Preliminares (parte II)

Universidad Internacional de Valencia

Máster Universitario en Inteligencia Artificial

02MIAR | Matemáticas:
Matemáticas para la Inteligencia Artificial

Profesor:
David Zorío Ventura

De:
Planeta Formación y Universidades

Definición

Un vector real v de dimensión n es una lista ordenada de n números reales:

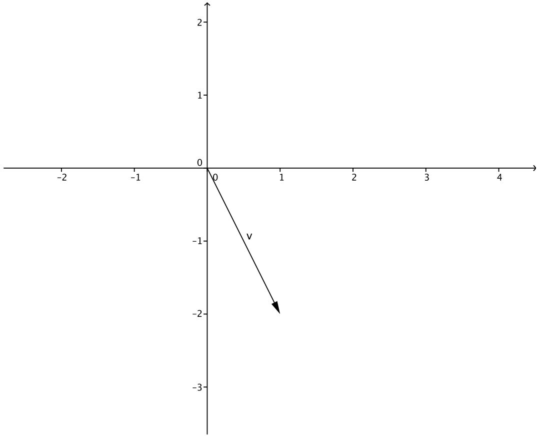
$$v = (v_1, v_2, \ldots, v_n), \quad v_i \in \mathbb{R}.$$

Se denota por $\mathbb{R}^n$ el conjunto formado por todos los vectores de dimensión n. Por tanto, podemos escribir $v \in \mathbb{R}^n$.

Ejemplos

1. $v = (1, -\frac{2}{3}) \in \mathbb{R}^2$.
2. $w = (\sqrt{2}, -1.2, \pi) \in \mathbb{R}^3$.

Los vectores en $\mathbb{R}^n$ se pueden interpretar de forma gráfica mediante un sistema de n coordenadas. En caso de tener una dimensión de $n = 3$ o inferior, éstos se pueden dibujar.



Definición

Dados dos vectores $v, w \in \mathbb{R}^n$, $v = (v_1, v_2, \ldots, v_n)$, $w = (w_1, w_2, \ldots, w_n)$, *podemos definir su suma como*

$$v + w = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n).$$

Definición

Dado un valor real $\lambda \in \mathbb{R}$ *(escalar) y un vector* $v \in \mathbb{R}^n$, $v = (v_1, v_2, \ldots, v_n)$, *podemos definir su producto como*

$$\lambda \cdot v = (\lambda v_1, \lambda v_2, \ldots, \lambda v_n).$$

La longitud de los vectores se puede medir a través de **normas**. Algunas normas vectoriales importantes son:

- Norma euclídea:

$$\|v\|_2 = \sqrt{\sum_{i=1}^{n} v_i^2} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

- Norma 1:

$$\|v\|_1 = \sum_{i=1}^{n} |v_i| = |v_1| + |v_2| + \cdots + |v_n|.$$

- Norma del máximo:

$$\|v\|_\infty = \max_{1 \leq i \leq n} |v_i|.$$

El producto escalar entre dos vectores $v, w \in \mathbb{R}^n$ se define como:

$$v \cdot w = \sum_{i=1}^{n} v_i w_i .$$

Ejemplo

Sean $v, w \in \mathbb{R}^3$*, con* $v = (2, -1, 0)$ *y* $w = (-3, -2, 1)$*. Entonces*

$$\begin{aligned} v \cdot w &= ({\color{red}2}, {\color{green}-1}, {\color{blue}0}) \cdot ({\color{red}-3}, {\color{green}-2}, {\color{blue}1}) \\ &= {\color{red}2} \cdot ({\color{red}-3}) + ({\color{green}-1}) \cdot ({\color{green}-2}) + {\color{blue}0} \cdot {\color{blue}1} = -6 + 2 + 0 = -4. \end{aligned}$$

Definición

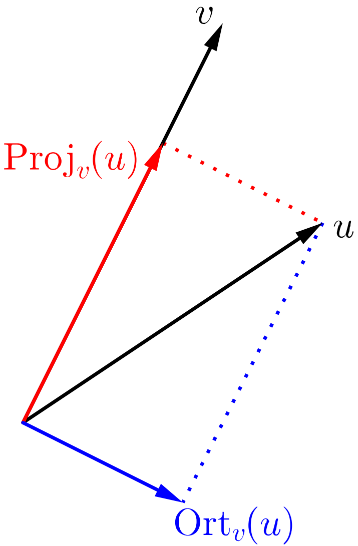
Sean $u, v \in \mathbb{R}^n$. Definimos la **proyección** *de u sobre v como*

$$\mathrm{Proj}_v(u) = \frac{u \cdot v}{v \cdot v} v.$$

Definición

Sean $u, v \in \mathbb{R}^n$. Definimos la **ortogonal** *de u sobre v como*

$$\mathrm{Ort}_v(u) = u - \mathrm{Proj}_v(u).$$



Definición

Diremos que dos vectores $u, v \in \mathbb{R}^n$ son **ortogonales** *si $u \cdot v = 0$.*

La definición de ortogonalidad entre dos vectores se corresponde con la noción geométrica de perpendicularidad.

Más concretamente,

$$u \cdot v = \|u\|_2 \|v\|_2 \cos(\widehat{uv}),$$

donde $\widehat{uv}$ es el ángulo que forman los vectores u y v.

Propiedades

Para $u, v, w \in \mathbb{R}^n$ *y* $a \in \mathbb{R}$:

- $u + v = v + u$.
- $u \cdot v = v \cdot u$.
- $u \cdot (v + w) = u \cdot v + u \cdot w$.
- $(a \cdot u) \cdot v = a \cdot (u \cdot v)$.
- $v \cdot v = \|v\|_2^2$.
- $u \cdot v = 0$ *si y sólo si* u *y* v *son ortogonales.*
- $v \cdot u = v \cdot \mathrm{Proj}_v(u)$.
- $\mathrm{Ort}_v(u) \cdot v = 0$

Universidad Internacional de Valencia

Definición

Dados $v_1, v_2, \ldots, v_k \in V$, *una* **combinación lineal** *de los* k *vectores anteriores es cualquier expresión de la forma*

$$a_1 v_1 + a_2 v_2 + \ldots + a_k v_k,$$

con $a_1, a_2, \ldots, a_k \in K$.

Definición

Sean $v_1, v_2, \ldots, v_k \in V$. *Diremos que los vectores anteriores son* **linealmente independientes** *si se cumple que la ecuación*

$$a_1 v_1 + a_2 v_2 + \ldots + a_k v_k = 0$$

tiene como única solución $a_1 = a_2 = \cdots = a_k = 0$. *En caso contrario, se dice que dichos vectores son* **linealmente dependientes**.

Definición

Una **matriz** *de tamaño $m \times n$ sobre el cuerpo de los números reales $\mathbb{R}$ es un conjunto formado por $m \cdot n$ valores reales, $a_{i,j}$, $1 \leq i \leq m$, $1 \leq j \leq n$, distribuidos en m filas y n columnas de la siguiente forma:*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}.$$

Escribimos $A \in \mathbb{R}^{m \times n}$, siendo $\mathbb{R}^{m \times n}$ el conjunto formado por todas las matrices $m \times n$. Asimismo, diremos que una matriz $A \in \mathbb{R}^{m \times n}$ es **cuadrada** *si $m = n$, es decir, si tiene el mismo número de filas y de columnas.*

Sean $A, B \in \mathbb{R}^{m\times n}$. Podemos definir la **suma** y la **resta** $A \pm B$ como:

$$A \pm B = \begin{pmatrix} a_{1,1} \pm b_{1,1} & a_{1,2} \pm b_{1,2} & \cdots & a_{1,n} \pm b_{1,n} \\ a_{2,1} \pm b_{2,1} & a_{2,2} \pm b_{2,2} & \cdots & a_{2,n} \pm b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} \pm b_{m,1} & a_{m,2} \pm b_{m,2} & \cdots & a_{m,n} \pm b_{m,n} \end{pmatrix}$$

Asimismo, dado un escalar $\lambda \in \mathbb{R}$, definimos el **producto** $\lambda \cdot A$ como

$$\lambda A = \begin{pmatrix} \lambda a_{1,1} & \lambda a_{1,2} & \cdots & \lambda a_{1,n} \\ \lambda a_{2,1} & \lambda a_{2,2} & \cdots & \lambda a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m,1} & \lambda a_{m,2} & \cdots & \lambda a_{m,n} \end{pmatrix}$$

Por último, si $A \in \mathbb{R}^{m\times s}$ y $B \in \mathbb{R}^{s\times n}$, entonces definimos el **producto matricial** $A \cdot B \in \mathbb{R}^{m\times n}$ como

$$A \cdot B = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,n} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m,1} & c_{m,2} & \cdots & c_{m,n} \end{pmatrix}$$

donde $c_{i,j} = \sum_{k=1}^{s} a_{i,k} b_{k,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots + a_{i,s}b_{s,j}$.

