

Repaso de álgebra de matrices y probabilidad

Definición (matriz)

Una *matriz* de dimensión $m \times n$ es un arreglo rectangular de números reales con m filas y n columnas:

$$\mathbf{A}_{m \times n} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mj} \end{pmatrix} = (a_{ij})_{m \times n}.$$

a_{ij} es el elemento en intersección de la i -ésima fila y la j -ésima columna.

Notación

Las matrices se escribirán en mayúsculas y negritas $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$

Definición (igualdad de matrices)

Sean $\mathbf{A} = (a_{ij})_{m \times n}$ y $\mathbf{B} = (b_{ij})_{p \times q}$ dos matrices. \mathbf{A} y \mathbf{B} son iguales si tienen la misma dimensión, $m = p$ y $n = q$, y sus entradas correspondientes son iguales, es decir, $a_{ij} = b_{ij}$.

Ejemplo.

La matriz

$$\mathbf{A} = \begin{pmatrix} 2 & 4 & 5 \\ 1 & 0 & 7 \end{pmatrix}$$

es de dimensión 2×3 , $a_{12} = 4$, $a_{21} = 1$ y a_{32} no tiene sentido.

Definiciones

- Si \mathbf{A} es una matriz de dimensión $m \times n$ con $m = n$, se dice que \mathbf{A} es *cuadrada*.
- Si $\mathbf{A} = (a_{ij})_{n \times n}$ y $a_{ij} = 0$ cuando $i \neq j$, se dice que \mathbf{A} es *diagonal*.
- Si $\mathbf{A} = (a_{ij})_{n \times n}$ y $a_{ij} = 0$ cuando $i > j$ ($i < j$), se dice que \mathbf{A} es *triangular superior (inferior)*.

Preliminares (cont.)

Sean:

$$\mathbf{A} = \begin{pmatrix} 2 & 4 & 5 & 8 \\ 1 & 0 & 7 & 3 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} 2 & 4 & 1 \\ 0 & 5 & 7 \\ 0 & 0 & 8 \end{pmatrix} \quad \mathbf{D} = \begin{pmatrix} 2 & 0 & 0 \\ 4 & 5 & 0 \\ 1 & 3 & 8 \end{pmatrix}$$

- **A** no es cuadrada.
- **B** es diagonal.
- **C** es triangular superior.
- **D** es triangular inferior.

Definición (vector)

- Un *vector columna* es una matriz con una sola columna.

$$\mathbf{v}_n = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

- Un *vector fila* es una matriz con una sola fila.

$$\mathbf{v}'_n = (v_1 \quad \cdots \quad v_n).$$

v_i , $i = 1, \dots, n$ es la i -ésima entrada o componente del vector \mathbf{v}_n (\mathbf{v}'_n).

Notación

- Los vectores se escribirán en minúsculas y negritas: $\mathbf{a}, \mathbf{b}, \dots, \mathbf{u}, \mathbf{v}, \dots$
- Todos los vectores representarán vectores columna.

Notación

Sea $\mathbf{A}_{m \times n}$ una matriz.

- $\mathbf{A}_{(i)}$ representa la i -ésima fila de la matriz \mathbf{A} , $i = 1, \dots, m$.
- $\mathbf{A}^{(j)}$ representa la j -ésima columna de la matriz \mathbf{A} , $j = 1, \dots, n$.
- Todos los vectores se representan como vectores columna, es decir, como matrices de dimensión $n \times 1$.

Definición (escalar)

Una matriz con una fila y una columna es llamada *escalar*.

Definición (trasposición)

Si \mathbf{A} es una matriz de dimensión $m \times n$, la matriz traspuesta de \mathbf{A} se denota por \mathbf{A}' , tiene dimensión $n \times m$ y es tal que las filas de \mathbf{A}' son las columnas de \mathbf{A} , es decir:

$$\mathbf{A}_{m \times n} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \Rightarrow \mathbf{A}'_{n \times m} = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

La ij -ésima entrada de \mathbf{A} es la ji -ésima entrada de \mathbf{A}' .

Ejemplo

$$\mathbf{A} = \begin{pmatrix} 2 & 4 & 5 \\ 1 & 7 & 3 \end{pmatrix} \Rightarrow \mathbf{A}' = \begin{pmatrix} 2 & 1 \\ 4 & 7 \\ 5 & 3 \end{pmatrix}$$

* El traspuesto de un vector columna es un vector fila y viceversa.

Matrices particionadas

Sea la matriz

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

Algunas veces es útil particionar las matrices. Por ejemplo:

$$\mathbf{A} = \begin{pmatrix} 1 & 6 & 8 & 9 & | & 3 & 8 \\ 2 & 4 & 1 & 6 & | & 1 & 1 \\ 4 & 3 & 6 & 1 & | & 2 & 1 \\ \hline 9 & 1 & 4 & 6 & | & 8 & 7 \\ 6 & 8 & 1 & 4 & | & 3 & 2 \end{pmatrix}.$$

Matrices particionadas (cont.)

Si definimos

$$\mathbf{A}_{11} = \begin{pmatrix} 1 & 6 & 8 & 9 \\ 2 & 4 & 1 & 6 \\ 4 & 3 & 6 & 1 \end{pmatrix} \quad \mathbf{A}_{12} = \begin{pmatrix} 3 & 8 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}$$
$$\mathbf{A}_{21} = \begin{pmatrix} 9 & 1 & 4 & 6 \\ 6 & 8 & 1 & 4 \end{pmatrix} \quad \mathbf{A}_{22} = \begin{pmatrix} 8 & 7 \\ 3 & 2 \end{pmatrix}$$

podemos escribir a \mathbf{A} en forma particionada como

$$\mathbf{A} = \left(\begin{array}{cc|cc} \mathbf{A}_{11} & \mathbf{A}_{12} & & \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} & & \end{array} \right)$$

Matrices particionadas (cont.)

