Matemáticas

Apuntes para el Grado de Biotecnología

Universidad de Zaragoza

Alberto Elduque

Departamento de Matemáticas
Universidad de Zaragoza
50009 Zaragoza, Spain

Las siguientes referencias han sido fundamentales en la preparación de estas notas:

- J.J. Lucas García, P. Lucas Saorín y J. Marín Fernández, *Matemáticas*, DM, Colección Texto–Guía, ICE Universidad de Murcia, 2a ed., 1999.
- C. Neuhauser: *Calculus for Biology and Medicine*, Pearson, 3rd ed., 2010.
- J. Stewart: *Calculus*, Brooks/Cole., 5th edition, 2003.

Las notas de Conchita Martínez Pérez, de quien "he heredado" la asignatura, han sido también de gran utilidad.

Índice general

Capítulo 1

Conjuntos de números. Combinatoria básica

La necesidad de contar y medir ha dado lugar a distintos sistemas de números.

§ 1. Números naturales

El primer conjunto de números con el que nos encontramos es el de los números naturales:

$$\mathbb{N} = \{1, 2, 3, 4, \ldots\}.$$

En $\mathbb{N}$ podemos sumar y multiplicar, y la propiedad siguiente nos permite probar muchos resultados:

Principio de Inducción: Si un subconjunto S de $\mathbb{N}$ verifica las dos propiedades siguientes:

- $1 \in S$,
- siempre que un número natural m pertenezca a S, también $m+1$ pertenece a S: $m \in S \Rightarrow m+1 \in S$,

Entonces $S = \mathbb{N}$.

Ejemplo. Para probar la fórmula $1 + 2 + \cdots + n = \dfrac{n(n+1)}{2}$, esto es, la suma de los n primeros números naturales es $\dfrac{n(n+1)}{2}$, consideramos el conjunto $S = \left\{n \in \mathbb{N} : 1 + \cdots + n = \dfrac{n(n+1)}{2}\right\}$. Deseamos probar la igualdad $S = \mathbb{N}$.

- $1 \in S$ pues $1 = \dfrac{1 \times 2}{2}$.
- Si $m \in S$, entonces $1 + \cdots + m = \dfrac{m(m+1)}{2}$, luego
$$1 + \cdots + m + (m+1) = \frac{m(m+1)}{2} + (m+1) = \frac{m(m+1) + 2(m+1)}{2} = \frac{(m+1)(m+2)}{2},$$
así también $m + 1 \in S$. Por el principio de inducción, $S = \mathbb{N}$. $\square$

§ 2. Números enteros

Añadimos a $\mathbb{N}$ el 0 y los opuestos de los números naturales para poder restar siempre, obteniendo el conjunto de los números enteros:

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$$

En $\mathbb{Z}$ podemos sumar, restar y multiplicar.

§ 3. Números racionales

Añadimos los inversos de los números enteros no nulos, y 'cerramos' para la multiplicación para obtener los números racionales

$$\mathbb{Q} = \left\{ \frac{m}{n} : m, n \in \mathbb{Z},\ n \neq 0 \right\}$$

donde $\dfrac{m}{n} = \dfrac{m'}{n'}$ si, y solo si, $mn' = nm'$. Las operaciones se efectúan como sigue:

$$\begin{aligned} \frac{a}{b} \pm \frac{c}{d} &= \frac{ad \pm bc}{bd}, \\ \frac{a}{b} \cdot \frac{c}{d} &= \frac{ac}{bd}, \\ \frac{a}{b} : \frac{c}{d} &= \frac{ad}{bc}. \end{aligned}$$

(Dividir un número por otro equivale a multiplicar el primer número por el inverso del segundo.)

§ 4. Combinatoria básica

En combinatoria estudiamos posibilidades de elegir elementos de conjuntos. Distinguimos entre:

Permutaciones: Modos de ordenar los elementos de un conjunto de n elementos:

$$n! = n \cdot (n-1) \cdot \cdots \cdot 2 \cdot 1$$

(esto se llama el *factorial de n*, que es el producto de los n primeros números naturales).

Así $1! = 1$, $2! = 2$, $3! = 3 \cdot 2 \cdot 1 = 6$, $4! = 24$, $5! = 120$, $6! = 720$, ... Además definimos $0! = 1$.

Variaciones: Secuencias ordenadas de p elementos distintos tomados de un conjunto de n elementos $(0 \leq p \leq n)$.

El número de variaciones es

$$V_n^p = n \cdot (n-1) \cdot \cdots \cdot (n-p+1) = \frac{n!}{(n-p)!}\,.$$

Por ejemplo, con 5 colores, ¿cuántas banderas de tres franjas horizontales distintas se pueden formar?

La solución es $V_5^3 = 5 \times 4 \times 3 = 60$.

Combinaciones: Subconjuntos de p elementos de un conjunto de n elementos ($0 \leq p \leq n$ y ¡el orden no importa!) El número de combinaciones viene dado por el *número combinatorio*

$$\binom{n}{p} := \frac{n \cdot (n-1) \cdot \dots \cdot (n-p+1)}{p!} = \frac{n!}{p! \cdot (n-p)!} .$$

Notemos la igualdad

$$\binom{n}{p} = \binom{n}{n-p} .$$

Por ejemplo, un estudiante tiene que responder a tres de las cinco preguntas propuestas en un examen. ¿Cuántas posibilidades tiene? Solución: $\binom{5}{3} = \binom{5}{2} = \frac{5 \times 4}{2} = 10$.

Las combinaciones aparecen al desarrollar el *binomio de Newton*[1]:

$$\begin{aligned}(a+b)^n &= (a+b) \cdot (a+b) \cdot \dots \cdot (a+b) \\ &= \sum_{p=0}^{n} \binom{n}{p} a^{n-p} b^p .\end{aligned}$$

Si deseamos elegir un subconjunto de $m+1$ elementos de un conjunto de $n+1$ elementos ($0 \leq m \leq n$) y fijamos uno de ellos, tenemos dos posibilidades, los subconjuntos que contienen al elemento fijado, obtenidos eligiendo m elementos de los n elementos restantes, y los que no lo contienen, obtenidos eligiendo $m+1$ elementos de los n elementos restantes:

$$\binom{n+1}{m+1} = \binom{n}{m} + \binom{n}{m+1} .$$

Utilizando esta igualdad es fácil calcular los números combinatorios con n pequeño, que forman el *triángulo de Pascal*, cuyas primeras filas son:

$$\begin{array}{ccccccccc} & & & & \binom{0}{0} & & & & \\ & & & \binom{1}{0} & & \binom{1}{1} & & & \\ & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\ & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\ \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} \end{array}$$

Utilizando la fórmula anterior, cada número del triángulo, salvo los de los laterales, es la suma de los dos números que tiene encima, a izquierda y derecha. Así el triángulo anterior es:

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \end{array}$$

[1] Del conjunto de n factores elegimos b en p factores y a en los restantes, de todos los modos posibles.

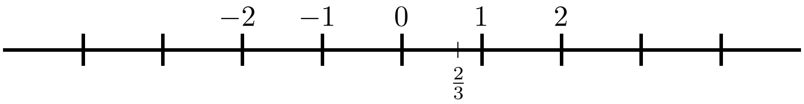
Ejemplo. La fórmula básica de la probabilidad nos dice que la probabilidad P de que un suceso ocurra es el cociente

$$P = \frac{\text{número de casos favorables}}{\text{número de casos posibles}}.$$

¿Cuál es la probabilidad de que al lanzar siete monedas salgan tres caras?
Solución: $P = \dfrac{\binom{7}{3}}{2^7} = \dfrac{35}{128}$.

§ 5. Números reales

Los matemáticos griegos pensaron que con los números racionales se podía medir cualquier cantidad. Representaron los números racionales como puntos de una recta:



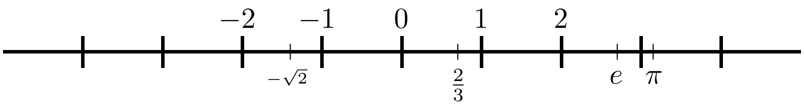
y pensaron que a todo punto de la recta le correspondía un número racional.

Pronto se dieron cuenta de que esto no es así: la longitud de la diagonal del cuadrado unidad mide $\sqrt{2}$, que no es un número racional.

En efecto, si fuera racional se tendría $\sqrt{2} = \dfrac{m}{n}$ con m, n números naturales primos entre sí. Pero entonces $2n^2 = m^2$, luego m es par: $m = 2m'$. Ahora nos queda $n^2 = 2(m')^2$, luego también n es par, llegando a una contradicción.

El sistema de números que se corresponde con los puntos de la recta es el de los números reales $\mathbb{R}$:

- $\mathbb{R}$ contiene raíces de números positivos, pero también muchos otros números, como π (básico en trigonometría), o e (base de los logaritmos naturales):



- En $\mathbb{R}$ podemos sumar, restar, multiplicar y dividir (por números distintos de 0) como en $\mathbb{Q}$, además verifica que para todo subconjunto *acotado superiormente* existe una cota superior mínima, llamada *supremo*. Del mismo modo, para todo subconjunto acotado inferiormente existe su *ínfimo*.

  Por ejemplo, en $\mathbb{Q}$ no existe el supremo del conjunto $\{x \in \mathbb{Q} : x^2 < 2\}$. Este supremo es $\sqrt{2}$. De hecho $\mathbb{R}$ se obtiene *completando* $\mathbb{Q}$ añadiendo supremos de subconjuntos de números racionales.

- Todo número real tiene una representación decimal $A'a_1a_2a_3\ldots$, con $A \in \mathbb{Z}$ y $a_n \in \mathbb{Z}$ con $0 \le a_n \le 9$ para todo $n \in \mathbb{N}$, que es única salvo las siguientes excepciones:

$$A'a_1a_2\ldots a_k99999\ldots = A'a_1a_2\ldots(a_k+1)00000\ldots \quad , \; a_k \neq 9.$$

Si, por ejemplo, $x = 2'479999\ldots$, entonces $10x = 24'79999\ldots$, luego $9x = 10x - x = 22'32$ y $x = \frac{1}{9}22'32 = 2'48$.

En ocasiones, para números pequeños o muy grandes, resulta útil la *notación científica*:

$$\pm A'a_1 \ldots a_k \times 10^r$$

con $1 \leq A \leq 9$, $0 \leq a_i \leq 9$, $i = 1, \ldots, k$, $r \in \mathbb{Z}$. Así, por ejemplo,

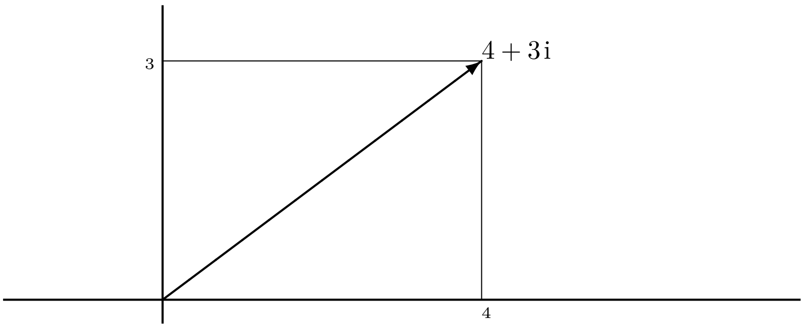
$$0'00000089 = 8'9 \times 10^{-7},$$
$$5000000 = 5 \times 10^6.$$

§ 6. Números complejos

En $\mathbb{R}$ no podemos resolver ecuaciones tan sencillas como $x^2 + 1 = 0$, por ello añadimos una *unidad imaginaria* i que verifique $\mathrm{i}^2 = -1$ y formamos el conjunto de los números complejos:

$$\mathbb{C} = \{a + b\mathrm{i} : a, b \in \mathbb{R}\}$$

cuyos 'números' podemos identificar con los puntos del plano $\mathbb{R}^2$: $a + b\mathrm{i} \leftrightarrow (a, b)$.



- En $\mathbb{C}$ podemos sumar, restar, multiplicar y dividir extendiendo ls operaciones de $\mathbb{R}$:

$$(a + b\mathrm{i}) \pm (c + d\mathrm{i}) = (a \pm c) + (b \pm d)\mathrm{i},$$
$$(a + b\mathrm{i})(c + d\mathrm{i}) = (ac - bd) + (ad + bc)\mathrm{i},$$
$$\frac{a + b\mathrm{i}}{c + d\mathrm{i}} = \frac{(a + b\mathrm{i})(c - d\mathrm{i})}{(c + d\mathrm{i})(c - d\mathrm{i})} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}\,\mathrm{i}$$

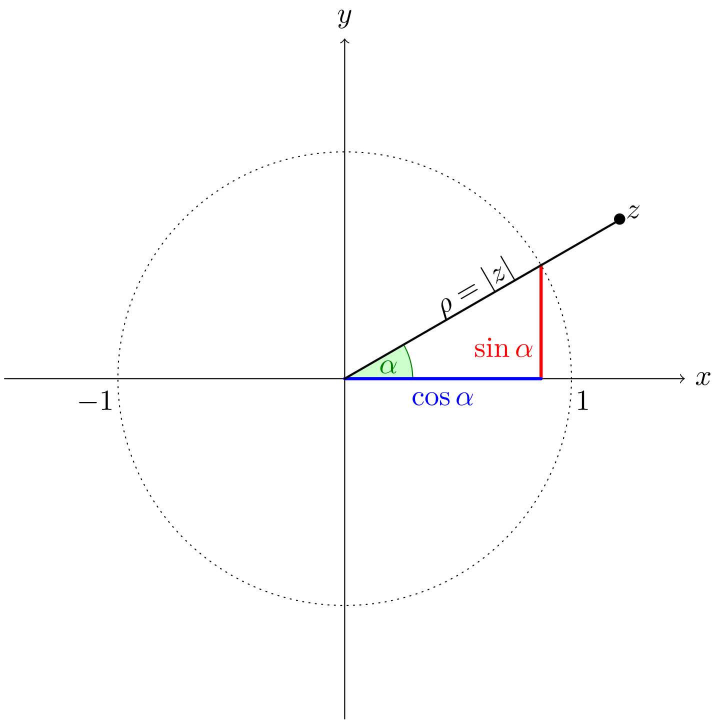
  (Para dividir multiplicamos numerador y denominador por el *conjugado* del denominador. El conjugado de $z = a + b\mathrm{i}$ es $\bar{z} = a - b\mathrm{i}$, de modo que $z\bar{z} = \bar{z}z = a^2 + b^2$.)

- Sorprendentemente, no solo la ecuación $x^2 + 1 = 0$ tiene solución en $\mathbb{C}$, sino toda "ecuación algebraica".

- Todo número complejo $z \in \mathbb{C}$ se puede representar en

forma cartesiana: $z = a + b\mathrm{i}$, donde $a \in \mathbb{R}$ se dice *parte real* de z y $b \in \mathbb{R}$ se dice *parte imaginaria*: $a = \mathrm{Re}(z)$, $b = \mathrm{Im}(z)$.

forma trigonométrica: $z = \rho(\cos\alpha + \mathrm{i}\,\mathrm{sen}\,\alpha)$, donde $\rho = |z| := \sqrt{a^2 + b^2}$ se dice *módulo* de z e indica la distancia al origen, y α es el ángulo que forma z con la parte positiva del eje de las abscisas OX, que se dice *argumento* de z.[2]



forma polar: $z = \rho e^{\mathrm{i}\alpha}$, donde[3] $e^{\mathrm{i}\alpha} := \cos\alpha + \mathrm{i}\,\mathrm{sen}\,\alpha$.

Esta es la forma más adecuada para calcular potencias y raíces, pues si $z = \rho e^{\mathrm{i}\alpha}$, $z' = \rho' e^{\mathrm{i}\alpha'}$, entonces

$$\begin{aligned} zz' &= \rho\rho'\Big((\cos\alpha + \mathrm{i}\,\mathrm{sen}\,\alpha)(\cos\alpha' + \mathrm{i}\,\mathrm{sen}\,\alpha')\Big) \\ &= \rho\rho'\Big((\cos\alpha\cos\alpha' - \mathrm{sen}\,\alpha\,\mathrm{sen}\,\alpha') + \mathrm{i}(\cos\alpha\,\mathrm{sen}\,\alpha' + \mathrm{sen}\,\alpha\cos\alpha')\Big) \\ &= \rho\rho'\Big(\cos(\alpha+\alpha') + \mathrm{i}\,\mathrm{sen}(\alpha+\alpha')\Big) \\ &= \rho\rho' e^{\mathrm{i}(\alpha+\alpha')}, \end{aligned}$$

luego $z^n = \rho^n e^{in\alpha}$ para todo $n \in \mathbb{N}$.

Así, las raíces n-ésimas de $z = \rho e^{\mathrm{i}\alpha} = \rho e^{\mathrm{i}(\alpha+2k\pi)}$ son

$$\sqrt[n]{\rho} e^{\mathrm{i}\left(\frac{\alpha}{n} + \frac{2k\pi}{n}\right)}, \quad k = 0, 1, \ldots, n-1.$$

[2]El argumento está definido salvo suma de múltiplos enteros de 2π radianes.

[3]Esta notación está justificada porque $\lim_{n\to\infty}\left(1 + \frac{x + y\,\mathrm{i}}{n}\right)^n = e^x(\cos y + \mathrm{i}\,\mathrm{sen}\,y)$.

Nota. Utilizando la fórmula

$$(\cos\alpha + \mathrm{i}\,\operatorname{sen}\alpha)(\cos\alpha' + \mathrm{i}\,\operatorname{sen}\alpha') = \cos(\alpha+\alpha') + \mathrm{i}\,\operatorname{sen}(\alpha+\alpha')$$

ya no hace falta recordar 'de memoria' las fórmulas trigonométricas del coseno y seno de la suma (o diferencia).

Capítulo 2

Funciones reales de una variable. Derivadas

§ 1. Definiciones básicas

Una *función* o aplicación f es una regla que asocia a cada elemento x de un conjunto A, llamado *dominio* de f, un único elemento, denotado por $f(x)$, de un conjunto B.

Escribimos

$$\begin{aligned} f : A &\longrightarrow B \\ x &\mapsto f(x). \end{aligned}$$

La *imagen* o *rango* de f es el subconjunto $\{y \in B : \exists x \in A \text{ tal que } f(x) = y\}$ de B.

Si el rango de una función g está contenido en el dominio de otra f, podemos considerar la función compuesta $g \circ f : x \mapsto g(f(x))$. Si f da una biyección entre A y su rango, podemos considerar la función inversa $f^{-1} : rango(f) \to A$.

En este capítulo nos restringiremos a funciones

$$\begin{aligned} f : I &\longrightarrow \mathbb{R}, \\ x &\mapsto f(x), \end{aligned}$$

con I un subconjunto no vacío de $\mathbb{R}$. Habitualmente, I será un *intervalo* de alguna de estas formas ($a, b \in \mathbb{R}$, $a < b$):

$$\begin{aligned} (a,b) &:= \{x \in \mathbb{R} \mid a < x < b\} && \text{intervalo abierto acotado,} \\ [a,b] &:= \{x \in \mathbb{R} \mid a \leq x \leq b\} && \text{intervalo cerrado acotado,} \\ [a,b) &:= \{x \in \mathbb{R} \mid a \leq x < b\} && \text{intervalo semiabierto acotado,} \\ (a,b] &:= \{x \in \mathbb{R} \mid a < x \leq b\} && \text{intervalo semiabierto acotado,} \end{aligned}$$

$$\begin{aligned}
(a,+\infty) &:= \{x \in \mathbb{R} \mid a < x\} \quad \text{intervalo abierto no acotado a derecha,}\\
[a,+\infty) &:= \{x \in \mathbb{R} \mid a \leq x\} \quad \text{intervalo cerrado no acotado a derecha,}\\
(-\infty,b) &:= \{x \in \mathbb{R} \mid x < b\} \quad \text{intervalo abierto no acotado a izquierda,}\\
(-\infty,b] &:= \{x \in \mathbb{R} \mid x \leq b\} \quad \text{intervalo cerrado no acotado a izquierda,}\\
(-\infty,+\infty) &:= \mathbb{R}.
\end{aligned}$$

Los puntos $c \in I$ para los que existe $\delta > 0$ con $(c-\delta, c+\delta) \subseteq I$ se dicen *puntos interiores* I, los puntos a y b arriba son *puntos frontera*.

§ 2. Ejemplos de funciones

A lo largo del curso haremos uso principalmente de las siguientes funciones:

- Funciones polinómicas: $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$.
- Funciones racionales: $f(x) = \dfrac{p(x)}{q(x)}$ con p y q polinómicas (su dominio es todo $\mathbb{R}$ salvo el conjunto (finito) de raíces del polinomio q).
- Exponencial: $f(x) = a^x$ con $a \in \mathbb{R}$, $a > 0$. En especial $\exp(x) := e^x$, donde e es la base de los logaritmos neperianos: $e = \text{lím}_{n\to+\infty} \left(1 + \dfrac{1}{n}\right)^n \simeq 2,7182818284....$
- Logaritmo (inversa de la exponencial): $\log_a x$, $x \in (0,+\infty)$, en especial $\ln x := \log_e x$ (*logaritmo neperiano*).

Todas estas funciones se pueden combinar mediante operaciones, composición, inversas, ...

Ejemplo. De acuerdo con la ley de desintegración radiactiva, si denotamos por $W(t)$ la cantidad de carbono 14 en el tiempo t, con $W(0) = W_0$, entonces

$$W(t) = W_0 e^{-\lambda t}, \quad t \geq 0,$$

con $\lambda > 0$ la *tasa de desintegración*.

Esta tasa se expresa en función de la *vida media* del material, que es el tiempo necesario T_h para que la mitad del material se desintegre. Así,

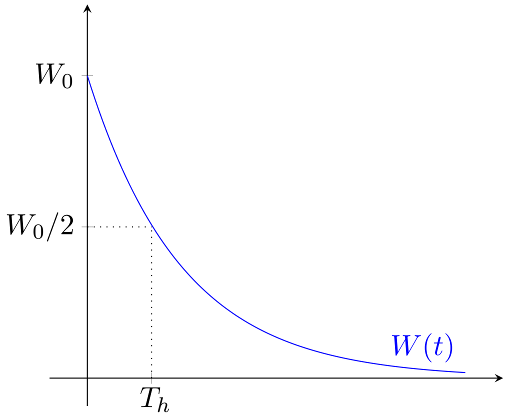
$$W(T_h) = \begin{cases} \dfrac{1}{2} W_0, \\ W_0 e^{-\lambda T_h}, \end{cases}$$

luego

$$e^{-\lambda T_h} = \frac{1}{2}, \quad e^{\lambda T_h} = 2, \quad \lambda T_h = \ln 2,$$

de donde obtenemos $\lambda = \dfrac{\ln 2}{T_h}$. Para el carbono 14, $T_H \simeq 5730$ años, luego la tasa de desintegración es $\lambda = \dfrac{\ln 2}{5370}$ si tomamos como unidad de tiempo los años. (¡Es muy importante tener claro qué unidades se toman!)