Nota

En general, el producto de matrices **NO** *es conmutativo. Por ejemplo:*

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

$$\begin{pmatrix} 0 & 1 \\ -3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ -7 & 8 \end{pmatrix}$$

Definición

Sea $A \in \mathbb{R}^{n\times n}$. *Diremos que* A *es* **invertible** *o* **regular** *si* $\exists B \in \mathbb{R}^{n\times n}$ *tal que* $A \cdot B = B \cdot A = I_n$. *La matriz* B *recibe el nombre de* **matriz inversa** *y se denota por* A^{-1}.

Definición

Sea $A \in \mathbb{R}^{n\times n}$ *una matriz cuadrada de la forma*

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}.$$

Definimos el **determinante** *de A de forma recurrente como sigue:*

- *Si* $n = 1$, $\det(A) = a_{1,1}$.
- *Para* $n > 1$,

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

donde $A_{i,j} \in \mathbb{R}^{(n-1)\times(n-1)}$ *es la matriz resultante de eliminar en A la fila i y la columna j.*

Teorema

El determinante de una matriz A puede calcularse de cualquiera de las siguientes formas:

- *Desarrollo por la fila i, $1 \leq i \leq n$:*

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

- *Desarrollo por la columna j, $1 \leq j \leq n$:*

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

donde $A_{i,j} \in \mathbb{R}^{(n-1)\times(n-1)}$ es la matriz resultante de eliminar en A la fila i y la columna j.

Teorema

$\forall A, B \in \mathbb{R}^{n\times n}$, se cumple

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

Teorema

$A \in \mathbb{R}^{n\times n}$ es regular si y sólo si $\det(A) \neq 0$.

Definición

*El **rango** de una matriz $A \in \mathbb{R}^{m\times n}$ es el número máximo de columnas que, como vectores, son linealmente independientes entre sí. Se denota por $\operatorname{rank}(A)$.*

Teorema

$A \in \mathbb{R}^{n\times n}$ es regular si y sólo si $\operatorname{rank}(A) = n$.

Definición

Sea $A \in \mathbb{R}^{m\times n}$*. La* **traspuesta** *de la matriz* A *se define como la matriz* $A' \in \mathbb{R}^{n\times m}$ *tal que* $A'(i,j) = A(j,i)$*,* $1 \leq i \leq n$*,* $1 \leq j \leq m$*.*

Teorema

Sea $A \in \mathbb{R}^{n\times n}$ *regular. Entonces la inversa de* A*,* A^{-1}*, viene dada por*

$$A^{-1} = \frac{1}{\det(A)} \cdot \operatorname{adj}(A),$$

donde $\operatorname{adj}(A)$ *es la* **matriz adjunta** *de* A*, dada por* $\operatorname{adj}(A) = C'$*, con* $C(i,j) = (-1)^{i+j}|A_{i,j}|$*,* $1 \leq i,j \leq n$*, con* $|A_{i,j}| = \det(A_{i,j})$*.*

Definición

Diremos que una aplicación $f : \mathbb{R}^n \to \mathbb{R}^m$ *es una* **aplicación lineal** *si* $\forall u_1, u_2 \in \mathbb{R}^n$ *y* $\forall a, b \in \mathbb{R}$ *se cumple*

$$f(au_1 + bu_2) = af(u_1) + bf(u_2).$$

Ejemplos

1. *La aplicación* $f : \mathbb{R}^3 \to \mathbb{R}^2$ *dada por* $f(x, y, z) = (x + y, 2z - x)$ *es lineal.*
2. *La aplicación* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *dada por* $f(x, y) = (x, x^2 - y)$ *no es lineal.*
3. *La aplicación* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *dada por* $f(x, y) = (x - y, x + 1)$ *no es lineal.*

Teorema

Toda aplicación lineal $f : \mathbb{R}^n \to \mathbb{R}^m$ *tiene una matriz que la representa, que además es única. Concretamente,* $\exists! M_f \in \mathbb{R}^{m\times n}$ *tal que* $\forall v \in \mathbb{R}^n$ *puede escribirse:*

$$f(v) = M_f v.$$

Ejemplo

Sea $f : \mathbb{R}^3 \to \mathbb{R}^4$ *dada por*

$$f(x,y,z) = \begin{pmatrix} x+y \\ y-z \\ 2x+y+z \\ 3z \end{pmatrix} \to M_f = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 2 & 1 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

Teorema

Sea $f : \mathbb{R}^n \to \mathbb{R}^s$ y $g : \mathbb{R}^s \to \mathbb{R}^m$ aplicaciones lineales, con M_f y M_g las matrices que las representan, respectivamente. Entonces $g \circ f : \mathbb{R}^n \to \mathbb{R}^m$ es una aplicación lineal, cuya matriz que la representa viene dada por $M_{g\circ f} = M_g M_f$.

Ejercicios

1. *Comprueba que el teorema se cumple para las dos ejemplos anteriores.*
2. *Probar el teorema para dos aplicaciones lineales cualesquiera. Es decir, probar lo siguiente:*
   2.1 *$g \circ f$ es una aplicación lineal si f y g lo son.*
   2.2 *La matriz que representa $g \circ f$, $M_{g\circ f}$, puede escribirse como el producto de las matrices que representan g y f, esto es, $M_g M_f$.*

Ejemplo

Un ejemplo clásico de aplicaciones lineales son las rotaciones. En el caso de $\mathbb{R}^2$ *vienen dadas por* $f_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ *definidas como:*

$$f_\theta(x,y) = \begin{pmatrix} \cos(\theta)x - \sin(\theta)y \\ \sin(\theta)x + \cos(\theta)y \end{pmatrix} \to M_{f_\theta} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

donde θ *representa el ángulo de rotación (en sentido antihorario). Por tanto, esta aplicación envía un vector* $(x,y) \in \mathbb{R}^2$ *a su correspondiente versión rotada* θ *grados:* $\big(\cos(\theta)x - \sin(\theta)y, \sin(\theta)x + \cos(\theta)y\big)$. *Asimismo, y como ya sabemos, podemos expresar la transformación matricialmente:*

$$f_\theta(x,y) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Ejemplo

Por ejemplo, la rotación de 30° (o $\pi/6$ rad) viene dada por:

$$f_{\pi/6} \equiv \begin{pmatrix} \cos\left(\frac{\pi}{6}\right) & -\sin\left(\frac{\pi}{6}\right) \\ \sin\left(\frac{\pi}{6}\right) & \cos\left(\frac{\pi}{6}\right) \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

Por otra parte, la rotación de 60° (o $\pi/3$ rad) viene dada por:

$$f_{\pi/3} \equiv \begin{pmatrix} \cos\left(\frac{\pi}{3}\right) & -\sin\left(\frac{\pi}{3}\right) \\ \sin\left(\frac{\pi}{3}\right) & \cos\left(\frac{\pi}{3}\right) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

Por tanto, si componemos sendas rotaciones queda:

$$f_{\pi/3} \circ f_{\pi/6} \equiv \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Ejemplo

Cabe esperar que el resultado sea una rotación de 90° *(o* $\pi/2$ *rad) con respecto al original:*

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

cumpliéndose $(x, y) \cdot (-y, x) = -xy + yx = 0$, *luego son ortogonales.*

Ejercicio

Utilizando las identidades trigonométricas que procedan, demuestra los siguientes enunciados:

(a) $\|f_\theta(v)\|_2 = \|v\|_2 \ \forall v \in \mathbb{R}^2, \forall \theta \in \mathbb{R}$.