Definición (Matriz particionada)

Si \mathbf{A} es una matriz de dimensión $m \times n$, se puede particionar en rc submatrices, tal que el \mathbf{A}_{ij} sea de dimensión $m_i \times n_j$ para $i = 1, \dots, r$ y $j = 1, \dots, c$, donde $\sum_{i=1}^r m_i = m$ y $\sum_{j=1}^c n_j = n$. \mathbf{A} se escribe entonces como

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1c} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{r1} & \mathbf{A}_{r2} & \cdots & \mathbf{A}_{rc} \end{pmatrix}.$$

En particular, si podemos escribir a \mathbf{A} como

$$\mathbf{A} = \left(\mathbf{A}^{(1)} \mid \cdots \mid \mathbf{A}^{(n)} \right) = \left(\mathbf{A}_{(1)} \mid \cdots \mid \mathbf{A}_{(m)} \right)'$$

Ejemplo

Si definimos

$$\mathbf{A} = \left(\begin{array}{cccc|cc} 1 & 6 & 8 & 9 & 3 & 8 \\ 2 & 4 & 1 & 6 & 1 & 1 \\ 4 & 3 & 6 & 1 & 2 & 1 \\ \hline 9 & 1 & 4 & 6 & 8 & 7 \\ 6 & 8 & 1 & 4 & 3 & 2 \end{array} \right) = \left(\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array} \right)$$

es fácil verificar que

$$\mathbf{A}' = \left(\begin{array}{c|c} \mathbf{A}'_{11} & \mathbf{A}'_{21} \\ \hline \mathbf{A}'_{12} & \mathbf{A}'_{22} \end{array} \right)$$

Resultado

$$\mathbf{A} = \left(\begin{array}{c|c|c|c} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1c} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{r1} & \mathbf{A}_{r2} & \cdots & \mathbf{A}_{rc} \end{array} \right) \Rightarrow \mathbf{A}' = \left(\begin{array}{c|c|c|c} \mathbf{A}'_{11} & \mathbf{A}'_{21} & \cdots & \mathbf{A}'_{r1} \\ \mathbf{A}'_{12} & \mathbf{A}'_{22} & \cdots & \mathbf{A}'_{r2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}'_{1c} & \mathbf{A}'_{2c} & \cdots & \mathbf{A}'_{rc} \end{array} \right).$$

Operaciones básicas (cont.)

Definición (traza)

Si \mathbf{A} es una matriz de dimensión $n \times n$, la traza de \mathbf{A} se denota por $tr(\mathbf{A})$ y se define como la suma de los elementos de la diagonal de \mathbf{A} , es decir:

$$\mathbf{A}_{n \times n} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \Rightarrow tr(\mathbf{A}) = \sum_{i=1}^n a_{ii}$$

Ejemplo

$$\mathbf{A} = \begin{pmatrix} 1 & 8 & 4 \\ 7 & 3 & -2 \\ 6 & 9 & -8 \end{pmatrix} \Rightarrow tr(\mathbf{A}) = 1 + 3 - 8 = -4.$$

Resultado

Si \mathbf{A} es una matriz cuadrada, entonces $tr(\mathbf{A}) = tr(\mathbf{A}')$.

Operaciones básicas (cont.)

Definición (suma)

Si **A** y **B** son matrices de dimensión $m \times n$, entonces se define la suma de **A** y **B** como

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix} \end{aligned}$$

Ejemplo

$$\begin{pmatrix} 9 & 2 & 4 \\ 3 & 1 & 2 \end{pmatrix} + \begin{pmatrix} 5 & 1 & 4 \\ 4 & 5 & 3 \end{pmatrix} = \begin{pmatrix} 14 & 3 & 8 \\ 7 & 6 & 5 \end{pmatrix}$$

Resultados

- Si \mathbf{A} y \mathbf{B} son matrices de la misma dimensión, entonces

$$(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$$

- Si además \mathbf{A} y \mathbf{B} son cuadradas, entonces,

$$tr(\mathbf{A} + \mathbf{B}) = tr(\mathbf{A}) + tr(\mathbf{B})$$

Operaciones básicas (cont.)

Definición (producto escalar)

Si \mathbf{A} es una matriz de dimensión $m \times n$ y c es un escalar, entonces se define el producto de c y \mathbf{A} como

$$c\mathbf{A} = c \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} ca_{11} & \dots & ca_{1n} \\ \vdots & \ddots & \vdots \\ ca_{m1} & \dots & ca_{mn} \end{pmatrix}$$

Ejemplo

$$2 \begin{pmatrix} 1 & 8 & 4 \\ 7 & 3 & -2 \\ 6 & 9 & -8 \end{pmatrix} = \begin{pmatrix} 2 & 16 & 8 \\ 14 & 6 & -4 \\ 12 & 18 & -16 \end{pmatrix}$$

Definición (producto de interior)

Si \mathbf{v} y \mathbf{u} son vectores de dimensión n , se define el producto interior de \mathbf{v} y \mathbf{u} , denotado $\mathbf{v}'\mathbf{u}$, como

$$\mathbf{v}'\mathbf{u} = \sum_{i=1}^n v_i u_i$$

El producto interior de dos vectores es un escalar.