La *gráfica* de $W(t)$, esto es, el conjunto de puntos $\{(t, W(t)) \mid t \geq 0\} \subseteq \mathbb{R}^2$ tiene la forma



- Funciones trigonométricas:

$\operatorname{sen} x$ (seno), que es una *función impar*: $\operatorname{sen}(-x) = -\operatorname{sen} x$,

$\cos x$ (coseno), que es una *función par*: $\cos(-x) = \cos x$,

$\operatorname{tg} x = \dfrac{\operatorname{sen} x}{\cos x}$ (tangente), $x \notin \dfrac{\pi}{2} + \pi\mathbb{Z}$,

$\operatorname{cotg} x = \dfrac{\cos x}{\operatorname{sen} x}$ (cotangente), $x \notin \pi\mathbb{Z}$,

$\sec x = \dfrac{1}{\cos x}$ (secante), $x \notin \dfrac{\pi}{2} + \pi\mathbb{Z}$,

$\operatorname{cosec} x = \dfrac{1}{\operatorname{sen} x}$ (cosecante), $x \notin \pi\mathbb{Z}$,

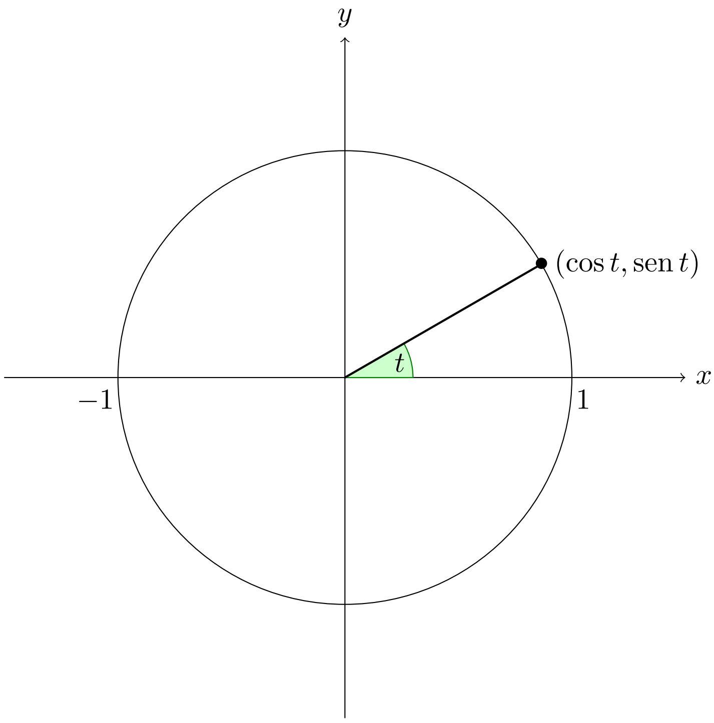
$\operatorname{arc\,sen} x$ (arco seno, inversa del seno), $x \in [-1, 1]$ con valores en $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$,

$\operatorname{arc\,cos} x$ (arco coseno, inversa del coseno), $x \in [-1, 1]$ con valores en $[0, \pi]$,

$\operatorname{arc\,tg} x$ (arco tangente, inversa de la tangente), $x \in \mathbb{R}$ con valores en $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

Las funciones seno y coseno nos ayudan a 'parametrizar' la circunferencia unidad:

$$\left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\right\} = \{(\cos t, \operatorname{sen} t) \mid t \in [0, 2\pi]\}$$



o, más generalmente, las elipses[1] de ecuación $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ $(0 < a, b \in \mathbb{R})$:

$$\left\{ (x,y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right\} = \{(a\cos t, b\operatorname{sen} t) \mid t \in [0, 2\pi]\} .$$

- Funciones trigonométricas hiperbólicas:

  Si consideramos la hipérbola H de ecuación $x^2 - y^2 = 1$, todo punto $(x,y) \in H$ con $x > 0$ verifica $(x+y)(x-y) = 1$ y existe un único $t \in \mathbb{R}$ tal que $x + y = e^t$, $x - y = e^{-t}$, luego

  $$(x,y) = (\cosh t, \operatorname{senh} t)$$

  con

  $$\cosh t := \frac{e^t + e^{-t}}{2} \text{ (coseno hiperbólico), que es función par,}$$
  $$\operatorname{senh} t := \frac{e^t - e^{-t}}{2} \text{ (seno hiperbólico), que es función impar.}$$

  con inversas

  $\arg\cosh x$ (argumento coseno hiperbólico) $x \in [1, +\infty)$,
  $\arg\operatorname{senh} x$ (argumento seno hiperbólico).

---

[1] La elipse de ecuación $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ $(a > b > 0)$ es el lugar geométrico de los puntos del plano cuya suma de distancias a los puntos $(c, 0)$ y $(-c, 0)$ es $2a$, siendo $c^2 = a^2 - b^2$.

§ 3. Límites de sucesiones

Una sucesión es una aplicación

$$
\begin{aligned}
a : \mathbb{N} &\longrightarrow \mathbb{R}\\
n &\mapsto a(n) = a_n.
\end{aligned}
$$

Hablaremos entonces de la sucesión $(a_n)_{n\in\mathbb{N}}$ o simplemente de (a_n).

Ejemplo. Los primeros términos de la sucesión $\left(\frac{1}{n}\right)$ son $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$, que cada vez son más pequeños 'tendiendo' a 0.

Decimos que la sucesión (a_n) tiene *límite* a, y escribiremos $\lim_{n\to\infty} a_n = a$, o $a_n \to a$, si para todo intervalo J alrededor de a existe un número natural n_0 tal que $a_n \in J$ para todo $n \geq n_0$. Equivalentemente:

$$\lim_{n\to\infty} a_n = a \quad \text{si} \quad \forall\epsilon > 0,\ \exists n_0 \in \mathbb{N} \text{ tal que } |a_n - a| < \epsilon\ \forall n \geq n_0.$$

En este caso decimos que la sucesión es *convergente*.

En el ejemplo anterior, para $\epsilon = 10^{-6}$ podemos tomar $n_0 = 10^6 + 1$.

La definición anterior vale para límites infinitos, si entendemos que los intervalos de la forma $(K, +\infty)$ (o $(-\infty, -K)$) son intervalos alrededor de $+\infty$ (o $-\infty$). Esto es:

$$\lim_{n\to\infty} a_n = +\infty \quad \text{si} \quad \forall K > 0,\ \exists n_0 \in \mathbb{N} \text{ tal que } a_n > K\ \forall n \geq n_0,$$

y análogamente para $\lim_{n\to\infty} a_n = -\infty$. En estos casos decimos que la sucesión es *divergente*.

Propiedades. Si existen $\lim_{n\to\infty} a_n$ y $\lim_{n\to\infty} b_n$, entonces:

- $\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$, salvo que aparezca la 'indeterminación' $\infty - \infty$ (esto es, que uno de los límites a la derecha sea $+\infty$ y el otro $-\infty$).
- Si $c \in \mathbb{R}$, $c \neq 0$, $\lim_{n\to\infty}(ca_n) = c \lim_{n\to\infty} a_n$.
- $\lim_{n\to\infty}(a_n b_n) = (\lim_{n\to\infty} a_n)(\lim_{n\to\infty} b_n)$, salvo que aparezca la 'indeterminación' $0 \cdot \infty$.
- $\lim_{n\to\infty} \dfrac{a_n}{b_n} = \dfrac{\lim_{n\to\infty} a_n}{\lim_{n\to\infty} b_n}$, salvo que aparezcan las 'indeterminaciones' $\dfrac{0}{0}$ o $\dfrac{\infty}{\infty}$.

Más propiedades:

- Si (a_n) es creciente: $a_n \leq a_{n+1}\ \forall n$, entonces existe $\lim_{n\to\infty} a_n$ y este límite es finito si la sucesión está acotada (esto es, $\exists M \in \mathbb{R}$ tal que $a_n < M\ \forall n$). En otro caso el límite es $+\infty$. (Análogamente para sucesiones decrecientes.)
- **Teorema del Sandwich:** Si existe $n_0 \in \mathbb{N}$ tal que $a_n \leq c_n \leq b_n$, $\forall n \geq n_0$, y $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = L$, entonces $\lim_{n\to\infty} c_n = L$.

Ejemplo. La *sucesión de Fibonacci* (1202) es la sucesión definida mediante:

$$a_0 = 1, \quad a_1 = 1, \quad a_n = a_{n-1} + a_{n-2} \ \forall n \geq 2.$$

Sus primeros términos son

$$1, 1, 2, 3, 5, 8, 13, 21, \ldots$$

Consideramos las nuevas sucesiones:

$$b_n = a_n^2 - a_{n+1}a_{n-1}, \qquad c_n = \frac{a_n}{a_{n-1}}, \quad n \geq 1.$$

Los primeros términos de c_n son:

$$1, 2, \frac{3}{2}, \frac{5}{3}, \frac{8}{5}, \frac{13}{8}, \ldots$$

Se tiene $b_1 = 1 - 2 = -1$ y

$$b_{n+1} = a_{n+1}^2 - a_{n+2}a_n = a_{n+1}(a_n + a_{n-1}) - (a_{n+1} + a_n)a_n = a_{n+1}a_{n-1} - a_n^2 = -b_n,$$

luego

$$b_n = (-1)^n \quad \forall n \in \mathbb{N},$$

que no es convergente (no tiene límite), oscilando entre 1 y -1. Además,

$$c_{n+1} - c_n = \frac{a_{n+1}}{a_n} - \frac{a_n}{a_{n-1}} = \frac{-b_n}{a_n a_{n-1}} = \frac{(-1)^{n+1}}{a_n a_{n-1}} \longrightarrow 0,$$

y también

$$c_{n+2} - c_n = c_{n+2} - c_{n+1} + c_{n+1} - c_n = \frac{(-1)^{n+2}}{a_{n+1}a_n} + \frac{(-1)^{n+1}}{a_n a_{n-1}} \begin{cases} < 0 & \text{si } n \text{ es par,} \\ > 0 & \text{si } n \text{ es impar.} \end{cases}$$

Por tanto, la sucesión (c_{2n}) es decreciente, la sucesión (c_{2n-1}) es creciente, luego ambas tienen límite, y la distancia entre pares e impares tiende a 0. Deducimos que tanto pares como impares convergen al mismo límite, luego existe $L = \lim_{n\to\infty} c_n$.

Además, de $a_{n+1} = a_n + a_{n-1}$ deducimos $c_{n+1} = 1 + \dfrac{1}{c_n}$. Al tomar límites obtenemos $L = 1 + \dfrac{1}{L}$ y concluimos que $L = \dfrac{1+\sqrt{5}}{2} \simeq 1'61803..$, pues L es positivo. Este número L se conoce como *razón áurea* y aparece con frecuencia en el arte.

Ejemplo. Calculemos $\lim_{n\to\infty} \dfrac{5^n}{n!}$, donde aparece la indeterminación $\dfrac{\infty}{\infty}$. Para ello notemos que para todo $n \geq 5$:

$$\frac{5^n}{n!} = \frac{5}{n}\frac{5}{n-1} \cdots \frac{5}{4}\frac{5}{3}\frac{5}{2}\frac{5}{1} \leq \frac{5}{n}\frac{5^4}{4!}$$

Así, para todo $n \geq 5$,

$$0 \leq \frac{5^n}{n!} \leq \frac{5}{n}\frac{625}{24},$$

y por el Teorema del Sandwich, $\lim_{n\to\infty} \dfrac{5^n}{n!} = 0$.

§ 4. Límites de funciones

Dada una función $f : I \longrightarrow \mathbb{R}$ y un punto c que o bien está en I o es un punto frontera de I, decimos que el límite de f cuando x tiende hacia c es L si $f(x)$ se acerca a L cuando x se acerca a c (siendo siempre $x \neq c$).

Escribiremos $\lim_{x\to c} f(x) = L$ o $f(x) \xrightarrow[x\to c]{} L$.

Formalmente, esto se concreta así:

- $\lim_{x\to c} f(x) = L \in \mathbb{R}$ si, para todo intervalo J alrededor de L, existe un intervalo $\tilde{J}$ alrededor de c tal que $f(x) \in J\ \forall x \in \tilde{J} \cap I \setminus \{c\}$; equivalentemente si $\forall \epsilon > 0\ \exists \delta > 0$ tal que $|f(x) - L| < \epsilon$ si $0 < |x - c| < \delta$, $x \in I$.
- $\lim_{x\to c} f(x) = \pm\infty$ (respectivamente $-\infty$) si $\forall M > 0\ \exists \delta > 0$ tal que $f(x) > M$ (resp. $f(x) < -M$) si $0 < |x - c| < \delta$, $x \in I$.
- $\lim_{x\to\pm\infty} f(x) = L$ si $\forall \epsilon > 0\ \exists M > 0$ tal que $|f(x) - L| < \epsilon$ para todo $x \in I$ con $x > M$ (o $x < -M$).
- $\lim_{x\to\pm\infty} f(x) = \pm\infty$ si ... (¡ejercicio!)

Propiedades.

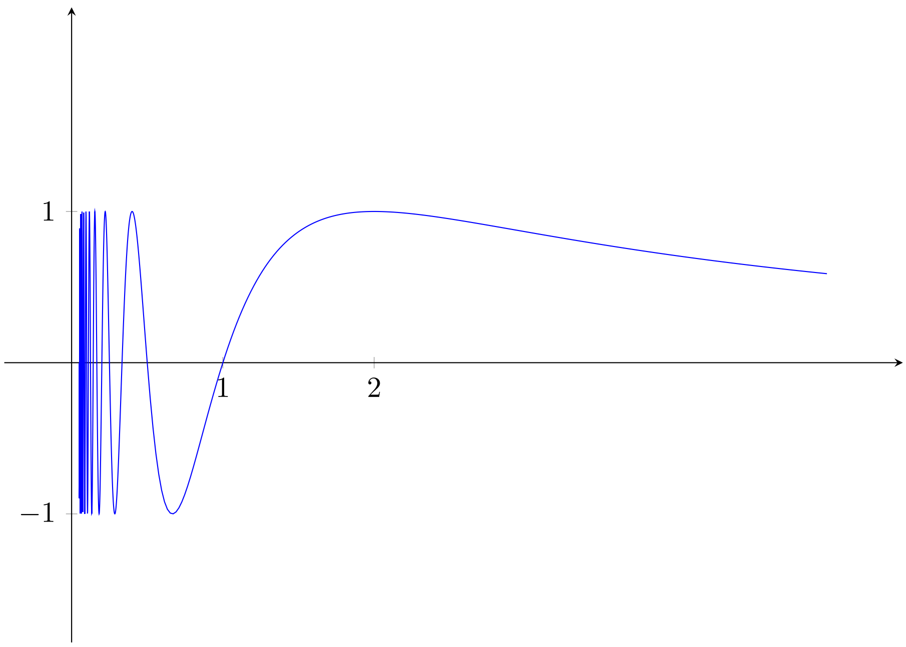
- $\lim_{x\to c} f(x) = L$ (c y L pueden ser $\pm\infty$) si para toda sucesión (x_n) con $x_n \in I \setminus \{c\}\ \forall n$) tal que $x_n \longrightarrow c$ se tiene $f(x_n) \longrightarrow L$.
- Si existen los límites $\lim_{x\to c} f(x)$ y $\lim_{x\to c} g(x)$ y $a \in \mathbb{R}$, entonces:
  - $\lim_{x\to c} af(x) = a \lim_{x\to c} f(x)$,
  - $\lim_{x\to c}\big(f(x) + g(x)\big) = \lim_{x\to c} f(x) + \lim_{x\to c} g(x)$ (salvo la indeterminación $\infty - \infty$),
  - $\lim_{x\to c}\Big(f(x)g(x)\Big) = \lim_{x\to c} f(x) \lim_{x\to c} g(x)$ (salvo $0 \cdot \infty$),
  - $\lim_{x\to c} \dfrac{f(x)}{g(x)} = \dfrac{\lim_{x\to c} f(x)}{\lim_{x\to c} g(x)}$ (salvo $\frac{0}{0}$ o $\frac{\infty}{\infty}$).
- **Teorema del Sandwich:** Si existe $\delta > 0$ tal que $f(x) \leq h(x) \leq g(x)\ \forall x \in I$ con $0 < |x-c| < \delta$ y $\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = L$, entonces $\lim_{x\to c} h(x) = L$.

Ejemplo. Consideremos la función:

$$\begin{aligned} f : (0, +\infty) &\longrightarrow \mathbb{R} \\ x &\mapsto \operatorname{sen} \frac{\pi}{x}. \end{aligned}$$

Notemos que $\dfrac{1}{n} \longrightarrow 0$, y $f\left(\dfrac{1}{n}\right) = \operatorname{sen}(n\pi) = 0 \longrightarrow 0$. Sin embargo $\dfrac{2}{4n+1} \longrightarrow 0$ también, pero $f\left(\dfrac{2}{4n+1}\right) = \operatorname{sen}\left((4n+1)\dfrac{\pi}{2}\right) = 1 \longrightarrow 1$ y $\dfrac{2}{4n+3} \longrightarrow 0$ y $f\left(\dfrac{2}{4n+3}\right) = \operatorname{sen}\left((4n+3)\dfrac{\pi}{2}\right) = -1 \longrightarrow -1$. Concluimos que no existe $\lim_{x\to 0} f(x)$.

En este caso, al tender x hacia 0, $f(x)$ oscila cada vez más deprisa entre -1 y 1:



También podemos hablar de límites por la derecha o por la izquierda:

- $\lim_{x\to c^+} f(x) = L$ si $\forall \epsilon > 0\ \exists \delta > 0$ tal que $|f(x) - L| < \epsilon\ \forall x \in (c, c+\delta)$,
- $\lim_{x\to c^-} f(x) = L$ si $\forall \epsilon > 0\ \exists \delta > 0$ tal que $|f(x) - L| < \epsilon\ \forall x \in (c-\delta, c)$,

y de modo parecido si $L = \pm\infty$.

Si c es un punto interior de I, para que exista $\lim_{x\to c} f(x)$ han de existir los dos límites laterales $\lim_{x\to c^+} f(x)$ y $\lim_{x\to c^-} f(x)$ y ser iguales.

§ 5. Continuidad

Dada una función $f : I \longrightarrow \mathbb{R}$ y un punto c,

- f se dice *continua* en c si $c \in I$ y $\lim_{x\to c} f(x) = f(c)$ (esto es, existe el límite y coincide con el valor $f(c)$).
- f se dice *discontinua* en c en otro caso.

  En el caso en que f sea discontinua en c pero exista $\lim_{x\to c} f(x)$ y sea finito, se dice que la discontinuidad es *evitable* pues podemos redefinir $f(c)$ como el valor del límite.
- Una función se dice *continua* si lo es en todo su dominio de definición.

Ejemplo. La función

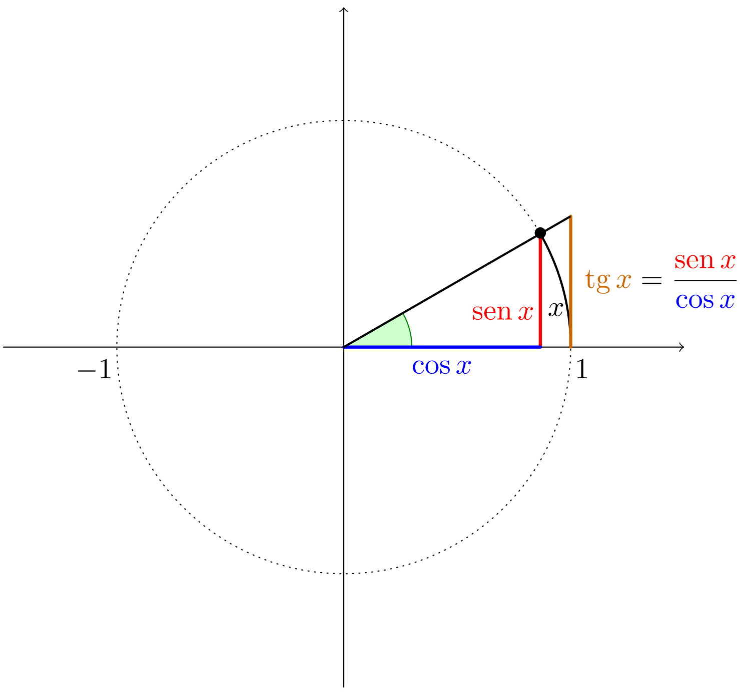
$$\begin{aligned} [0,+\infty) &\longrightarrow \mathbb{R} \\ x &\mapsto \lfloor x \rfloor \quad \text{(parte entera de } x\text{)}, \end{aligned}$$

es discontinua en los números naturales, pues $\lim_{x\to n^-} \lfloor x \rfloor = n-1 \neq n = \lim_{x\to n^+} \lfloor x \rfloor$.

- Las funciones $\operatorname{cotg} x$, $\operatorname{sen} \dfrac{\pi}{x}$ son discontinuas en $x = 0$.
- La función

$$\begin{aligned} \mathbb{R}\setminus\{0\} &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{\operatorname{sen} x}{x}, \end{aligned}$$

verifica $\lim_{x\to 0} \dfrac{\operatorname{sen} x}{x} = 1$. En efecto,



para valores de x pequeños y positivos, el ángulo (medido en radianes), su seno y tangente, satisfacen

$$\operatorname{sen} x \leq x \leq \operatorname{tg} x$$

de donde obtenemos

$$\cos x \leq \frac{\operatorname{sen} x}{x} \leq 1,$$

que vale también para ángulos negativos (pequeños) por ser $\cos x$ par y $\operatorname{sen} x$ y x impares.

Como $\lim_{x\to 0} \cos x = 1$[2] deducimos, por el Teorema del Sandwich,

$$\lim_{x\to 0} \frac{\operatorname{sen} x}{x} = 1,$$

esto es, la discontinuidad en 0 es evitable y podemos redefinir f, convirtiéndola en una función continua:

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} \dfrac{\operatorname{sen} x}{x} & \text{si } x \neq 0, \\ 1 & \text{si } x = 0. \end{cases} \end{aligned}$$

Propiedades.