(b) $f_\beta \circ f_\alpha = f_{\alpha+\beta} \ \forall \alpha, \beta \in \mathbb{R}$.

Funciones reales de variable real

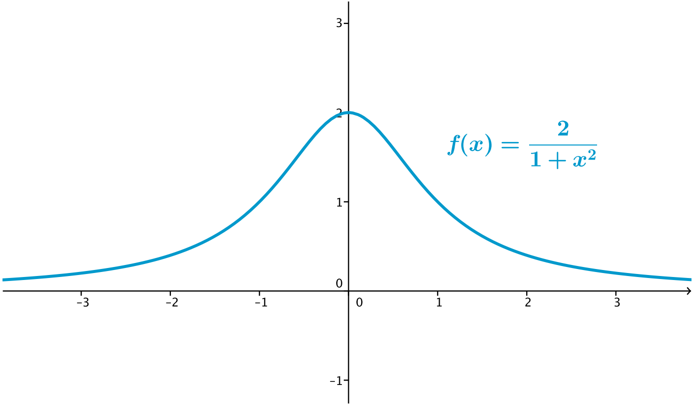
Cuando se especifique sólo la regla y no el dominio de la función, se entenderá que éste es el mayor subconjunto de los números reales posible donde dicha regla tiene sentido.

Ejemplos

1. *Tiene sentido definir la función $f(x) = x^2$ sobre todos los números reales, es decir, $f : \mathbb{R} \to \mathbb{R}$.*
2. *La función $f(x) = \sqrt{x}$ sólo tiene sentido tomarla sobre los números no negativos, luego $f : [0, +\infty) \to \mathbb{R}$.*
3. *La función $f(x) = \log(x)$ sólo tiene sentido tomarla sobre los números positivos, luego $f : (0, +\infty) \to \mathbb{R}$.*
4. *Tiene sentido definir la función $f(x) = \frac{1}{x}$ sobre todos los reales salvo el 0, luego $f : (-\infty, 0) \cup (0, +\infty) \to \mathbb{R}$.*

Funciones reales de variable real

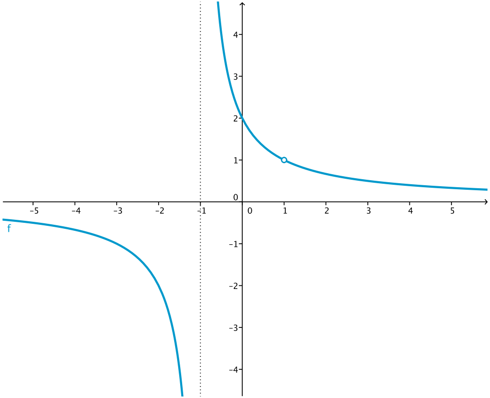
Una función real de variable real $f : A \to \mathbb{R}$ puede representarse gráficamente en dos dimensiones, considerando como coordenada X (abscisas) los diferentes valores del dominio y como coordenada Y (ordenadas) los valores de f asociados: $\{(x, f(x)) \in \mathbb{R}^2 : x \in A\}$.



Consideremos la función siguiente:

$$f(x) = \frac{2x - 2}{x^2 - 1}.$$

¿Qué ocurre en $x = -1$? ¿Y en $x = 1$?



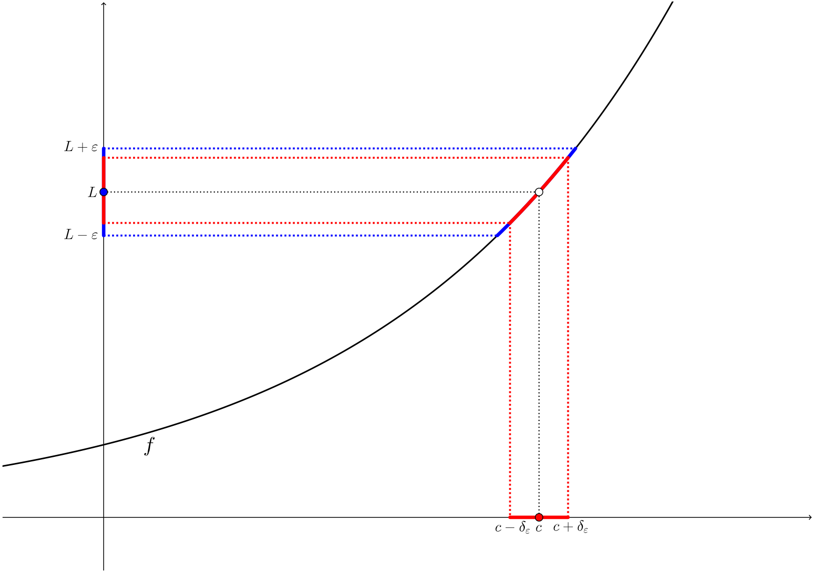
Definición

Sea $f : A \to \mathbb{R}$ *y* $c \in A$*, de forma que* $(c-\rho, c+\rho) \subseteq A$ *para algún* $\rho > 0$*. Diremos que existe el* **límite** $L \in \mathbb{R}$ *de* f *cuando* x *tiende a* c*, que denotaremos por* $\lim\limits_{x \to c} f(x) = L$*, si*

$$\forall \varepsilon > 0, \exists \delta_\varepsilon > 0 : 0 < |x-c| < \delta_\varepsilon \to |f(x) - L| < \varepsilon.$$

Intuitivamente, lo que indica la definición anterior es el siguiente enunciado: *dado un margen de error cualquiera, podemos encontrar un intervalo lo suficientemente pequeño centrado en* c *(asociado al margen de error proporcionado), de forma que la distancia entre* $f(x)$ *y* L *está siempre por debajo de ese margen de error dado en el intervalo considerado.*

Funciones reales de variable real



Definición

- *Diremos que* $\lim\limits_{x\to c^{\pm}} f(x) = L$ *si*
$$\forall\varepsilon > 0, \exists\delta_\varepsilon > 0 : 0 < \pm(x-c) < \delta_\varepsilon \to |f(x) - L| < \varepsilon.$$
- *Diremos que* $\lim\limits_{x\to c^{\pm}} f(x) = \pm\infty$ *si*
$$\forall M > 0, \exists\delta_M > 0 : 0 < \pm(x-c) < \delta_M \to \pm f(x) > \pm M.$$
- *Diremos que* $\lim\limits_{x\to c} f(x) = \infty$ *si*
$$\forall M > 0, \exists\delta_M > 0 : 0 < |x-c| < \delta_M \to |f(x)| > M.$$
- *Diremos que* $\lim\limits_{x\to \pm\infty} f(x) = L$ *si*
$$\forall\varepsilon > 0, \exists K_\varepsilon > 0 : \pm x > \pm K_\varepsilon \to |f(x) - L| < \varepsilon.$$
- *Diremos que* $\lim\limits_{x\to \pm\infty} f(x) = \pm\infty$ *si*
$$\forall M > 0, \exists K_M > 0 : \pm x > \pm K_M \to \pm f(x) > \pm M.$$

Universidad Internacional de Valencia

Teorema

Se cumple:

$$\exists \lim_{x\to c} f(x) = L \leftrightarrow \exists \lim_{x\to c^-} f(x) = L \wedge \exists \lim_{x\to c^+} f(x) = L.$$

Propiedades

Si $\exists \lim_{x\to c} f(x) = L_1$ *y* $\exists \lim_{x\to c} g(x) = L_2$, *entonces:*

- $\exists \lim_{x\to c} \left(f(x) \pm g(x)\right) = L_1 \pm L_2.$
- $\exists \lim_{x\to c} \left(f(x)g(x)\right) = L_1 L_2.$
- *Si* $L_2 \neq 0$, $\exists \lim_{x\to c} \dfrac{f(x)}{g(x)} = \dfrac{L_1}{L_2}.$
- *Si* $\exists \lim_{x\to L_1} g(x) = L_2$, *entonces* $\exists \lim_{x\to c} g(f(x)) = L_2.$

Definición

Diremos que una función f *es* **continua en un punto** c *si*

$$\exists \lim_{x\to c} f(x) = f(c).$$

En caso de que la continuidad se dé en todos los puntos de su dominio, diremos además que f *es* **continua**.