Ejemplo

$$\mathbf{v} = \begin{pmatrix} 1 \\ 7 \\ 6 \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} 8 \\ 3 \\ 2 \end{pmatrix} \Rightarrow \mathbf{v}'\mathbf{u} = (1)(8) + (7)(3) + (6)(2) = 41$$

Producto de matrices (cont.)

Definición (producto de una matriz y un vector)

Si \mathbf{A} es una matriz de dimensión $m \times n$ y \mathbf{v} es un vector columna de dimensión n , entonces se define el producto de \mathbf{A} y \mathbf{v} como

$$\mathbf{A}\mathbf{v} = \begin{pmatrix} \mathbf{A}'_{(1)} \\ \mathbf{A}'_{(2)} \\ \vdots \\ \mathbf{A}'_{(m)} \end{pmatrix} \mathbf{v} = \begin{pmatrix} \mathbf{A}'_{(1)}\mathbf{v} \\ \mathbf{A}'_{(2)}\mathbf{v} \\ \vdots \\ \mathbf{A}'_{(m)}\mathbf{v} \end{pmatrix}.$$

El producto de una matriz de dimensión $m \times n$ y un vector columna de dimensión n es un vector columna de dimensión m .

Ejemplo

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \Rightarrow \mathbf{A}\mathbf{v} = \begin{pmatrix} (1)(3) + (3)(2) \\ (2)(3) + (4)(2) \end{pmatrix} = \begin{pmatrix} 9 \\ 14 \end{pmatrix}$$

Producto de matrices (cont.)

Definición (producto de dos matrices)

Si \mathbf{A} es una matriz de dimensión $m \times n$ y \mathbf{B} es una matriz de dimensión $n \times p$, entonces se define el producto de \mathbf{A} y \mathbf{B} como

$$\begin{aligned}\mathbf{AB} &= \mathbf{A} \left(\mathbf{B}^{(1)} \mid \mathbf{B}^{(2)} \mid \dots \mid \mathbf{B}^{(p)} \right) \\ &= \left(\mathbf{AB}^{(1)} \mid \mathbf{AB}^{(2)} \mid \dots \mid \mathbf{AB}^{(p)} \right) \\ &= \begin{pmatrix} \mathbf{A}'_{(1)} \mathbf{B}^{(1)} & \mathbf{A}'_{(1)} \mathbf{B}^{(2)} & \dots & \mathbf{A}'_{(1)} \mathbf{B}^{(p)} \\ \mathbf{A}'_{(2)} \mathbf{B}^{(1)} & \mathbf{A}'_{(2)} \mathbf{B}^{(2)} & \dots & \mathbf{A}'_{(2)} \mathbf{B}^{(p)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}'_{(m)} \mathbf{B}^{(1)} & \mathbf{A}'_{(m)} \mathbf{B}^{(2)} & \dots & \mathbf{A}'_{(m)} \mathbf{B}^{(p)} \end{pmatrix}\end{aligned}$$

El producto de una matriz de dimensión $m \times n$ y una matriz de dimensión $n \times p$ es una matriz de dimensión $m \times p$.

Ejemplo

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 1 \\ 2 & 3 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 3 & 1 & 2 \\ 2 & 3 & 1 & 3 \end{pmatrix} \Rightarrow$$
$$\mathbf{AB} = \begin{pmatrix} 5 & 9 & 3 & 8 \\ 5 & 12 & 4 & 9 \\ 8 & 15 & 5 & 13 \end{pmatrix}.$$

Observación

Si \mathbf{A} es una matriz de dimensión $m \times n$ y \mathbf{B} es una matriz de dimensión $n \times p$, entonces:

$$\mathbf{AB} = \left(\mathbf{A}'_{(i)} \mathbf{B}^{(j)} \right)_{m \times p} = \left(\sum_{k=1}^n a_{ik} b_{kj} \right)_{m \times p}$$

Producto de matrices (cont.)

Los siguientes resultados serán de utilidad para hacer inferencias en el modelo de regresión lineal expresado en forma matricial.

Proposición

Si \mathbf{A} y \mathbf{B} son matrices de las dimensiones adecuadas, se cumple

- $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.
- $tr(\mathbf{AB}) = tr(\mathbf{BA})$.

Corolario

Si \mathbf{A} , \mathbf{B} y \mathbf{C} son matrices de las dimensiones adecuadas, entonces

$$tr(\mathbf{ABC}) = tr(\mathbf{BCA}) = tr(\mathbf{CAB}).$$

Este resultado generalmente se enuncia como sigue: *la traza de un producto es invariante ante permutaciones cíclicas.*

Comentarios sobre las operaciones matriciales

- La traspuesta de una matriz siempre está definida.
- La traza sólo está definida para matrices cuadradas, es decir, con igual número de filas y columnas.
- La suma de matrices está definida sólo para matrices de la misma dimensión.
- El producto dos matrices \mathbf{A} y \mathbf{B} está definido sólo si el número de columnas de \mathbf{A} es igual al número de filas de \mathbf{B} .
- La dimensión del producto \mathbf{AB} es igual al número de filas de \mathbf{A} por el número de columnas de \mathbf{B} .
- La ij -ésima entrada del producto \mathbf{AB} es el producto interior de la i -ésima fila de \mathbf{A} , $\mathbf{A}'_{(i)}$, y la j -ésima columna de \mathbf{B} , $\mathbf{B}^{(j)}$.