- Una función $f : I \to \mathbb{R}$ es continua en $c \in I$ si para toda sucesión (x_n) en I con $x_n \longrightarrow c$ se tiene $f(x_n) \longrightarrow f(c)$.
- La suma, resta, producto y cociente de funciones continuas son continuas, salvo en los 'ceros del denominador'.
- Las funciones polinomiales, racionales, exponenciales, logarítmicas, trigonométricas y trigonométricas hiperbólicas son continuas en sus dominios.
- La composición de funciones continuas es continua, esto es, si $g(x)$ es continua en c y $f(x)$ es continua en $g(c)$, entonces $f \circ g(x)$ es continua en c:

$$g(x) \xrightarrow[x\to c]{} g(c), \; f(y) \xrightarrow[y\to g(c)]{} f(g(c)) \quad \Longrightarrow \quad f(g(x)) \xrightarrow[x\to c]{} f(g(c)).$$

- **Teorema de Bolzano** (o del *valor intermedio*): Si $f : [a, b] \to \mathbb{R}$ es continua y $L \in \mathbb{R}$ verifica que $f(a) < L < f(b)$, o $f(a) > L > f(b)$, entonces existe $c \in (a, b)$ tal que $f(c) = L$.

Ejemplo. Aproximemos la raiz del polinomio $f(x) = x^3 + 3x + 2$.

$$f(0) = 2 > 0, \quad f(-1) = -1 - 3 + 2 = -2 < 0,$$

luego $f(x)$ tiene un cero entre -1 y 0 por el Teorema de Bolzano. Ahora,

$$f\left(\frac{-1}{2}\right) = -\frac{1}{8} - \frac{3}{2} + 2 > 0,$$

luego $f(x)$ tiene un cero entre -1 y $\dfrac{-1}{2}$, ...

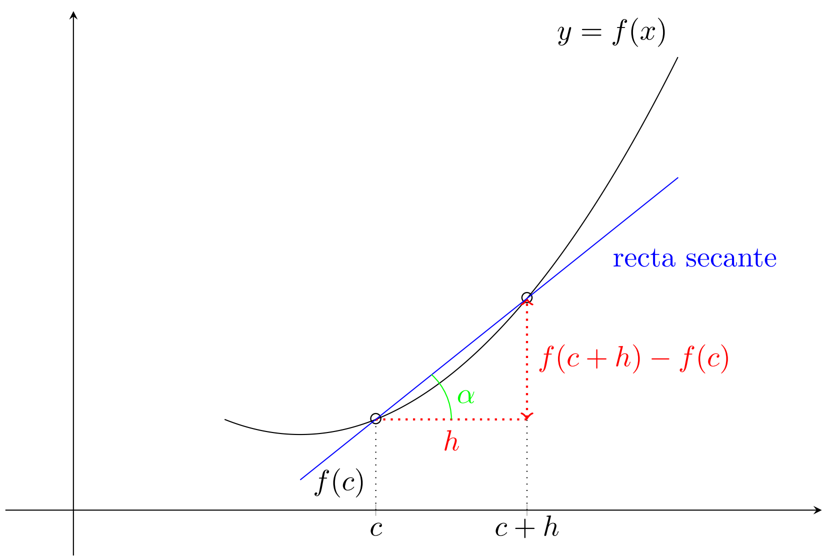
[2] Para x pequeño, $1 - |x| \leq \cos x \leq 1$.

§ 6. Derivabilidad

La razón de cambio de una función $f : I \to \mathbb{R}$ entre los puntos c y $c+h$ es

$$\frac{f(c+h)-f(c)}{h}$$

y esta razón de cambio nos da la tangente del ángulo que forma la recta secante a la gráfica de f en los puntos $(c, f(c))$ y $(c+h, f(c+h))$ con el eje de las abscisas.



La *derivada* de f en c es la *razón instantánea de cambio*:

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

que también se denota por $\dfrac{df}{dx}(c)$ o $\left.\dfrac{d}{dx}\right|_{x=c}(f)$.

Este límite puede no existir. Si existe y es $\neq \pm\infty$, decimos que f es *derivable* en c. En este caso $f'(c)$ es la tangente del ángulo que forma la recta tangente (límite de las secantes) a la gráfica en el punto $(c, f(c))$ con el eje de las abscisas. Dicha recta tangente tiene, por tanto, como ecuación:

$$y - f(c) = f'(c)(x-c).$$

La función f se dice *derivable* si lo es en todos los puntos de su dominio.

Si $f : I \longrightarrow \mathbb{R}$ es derivable en todos los puntos de un subintervalo $\tilde{I} \subseteq I$, entonces consideraremos la función derivada $f' : \tilde{I} \longrightarrow \mathbb{R}$. Esta función puede ser a su vez derivable en un subintervalo, pudiendo considerar las *derivadas sucesivas* $f^{(2)} = f''$, $f^{(3)} = \left(f^{(2)}\right)'$, ...

Propiedades.

- Si f es derivable en c, f es continua en c. En efecto,

$$\lim_{x\to c} f(x) = \lim_{x\to c}\left(\frac{f(x)-f(c)}{x-c}(x-c)+f(c)\right) = f(c),$$

  pues

$$\lim_{x\to c}\left(\frac{f(x)-f(c)}{x-c}(x-c)\right) = \lim_{x\to c}\frac{f(x)-f(c)}{x-c}\lim_{x\to c}(x-c) = f'(c)\cdot 0 = 0.$$

- Si f y g son derivables en x y $a \in \mathbb{R}$:
  - af es derivable en x y $\big(af\big)'(x) = af'(x)$,
  - $f+g$ es derivable en x y $\big(f+g\big)'(x) = f'(x)+g'(x)$,
  - fg es derivable en x y $\big(fg\big)'(x) = f'(x)g(x)+f(x)g'(x)$, ya que

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

  - Si $g(x)\neq 0$, entonces $\dfrac{f}{g}$ es derivable en x y $\left(\dfrac{f}{g}\right)'(x) = \dfrac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$.

- **Regla de la cadena:** Si f es derivable en $g(x)$ y g es derivable en x, entonces la composición $f\circ g$ es derivable en x y $(f\circ g)'(x) = f'\big(g(x)\big)g'(x)$.

  En efecto, se tiene

$$\lim_{x\to c}\frac{f\big(g(x)\big)-f\big(g(c)\big)}{x-c} = \lim_{x\to c}\varphi(x)\,\frac{g(x)-g(c)}{x-c}$$

  donde la función φ, definida en un intervalo alrededor de c y continua en c, viene dada por

$$\varphi(x) = \begin{cases} \dfrac{f\big(g(x)\big)-f\big(g(c)\big)}{g(x)-g(c)} & \text{si } g(x)\neq g(c),\\ f'\big(g(c)\big) & \text{si } g(x)=g(c),\end{cases}$$

  luego

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

Las derivadas de las funciones más usadas aparecen en el Cuadro 2.1

Cuadro 2.1: Tabla de derivadas

$f(x) = k$	$f'(x) = 0$
$f(x) = x$	$f'(x) = 1$
$f(x) = x^a$	$f'(x) = ax^{a-1},\ a \in \mathbb{R}$
$f(x) = \ln x$	$f'(x) = \dfrac{1}{x}$
$f(x) = \lg_a x$	$f'(x) = \dfrac{1}{x \ln a}$
$f(x) = e^x$	$f'(x) = e^x$
$f(x) = a^x$	$f'(x) = a^x \ln a$
$f(x) = g(x)^{h(x)}$	$f'(x) = h(x)g(x)^{h(x)-1}g'(x) + g(x)^{h(x)}h'(x) \ln g(x)$
$f(x) = \operatorname{sen} x$	$f'(x) = \cos x$
$f(x) = \cos x$	$f'(x) = -\operatorname{sen} x$
$f(x) = \operatorname{tg} x$	$f'(x) = \dfrac{1}{\cos^2 x} = 1 + \operatorname{tg}^2 x$
$f(x) = \operatorname{arc\,sen} x$	$f'(x) = \dfrac{1}{\sqrt{1-x^2}}$
$f(x) = \arccos x$	$f'(x) = \dfrac{-1}{\sqrt{1-x^2}}$
$f(x) = \operatorname{arc\,tg} x$	$f'(x) = \dfrac{1}{1+x^2}$
$f(x) = \operatorname{senh} x$	$f'(x) = \cosh x$
$f(x) = \cosh x$	$f'(x) = \operatorname{senh} x$
$f(x) = \tanh x$	$f'(x) = \dfrac{1}{\cosh^2 x} = 1 - \tanh^2 x$
$f(x) = \mathrm{arcsenh} x$	$f'(x) = \dfrac{1}{\sqrt{1+x^2}}$
$f(x) = \mathrm{argcosh} x$	$f'(x) = \dfrac{1}{\sqrt{x^2-1}}$
$f(x) = \mathrm{argtgh} x$	$f'(x) = \dfrac{1}{1-x^2}$

§ 7. Crecimiento y concavidad

7.1. Extremos

Dada una función $f : I \longrightarrow \mathbb{R}$ y un punto $c \in I$, se dice que f tiene un **máximo** (respectivamente **mínimo**) **absoluto** en c si $f(c) \geq f(x)$ (resp. $f(c) \leq f(x)$) para todo $x \in I$.

Proposición. *Toda función continua $f : [a, b] \to \mathbb{R}$ $(a, b \in \mathbb{R})$ tiene máximos y mínimos absolutos.*

Una función $f : I \longrightarrow \mathbb{R}$ tiene un **máximo** (respectivamente **mínimo**) **relativo** en c si existe $\delta > 0$ tal que $f(c) \geq f(x)$ (resp. $f(c) \leq f(x)$) para todo $x \in I \cap (c - \delta, c + \delta)$.

Proposición. *Si f tiene un extremo (máximo o mínimo) relativo en un punto interior c de I y es derivable en c, entonces $f'(c) = 0$.*

En efecto, si f tiene un máximo relativo en c, la razón de cambio $\dfrac{f(x) - f(c)}{x - c}$ es ≥ 0 a la izquierda de c y ≤ 0 a la derecha de c, luego si existe el límite $f'(c)$, debe de ser igual a 0.

Atención: Un punto interior c con $f'(c) = 0$ (*punto crítico*) no es necesariamente un extremo relativo.

7.2. Teorema del Valor Medio

Teorema (Rolle)**.** *Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función continua en $[a, b]$ y derivable en (a, b) y tal que $f(a) = f(b)$. Entonces existe un $c \in (a, b)$ tal que $f'(c) = 0$.*

Si f es constante, entonces $f'(x) = 0$ para todo x. En otro caso f tiene un máximo y mínimo absolutos distintos, y al menos uno de ellos es un punto interior. Ahora basta aplicar el resultado anterior.

Teorema (del Valor Medio (TVM))**.** *Sean $f, g : [a, b] \longrightarrow \mathbb{R}$ dos funciones continuas en $[a, b]$ y derivables en (a, b). Entonces existe un $c \in (a, b)$ tal que*

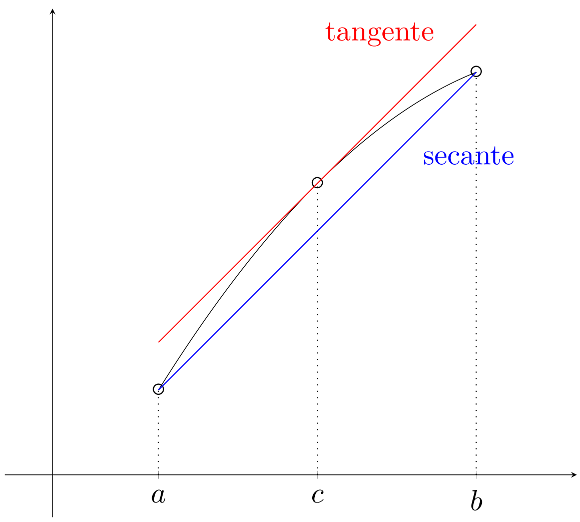
$$f'(c)\big(g(b) - g(a)\big) = g'(c)\big(f(b) - f(a)\big).$$

Basta considerar la función $H(x) = f(x)\big(g(b) - g(a)\big) - g(x)\big(f(b) - f(a)\big)$, que verifica $H(a) = f(a)g(b) - g(a)f(b) = H(b)$ y aplicar el Teorema de Rolle.

En particular, si $f : [a, b] \longrightarrow \mathbb{R}$ es continua en $[a, b]$ y derivable en (a, b), entonces existe un $c \in (a, b)$ tal que

$$f'(c) = \frac{f(b) - f(a)}{b - a},$$

esto es, la recta tangente en c es paralela a la recta secante en a y b.



7.3. Regla de L'Hôpital

Esta regla es consecuencia del Teorema del Valor Medio y es una herramienta muy útil para el cálculo de límites indeterminados de cocientes.

Sean $f, g : I \longrightarrow \mathbb{R}$ dos funciones derivables en un intervalo abierto I. Sea a un punto que es interior o frontera de I, incluyendo la posibilidad $a = \pm\infty$, con

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0 \text{ o } \pm\infty.$$

Entonces, si existe

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

y existe un entorno[3] J de a donde $g'(x) \neq 0 \ \forall a \neq x \in J \cap I$, se tiene también

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

Ejemplos.

- $\lim_{x\to 0} \dfrac{\operatorname{sen} x}{x} = \lim_{x\to 0} \dfrac{\cos x}{1} = 1$. (Indeterminación $\frac{0}{0}$.)

  Escribiremos $\operatorname{sen} x \sim x$ si $x \to 0$ y diremos que $\operatorname{sen} x$ es *equivalente* a x si $x \to 0$.

[3] $J = (a-\delta, a+\delta)$ para algún $\delta > 0$ si $a \neq \pm\infty$, y $J = (r, +\infty)$ (resp. $(-\infty, r)$) para algún $r \in \mathbb{R}$ si $a = +\infty$ (resp. $a = -\infty$).

Si tenemos un límite de la forma $\lim_{x\to a}(\operatorname{sen} f(x))g(x)$, y sabemos que $\lim_{x\to a} f(x) = 0$, entonces

$$\lim_{x\to a}(\operatorname{sen} f(x))g(x) = \lim_{x\to a} \frac{\operatorname{sen} f(x)}{f(x)} f(x)g(x) = \lim_{x\to a} f(x)g(x)$$

ya que $\lim_{x\to a} \dfrac{\operatorname{sen} f(x)}{f(x)} = 1$. Esto es, podemos sustituir $\operatorname{sen} f(x)$ por $f(x)$ *¡cuando aparece como factor!*

- $\lim_{x\to 0} \dfrac{1-\cos x}{x^2} = \lim_{x\to 0} \dfrac{\operatorname{sen} x}{2x} = \dfrac{1}{2}$, luego $\lim_{x\to 0} \dfrac{1-\cos x}{x^2/2} = 1$. Por tanto $1-\cos x \sim \dfrac{x^2}{2}$ si $x \to 0$.

  Análogamente se tiene, si $x \to 0$,

$$x \sim \operatorname{arc\,sen} x \sim \operatorname{tg} x \sim \ln(1+x) \sim e^x - 1.$$

- $\lim_{x\to +\infty} \dfrac{\ln x}{x} = \lim_{x\to +\infty} \dfrac{1/x}{1} = 0$. (Indeterminación $\frac{\infty}{\infty}$.)

- $\lim_{x\to 0^+} x \ln x = \lim_{x\to 0^+} \dfrac{\ln x}{1/x} = \lim_{x\to 0^+} \dfrac{1/x}{-1/x^2} = \lim_{x\to 0^+}(-x) = 0$. (Indeterminación $0\cdot\infty$.)

- $\lim_{x\to 0^+} x^x = \lim_{x\to 0^+} \exp(\ln x^x) = \exp\left(\lim_{x\to 0^+} x\ln x\right) = \exp(0) = 1$. Aquí hemos utilizado que exp es continua. (Indeterminación 0^0.)

- $\lim_{x\to 0} \left(\dfrac{\operatorname{sen} x}{x}\right)^{1/x} = \exp\left(\lim_{x\to 0} \dfrac{\ln(\operatorname{sen} x/x)}{x}\right)$ y de este modo transformamos la indeterminación 1^∞ en una indeterminación $\frac{0}{0}$. Ahora[4]

$$\begin{aligned}
\lim_{x\to 0} \frac{\ln(\operatorname{sen} x/x)}{x} &= \lim_{x\to 0} \frac{\ln(\operatorname{sen} x) - \ln(x)}{x} = \lim_{x\to 0} \frac{\frac{\cos x}{\operatorname{sen} x} - \frac{1}{x}}{1} = \lim_{x\to 0} \frac{x\cos x - \operatorname{sen} x}{x \operatorname{sen} x} \\
&= \lim_{x\to 0} \frac{x\cos x - \operatorname{sen} x}{x^2} \quad \text{utilizando } \operatorname{sen} x \sim x \\
&= \lim_{x\to 0} \frac{\cos x - x\operatorname{sen} x - \cos x}{2x} = \lim_{x\to 0} \frac{-x\operatorname{sen} x}{2x} = 0.
\end{aligned}$$

  luego $\lim_{x\to 0} \left(\dfrac{\operatorname{sen} x}{x}\right)^{1/x} = \exp(0) = 1$.

[4]Alternativamente, usando la equivalencia $\ln(1+x) \sim x$ si $x \to 0$, podemos sustituir $\ln \frac{\operatorname{sen} x}{x}$ por $\frac{\operatorname{sen} x}{x} - 1$ y calcular $\lim_{x\to 0} \dfrac{\frac{\operatorname{sen} x}{x} - 1}{x}$.

7.4. Crecimiento

Una función $f : [a, b] \longrightarrow \mathbb{R}$ se dice **creciente** (resp. **decreciente**) si para cualesquiera $x_1, x_2 \in [a, b]$ con $x_1 < x_2$ se tiene $f(x_1) \leq f(x_2)$ (resp. $f(x_1) \geq f(x_2)$).

Si esta última desigualdad es estricta, diremos que f es **estrictamente creciente** (resp. **estrictamente decreciente**).

Si f es continua en $[a, b]$ y derivable en (a, b), entonces f es creciente (resp. decreciente) si y solo si $f'(x) \geq 0$ (resp. $f'(x) \leq 0$) $\forall x \in (a, b)$. (Esto es consecuencia del Teorema del Valor Medio.)

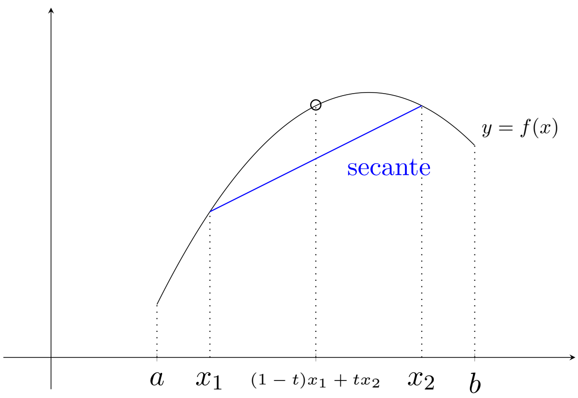
7.5. Concavidad

Una función $f : [a, b] \longrightarrow \mathbb{R}$ se dice **cóncava** (resp. **convexa**) si $\forall x_1 < x_2$ en $[a, b]$, la gráfica de f está por encima (resp. por debajo) del segmento que une $(x_1, f(x_1))$ y $(x_2, f(x_2))$. Así, f es cóncava si $\forall x_1 < x_2$ en $[a, b]$ se tiene

$$f(x_1 + t(x_2 - x_1)) > f(x_1) + t\big(f(x_2) - f(x_1)\big) \quad \forall 0 < t < 1,$$

esto es,

$$f((1-t)x_1 + tx_2) > (1-t)f(x_1) + tf(x_2) \quad \forall 0 < t < 1.$$



Así, si f es cóncava y $a \leq x_1 < x_2 \leq b$ tenemos

$$\frac{f(x_1 + t(x_2 - x_1)) - f(x_1)}{t(x_2 - x_1)}(x_2 - x_1) > f(x_2) - f(x_1).$$

Si f es derivable en x_1, haciendo el límite con $t \to 0$ nos queda ($h = t(x_2 - x_1)$)

$$f'(x_1)(x_2 - x_1) \geq f(x_2) - f(x_1),$$

luego

$$f(x_2) \leq f(x_1) + f'(x_1)(x_2 - x_1),$$

y de hecho la desigualdad es estricta, pues la tangente en x_1 no coincide con la secante. De modo análogo, si f es derivable en x_2, cambiando t por $1 - t$, se tiene $f(x_1) < f(x_2) + f'(x_2)(x_1 - x_2)$. Por tanto, *La gráfica de una función cóncava derivable está por debajo de las rectas tangentes.*[5]

Si la función $f : [a,b] \longrightarrow \mathbb{R}$ es continua en $[a,b]$ y derivable en (a,b), entonces f es cóncava (resp. convexa) si y solo si f' es estrictamente decreciente (resp. creciente) en (a,b).

Si f es dos veces derivable en (a,b) y $f''(x) < 0$ (resp. $f''(x) > 0$) $\forall x \in (a,b)$, entonces f es cóncava (resp. convexa).

Dada una función $f : [a,b] \longrightarrow \mathbb{R}$, un punto $c \in (a,b)$ se dice **punto de inflexión** de f si $\exists \delta > 0$ tal que f es cóncava en $[c-\delta, c]$ y convexa en $[c, c+\delta]$, o al revés.

Si f es dos veces derivable en un punto de inflexión c, entonces $f''(c) = 0$.

Proposición. *Sea $f : [a,b] \longrightarrow \mathbb{R}$ una función continua en $[a,b]$ y dos veces derivable en (a,b), y sea $c \in (a,b)$.*

Si $f'(c) = 0$ y $f''(c) < 0$ (resp. $f''(c) > 0$), entonces c es un máximo (resp. mínimo) relativo de f.

A modo de resumen:

- Si $f'(x) > 0$ (resp. $f'(x) < 0$) $\forall x \in (a,b)$, entonces f es creciente (resp. decreciente) en $[a,b]$.
- Si $f''(x) > 0$ (resp. $f''(x) < 0$) $\forall x \in (a,b)$, entonces f es convexa (resp. cóncava) en $[a,b]$.
- Si $f'(c) = 0$ para algún $c \in (a,b)$, c es candidato a ser extremo relativo. Si además $f''(c) > 0$ (resp. $f''(c) < 0$), entonces c es un mínimo (resp. máximo) relativo.
- Si $f''(c) = 0$ para algún $c \in (a,b)$, c es candidato a ser un punto de inflexión.

7.6. Ejemplos

Ejemplo. $f(x) = \dfrac{2x}{1+x^2}$, $x \in \mathbb{R}$.

