vuelvo mañana”**

Secuencia que le corresponde: **23 21 5 12 23 16 28 13 1 15 1 14 1**

Para la Matriz de código, se elige cualquiera que sea invertible, debe ser conocida por el emisor y el receptor: $C = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$

Construcción de la matriz A (tendrá dos filas): $\begin{pmatrix} 23 & 5 & 23 & 28 & 1 & 1 & 1 \\ 21 & 12 & 16 & 13 & 15 & 14 & 28 \end{pmatrix}$

Multiplicando CA, se obtiene: $B = \begin{pmatrix} 86 & 22 & 62 & 69 & 17 & 16 & 30 \\ 67 & 41 & 71 & 67 & 46 & 43 & 85 \end{pmatrix}$

El receptor es quien recibe esta matriz y debe conocer la inversa de la matriz de código para recuperar el mensaje. En este caso $C^{-1} = \begin{pmatrix} \frac{-1}{5} & \frac{3}{5} \\ \frac{2}{5} & -\frac{1}{5} \end{pmatrix}$ entonces al realizar el producto $C^{-1}$ .B= $C^{-1}$ . C. A= I. A=A y se recupera el mensaje en forma matricial. Es evidente que el receptor debe conocer tanto la primera fase de la codificación, es decir que número le corresponde a cada letra, como la segunda fase, es decir la matriz de código.

Basado en el artículo "Aplicación de las matrices invertibles en criptografía", Juan Carlos Cortés Lopez, Gema Carlo Sanjuán.

---

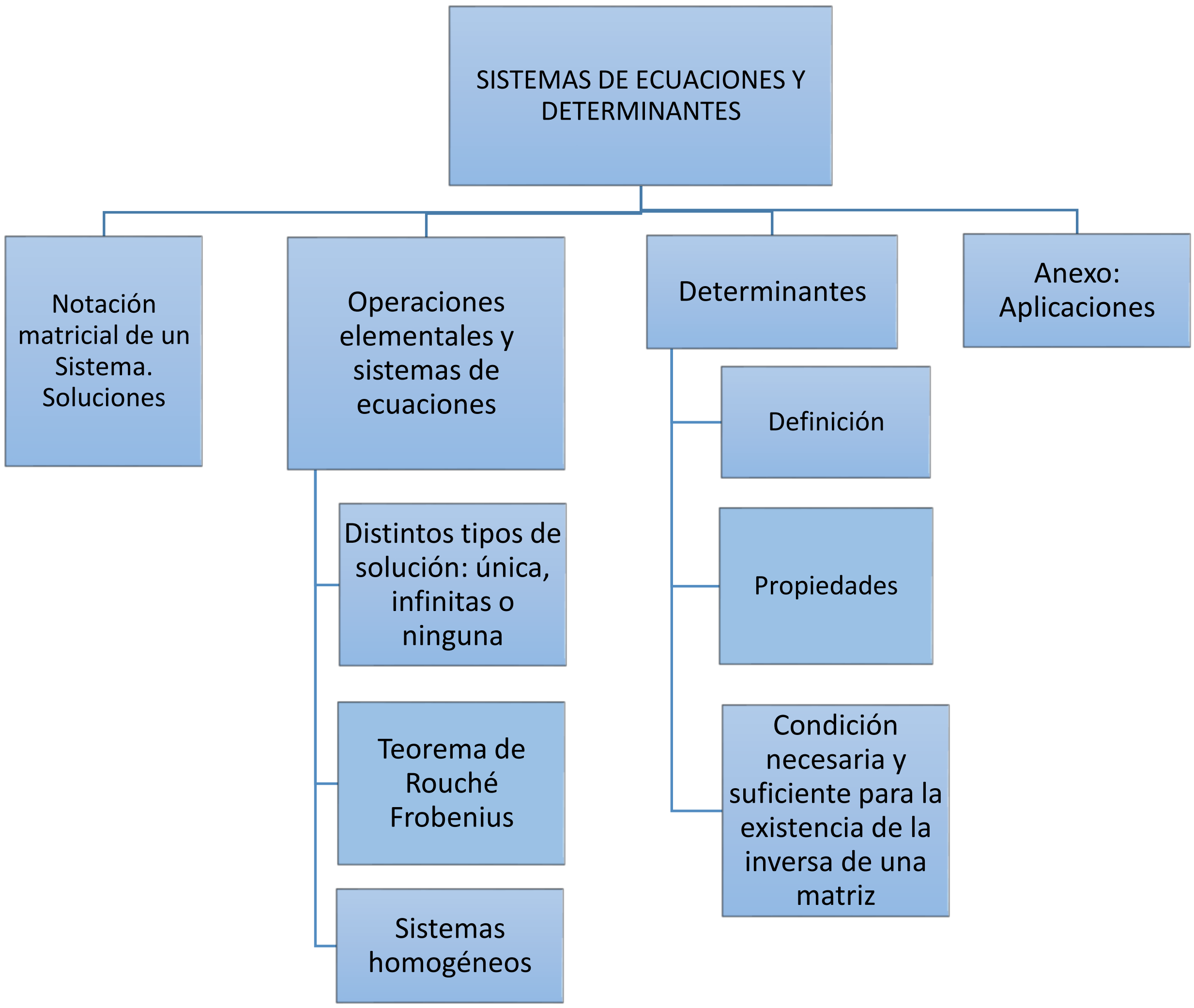
Bibliografía

- R. Espinosa Armenta, **Matemáticas discretas**, Editorial Alfaomega, Mexico, 2010
- Smith, et al , **Algebra, trigonometría y geometría analítica**, Pearson-Addison Wesley Longman, 1998
- Swokoski, Earl W. y Cole, Jeffery A., **Algebra y trigonometría con geometría analítica,** 11ma ed., Editorial  Thomson, 2006

Capítulo 7

# SISTEMAS DE ECUACIONES Y DETERMINANTES

CONTENIDOS:




**Matemático invitado: Leonardo Pisano**

Los sistemas de ecuaciones lineales modelizan muchos problemas de la realidad y la preocupación por resolverlos y por encontrar métodos para su resolución data de muchos años.

Leonardo Pisano, matemático italiano que vivió aproximadamente entre los años 1175 y 1250 d.c, más conocido como Fibonacci, y que se conoce sobre todo por su famosa sucesión, viajaba habitualmente por cuestiones de comercio. Durante sus viajes aprendió la nueva aritmética árabe, que después presentó al occidente en su conocido libro Liber Abaci. Cuenta la leyenda que el Emperador Federico II de Sicilia invitó a Fibonacci y a otros sabios a participar en una especie de torneo de matemáticas, en el que plantearon varios problemas. Uno de ellos era el siguiente:

"Tres hombres poseen una sola pila de monedas, y sus partes son 1/2, 1/3 y 1/6. Cada uno toma algo de dinero de la pila hasta que no queda nada. El primero regresa 1/2 de lo que tomó, el segundo 1/3 y el tercero 1/6. Cuando el total reintegrado se divide por igual entre los tres, se descubre que cada uno posee lo que le corresponde. ¿Cuánto dinero había en la pila original, y cuánto tomó cada uno de esa pila? "

Fibonacci llegó a la solución: la cantidad total era 47 y las cantidades que tomaron cada uno, fueron 33, 13 y 1. (Extraído del libro "Algebra Lineal con Aplicaciones",de Nakos y Joyner, 1999)

El desafío será resolver el problema al final del capítulo y analizar si es correcta la solución de Fibonacci.

---

**1. Notación matricial de un Sistema. Soluciones.**

Aplicaremos los contenidos vistos en el capítulo anterior para resolver sistemas de ecuaciones lineales.

Ejemplos de sistemas de ecuaciones lineales:

a) $\begin{cases} 2x - y = 4 \\ 3x + 2y = 1 \end{cases}$

Sistema de 2 ecuaciones con 2 incógnitas

b) $\begin{cases} x - 2y + z = 4 \\ 5x - y + 3z = 3 \end{cases}$

Sistema de 2 ecuaciones con 3 incógnitas

c) $\begin{cases} x - y + z = 2 \\ 4x - 2y + z = 1 \\ x + y - 2z = 3 \end{cases}$

Sistema de 3 ecuaciones con 3 incógnitas

**Encontrar soluciones a estos sistemas implica hallar valores de las variables que cumplan todas las ecuaciones a la vez.**

Podemos representar los sistemas anteriores, como un producto matricial, con una matriz **A** llamada de los coeficientes, una matriz **X**, llamada matriz de las incógnitas y una matriz **b**, llamada matriz del término independiente:

a) $\underbrace{\begin{pmatrix} 2 & -1 \\ 3 & 2 \end{pmatrix}}_{A}\underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{X} = \underbrace{\begin{pmatrix} 4 \\ 1 \end{pmatrix}}_{b}$ Expresado también como: AX=b

Si hacemos el producto nos queda: $\begin{pmatrix} 2x - y \\ 3x + 2y \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$

Que por igualdad de matrices resulta: $2x - y = 4$

$3x + 2y = 1$ que es el sistema original.

b) $\begin{pmatrix} 1 & -2 & 1 \\ 5 & -1 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$

Si hacemos el producto nos queda: $\begin{pmatrix} x - 2y + z \\ 5x - y + 3z \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$

Que por igualdad de matrices resulta: $x - 2y + z = 4$

$5x - y + 3z = 3$ que es el sistema original.

c) $\begin{pmatrix} 1 & -1 & 1 \\ 4 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$

Si hacemos el producto nos queda: $\begin{pmatrix} x - y + z \\ 4x - 2y + z \\ x + y - 2z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$

Que por igualdad de matrices resulta: $x - y + z = 2$

$4x - 2y + z = 1$

$x + y - 2z = 3$ que es el sistema original.

Esto justifica que podamos tratar los sistemas como matrices, donde es evidente que si cambiamos los nombres de las incógnitas los sistemas son los mismos, es decir que la información de cada sistema está en los números, podemos entonces escribir las matrices que contienen esa información:

a) $\left(\begin{array}{cc|c} 2 & -1 & 4 \\ 3 & 2 & 1 \end{array}\right)$ b) $\left(\begin{array}{ccc|c} 1 & -2 & 1 & 4 \\ 5 & -1 & 3 & 3 \end{array}\right)$ c) $\left(\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 4 & -2 & 1 & 1 \\ 1 & 1 & -2 & 3 \end{array}\right)$

Veremos en lo que sigue como resolver los sistemas en forma matricial.

En general un sistema de *m* ecuaciones lineales con *n* incógnitas puede escribirse como:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots\ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots\ldots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots\ldots + a_{3n}x_n = b_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots\ldots + a_{mn}x_n = b_m \end{cases}$$

Podemos expresar las matrices del sistema:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix} \in \mathbb{R}^{mxn} \quad , \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \ldots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \ldots \\ b_m \end{pmatrix} \in \mathbb{R}^{mx1}$$

$A$ es la matriz de coeficientes, $X$ es la matriz de las incógnitas y $b$ la matriz de términos independientes

El sistema lo escribimos entonces en forma equivalente como:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} \end{pmatrix} . \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \ldots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \ldots \\ b_m \end{pmatrix}$$

que indicamos en forma breve como: $AX = b$

Llamamos ***matriz ampliada*** del sistema $(A|b)$, a la matriz $A$ de coeficientes a la que se agrega como última columna la matriz $b$ de los términos independientes.

Una **solución** del sistema es una n-upla $(c_1, c_2, \ldots, c_n)$ de números tales que reemplazando respectivamente $x_1$ por $c_1$, $x_2$ por $c_2, \ldots\ldots, x_n$ por $c_n$ en **cada** **una** de las ecuaciones, se cumple la igualdad en **todas** ellas.

---

### 2. Operaciones elementales y sistemas de ecuaciones lineales

Si $AX = b$ es un sistema de ecuaciones, aplicando operaciones elementales de fila a la matriz $(A|b)$ ampliada con la columna ***b*** de términos independientes, se obtiene una matriz equivalente por filas, el sistema de ecuaciones correspondiente tiene las mismas soluciones que el de partida.