Ejemplos

1. *Todas las funciones polinómicas son continuas.*
2. $f(x) = |x|$ *es continua.*
3. $f(x) = \log(x)$ *es continua en todo su dominio.*
4. $f(x) = \sqrt{x}$ *es continua en todo su dominio.*

Ejemplos

5. *Sea f dada por*

$$f(x) = \begin{cases} x^2 + 1 & x < 0, \\ x - 1 & x \geq 0. \end{cases}$$

Entonces

$$\lim_{x \to 0^-} f(x) = \lim_{x \to 0^-} (x^2 + 1) = 1 \neq \lim_{x \to 0^+} f(x) = \lim_{x \to 0^+} (x - 1) = -1,$$

luego $\nexists \lim_{x \to 0} f(x)$, *por lo que f no es continua en* 0.

Ejemplos

6. *Sea f dada por*

$$f(x) = \begin{cases} x^3 + 2 & x < 1, \\ 4 & x = 1, \\ 5x - 2 & x > 1. \end{cases}$$

Entonces

$$\lim_{x \to 1^-} f(x) = \lim_{x \to 1^-} (x^3 + 2) = 3 = \lim_{x \to 1^+} f(x) = \lim_{x \to 1^+} (5x - 2) = 3.$$

Luego $\exists \lim_{x \to 1} f(x) = 3$. *Sin embargo,* $f(1) = 4 \neq 3$, *por lo que f no es continua en* 1.

Ejemplos

7. *Sea f dada por*

$$f(x) = \begin{cases} \frac{2x-2}{x^2-1} & x \notin \{-1, 1\}, \\ 0 & x = -1, \\ 1 & x = 1. \end{cases}$$

Entonces

- $\lim\limits_{x \to 1} f(x) = \lim\limits_{x \to 1} \dfrac{2x-2}{x^2-1} = 1 = f(1)$, *luego f es continua en 1.*
- $\lim\limits_{x \to -1^+} f(x) = \lim\limits_{x \to -1^+} \dfrac{2x-2}{x^2-1} = +\infty \neq 0 = f(-1)$, *luego f no es continua en -1.*

Propiedades

Sean f y g funciones continuas en $c \in \mathbb{R}$. Entonces:

- $f \pm g$ *es continua en* c.
- fg *es continua en* c.
- *Si* $g(c) \neq 0$, *entonces* $\dfrac{f}{g}$ *es continua en* c.
- $g \circ f$ *es continua en* c.

Teorema (Teorema de Bolzano)

Sea $f : A \to \mathbb{R}$ *continua y* $a, b \in \mathbb{R}$, *con* $a < b$ *y* $[a, b] \subseteq A$. *Si* $f(a)f(b) < 0$, *entonces* $\exists c \in (a, b)$ *tal que* $f(c) = 0$.

Ejemplo

Probemos que la ecuación $\sin(x) = \cos(x)$ *tiene una solución entre* 0 *y* $\frac{\pi}{2}$. *En efecto, si definimos* $f(x) = \sin(x) - \cos(x)$, *se tiene por una parte que* $f(0) = -1$ *y, por otra parte,* $f\left(\frac{\pi}{2}\right) = 1$, *luego* $f(0)f(\frac{\pi}{2}) < 0$, *por lo que* $\exists c \in (0, \frac{\pi}{2})$ *tal que* $f(c) = 0$, *es decir:*

$$f(c) = 0 \leftrightarrow \sin(c) - \cos(c) = 0 \leftrightarrow \sin(c) = \cos(c).$$

Definición

Sea $f : A \subseteq \mathbb{R} \to \mathbb{R}$*. Diremos que* f *es* **derivable en un punto** $c \in \mathrm{int}(A)$ *(*$\exists \rho > 0$*:* $(c-\rho, c+\rho) \subseteq A$*) si*

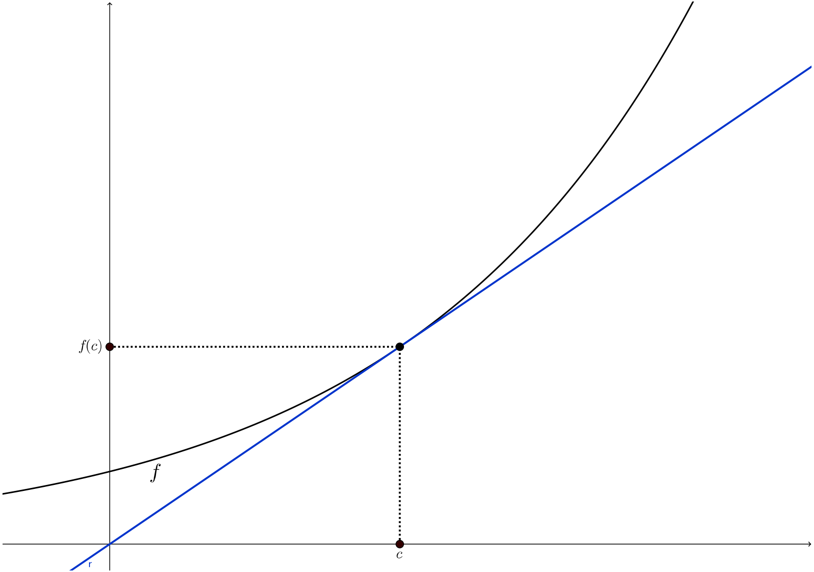
$$\exists \lim_{h \to 0} \frac{f(c+h) - f(c)}{h}.$$

En tal caso, llamaremos **derivada de** f **en** c *al valor anterior y lo denotaremos por*

$$f'(c) = \frac{df}{dx}(c) = \lim_{h \to 0} \frac{f(c+h) - f(c)}{h}.$$

Además, si f *es derivable en* x*,* $\forall x \in \mathrm{int}(A)$*, diremos que* f *es* **derivable***.*

Funciones reales de variable real



Funciones reales de variable real

Universidad Internacional de Valencia

Algunas derivadas de funciones elementales:

$f(x)$	$f'(x)$
x^r	rx^{r-1}
$\sum_{k=0}^{n} a_k x^k$	$\sum_{k=1}^{n} k a_k x^{k-1}$
$\sin(x)$	$\cos(x)$
$\cos(x)$	$-\sin(x)$
$\tan(x)$	$1+\tan^2(x)$
e^x	e^x
a^x	$\ln(a)a^x$

$f(x)$	$f'(x)$
$\ln(x)$	$\frac{1}{x}$
$\log_a(x)$	$\frac{1}{\ln(a)x}$
$\arcsin(x)$	$\frac{1}{\sqrt{1-x^2}}$
$\arccos(x)$	$-\frac{1}{\sqrt{1-x^2}}$
$\arctan(x)$	$\frac{1}{1+x^2}$