Vectores y matrices especiales

- *Vector de unos* de dimensión n :

$$\mathbf{1}_n = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

- *Vector de ceros* de dimensión n :

$$\mathbf{0}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

- *Matriz identidad* de dimensión n :

$$\mathbf{I}_{n \times n} = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}_{n \times n}.$$

- *Matriz de unos* de dimensión $m \times n$:

$$\mathbf{J}_{m \times n} = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}_{m \times n} .$$

- *Matriz de ceros* de dimensión $m \times n$:

$$\mathbf{0}_{m \times n} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}_{m \times n} .$$

Leyes del álgebra de matrices

La suma y producto matricial cumplen con las siguientes propiedades, en todos los casos **A**, **B**, **C**, **I** y **0** son matrices de las dimensiones adecuadas

- $\mathbf{A} + \mathbf{0} = \mathbf{A}$.
- $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$.
- $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$.
- $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$.
- $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$.
- $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$.
- $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.
- En general, $\mathbf{AB} \neq \mathbf{BA}$.
- Para factorizar expresiones es importante el orden de los productos, por ejemplo $\mathbf{XA} + \mathbf{XB} = \mathbf{X}(\mathbf{A} + \mathbf{B})$ y $\mathbf{AX} + \mathbf{BX} = (\mathbf{A} + \mathbf{B})\mathbf{X}$, pero $\mathbf{XA} + \mathbf{BX}$ y $\mathbf{AX} + \mathbf{XB}$ no son factorizables.

Definición (matriz simétrica)

Si \mathbf{A} es una matriz cuadrada y $\mathbf{A} = \mathbf{A}'$, se dice que \mathbf{A} es simétrica.

Ejemplo

$$\mathbf{A} = \begin{pmatrix} 49 & 6 & 7 \\ 6 & 20 & 17 \\ 7 & 17 & 35 \end{pmatrix}, \quad \mathbf{A}' = \begin{pmatrix} 49 & 6 & 7 \\ 6 & 20 & 17 \\ 7 & 17 & 35 \end{pmatrix} \Rightarrow$$
$$\mathbf{A} = \mathbf{A}'$$

Matrices especiales (continuación)

Definición (matriz idempotente)

Si \mathbf{A} es una matriz cuadrada y $\mathbf{AA} = \mathbf{A}$, se dice que \mathbf{A} es idempotente.

Ejemplo

$$\begin{aligned}\mathbf{A} &= \begin{pmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{pmatrix} \Rightarrow \mathbf{A}^2 = \mathbf{AA} = \\ &= \begin{pmatrix} 2(2) + 1(3) - (1)(5) & -2(3) - 3(4) + 3(5) & -2(5) - 3(5) + 5(4) \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{pmatrix} \\ &\Rightarrow \mathbf{A}^2 = \mathbf{A}\end{aligned}$$

Matrices especiales (continuación)

Definición (matriz nilpotente)

Si \mathbf{A} es una matriz cuadrada y $\mathbf{AA} = \mathbf{0}$, se dice que \mathbf{A} es nilpotente.

Ejemplo

$$\begin{aligned}\mathbf{A} &= \begin{pmatrix} 5 & -3 & 2 \\ 15 & -9 & 6 \\ 10 & -6 & 4 \end{pmatrix} \Rightarrow \mathbf{A}^2 = \mathbf{AA} = \\ &= \begin{pmatrix} 5(5) - 15(3) + 2(10) & -3(5) + 3(9) - 2(6) & 2(5) - 3(6) + 2(4) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &\Rightarrow \mathbf{A}^2 = \mathbf{0}\end{aligned}$$

Matrices especiales (continuación)

Definición (matriz ortogonal)

Si \mathbf{A} es una matriz cuadrada y $\mathbf{AA}' = \mathbf{I}$, se dice que \mathbf{A} es ortogonal.

Ejemplo

$$\begin{aligned}\mathbf{A} &= \begin{pmatrix} 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix} \Rightarrow \mathbf{AA}' = \\ &= \begin{pmatrix} \frac{2}{\sqrt{6}}(\frac{2}{\sqrt{6}}) + \frac{1}{\sqrt{3}}(\frac{1}{\sqrt{3}}) & \frac{2}{\sqrt{6}}(\frac{1}{\sqrt{6}}) - \frac{1}{\sqrt{3}}(\frac{1}{\sqrt{3}}) & -\frac{2}{\sqrt{6}}(\frac{1}{\sqrt{6}}) + \frac{1}{\sqrt{3}}(\frac{1}{\sqrt{3}}) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &\Rightarrow \mathbf{AA}' = \mathbf{I}\end{aligned}$$

Definición (permutación)

Sea $\mathcal{U}_n = \{1, \dots, n\}$. Una *permutación* de \mathcal{U}_n es una función biyectiva $\sigma : \mathcal{U}_n \rightarrow \mathcal{U}_n$. Se denotará por S_n al conjunto de todas las permutaciones de \mathcal{U}_n .

Ejemplo

Los elementos S_3 son:

$$\sigma_1 = (1, 2, 3) \quad \sigma_2 = (1, 3, 2) \quad \sigma_3 = (3, 2, 1)$$

$$\sigma_4 = (2, 1, 3) \quad \sigma_5 = (2, 3, 1) \quad \sigma_6 = (3, 1, 2)$$

Es fácil mostrar que el número de elementos de S_n es $n!$.