Esta función es impar ($f(-x) = -f(x)$ $\forall x \in \mathbb{R}$), luego bastaría estudiarla en el intervalo $[0, +\infty)$.

$$f'(x) = \frac{2(1+x^2) - (2x)^2}{(1+x^2)^2} = \frac{2(1-x^2)}{(1+x^2)^2},$$

[5] Y al revés, si la gráfica de f está por debajo de las rectas tangentes, para todo $x_0 \neq x_1$ en $[a,b]$ y $0 < t < 1$, sea $x_t = (1-t)x_0 + tx_1$. Entonces $f(x_0) < f(x_t) + f'(x_t)(x_0 - x_t)$ y $f(x_1) < f(x_t) + f'(x_t)(x_1 - x_t)$, pues tanto $f(x_0)$ como $f(x_1)$ están por debajo de la recta tangente a f en x_t. De aquí deducimos

$$(1-t)f(x_0) + tf(x_1) < f(x_t) + f'(x_t)\big((1-t)(x_0 - x_t) + t(x_1 - x_t)\big) = f(x_t),$$

luego obtenemos

$$(1-t)f(x_0) + tf(x_1) < f\big((1-t)x_0 + tx_1\big)$$

y f es cóncava.

luego

$$f'(x) \text{ es } \begin{cases} < 0 & \text{en } (-\infty,-1),\ f \text{ decrece,} \\ = 0 & \text{en } -1, \text{ mínimo relativo,} \\ > 0 & \text{en } (-1,1),\ f \text{ crece,} \\ = 0 & \text{en } 1, \text{ máximo relativo,} \\ < 0 & \text{en } (1,+\infty)\ f \text{ decrece.} \end{cases}$$

Además,

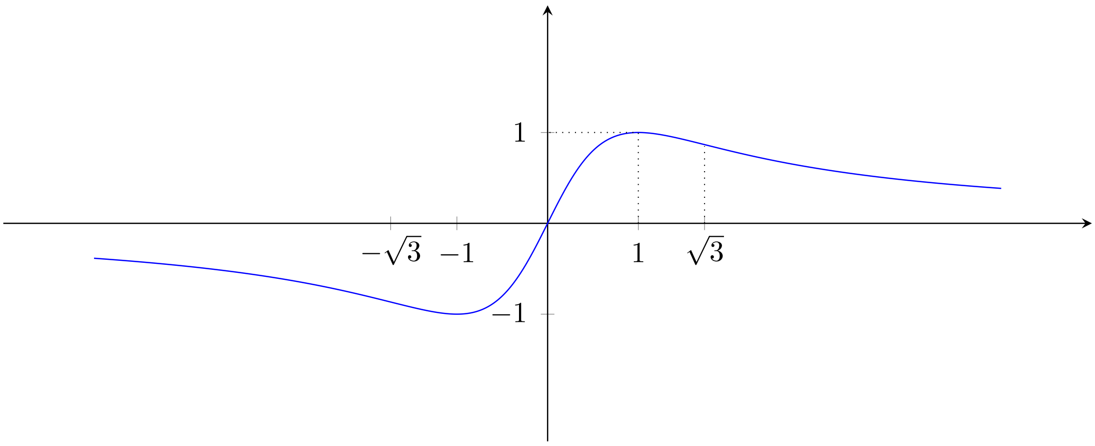
$$\lim_{x\to+\infty} \frac{2x}{1+x^2} = \lim_{x\to+\infty} \frac{\frac{2}{x}}{\frac{1}{x^2}+1} = \frac{0}{0+1} = 0.$$

Decimos entonces que $y=0$ es una *asíntota horizontal.*

$$f''(x) = 2\frac{-2x(1+x^2)^2-(1-x^2)4x(1+x^2)}{(1+x^2)^4} = 4\frac{-x(1+x^2)-2x(1-x^2)}{(1+x^2)^3} = \frac{-4x(3-x^2)}{(1+x^2)^3},$$

así,

$$f''(x) \text{ es } \begin{cases} < 0 & \text{en } (-\infty,-\sqrt{3}) \text{ (cóncava),} \\ = 0 & \text{en } -\sqrt{3} \text{ (punto de inflexión),} \\ > 0 & \text{en } (-\sqrt{3},0) \text{ (convexa),} \\ = 0 & \text{en } 0 \text{ (punto de inflexión),} \\ < 0 & \text{en } (0,\sqrt{3}) \text{ (cóncava),} \\ = 0 & \text{en } \sqrt{3} \text{ (punto de inflexión),} \\ > 0 & \text{en } (\sqrt{3},\infty) \text{ (convexa).} \end{cases}$$



El punto -1 es el mínimo absoluto de f y 1 es el máximo absoluto.

Ejemplo. $f:(0,+\infty)\longrightarrow\mathbb{R}$, $f(x)=\dfrac{\ln x}{x}$.

Estudiamos primero su comportamiento en los ‘extremos’: al tender x a 0 y a $+\infty$:

$$\lim_{x\to0^+} \frac{\ln x}{x} = \lim_{x\to0^+} \frac{1}{x}\ln x = -\infty,$$

luego $x = 0$ es una asíntota vertical. Además,

$$\lim_{x\to+\infty} \frac{\ln x}{x} = 0$$

(basta aplicar la Regla de L'Hôpital, pero se obtiene también por el Teorema del Sandwich teniendo en cuenta que $e^x \geq x^2$ $\forall x \geq 0$, luego $\ln x \leq \sqrt{x}$ y $0 \leq \frac{\ln x}{x} \leq \frac{1}{\sqrt{x}}$) y, por tanto, $y = 0$ es una asíntota horizontal.

Ahora,

$$f'(x) = \frac{1 - \ln x}{x^2},$$

luego

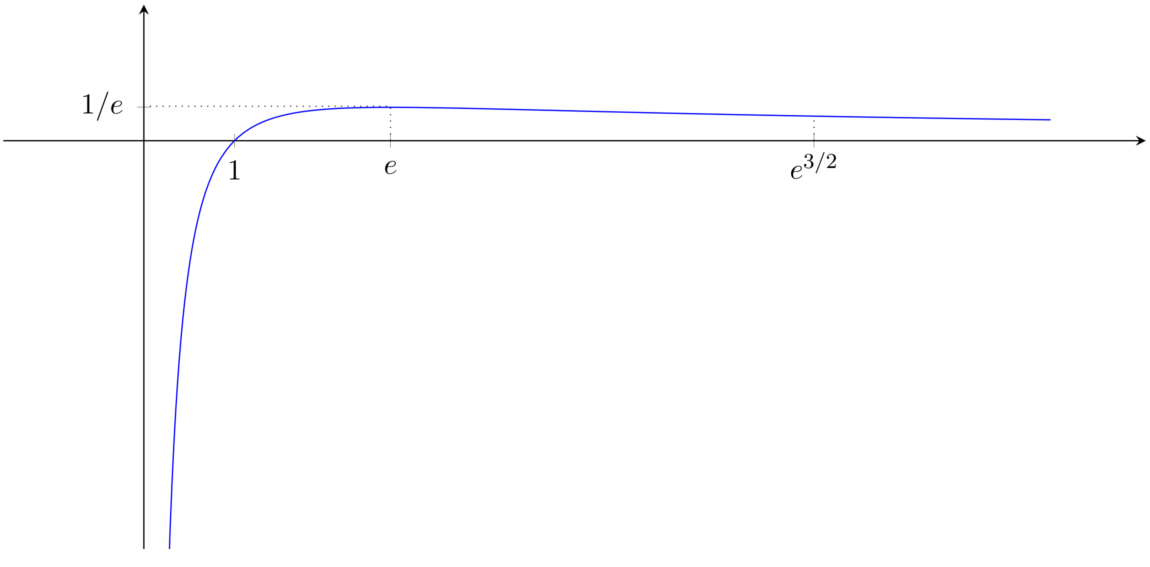
$$f'(x) \text{ es } \begin{cases} > 0 & \text{en } (0, e) \text{ (creciente)}, \\ = 0 & \text{en } e \text{ (máximo global)}, \\ < 0 & \text{en } (e, +\infty) \text{ (decrece)}. \end{cases}$$

Además,

$$f''(x) = \frac{-3 + 2\ln x}{x^3},$$

y, por tanto,

$$f''(x) \text{ es } \begin{cases} < 0 & \text{en } (0, e^{3/2}) \text{ (cóncava)}, \\ = 0 & \text{en } e^{3/2} \text{ (punto de inflexión)}, \\ > 0 & \text{en } (e^{3/2}, +\infty) \text{ (convexa)}. \end{cases}$$



Nota. En ocasiones aparecen asíntotas oblicuas: la recta $y = ax + b$ $(a \neq 0)$ es *asíntota oblicua* de la función $f(x)$ si

$$\lim_{x\to\pm\infty} \Big(f(x) - (ax + b) \Big) = 0.$$

Esto fuerza $a = \lim_{x\to\pm\infty} \dfrac{f(x)}{x}$ y $b = \lim_{x\to\pm\infty}\Big(f(x) - ax\Big)$.

Capítulo 3

Aproximación lineal y polinómica

§ 1. Aproximación lineal

1.1. Recta tangente

Si $f : I \longrightarrow \mathbb{R}$ es una función definida en un intervalo I y derivable en un punto a interior de I, la *recta tangente* a la gráfica de f en el punto $\big(a, f(a)\big)$ es la recta de ecuación

$$\boxed{y = f(a) + f'(a)(x - a)\,.}$$

Como $\lim_{x\to a} \dfrac{f(x) - f(a) - f'(a)(x-a)}{x-a} = 0$, el *error*

$$\varepsilon(x) := f(x) - f(a) - f'(a)(x-a)$$

cometido al aproximar $f(x)$ por $f(a) + f'(a)(x-a)$ es pequeño comparado con $x - a$ cuando x está próximo al punto a:

$$\lim_{x\to a} \frac{\varepsilon(x)}{x-a} = 0.$$

Si f es derivable dos veces en el interior de I y $b \in I$, por el Teorema del Valor Medio existe c entre a y b tal que

$$\boxed{\varepsilon(b) = f(b) - f(a) - f'(a)(b-a) = \frac{f''(c)}{2}(b-a)^2.}$$

En efecto, sea

$$F(x) = f(b) - f(x) - f'(x)(b-x) \quad \text{y} \quad G(x) = (b-x)^2\,.$$

Se tiene $F(b) = G(b) = 0$, luego por el Teorema del Valor Medio

$$\frac{F(a)}{G(a)} = \frac{F(b) - F(a)}{G(b) - G(a)} = \frac{F'(c)}{G'(c)} = \frac{-f''(c)(b-c)}{-2(b-c)}$$

para un c entre a y b. Así obtenemos

$$\varepsilon(b) = F(a) = \frac{f''(c)}{2}(b-a)^2.$$

Si f no es derivable pero sí continua en a y $\lim_{x\to a} \dfrac{f(x)-f(a)}{x-a} = \pm\infty$, entonces la recta tangente es vertical: $x = a$.

1.2. Método de Newton (o de las tangentes)

No siempre es posible resolver de forma exacta ecuaciones de la forma

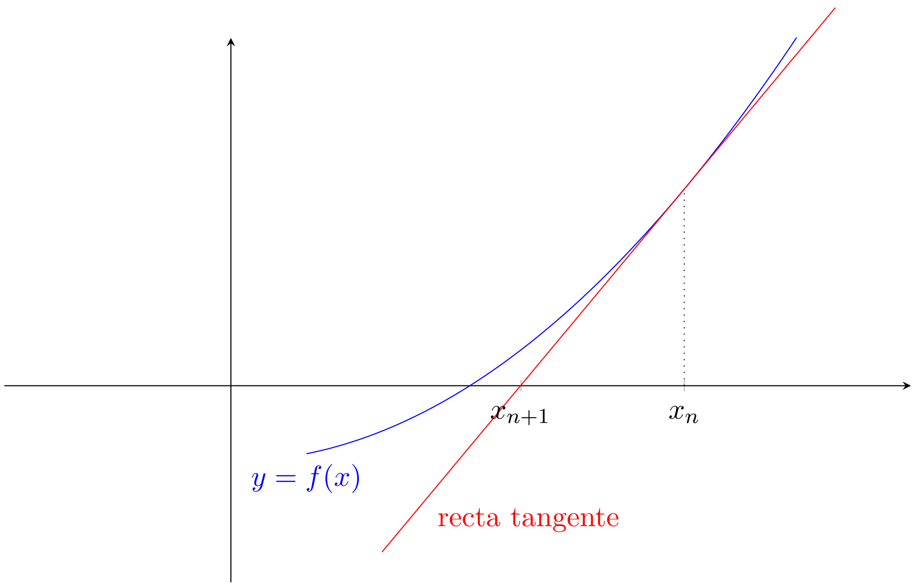
$$f(x) = 0\,,$$

pero es suficiente encontrar buenas aproximaciones a las soluciones.

Sea $f : I \longrightarrow \mathbb{R}$ derivable en el intervalo I, sea x_0 una primera aproximación a una solución de la ecuación $f(x) = 0$. El *Método de Newton* obtiene una solución aproximada mediante la siguiente iteración:

$$\boxed{x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}}$$

Esto es, x_{n+1} es el punto de corte de la recta tangente a f en x_n con el eje OX.



Bajo determinadas condiciones, la sucesión (x_n) converge a una solución de $f(x) = 0$ con convergencia muy rápida.

Ejemplo. Sea $f(x) = x^2 - 3$ y $x_0 = 2$ (nuestra primera aproximación a $\sqrt{3}$). El método de Newton nos da el esquema iterativo

$$x_{n+1} = x_n - \frac{x_n^2 - 3}{2x_n} = \frac{x_n}{2} + \frac{3}{2x_n}, \qquad n = 0, 1, 2, \ldots$$

Los primeros términos son:

$$\begin{aligned}
n = 0: &\quad x_0 = 2, \\
n = 1: &\quad x_1 = 1 + \frac{3}{4} = 1,75, \\
n = 2: &\quad x_2 = \frac{7}{8} + \frac{6}{7} = 1,7321479.. \quad \text{¡¡error} < 10^{-4}\text{!!}
\end{aligned}$$

§ 2. Aproximación polinomial

2.1. Interpolación

Si $(a_0, b_0), \dots, (a_n, b_n)$ son $n+1$ puntos en $\mathbb{R}^2$ con los a_i's distintos, existe un único polinomio $p(x)$ de grado $\le n$ tal que $p(a_0) = b_0$, $p(a_1) = b_1$, ..., $p(a_n) = b_n$. Dicho polinomio se llama **polinomio de interpolación**.

En efecto, $\forall j = 0, \dots, n$, consideremos el polinomio de grado n:

$$h_j(x) = \prod_{i \neq j} \frac{x - a_i}{a_j - a_i}$$

que verifica

$$h_j(a_j) = 1, \quad h_j(a_i) = 0 \text{ si } i \neq j.$$

Así el polinomio

$$\boxed{p(x) := \sum_{j=0}^{n} b_j h_j(x)}$$

tiene grado $\le n$ y verifica $p(a_j) = b_j$, $\forall j = 0, 1, \dots, n$.

Además, si $p(x)$ y $q(x)$ son dos polinomios de grado $\le n$ con $p(a_j) = b_j = q(a_j)$ $\forall j = 0, 1, \dots, n$, entonces su diferencia $d(x) = p(x) - q(x)$ tiene grado $\le n$ y, al menos, $n+1$ raíces, ya que $d(a_j) = 0$ $\forall j$. Por tanto $d(x) = 0$.

2.2. Polinomio de Taylor

Dado un polinomio $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ ($n \in \mathbb{N}$, $a_0, \dots, a_n \in \mathbb{R}$), y un número real $a \in \mathbb{R}$, podemos expresar $p(x)$ como combinación de potencias de $x - a$:

$$\begin{aligned}
p(x) &= a_0 + a_1\big((x-a)+a\big) + a_2\big((x-a)+a\big)^2 + \cdots + a_n\big((x-a)+a\big)^n \\
&\quad \text{desarrollamos los binomios ...} \\
&= b_0 + b_1(x-a) + b_2(x-a)^2 + \cdots + b_n(x-a)^n
\end{aligned}$$

En esta situación

$$
\begin{array}{llll}
b_0 = p(a), & & p'(x) = b_1 + 2b_2(x-a) + \cdots + nb_n(x-a)^{n-1}, & \\
b_1 = p'(a), & & p''(x) = 2b_2 + 3\cdot 2b_3(x-a) + \cdots + n(n-1)b_n(x-a)^{n-2}, & \\
b_2 = \frac{1}{2}p''(a), & & \cdots \quad \cdots & \\
\vdots \quad \vdots & & \vdots & \ddots
\end{array}
$$

Obtenemos

$$b_m = \frac{p^{(m)}(a)}{m!} \quad \forall m = 0, 1, \ldots, n.$$

Si $f : I \longrightarrow \mathbb{R}$ es una función derivable n veces en $a \in I$ (luego f ha de ser derivable $n-1$ veces en un entorno $(a-\delta, a+\delta)$), se llama **Polinomio de Taylor** de grado n de $f(x)$ en a al polinomio:

$$\boxed{P_n(x) := f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n\,.}$$

Esto es, $P_n(x)$ *es el polinomio de grado* $\leq n$ *tal que* $f^{(m)}(a) = P_n^{(m)}(a)$, $\forall m = 0, 1, \ldots, n$.

Nota. $P_1(x)$ nos da la recta tangente a $f(x)$ en el punto a.

Ejemplo. $f(x) = e^x$, $a = 0$, entonces

$$P_n(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} = \sum_{i=0}^{n} \frac{x^i}{i!}.$$

La diferencia (o “error”)

$$R_n(x) := f(x) - P_n(x)$$

se dice **Resto de Taylor**, y la expresión

$$f(x) = P_n(x) + R_n(x)$$

se dice **Fórmula de Taylor**.

Propiedades.

- $\lim_{x\to a} \dfrac{f(x) - P_n(x)}{(x-a)^n} = \lim_{x\to a} \dfrac{R_n(x)}{(x-a)^n} = 0$, luego $R_n(x)$ es “pequeño” comparado con $(x-a)^n$ en las cercanías de a. De hecho, $P_n(x)$ es el único polinomio de grado $\leq n$ con esta propiedad.

  Para probar esta propiedad basta aplicar la regla de L’Hôpital reiteradamente.

- Si f es derivable n veces con derivadas continuas en I y existe la derivada $f^{(n+1)}$ en el interior de I, y $a, b \in I$, entonces existe c entre a y b tal que
$$R_n(b) = \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}.$$
Esto es consecuencia del Teorema del Valor Medio.

Ejemplos.

- $f(x) = \operatorname{sen} x$, $a = 0$ y $n = 8$. Como $f'(x) = \cos x$, $f''(x) = -\operatorname{sen} x$, ..., se tiene:
$$P_8(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!}\,.$$
Además $R_8(x) = \dfrac{\cos c}{9!}x^9$, con c entre 0 y x, luego
$$|\operatorname{sen} x - P_8(x)| = |R_8(x)| \leq \frac{|x|^9}{9!}.$$
En general,
$$P_{2n+1}(x) = P_{2n+2}(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!}\,.$$

- $f(x) = \cos x$:
$$P_{2n}(x) = P_{2n+1}(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - + \cdots + (-1)^n \frac{x^{2n}}{(2n)!}\,.$$

- $f(x) = (1+x)^\alpha$. En este caso, definiendo los *números combinatorios generalizados*
$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}\,,$$
se tiene
$$P_n(x) = 1 + \alpha x + \binom{\alpha}{2}x^2 + \cdots + \binom{\alpha}{n}x^n\,.$$

Polinomio de Taylor, máximos, mínimos, concavidad.

Si $f : I \longrightarrow \mathbb{R}$ es una función $n+1$ veces derivable en un entorno de un punto interior $a \in I$, con derivadas continuas. Entonces

- Si n es impar, $f^{(i)}(a) = 0$ $\forall i = 1, \ldots, n$ y $f^{(n+1)}(a) > 0$ (resp. $f^{(n+1)}(a) < 0$), entonces f tiene un mínimo (resp. máximo) relativo en a.
En efecto, si $f^{(n+1)}(a) > 0$, por continuidad existe un entorno $(a-\delta, a+\delta)$ con $f^{(n+1)}(x) > 0$ $\forall x \in (a-\delta, a+\delta)$, luego la fórmula de Taylor nos da
$$f(x) = f(a) + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$
con c entre a y x. Así, $f(x) > f(a)$ $\forall a \neq x \in (a-\delta, a+\delta)$.

- Si n es par, $f^{(i)}(a) = 0$ $\forall i = 2, \dots, n$ y $f^{(n+1)}(a) \neq 0$, entonces a es un punto de inflexión, pues en un entorno de a

$$f(x) = f(a) + f'(a)(x-a) + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

luego la diferencia entre $f(x) - (f(a) + f'(a)(x-a))$ (diferencia entre la gráfica y la tangente en el punto a) cambia de signo al pasar de $x < a$ a $x > a$.

- Si n es impar, $f^{(i)}(a) = 0$ $\forall i = 2, \dots, n$ y $f^{(n+1)}(a) > 0$ (resp. $f^{(n+1)}(a) < 0$), entonces f es convexa (resp. cóncava) en el punto a.

Ejercicio. La función $f : (0, 2) \longrightarrow \mathbb{R}$ es 5 veces derivable con derivadas continuas y verifica

$$f(1) = 3, \quad f'(1) = f''(1) = f'''(1) = f^{(4)}(1) = 0, \quad f^{(5)}(1) > 0.$$

¿Tiene f un máximo o mínimo relativo en 1?

§ 3. Series de Taylor

3.1. Series

Nos planteamos aquí qué se entiende por la suma de los elementos de una sucesión.

Dada una sucesión $(a_n)_{n\in\mathbb{N}}$ en $\mathbb{R}$, se llama **serie** asociada a esta sucesión a la nueva sucesión (s_n) (llamada *sucesión de sumas parciales*) con

$$s_n = a_1 + \cdots + a_n$$

para todo $n \in \mathbb{N}$. Esta serie se denota $\sum_{n=1}^{\infty} a_n$ (o $\sum_{n=0}^{\infty} a_n$ si comenzamos con una sucesión $(a_n)_{n\geq 0}$).

- La serie $\sum_{n=1}^{\infty} a_n$ se dice **convergente** si existe el límite $\lim_{n\to\infty} s_n$. Si este límite es L, escribimos $\sum_{n=1}^{\infty} a_n = L$.

- La serie $\sum_{n=1}^{\infty} a_n$ se dice **divergente** si $\lim_{n\to\infty} s_n = \pm\infty$, y escribimos entonces $\sum_{n=1}^{\infty} a_n = \pm\infty$.

- La serie $\sum_{n=1}^{\infty} a_n$ se dice **oscilante** si no existe $\lim_{n\to\infty} s_n$.