**Ejemplo 2.1:**

Supongamos que queremos resolver el siguiente sistema:

$$\begin{cases} 2x+4y=2 \\ x-y=3 \end{cases}$$

Vamos a resolverlo mostrando que las operaciones elementales sobre las filas de la matriz ampliada del sistema no son más que las operaciones que hacemos habitualmente en las ecuaciones. Resolveremos en paralelo con las ecuaciones y como matriz:

**Ecuaciones**

$$\begin{cases} 2x+4y=2 \\ x-y=3 \end{cases}$$

Multiplicamos la primer ecuación por $\frac{1}{2}$ , a ambos lados de la igualdad:

$$\begin{cases} x+2y=1 \\ x-y=3 \end{cases}$$

A la segunda ecuación le restamos la primera, ya que si tenemos H=M y W=P, también es cierto que W-H=P-M:

$$\begin{cases} x+2y=1 \\ x-x-y-2y=3-1 \end{cases} \longrightarrow \begin{cases} x+2y=1 \\ -3y=2 \end{cases}$$

Multiplicamos la segunda ecuación por $-\frac{1}{3}$ , a ambos lados de la igualdad:

$$\begin{cases} x+2y=1 \\ y=-\dfrac{2}{3} \end{cases}$$

A la primera ecuación le restamos la segunda multiplicada por 2:

$$\begin{cases} x+2y-2y=1-2(-\frac{2}{3}) \\ y=-\frac{2}{3} \end{cases} \longrightarrow \begin{cases} x=\frac{7}{3} \\ y=-\frac{2}{3} \end{cases}$$

Hemos encontrado la solución: $\boxed{x=\frac{7}{3},\ y=-\frac{2}{3}}$

**Forma matricial**

$$(A|b)=\left(\begin{matrix} 2 & 4 \\ 1 & -1 \end{matrix}\left|\begin{matrix} 2 \\ 3 \end{matrix}\right.\right)$$

$F_1 \leftarrow F_1.\frac{1}{2}$

$$\left(\begin{matrix} 1 & 2 \\ 1 & -1 \end{matrix}\left|\begin{matrix} 1 \\ 3 \end{matrix}\right.\right)$$

$F_2 \leftarrow F_2 - F_1$

$$\left(\begin{matrix} 1 & 2 \\ 0 & -3 \end{matrix}\left|\begin{matrix} 1 \\ 2 \end{matrix}\right.\right)$$

$F_2 \leftarrow F_2(-\frac{1}{3})$

$$\left(\begin{matrix} 1 & 2 \\ 0 & 1 \end{matrix}\left|\begin{matrix} 1 \\ -\frac{2}{3} \end{matrix}\right.\right)$$

$F_1 \leftarrow F_1 - 2F_2$

$$\left(\underbrace{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}}_{A_R}\left|\begin{matrix} \frac{7}{3} \\ -\frac{2}{3} \end{matrix}\right.\right)$$

Llegamos a la escalonada y reducida por filas equivalente con $A$, entonces:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} \frac{7}{3} \\ -\frac{2}{3} \end{pmatrix} \longrightarrow \begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} \frac{7}{3} \\ -\frac{2}{3} \end{pmatrix}$$

Hemos encontrado la solución: $\boxed{x=\frac{7}{3},\ y=-\frac{2}{3}}$

El ejemplo anterior justifica que podemos resolver un sistema de ecuaciones llevándolo a su forma matricial y realizando operaciones elementales. La siguiente proposición lo enuncia formalmente.

**Proposición 2.1:** Si $AX = b$ es un sistema de ecuaciones lineales y **e** es una operación elemental de fila, entonces el sistema *e(A). X= e(b)* tiene las mismas soluciones que $AX = b.$

**Demostración:**

Sea $AX = b$ el siguiente sistema:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots. \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \longrightarrow \quad (A|b) = \left(\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} & b_2 \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} & b_3 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} & b_m \end{array}\right)$$

Sea **c** = $(c_1, c_2, \ldots, c_n)$ una solución de $AX = b$ , esto quiere decir que si reemplazamos las incógnitas $(x_1, x_2, \ldots, x_n)$ por estos números $(c_1, c_2, \ldots, c_n)$, se cumplen las igualdades*:*

$$\begin{cases} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n = b_1 \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots. \\ a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n = b_m \end{cases}$$

**Caso 1:**

Sea **e** la operación que multiplica la fila i de $(A|b)$ por un escalar $h \in \mathbb{R}$, $h \neq 0$:

$$\left(\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ ha_{i1} & ha_{i2} & ha_{i3} & \ldots & \ldots & hb_i \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} & b_m \end{array}\right)$$

Y recuperando las ecuaciones tenemos: $\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots. \\ ha_{i1}x_1 + ha_{i2}x_2 + \cdots + ha_{in}x_n = hb_i \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$

Veamos ahora que $(c_1, c_2, \ldots, c_n)$ es solución de este nuevo sistema reemplazando $(x_1, x_2, \ldots, x_n)$ por $(c_1, c_2, \ldots, c_n)$ :

$$\begin{cases} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n = b_1 \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots. \\ ha_{i1}c_1 + ha_{i2}c_2 + \cdots + ha_{in}c_n = hb_i \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots.. \\ a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n = b_m \end{cases}$$

Sólo hay que mirar la ecuación i, ya que las otras son iguales a las del sistema original y por lo tanto se cumplen las igualdades.

Ecuación i: $ha_{i1}c_1 + ha_{i2}c_2 + \cdots + ha_{in}c_n =$

$= h\,\underbrace{(a_{i1}c_1 + a_{i2}c_2 + \cdots + a_{in}c_n)}_{b_i \text{ porque es la ecuación original}} = hb_i$

Luego $(c_1, c_2, \dots, c_n)$ es solución del sistema e(A).x = e(b), para este caso.

**Caso 2:**

Sea **e** la operación que reemplaza la fila i ($F_i$) por $F_i + hF_j$ :

$$\left(\begin{array}{cccccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & & b_2 \\ \dots & \dots & \dots & \dots & \dots & & \dots \\ a_{i1} + ha_{j1} & a_{i2} + ha_{j2} & a_{i3} + ha_{j3} & \dots & a_{in} + ha_{jn} & & b_i + hb_j \\ \dots & \dots & \dots & \dots & \dots & & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} & & b_m \end{array}\right)$$

Y recuperando las ecuaciones tenemos:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots. \\ (a_{i1} + ha_{j1})x_1 + (a_{i2} + ha_{j2})x_2 + \cdots + (a_{in} + ha_{jn})x_n = b_i + hb_j \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots.. \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

Veamos ahora que $(c_1, c_2, \dots, c_n)$ es solución:

$$\begin{cases} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n = b_1 \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots. \\ (a_{i1} + ha_{j1})c_1 + (a_{i2} + ha_{j2})c_2 + \cdots + (a_{in} + ha_{jn})c_n = b_i + hb_j \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots.. \\ a_{m1}c_1 + a_{m2}c_2 + \cdots + a_{mn}c_n = b_m \end{cases}$$

Sólo hay que mirar la ecuación i, ya que las otras son iguales a las del sistema original y por lo tanto se cumplen las igualdades.

Ecuación i: $(a_{i1} + ha_{j1})c_1 + (a_{i2} + ha_{j2})c_2 + \cdots + (a_{in} + ha_{jn})c_n =$

$$= a_{i1}c_1 + ha_{j1}c_1 + a_{i2}c_2 + ha_{j2}c_2 + \cdots + a_{in}c_n + ha_{jn}c_n =$$

$$\underbrace{(a_{i1}c_1 + a_{i2}c_2 + \cdots + a_{in}c_n)}_{b_i} + h\left(\underbrace{a_{j1}c_1 + a_{j2}c_2 + \cdots + a_{jn}c_n}_{b_j}\right) = b_i + hb_j$$

Luego $(c_1, c_2, \ldots, c_n)$ es solución del sistema e(A).x = e(b), para este caso.

**Caso3:**

Sea **e** la operación de permutar $F_i$ con $F_k$ . Este caso es inmediato ya que es solamente permutar dos ecuaciones de lugar.

El enunciado quedó probado para cuando se aplica **una** operación elemental de cualquiera de los tres tipos que hay. También vale cuando se aplica, a las matrices A y b del sistema, un número finito de operaciones elementales de fila ya que es consecuencia de la aplicación iterada de este resultado.

---

En consecuencia, el sistema: $e_1$(**A**).x = $e_1$(**b**) tiene las mismas soluciones que **A**.x=**b**, es decir, son sistemas equivalentes.

Si se aplica una segunda operación elemental, resulta:

$e_2 \circ e_1(A)X = e_2 \circ e_1(b)$ equivalente a $e_1(A)X = e_1(b)$, y también equivalente a $AX = b$

Así, si se aplican k operaciones elementales hasta llegar a la escalonada y reducida por filas equivalente con A:

$$\underbrace{e_k \circ e_{k-1} \circ \ldots e_2 \circ e_1(A)}_{A_R} X = \underbrace{e_k \circ e_{k-1} \circ \ldots e_2 \circ e_1(b)}_{b^*},$$

el sistema $\boldsymbol{A_R X = b^*}$, tiene las mismas soluciones que $\boldsymbol{AX = b}$**.**

**Aplicación a los distintos tipos de Sistemas lineales**

El objetivo es encontrar las soluciones (o bien decidir que no hay) del sistema dado $\boldsymbol{AX = b}$ . El método de aplicar operaciones elementales a la matriz ampliada $(A|b)$, hasta llegar a $(A_R|b^*)$ se conoce con el nombre de Método de Gauss – Jordan:

1) Llevamos el sistema $\boldsymbol{AX = b}$ a su forma matricial, escribiendo la matriz ampliada $(A|b)$
2) Hacemos operaciones elementales por filas en la matriz hasta llegar a $(A_R|b^*)$ .
3) A partir de la última matriz recuperamos las ecuaciones haciendo el producto : $A_R.X = b^*$
4) A partir de las nuevas ecuaciones tenemos un sistema equivalente con el original y encontramos la o las soluciones o concluimos que no las tiene.

**Observación:** un sistema de dos ecuaciones con dos incógnitas puede pensarse geométricamente como el problema de hallar la intersección entre dos rectas. Recordemos que dos rectas pueden cortarse en un punto, pueden ser coincidentes y en consecuencia tener infinitos puntos en común o pueden ser paralelas y no tener ningún punto en común. Un sistema cualquiera con más ecuaciones y más incógnitas, también presentará sólo estas 3 situaciones.

Mostraremos a continuación 3 sistemas de dos ecuaciones con dos incógnitas en los que veremos los 3 casos posibles: que tenga solución única, que tenga infinitas soluciones o que no tenga solución.

**Ejemplo 2.2:** Resolver los siguientes sistemas:

a) $\begin{cases} x+y=2 \\ 2x+3y=3 \end{cases}$

Escribimos la matriz:

$$(A|b)=\left(\begin{array}{cc|c}1 & 1 & 2\\ 2 & 3 & 3\end{array}\right)$$

$F_2 \leftarrow F_2 - 2F_1$

$$\left(\begin{array}{cc|c}1 & 1 & 2\\ 0 & 1 & -1\end{array}\right)$$

$F_1 \leftarrow F_1 - F_2$

$$\left(\begin{array}{c|c}\underbrace{\begin{matrix}1 & 0\\ 0 & 1\end{matrix}}_{A_R} & \underbrace{\begin{matrix}3\\ -1\end{matrix}}_{b^*}\end{array}\right)$$

Entonces:

$$\underbrace{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}} = \begin{pmatrix}3\\ -1\end{pmatrix}$$

$$\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}3\\ -1\end{pmatrix}$$

Por lo tanto, hay solución única: $x=3,\quad y=-1$

b) $\begin{cases} x+y=2 \\ 2x+2y=4 \end{cases}$

Escribimos la matriz:

$$(A|b)=\left(\begin{array}{cc|c}1 & 1 & 2\\ 2 & 2 & 4\end{array}\right)$$

$F_2 \leftarrow F_2 - 2F_1$

$$\left(\begin{array}{c|c}\underbrace{\begin{matrix}1 & 1\\ 0 & 0\end{matrix}}_{A_R} & \underbrace{\begin{matrix}2\\ 0\end{matrix}}_{b^*}\end{array}\right)$$

Entonces:

$$\underbrace{\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}} = \begin{pmatrix}2\\ 0\end{pmatrix}$$

$$\begin{pmatrix}x+y\\ 0\end{pmatrix} = \begin{pmatrix}2\\ 0\end{pmatrix}$$

Entonces:

$x+y=2$

$0=0$

Por lo tanto, hay infinitas soluciones:

$S\{(x,y): x=2-y, y\in\mathbb{R}\}$

c) $\begin{cases} x+y=2 \\ 2x+2y=3 \end{cases}$

Escribimos la matriz:

$$(A|b)=\left(\begin{array}{cc|c}1 & 1 & 2\\ 2 & 2 & 3\end{array}\right)$$

$F_2 \leftarrow F_2 - 2F_1$

$$\left(\begin{array}{c|c}\underbrace{\begin{matrix}1 & 1\\ 0 & 0\end{matrix}}_{A_R} & \underbrace{\begin{matrix}2\\ -1\end{matrix}}_{b^*}\end{array}\right)$$

Entonces:

$$\underbrace{\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}} = \begin{pmatrix}2\\ -1\end{pmatrix}$$

$$\begin{pmatrix}x+y\\ 0\end{pmatrix} = \begin{pmatrix}2\\ -1\end{pmatrix}$$

Entonces:

$x+y=2$

$0=-1$ ABSURDO

Por lo tanto, no hay solución.

El siguiente teorema permite establecer comparando el rango de la matriz de coeficientes con el de la matriz ampliada si el sistema es compatible o incompatible.

**Teorema (Rouché, Frobenius):**

Sea $AX = b$ un sistema de m ecuaciones lineales con n incógnitas,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots\ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots\ldots + a_{2n}x_n = b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots\ldots + a_{3n}x_n = b_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \ldots\ldots + a_{mn}x_n = b_m \end{cases}$$

**1)** $AX = b$ es compatible si y sólo si $R(A) = R(A|b)$

($R(A)$indica el rango de A y $R(A|b)$ indica el rango de la matriz ampliada del sistema)

**2)** Además, si $R(A) = R(A|b) = n$ ($n$ = número de incógnitas), entonces la solución es única, (sistema compatible determinado), si $R(A) = R(A|b) < n$ , entonces tiene infinitas soluciones (sistema compatible indeterminado).

**Observación:** Notemos que este teorema permite clasificar el sistema, no obtener las soluciones en caso que existan.

Si miramos los ejemplos anteriores tenemos:

a) $R(A) = R(A|b) = 2$ , los rangos coinciden y además coinciden con el número de incógnitas, por lo tanto hay solución única.

b) $R(A) = R(A|b) = 1$ , los rangos coinciden pero el número de incógnitas es mayor, por lo tanto hay infinitas soluciones.}

c) $R(A) \neq R(A|b)$ , la matriz de los coeficientes y la matriz ampliada tienen distinto rango, por lo tanto no hay solución.