Definición (inversiones, paridad y signo)

Dada una permutación σ del conjunto \mathcal{U}_n , se define el número de *inversiones* de índices i como el número de índices $j > i$ tales que $\sigma(i) > \sigma(j)$. El número total de inversiones de σ se denota por $\varphi(\sigma)$ y se puede calcular como

$$\varphi(\sigma) = \sum_{i=1}^n \sum_{j>i}^n 1_{\{\sigma(i) > \sigma(j)\}}$$

Si $\varphi(\sigma)$ es impar, se dice que σ es impar, en caso contrario, es par. Por último, El signo se denota por $sign(\sigma)$ y se define como

$$sign(\sigma) = (-1)^{\varphi(\sigma)}$$

Ejemplo

Si consideramos el conjunto S_3 , podemos verificar que

$$\begin{array}{lll} \varphi(\sigma_1) = 0 & \varphi(\sigma_2) = 1 & \varphi(\sigma_3) = 3 \\ \varphi(\sigma_4) = 1 & \varphi(\sigma_5) = 2 & \varphi(\sigma_6) = 2 \end{array}$$

Entonces

$$\begin{array}{lll} \text{sign}(\sigma_1) = 1 & \text{sign}(\sigma_2) = -1 & \text{sign}(\sigma_3) = -1 \\ \text{sign}(\sigma_4) = -1 & \text{sign}(\sigma_5) = 1 & \text{sign}(\sigma_6) = 1 \end{array}$$

Determinantes (cont.)

Definición (determinante)

Si \mathbf{A} es una matriz cuadrada de dimensión n , se define el determinante de $\mathbf{A} = (a_{ij})_{n \times n}$, denotado por $|\mathbf{A}|$, como:

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

Ejemplo

$$\left| \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right| = a_{11}a_{22} - a_{21}a_{12}$$

Determinantes (cont.)

Ejemplo

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} \\ - a_{21}a_{21}a_{33} + a_{12}a_{32}a_{31} + a_{13}a_{21}a_{32}$$

Definición (menor)

Sea \mathbf{A} una matriz cuadrada de dimensión n . Se define el ij -ésimo *menor* de \mathbf{A} como la matriz que resulta de eliminar la i -ésima fila y la j -ésima columna de \mathbf{A} .

Definición (cofactor)

Sea \mathbf{A} una matriz cuadrada de dimensión n . Se define el ij -ésimo *cofactor* de \mathbf{A} como $(-1)^{i+j}$ por el determinante del ij -ésimo menor de \mathbf{A} .

Determinantes (cont.)

Resultado

Sea \mathbf{A} una matriz cuadrada de dimensión n . Para $i, j = 1, \dots, n$ se cumple

$$|\mathbf{A}| = \sum_{k=1}^n a_{ik} C_{ik} = \sum_{k=1}^n a_{kj} C_{jk}$$

donde C_{ij} son los cofactores de \mathbf{A} .

Ejemplo

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11} (a_{22}a_{33} - a_{32}a_{23}) - a_{21} (a_{12}a_{33} - a_{32}a_{13}) \\ &\quad + a_{31} (a_{12}a_{23} - a_{22}a_{13}) \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} \\ &\quad + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} \end{aligned}$$

Propiedades de los determinantes

Si \mathbf{A} es una matriz cuadrada de dimensión n y c es un escalar, entonces se cumple

- $|\mathbf{A}| = |\mathbf{A}'|$.
- Si se intercambian de lugar dos columnas o dos filas de \mathbf{A} , $|\mathbf{A}|$ no se modifica.
- Si dos columnas o dos filas de \mathbf{A} son iguales, entonces $|\mathbf{A}| = 0$.
- $|c\mathbf{A}| = c^n |\mathbf{A}|$.
- Si se multiplica por c una columna o fila de \mathbf{A} , entonces el determinante se también se multiplica por c .
- Si a una columna o fila de \mathbf{A} se le suma un múltiplo de otra columna o fila de \mathbf{A} , el determinante no se multiplica.
- Si \mathbf{B} es una matriz cuadrada de dimensión n , entonces $|\mathbf{AB}| = |\mathbf{A}| |\mathbf{B}|$.
- Si $\mathbf{A} = \text{diag}(a_1, \dots, a_n)$, entonces $|\mathbf{A}| = a_1 \dots a_n$.

Definición (matriz inversa)

Sea \mathbf{A} una matriz cuadrada de dimensión n . Si existe una matriz \mathbf{B} de dimensión n tal que

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n.$$

se dice que \mathbf{A} es invertible y que \mathbf{B} es la matriz inversa de \mathbf{A} , que se denota \mathbf{A}^{-1} .

Ejemplo

Si

$$\mathbf{A} = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}$$

es fácil verificar que

$$\mathbf{A}^{-1} = \frac{1}{2} \begin{pmatrix} 3 & -4 \\ -1 & 2 \end{pmatrix}$$

Cálculo de la inversa de una matriz

Definición (matriz adjunta)

Sea \mathbf{A} una matriz cuadrada de dimensión n , se define la matriz adjunta de \mathbf{A} , denotada por $adj(\mathbf{A})$, como

$$adj(\mathbf{A}) = (C_{ji})_{n \times n}$$

donde C_{ji} es el ji -ésimo cofactor de \mathbf{A} . Es decir, la matriz adjunta de \mathbf{A} es la traspuesta de la matriz de cofactores de \mathbf{A} .