Si una serie $\sum_{n=0}^{\infty} a_n$ es convergente, entonces necesariamente $\lim_{n\to\infty} a_n = 0$ (ya que $a_n = s_n - s_{n-1}$ $\forall n \geq 2$).

Ejemplo. (Series geométricas). Dado un número real $r \in \mathbb{R}$, consideramos la serie $\sum_{n=0}^{\infty} r^n$. Si $r = 1$ la serie diverge, y si $r \neq 1$ se tiene:

$$1 + r + \cdots + r^n = \frac{1 - r^{n+1}}{1 - r} \begin{cases} \xrightarrow[n\to\infty]{} \dfrac{1}{1-r} & \text{si } |r| < 1, \\ \xrightarrow[n\to\infty]{} +\infty & \text{si } r \geq 1, \\ \text{no tiene límite} & \text{si } r \leq -1. \end{cases}$$

En particular,

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r} \qquad \text{si } |r| < 1.$$

3.2. Series de Taylor

Sea $f : I \longrightarrow \mathbb{R}$ una función continua e indefinidamente derivable en el interior de I. Sea a un punto interior de I.

La **serie de Taylor** de f en a es la serie

$$\boxed{\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n}$$

(¡para cada x se tiene una serie!)

La suma parcial n-ésima

$$\sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i = P_n(x)$$

es el polinomio de Taylor de f en a de grado n. Por tanto la serie de Taylor no es más que el límite de los polinomios de Taylor:

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = \lim_{n\to\infty} P_n(x).$$

Proposición. *Si existe $M > 0$ tal que $\forall n \in \mathbb{N}$ y $\forall x \in I$ se tiene $|f^{(n)}(x)| < M$, entonces $\forall x \in I$ se tiene que la serie de Taylor de f en a coincide con f:*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

para todo $x \in I$.

En efecto, el resto de Taylor satisface

$$|R_n(x)| = \frac{|f^{(n+1)}(c)|}{(n+1)!}|x-a|^{n+1} < \frac{M|x-a|^{n+1}}{(n+1)!} \xrightarrow[n\to\infty]{} 0,$$

luego se tiene

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = \lim_{n\to\infty} P_n(x) = \lim_{n\to\infty} \Big(f(x) - R_n(x)\Big) = f(x) - \lim_{n\to\infty} R_n(x) = f(x).$$

Ejemplos.

- e^x está acotada en todo intervalo finito, luego para todo x se tiene:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} .$$

- $\operatorname{sen} x = \displaystyle\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad \cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \ \forall x \in \mathbb{R}.$

- $(1+x)^\alpha = \displaystyle\sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$, si $|x| < 1$. En otro caso la serie puede ser divergente u oscilante. En particular,

$$(1+x)^{-1} = \frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n$$

  si $|x| < 1$.

3.3. Series de potencias (opcional)

Las series de la forma $\sum_{n=0}^{\infty} a_n x^n$ se dicen **series de potencias**.

Propiedades.

- Si $\sum_{n=0}^{\infty} a_n x^n$ converge para $x = b$, entonces *converge absolutamente* en el intervalo $(-b, b)$, esto es, la serie de los valores absolutos $\sum_{n=0}^{\infty} |a_n x^n|$ converge.

  Esto se debe a que si $\sum_{n=0}^{\infty} a_n b^n$ converge, entonces $\lim_{n\to\infty} a_n b^n = 0$, y la sucesión $(a_n b^n)$ está acotada: existe $M > 0$ tal que $|a_n b^n| < M \ \forall n$. De aquí se obtiene

$$\sum_{n=0}^{\infty} |a_n x^n| \le \sum_{n=0}^{\infty} |a_n b^n| \left|\frac{x}{b}\right|^n \le M \sum_{n=0}^{\infty} \left|\frac{x}{b}\right|^n$$

  y la última serie es convergente (geométrica) si $\left|\dfrac{x}{b}\right| < 1$.

- O bien $\sum_{n=0}^{\infty} a_n x^n$ sólo converge para $x = 0$ o existe $R \in \mathbb{R}^+ \cup \{+\infty\}$ tal que $\sum_{n=0}^{\infty} a_n x^n$ converge en $(-R, R)$ y no converge en $\mathbb{R} \setminus [-R, R]$. La situación para $x = R$ y $x = -R$ hay que estudiarla de modo separado. R se dice *radio de convergencia* de la serie.

- Si R es el radio de convergencia de $\sum_{n=0}^{\infty} a_n x^n$ y definimos $f(x) := \sum_{n=0}^{\infty} a_n x^n$ para todo $x \in (-R, R)$, entonces la función f es derivable indefinidamente en $(-R, R)$, con $f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$.

  En particular, la serie de Taylor de $f(x)$ en $x = 0$ es precisamente $\sum_{n=0}^{\infty} a_n x^n$, pues $a_n = \dfrac{f^{(n)}(a)}{n!}$ para todo $n \in \mathbb{N}$.

Capítulo 4

Integración en una variable

§ 1. Primitivas

Dadas dos funciones $F, f : I \longrightarrow \mathbb{R}$, definidas en un intervalo I, si $F'(x) = f(x)\ \forall x \in I$, decimos que F es una **primitiva** (o *integral* o *antiderivada*) de f.

Si G es otra primitiva de f, entonces $G'(x) - F'(x) = 0\ \forall x \in I$, luego $G(x) = F(x) + C$, donde C es una constante. Escribiremos,

$$\int f(x)\,dx = F(x) + C.$$

1.1. Lista de integrales elementales

1. $\displaystyle\int (x+a)^r\,dx = \frac{(x+a)^{r+1}}{r+1} + C$, si $r \neq -1$

2. $\displaystyle\int \frac{dx}{x+a} = \ln|x+a| + C$

3. $\displaystyle\int e^x\,dx = e^x + C$

4. $\displaystyle\int \cos x\,dx = \operatorname{sen} x + C$

5. $\displaystyle\int \operatorname{sen} x\,dx = -\cos x + C$

6. $\displaystyle\int \cosh x\,dx = \operatorname{senh} x + C$

7. $\displaystyle\int \operatorname{senh} x\,dx = \cosh x + C$

8. $\displaystyle\int \frac{dx}{\cos^2 x} = \int (1 + \operatorname{tg}^2 x)\,dx = \operatorname{tg} x + C$

9. $\displaystyle\int \frac{dx}{\operatorname{sen}^2 x} = -\operatorname{ctg} x + C$

10. $\displaystyle\int \frac{dx}{(x+a)^2+b^2} = \frac{1}{b}\operatorname{arc\,tg}\frac{x+a}{b} + C = -\frac{1}{b}\operatorname{arc\,ctg}\frac{x+a}{b} + D$, si $b \neq 0$

11. $\displaystyle\int \frac{2(x+a)}{(x+a)^2+r}\,dx = \ln|(x+a)^2+r| + C$

12. $\displaystyle\int \frac{2(x+a)}{[(x+a)^2+r]^n}\,dx = \frac{1/(1-n)}{[(x+a)^2+r]^{n-1}} + C$, si $n \neq 1$

13. $\displaystyle\int \frac{dx}{\sqrt{(x+a)^2+r}} = \ln\left|x+a+\sqrt{(x+a)^2+r}\right| + C$

14. $\displaystyle\int \frac{dx}{\sqrt{b^2-(x+a)^2}} = \operatorname{arc\,sen}\frac{x+a}{b} + C = -\arccos\frac{x+a}{b} + D$, si $b > 0$

15. $\displaystyle\int \frac{dx}{x^2+1} = \operatorname{arc\,tg} x + C = -\operatorname{arc\,ctg} x + D$ (es un caso particular de 10)

16. $\displaystyle\int \frac{dx}{\sqrt{x^2+1}} = \operatorname{arg\,senh} x + C = \ln(x+\sqrt{x^2+1}) + C$ (caso particular de 13)

17. $\displaystyle\int \frac{dx}{\sqrt{x^2-1}} = \operatorname{arg\,cosh} x + C = \ln\left|x+\sqrt{x^2-1}\right| + C$ (caso particular de 13)

18. $\displaystyle\int \frac{dx}{\sqrt{1-x^2}} = \operatorname{arc\,sen} x + C = -\arccos x + D$ (es un caso particular de 14)

1.2. Técnicas de integración

Cambio de variable:

Si F es una primitiva de f: $\int f(t)\,dt = F(t) + C$, por la regla de la cadena tenemos

$$\Big(F(g(x))\Big)' = F'(g(x))g'(x) = f(g(x))g'(x)$$

y, por tanto,

$$\int f(g(x))g'(x)\,dx = F(g(x)) + C.$$

Esto es, para calcular la integral $\int f(g(x))g'(x)\,dx$, sustituimos $t = g(x)$ y $dt = g'(x)\,dx$.

Ejemplo. $\int 4x\sqrt{x^2+1}\,dx$

- Elegimos una sustitución adecuada $t = g(x)$, en nuestro caso $t = x^2 + 1$,
- calculamos $dt = g'(x)\,dx$: $\quad dt = 2x\,dx$,
- reescribimos en términos de t: $\displaystyle\int 4x\sqrt{x^2+1}\,dx = \int 2\sqrt{t}\,dt$,
- calculamos la integral en t: $\displaystyle\int 2\sqrt{t}\,dt = \frac{4}{3}t^{3/2} + C$,
- cambiamos t por su valor: $\displaystyle\int 4x\sqrt{x^2+1}\,dx = \frac{4}{3}(x^2+1)^{3/2} + C$.

Integración por partes:

Utilizando la derivada del producto:

$$\big(u(x)v(x)\big)' = u'(x)v(x) + u(x)v'(x)$$

y escribiendo $du = u'(x)\,dx$ y $dv = v'(x)\,dx$, obtenemos

$$\int u\,dv = uv - \int v\,du$$

Ejemplo. Con $u(x) = x$ y $v(x) = e^x$ obtenemos:

$$\int xe^x\,dx = xe^x - \int e^x\,dx = (x-1)e^x + C.$$

Integración de funciones racionales:

Las funciones racionales son las que aparecen como cociente de polinomios: $\displaystyle\int \frac{f(x)}{g(x)}\,dx$, con $f(x)$ y $g(x)$ polinomios con coeficientes reales.

Mediante el algoritmo de la división, dividimos $f(x)$ por $g(x)$, obteniendo $f(x) = q(x)g(x)+r(x)$, con el grado de $r(x)$ menor que el de $g(x)$. Así

$$\int \frac{f(x)}{g(x)}\,dx = \int q(x)\,dx + \int \frac{r(x)}{g(x)}\,dx$$

y el problema se reduce al caso en que el grado de $f(x)$ es menor que el de $g(x)$.

Factorizamos entonces el denominador:

$$g(x) = a(x-a_1)^{k_1}\cdots(x-a_r)^{k_r}(x^2+b_1x+c_1)^{l_1}\cdots(x^2+b_sx+c_s)^{l_s},$$

con $b_i^2 - 4c_i < 0\ \forall i = 1, \ldots, s$.

- Si todos los factores tienen exponente 1, descomponemos en *factores simples*. Esto es, obtenemos $A_1, \dots, A_r, B_1, C_1, \dots, B_s, C_s$ tales que

$$\frac{f(x)}{g(x)} = \frac{A_1}{x-a_1} + \cdots + \frac{A_r}{x-a_r} + \frac{B_1 x + C_1}{x^2 + b_1 x + c_1} + \cdots + \frac{B_s x + C_s}{x^2 + b_s x + c_s}.$$

  Las integrales que se obtienen son sencillas:

$$\int \frac{dx}{x-2} = \ln|x-2| + C,$$

$$\begin{aligned}\int \frac{x+2}{x^2+2x+2}\,dx &= \int \frac{x+2}{(x+1)^2+1}\,dx \\ &= \frac{1}{2}\int \frac{2(x+1)}{(x+1)^2+1}\,dx + \int \frac{1}{(x+1)^2+1}\,dx \\ &= \frac{1}{2}\ln\Big((x+1)^2+1\Big) + \operatorname{arc\,tg}(x+1) + C.\end{aligned}$$

- Si hay factores con exponente > 1, utilizaremos el *método de Hermite-Ostrogradsky*:

$$\int \frac{f(x)}{g(x)}\,dx = \frac{p(x)}{\tilde{g}(x)} + \int \frac{q(x)}{\hat{g}(x)}\,dx,$$

  con

$$\begin{aligned}\tilde{g}(x) &= a(x-a_1)^{k_1-1}\cdots(x-a_r)^{k_r-1}(x^2+b_1x+c_1)^{l_1-1}\cdots(x^2+b_sx+c_s)^{l_s-1}, \\ \hat{g}(x) &= a(x-a_1)\cdots(x-a_r)(x^2+b_1x+c_1)\cdots(x^2+b_sx+c_s),\end{aligned}$$

  con el grado de $p(x)$ menor que el de $\tilde{g}(x)$ y el de $q(x)$ menor que el de $\hat{g}(x)$. Calculamos los coeficientes de $p(x)$ y $q(x)$ resolviendo la igualdad

$$\frac{f(x)}{g(x)} = \frac{p'(x)\tilde{g}(x) - p(x)\tilde{g}'(x)}{\tilde{g}(x)^2} + \frac{q(x)}{\hat{g}(x)}\ .$$

  De este modo pasamos al caso en el que los factores tienen exponente 1.

Funciones trigonométricas:

1. $\displaystyle\int R(\operatorname{sen} x)\cos x\,dx$, donde R es una función racional. Se reduce a la integral de una función racional con el cambio $t = \operatorname{sen} x$.

2. $\displaystyle\int R(\cos x)\operatorname{sen} x\,dx$, donde R es una función racional. Se reduce a la integral de una función racional con el cambio $t = \cos x$.

3. $\displaystyle\int R(\operatorname{tg} x)\,dx$, donde R es una función racional. Se reduce a la integral de una función racional con el cambio $t = \operatorname{tg} x$.

4. $\int R(\operatorname{sen} x, \cos x)\, dx$, donde R es una función racional. Se reduce a la integral de una función racional con el cambio

$$t = \operatorname{tg}\frac{x}{2}, \qquad dx = \frac{2\,dt}{1+t^2}, \qquad \cos x = \frac{1-t^2}{1+t^2}, \qquad \operatorname{sen} x = \frac{2t}{1+t^2}.$$

5. Los productos de funciones trigonométricas se pueden transformar en sumas, mediante las fórmulas siguientes[1]:

$$\begin{aligned} 2\operatorname{sen} a \operatorname{sen} b &= \cos(a-b) - \cos(a+b) \\ 2\cos a \cos b &= \cos(a-b) + \cos(a+b) \\ 2\operatorname{sen} a \cos b &= \operatorname{sen}(a-b) + \operatorname{sen}(a+b) \end{aligned}$$

En particular: $\cos^2 a = \dfrac{1+\cos 2a}{2}$; $\operatorname{sen}^2 a = \dfrac{1-\cos 2a}{2}$.

Algunas funciones irracionales:

1. $\int R(x, x^{m/n}, \ldots, x^{r/s})\, dx$, donde R es una función racional. Se reduce a la integral de una función racional mediante el cambio $x = t^k$, donde k es un común múltiplo de los denominadores $n, \ldots, s$.

2. $\int R\left(x, \left(\dfrac{ax+b}{cx+d}\right)^{1/n}\right) dx$, donde R es una función racional. Se reduce a la integral de una función racional con el cambio $\dfrac{ax+b}{cx+d} = t^n$.

3. $\int \dfrac{dx}{\sqrt{ax^2+bx+c}}$. Si $a=0$ es inmediata. Y si $a \neq 0$ también, ya que

$$\int \frac{dx}{\sqrt{ax^2+bx+c}} = \int \frac{dx}{\sqrt{a\left(x+\frac{b}{2a}\right)^2 + c - \frac{b^2}{4a}}}.$$

4. $\int \dfrac{P(x)}{\sqrt{ax^2+bx+c}}\, dx$, donde P es un polinomio. Se hallan una constante K y un polinomio Q, de grado menor que P, tales que

$$\int \frac{P(x)}{\sqrt{ax^2+bx+c}}\, dx = Q(x)\sqrt{ax^2+bx+c} + K\int \frac{dx}{\sqrt{ax^2+bx+c}}.$$

5. $\int \dfrac{dx}{(x-u)^m\sqrt{ax^2+bx+c}}$. Se hace el cambio $x - u = \dfrac{1}{t}$ y se reduce a una de las anteriores.

[1]Estas fórmulas se obtienen fácilmente usando números complejos.

6. $\displaystyle\int R\Big(x, \sqrt{a^2 - (x+b)^2}\Big)\, dx$, donde R es una función racional. Se reduce a la integral de una función de tipo trigonométrico mediante uno de los dos cambios

$$x + b = a\cos t \qquad \text{o} \qquad x + b = a \operatorname{sen} t.$$

7. $\displaystyle\int R\Big(x, \sqrt{(x+b)^2 - a^2}\Big)\, dx$, donde R es una función racional. Se reduce a la integral de una función de tipo trigonométrico mediante uno de los dos cambios

$$x + b = \frac{a}{\cos t} \qquad \text{o} \qquad x + b = \frac{a}{\operatorname{sen} t}.$$

8. $\displaystyle\int R\Big(x, \sqrt{a^2 + (x+b)^2}\Big)\, dx$, donde R es una función racional. Se reduce a la integral de una función de tipo trigonométrico mediante uno de los dos cambios

$$x + b = a \operatorname{tg} t \qquad \text{o} \qquad x + b = \frac{a}{\operatorname{tg} t}.$$

9. $\displaystyle\int R\Big(x, \sqrt{ax^2 + bx + c}\Big)\, dx$, donde R es una función racional. O bien se expresa como uno de los tres tipos anteriores, o bien se reduce a la integral de una función racional mediante un *cambio de variable de Euler*:

   a) $\sqrt{ax^2 + bx + c} = t \pm x\sqrt{a}$, si $a > 0$;

   b) $\sqrt{ax^2 + bx + c} = tx \pm \sqrt{c}$, si $c > 0$;

   c) $\sqrt{ax^2 + bx + c} = t(x - u)$, si $au^2 + bu + c = 0$.

§ 2. Integral definida

El concepto de integral no surgió a partir del problema de cálculo de primitivas, sino del problema de cálculo de áreas, volúmenes, ...

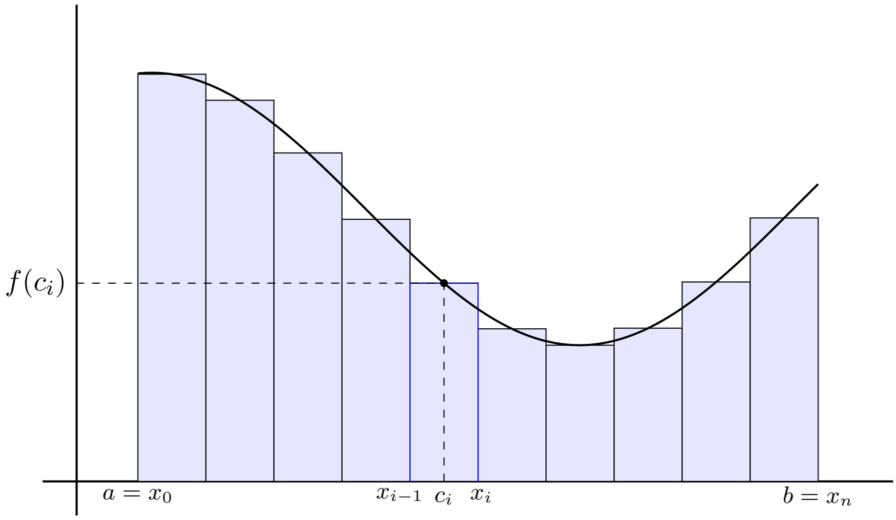
Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función continua con $f(x) > 0$ $\forall x \in [a, b]$.

Podemos aproximar el área bajo la gráfica de f tomando *particiones*

$$P = \{a = x_0 < x_1 < x_2 < \ldots < x_n = b\}$$

y puntos $c_i \in [x_{i-1}, x_i]$ $\forall i$, y considerando sumas (llamadas *Sumas de Riemann*):

$$S_P = \sum_{i=1}^{n} f(c_i)(x_i - x_{i-1}).$$



La *norma* de P es $\|P\| = \text{máx}\{x_i - x_{i-1} : 1 \leq i \leq n\}$.

Al tomar particiones cada vez "más finas" (esto es, cuya norma es cada vez más pequeña), las sumas S_P convergen hacia el área bajo la gráfica de f.

Una función arbitraria $f : [a, b] \longrightarrow \mathbb{R}$ se dice *integrable* si existe el límite[2] de sus sumas de Riemann al hacer la norma de las particiones tender a 0. Dicho límite se dice **integral definida** de f y se denota:

$$\boxed{\int_a^b f(x)\,dx\,.}$$

Las funciones continuas son siempre integrables.

Por convenio,

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx, \qquad \int_a^a f(x)\,dx = 0\,.$$

Algunas propiedades elementales son:

$$\int_a^b kf(x)\,dx = k\int_a^b f(x)\,dx, \quad (k \in \mathbb{R})$$
$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx,$$
$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx, \quad a < c < b.$$

Teorema Fundamental del Cálculo Integral. *Si $f : [a, b] \longrightarrow \mathbb{R}$ es una función continua, la función*

$$\begin{aligned} F : [a, b] &\longrightarrow \mathbb{R} \\ x &\mapsto F(x) = \int_a^x f(t)\,dt, \end{aligned}$$

[2]Estos límites se definen de una manera muy precisa.

es una primitiva de f.

Si G es cualquier primitiva de f, entonces

$$\int_a^b f(t)\,dt = \Big[G(x)\Big]_a^b := G(b) - G(a).$$

En efecto, $F(x+h)-F(x) = \int_x^{x+h} f(t)\,dt$, y al hacer h pequeño, el valor de $\int_x^{x+h} f(t)\,dt$ tiende hacia $f(x)h$, luego

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

Si $G(x)$ es una primitiva cualquiera de $f(x)$, entonces $G(x) = F(x)+C$, luego $G(a) = F(a)+C = 0+C = C$ y

$$\int_a^b f(t)\,dt = F(b) = G(b) - C = G(b) - G(a).$$

Cambio de variables. Para integrales definidas, el cambio $t = g(x)$ queda así:

$$\boxed{\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(t)\,dt\,.}$$

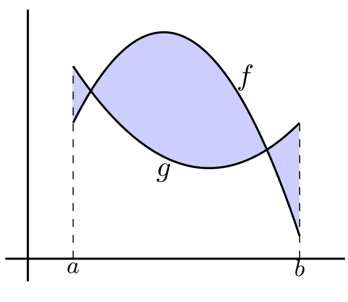
Notemos que si F es primitiva de f, entonces $F(g(x))$ es primitiva de $f\big(g(x)\big)g'(x)$, luego ambos lados son iguales a $F\big(g(b)\big) - F\big(g(a)\big)$.