**Ejemplo 2.3:**

Resolver el sistema $\begin{cases} 2x + 2y + 2z = 4 \\ 2x - y - z = 1 \\ x + 2y - z = -3 \end{cases}$

Se forma la matriz ampliada $(A\ \ b)=\left(\begin{array}{ccc|c} 2 & 2 & 2 & 4 \\ 2 & -1 & -1 & 1 \\ 1 & 2 & -1 & -3 \end{array}\right)$ y se le aplica una sucesión de operaciones elementales (no necesariamente única) para llevarla a $(A_R|b^*)$ , donde $A_R$ es la reducida equivalente con $A$ y $b^*$, la columna resultante de esa misma sucesión de operaciones elementales.

$(A_R|b^*)$ es **única** cualesquiera sean las operaciones elementales elegidas para reducir ***A.***

$$\left(\begin{array}{ccc|c} 2 & 2 & 2 & 4 \\ 2 & -1 & -1 & 1 \\ 1 & 2 & -1 & -3 \end{array}\right) \quad \frac{1}{2}F_1 \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 2 & -1 & -1 & 1 \\ 1 & 2 & -1 & -3 \end{array}\right) \quad \begin{array}{c} F_2-2F_1 \\ F_3-F_1 \end{array} \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & -3 & -3 & -3 \\ 0 & 1 & -2 & -5 \end{array}\right),$$

$$\frac{-1}{3}F_2 \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & -2 & -5 \end{array}\right), \quad \begin{array}{c} F_1-F_2 \\ F_3-F_2 \end{array} \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -3 & -6 \end{array}\right), \quad \frac{-1}{3}F_3 \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right),$$

$$F_2-F_3 \rightarrow \left(\begin{array}{c|c} \underbrace{\begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}}_{A_R} & \underbrace{\begin{matrix} 1 \\ -1 \\ 2 \end{matrix}}_{b^*} \end{array}\right)$$

La matriz $A_R$ reducida equivalente a ***A*** en este caso es la identidad.

El sistema así obtenido da las soluciones del sistema dado.

El sistema equivalente al de partida con notación usual y con la matricial es el siguiente:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \qquad \begin{cases} x+0y+0z=1 \\ 0x+y+0z=-1 \\ 0x+0y+z=2 \end{cases}$$

Observemos que $R(A) = R(A|b) = 3$, por lo tanto por el Teorema de Rouché Frobenius hay **solución única.** Esa solución es : $x=1,\ \ y=\text{-}1,\ \ z=2$.

**Ejemplo 2.4:**

Este es un ejemplo de un sistema con infinitas soluciones (compatible indeterminado), se encontrará el conjunto de sus infinitas soluciones:

$$\begin{cases} x+y-2z+4t=5 \\ 2x+2y-3z+t=3 \\ 3x+3y-4z-2t=1 \end{cases}$$

Se construye la correspondiente matriz ampliada $(A|b)$ y se le aplican operaciones elementales hasta obtener $(A_R|b^*)$:

$$(A|b) = \left(\begin{array}{cccc|c} 1 & 1 & -2 & 4 & 5 \\ 2 & 2 & -3 & 1 & 3 \\ 3 & 3 & -4 & -2 & 1 \end{array}\right) \quad F_2 \leftarrow (F_2 - 2F_1) \longrightarrow \left(\begin{array}{cccc|c} 1 & 1 & -2 & 4 & 5 \\ 0 & 0 & 1 & -7 & -7 \\ 3 & 3 & -4 & -2 & 1 \end{array}\right)$$

$$F_3 \leftarrow (F_3 - 3F_1) \longrightarrow \left(\begin{array}{cccc|c} 1 & 1 & -2 & 4 & 5 \\ 0 & 0 & 1 & -7 & -7 \\ 0 & 0 & 2 & -14 & -14 \end{array}\right) F_1 \leftarrow (F_1 + 2F_2) \longrightarrow \left(\begin{array}{cccc|c} 1 & 1 & 0 & -10 & -9 \\ 0 & 0 & 1 & -7 & -7 \\ 0 & 0 & 2 & -14 & -14 \end{array}\right)$$

$$F_3 \leftarrow (F_3 - 2F_2) \longrightarrow \left(\begin{array}{cccc|c} 1 & 1 & 0 & -10 & -9 \\ 0 & 0 & 1 & -7 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) = (A_R|b^*)$$

En este, como en todos los casos, la sucesión de operaciones elementales aplicada **no** es única, no debe necesariamente ser la que se muestra ahí arriba, pero cualesquiera sean las operaciones elementales y el orden en que se apliquen, **la matriz final obtenida $(\boldsymbol{A_R}|\boldsymbol{b}^*)$ es siempre la misma, ésa es única.**

Observemos que $R(A) = R(A|b) = 2 < 3$, los rangos coinciden pero es menor al número de incógnitas, por lo tanto ya sabemos que tiene infinitas soluciones, debemos encontrar la expresión de ellas.

La matriz $(A_R|b^*)$ es la matriz ampliada correspondiente a un sistema equivalente (es decir: con las mismas soluciones) que el enunciado. Expresado como producto matricial:

$$\begin{pmatrix} 1 & 1 & 0 & -10 \\ 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 \end{pmatrix} . \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} -9 \\ -7 \\ 0 \end{pmatrix}$$

Entonces: $\begin{cases} x+y-10t=-9 \\ z-7t=-7 \\ 0=0 \end{cases}$ del que se obtiene $\begin{cases} x=-9-y+10t \\ z=-7+7t \end{cases}$

Las letras $y, t$ son variables libres, pueden adoptar cualquier número real, esto se indica con: $y, t \in \mathbb{R}$. En cambio, $x, z$ son dependientes (dependen de los valores elegidos para $y$ y para $t$).

El hecho de que haya al menos una variable libre, que recorre todos los números reales, hace que el conjunto de soluciones del sistema sea infinito.

En este ejemplo el **conjunto de infinitas soluciones** del sistema se expresa como:

$$\boxed{S=\{(x,y,z,t):\ x=-9-y+10t\,, \quad z=-7+7t,\ \ y,t\in\mathbb{R}\}}$$

O bien $S=\{(-9-y+10t, y, -7+7t, t):,\ \ y,t\in\mathbb{R}\}$

Si se quiere obtener una solución particular, se asignan números determinados a las variables libres (en este ejemplo *y, t)* y con esos valores se obtienen los de *x* y *z.*

Si elegimos $y=0, t=1$ obtenemos la solución: $x=1, y=0, z=0, t=1$.

Si elegimos $y=2, t=-1$ obtenemos la solución: $x=-21, y=2, z=-14, t=-1$.

**Observación**: La palabra "indeterminado" sugiere que las soluciones no pueden conocerse, esto no es así: las infinitas soluciones de un sistema compatible indeterminado siguen todas una misma "ley" (que suele llamarse solución general), a partir de la cual se puede obtener cualquier solución particular, como las que se muestran en el ejemplo anterior.

**Ejemplo 2.5:**

Resolver el siguiente sistema:

$$\begin{cases} x+y+2z=1 \\ 3x+2y+z=5 \\ 2x+y-z=3 \end{cases}$$

$$(A|b)=\left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 3 & 2 & 1 & 5 \\ 2 & 1 & -1 & 3 \end{array}\right) \quad F_2 \leftarrow (F_2-3F_1) \longrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & -1 & -5 & 2 \\ 2 & 1 & -1 & 3 \end{array}\right)$$

$$F_3 \leftarrow (F_3-2F_1) \longrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & -1 & -5 & 2 \\ 0 & -1 & -5 & 1 \end{array}\right) F_2 \leftarrow (F_2.(-1) \longrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & 5 & -2 \\ 0 & -1 & -5 & 1 \end{array}\right)$$

$$F_1 \leftarrow (F_1 - F_2) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & -3 & 3 \\ 0 & 1 & 5 & -2 \\ 0 & -1 & -5 & 1 \end{array}\right) F_3 \leftarrow (F_3 + F_2) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & -3 & 3 \\ 0 & 1 & 5 & -2 \\ 0 & 0 & 0 & -1 \end{array}\right) = (A_R|b^*)$$

En el sistema equivalente que se obtiene, queda la última ecuación $0x + 0y + 0z = -1$, imposible de resolver, equivale (cualesquiera sean x, y , z ) a la igualdad falsa 0 = -1.

Por lo tanto, **el sistema no tiene solución, también decimos que es un sistema incompatible.**

Observemos que $R(A) \neq R(A|b)$, por lo tanto también podemos usar el Teorema de Rouché Frobenius para justificar que el sistema no tiene solución.

---

Los sistemas de ecuaciones son útiles para representan una gran cantidad de problemas. Veremos algunos ejemplos.

**Ejemplo 2.6:**

Ana tiene el triple de edad que su hijo Jaime. Dentro de 15 años, la edad de Ana será el doble que la de su hijo. ¿Cuántos años más que Jaime tiene su madre?

Para resolverlo llamamos $A$ a la edad de Ana y $J$ a la edad de Jaime y planteamos las ecuaciones:

$$\begin{cases} A = 3J \\ A + 15 = 2(J + 15) \end{cases}$$

Ana tiene el triple de edad que su hijo Jaime

Dentro de 15 años, la edad de Ana será el doble que la de su hijo

Reacomodamos las ecuaciones, dejando las incógnitas de un lado de la igualdad y los números del otro:

$$\begin{cases} A - 3J = 0 \\ A - 2J = 15 \end{cases}$$

Ya estamos en condiciones de escribir la matriz ampliada y resolver:

$$\left(\begin{array}{cc|c} 1 & -3 & 0 \\ 1 & -2 & 15 \end{array}\right) \quad F_2 \leftarrow (F_2 - F_1) \rightarrow \left(\begin{array}{cc|c} 1 & -3 & 0 \\ 0 & 1 & 15 \end{array}\right) F_1 \leftarrow (F_1 + 3F_2) \rightarrow \left(\begin{array}{cc|c} 1 & 0 & 45 \\ 0 & 1 & 15 \end{array}\right)$$

Hemos llegado a la escalonada y reducida por filas, esto nos dice que $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} A \\ J \end{pmatrix} = \begin{pmatrix} 45 \\ 15 \end{pmatrix}$

Entonces $\begin{pmatrix} A \\ J \end{pmatrix} = \begin{pmatrix} 45 \\ 15 \end{pmatrix}$ y por lo tanto el sistema tiene solución única $A = 45,\ J = 15$

Esas son las edades actuales de Ana y de su hijo Jaime, por lo tanto Ana tiene 30 años más que su hijo.

**Ejemplo 2.7:**

Una empresa de transportes gestiona una flota de 60 camiones, entre grandes, medianos y pequeños. Los grandes transportan una media diaria de 15000kg y recorren 400km. Los medianos transportan 10000kg y recorren 300km y los pequeños transportan 5000kg y recorren 100km. Diariamente los camiones transportan 475000kg y recorren 12500km entre todos. ¿Cuántos camiones de cada tipo gestiona la empresa?

Para resolverlo llamamos $G$ a la cantidad de camiones grandes, $M$ a la cantidad de camiones medianos y $P$ a la cantidad de camiones pequeños y planteamos las ecuaciones:

$$\begin{cases} G + M + P = 60 \\ 15000G + 10000M + 5000P = 475000 \\ 400G + 300M + 100P = 12500 \end{cases}$$

Tenemos entonces un sistema de 3 ecuaciones con 3 incógnitas:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 60 \\ 15000 & 10000 & 5000 & 475000 \\ 400 & 300 & 100 & 12500 \end{array}\right)$$

$$F_2 \leftarrow (F_2 - 15000F_1) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 60 \\ 0 & -5000 & -10000 & -425000 \\ 400 & 300 & 100 & 12500 \end{array}\right)$$

$$F_3 \leftarrow (F_3 - 400F_1) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 60 \\ 0 & -5000 & -10000 & -425000 \\ 0 & -100 & -300 & -11500 \end{array}\right)$$

$$F_2 \leftarrow \left(F_2.(-\frac{1}{5000})\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 60 \\ 0 & 1 & 2 & 85 \\ 0 & -100 & -300 & -11500 \end{array}\right)$$

$$F_1 \leftarrow (F_1 - F_2) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1 & -25 \\ 0 & 1 & 2 & 85 \\ 0 & -100 & -300 & -11500 \end{array}\right)$$

$$F_3 \leftarrow (F_3 + 100F_2) \quad \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1 & -25 \\ 0 & 1 & 2 & 85 \\ 0 & 0 & -100 & -3000 \end{array}\right)$$

$$F_3 \leftarrow \left(F_3.\left(-\frac{1}{100}\right)\right) \quad \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1 & -25 \\ 0 & 1 & 2 & 85 \\ 0 & 0 & 1 & 30 \end{array}\right)$$

$$F_1 \leftarrow (F_1 + F_3) \quad \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 2 & 85 \\ 0 & 0 & 1 & 30 \end{array}\right) \qquad F_2 \leftarrow (F_2 - 2F_3) \quad \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 25 \\ 0 & 0 & 1 & 30 \end{array}\right)$$

Hemos llegado a la escalonada y reducida por filas, esto nos dice que:

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} G \\ M \\ P \end{pmatrix} = \begin{pmatrix} 5 \\ 25 \\ 30 \end{pmatrix}$ entonces $\begin{pmatrix} G \\ M \\ P \end{pmatrix} = \begin{pmatrix} 5 \\ 25 \\ 30 \end{pmatrix}$ , el sistema tiene solución única:

La empresa tiene 5 camiones grandes, 25 medianos y 30 pequeños.

---

**Sistemas homogéneos de ecuaciones lineales**

**Definición:** Un sistema de ecuaciones lineales es **homogéneo** si todos los términos independientes son ceros.

Su notación matricial es $AX = O$ donde $O$ indica una matriz columna nula.

**Observación:**

La importancia de los sistemas homogéneos reside en que son sistemas siempre compatibles.