Universidad Internacional de Valencia

Ejemplos

1. *Sea* $f(x) = 2x^5 + 4x^3 - x^2 + x - 7$. *Entonces* $f'(x) = 10x^4 + 12x^2 - 2x + 1$.
2. *Sea* $f(x) = \sqrt{x}$. *Podemos escribir* $f(x) = x^{\frac{1}{2}}$. *Por tanto:*
$$f'(x) = \frac{1}{2}x^{\frac{1}{2}-1} = \frac{1}{2}x^{-\frac{1}{2}} = \frac{1}{2}\frac{1}{x^{\frac{1}{2}}} = \frac{1}{2\sqrt{x}}.$$
3. *Sea* $f(x) = \sqrt[3]{x^4}$. *Podemos escribir* $f(x) = x^{\frac{4}{3}}$, *luego:*
$$f'(x) = \frac{4}{3}x^{\frac{4}{3}-1} = \frac{4}{3}x^{\frac{1}{3}} = \frac{4}{3}\sqrt[3]{x}.$$
4. *Sea* $f(x) = \dfrac{1}{x}$. *Entonces podemos escribir* $f(x) = x^{-1}$, *por lo que:*
$$f'(x) = (-1)\cdot x^{-1-1} = -x^{-2} = -\frac{1}{x^2}.$$

Propiedades

Sean f, g funciones derivables en x. Entonces:

1. $[f(x) \pm g(x)]' = f'(x) \pm g'(x)$.
2. $[f(x)g(x)]' = f'(x)g(x) + f(x)g'(x)$ *(**regla del producto**).*
3. *En particular, si $a \in \mathbb{R}$, $[af(x)]' = af'(x)$.*
4. $[g(f(x))]' = g'(f(x))f'(x)$ *(**regla de la cadena**).*
5. *En particular, si $g(x) \neq 0$:* $\left[\dfrac{1}{g(x)}\right]' = -\dfrac{g'(x)}{g(x)^2}$
6. *En particular, combinando 2 y 5, si $g(x) \neq 0$ entonces* $\left[\dfrac{f(x)}{g(x)}\right]' = \dfrac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$.

Ejemplos

Calculamos $f'(x)$ para los siguientes casos:

1. $f(x) = x^2 \ln(x)$.
$$f'(x) = 2x\ln(x) + x.$$
2. $f(x) = \sin(\cos(x))$.
$$f'(x) = -\cos(\cos(x))\sin(x).$$
3. $f(x) = \sin(e^{\cos(x)})$.
$$f'(x) = -\cos(e^{\cos(x)})e^{\cos(x)}\sin(x).$$
4. $f(x) = x\sin(x)\ln(x^2)$.
$$f'(x) = \sin(x)\ln(x^2) + x\cos(x)\ln(x^2) + 2\sin(x).$$

De la regla de la cadena se deducen las siguientes fórmulas generalizadas:

$F(x)$	$F'(x)$
$f(x)^r$	$rf(x)^{r-1}f'(x)$
$\sin(f(x))$	$\cos(f(x))f'(x)$
$\cos(f(x))$	$-\sin(f(x))f'(x)$
$\tan(f(x))$	$[1+\tan^2(f(x))]f'(x)$
$e^{f(x)}$	$e^{f(x)}f'(x)$
$a^{f(x)}$	$\ln(a)a^{f(x)}f'(x)$
$\ln(f(x))$	$\dfrac{f'(x)}{f(x)}$

$F(x)$	$F'(x)$
$\log_a(f(x))$	$\dfrac{f'(x)}{\ln(a)f(x)}$
$\arcsin(f(x))$	$\dfrac{f'(x)}{\sqrt{1-f(x)^2}}$
$\arccos(f(x))$	$-\dfrac{f'(x)}{\sqrt{1-f(x)^2}}$
$\arctan(f(x))$	$\dfrac{f'(x)}{1+f(x)^2}$

Teorema (Teorema de Rolle)

Sea $f : A \to \mathbb{R}$, $a, b \in \mathbb{R}$, $a < b$, *tales que* $[a, b] \subseteq A$, *con* f *derivable en* (a, b). *Si* $f(a) = f(b)$, *entonces* $\exists c \in (a, b)$ *tal que* $f'(c) = 0$.

Teorema (Regla de l'Hôpital)

Sean f, g *funciones reales de variable real, cumpliendo* $\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0$ *o* $\pm\infty$. *Si* $\exists \rho > 0$ *tal que* $\forall x \in (c - \rho, c) \cup (c, c + \rho)$ f *y* g *son derivables, con* $g(x) \neq 0$, *y* $\exists \lim\limits_{x\to c} \dfrac{f'(x)}{g'(x)}$, *entonces*

$$\exists \lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}.$$

Ejemplo

Obtengamos $\lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^x$. *Para ello, definimos* $L=\lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^x$ *y, aprovechando la continuidad de* f*:*

$$\begin{aligned}
\ln(L) &= \ln\left[\lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^x\right] = \lim_{x\to+\infty}\ln\left[\left(1+\frac{1}{x}\right)^x\right] \\
&= \lim_{x\to+\infty} x\ln\left(1+\frac{1}{x}\right) = \lim_{x\to+\infty}\frac{\ln\left(1+\frac{1}{x}\right)}{\frac{1}{x}} \\
&= \lim_{x\to+\infty}\frac{\frac{-\frac{1}{x^2}}{1+\frac{1}{x}}}{-\frac{1}{x^2}} = \lim_{x\to+\infty}\frac{1}{1+\frac{1}{x}} = 1.
\end{aligned}$$

Por tanto, $\ln(L)=1$, *luego* $L=e^1=e$.

Definición

Diremos que f es derivable n veces en c si $\exists f^{(n)}(c)$*, donde* $f^{(0)}(c) = f(c)$ *y* $f^{(k)}(c) = (f^{(k-1)})'(c)$, $1 \leq k \leq n$.

Si una función es derivable n veces en todo su dominio, diremos que f es n veces derivable, y llamaremos a $f^{(n)}$ **derivada de orden n de f**.

Ejemplos

1. *Sea* $f(x) = \sin(x)$*. Calculemos* $f''(x)$. $f'(x) = \cos(x) \rightarrow f''(x) = -\sin(x)$.
2. *Sea* $f(x) = x^3 - x + 2$*. Obtengamos* $f'''(x)$. $f'(x) = 3x^2 - 1 \rightarrow f''(x) = 6x \rightarrow f'''(x) = 6$.

Universidad Internacional de Valencia

Definición

Sea $f : A \subseteq \mathbb{R} \to \mathbb{R}$ *y* $F : A \subseteq \mathbb{R} \to \mathbb{R}$. *Diremos que* F *es una* **primitiva** *de* f *si* $F' = f$.

Ejemplos

1. *Una primitiva de* $6x^2$ *es* $2x^3$.
2. *Una primitiva de* e^{2x+1} *es* $\dfrac{1}{2}e^{2x+1}$.
3. *Otra primitiva de* $6x^2$ *es* $2x^3 + 1$.
4. *Otra primitiva de* e^{2x+1} *es* $\dfrac{1}{2}e^{2x+1} - 7$.
5. $2x^3 + C$ *son primitivas de* $6x^2$, $\forall C \in \mathbb{R}$.
6. $\dfrac{1}{2}e^{2x+1} + C$ *son primitivas de* e^{2x+1}, $\forall C \in \mathbb{R}$.

Teorema

Si f es continua y F_1, F_2 son primitivas de f, entonces $\exists C \in \mathbb{R}$ tal que $F_2 - F_1 = C$.

Esto motiva la siguiente definición:

Definición

Definimos

$$\int f(x)dx$$

como la familia formada por todas las primitivas de f.

Por el teorema anterior, bastará con encontrar una única primitiva F cumpliendo $F' = f$, puesto que el resto de primitivas posibles vendrá unívocamente determinado por las funciones de la forma $F + C$, con $C \in \mathbb{R}$.