§ 3. Aplicaciones del cálculo integral

Sean $f, g : [a,b] \longrightarrow \mathbb{R}$ dos funciones continuas.

3.1. Áreas

El área de la región limitada por las gráficas de f y g



viene dado por

$$\boxed{\int_a^b |f(x) - g(x)|\,dx.}$$

Ejemplo. El área encerrada por la función $(1+\cos^2 x)\operatorname{sen} x$, $0 \le x \le 2\pi$ y el eje OX es, por simetría,

$$\begin{aligned} 2\int_0^\pi (1+\cos^2 x)\operatorname{sen} x\,dx &= -2\int_1^{-1}(1+t^2)\,dt \quad (\text{cambio } t=\cos x) \\ &= 2\int_{-1}^{1}(1+t^2)\,dt = 2\Big[t+\frac{t^3}{3}\Big]_{-1}^{1} = \frac{16}{3}. \end{aligned}$$

Ejemplo. El área encerrada por la gráfica[3] de xe^{-x}, $x \in [0,+\infty)$ es

$$\begin{aligned} \int_0^{+\infty} xe^{-x}\,dx &= \Big[-xe^{-x}\Big]_0^{+\infty} + \int_0^{+\infty} e^{-x}\,dx \quad (\text{partes: } u=x,\ v=-e^{-x}) \\ &= -\Big[(x+1)e^{-x}\Big]_0^{+\infty} = 1. \end{aligned}$$

3.2. Volúmenes de revolución

Si hacemos girar el área determinada por la gráfica de f alrededor del eje OX, el volumen obtenido es el límite de las sumas

$$\sum_{i=1}^{n} \pi f(c_i)^2 (x_i - x_{i-1}),$$

que son las sumas de Riemann de la función $\pi f(x)^2$. Por tanto el volumen viene dado por

$$\boxed{\pi\int_a^b f(x)^2\,dx\,.}$$

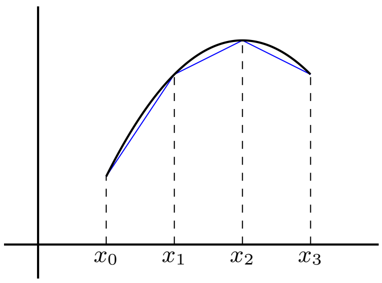
Si $0 \le a < b$ y giramos alrededor del eje OY, el volumen generado viene dado por

$$\boxed{2\pi\int_a^b x|f(x)|\,dx\,.}$$

3.3. Longitud de una curva

La longitud de la gráfica de la función f viene dada por el límite de sumas de la forma

$$\sum_{i=1}^{n}\sqrt{(x_i-x_{i-1})^2+\big(f(x_i)-f(x_{i-1})\big)^2} = \sum_{i=1}^{n}\sqrt{1+\left(\frac{f(x_i)-f(x_{i-1})}{x_i-x_{i-1}}\right)^2}\,(x_i-x_{i-1})$$



[3]Las *integrales impropias* $\int_a^{+\infty} f(x)\,dx$ deben de interpretarse como $\lim_{t\to+\infty}\int_a^t f(x)\,dx$, y análogamente para las integrales de la forma $\int_a^b f(x)\,dx$ si $\lim_{x\to b^-} f(x) = \pm\infty$ o $\lim_{x\to a^-} f(x) = \pm\infty$.

que, por el Teorema del Valor Medio, es una suma de Riemann

$$\sum_{i=1}^{n} \sqrt{1+\left(f'(c_i)\right)^2}(x_i - x_{i-1})$$

de la función $\sqrt{1+\left(f'(x)\right)^2}$. Por tanto, la longitud es

$$\boxed{\int_a^b \sqrt{1+\left(f'(x)\right)^2}\, dx\,.}$$

3.4. Áreas de revolución

El área de la superficie generada al girar la gráfica de f alrededor de OX es

$$\boxed{2\pi \int_a^b |f(x)|\sqrt{1+\left(f'(x)\right)^2}\, dx\,.}$$

Capítulo 5

Curvas en paramétricas

En este capítulo consideraremos funciones definidas en un intervalo I de $\mathbb{R}$, pero con valores en los espacios $\mathbb{R}^2$ o $\mathbb{R}^3$. En $\mathbb{R}^3$ (y más sencillamente en $\mathbb{R}^2$) sus elementos son vectores $\mathbf{u} = (u_1, u_2, u_3)$, con los que podemos calcular su:

- *producto escalar:* $\mathbf{u} \bullet \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3$,
- *norma:* $\|\mathbf{u}\| = \sqrt{\mathbf{u} \bullet \mathbf{u}} = \sqrt{u_1^2 + u_2^2 + u_3^2}$,
- *distancia:* $\mathrm{d}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2}$.

§ 1. Definiciones y ejemplos

Una *curva en paramétricas* en $\mathbb{R}^3$ (análogamente para $\mathbb{R}^2$) es una función continua:

$$\begin{aligned} \sigma : I &\longrightarrow \mathbb{R}^3 \\ t &\mapsto \sigma(t) = \big(x(t), y(t), z(t)\big) \end{aligned}$$

donde I es un intervalo en $\mathbb{R}$.

Ejemplos. **1.** Una recta es un ejemplo de curva. La ecuación en paramétricas de la recta en $\mathbb{R}^3$ que pasa por el punto de coordenadas $P = (p_1, p_2, p_3)$ y tiene como vector director $\mathbf{u} = (u_1, u_2, u_3)$ viene dada por

$$\begin{aligned} \sigma : \mathbb{R} &\longrightarrow \mathbb{R}^3 \\ t &\mapsto \sigma(t) = P + t\mathbf{u}, \end{aligned}$$

Esto es:

$$\begin{cases} x(t) = p_1 + tu_1, \\ y(t) = p_2 + tu_2, \\ z(t) = p_3 + tu_3. \end{cases}$$

2. La circunferencia unidad en $\mathbb{R}^2$, de ecuación $x^2 + y^2 = 1$ (*ecuación implícita*), la podemos ver como la curva en paramétricas:

$$\begin{cases} x(t) = \cos t, \\ y(t) = \operatorname{sen} t, \end{cases}$$

con $t \in [0, 2\pi]$.

3. La elipse de ecuación implícita $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ $(a > b)$ es el lugar geométrico de los puntos cuya suma de distancias a los puntos (*focos*) $(\pm\sqrt{a^2 - b^2}, 0)$ es constante e igual a $2a$. En paramétricas sus ecuaciones son:

$$\begin{cases} x(t) = a\cos t, \\ y(t) = b\operatorname{sen} t, \end{cases}$$

con $t \in [0, 2\pi]$.

4. La parte de la hipérbola de ecuación $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ con $x > 0$, tiene como ecuaciones paramétricas:

$$\begin{cases} x(t) = a\cosh t, \\ y(t) = b\operatorname{senh} t, \end{cases}$$

con $t \in \mathbb{R}$.

5. La ecuación implícita de la *Lemniscata de Bernoulli* es:

$$(x^2 + y^2)^2 = 2a^2(x^2 - y^2)$$

(donde $a \in \mathbb{R}$, $a > 0$).

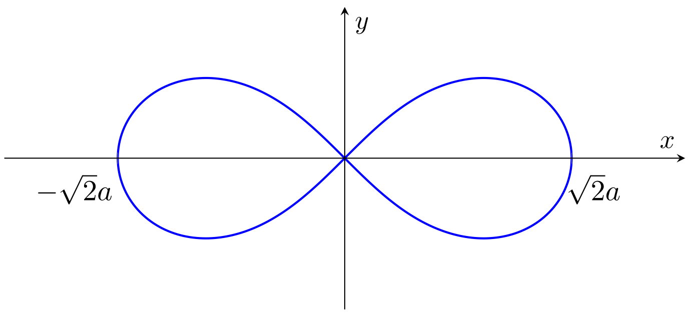
Si cortamos la lemniscata con la recta $y = sx$ obtenemos $(1 + s^2)^2 x^4 = 2a^2(1 - s^2)x^2$ y, suponiendo $x > 0$, deducimos las ecuaciones:

$$x = \frac{\sqrt{2}a\sqrt{1 - s^2}}{1 + s^2}, \quad y = \frac{\sqrt{2}as\sqrt{1 - s^2}}{1 + s^2}.$$

Haciendo $s = \cos t$ obtenemos las ecuaciones paramétricas:

$$\begin{cases} x(t) = \dfrac{\sqrt{2}a\operatorname{sen} t}{1 + \cos^2 t}, \\ y(t) = \dfrac{\sqrt{2}a\operatorname{sen} t\cos t}{1 + \cos^2 t}, \end{cases}$$

con $t \in [-\pi, \pi]$.

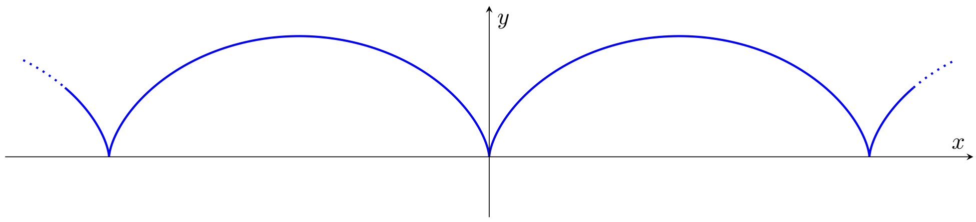


El nombre viene del latín `lemniscus` (cinta o lazo). Está inspirada en la elipse: es el lugar geométrico de los puntos de forma que el producto de la distancia a dos puntos prefijados es constante. Tiene forma de ∞ y de hecho parece ser que esta curva es el origen del símbolo ∞.

6. La *Cicloide* es la curva con ecuación en paramétricas

$$\begin{cases} x(t) = t - \operatorname{sen} t, \\ y(t) = 1 - \cos t, \end{cases}$$

$t \in \mathbb{R}$.



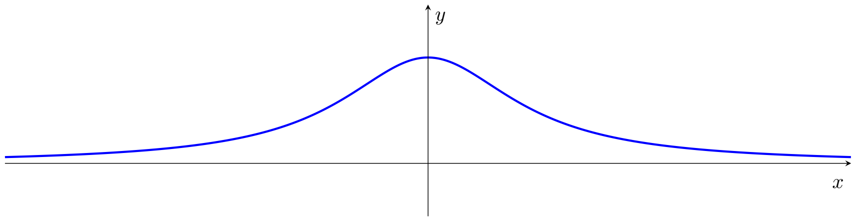
7. La *Bruja de Agnesi*[1] es la curva de ecuación

$$y = \frac{a^3}{x^2 + a^2}.$$

Haciendo $t = \dfrac{x}{a}$, obtenemos sus ecuaciones paramétricas:

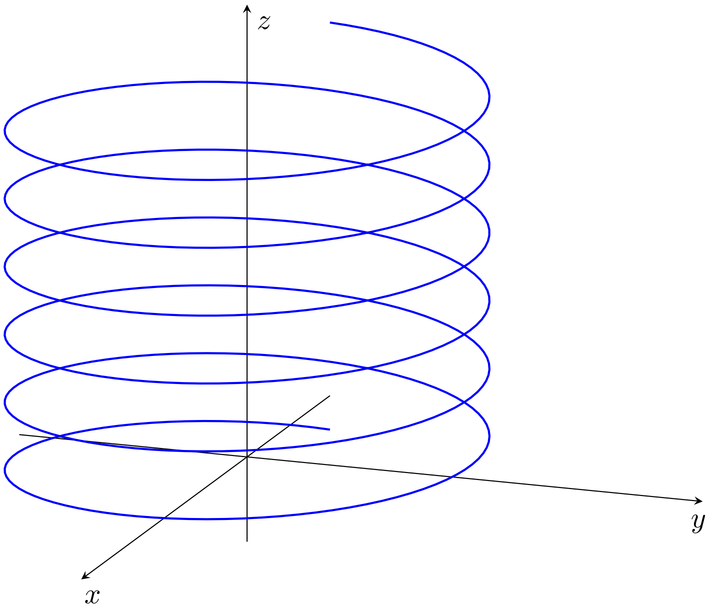
$$\begin{cases} x = at, \\ y = \dfrac{a}{1 + t^2}. \end{cases}$$

[1]El nombre de esta curva se debe a una mala traducción: aparece (no por primera vez) en un tratado de análisis escrito en 1748 por María Cayetana Agnesi. El nombre que ella le da es el de *curva versiera* (término naval). El traductor al inglés confundió esta palabra con *avversiera*, palabra con la misma raíz que averno y que significa diablesa y lo tradujo por *witch*.



8. La *hélice circular* es la curva de $\mathbb{R}^3$ de ecuaciones paramétricas:

$$\begin{cases} x(t) = \cos t, \\ y(t) = \operatorname{sen} t, \\ z(t) = at. \end{cases}$$



§ 2. Recta tangente

Dada una curva en paramétricas:

$$\begin{aligned} \sigma : I &\longrightarrow \mathbb{R}^3 \\ t &\mapsto \sigma(t) = \big(x(t), y(t), z(t)\big) \end{aligned}$$

donde I es un intervalo en $\mathbb{R}$, si las coordenadas de σ son derivables en un punto t_0, entonces σ *es derivable* en t_0, y

$$\sigma'(t_0) := \lim_{t \to t_0} \frac{\sigma(t) - \sigma(t_0)}{t - t_0} = \big(x'(t_0), y'(t_0), z'(t_0)\big).$$

Esto es, la derivada se obtiene derivando las componentes, que son funciones de una variable $I \to \mathbb{R}$.

La derivada $\sigma'(t_0)$, si es no nula, nos da un vector tangente a la curva en el punto $\sigma(t_0)$, de modo que la recta tangente a la curva σ en el punto t_0 es

$$\begin{cases} x = x(t_0) + tx'(t_0), \\ y = y(t_0) + ty'(t_0), \\ z = z(t_0) + tz'(t_0). \end{cases}$$

Ejemplo. Un vector tangente a la circunferencia unidad: $\sigma(t) = (\cos t, \operatorname{sen} t)$, en el punto $(\cos t_0, \operatorname{sen} t_0)$ es $\sigma'(t_0) = (-\operatorname{sen} t_0, \cos t_0)$, por lo que la recta tangente es:

$$\begin{cases} x = \cos t_0 - t \operatorname{sen} t_0, \\ y = \operatorname{sen} t_0 + t \cos t_0. \end{cases}$$

Notemos que si utilizamos la ecuación implícita de la circunferencia: $x^2 + y^2 = 1$, podemos calcular un vector tangente en un punto $x_0 = (\cos t_0, \operatorname{sen} t_0)$ donde la tangente no sea vertical despejando y en función de x y derivando. Así tenemos:

$$2x + 2yy' = 0,$$

luego $y' = -\dfrac{x}{y}$. Así, en el punto de coordenadas $(\cos t_0, \operatorname{sen} t_0)$, un vector tangente $(1, y')$ es $-\left(1, -\dfrac{x}{y}\right) = \left(-1, \dfrac{\cos t_0}{\operatorname{sen} t_0}\right)$, que es proporcional al obtenido anteriormente.

§ 3. Longitudes, áreas, trabajo

3.1. Integral de línea

Dada una función $f : D \to \mathbb{R}$, definida en un subconjunto D de $\mathbb{R}^3$, y una curva $\sigma(t) = (x(t), y(t), z(t))$, $t \in [t_0, t_1]$, de modo que $\sigma(t) \in D$ $\forall t$, se llama **integral de línea** de f a lo largo de σ a la integral

$$\begin{aligned} \int_\sigma f\, ds &:= \int_{t_0}^{t_1} f(\sigma(t)) \|\sigma'(t)\|\, dt \\ &= \int_{t_0}^{t_1} f(x(t), y(t), z(t)) \sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2 + \big(z'(t)\big)^2}\, dt\,. \end{aligned}$$

En particular, para la función constante igual a 1, la integral de línea nos da la longitud de la curva:

$$\boxed{L = \int_{t_0}^{t_1} \sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2 + \big(z'(t)\big)^2}\, dt\,.}$$

Las fórmulas para curvas planas: $\sigma(t) = (x(t), y(t))$, son las mismas eliminando las z's.

Ejemplo. La longitud de la circunferencia unidad ($\sigma(t) = (\cos t, \operatorname{sen} t)$, $t \in [0, 2\pi]$) es

$$L = \int_0^{2\pi} \sqrt{(-\operatorname{sen} t)^2 + (\cos t)^2}\, dt = \int_0^{2\pi} dt = 2\pi\,.$$

3.2. Área limitada por una curva plana

El área limitada por una curva plana $\sigma(t) = (x(t), y(t))$, y con $x(t)$ derivable y estrictamente creciente o decreciente, y el eje OX entre $t = t_0$ y $t = t_1$ se deduce de la fórmula $\int_a^b |f(x)|dx$, donde $y = f(x)$, mediante el cambio de variables $x = x(t)$, obteniendo:

$$A = \int_{t_0}^{t_1} |y(t)x'(t)|dt\,.$$

El área limitada por una curva plana *simple* (esto es, sin autointersecciones) $\sigma(t) = \big(x(t), y(t)\big)$, con $t \in [t_0, t_1]$ y $\sigma(t_0) = \sigma(t_1)$ es

$$\boxed{A = \left|\int_{t_0}^{t_1} y(t)x'(t)dt\right|\,.}$$

Ejemplo. El área del círculo unidad es

$$A = \left|\int_0^{2\pi} \operatorname{sen} t(-\operatorname{sen} t)\,dt\right| = \int_0^{2\pi} \operatorname{sen}^2 t\,dt = \int_0^{2\pi} \frac{1-\operatorname{sen} 2t}{2}\,dt = \frac{1}{2}\left[t + \frac{\cos 2t}{2}\right]_0^{2\pi} = \pi\,.$$

3.3. Trabajo realizado por una fuerza al recorrer una curva

El trabajo realizado por una fuerza (o *campo vectorial*)

$$\begin{aligned}\mathbf{F} : D &\longrightarrow \mathbb{R}^3 \qquad (D \subseteq \mathbb{R}^3)\\ (x, y, z) &\mapsto \Big(f(x, y, z), g(x, y, z), h(x, y, z)\Big)\\ &\qquad\qquad = f(x, y, z)\mathbf{i} + g(x, y, z)\mathbf{j} + h(x, y, z)\mathbf{k}\\ &\qquad\qquad\quad (\mathbf{i} = (1, 0, 0),\ \mathbf{j} = (0, 1, 0),\ \mathbf{k} = (0, 0, 1))\end{aligned}$$

a lo largo de una curva $\sigma : [t_0, t_1] \to \mathbb{R}^3$, $t \mapsto \big(x(t), y(t), z(t)\big)$, con imagen contenida en D, es la integral de la componente de la fuerza en la dirección de $\sigma(t)$:

$$\boxed{\begin{aligned}W &= \int_{t_0}^{t_1} \Big(\mathbf{F}(\sigma(t)) \bullet \sigma'(t)\Big)\,dt\\ &= \int_{t_0}^{t_1} \Big(f\big(x(t), y(t), z(t)\big)x'(t) + g\big(x(t), y(t), z(t)\big)y'(t) + h\big(x(t), y(t), z(t)\big)z'(t)\Big)\,dt\,.\end{aligned}}$$

(Se suele escribir simplemente $W = \int_\sigma \mathbf{F} \bullet d\sigma$.)

Análogamente en el plano se tiene simplemente

$$W = \int_{t_0}^{t_1} \Big(f\big(x(t), y(t)\big)x'(t) + g\big(x(t), y(t)\big)y'(t)\Big)\,dt\,.$$

Ejemplo. El trabajo realizado por la fuerza $F(x,y) = (x,-y)$ (o $F = x\mathbf{i} - y\mathbf{j}$) al recorrer la circunferencia unidad es

$$W = \int_0^{2\pi} \Big(\cos t(-\operatorname{sen} t) - \operatorname{sen} t \cos t\Big)\, dt = -\int_0^{2\pi} 2\cos t \operatorname{sen} t\, dt$$

$$= -\int_0^{2\pi} \operatorname{sen} 2t\, dt = \left[\frac{\cos 2t}{2}\right]_0^{2\pi} = 0\,.$$

Capítulo 6

Álgebra lineal

El álgebra lineal es, a 'grosso modo', la matemática de los vectores y las matrices.

Trabajaremos con vectores, matrices, ..., sobre $\mathbb{R}$, pero todo tiene sentido para vectores, matrices, ..., sobre $\mathbb{C}$.[1]

§ 1. Vectores, matrices, aplicaciones lineales

Los elementos $\mathbf{x} = (x_1, \dots, x_n)$ de $\mathbb{R}^n$ $(n \in \mathbb{N})$ se dicen *vectores*. Los escribiremos en ocasiones como columnas $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

1.1. Vectores

Podemos operar vectores de modo que dados dos vectores $\mathbf{u} = (u_1, \dots, u_n),\ \mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n$ y un escalar $t \in \mathbb{R}$, tenemos:

$$\begin{aligned}
&\mathbf{u} \pm \mathbf{v} = (u_1 \pm v_1, \dots, u_n \pm v_n) && \text{(suma/diferencia)}, \\
&t\mathbf{u} = (tu_1, \dots, tu_n) && \text{(producto por escalar)}, \\
&\mathbf{u} \bullet \mathbf{v} = u_1 v_1 + \cdots + u_n v_n && \text{(producto escalar)}, \\
&\|\mathbf{u}\| = \sqrt{\mathbf{u} \bullet \mathbf{u}} = \sqrt{u_1^2 + \cdots + u_n^2} && \text{(norma)}.
\end{aligned}$$

1.2. Matrices

Una *matriz* $m \times n$ es una disposición rectangular de números con m filas y n columnas:

$$A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m_1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

[1] ¡Y sobre *cuerpos arbitrarios*!

El espacio de las matrices $m \times n$ se denota por $\mathrm{Mat}_{m\times n}(\mathbb{R})$ o $\mathbb{R}^{m\times n}$.

Podemos operar con matrices del modo siguiente:

Suma: Si $A = (a_{ij})$ y $B = (b_{ij})$ son matrices $m \times n$, su suma es $A + B = (a_{ij} + b_{ij})$.

Producto por escalares: $tA = (ta_{ij})$ $(t \in \mathbb{R},\ A \in \mathbb{R}^{m\times n})$.