Por ejemplo, en el sistema $\begin{cases} 3x + 4y - 6z = 0 \\ 2x + y + 10z = 0 \\ 70x - \sqrt{2}y + 4z = 0 \end{cases}$ , si reemplazamos las incógnitas por 0, seguro satisfacen las ecuaciones, por lo tanto, en cualquier sistema homogéneo, esa es una solución.

***Todo sistema homogéneo es compatible***: la solución **trivial** es aquella en la que se reemplazan todas las incógnitas por ceros, esta solución nula la tienen todos los sistemas homogéneos, por esa razón se llama trivial.

Si el sistema homogéneo tiene una única solución debe necesariamente ser la trivial. Si tiene infinitas soluciones, una de ellas debe ser la trivial.

Aplicando el teorema de Rouché Frobenius a un sistema homogéneo los rangos de ambas matrices siempre coinciden debido a que la columna que se agrega para construir la matriz ampliada es nula. Este número puede coincidir con el número n de incógnitas (compatible determinado) o ser menor que n (compatible indeterminado).

Los sistemas homogéneos se resuelven mediante operaciones elementales como los anteriores.

---

**Ejercicios:**

**1)** Representar en el plano los siguientes sistemas, indicar qué tipo de solución hay en cada caso: $\begin{cases} x + 2y = 0 \\ 2x + 3y = 1 \end{cases}$ $\begin{cases} 2x + 6y = 8 \\ x + 3y = 4 \end{cases}$ $\begin{cases} 2x - y = 3 \\ 6x - 3y = -5 \end{cases}$

**2)** Verificar que los valores dados son *soluciones* de los sistemas planteados:

a) x=1, y =2 para $\begin{cases} 3x + y = 5 \\ x - y = -1 \end{cases}$ b) x=1, y=1 para $\begin{cases} x - y = 0 \\ x + y = 2 \\ 3x - y = 2 \end{cases}$

c) {(x,y) ; x = -3y} para $\begin{cases} x + 3y = 0 \\ -3x - 9y = 0 \\ 2x + 6y = 0 \end{cases}$ d) x=15/9, y=8/9, z= -11/9 para $\begin{cases} 2x + y + z = 3 \\ y + 4z = -4 \\ x - y - z = 2 \end{cases}$

e) {(x, y, z) ; x =1+y, z=3y} para $\begin{cases} x + 2y - z = 1 \\ x - y = 1 \end{cases}$

**3)** Resolver los siguientes sistemas por operaciones elementales y expresar la o las soluciones en caso de que existan. Clasificarlos por el teorema de Rouché-Frobenius. En los casos que haya infinitas soluciones dar dos soluciones particulares:

a) $\begin{cases} x_1 - x_2 + x_3 = 2 \\ x_1 \quad - x_3 = 6 \\ \quad x_2 + x_3 = 4 \end{cases}$ b) $\begin{cases} u_1 + u_2 + u_3 = 1 \\ \quad u_2 - u_3 = 1 \end{cases}$ c) $\begin{cases} 2x + 3y - z = 1 \\ x + y + z = 0 \\ x + 2y + 3z = 4 \end{cases}$ d) $\begin{cases} x + y = 1 \\ x - 2y = 3 \\ 3x - y = 0 \end{cases}$

e) $\begin{cases} x + 2y + z = 7 \\ 3x \quad + 5z = 4 \\ 2x + y + 3z = -1 \end{cases}$ f) $\begin{cases} x_1 - 2x_2 + 3x_3 = 1 \\ 2x_1 + 5x_2 - x_3 + 2x_4 = 2 \end{cases}$ g) $\begin{cases} x - 2y + 3z - 2w = 0 \\ 3x - 7y - 2z + 4w = 0 \\ 4x + 3y + 5z + 2w = 0 \end{cases}$

**4)** Analizar la ***verdad*** o ***falsedad*** de las siguientes afirmaciones. Fundamentar la respuesta

a) Todo sistema homogéneo tiene al menos una solución.

b) Los sistemas homogéneos tienen siempre infinitas soluciones.

c) Un sistema homogéneo que no tiene una única solución, tiene infinitas soluciones.

d) Si un sistema no homogéneo no tiene solución única, debe tener infinitas soluciones.

e) Si un sistema tiene más de una solución, entonces tiene infinitas.

f) La ecuación x + y = 0 no tiene solución.

g) Si para cada ecuación del sistema hay alguna solución, entonces el sistema tiene solución.

h) Si un sistema es incompatible, entonces cada ecuación del mismo tampoco tiene solución.

**5)** Determinar (si existen) los valores de ***b*** para que los siguientes sistemas sean

***i)*** compatible (en tal caso resolverlo, expresar la solución en la forma adecuada)

***ii)*** incompatible

a) $\begin{cases} x_1 + x_2 - x_3 + 3x_4 = 1 \\ 3x_1 \quad + 2x_3 + 7x_4 = 5 \\ 2x_1 - x_2 + 3x_3 + 4x_4 = b \end{cases}$ b) $\begin{cases} x + 5y - z = 7 \\ \quad 4y + 3z = b \\ -x + 3y + 7z = 5 \end{cases}$ c) $\begin{cases} x - y + \frac{1}{2}z - 3w = 8 \\ -2x + 2y - z + 6w = b \end{cases}$

**6)** Determinar qué relación debe haber entre *a, b* y *c* para que este sistema sea compatible:

$$\begin{cases} x + 2y - 3z = a \\ 2x + 6y - 11z = b \\ x - 2y + 7z = c \end{cases}$$

**7)** Si la terna $(2,1,-1)$ es <u>**una *solución***</u> de $\begin{cases} x_1 - x_2 + x_3 = a \\ 2x_1 \quad + x_3 = b \\ x_1 + x_2 \quad = c \end{cases}$ hallar todas las soluciones del sistema. (piense dónde debe reemplazar 2, 1 y -1)

**8)** En cada caso determinar, si existen, los valores de ***k*** tales que el sistema resulte, respectivamente:

i) compatible determinado

ii) compatible indeterminado

iii) incompatible

(a) $\begin{cases} x_1 + 3x_2 = 4 \\ 2x_1 + k.x_2 = k+2 \end{cases}$ (b) $\begin{cases} x_1 + 2x_2 = 4 \\ 3x_1 + (k-1).x_2 = k \end{cases}$ (c ) $\begin{cases} x_1 + x_2 + x_3 = k \\ x_1 + k.x_2 + x_3 = 1 \end{cases}$

**9)** Resuelve el problema de la Pila de monedas de Fibonacci planteado al principio del capítulo y analiza si la solución hallada por Fibonacci es correcta.

---

**3. Determinante de una matriz.**

El determinante de una matriz es un número real que se asigna SOLO A MATRICES CUADRADAS. Este número fue primeramente calculado para encontrar condiciones para que un sistema de ecuaciones de 2x2 tuviera solución única, esto motiva el estudio de los determinantes y nos dará una poderosa herramienta para analizar los sistemas de ecuaciones cuya matriz de coeficientes sea cuadrada.

Si A está en $\mathbb{R}^{nxn}$ el determinante de A, que se nota **det(A)** o $|\boldsymbol{A}|$ , es un número real que se puede pensar como una función, con dominio en $\mathbb{R}^{nxn}$ y codominio en $\mathbb{R}$:
det**:** $\mathbb{R}^{nxn} \to \mathbb{R}$ que a toda matriz A en $\mathbb{R}^{nxn}$ le asigna el único número **det(A)** $\in \mathbb{R}$.

**Ejemplo 3.1:**

Sea $A = (3)$ una matriz de 1x1, el determinante de $A$, que se escribe det(A) o $|\boldsymbol{A}|$ es 3.
Es decir es el único número que tiene la matriz.

**Ejemplo 3.2:**

Sea $A = \begin{pmatrix} 2 & 3 \\ 5 & 4 \end{pmatrix}$ una matriz de 2x2.

Para calcular el determinante, elegimos cualquier fila o cualquier columna de la matriz, elijamos la fila 1, tenemos entonces:

det(A)= $2.\ (-\mathbf{1})^{\mathbf{1+1}}\ .\boldsymbol{det}(\mathbf{4}) \quad + \quad \mathbf{3}.\ (-\mathbf{1})^{\mathbf{1+2}}.\boldsymbol{det}(\mathbf{5})$=

$a_{11}$

(-1) elevado al número de fila más columna

Determinante de la matriz que queda sacando fila1, columna1

$a_{12}$

(-1) elevado al número de fila más columna

Determinante de la matriz que queda sacando fila1, columna2

det(A)= 2.4 – 3.5 = 8-15= - 7

Es decir que de manera práctica podemos calcularlo como el producto de los elementos de la diagonal principal menos el producto de los elementos de la diagonal secundaria.

**Ejemplo 3.3:**

Sea $A = \begin{pmatrix} 1 & 3 & -1 \\ 2 & 0 & 2 \\ 1 & 4 & 1 \end{pmatrix}$una matriz de 3x3.

Para calcular el determinante, elegimos cualquier fila o cualquier columna de la matriz, elijamos la fila 1, tenemos entonces:

det(A)= $1.\ (-\mathbf{1})^{\mathbf{1+1}}\ .\boldsymbol{det}\begin{pmatrix} \mathbf{0} & \mathbf{2} \\ \mathbf{4} & \mathbf{1} \end{pmatrix} + \mathbf{3}.\ (-\mathbf{1})^{\mathbf{1+2}}.\boldsymbol{det}\begin{pmatrix} \mathbf{2} & \mathbf{2} \\ \mathbf{1} & \mathbf{1} \end{pmatrix} + (-\mathbf{1}).\ (-\mathbf{1})^{\mathbf{1+3}}.\ \boldsymbol{det}\begin{pmatrix} \mathbf{2} & \mathbf{0} \\ \mathbf{1} & \mathbf{4} \end{pmatrix}$=

$a_{11}$

(-1) elevado al número de fila más columna

Determinante de la matriz que queda sacando fila1, columna1

$a_{12}$

(-1) elevado al número de fila más columna

Determinante de la matriz que queda sacando fila1, columna2

$a_{13}$

(-1) elevado al número de fila más columna

Determinante de la matriz que queda sacando fila1, columna3

=1.1 $.\,(\mathbf{0.1-4.2}) + \ \mathbf{3}.(-\mathbf{1}).(\mathbf{2.1-1.2}) + (-\mathbf{1}).\mathbf{1}.(\mathbf{2.4-1.0}) = -\mathbf{8}+\mathbf{0}-\mathbf{8} = -\mathbf{16}$

Si eligiéramos otra fila o cualquier columna el resultado sería el mismo. Desarrollemos el determinante por la columna 2:

det(A)= $3.\ (-\mathbf{1})^{1+2}\ .\ \boldsymbol{det}\begin{pmatrix}\mathbf{2} & \mathbf{2}\\ \mathbf{1} & \mathbf{1}\end{pmatrix} + \mathbf{0}.\ (-\mathbf{1})^{2+2}\ .\ \boldsymbol{det}\begin{pmatrix}\mathbf{1} & -\mathbf{1}\\ \mathbf{1} & \mathbf{1}\end{pmatrix} + \mathbf{4}.\ (-\mathbf{1})^{3+2}.\ \boldsymbol{det}\begin{pmatrix}\mathbf{1} & -\mathbf{1}\\ \mathbf{2} & \mathbf{2}\end{pmatrix}$=

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $a_{12}$ | (-1) elevado al número de fila más columna | Determinante de la matriz que queda sacando fila1, columna2 | $a_{22}$ | (-1) elevado al número de fila más columna | Determinante de la matriz que queda sacando fila2, columna2 | $a_{32}$ | (-1) elevado al número de fila más columna | Determinante de la matriz que queda sacando fila3, columna2 |

$=3.(-1).(\mathbf{2.1-1.2})+\mathbf{0.1}.\big(\mathbf{1.1-1.(-1)}\big)+\mathbf{4}.(-\mathbf{1}).(\mathbf{1.2-2.(-1)}=\mathbf{0+0-16=-16}$

Se define a continuación, por recurrencia, el determinante de cualquier matriz **cuadrada**:

**Definición:** Sea A matriz de $\mathbb{R}^{nxn}$,

1) Si *n= 1*, A = $(a_{11})$, se define **det(A) =** $a_{11}$ **o** $|A|$**=**$a_{11}$

2) Si *n es cualquier número natural mayor o igual que 2*, se define, dada cualquier fila $\boldsymbol{i}$**:**

**det(A)=** $\boldsymbol{a_{i1}}.(-\mathbf{1})^{\boldsymbol{i}+\mathbf{1}}\boldsymbol{det}\big(\boldsymbol{A}(\boldsymbol{i}|\mathbf{1})\big)+\boldsymbol{a_{i2}}.(-\mathbf{1})^{\boldsymbol{i}+\mathbf{2}}\boldsymbol{det}\big(\boldsymbol{A}(\boldsymbol{i}|\mathbf{2})\big)+\cdots+\boldsymbol{a_{in}}.(-\mathbf{1})^{\boldsymbol{i}+\boldsymbol{n}}\boldsymbol{det}\big(\boldsymbol{A}(\boldsymbol{i}|\boldsymbol{n})\big)$

o también, dada cualquier columna $\boldsymbol{j}$**:**

**det(A)=** $\boldsymbol{a_{1j}}.(-\mathbf{1})^{\mathbf{1}+\boldsymbol{j}}\boldsymbol{det}\big(\boldsymbol{A}(\mathbf{1}|\boldsymbol{j})\big)+\boldsymbol{a_{2j}}.(-\mathbf{1})^{\mathbf{2}+\boldsymbol{j}}\boldsymbol{det}\big(\boldsymbol{A}(\mathbf{2}|\boldsymbol{j})\big)+\cdots+\boldsymbol{a_{nj}}.(-\mathbf{1})^{\boldsymbol{n}+\boldsymbol{j}}\boldsymbol{det}\big(\boldsymbol{A}(\boldsymbol{n}|\boldsymbol{j})\big)$

Por comodidad escribimos ambas expresiones con la notación sigma:

$$\boldsymbol{det}(\boldsymbol{A})=\underbrace{\sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{n}}\boldsymbol{a_{ik}}.\overbrace{(-\mathbf{1})^{\boldsymbol{i}+\boldsymbol{k}}\boldsymbol{det}\big(\boldsymbol{A}(\boldsymbol{i}|\boldsymbol{k})\big)}^{\alpha_{ik}}}_{\text{Está fija la fila i, va variando la columna}}=\underbrace{\sum_{\boldsymbol{k}=\mathbf{1}}^{\boldsymbol{n}}\boldsymbol{a_{kj}}.\overbrace{(-\mathbf{1})^{\boldsymbol{k}+\boldsymbol{j}}\boldsymbol{det}\big(\boldsymbol{A}(\boldsymbol{k}|\boldsymbol{j})\big)}^{\alpha_{kj}}}_{\text{Está fija la columna j, va variando la fila}}$$

Se llama ***adjunto*** (o ***cofactor***) del coeficiente $a_{ij}$ al valor $\alpha_{i\,j} = (-1)^{i+j}$.**det** $A(i\,/\,j)$ **.**

El ***determinante*** de A es la suma de los productos de los coeficientes de una fila (o una columna) <u>cualquiera</u> de A por sus respectivos cofactores, es decir, desarrollado por la fila *i* :

$$\det(A)=\sum_{j=1}^{n}a_{ij}.\alpha_{ij}=a_{i1}.\alpha_{i\,1}+a_{i2}.\alpha_{i\,2}+\ldots+a_{in}.\alpha_{i\,n}$$

**Observación:** El determinante está bien definido: el valor det(A) es único cualquiera sea la fila o la columna de A que se elija para calcularlo.