Proposición

Si \mathbf{A} una matriz cuadrada de dimensión n y $adj(\mathbf{A})$ su matriz adjunta, entonces

$$\mathbf{A}adj(\mathbf{A}) = adj(\mathbf{A})\mathbf{A} = diag(|\mathbf{A}|, \dots, |\mathbf{A}|)$$

Cálculo de la inversa de una matriz (cont.)

Resultado

Si \mathbf{A} es cuadrada de dimensión n , entonces \mathbf{A} es invertible si y sólo si $|\mathbf{A}| \neq 0$, además

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \text{adj}(\mathbf{A}).$$

Ejemplo

Sea

$$\mathbf{A} = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}$$

Se puede verificar que $|\mathbf{A}| = 2$. Entonces, \mathbf{A} es invertible. $M_{11} = 2$, $M_{12} = 1$, $M_{21} = 4$ y $M_{22} = 3$. Luego, la matriz de cofactores de \mathbf{A} es

$$\begin{pmatrix} 2 & -1 \\ -4 & 3 \end{pmatrix}$$

Cálculo de la inversa de una matriz (cont.)

Por lo tanto,

$$\mathbf{A}^{-1} = \frac{1}{2} \begin{pmatrix} 3 & -4 \\ -1 & 2 \end{pmatrix}$$

Proposición

Sea \mathbf{A} una matriz cuadrada no singular. \mathbf{A}^{-1} cumple con las siguientes propiedades.

- a) \mathbf{A}^{-1} es única.
- b) $|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1}$.
- c) \mathbf{A}^{-1} es no singular.
- d) La inversa de \mathbf{A}^{-1} es \mathbf{A} .
- e) $(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}$.
- f) Si \mathbf{A} es simétrica, entonces \mathbf{A}^{-1} también lo es.
- g) Si \mathbf{B} es no singular, entonces $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Casos especiales

-

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

- Si $\mathbf{D} = \text{diag}(d_1, \dots, d_n)$, con $d_i \neq 0$, entonces $\mathbf{D}^{-1} = \text{diag}(d_1^{-1}, \dots, d_n^{-1})$.
- Si \mathbf{P} es una matriz ortogonal, entonces $\mathbf{P}^{-1} = \mathbf{P}'$.
- Si $a \neq 0$ y $a + bn \neq 0$, entonces

$$(a\mathbf{I}_n + b\mathbf{J}_n)^{-1} = \frac{1}{a} \left(\mathbf{I}_n - \frac{b}{a + bn} \mathbf{J}_n \right)$$

Definición (combinación lineal)

Sean $\mathbf{v}_1, \dots, \mathbf{v}_p$ vectores de la misma dimensión. Se dice que el vector \mathbf{u} es *combinación lineal* CL de $\mathbf{v}_1, \dots, \mathbf{v}_p$ si existen escalares a_1, \dots, a_p tales que

$$\mathbf{u} = a_1\mathbf{v}_1 + \dots + a_p\mathbf{v}_p.$$

Ejemplo

Si \mathbf{X} es una matriz de dimensión $m \times n$, \mathbf{a} un vector de dimensión n y \mathbf{b} un vector dimensión m , entonces \mathbf{Xa} es CL de las columnas de \mathbf{X} y $\mathbf{b}'\mathbf{X}$ es (CL) de las filas de \mathbf{X} .

Definición (independencia lineal)

Sean $\mathbf{v}_1, \dots, \mathbf{v}_p$ vectores de la misma dimensión. Se dice que el conjunto $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ es *linealmente independiente* (LI) si y sólo si

$$a_1\mathbf{v}_1 + \dots + a_p\mathbf{v}_p = \mathbf{0}$$

si y sólo si $a_1 = \dots = a_p = 0$. Un conjunto de vectores es *linealmente dependiente* (LD) si no es LI.

Si un conjunto de vectores es LD, entonces, alguno de ellos es $\mathbf{0}$ o es CL del resto de los vectores.

Proposición

Sea \mathbf{A} una matriz cuadrada de dimensión n . Si $\mathbf{Ax} = \mathbf{0}_n$, para algún vector \mathbf{x} de dimensión n distinto de $\mathbf{0}_n$, entonces $|A| = 0$.

La demostración de la proposición anterior es muy sencilla. Si $\mathbf{Ax} = \mathbf{0}_n$ para algún $\mathbf{x} \neq \mathbf{0}_n$, significa que alguna de las columnas de \mathbf{A} es igual $\mathbf{0}_n$ o bien es CL del resto de las columnas.

Proposición

Sea \mathbf{A} una matriz de dimensión $m \times n$. El máximo número de columnas LI de \mathbf{A} es igual al máximo número de filas LI de \mathbf{A} .

Definición (rango)

Sea \mathbf{A} una matriz de dimensión $m \times n$, se define el rango de \mathbf{A} como el máximo número de columnas (o filas) LI de \mathbf{A} y se denota por $r(\mathbf{A})$. Se define $r(\mathbf{0}) = 0$.

Propiedades del rango

- $r(\mathbf{A})$ es un entero positivo, excepto si $\mathbf{A} = \mathbf{0}$, en cuyo caso $r(\mathbf{A}) = 0$.
- Si \mathbf{A} es de dimensión $m \times n$, entonces $r(\mathbf{A}) \leq m, n$.
- Si \mathbf{A} es cuadrada de dimensión n y $r(\mathbf{A}) = n$, entonces \mathbf{A} es no singular y por lo tanto invertible.
- Si \mathbf{A} es cuadrada de dimensión n y $r(\mathbf{A}) < n$, entonces \mathbf{A} es singular y no invertible.
- Si \mathbf{A} es de dimensión $m \times n$ y $r(\mathbf{A}) = m < n$, se dice que \mathbf{A} es de rango completo por filas.
- Si \mathbf{A} es de dimensión $m \times n$ y $r(\mathbf{A}) = n < m$, se dice que \mathbf{A} es de rango completo por columnas.
- Si \mathbf{A} es cuadrada de dimensión n y $r(\mathbf{A}) = n$, se dice que \mathbf{A} es de rango completo.