Producto de matrices: El producto de $A = (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}} \in \mathbb{R}^{m\times n}$ y $B = (b_{ij})_{\substack{1\le i\le n\\ 1\le j\le p}} \in \mathbb{R}^{n\times p}$ (¡el número de columnas de A debe de ser igual al número de filas de B!) es la matriz $AB = (c_{ij})_{\substack{1\le i\le m\\ 1\le j\le p}}$, donde

$$c_{ij} = \sum_{k=1^n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

El producto, cuando existe, es asociativo y distributivo respecto a la suma.

Matriz traspuesta: La *traspuesta* de la matriz $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ es la matriz, que se denota por A' o $A^\top$, en $\mathbb{R}^{n\times n}$ cuya componente (i, j) es la componente (j, i) de A:

$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix}^\top = \begin{pmatrix} a & c & e \\ b & d & f \end{pmatrix}.$$

Dados dos vectores columna, tenemos $\mathbf{u} \bullet \mathbf{v} = \mathbf{u}^\top \mathbf{v}$.

Matriz inversa: Dada una matriz *cuadrada* $(n = m)$ $A \in \mathbb{R}^{n\times n}$, se dice *inversa* de A a la matriz $A^{-1} \in \mathbb{R}^{n\times n}$ tal que

$$AA^{-1} = A^{-1}A = I_n$$

donde I_n es la *matriz identidad de grado* n:

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

No toda matriz cuadrada tiene inversa.

1.3. Aplicaciones lineales

Toda matriz $A \in \mathbb{R}^{m\times n}$ induce una aplicación

$$T_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto T_A(\mathbf{x}) = A\mathbf{x}.$$

Se tienen las siguientes propiedades:

$$T_A(\mathbf{x}+\mathbf{y}) = T_A(\mathbf{x}) + T_A(\mathbf{y}),$$
$$T_A(t\mathbf{x}) = tT_A(\mathbf{x}).$$

para todo $t \in \mathbb{R}$ y $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Las aplicaciones con estas propiedades se llaman *aplicaciones lineales*.

Recíprocamente, si $f : \mathbb{R}^n \to \mathbb{R}^m$ es una aplicación lineal y $\mathbf{a}_i$ es el vector columna imagen del vector $\mathbf{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$ (1 en el lugar i y 0's en el resto), entonces para todo $\mathbf{x} \in \mathbb{R}^n$,

$$f(\mathbf{x}) = f\left(\sum_{i=1}^{n} x_i \mathbf{e}_i\right) = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a_n} = T_A(\mathbf{x})$$

donde $A = \Big(\mathbf{a}_1 \mid \cdots \mid \mathbf{a}_n\Big)$ es la matriz cuyas columnas son los vectores $\mathbf{a}_i$, $i = 1, \ldots, n$.

La matriz A se dice *matriz coordenada* de la aplicación lineal f.

§ 2. Sistemas de ecuaciones lineales. Rangos de matrices

2.1. Método de Gauss

Dado el sistema de ecuaciones:

$$\begin{cases} -2x_1 + 2x_2 - 4x_3 - 6x_4 = -4 \\ -3x_1 + 6x_2 + 3x_3 - 15x_4 = -3 \\ 5x_1 - 8x_2 - x_3 + 17x_4 = 9 \\ x_1 + x_2 + 11x_3 + 7x_4 = 7 \end{cases}$$

la matriz

$$A = \begin{pmatrix} -2 & 2 & -4 & -6 \\ -3 & 6 & 3 & -15 \\ 5 & -8 & -1 & 17 \\ 1 & 1 & 11 & 7 \end{pmatrix}$$

se dice *matriz de coeficientes* del sistema, y la matriz

$$\Big(A \mid \mathbf{b}\Big) = \left(\begin{array}{cccc|c} -2 & 2 & -4 & -6 & -4 \\ -3 & 6 & 3 & -15 & -3 \\ 5 & -8 & -1 & 17 & 9 \\ 1 & 1 & 11 & 7 & 7 \end{array}\right)$$

es la *matriz ampliada*.

La idea general para hallar las soluciones de este sistema consiste en utilizar alguna de las ecuaciones para despejar una de las variables y sustituir el resultado en las demás, de modo que se obtiene un sistema equivalente más sencillo.

En nuestro ejemplo podemos utilizar la primera ecuación para despejar x_1. Para ello multiplicamos la primera ecuación por $-\frac{1}{2}$:

$$x_1 - x_2 + 2x_3 + 3x_4 = 2 \tag{6.1}$$

luego $x_1 = 2 + x_2 - 2x_3 - 3x_4$. Sustituyendo en las demás ecuaciones obtenemos:

2.ª ecuación: $-3(2 + x_2 - 2x_3 + 3x_4) + 6x_2 + 3x_3 - 15x_4 = -3$, o bien

$$(6-3)x_2 + (3+6)x_3 + (-15+9)x_4 = -3 + 6, \qquad \text{esto es,} \quad 3x_2 + 9x_3 - 6x_4 = 3,$$

que es igual a la 2.ª ecuación inicial más la ecuación (6.1) multiplicada por 3.

3.ª ecuación: $5(2 + x_2 - 2x_3 - 3x_4) - 8x_2 - x_3 + 17x_4 = 9$, o bien

$$(-8+5)x_2 + (-1-10)x_3 + (17-15)x_4 = 9 - 10, \qquad \text{esto es,} \quad -3x_2 - 11x_3 + 2x_4 = -1,$$

que es igual a la 3.ª ecuación inicial menos (6.1) multiplicada por 5.

4.ª ecuación: $(2 + x_2 - 2x_3 - 3x_4) + x_2 + 11x_3 + 7x_4$, o bien

$$(1+1)x_2 + (11-2)x_3 + (7-3)x_4 = 7 - 2, \qquad \text{esto es,} \quad 2x_2 + 9x_3 + 4x_4 = 5,$$

que es la 4.ª ecuación inicial menos (6.1).

Obtenemos pues el sistema equivalente

$$\begin{cases} x_1 - x_2 + 2x_3 + 3x_4 = 2 & -\frac{1}{2}(1.^{\text{a}}\text{ ecuación}) \\ 3x_2 + 9x_3 - 6x_4 = 3 & (2.^{\text{a}}\text{ ecuación}) - \frac{3}{2}(1.^{\text{a}}\text{ ecuación}) \\ -3x_2 - 11x_3 + 2x_4 = -1 & (3.^{\text{a}}\text{ ecuación}) + \frac{5}{2}(1.^{\text{a}}\text{ ecuación}) \\ 2x_2 + 9x_3 + 4x_4 = 5 & (4.^{\text{a}}\text{ ecuación}) + \frac{1}{2}(1.^{\text{a}}\text{ ecuación}) \end{cases}$$

Vemos, por tanto, que el proceso de despejar una variable en una de las ecuaciones y sustituir su valor en las demás equivale a sumar múltiplos escalares de la ecuación usada para despejar a las demás ecuaciones. Las operaciones efectuadas afectan solo a los términos de la matriz ampliada:

$$\left(\begin{array}{cccc|c} -2 & 2 & -4 & -6 & -4 \\ -3 & 6 & 3 & -15 & -3 \\ 5 & -8 & -1 & 17 & 9 \\ 1 & 1 & 11 & 7 & 7 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ -3 & 6 & 3 & -15 & -3 \\ 5 & -8 & -1 & 17 & 9 \\ 1 & 1 & 11 & 7 & 7 \end{array}\right) \quad (1.^{\text{a}}\text{ fila}) \text{ por } \frac{1}{2}$$

$$\longrightarrow \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 3 & 9 & -6 & 3 \\ 0 & -3 & -11 & 2 & -1 \\ 0 & 2 & 9 & 4 & 5 \end{array}\right) \quad \begin{array}{c} (2.^{\text{a}}\text{ fila}) + 3(1.^{\text{a}}\text{ fila}) \\ (3.^{\text{a}}\text{ fila}) - 5(1.^{\text{a}}\text{ fila}) \\ (4.^{\text{a}}\text{ fila}) - (1.^{\text{a}}\text{ fila}) \end{array}$$

Ahora procedemos de modo análogo a despejar x_2 de la 2.ª ecuación y sustituir en las siguientes (llamaremos f_i's a las filas):

$$\left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 3 & 9 & -6 & 3 \\ 0 & -3 & -11 & 2 & -1 \\ 0 & 2 & 9 & 4 & 5 \end{array}\right) \xrightarrow{\frac{1}{3}\times f_2} \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & -3 & -11 & 2 & -1 \\ 0 & 2 & 9 & 4 & 5 \end{array}\right) \xrightarrow[f_4-2\times f_2]{f_3+3\times f_2} \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & -2 & -4 & 2 \\ 0 & 0 & 3 & 8 & 3 \end{array}\right)$$

y a despejar x_3 de la 3.ª ecuación y sustituir en la 4.ª:

$$\left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & -2 & -4 & 2 \\ 0 & 0 & 3 & 8 & 3 \end{array}\right) \xrightarrow{\frac{-1}{2}\times f_3} \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 3 & 8 & 3 \end{array}\right) \xrightarrow{f_4-3\times f_3} \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 2 & 6 \end{array}\right)$$

Despejamos, por último, x_4 de la última ecuación:

$$\left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 2 & 6 \end{array}\right) \xrightarrow{\frac{1}{2}f_4} \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

Las soluciones de este sistema equivalente se obtienen fácilmente por *sustitución regresiva*:

$$\begin{aligned} x_4 &= 3, \\ x_3 &= -1 - 2x_4 = -7, \\ x_2 &= 1 - 3x_3 + 2x_4 = 28, \\ x_1 &= 2 + x_2 - 2x_3 - 3x_4 = 35. \end{aligned}$$

Ejemplo. Consideramos el siguiente sistema de ecuaciones lineales:

$$\begin{cases} 2x + 5y + 2z = -1 \\ x + 2y - z = 2 \\ 7x + 17y + 5z = -1 \end{cases}$$

Operando como antes, intercambiando primero las dos primeras filas obtenemos (señalamos con un cuadrado los *pivotes* usados para obtener 0's debajo):

$$\left(\begin{array}{ccc|c} 2 & 5 & 2 & -1 \\ 1 & 2 & -1 & 2 \\ 7 & 17 & 5 & -1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} \boxed{1} & 2 & -1 & 2 \\ 2 & 5 & 2 & -1 \\ 7 & 17 & 5 & -1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & \boxed{1} & 4 & -5 \\ 0 & 3 & 12 & -15 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & 1 & 4 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

El sistema inicial es pues equivalente a

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

que tiene como soluciones

$$z \text{ arbitrario}, \quad y = -5 - 4z, \quad x = 2 - 2y + z = 12 + 9z.$$

De otro modo, el conjunto de soluciones depende de un parámetro $t = z$:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 12 \\ -5 \\ 0 \end{pmatrix} + t \begin{pmatrix} 9 \\ -4 \\ 1 \end{pmatrix}$$

que es la recta que pasa por el punto $(12, -5, 0)$ y tiene como vector director $(9, -4, 1)$.

Ejemplo.

$$\begin{cases} x \; + \; 10z = 5 \\ 3x + y - 4z = -1 \\ 4x + y + 6z = 1 \end{cases}$$

Operando como antes, obtenemos:

$$\left(\begin{array}{ccc|c} \boxed{1} & 0 & 10 & 5 \\ 3 & 1 & -4 & -1 \\ 4 & 1 & 6 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & 10 & 5 \\ 0 & \boxed{1} & -34 & -16 \\ 0 & 1 & -34 & -19 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|c} 1 & 0 & 10 & 5 \\ 0 & 1 & -34 & -16 \\ 0 & 0 & 0 & -3 \end{array}\right)$$

El sistema inicial es pues equivalente al sistema

$$\begin{cases} x + 10z = 5 \\ y - 34z = -16 \\ \qquad 0 = -3 \end{cases}$$

que no tiene solución, pues para ningún valor de x, y, z, se verifica $0 = -3$.

(Multiplicando la última fila por $\dfrac{-1}{3}$ podemos cambiar -3 por 1 en la última columna.)

El proceso seguido en estos ejemplos comporta solo tres *operaciones elementales*:

(i) Intercambio de filas.

(ii) Multiplicación de una fila por un escalar no nulo.

(iii) Adición a una fila de un múltiplo escalar de otra.

Estas tres operaciones nos llevan de un sistema de ecuaciones lineales a otro equivalente. Notemos que todas las operaciones las podemos efectuar sobre la matriz ampliada del sistema, luego lo podemos plantear en términos matriciales:

Método de Gauss:

1. Si la matriz es la matriz nula, paramos.

2. En otro caso buscamos la primera columna por la izquierda con algún elemento no nulo α (que llamaremos *pivote*) e intercambiamos la primera fila con la que contiene el pivote.

3. Multiplicamos ahora la primera fila por $\dfrac{1}{\alpha}$ haciendo que el pivote valga 1.

4. Añadiendo a cada fila distinta de la primera un múltiplo escalar de la primera obtenemos 0's debajo del pivote.

5. Repetimos los pasos anteriores con las restantes filas, a partir de la columna del pivote.

Aplicando este algoritmo a cualquier matriz obtenemos una nueva matriz obtenemos una matriz *escalonada por filas* de manera que el primer elemento no nulo de cada fila es 1.

Si además deseamos hacer la sustitución regresiva más cómoda podemos hacer 0's encima de los pivotes:

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 2 \\ 0 & 1 & 4 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right) \xrightarrow{f_1-2\times f_2} \left(\begin{array}{ccc|c} 1 & 0 & -9 & 12 \\ 0 & 1 & 4 & -5 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

De este modo, las variables correspondientes a los pivotes quedan despejadas en función de las demás, que aparecen como parámetros.

2.2. Existencia de soluciones. Rango de matrices

Dado un sistema de ecuaciones $A\mathbf{x} = \mathbf{b}$:

$$\left\{\begin{array}{r} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \quad \vdots \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_n \end{array}\right.$$

realizamos operaciones elementales sobre la matriz ampliada para convertirla en una matriz escalonada:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix} \xrightarrow[\text{elementales}]{\text{operaciones}} \left(\begin{array}{ccccccccccc|c} 1 & \cdots & \cdot & \cdots & \cdot & \cdots & \cdots & \cdot & \cdots & & & \cdot \\ & & 1 & \cdots & \cdot & \cdots & \cdots & \cdot & \cdots & & & \cdot \\ & & & & 1 & \cdots & \cdots & \cdot & \cdots & & & \cdot \\ & & & & & & \ddots & \vdots & \vdots & & & \vdots \\ & & & & & & & 1 & \cdots & & & \cdot \\ & & & \text{0's} & & & & & & & & \alpha \\ & & & & & & & & & & & 0 \\ & & & & & & & & & & & \vdots \end{array}\right)$$

donde α es 0 o 1. Sean $1 \leq j_1 < \cdots < j_r$ los índices de las columnas donde aparecen los pivotes.

Aparecen las siguientes posibilidades para nuestro sistema $A\mathbf{x} = \mathbf{b}$:

Sistema compatible determinado: $\alpha = 0$ y $r = n$. En este caso la solución es única, al no quedar ninguna variable libre como parámetro.

Sistema compatible indeterminado: $\alpha = 0$ y $r < n$. En este caso hay infinitas soluciones. Las variables $x_{j_1}, \ldots, x_{j_r}$ quedan despejadas en función del resto de variables, que quedan como parámetros. En este caso, si $\mathbf{x}'$ y $\mathbf{x}''$ son soluciones, $A(\mathbf{x}' - \mathbf{x}'') = \mathbf{b} - \mathbf{b} = 0$, luego el conjunto de soluciones es:

$$\mathbf{x}' + \{\mathbf{y} \in \mathbb{R}^n : A\mathbf{y} = 0\}.$$

Esto es, toda solución es la suma de una solución particular dada y una solución del *sistema homogéneo* $A\mathbf{x} = 0$. Las soluciones del sistema homogéneo dependen de $n - r$ parámetros. Si $n - r = 1$, el conjunto de soluciones es una recta, si $n - r = 2$ es un plano, ...

Sistema incompatible: $\alpha = 1$ y no hay soluciones.

En el caso en que $b_1 = \cdots = b_m = 0$, el sistema se dice *homogéneo*. En este caso la columna derecha de la matriz ampliada es nula, luego tenemos solamente dos posibilidades:

(i) $r = n$ y la única solución es la solución trivial: $x_1 = \cdots = x_n = 0$,

(ii) $r < n$ y existen infinitas soluciones.

Si $\mathbf{s}_1, \mathbf{s_2} \in \mathbb{R}^n$ son dos soluciones del sistema homogéneo y $t \in \mathbb{R}$, también $\mathbf{s_1} + \mathbf{s_2}$ y $t\mathbf{s_1}$ son soluciones. Esto es, el conjunto de soluciones es cerrado para la suma y para el producto por escalares. Esto merece un nombre:

Definición. Los subconjuntos no vacíos de $\mathbb{R}^n$ cerrados para la suma y para el producto por escalares se llaman **subespacios**.

El algoritmo de Gauss implica que toda matriz $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ puede reducirse, mediante operaciones elementales de filas, a una matriz escalonada:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \xrightarrow[\text{elementales}]{\text{operaciones}} \begin{pmatrix} 1 & \cdots & \cdot & \cdots & \cdot & \cdots & \cdots & \cdot & \cdots \\ & & 1 & \cdots & \cdot & \cdots & \cdots & \cdot & \cdots \\ & & & & 1 & \cdots & \cdots & \cdot & \cdots \\ & & \text{0's} & & & & \ddots & \vdots & \vdots \\ & & & & & & & 1 & \cdots \end{pmatrix}$$

El número de filas no nulas r de esta matriz escalonada se llama **rango** de A: $r = \operatorname{rango}(A)$. Este número no depende de las operaciones elementales realizadas[2].

Las posibilidades contempladas antes las podemos reescribir así:

Teorema (Rouché-Frobenius)**.** *Dado el sistema de ecuaciones lineales* $A\mathbf{x} = \mathbf{b}$, $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, *se tiene:*

(i) *Si* $\operatorname{rango}(A) = \operatorname{rango}(A \mid \mathbf{b}) = n$ *el sistema es compatible determinado.*

(ii) *Si* $\operatorname{rango}(A) = \operatorname{rango}(A \mid \mathbf{b}) < n$ *el sistema es compatible indeterminado.*

(iii) *Si* $\operatorname{rango}(A) < \operatorname{rango}(A \mid \mathbf{b})$ *el sistema es incompatible.*

Corolario. *Una matriz* $A \in \mathbb{R}^{n\times n}$ *tiene inversa si y solo si* $\operatorname{rango}(A) = n$.

Demostración. Si $\operatorname{rango}(A) = n$ todo sistema de ecuaciones $A\mathbf{x} = \mathbf{b}$ tiene solución única. En particular podemos encontrar $B \in \mathbb{R}^{n\times n}$ tal que $AB = I_n$, resolviendo un sistema para encontrar cada columna de B. Además, de $A(BA - I_n) = (AB)A - A = I_nA = 0$, concluimos que cada columna de $BA - I_n$ es nula (de nuevo por la unicidad de soluciones), luego también $BA = I_n$, y A tiene inversa B.

Al revés, si A tiene inversa A^{-1}, la única solución de $A\mathbf{x} = 0$ es $\mathbf{x} = A^{-1}0 = 0$ y, por tanto, $\operatorname{rango}(A) = n$. □

Para obtener la inversa de una matriz A, debemos resolver un sistema para cada columna de A^{-1}, pero podemos resolver todos estos sistemas simultáneamente: basta efectuar operaciones elementales a las filas de A para reducirla a I_n (esto es posible pues $\operatorname{rango}(A) = n$), pero haciendo a la vez estas operaciones elementales a I_n:

$$(A \mid I_n) \xrightarrow[\text{elementales}]{\text{operaciones}} (I_n \mid A^{-1}).$$

[2] Si $1 \le j_1 < \cdots < j_r \le n$ son los índices de las columnas de los pivotes, entonces si $A^{j_1,\dots,j_r}$ es la matriz $m \times r$ formada por las columnas $j_1, \dots, j_r$ de A, el sistema homogéneo $A^{j_1,\dots,j_r}\mathbf{z} = 0$ solo tiene la solución trivial, mientras que el sistema homogéneo correspondiente a la matriz formada por cualesquiera $r + 1$ columnas de A tiene infinitas soluciones. Esto es, $r = \operatorname{rango}(A)$ si, y solamente si, hay r columnas de A tales que el correspondiente sistema homogéneo solo tiene la solución trivial, pero el sistema asociado a cualesquiera $r+1$ columnas tiene infinitas soluciones.

2.3. Rango y dimensión

Dado un subespacio S de $\mathbb{R}^n$, y vectores $\mathbf{v}_1, \dots, \mathbf{v}_r$ de S:

1. Decimos que estos vectores forman un *sistema generador* de S (o que generan S), si todo vector $\mathbf{v}$ de S es una 'combinación lineal' de $\mathbf{v}_1, \dots, \mathbf{v}_r$, esto es, si existen escalares $t_1, \dots, t_r \in \mathbb{R}$ tales que
$$\mathbf{v} = t_1\mathbf{v}_1 + \cdots + t_r\mathbf{v}_r.$$
Escribimos entonces $S = \langle \mathbf{v}_1, \dots, \mathbf{v}_r \rangle$

2. Decimos que estos vectores son *linealmente independientes* si la única combinación lineal de ellos que da el vector nulo es la que tiene todos los escalares nulos:
$$t_1\mathbf{v}_1 + \cdots + t_r\mathbf{v}_r = \mathbf{0} \quad \Longrightarrow \quad t_1 = \cdots = t_r = 0.$$
Esto equivale a que el sistema homogéneo[3]
$$\Big(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_r\Big)\mathbf{x} = 0$$
sólo tiene la solución trivial, es decir, a que $\text{rango}\Big(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_r\Big) = r$.

3. Decimos que estos vectores forman una *base* de S si son linealmente independientes y forman un sistema generador de S. Esto es, si todo vector de S se puede escribir de modo único como una combinación lineal de $\mathbf{v}_1, \dots, \mathbf{v}_n$.

Propiedades:

- Si S es un subespacio de $\mathbb{R}^n$ (por columnas), $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ es un sistema generador de S, y $\mathbf{v}_1, \dots, \mathbf{v}_s$ son vectores linealmente independientes de S, entonces $s \leq r$.