**Ejemplo 3.4:**

Calcular det(A) siendo A = $\begin{pmatrix} 3 & -1 & 2 \\ 2 & -1 & -2 \\ 0 & 1 & -3 \end{pmatrix}$

Calcularemos por la fila 2: det(A) = $a_{21} . \alpha_{21} + a_{22} . \alpha_{22} + a_{23} . \alpha_{23}$ ,

calculando primero los cofactores $\alpha_{2j}$ :

$\alpha_{21} = (-1)^{2+1}$ . det A(2/1) = ( - 1) . $\det\begin{pmatrix} -1 & 2 \\ 1 & -3 \end{pmatrix}$ **= -1**

$\alpha_{22} = (-1)^{2+2}$ .det A(2/2) = $\det\begin{pmatrix} 3 & 2 \\ 0 & -3 \end{pmatrix}$ = **- 9**

$\alpha_{23} = (-1)^{2+3}$ . det A(2/3) = $(-1). \det\begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix}$ = **- 3**

Entonces:

**det(A) = 2 . (- 1) + (- 1). ( - 9) + (- 2 ) . (- 3 ) = 13**

El mismo valor 13 se obtiene calculando det(A) por cualquier otra fila o por cualquier columna. Tomando por ejemplo la primera columna:

det (A) = $3.(-1)^{1+1}.\det\begin{pmatrix} -1 & -2 \\ 1 & -3 \end{pmatrix}$ + $2.(-1)^{2+1}.\det\begin{pmatrix} -1 & 2 \\ 1 & -3 \end{pmatrix}$ **+** 0 = 13

**Ejemplo 3.5:**

Calcular det (A) siendo A = $\begin{pmatrix} 2 & 3 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 4 & -1 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{pmatrix}$

Calcularemos por la fila 1:

det (A) = $\underbrace{2.(-1)^{1+1}}_{2}.\underbrace{det\begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}}_{A(1/1)} + \underbrace{3.(-1)^{1+2}}_{-3}.\underbrace{det\begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}}_{A(1/2)}$ + 0 + 0 =

$$= 2.\left[\overbrace{\underbrace{2.(-1)^{1+1}}_{2}.\underbrace{det\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}}_{2} + 0 + 0}^{\det(A(1/1))\text{ por fila 1}}\right] - 3\left[\overbrace{\underbrace{1.(-1)^{1+1}}_{1}.\underbrace{det\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}}_{2} + 0 + 0}^{\det(A(1/2))\text{ por fila 1}}\right] =$$

$$2.2.2 - 3.1.2 = 8 - 6 = 2$$

Entonces: **det(A) = 2**

Observemos que siempre nos conviene elegir la fila o la columna que tenga más ceros para hacer menos cuentas. En este ejemplo podríamos haber elegido la columna 4:

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

Calcularemos por la columna 4:

det (A) = $0 + 0 + 0 + \underbrace{2.(-1)^{4+4}}_{2}.\underbrace{det\begin{pmatrix} 2 & 3 & 0 \\ 1 & 2 & 0 \\ 4 & -1 & 1 \end{pmatrix}}_{A(4/4)}$ =

$$= 2.\left[\overbrace{0 + 0 + \underbrace{1.(-1)^{3+3}}_{1}.\underbrace{det\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}}_{1}}^{\det(A(4/4))\text{ por columna 3}}\right] = 2.1 = 2$$

Entonces: **det(A) = 2**

### Algunas propiedades del determinante

A continuación se enuncian algunas propiedades de los determinantes. Existen además otras propiedades no incluidas en los contenidos de esta asignatura. Demostraremos sólo algunas. Por la definición del determinante, en la que hemos mencionado que puede calcularse por fila o por columna, todas las propiedades enunciadas a continuación para las filas vale también para las columnas de la matriz.

Sea $A \in \mathbb{R}^{\boldsymbol{nxn}}$**:**

**Propiedad 1)** $\mathbf{det}\left(\boldsymbol{F_i}(\boldsymbol{c})(\boldsymbol{A})\right) = \boldsymbol{c}.\mathbf{det}(\boldsymbol{A})$

Si $B$ es la matriz que se obtiene de $A$ luego de aplicar la operación elemental de multiplicar todos los coeficientes de **una** fila (respectivamente una columna) por un escalar *c* entonces el $\det(B) = c.\det(A)$

**Demostración:**

Llamemos *B* a la matriz que se obtiene multiplicando la fila k de *A* por c, los elementos de *B*, son iguales a los de *A*, salvo los de la fila k, que son los de *A* multiplicados por el número c, entonces : $B = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ ca_{k1} & ca_{k2} & \dots & ca_{kn} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$

Calculamos el det(*B)* por la fila k,

$$\boldsymbol{det}(\boldsymbol{B}) = \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{b_{kj}}.(-\mathbf{1})^{\boldsymbol{k}+\boldsymbol{j}}\boldsymbol{det}\left(\boldsymbol{B}(\boldsymbol{k}|\boldsymbol{j})\right)$$

Por la definición de $B,$ $det\left(B(k|j)\right) = det\left(A(k|j)\right)$, porque en ambas se suprime la fila *k* que es la única en la que difieren, y $\boldsymbol{b_{kj}} = \boldsymbol{c}.\boldsymbol{a_{kj}}$, entonces se tiene:

$$\boldsymbol{det}(\boldsymbol{B}) = \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{c}.\boldsymbol{a_{kj}}.(-\mathbf{1})^{\boldsymbol{k}+\boldsymbol{j}}\boldsymbol{det}\left(\boldsymbol{A}(\boldsymbol{k}|\boldsymbol{j})\right) = \boldsymbol{c}.\left(\sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{a_{kj}}.(-\mathbf{1})^{\boldsymbol{k}+\boldsymbol{j}}\boldsymbol{det}\left(\boldsymbol{A}(\boldsymbol{k}|\boldsymbol{j})\right)\right) = \boldsymbol{c}.\boldsymbol{det}(\boldsymbol{A})$$

**Corolario de la Propiedad 1** Si consideramos la matriz $c.A$ (producto del escalar *c* por la matriz *A*), entonces: $\det(c.A) = c^n.\det(A)$

**Propiedad 2)** Si una fila (respectivamente una columna) de *A* tiene todos sus coeficientes iguales a 0 entonces det(*A)* = 0.

**Demostración:**

Sea $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{array}{l} \\ \\ \longrightarrow \text{fila k} \\ \\ \\ \end{array}$

Desarrollamos el determinante por la fila k de A:

$$det(A) = \sum_{j=1}^{n} a_{kj}.(-1)^{k+j} det\big(A(k|j)\big) = \sum_{j=1}^{n} 0.(-1)^{k+j} det\big(A(k|j)\big) = 0$$

**Propiedad 3)** Sean $A$ y $B$ matrices de $\mathbb{R}^{nxn}$ entonces $\det(A.B) = \det A . \det B$

**Observaciones**

1) Dado un $n$ fijo, la propiedad 4 se generaliza a cualquier número finito de matrices $n \times n$

Si $M_1, M_2, \dots, M_k$ son matrices $n \times n$, entonces
$\det(M_1.M_2.....M_k) = \det M_1 . \det M_2 ..... \det M_k$.

2) Esta propiedad del determinante es, tal como se indica, para el **producto** de matrices, **no** existe ninguna con relación a la suma de matrices

**Propiedad 4)** Para la matriz identidad $I_n$ , $\det(I_n) = 1$, cualquiera sea $n$.

**Demostración:**

Lo demostraremos por inducción sobre el número de filas de la matriz, el enunciado entonces es: P(n): "$\det(I_n) = 1$ " para todo n, natural mayor o igual a 1.

1) Para n=1 es trivial por la definición. Por lo tanto vale P(1).

2) Si $\underbrace{\mathrm{P(k)}: \det(I_k) = 1}_{Hipótesis\ inductiva}$ **entonces** $\underbrace{\mathrm{P(k+1)}: \ \det(I_{k+1}) = 1}_{Tesis\ inductiva}$

Para calcular $\det(I_{k+1})$ lo desarrollamos por la columna 1, entonces:

$$\det(I_{k+1}) = 1.(-1)^{1+1}.\det(I_k) + 0 + 0 + \cdots + 0 = 1.(-1)^{1+1}.\det(I_k)$$

Columna con un solo elemento no nulo $\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$ $\det(I_k) = 1$ por hipótesis

Por lo tanto $\det(I_{k+1}) = 1.(-1)^{1+1}.\det(I_k) = 1.1.1. = 1$

Hemos probado los dos pasos de la inducción, por lo tanto $\det(I_n) = 1$ para todo natural.

**Propiedad 5)** $\mathbf{det}\big((F_i \leftarrow F_i + cF_k)(A)\big) = \mathbf{det}(A)$

Si $B$ es la matriz que se obtiene de $A$ luego de aplicar la operación elemental de sumarle a la fila i la fila k multiplicada por un escalar *c* , distinto de 0, entonces el $\det(B) = \det(A)$

**Propiedad 6)** $\mathbf{det}\big((F_i \leftrightarrow F_k)(A)\big) = -\mathbf{det}(A)$

Si $B$ es la matriz que se obtiene de $A$ luego de aplicar la operación elemental de permutar la fila i con la fila k, entonces $\det(B) = -\det(A)$

**Propiedad 7) Condición necesaria y suficiente para la existencia de la inversa:**

Sea $A \in \mathbb{R}^{nxn}$,

A tiene inversa **si y sólo si** $\det(A) \neq 0$.

El enunciado afirma dos cosas:

1) Si A tiene inversa **entonces** su determinante es distinto de 0.

2) Si el determinante de A es distinto de 0 **entonces** A tiene inversa.

Demostración de 1):

Sabemos que A tiene inversa, entonces existe $A^{-1}$ tal que $A.A^{-1} = I$

Cuando dos matrices son iguales su determinante también lo es, (notar que no es cierto la recíproca, es decir si dos matrices tienen el mismo determinante no tienen por qué ser iguales), por lo tanto:

$\det(A.A^{-1}) = \det(I)$ y por las propiedades 4 y 5 se tiene que $\det(A).\det(A^{-1}) = 1$

Tenemos entonces un producto de números reales igual a 1, por lo tanto ninguno de los dos factores puede ser 0, en particular **determinante de A es distinto de 0**.

Demostración de 2):

Sabemos que el determinante de A es distinto de 0, queremos ver que A es invertible.

Para eso alcanza con probar que A es equivalente por filas con la identidad.

Si le aplicamos un número finito de operaciones elementales a A para llegar a la escalonada y reducida por filas equivalente con A, el $\det(A_R)$ será:

* igual al de A, o
* cambiará de signo con el de A o
* será un número, no nulo, multiplicado por el determinante de A, por las propiedades 1, 5 y 6.

Es decir que si $\det(A) \neq 0$ entonces $\det(A_R) \neq 0$ , entonces $A_R$ no puede tener una fila de 0, porque si así fuera su determinante valdría 0, esto quiere decir que $A_R = I$, por lo tanto **A tiene inversa**.

Esta importante propiedad nos garantiza una condición necesaria y suficiente para la existencia de la inversa.

---

**Propiedad 8)** Si la matriz $A$ tiene inversa entonces $\det(A^{-1}) = \frac{1}{\det(A)}$

**Demostración:**

Sabemos que $A$ tiene inversa , entonces : $A.A^{-1} = A^{-1}.A = I_n$

Tomemos una igualdad: $A^{-1}.A = I_n$

Aplicamos determinante a ambos lados, ya que si las matrices son iguales sus determinantes también:

$$\underbrace{\det(A^{-1}.A)}_{\boxed{\det(A^{-1}).\det(A) \text{ por propiedad } 3}} = \underbrace{\det(I_n)}_{\boxed{\det(I_n) = 1 \text{ por propiedad } 4}}$$

Entonces:

$$\det(A^{-1}).\underbrace{\det(A)}_{\boxed{\text{Es un número real distinto de 0, por propiedad 7}}} = 1$$

Entonces:

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

**Propiedad 9)** Dado un sistema de $n$ ecuaciones con $n$ incógnitas $AX = b,$ tiene solución única **si y sólo si** $A$ tiene inversa.