Para calcular primitivas, basta con fijarse en las tablas de las derivadas en sentido inverso:

$f(x)$	$\int f(x)\mathrm{d}x$
x^r	$\frac{1}{r+1}x^{r+1} + C$
$\cos(x)$	$\sin(x) + C$
$\sin(x)$	$-\cos(x) + C$
$1 + \tan^2(x)$	$\tan(x) + C$
e^x	$e^x + C$
a^x	$\frac{1}{\ln(a)}a^x + C$

$f(x)$	$\int f(x)\mathrm{d}x$
$\dfrac{1}{x}$	$\ln(x) + C$
$\dfrac{1}{\sqrt{1-x^2}}$	$\arcsin(x) + C$
$-\dfrac{1}{\sqrt{1-x^2}}$	$\arccos(x) + C$
$\dfrac{1}{1+x^2}$	$\arctan(x) + C$

Funciones reales de variable real

$f(x)$	$\int f(x)\mathrm{d}x$
$g(x)^r g'(x)$	$\frac{1}{r+1} g(x)^{r+1} + C$
$\cos(g(x))g'(x)$	$\sin(g(x)) + C$
$\sin(g(x))g'(x)$	$-\cos(g(x)) + C$
$[1 + \tan^2(g(x))]g'(x)$	$\tan(g(x)) + C$
$e^{g(x)} g'(x)$	$e^{g(x)} + C$
$a^{g(x)} g'(x)$	$\frac{1}{\ln(a)} a^{g(x)} + C$

$f(x)$	$\int f(x)\mathrm{d}x$
$\frac{g'(x)}{g(x)}$	$\ln(g(x)) + C$
$\frac{g'(x)}{\sqrt{1 - g(x)^2}}$	$\arcsin(g(x)) + C$
$-\frac{g'(x)}{\sqrt{1 - g(x)^2}}$	$\arccos(g(x)) + C$
$\frac{g'(x)}{1 + g(x)^2}$	$\arctan(g(x)) + C$

Ejemplos

1. $$\int 2x\cos(x^2)\mathrm{d}x = \sin(x^2) + C.$$
2. $$\int x\sqrt{x^2+1}\mathrm{d}x = \frac{1}{2}\int 2x(x^2+1)^{\frac{1}{2}}\mathrm{d}x = \frac{1}{2}\cdot\frac{2}{3}(x^2+1)^{\frac{3}{2}} + C = \frac{1}{3}\sqrt{(x^2+1)^3} + C.$$
3. $$\int \tan(x)\mathrm{d}x = \int \frac{\sin(x)}{\cos(x)}\mathrm{d}x = -\int \frac{-\sin(x)}{\cos(x)}\mathrm{d}x = -\ln\big(\cos(x)\big) + C.$$
4. $$\int \frac{1}{x^2-4x+5}\mathrm{d}x = \int \frac{1}{1+(x-2)^2}\mathrm{d}x = \arctan(x-2) + C.$$

- Algunas integrales no pueden realizarse de forma inmediata a través de los métodos elementales, como por ejemplo

$$\int x \sin(x) \mathrm{d}x.$$

- Un método para calcular algunas primitivas como la anterior es el conocido como **integración por partes**, que se basa en la fórmula de la derivada de un producto:

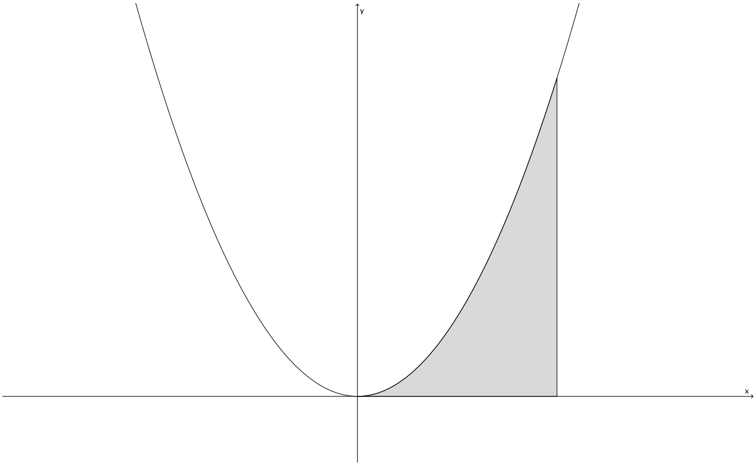
$$(uv)' = u'v + uv'.$$

  Tomando primitivas:

$$uv = \int u'v + \int uv' \rightarrow \int uv' = uv - \int vu'.$$

Funciones reales de variable real

Nuestro objetivo es aproximar el área que encierra una función f en un determinado intervalo $[a, b]$:



- Una forma de obtener una solución aproximada es considerar una **partición** del intervalo $[a, b]$ en n subintervalos con extremos $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b\}$ y considerar en cada caso el área de rectángulo con base en el subintervalo $[x_{i-1}, x_i]$ y altura $f(\xi_i)$, para algún conjunto de puntos intermedios $\mathcal{R} = \{\xi_i \in [x_{i-1}, x_i], 1 \leq i \leq n\}$, cuya área será por tanto $(x_i - x_{i-1})f(\xi_i)$, $1 \leq i \leq n$.
- Por tanto, una aproximación del área que encierra f en $[a, b]$ asociada a dicha partición consiste en sumar las áreas de dichos rectángulos:

$$\sigma_f(\mathcal{P}, \mathcal{R}) = \sum_{i=1}^{n} (x_i - x_{i-1})f(\xi_i).$$

- En los casos en los que f tenga un comportamiento "razonable" en $[a, b]$, cabe esperar que al aumentar n (el número de intervalos) la expresión anterior se irá aproximando cada vez más al área que se desea calcular.

Definición

Sea $f : A \subseteq \mathbb{R} \to \mathbb{R}$ *y* $a, b \in \mathbb{R}$ *tal que* $a < b$ *y* $[a, b] \subseteq A$*. Sea* $\mathcal{P} = \{a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b\}$ *una partición asociada a dicho intervalo. Llamamos* **suma inferior de Riemann** *asociada a dicha partición al resultado de calcular la siguiente suma de áreas de rectángulos:*

$$\sigma_f(\mathcal{P}) = \sum_{i=1}^{n} (x_i - x_{i-1}) \inf_{x \in [x_{i-1}, x_i]} f(x).$$

Definición

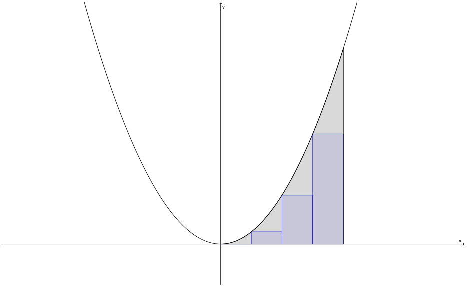
En las condiciones anteriores, llamamos **suma superior de Riemann** *al resultado de calcular la siguiente suma de áreas de rectángulos:*

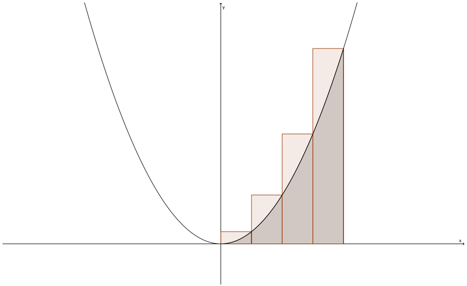
$$\Sigma_f(\mathcal{P}) = \sum_{i=1}^{n} (x_i - x_{i-1}) \sup_{x \in [x_{i-1}, x_i]} f(x).$$

Funciones reales de variable real

- En el caso particular en que los puntos de una partición $\mathcal{P}$ estén igualmente equiespaciados, diremos que la partición es **uniforme**, en cuyo caso se cumple $x_i - x_{i-1} = \frac{b-a}{n}$, $1 \leq i \leq n$.
- En tal caso, podemos escribir

$$\sigma_f(\mathcal{P}) = \frac{b-a}{n} \sum_{i=1}^{n} \inf_{x \in [x_{i-1}, x_i]} f(x), \; \Sigma_f(\mathcal{P}) = \frac{b-a}{n} \sum_{i=1}^{n} \sup_{x \in [x_{i-1}, x_i]} f(x).$$