Proposición

- Si \mathbf{A} es de rango completo por columnas, entonces $\mathbf{A}'\mathbf{A}$ es de rango completo.
- Si \mathbf{B} es de rango completo por filas, entonces $\mathbf{B}\mathbf{B}'$ es de rango completo.

Definición (matrices definidas)

Sea \mathbf{A} una matriz real simétrica de dimensión n . Si para todo \mathbf{x} de dimensión n y $\mathbf{x} \neq \mathbf{0}$:

- 1 $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$, entonces se dice que \mathbf{A} es positiva definida;
- 2 $\mathbf{x}'\mathbf{A}\mathbf{x} < 0$, entonces \mathbf{A} es negativa definida;

Si $\mathbf{x}'\mathbf{A}\mathbf{x}$ es positivo para algunos valores de \mathbf{x} y negativo para otros, se dice que \mathbf{A} es indefinida.

Proposición

Sean $f : \mathbb{R}^n \rightarrow \mathbb{R}$ y $g : \mathbb{R}^n \rightarrow \mathbb{R}$ funciones.

- Si $f(\mathbf{x}) = \mathbf{a}'\mathbf{x}$, con $\mathbf{a} \in \mathbb{R}^n$, entonces

$$\nabla f(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}} f(\mathbf{x}) = \mathbf{a}.$$

- $\nabla (f + g)(\mathbf{x}) = \nabla f(\mathbf{x}) + \nabla g(\mathbf{x})$.
- $f(\mathbf{x}) = \mathbf{x}'\mathbf{x}$, entonces

$$\nabla f(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}} f(\mathbf{x}) = 2\mathbf{x}.$$

- Si $f(\mathbf{x}) = \mathbf{a}'\mathbf{A}\mathbf{x}$, con $\mathbf{a} \in \mathbb{R}$ y \mathbf{A} una matriz cuadrada de orden n , entonces:

$$\nabla f(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}} f(\mathbf{x}) = \mathbf{A}'\mathbf{a}.$$

- $\nabla (fg)(\mathbf{x}) = f(\mathbf{x}) \nabla g(\mathbf{x}) + g(\mathbf{x}) \nabla f(\mathbf{x})$.

- Si $f(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$, entonces

$$\nabla f(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}} f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{A}'\mathbf{x}.$$

- En el resultado anterior, si \mathbf{A} es simétrica, entonces

$$\nabla f(\mathbf{x}) = \frac{\partial}{\partial \mathbf{x}} f = 2\mathbf{A}\mathbf{x}.$$

Definiciones

Sea $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ y $\mathbf{z} \in D$.

- 1 Si existe una vecindad $N_{\mathbf{z}} \subset D$ tal que $f(\mathbf{x}) \leq f(\mathbf{z})$, para todo $\mathbf{x} \in N_{\mathbf{z}}$, se dice que f tiene en \mathbf{z} un máximo local.
- 2 Si existe una vecindad $N_{\mathbf{z}} \subset D$ tal que $f(\mathbf{x}) \geq f(\mathbf{z})$, para todo $\mathbf{x} \in N_{\mathbf{z}}$, se dice que f tiene en \mathbf{z} un mínimo local.
- 3 Si $f(\mathbf{x}) \leq f(\mathbf{z})$, para todo $\mathbf{x} \in D$, se dice que f tiene un máximo absoluto en \mathbf{z} .
- 4 Si $f(\mathbf{x}) \geq f(\mathbf{z})$, para todo $\mathbf{x} \in D$, se dice que f tiene un mínimo absoluto en \mathbf{z} .

Definición (puntos críticos)

Sea $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, con derivadas parciales en $D' \subset D$ tal que

$$\nabla f(\mathbf{z}) = \mathbf{0}_n.$$

con $\mathbf{z} \in D'$. Se dice que \mathbf{z} es un punto crítico (o estacionario) de f .

Teorema (condiciones necesarias)

Sea f como en la definición anterior. Si f tiene un valor extremo en $\mathbf{z} \in D'$, entonces \mathbf{z} es un punto crítico de f .

El recíproco del teorema anterior no es verdadero, pero se pueden dar condiciones en las que en punto crítico una función tiene un valor extremo.

Definición (matriz hessiana)

Sea $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ con segundas derivadas parciales en $D' \subset D$. Se define la *matriz hessiana* de f como

$$H(\mathbf{x}) = \left(\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} \right)_{n \times n}.$$

Teorema (criterio de las segundas derivadas parciales)

Sea f como en la definición anterior y $\mathbf{z} \in D'$ un punto crítico de f . Entonces:

- 1 Si $H(\mathbf{z})$ es positiva definida, entonces f tiene un mínimo en \mathbf{z} .
- 2 Si $H(\mathbf{z})$ es negativa definida, entonces f tiene un máximo en \mathbf{z} .
- 3 Si $H(\mathbf{z})$ es indefinida, entonces f tiene un punto silla (*saddle point*) en \mathbf{z} .