**Demostración: 1)** $\underbrace{AX = b \text{ tiene solución única}}_{hipótesis}$ entonces $\underbrace{A \text{ tiene inversa}}_{tesis}$

Como $AX = b$ tiene solución única, sabemos por el Teorema de Rouché Frobenius que $R(A) = R(A|b) = n$ , es decir que el rango de A es igual al rango de la matriz ampliada e igual al número de incógnitas.

Como el sistema tiene la misma cantidad de incógnitas que de ecuaciones, si $R(A) = n$ la matriz $A_R$ no tiene filas de ceros, por lo tanto $A_R = I_n$ , y por lo tanto $\boldsymbol{A}$ **tiene inversa**.

**2)** $\underbrace{A\ tiene\ inversa}_{hipótesis}$ entonces $\underbrace{AX = b \text{ tiene solución única}}_{tesis}$

Como $A$ tiene inversa entonces $A$ es equivalente por filas con la identidad.

Entonces $R(A) = n$ y como $(A|b)$ tiene n filas, si $A_R$ no tiene filas nulas $(A_R|b^*)$ tampoco tiene filas nulas y por lo tanto $R(A|b) = n$ .

Los rangos coinciden y coinciden con el número de incógnitas por lo tanto, por el Teorema de Rouche Frobenius, **el sistema $\boldsymbol{AX = b}$ tiene solución única.**

---

**Ejercicios:**

**10)** Calcular los siguientes determinantes:

a) $\begin{pmatrix} 3 & 2 \\ 7 & 1 \end{pmatrix}$ b) $\begin{pmatrix} 2 & 0 & 3 \\ 0 & -2 & 1 \\ 1 & 0 & 2 \end{pmatrix}$ c) $\begin{pmatrix} 2 & 0 & 3 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & -2 & 1 & 0 \\ 1 & 0 & 2 & 3 \end{pmatrix}$

**11)** Sea $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$ tal que det(A)= 5. Calcular los determinantes de las siguientes matrices, indicando las propiedades usadas:

***a)*** $\begin{pmatrix} -a & -b & -c \\ 2.d & 2.e & 2.f \\ 3.g & 3.h & 3.i \end{pmatrix}$ ***b)*** $\begin{pmatrix} d & e & f \\ a & b & c \\ g & h & i \end{pmatrix}$ ***c)*** $\begin{pmatrix} a+2g & b+2h & c+2i \\ d & e & f \\ g & h & i \end{pmatrix}$

***d)*** $3A$ ***e)*** $A^2$ ***f)*** $\frac{1}{2}(A^{-1})$

**12)** Sea $A = \begin{pmatrix} a & 0 & 0 & 0 \\ b & c & 0 & 0 \\ d & e & f & 0 \\ g & h & i & j \end{pmatrix}$ una matriz triangular inferior. Demostrar que $\det(A) = a.c.f.j$

**13)** Sean A, B, C matrices *nxn*, tales que C tiene inversa y **A = C.B.C** $^{-1}$**.**

Probar que det(A) = det(B). Fundamentar cada paso de la prueba.

**14)** Aplicación importante de la propiedad 5):

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

Si queremos calcular su determinante tendremos que calcular 4 determinantes de 3x3. Podemos reducir el problema aplicando la única operación elemental que no cambia el valor del determinante. **La matriz obtenida será distinta, será equivalente por filas con la original, pero su determinante es el mismo.**

$$A = \begin{pmatrix} 3 & 2 & 4 & 1 \\ 0 & -2 & 1 & 0 \\ 1 & 0 & 1 & 2 \\ -1 & 1 & -1 & -2 \end{pmatrix} \begin{matrix} F_1 \leftarrow (F_1 - 3F_3) \rightarrow B = \\ F_4 \leftarrow (F_4 + F_3) \end{matrix} \begin{pmatrix} 0 & 2 & -1 & -5 \\ 0 & -2 & 1 & 0 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Por las operaciones que hemos aplicado, sabemos que $\det(A) = \det(B)$ , calculamos entonces $\det(B)$ por columna 1:

$$\det(B) = 0 + 0 + \underbrace{1.(-1)^{1+3}}_{1} \underbrace{.det\begin{pmatrix} 2 & -1 & -5 \\ -2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}}_{\det(A(1/3))} + 0 =$$

$$= 1.\left[\overbrace{1.(-1)^{3+1}.det\begin{pmatrix} -1 & -5 \\ 1 & 0 \end{pmatrix} + 0 + 0}^{\det(A(1/3)) \text{ por fila 3}}\right] = 1.1.1(0+5) = 5$$

Hemos reducido el cálculo del determinante de una matriz de 4x4 al cálculo de un determinante de una matriz de 3x3.

a) Utilice las propiedades para calcular los siguientes determinantes:

$$A=\begin{pmatrix}3&0&4&0\\2&-2&0&5\\1&0&1&2\\-1&1&-1&-2\end{pmatrix}\quad B=\begin{pmatrix}-2&8&5\\1&5&1\\3&0&3\end{pmatrix}\quad C=\begin{pmatrix}6&3&4&1\\2&0&5&1\\1&2&1&2\\0&1&2&1\end{pmatrix}\quad D=\begin{pmatrix}3&1&4\\1&5&-7\\2&1&3\end{pmatrix}$$

b) Utilice las propiedades para llevar la siguiente matriz a una triangular superior y calcule su determinante: $B=\begin{pmatrix}1&2&4&1\\0&1&1&0\\1&2&1&2\\0&0&3&-2\end{pmatrix}$

En el ejercicio 3 probó que si una matriz es triangular inferior su determinante es el producto de los elementos de la diagonal. ¿Si la matriz es triangular superior vale lo mismo? Justifique.

**15)** Si A es una matriz 5x5 y el det (A) **= k**, hallar y justificar:

i) det (8.A) ii) $(6.A)^9$

**16)** Si B es una matriz *nxn* y el det B **= 10**, hallar det ($\frac{3}{4}$.B). Justificar.

**17)** Si *A*, *B*, C son matrices 5x5, det(A)=3, det(*B*)=2 y det(*C*)=6, indicar cuánto valen:

***a)*** $det\left(A.\left(\frac{1}{3}B\right).A^3(5B^{-1})\right)$

***b)*** $det\left(\left(\frac{1}{2}B\right).A^4(B.A)^{-1}\right)$

***c)*** $det\left(\left(\frac{5}{3}(B.C)\right)^3.\left((2.B)A^{-1}\right)^4(C.B.A)^{-1}\right)$

<u>Mencionar todas las propiedades usadas en cada paso. Justificar todas las respuestas.</u>

**18)** Decidir si las siguientes matrices tienen o no inversa:

$$A=\begin{pmatrix}2&-3\\1&-1\end{pmatrix}\qquad B=\begin{pmatrix}1&0&2\\0&2&1\\-3&-6&-9\end{pmatrix}\qquad C=\begin{pmatrix}3&0&0\\1&2&5\\0&1&-1\end{pmatrix}$$

**19)** Hallar los valores de *k* para que las siguientes matrices tengan inversa

$$A=\begin{pmatrix}1&k+2\\-(k-2)&1\end{pmatrix}\qquad B=\begin{pmatrix}k-5&0&0\\0&1&0\\-2&0&k+5\end{pmatrix}$$

**20)** Sean A, B matrices *nxn*. Decidir por propiedades del determinante si las siguientes afirmaciones son V o F. Justificar.

a) Si A no tiene inversa, entonces A.B no tiene inversa

b) Si A tiene inversa y B no, entonces A.B no tiene inversa.

c) Si A.B no tiene inversa, entonces ni A ni B tienen inversa.

d) Si A.B no tiene inversa, entonces al menos una de las dos, A o B, no tiene inversa.

e) Si det(A)=det(B) entonces A=B

**21)** Decidir si hay valores de *k* (y encontrarlos) para los que el siguiente sistema sea compatible determinado, justificar la respuesta.

$$\begin{cases} 3(k+5).x_1 + 9.x_2 - x_3 = b_1 \\ \qquad\qquad 6.x_2 + 12.x_3 = b_2 \\ 5.x_1 \quad + (1-k).x_2 + x_3 = b_3 \end{cases}$$

**22)** Calcular el valor de los siguientes determinantes:

a) $\begin{pmatrix} 3 & 4 \\ 0 & 2 \end{pmatrix}$ b) $\begin{pmatrix} -2 & 78 & 45 \\ 0 & 5 & -87 \\ 0 & 0 & 3 \end{pmatrix}$ c) $\begin{pmatrix} -2 & 0 & 0 \\ 230 & 5 & 0 \\ -80 & 32 & 3 \end{pmatrix}$

---

## 4. ANEXO: APLICACIONES

Un gran número de problemas se representan mediante sistemas de ecuaciones lineales como hemos visto en los ejemplos 2.6 y 2.7.

En los siguientes problemas sólo nos ocuparemos de hallar las ecuaciones que modelizan el problema, su resolución la dejamos como ejercicio:

1) Se tiene un rectángulo cuya altura mide 2cm más que su base y cuyo perímetro es igual a 24cm. Calcular las dimensiones del rectángulo.

Podemos representar gráficamente la situación:


Perímetro=2.altura+2.base

base

Llamando $a$ a la altura y $b$ a la base del rectángulo, tenemos que:

$\begin{cases} a = 2 + b \\ 2a + 2b = 24 \end{cases}$ o equivalentemente $\begin{cases} a - b = 2 \\ 2a + 2b = 24 \end{cases}$

Por lo que queda por resolver un sistema de dos ecuaciones con dos incógnitas.

2) Una editorial edita 3 calidades de libros: encuadernación rústica, con pasta dura y de lujo. Para los rústicos, la empresa gasta en promedio \$50 en papel, \$20 en ilustraciones y \$30 en las pastas, para los de pasta dura gasta \$100 en papel, \$40 en ilustraciones y \$80 en pastas y para los de lujo, \$200 en papel, \$120 en ilustraciones y \$240 en pastas. Si el presupuesto permite \$2 350 000 en papel, \$1 100 000 en ilustraciones y \$2 050 000 es pastas. ¿Cuántos libros de cada categoría pueden producirse?

Si llamamos $R$ a la cantidad de libros de encuadernación rústica, $P$ a los de pasta dura y $L$ a los de lujo, tenemos que:

$$\begin{cases} 50R + 100P + 200L = 2\,350\,000 \\ 20R + 40P + 120L = 1\,100\,000 \\ 30R + 80P + 240L = 2\,050\,000 \end{cases}$$

Luego, la respuesta se obtiene resolviendo este sistema de ecuaciones.

Este problema fue extraído del libro “Algebra lineal con aplicaciones” de George Nakos y David Joyner.

Bibliografía

- R. Espinosa Armenta, **Matemáticas discretas**, Editorial Alfaomega, Mexico, 2010
- Smith, et al , **Algebra, trigonometría y geometría analítica**, Pearson-Addison Wesley Longman, 1998
- Swokoski, Earl W. y Cole, Jeffery A., **Algebra y trigonometría con geometría analítica,** 11ma ed., Editorial Thomson, 2006

# Preliminares - Vectores

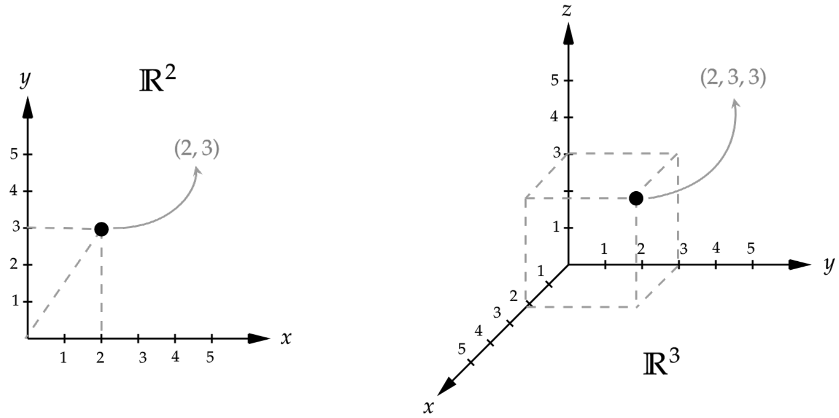
Este documento tiene el objetivo de proveer un acercamiento inicial al concepto de vector y su interpretación geométrica en el plano $\mathbb{R}^2$ y en el espacio $\mathbb{R}^3$.

## 1 Puntos y Vectores

Primeramente recordemos el concepto de **punto** en el plano cartesiano ($\mathbb{R}^2$) y en el espacio ($\mathbb{R}^3$)

**Definición 1.1.** *Definimos a un **punto** $P$ del plano $\mathbb{R}^2$ como el par ordenado de números reales $(a_1, a_2)$. Tanto a $a_1$ como a $a_2$ se los conoce como **coordenadas cartesianas de** $P$. Si consideramos a los ejes cartesianos $x$ e $y$, y a cuya intersección llamamos **origen**, podemos renombrar a $a_1$ como la **componente x** de $P$ y a $a_2$ como la **componente x** de $P$.*
*Pensando ahora en el espacio $\mathbb{R}^3$, podemos extender la definición anterior de punto, considerando que un **punto** $Q$ en el espacio es una terna ordenada $(a_1, a_2, a_3)$ de valores reales. $a_3$ representa entonces la 3ra componente o coordenada de $Q$ que se encuentra sobre un nuevo eje, al cual solemos llamar $z$.*

Gráficamente podemos ver a los puntos en $\mathbb{R}^2$ y $\mathbb{R}^3$ de la siguiente forma:



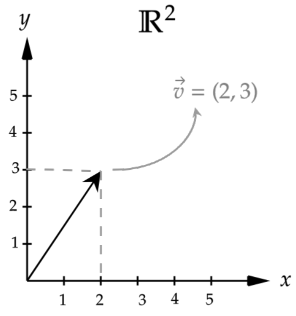
Teniendo esto en cuenta introduciremos a continuación la definición geométrica de un vector.