Universidad Internacional de Valencia

Definición

Diremos que f es **integrable Riemann** *en $[a,b]$ si*

$$\sup_{\mathcal{P}} \sigma_f(\mathcal{P}) = \inf_{\mathcal{P}} \Sigma_f(\mathcal{P}).$$

En tal caso, llamaremos

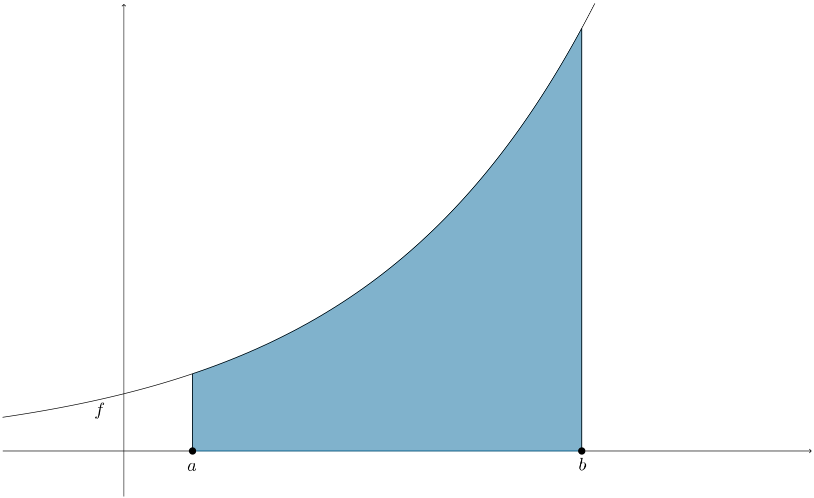
$$\int_a^b f(x)dx = \sup_{\mathcal{P}} \sigma_f(\mathcal{P}) = \inf_{\mathcal{P}} \Sigma_f(\mathcal{P}),$$

que representa el área encerrada por f entre a y b, y que recibe el nombre de **integral definida** *de f en $[a,b]$.*

Teorema (Regla de Barrow)

Si F es una primitiva de f, entonces $\int_a^b f(x)dx = F(b) - F(a)$.

Funciones reales de variable real



Ejemplos

1. $\displaystyle\int_0^1 2x dx = \left[x^2\right]_0^1 = 1^2 - 0^2 = 1.$
2. $\displaystyle\int_0^1 e^x dx = [e^x]_0^1 = e^1 - e^0 = e - 1.$
3. $\displaystyle\int_0^\pi \sin(x) dx = [-\cos(x)]_0^\pi = -\cos(\pi) - (-\cos(0)) = -(-1) - (-1) = 2.$
4. $\displaystyle\int_{-1}^1 \frac{1}{1+x^2} dx = [\arctan(x)]_{-1}^1 = \arctan(1) - \arctan(-1) = \frac{\pi}{4} - \left(-\frac{\pi}{4}\right) = \frac{\pi}{2}.$

Introducción a la probabilidad

- La probabilidad es una medida que asigna un valor de certidumbre a la ocurrencia de un suceso determinado.
- Habitualmente toma valores comprendidos entre 0 y 1, donde valores próximos a 0 indica poca probabilidad y los próximos a 1 bastante probabilidad.
- Por ejemplo, si lanzamos una moneda equilibrada, es razonable asignar $P(\text{cara}) = 0.5$ y $P(\text{cruz}) = 0.5$.
- Otro ejemplo: si tenemos una urna con 30 bolas blancas y 70 negras y extraemos una al azar, también es razonable asignar $P(\text{blanca}) = 0.3$ y $P(\text{negra}) = 0.7$.
- La probabilidad pretende asignar la proporción a la que tienden a darse los resultados tras muchos experimentos:
  - En el ejemplo de las monedas: 50 % caras y 50 % cruces.
  - En el ejemplo de la urna: 30 % blancas y 70 % negras.

Definición

El conjunto formado por todos los resultados posibles, Ω, recibe el nombre de **espacio muestral**. *Cada subconjunto formado por esas posibles ocurrencias, $A \subseteq \Omega$, reciben el nombre de* **sucesos** *o* **eventos**.

Ejemplos

1. *El espacio muestral asociado al lanzamiento de una moneda es $\Omega = \{cara, cruz\}$.*
2. *Extracción en una urna con bolas blancas y negras: $\Omega = \{blanca, negra\}$.*
3. *Lanzamiento de un dado: $\Omega = \{1, 2, 3, 4, 5, 6\}$.*
4. *Lanzamiento de dos monedas:*
   $\Omega = \{(cara, cara), (cara, cruz), (cruz, cara), (cruz, cruz)\}$.

Ejemplos

1. *Lanzamiento de una moneda equilibrada:* $\Omega = \{cara, cruz\}$. *Posibles sucesos:*
   - *No ocurre nada:* $A = \emptyset$. $P(A) = 0$.
   - *Sale cara:* $A = \{cara\}$. $P(A) = 0.5$.
   - *Sale cruz:* $A = \{cruz\}$. $P(A) = 0.5$.
   - *Sale alguna de las dos caras:* $A = \{cara, cruz\}$. $P(A) = 1$.
2. *Lanzamiento de un dado:* $\Omega = \{1, 2, 3, 4, 5, 6\}$. *Ejemplos de sucesos:*
   - *Sale un número par:* $A = \{2, 4, 6\}$. $P(A) = 0.5$.
   - *Sale un número impar:* $A = \{1, 3, 5\}$. $P(A) = 0.5$.
   - *Sale un número menor que 5:* $A = \{1, 2, 3, 4\}$. $P(A) = \frac{2}{3}$.
   - *Sale un número que empieza por "c":* $A = \{4, 5\}$. $P(A) = \frac{1}{3}$.

Ejemplos

3. *Lanzamiento de dos monedas:*
   $\Omega = \{(cara, cara), (cara, cruz), (cruz, cara), (cruz, cruz)\}$. *Ejemplos de sucesos:*
   - *Obtener el mismo resultado:* $A = \{(cara, cara), (cruz, cruz)\}$. $P(A) = 0.5$.
   - *No obtener dos caras:* $A = \{(cara, cruz), (cruz, cara), (cruz, cruz)\}$. $P(A) = 0.75$.
4. *Lanzamiento de dos dados:* $\Omega = \{(i, j) \mid 1 \leq i, j \leq 6\}$. *Ejemplos de sucesos:*
   - *La suma de los resultados es* 10*:* $A = \{(4, 6), (5, 5), (6, 4)\}$. $P(A) = \frac{1}{12}$.
   - *Ambos resultados son pares:*
     $A = \{(2, 2), (2, 4), (2, 6), (4, 2), (4, 4), (4, 6), (6, 2), (6, 4), (6, 6)\}$. $P(A) = \frac{1}{4}$.

Los sucesos pueden combinarse entre sí mediante las operaciones básicas entre conjuntos (unión, intersección, complementario, etc.).

Ejemplos

1. *Lanzamos un dado y consideramos el siguiente suceso: "sacar un número par o que sea mayor que 3". Llamamos A al suceso "sacar un número par" (por lo que $A = \{2, 4, 6\}$) y B al suceso "sacar un número mayor que 3" (luego $B = \{4, 5, 6\}$). Por tanto, el suceso que buscamos es*

$$A \cup B = \{2, 4, 6\} \cup \{4, 5, 6\} = \{2, 4, 5, 6\}.$$

2. *Bajo las condiciones anteriores, consideramos el suceso "sacar un número par mayor que 3". En este caso la operación que debemos realizar es*

$$A \cap B = \{2, 4, 6\} \cap \{4, 5, 6\} = \{4, 6\}.$$

Definición

Dos sucesos $A, B \subseteq \Omega$ son **mutuamente excluyentes** *si $A \cap B = \emptyset$.*

Ejemplo

Al lanzar un dado, $A = \{1, 3\}$, $B = \{2, 4, 6\}$. Se cumple $A \cap B = \emptyset$.