Definición (Vector aleatorio)

Se dice que $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ es un vector aleatorio en \mathbb{R}^n si cada componente de este vector X_i es una variable aleatoria real

Definición (Función de distribución)

Sea $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ un vector aleatorio en \mathbb{R}^n . La función de distribución de \mathbf{X} se define como

$$F_{\mathbf{X}}(x_1, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

Función de densidad

Sea $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$ un vector aleatorio en \mathbb{R}^n .

- Si cada componente de \mathbf{X} es una v.a. discreta, se define la densidad de \mathbf{X} como la función $p : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que

$$f_{\mathbf{X}}(x_1, \dots, x_n) = P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n).$$

para cualquier $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$.

- Si cada componente de \mathbf{X} es una v. a. (absolutamente) continua, se define la densidad de \mathbf{X} como la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ tal que

$$P(\mathbf{X} \in D) = \int_D f_{\mathbf{X}}(x_1, x_2, \dots, x_n) d\mathbf{x},$$

para cualquier* conjunto $D \subset \mathbb{R}^n$.

Definición

Sea $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ un vector aleatorio en \mathbb{R}^n . Se define la esperanza de \mathbf{X} como el vector

$$E(\mathbf{X}) = (E(X_1), E(X_2), \dots, E(X_n))'$$

Se define la varianza del vector \mathbf{X} como matriz

$$\begin{aligned} V(\mathbf{X}) &= E((\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))') \\ &= \begin{pmatrix} V(X_1) & \text{Cov}(X_1, X_2) & \dots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1) & V(X_2) & \dots & \text{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_n, X_1) & \text{Cov}(X_n, X_2) & \dots & V(X_n) \end{pmatrix} \end{aligned}$$

Observación: Se puede mostrar que $V(\mathbf{X})$ es una matriz simétrica y positiva definida.

Proposición

Sea \mathbf{A} una matriz cuadrada constante de dimensión n , \mathbf{a} un vector constante de dimensión n y \mathbf{X} un vector aleatorio en \mathbb{R}^n tal que $E(\mathbf{X}) = \boldsymbol{\mu}$ y $V(\mathbf{X}) = \boldsymbol{\Sigma}$

- $E(\mathbf{a}'\mathbf{X}) = \mathbf{a}'\boldsymbol{\mu}$.
- $E(\mathbf{A}\mathbf{X}) = \mathbf{A}\boldsymbol{\mu}$.
- $V(\mathbf{a}'\mathbf{X}) = \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}$.
- $V(\mathbf{A}\mathbf{X}) = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'$.
- $E(\mathbf{X}'\mathbf{A}\mathbf{X}) = \text{tr}(\mathbf{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}$.

La distribución normal multivariada

Distribución normal univariada

Una variable aleatoria X tiene una distribución normal con media μ y varianza σ^2 , con $\mu \in \mathbb{R}$ y $\sigma^2 \in \mathbb{R}^+$, si su densidad está dada por

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (x - \mu)^2\right), \quad x \in \mathbb{R}.$$

En símbolos $X \sim N(\mu, \sigma)$.

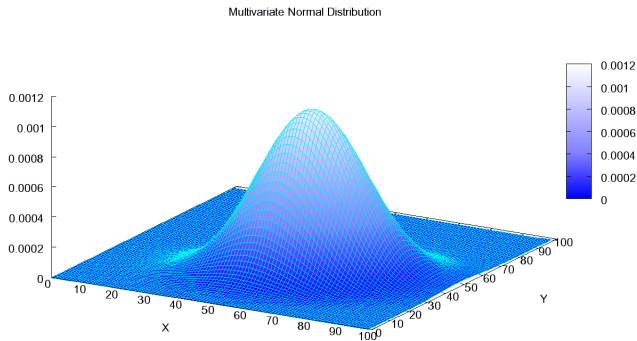
Distribución normal multivariada

Un vector aleatorio \mathbf{X} en \mathbb{R}^p tiene una distribución normal p -variada con vector de medias $\boldsymbol{\mu}$ y matriz de covarianzas $\boldsymbol{\Sigma}$, con $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma}$ simétrica y positiva definida de dimensión p , si su densidad está dada por

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{p/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right), \quad \mathbf{x} \in \mathbb{R}^p.$$

En símbolos, $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Normal multivariada y sus propiedades



Resultado

Si $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\mathbf{a} \in \mathbb{R}^p$ y \mathbf{B} es una matriz de dimensión $m \times p$ de rango completo, entonces

$$\mathbf{a} + \mathbf{B}\mathbf{X} \sim N_p(\mathbf{a} + \mathbf{B}\boldsymbol{\mu}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}').$$

Del resultado anterior se siguen las siguiente conclusiones:

- La i -ésima componente de \mathbf{X} sigue una distribución normal univariada con media μ_i y varianza σ_i^2 , donde μ_i es la i -ésima componente $\boldsymbol{\mu}$ y σ_i^2 es la i -ésima componente de la diagonal de $\boldsymbol{\Sigma}$.
- Si $\mathbf{a} \in \mathbb{R}^p$, entonces $\mathbf{a}'\mathbf{X}$ tiene distribución normal univariada con media $\mathbf{a}'\boldsymbol{\mu}$ y varianza $\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}$.
- Si \mathbf{A} es de dimensión $q \times p$ con $1 < q < p$ tal que

$$\mathbf{A} = \left(\mathbf{I}_{q \times q} \mid \mathbf{0}_{q \times (p-q)} \right),$$

entonces $\mathbf{X}_1 = \mathbf{A}\mathbf{X} \sim N_q(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$, donde

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{y} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$