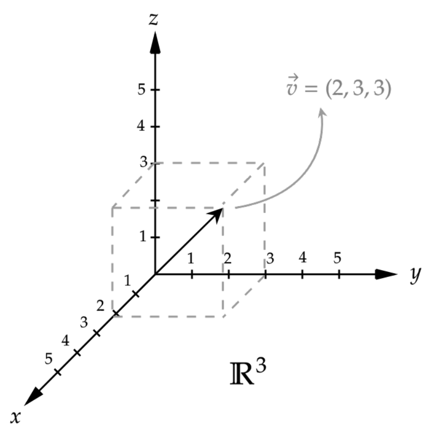
**Definición 1.2.** *Definimos a un* ***vector*** $\mathbf{v}$ *o* $\vec{v}$ *en el plano* $\mathbb{R}^2$*, como una tupla ordenada* $\mathbf{v} = (v_1, v_2)$*, donde* $v_1$ *y* $v_2$ *son numeros reales. A diferencia de los puntos, representamos graficamente a un vector mediante un segmento de recta dirigido ("flecha") que parte de un punto inicial (cola) hasta un punto final (punta). La* ***magnitud****,* ***sentido*** *y* ***dirección*** *del vector lo caracterizan y diferencian del resto. Si se elige al origen de coordenadas como punto inicial del vector, se dice que es un* ***vector en posición canónica****. Estos vectores tienen como punto inicial a* $O = (0,0)$ *y como punto terminal a* $P = (x,y)$ *y se los anota como* $\mathbf{v} = \overrightarrow{OP}$*. De esta forma existe una relación bi-unívoca entre vectores y los puntos del plano, por lo que identificaremos a cada vector con las coordenadas del punto donde termina, es decir,* $\mathbf{v} = (x,y)$

*Particularmente cuando hablamos de vectores, si* $\mathbf{v} = (a_1, a_2)$*, decimos que* $a_1$ *y* $a_2$ *son las* ***componentes*** *del vector* $\mathbf{v}$*.*

*Podemos extender la noción de vector en el plano al espacio* $\mathbb{R}^3$*, considerando una tercera componente o coordenada. Por ejemplo:* $\mathbf{v} = (a_1, a_2, a_3)$

De forma gráfica, podemos ver a dos vectores que tienen como comienzo al origen de coordenadas:



**Definición 1.3.** *Considerando los vectores que* ***no*** *parten del origen de coordenadas, es decir, que no estan en posición canónica, podemos definirlos como aquellos vectores* $\mathbf{v} = (v_1, v_2)$ *que tienen un punto inicial* $A = (a_1, a_2)$ *y uno terminal* $B = (b_1, b_2)$*, y donde* $v_1 = b_1 - a_1$ *y* $v_2 = b_2 - a_2$*. Se los anota de la forma* $\mathbf{v} = \overrightarrow{AB}$
*En el caso de considerar el espacio, tenemos que un vector es de la forma* $\mathbf{v} = (v_1, v_2, v_3)$ *y parte de un punto* $A = (a_1, a_2, a_3)$ *y otro terminal* $B = (b_1, b_2, b_3)$*. De forma análoga, también tenemos que* $v_3 = b_3 - a_3$*.*
*Notamos en estos casos que no podemos representar al vector simplemente con su punto terminal, sino que debemos proveer el inicial también. En efecto, los vectores en posición canónica son un caso especial de estos vectores "generales" o libres.*

## 2 Operaciones algebraicas con vectores

En esta sección definiremos las operaciones básicas que pueden realizarse con los vectores y sus interpretaciones geométricas en el plano (por simplicidad no contemplaremos las interpretaciones en el espacio, pero resultan análogas).

Comencemos con la relación elemental de **igualdad** entre dos vectores:

**Definición 2.1.** *Dados dos vectores* $\mathbf{u} = (u_1, u_2)$ *y* $\mathbf{v} = (v_1, v_2)$*, decimos que son* ***iguales****, y lo anotamos como* $\mathbf{u} = \mathbf{v}$*, si y sólo si* $u_1 = v_1$*,* $u_2 = v_2$*. Es decir, dos vectores son iguales si y sólo si son iguales componente a componente.*

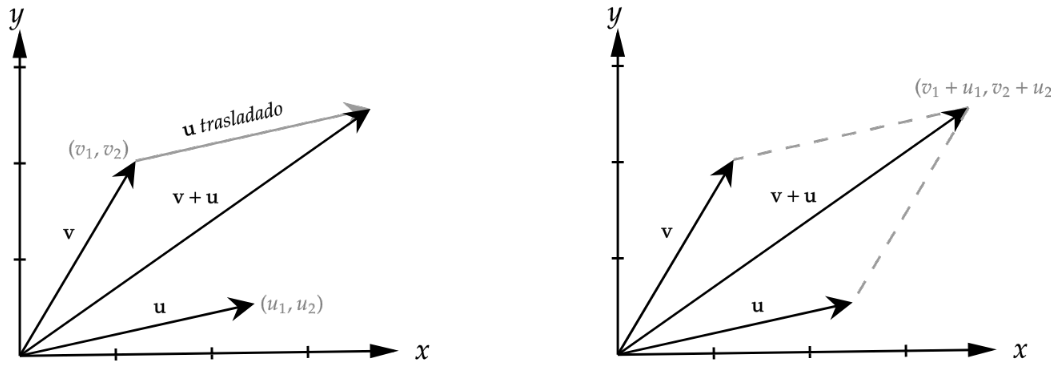
Tomando en cuenta esta relación, pasaremos a definir las distintas operaciones algebráicas básicas para vectores en el plano. Estas definiciones se extienden de forma análoga para los vectores del espacio

## Suma:

Dados dos vectores del plano $\mathbf{v} = (v_1, v_2)$ y $\mathbf{u} = (v_1, v_2)$, definimos la suma entre ellos como:

$$\mathbf{v} + \mathbf{u} = (v_1 + u_1, v_2 + u_2)$$

Gráficamente, si trasladamos el vector $\mathbf{u}$ de forma tal que su cola quede en la punta del vector $\mathbf{v}$, el vector suma es aquel con punto de inicio en la cola de $\mathbf{v}$ y punto terminal en el nuevo vector trasladado de $\mathbf{u}$:




Podemos observar, que el resultado de la suma de dos vectores es otro vector que tiene la misma magnitud que la diagonal del paralelogramo que se forma al trasladar los vectores originales.

## Producto por escalar:

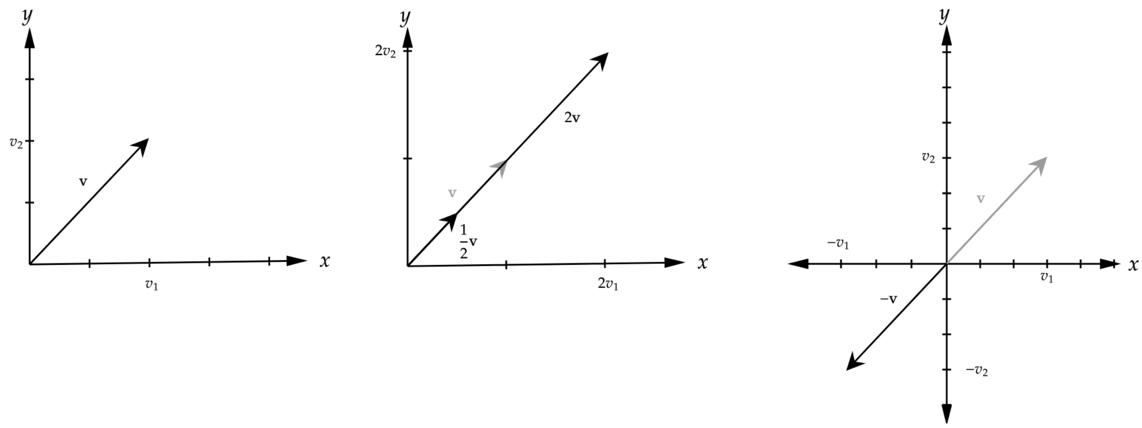
Dados un vector en el plano $\mathbf{v} = (v_1, v_2)$ y un escalar real $\alpha \in \mathbb{R}$ definimos al **producto por escalar** como:

$$\alpha.\mathbf{v} = \alpha.(v_1, v_2) = (\alpha.v_1, \alpha.v_2)$$

De forma geométrica, podemos interpretar al producto por escalar como una contracción o extensión del vector $\mathbf{v}$ en $|\alpha|$ unidades.

- Si $|\alpha| > 1$, equivale a "estirar" el vector $|\alpha|$ unidades en ambas coordenadas.
- Si $0 < |\alpha| < 1$, equivale a "contraer" el vector $|\alpha|$ unidades en ambas coordenadas.

- Si $\alpha < 0$, equivale a la extension o contraccion del vector por $|\alpha|$ unidades pero en sentido contrario, con la misma direccion.




## Resta:

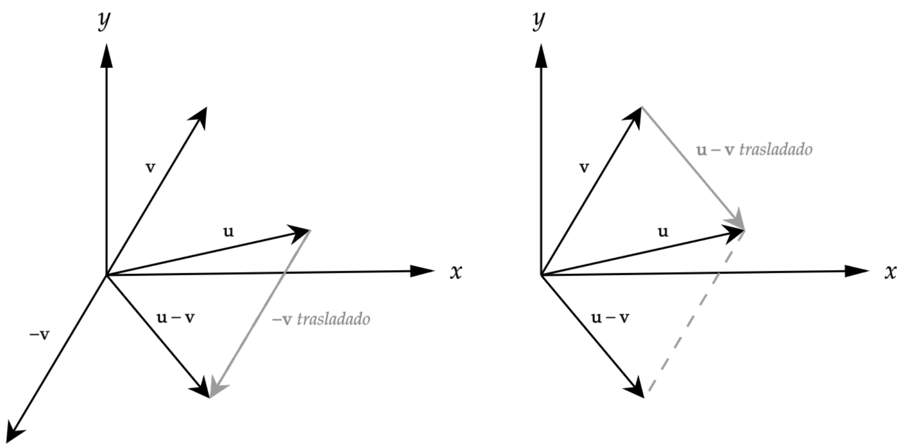
Dados dos vectores en el plano $\mathbf{v} = (v_1, v_2)$ y $\mathbf{u} = (v_1, v_2)$, definimos la resta entre ellos como:

$$\mathbf{v} - \mathbf{u} = (v_1 - u_1, v_2 - u_2)$$

De forma geométrica, podemos dar dos interpretaciones que estan relacionadas. En primer lugar, y de forma más intuitiva, podemos pensar a la resta como la suma de un vector $\mathbf{v}$ y el opuesto de $\mathbf{u}$, esto es, $-\mathbf{u}$. Como hemos definido para el producto, multiplicar un vector por el escalar $-1$ equivale a invertir el sentido del vector. Luego podemos proceder con la misma interpretacion geometrica de la suma (posicionar la cola del segundo en la punta del primero).

También podemos pensarlo de la siguiente forma: Como $\mathbf{a} + (\mathbf{b} - \mathbf{a}) = \mathbf{b}$ quiere decir que, según la interpretación geométrica de la suma de vectores, el vector resta $\mathbf{b} - \mathbf{a}$ es aquel que se le suma a $\mathbf{a}$ para llegar a $\mathbf{b}$. Finalmente, podríamos pensar a $\mathbf{a} - \mathbf{b} = -(\mathbf{b} - \mathbf{a})$, con lo que podemos obtener el vector $\mathbf{a} - \mathbf{b}$ invirtiendo el sentido de $\mathbf{b} - \mathbf{a}$

A continuación mostramos de forma geométrica la primera y segunda interpretación (a la izquierda y derecha respectivamente) de la resta de vectores.