Definición

Una probabilidad P definida sobre el conjunto de sucesos asociado a un espacio muestral Ω es una función real cumpliendo:

1. *$P(A) \geq 0$ para cualquier suceso $A \subseteq \Omega$.*
2. *$P(\Omega) = 1$.*
3. *Para cada par de sucesos $A, B \subseteq \Omega$* **mutuamente excluyentes** *(es decir, $A \cap B = \emptyset$), se tiene $P(A \cup B) = P(A) + P(B)$.*

Introducción a la probabilidad

De los axiomas asociados a una probabilidad se pueden deducir las siguientes propiedades:

Propiedades

Sea Ω un espacio muestral, P una probabilidad y $A, B \subseteq \Omega$ sucesos. Entonces se cumple:

1. $P(A^c) = 1 - P(A)$.
2. $P(\emptyset) = 0$.
3. *Si* $A \subseteq B$, *entonces* $P(A) \leq P(B)$. *En particular,* $P(A) \leq 1$.
4. *Si* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
5. *En particular,* $P(A \cup B) \leq P(A) + P(B)$.

Ejercicio

Demuéstrense las propiedades anteriores utilizando los tres axiomas de la probabilidad.

Ejemplo

En un grupo de amigos, un 40 % de ellos estudia Ingeniería Informática, un 30 % Matemáticas y un 10 % ambas cosas a la vez. ¿Cuál es la probabilidad de que una persona elegida al azar estudie alguna de las dos carreras? Si $A = \{\text{estudiar Ingeniería Informática}\}$ *y* $B = \{\text{estudiar Matemáticas}\}$, *se nos pide*

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.4 + 0.3 - 0.1 = 0.6.$$

La ocurrencia de determinados sucesos (o el conocimiento de determinada información) puede *condicionar* la ocurrencia de otros sucesos.

Ejemplo

Lanzamos un dado equilibrado y consideramos el suceso $A = \{\textit{obtener un } 2\}$. *En ese caso, y sin saber nada más, sabemos que la probabilidad de que ocurra ese suceso es* $\frac{1}{6}$. *Sin embargo, esto puede cambiar si conocemos de antemano que ha ocurrido otro suceso* B.

- *Por ejemplo, si* $B = \{\textit{se ha obtenido un número par}\}$, *entonces la probabilidad de* A *condicionada a* B *es* $\frac{1}{3}$.
- *Por otra parte, si* $B = \{\textit{se ha obtenido un número impar}\}$, *entonces dicha probabilidad es* 0.

Definición

Sean A, B sucesos de un espacio muestral Ω y P una probabilidad asociada a éste. Definimos la probabilidad de A **condicionada** *a B como*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Ejemplo

En el ejemplo anterior del dado, y tomando $A = \{obtener\ un\ 2\}$:

- *Si $B = \{obtener\ un\ número\ par\}$, entonces*
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}.$$
- *Si $B = \{obtener\ un\ número\ impar\}$, entonces*
$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\emptyset)}{P(B)} = \frac{0}{\frac{1}{2}} = 0.$$

Teorema (Teorema de Bayes)

Sean A, B sucesos de un espacio muestral Ω y P una probabilidad asociada. Entonces

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}.$$

Ejemplo

Consideramos el lanzamiento de un dado y los sucesos $A = \{1, 2, 4, 5\}$ y $B = \{2, 4, 6\}$. Entonces

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}.$$

Por otra parte, utilizando el teorema de Bayes:

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} = \frac{\frac{2}{3} \cdot \frac{1}{2}}{\frac{2}{3}} = \frac{1}{2}.$$

Universidad Internacional de Valencia

Definición

Diremos que dos sucesos $A, B \subseteq \Omega$ son independientes si $P(A|B) = P(A)$.

En general, de la definición de probabilidad condicionada, se deduce que para cualquier par de sucesos $A, B \subseteq \Omega$ (independientes o no) se cumple

$$P(A \cap B) = P(A|B)P(B).$$

Por tanto:

Teorema

Dos sucesos $A, B \subseteq \Omega$ son independientes si y sólo si se cumple $P(A \cap B) = P(A)P(B)$.

Ejemplo

Lanzamos una moneda y un dado, ambos equilibrados. ¿Cuál es la probabilidad de obtener una cara en la moneda y, simultáneamente, un 3 en el dado?

Si llamamos A al suceso "obtener una cara al lanzar la moneda" y B al suceso "obtener un 3 al lanzar el dado", entonces se nos pide $P(A \cap B)$.

Dado que los sucesos A y B son independientes, se tiene:

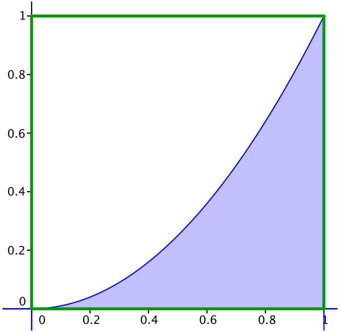
$$P(A \cap B) = P(A)P(B) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}.$$

Introducción a la probabilidad

En ocasiones nos podemos encontrar con espacios muestrales infinitos.

Ejemplo

Sea $\Omega = [0,1] \times [0,1]$ *y la función* $f : [0,1] \to [0,1]$ *dada por* $f(x) = x^2$. *¿Cuál es la probabilidad de que un punto elegido al azar en el cuadrado* $[0,1] \times [0,1]$ *esté por debajo de la gráfica de* f*?*



El área encerrada por la región es:

$$\int_0^1 x^2 \mathrm{d}x = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}.$$

Por tanto, la probabilidad es:

$$\frac{\text{Área región}}{\text{Área cuadrado}} = \frac{\frac{1}{3}}{1} = \frac{1}{3}.$$

Definición

Una **partición** $A_1, A_2, \ldots, A_n$ *sobre un espacio muestral* Ω *es un conjunto de sucesos cumpliendo* $A_i \cap A_j = \emptyset$ *para* $1 \leq i, j \leq n$, $i \neq j$ *y además*

$$\bigcup_{i=1}^{n} A_i = \Omega.$$

Teorema (Teorema de la probabilidad total)

Sea $A_1, A_2, \ldots, A_n$ *una* **partición** *sobre un espacio muestral* Ω *y* B *un suceso cualquiera del que se conoce* $P(B|A_i)$, *para* $1 \leq i \leq n$. *Entonces*

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i).$$

Universidad Internacional de Valencia

Ejemplo

En un determinado grupo de personas, un 1 % de las mismas padece una enfermedad específica. Existe un test de detección de la misma, el cual tiene un 1 % de probabilidad de obtener un falso positivo, así como un 1 % de obtener un falso negativo. ¿Cuál es la probabilidad de que un individuo que ha dado positivo en el test padezca realmente la enfermedad?

Sea E el suceso "sufrir la enfermedad" y T el suceso "obtener positivo". Se sabe que $P(E) = 0.01$, $P(T|\overline{E}) = 0.01$ y $P(\overline{T}|E) = 0.01$. Se pide obtener $P(E|T)$.
Aplicando el teorema de Bayes y el teorema de la probabilidad total:

$$p(E|T) = \frac{P(T|E)P(E)}{P(T)} = \frac{P(T|E)P(E)}{P(T|E)P(E) + P(T|\overline{E})P(\overline{E})}$$
$$= \frac{0.99 \cdot 0.01}{0.99 \cdot 0.01 + 0.01 \cdot 0.99} = \frac{1}{2} = 0.5.$$

¡Muchas gracias!

Universidad Internacional de Valencia

Contacto:

david.zorio@campusviu.es

De:
Planeta Formación y Universidades