De esta última operación podemos destacar una gran utilidad, la definición de un vector entre dos puntos dados. Como hemos interpretado geométricamente, y como podemos ver en el gráfico de la derecha, dados dos vectores $\mathbf{a}$ y $\mathbf{b}$, la resta del segundo con el primero $\mathbf{b} - \mathbf{a}$ da como resultado el vector con punto inicial en $\mathbf{a}$ y terminal en $\mathbf{b}$. Consecuentemente, podemos definir lo siguiente:

**Definición 2.2.** *Dados dos puntos en el espacio $P = (x, y, z)$ y $P' = (x', y', z')$, podemos definir al* ***vector entre los puntos*** *$P$* ***y*** *$P'$ (comienza en $P$ y termina en $P'$) como la resta de los vectores en posición canónica asociados a ambos puntos. Sean $\mathbf{a}$ y $\mathbf{a}'$ los vectores con inicio en el origen de coordenadas y punto terminal en $P$ y $P'$ respectivamente, la resta:*

$$\mathbf{a}' - \mathbf{a} = (x', y', z') - (x, y, z) = (x' - x, y' - y, z' - z)$$

*resulta ser el vector $\overrightarrow{PP'}$.*

***Nota:*** *De forma análoga ocurre para los vectores en el plano. Simplemente los puntos y vectores tienen una coordenada menos.*

# 3 Vectores base unitarios

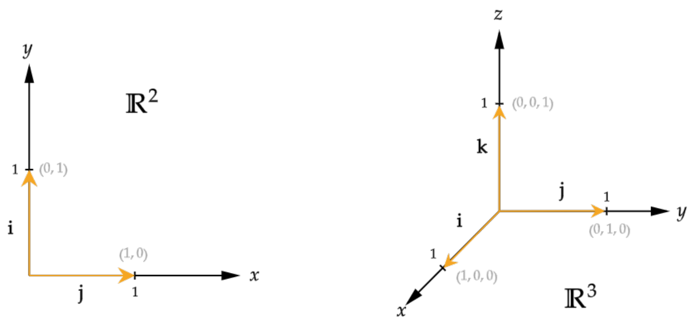
A la hora de describir los vectores del espacio muchas veces es conveniente expresarlos en términos de sumas y productos por escalar de 3 vectores base o también conocidos como **versores**: **i**, **j** y **k**. Estos 3 segmentos de recta dirigidos tienen longitud 1 y son paralelos a cada uno de los ejes, $x$, $y$ y $z$ respectivamente:

- $\mathbf{i} = (1, 0, 0)$ (Paralelo al eje x)
- $\mathbf{j} = (0, 1, 0)$ (Paralelo al eje y)
- $\mathbf{k} = (0, 0, 1)$ (Paralelo al eje z)

De forma análoga se pueden definir a estos vectores unitarios para el plano, donde solo se toman 2 ejes coordenados,y :

- $\mathbf{i} = (1, 0)$ (Paralelo al eje x)
- $\mathbf{j} = (0, 1)$ (Paralelo al eje y)

Podemos ver su representación gráfica a continuación:




Notemos que dado un vector cualquiera en posición canónica $\mathbf{a} = (a_1, a_2, a_3)$, podemos expresarlo en términos de sumas y productos por escalar de la siguiente manera:

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$$

ya que se cumple

$$\mathbf{a} = (a_1, a_2, a_3) = (a_1, 0, 0) + (0, a_2, 0) + (0, 0, a_3) = a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$$

Con esta forma de expresar a los vectores, es posible pensar a un vector cualquiera primero como extensiones o contracciones de los vectores unitarios **i**, **j** y **k** , y luego una suma de ellos.

# 4 Magnitud y Distancia

Cuando hemos nombrado a **i**, **j** y **k** como vectores **unitarios**, hemos hecho referencia su **magnitud**. Esto es la *longitud* del vector en el espacio. La semirecta que une dos puntos siempre puede pensarse como la hipotenusa de un triángulo rectángulo y, como sabemos del Teorema de Pitágoras, es posible calcular su longitud (la distancia euclidea entre los puntos) como:

$$hipotenusa = \sqrt{cateto_1^2 + cateto_2^2}$$

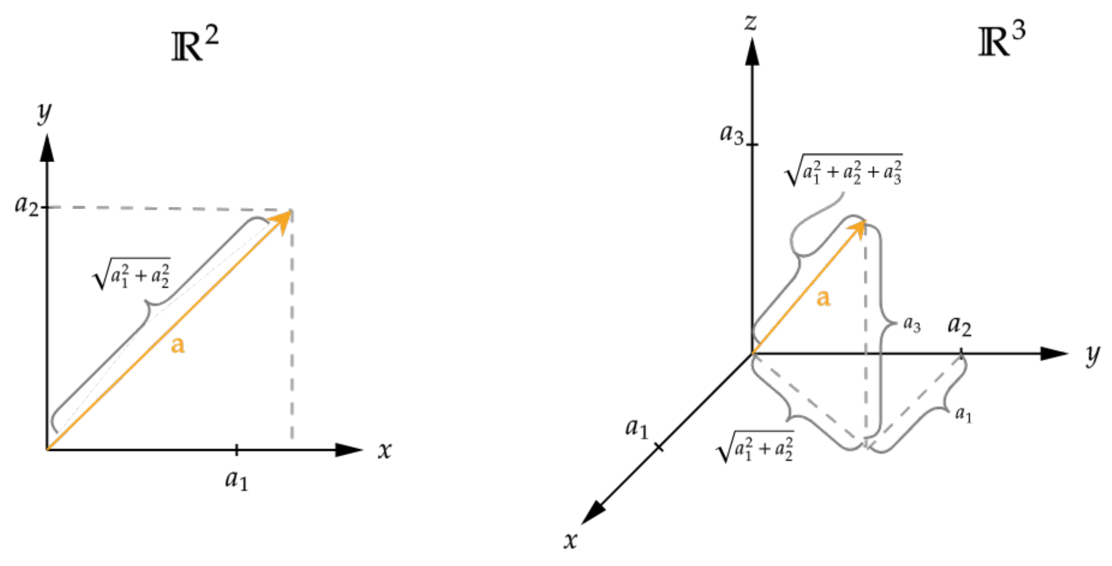
En términos de vectores:

**Definición 4.1.** *Dado un vector en el plano* $\mathbf{a} = (a_1, a_2)$*, definimos a su* ***magnitud***, ***longitud*** *o* ***norma*** *como:*

$$||\mathbf{a}|| = \sqrt{a_1^2 + a_2^2}$$

*De forma análoga, y extendiendo el Teorema de Pitágoras a 3 dimensiones, definimos la* ***magnitud*** *de un vector* $\mathbf{b} = (b_1, b_2, b_3)$ *en el espacio como:*

$$||\mathbf{b}|| = \sqrt{b_1^2 + b_2^2 + b_3^2}$$




**Definición 4.2.** *A aquellos vectores cuya magnitud es igual a 1, es decir,* $||\mathbf{a}|| = 1$*, decimos que son* ***vectores unitarios***.

A cualquier vector dado **a** podemos **normalizarlo**, esto es, obtener un vector con la misma direccion, punto de origen y sentido, pero con magnitud igual a 1. Para convertir a un vector en unitario, o normalizarlo, solo basta con dividirlo por su magnitud (aplicando la operación de producto por escalar):

$$\text{Sea } \mathbf{a} \text{ un vector cualquiera} \quad \Rightarrow \quad \frac{\mathbf{a}}{||\mathbf{a}||} \text{ es un vector unitario}$$

Veamos una demostración sencilla de esta propiedad para los vectores del plano: Sea $\mathbf{a} = (a_1, a_2)$ un vector en el plano y $||\mathbf{a}||$ su magnitud, calculemos la norma del nuevo vector

$$\frac{\mathbf{a}}{||\mathbf{a}||} = \left(\frac{a_1}{||\mathbf{a}||}, \frac{a_2}{||\mathbf{a}||}\right)$$

$$\left|\left|\frac{\mathbf{a}}{||\mathbf{a}||}\right|\right| = \sqrt{\left(\frac{a_1}{||\mathbf{a}||}\right)^2 + \left(\frac{a_2}{||\mathbf{a}||}\right)^2} = \sqrt{\frac{a_1^2}{(\sqrt{a_1^2 + a_2^2})^2} + \frac{a_2^2}{(\sqrt{a_1^2 + a_2^2})^2}} =$$

$$= \sqrt{\frac{a_1^2}{a_1^2 + a_2^2} + \frac{a_2^2}{a_1^2 + a_2^2}} = \sqrt{\frac{a_1^2 + a_2^2}{a_1^2 + a_2^2}} = \sqrt{1} = 1$$

Una vez introducida la noción de magnitud de un vector y ayudandonos con la definición de resta, podemos definir la **distancia entre los puntos terminales de dos vectores a** y **b**. Como hemos visto en la sección 2, dados dos vectores cualesquiera **a** y **b** podemos obtener un vector paralelo a aquel que une los puntos terminales de estos dos primeros. Vimos que $\mathbf{b} - \mathbf{a}$ es el vector paralelo al que tiene punto inicial en el terminal de **a** y punto terminal en el de **b**. Luego, como la magnitud no varía al trasladar un vector, podemos interpretar a $||\mathbf{b} - \mathbf{a}||$ como la distancia entre los puntos terminales de los vectores **a** y **b**

# 5 Producto interno y Ángulo entre vectores

Supongamos que queremos hallar el ángulo mínimo que se forma entre dos vectores del espacio $\mathbf{a}$ y $\mathbf{b}$ en $\mathbb{R}^3$. En esta última sección introduciremos la operación de **producto punto** o **producto escalar**, que son una forma de **producto interno**, que hará posible el cálculo del ángulo.

**Definición 5.1.** *Sean dos vectores $\mathbf{a} = (a_1, a_2, a_3)$ y $\mathbf{b} = (b_1, b_2, b_3)$ en $\mathbb{R}^3$, definimos al* ***producto escalar*** *o* ***producto punto*** *entre ellos como:*

$$\mathbf{a} \cdot \mathbf{b} = a_1.b_1 + a_2.b_2 + a_3.b_3$$

*Esta operacion de producto escalar es definida para vectores en general $\mathbf{v}$ y $\mathbf{u}$ con cualquier cantidad de componentes como sigue:*

$$\mathbf{v} \cdot \mathbf{u} = \sum_{i=1}^{n} v_i.u_i = v_1.u_1 + \cdots + v_n.u_n$$

Notar que el resultado de esta operación entre vectores da como resultado **un escalar** en $\mathbb{R}$, no un vector. Luego, destaquemos algunas propiedades de la operación que surgen de su definición:

**Propiedades 5.2.** *Sean $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ vectores en $\mathbb{R}^n$ y $\alpha, \beta$ escalares en $\mathbb{R}$, entonces se cumplen las siguientes propiedades:*

1. $\mathbf{a} \cdot \mathbf{a} \geq 0 \qquad \forall \mathbf{a} \in \mathbb{R}^n$
2. $\mathbf{a} \cdot \mathbf{a} = 0 \qquad \Longleftrightarrow \qquad \mathbf{a} = \mathbf{0} \qquad \forall \mathbf{a} \in \mathbb{R}^n$
3. $\alpha.(\mathbf{a} \cdot \mathbf{b}) = \alpha.\mathbf{a} \cdot \mathbf{b} \qquad y \qquad \beta.(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot \beta.\mathbf{b}$
4. $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = (\mathbf{a} \cdot \mathbf{b}) + (\mathbf{a} \cdot \mathbf{c}) \qquad y \qquad (\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = (\mathbf{a} \cdot \mathbf{c}) + (\mathbf{b} \cdot \mathbf{c})$
5. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$

Veamos una relación importante entre la norma de un vector y la operación de producto punto. Recordando ambas definiciones, podemos observar que dado un vector cualquiera del espacio (por simplicidad) $\mathbf{a} = (a_1, a_2, a_3)$:

$$\mathbf{a}\cdot\mathbf{a} = a_1.a_1 + a_2.a_2 + a_3.a_3 = a_1^2 + a_2^2 + a_3^3 = \left(\sqrt{a_1^2 + a_2^2 + a_3^3}\right)^2 = ||\mathbf{a}||^2$$

De esta forma llegamos a que, siempre podemos calcular la magnitud de un vector **a** en terminos del producto escalar:

$$||\mathbf{a}|| = \sqrt{\mathbf{a}\cdot\mathbf{a}}$$

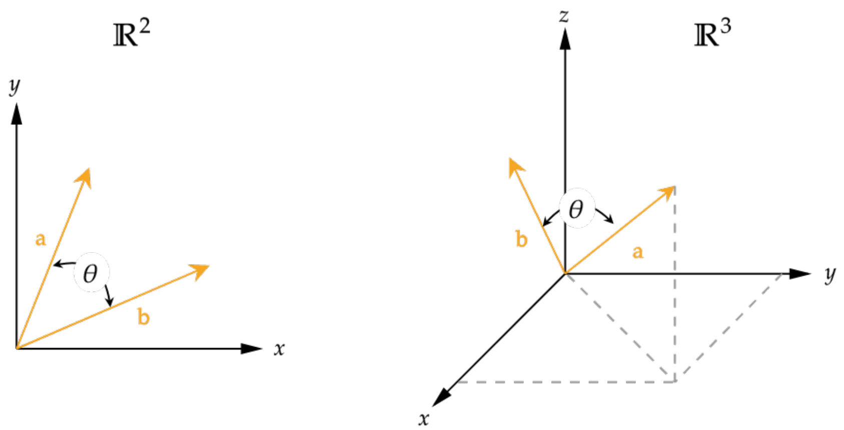
Presentemos finalmente la razón por la que el producto escalar entre vectores nos sirve para calcular el ángulo mínimo entre ellos:

**Teorema 5.3.** *Sean* **a** *y* **b** *dos vectores en el espacio* $\mathbb{R}^3$ *(también vale para cualquier dimensión), y sea* $\theta$ *el ángulo entre dichos vectores, donde* $0 \leq \theta \leq \pi$*, se cumple la siguiente igualdad:*

$$\mathbf{a}\cdot\mathbf{b} = ||\mathbf{a}||\,||\mathbf{b}||\,\cos(\theta)$$

Tomando esta igualdad, podemos derivar que el ángulo entre dos vectores tiene la forma:

$$\theta = \cos^{-1}\left(\frac{\mathbf{a}\cdot\mathbf{b}}{||\mathbf{a}||\,||\mathbf{b}||}\right)$$




Una última observación con respecto al producto escalar, es que el signo del resultado de esta operación siempre dependerá del signo del resultado del coseno del ángulo y se cumple que:

**Propiedades 5.4.** *Sean* **a**, **b** *dos vectores:*

- $\mathbf{a} \cdot \mathbf{b} > 0 \quad si \quad \theta < \frac{\pi}{2} = 90^o$
- $\mathbf{a} \cdot \mathbf{b} = 0 \quad si \quad \theta = \frac{\pi}{2} = 90^o$
- $\mathbf{a} \cdot \mathbf{b} < 0 \quad si \quad \theta > \frac{\pi}{2} = 90^o$