  En efecto, existen escalares tales que
$$\begin{cases} \mathbf{v}_1 = a_{11}\mathbf{u}_1 + \cdots + a_{r1}\mathbf{u}_r, \\ \vdots \qquad \vdots \\ \mathbf{v}_s = a_{1s}\mathbf{u}_1 + \cdots + a_{rs}\mathbf{u}_r, \end{cases}$$
  que podemos escribir en forma matricial como
$$\Big(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_s\Big) = \Big(\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_r\Big)\, A$$
  con $A = \big(a_{ij}\big) \in \mathbb{R}^{r \times s}$. Si $\text{rango}(A) < s$, entonces el sistema homogéneo $A\mathbf{x} = 0$ tiene alguna solución no trivial, y para cualquiera de estas soluciones $\mathbf{x}'$,
$$\Big(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_s\Big)\mathbf{x}' = \Big(\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_r\Big)\, A\mathbf{x}' = 0,$$
  luego $t_1\mathbf{v}_1 + \cdots + t_s\mathbf{v}_s = 0$, con $t_1, \dots, t_s$ las componentes de $\mathbf{x}'$. Esto contradice la independencia lineal. Así pues, se tiene necesariamente $\text{rango}(A) = s$, luego $s = \text{rango}(A) \leq r$, pues r es el número de filas de A.

[3] $\Big(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_r\Big)$ denota la matriz cuyas columnas vienen dadas por los vectores $\mathbf{v}_1, \dots, \mathbf{v}_r$.

- Como consecuencia, todas las bases de S tienen el mismo número de elementos, y este número se llama **dimensión** de S y se denota $\dim S$. Además las bases de S coinciden con los sistemas generadores de S con el menor número posible de elementos, y con las familias linealmente independientes de vectores de S con el mayor número posible de vectores.

- Dada una matriz $A \in \mathbb{R}^{m\times n}$ el *espacio de filas* de A es el subespacio de $\mathbb{R}^n$ generado por sus filas. Análogamente tenemos el *espacio de columnas* de A, que es subespacio de $\mathbb{R}^m$. Entonces

$$\text{rango}(A) = \dim\big(\text{espacio de filas de } A\big)$$

pues al hacer operaciones elementales el espacio de filas no cambia, y **las filas no nulas de la matriz escalonada obtenida por el Método de Gauss forman una base de este espacio**. Además, se tiene también:

$$\text{rango}(A) = \dim\big(\text{espacio de columnas de } A\big)$$

pues **las columnas de A donde aparecen los pivotes forman una base de este espacio**.

Este último resultado nos da una estrategia para **calcular bases de subespacios**, una vez que tenemos un sistema generador:

Formamos la matriz cuyas filas son los vectores del sistema generador y la escalonamos mediante operaciones elementales por el Método de Gauss.

§ 3. Ajuste de datos por mínimos cuadrados

Dado un subespacio S de $\mathbb{R}^n$, el subespacio

$$S^\perp := \{\mathbf{v} \in \mathbb{R}^n : \mathbf{u} \bullet \mathbf{v} = 0 \ \forall \mathbf{u} \in S\}$$

se dice **subespacio ortogonal** a S.

Propiedades:

- Todo vector $\mathbf{a} \in \mathbb{R}^n$ (por columnas) se descompone de modo único como $\mathbf{a} = \mathbf{u} + \mathbf{v}$ con $\mathbf{u} \in S$ y $\mathbf{v} \in S^\perp$.

Demostración. Sea $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$ una base de S. Buscamos un vector $\mathbf{u} = t_1\mathbf{u}_1 + \cdots + t_r\mathbf{u}_r$ de modo que $\mathbf{a} - \mathbf{u} \in S^\perp$, esto es:

$$\mathbf{u}_i \bullet \big(t_1\mathbf{u}_1 + \cdots + t_r\mathbf{u}_r - \mathbf{a}\big) = 0$$

es decir, debemos resolver el sistema:

$$\begin{pmatrix} \mathbf{u}_1^\top \\ \hline \vdots \\ \hline \mathbf{u}_r^\top \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \mid \cdots \mid \mathbf{u}_r \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} \mathbf{u}_1^\top \\ \hline \vdots \\ \hline \mathbf{u}_r^\top \end{pmatrix} \mathbf{a}$$

El correspondiente sistema homogéneo solo tiene la solución trivial, pues

$$\begin{aligned}
&\begin{pmatrix} \mathbf{u}_1^\top \\ \hline \vdots \\ \hline \mathbf{u}_r^\top \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \mid \cdots \mid \mathbf{u}_r \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = 0 \\
&\iff \big(x_1\mathbf{u}_1 + \cdots + x_r\mathbf{u}_r\big) \bullet \big(x_1\mathbf{u}_1 + \cdots + x_r\mathbf{u}_r\big) \\
&\iff x_1\mathbf{u}_1 + \cdots + x_r\mathbf{u}_r = 0 \\
&\iff x_1 = \cdots = x_r = 0.
\end{aligned}$$

luego el rango de la matriz es r y el sistema es compatible determinado. □

- El vector $\mathbf{u}$ anterior se llama la *proyección ortogonal* de $\mathbf{a}$ en S: $\mathbf{u} = \operatorname{proy}_S(\mathbf{a})$. Este vector verifica que $\|\mathbf{a}-\mathbf{u}\| < \|\mathbf{a}-\mathbf{u}'\|$ para todo $\mathbf{u}' \in S \setminus \{\mathbf{u}\}$. Es decir, $\mathbf{u}$ es el *vector de S más cercano al vector* $\mathbf{a}$.

  Demostración.

$$\begin{aligned}
\|\mathbf{a}-\mathbf{u}'\|^2 &= \|(\mathbf{a}-\mathbf{u}) + (\mathbf{u}-\mathbf{u}')\| = \|\mathbf{v} - (\mathbf{u}-\mathbf{u}')\| \\
&= \big(\mathbf{v} - (\mathbf{u}-\mathbf{u}')\big) \bullet \big(\mathbf{v} - (\mathbf{u}-\mathbf{u}')\big) \\
&= \|\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{u}'\|^2 \quad \text{ya que } \mathbf{v} \bullet (\mathbf{u}-\mathbf{u}') = 0 \\
&> \|\mathbf{v}\|^2 = \|\mathbf{a}-\mathbf{u}\|^2.
\end{aligned}$$

  □

- Si $A \in \mathbb{R}^{m\times n}$ y $\mathbf{b} \in \mathbb{R}^m$, entonces un vector $\mathbf{x} \in \mathbb{R}^n$ verifica que $\|A\mathbf{x} - \mathbf{b}\|$ es mínima si y solo si $\mathbf{x}$ es solución del sistema de ecuaciones lineales $A^\top A\mathbf{x} = A^\top\mathbf{b}$.

  Demostración. Consideremos las columnas de A: $A = \big(\mathbf{a}_1 \mid \cdots \mid \mathbf{a}_n\big)$. Sea S el correspondiente espacio de columnas. Dado un vector $\mathbf{x} \in \mathbb{R}^n$ se tiene

$$\begin{aligned}
\|A\mathbf{x} - b\| \text{ es mínimo} \quad &\iff \quad x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \operatorname{proy}_S(\mathbf{b}) \\
&\iff \quad x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n - \mathbf{b} \in S^\perp \\
&\iff \quad \mathbf{a}_i \bullet (x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n - \mathbf{b}) = 0 \ \forall i = 1, \dots, n \\
&\iff \quad A^\top A\mathbf{x} = A^\top\mathbf{b}.
\end{aligned}$$

  (Notemos que estos sistemas son siempre compatibles, pues la proyección ortogonal existe siempre.) □

Ejemplo. Tenemos una serie de datos (x_i, y_i), $i = 1, \dots, N$ que relacionan dos variables x e y y sospechamos que estas variables están relacionadas por una ecuación de la forma $y = a + bx + cx^2$ (parábola). El objetivo es encontrar la parábola que más se ajuste a nuestros datos.

Si hubiera una parábola que pasara exactamente por nuestros datos, sus coeficientes verificarían

$$\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_N & x_N^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}$$

Es pues razonable buscar los valores de a, b, c que hagan mínima la norma

$$\left\| \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_N & x_N^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} \right\|$$

esto es, que hagan mínima la suma de cuadrados:

$$\sum_{i=1}^{N} \Big(y_i - (a + bx_i + cx_i^2) \Big)^2 .$$

La solución se obtiene, por tanto, resolviendo el sistema (con incógnitas a, b, c):

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_N \\ x_1^2 & x_2^2 & \cdots & x_N^2 \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_N & x_N^2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_N \\ x_1^2 & x_2^2 & \cdots & x_N^2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}$$

§ 4. Determinantes

Definimos la aplicación *determinante*

$$\begin{aligned} \det : \mathbb{R}^{n \times n} &\longrightarrow \mathbb{R} \\ A &\mapsto \det(A) = |A|, \end{aligned}$$

inductivamente mediante:

$n = 1$: $A = (a)$, $|A| := a$.

$n > 1$: $A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}}$, suponiendo que ya hemos definido el determinante para matrices en $\mathbb{R}^{(n-1)\times(n-1)}$, llamamos *adjunto* (ij) de A al número:

$$A_{ij} := (-1)^{i+j} \begin{vmatrix} a_{11} & \cdots & a_{1\,j-1} & a_{1\,j+1} & \cdots a_{1n} & \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1\,1} & \cdots & a_{i-1\,j-1} & a_{i-1\,j+1} & \cdots & a_{i-1\,n} \\ a_{i+1\,1} & \cdots & a_{i+1\,j-1} & a_{i+1\,j+1} & \cdots & a_{i+1\,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n\,j-1} & a_{n\,j+1} & \cdots & a_{nn} \end{vmatrix}$$

(esto es $(-1)^{i+j}$ veces el determinante de la matriz que resulta al suprimir la fila i y la columna j de A)), y definimos

$$|A| := a_{11}A_{11} + \cdots + a_{n1}A_{n1}$$

(elementos de la primera columna por sus adjuntos).

Así

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12},$$

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

Propiedades:

1. $|I_n| = 1, \quad \begin{vmatrix} \alpha_1 & * & * & \cdots & * \\ 0 & \alpha_2 & * & \cdots & * \\ 0 & 0 & \alpha_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_n \end{vmatrix} = \alpha_1\alpha_2\alpha_3 \cdots \alpha_n.$

   Esto es, el determinante de una matriz triangular superior es el producto de los elementos de su diagonal.

2. Sea P_{ij} la matriz obtenida permutando las filas i y j de I_n. Para toda matriz $A \in \mathbb{R}^{n\times n}$, $P_{ij}A$ es la matriz obtenida permutando las filas i y j de A. Entonces,

   $$|P_{ij}A| = |P_{ij}||A| = -|A|,$$

   es decir, al intercambiar dos filas de A el determinante cambia de signo. En particular, si dos filas de A son iguales, su determinante es 0.

3. Sea $P_i(t)$ $(0 \neq t \in \mathbb{R})$ la matriz obtenida multiplicando la fila i de I_n por t. Para toda matriz $A \in \mathbb{R}^{n\times n}$, $P_i(t)A$ es la matriz obtenida multiplicando la fila i de A por t. Entonces,

   $$|P_i(t)A| = |P_i(t)||A| = t|A|,$$

   es decir, al multiplicar una fila por t, el determinante queda multiplicado por t. Por tanto, podemos sacar 'factor común' en cada fila. En particular, si A tiene una fila nula, su determinante es 0.

4. Si una fila de A es suma de dos vectores, el determinante de A es la correspondiente suma:

   $$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a'_{i1} + a''_{i1} & \cdots & a'_{in} + a''_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a'_{i1} & \cdots & a'_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a''_{i1} & \cdots & a''_{in} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

5. Sea $P_{ij}(t)$ ($t \in \mathbb{R}$) la matriz obtenida sumando a la fila i de I_n la fila j multiplicada por t. Para toda matriz $A \in \mathbb{R}^{n\times n}$, $P_{ij}(t)A$ es la matriz obtenida sumando a la fila i de A la j multiplicada por t. Entonces,

$$|P_{ij}(t)A| = |P_{ij}(t)||A| = |A|,$$

Las matrices P_{ij}, $P_i(t)$ y $P_{ij}(t)$ de los apartados 2, 3 y 5 se llaman *matrices elementales*. Todas ellas tienen inversa: $P_{ij}^{-1} = P_{ij}$, $P_i(t)^{-1} = P_i(t^{-1})$, $P_{ij}(t) = P_{ij}(-t)$. Realizar operaciones elementales con las filas de una matriz equivale a multiplicar por la izquierda por matrices elementales, luego si

$$A \xrightarrow[\text{elementales}]{\textit{operaciones}} B$$

entonces (usando las inversas), se tiene que $A = Q_1 \cdots Q_s B$, para algunas matrices elementales $Q_1, \ldots, Q_s$, y $|A| = |Q_1| \cdots |Q_s||B|$.

6. A es *regular* (esto es, tiene inversa) si y solo si $|A| \neq 0$.

Si A es regular, entonces la podemos transformar en I_n mediante operaciones elementales, luego $|A|$ es un producto de determinantes de matrices elementales y, por tanto, es no nulo. Por el contrario, si A no es regular la podemos transformar mediante operaciones elementales en una matriz con la última fila nula, luego su determinante es 0.

7. $|AB| = |A||B|$.

Si A es regular, es un producto de matrices elementales y podemos aplicar los puntos 2,3 y 5. Si A no es regular, multiplicando a izquierda por matrices elementales la podemos convertir en una matriz U con la última fila nula. Al multiplicar AB por las mismas matrices elementales obtenemos UB que también tiene la última fila nula. Así pues, AB tampoco es regular y tenemos $|AB| = 0 = |A||B|$.

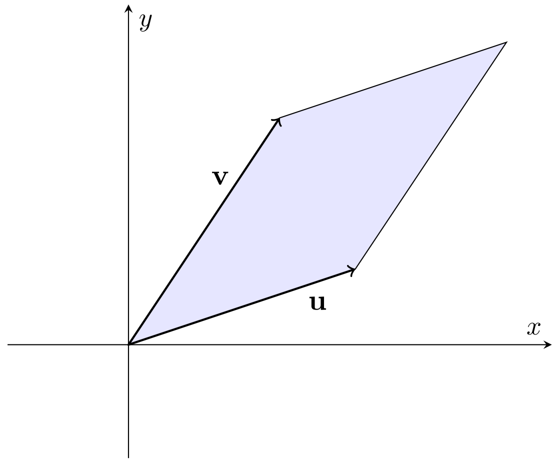
8. $|A| = |A^\top|$.

Si A es regular, se deduce de que el resultado es cierto trivialmente para las matrices elementales. Si A no es regular, tampoco lo es $A^\top$ y $|A| = 0 = |A^\top|$.

Por tanto, las propiedades de los apartados 1–5 se verifican cambiando filas por columnas, y podemos 'desarrollar' un determinante por cualquier fila o columna.

La estrategia adecuada para el cálculo de determinantes es la utilización de operaciones elementales por filas o columnas para simplificar la matriz.

9. El área del paralelogramo determinado por dos vectores no nulos $\mathbf{u}, \mathbf{v}$ de $\mathbb{R}^2$ es el valor absoluto del determinante de $\big(\mathbf{u} \mid \mathbf{v}\big)$.



Análogamente, el volumen del paralelepípedo determinado por tres vectores no nulos $\mathbf{u}, \mathbf{v}, \mathbf{w}$ de $\mathbb{R}^3$ es el valor absoluto del determinante de $\begin{pmatrix}\mathbf{u} \mid \mathbf{v} \mid \mathbf{w}\end{pmatrix}$.[4]

Los vectores $\mathbf{u}$ y $\mathbf{v}$ son linealmente dependientes si y solo si el área del paralelogramo generado es 0, y también si y solo si el determinante de $\begin{pmatrix}\mathbf{u} \mid \mathbf{v}\end{pmatrix}$ es 0.

Si los vectores son linealmente independientes, denotemos por $A(\mathbf{u}, \mathbf{v})$ el área del paralelogramo generado. Puesto que la altura sobre $\mathbf{u}$ de este paralelogramo es la misma que la altura del paralelogramo generado por $\mathbf{u}$ y $\mathbf{v} + t\mathbf{u}$ para todo $t \in \mathbb{R}$, y análogamente si tomamos alturas sobre $\mathbf{v}$, tenemos

$$A(\mathbf{u}, \mathbf{v}) = A(\mathbf{u}, \mathbf{v} + t\mathbf{u}) = A(\mathbf{u} + s\mathbf{v}, \mathbf{v})$$

para cualesquiera $t, s \in \mathbb{R}$. Si $\mathbf{u} = \begin{pmatrix} a \\ b \end{pmatrix}$, $\mathbf{v} = \begin{pmatrix} c \\ d \end{pmatrix}$ y, por ejemplo, $a \neq 0$ (en otro caso $c \neq 0$ y se procede análogamente), nos queda

$$\begin{aligned} A(\mathbf{u}, \mathbf{v}) &= A\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right) = A\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} - \frac{c}{a}\begin{pmatrix} a \\ b \end{pmatrix}\right) = A\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} 0 \\ d - \frac{c}{a}b \end{pmatrix}\right) \\ &= A\left(\begin{pmatrix} a \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ d - \frac{c}{a}b \end{pmatrix}\right) = \left|a\left(d - \frac{c}{a}b\right)\right| = |ad - bc|, \end{aligned}$$

que es el valor absoluto del determinante de $\begin{pmatrix}\mathbf{u} \mid \mathbf{v}\end{pmatrix}$.

Nota. Hay otro número importante, además del determinante, asociado a toda matriz cuadrada, y es la suma de los elementos de la diagonal, que se llama **traza** de la matriz:

$$\operatorname{traza}(A) = a_{11} + \cdots + a_{nn}.$$

[4] Para dimensiones superiores, estas propiedades se emplean para definir adecuadamente el ‘hipervolumen’.

§ 5. Valores y vectores propios. Aplicaciones

5.1. Valores y vectores propios

Definición. Sea $A \in \mathbb{R}^{n\times n}$ y $\lambda \in \mathbb{R}$. λ se dice **valor propio** de A si existe un vector no nulo $\mathbf{x} \in \mathbb{R}^n$ tal que $A\mathbf{x} = \lambda\mathbf{x}$. En este caso, $\mathbf{x}$ se dice **vector propio**.

Teorema. *λ es valor propio de la matriz A si y solo si $|\lambda I_n - A| = 0$, esto es, λ es una raíz del **polinomio característico** de A: $|xI_n - A|$.*

Demostración. λ es valor propio de A si y solo si el sistema homogéneo $(\lambda I_n - A)\mathbf{x} = 0$ tiene una solución no trivial, y esto ocurre si y solo si rango$(\lambda I_n - A) < n$, si y solo si $|\lambda I_n - A| = 0$. □

Ejemplo. Los valores propios de $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ son las raíces de $\begin{vmatrix} x-1 & -2 \\ -3 & x-2 \end{vmatrix} = (x-1)(x-2) - 6 = x^2 - 3x - 4 = (x+1)(x-4)$, luego los valores propios son -1 y 4.

Para el valor propio $\lambda = -1$, los correspondientes vectores propios son las soluciones del sistema homogéneo

$$\begin{pmatrix} -2 & -2 \\ -3 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

esto es, los vectores $\begin{pmatrix} t \\ -t \end{pmatrix}$, $0 \neq t \in \mathbb{R}$. Para $\lambda = 4$ son las soluciones de

$$\begin{pmatrix} 3 & -2 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

esto es, los vectores $\begin{pmatrix} 2t \\ 3t \end{pmatrix}$, $0 \neq t \in \mathbb{R}$.

Propiedades: Sea $A \in \mathbb{R}^{n\times n}$ y $\lambda \in \mathbb{R}$ un valor propio de A.

- $S(A,\lambda) := \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \lambda\mathbf{x}\}$ es un subespacio (solución de un sistema de ecuaciones homogéneo), y su dimensión es

  $$\dim S(A,\lambda) = n - \text{rango}(\lambda I_n - A).$$

  $S(A,\lambda)$ se dice **subespacio fundamental** de A de valor propio λ.

- Vectores propios correspondientes a valores propios distintos son siempre linealmente independientes.

  En efecto, si $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1, \ldots, A\mathbf{v}_s = \lambda_s\mathbf{v}_s$, con los $\mathbf{v}_i$'s no nulos y los λ_i's distintos y si tenemos una combinación lineal

  $$\mu_1\mathbf{v}_1 + \cdots + \mu_s\mathbf{v}_s = 0,$$

  debemos probar que los μ_i's son todos 0. Esto es trivial si $s = 1$. Supongamos que lo podemos probar para $s-1$ vectores propios. Multiplicando nuestra combinación lineal por A, obtenemos

  $$\mu_1\lambda_1\mathbf{v}_1 + \cdots + \mu_s\lambda_s\mathbf{v}_s = 0,$$

  y multiplicando la primera combinación lineal por λ_1 y restándole la segunda:

  $$\mu_2(\lambda_1 - \lambda_2)\mathbf{v}_2 + \cdots + \mu_s(\lambda_1 - \lambda_s)\mathbf{v}_s = 0,$$

  de donde obtenemos (por ser $\lambda_1 \neq \lambda_i \ \forall i = 2, \ldots, s$) $\mu_2 = \cdots = \mu_s = 0$. Esto implica también $\mu_1 = 0$.

- $\dim S(A, \lambda)$ es menor o igual que la multiplicidad de λ como raíz del polinomio característico $|xI_n - A|$.

  Demostración. Sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ una base de $S(A, \lambda)$, y completamos (añadiendo en cada paso un vector fuera del subespacio generado por los anteriores) hasta una base de $\mathbb{R}^n$: $\{\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{u}_{r+1}, \ldots, \mathbf{u}_n\}$. La matriz

  $$P = \Big(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n\Big)$$

  es regular y tenemos

  $$AP = A\Big(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n\Big) = \Big(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n\Big) \begin{pmatrix} \lambda & 0 & \cdots & 0 & * & \cdots & * \\ 0 & \lambda & \cdots & 0 & * & \cdots & * \\ & & \ddots & \vdots & \vdots & \ddots & \vdots \\ & & & \lambda & * & \cdots & * \\ & \text{0's} & & & * & \cdots & * \\ & & & & \vdots & \vdots & \vdots \end{pmatrix}$$

  donde la última matriz tiene exactamente r λ's en la diagonal. Así

  $$A = P \begin{pmatrix} \lambda & 0 & \cdots & 0 & * & \cdots & * \\ 0 & \lambda & \cdots & 0 & * & \cdots & * \\ & & \ddots & \vdots & \vdots & \ddots & \vdots \\ & & & \lambda & * & \cdots & * \\ & \text{0's} & & & * & \cdots & * \\ & & & & \vdots & \vdots & \vdots \end{pmatrix} P^{-1}$$

y, por tanto,

$$
\begin{aligned}
|xI_n - A| &= \left| xI_n - P \begin{pmatrix} \lambda & 0 & \cdots & 0 & * & \cdots & * \\ 0 & \lambda & \cdots & 0 & * & \cdots & * \\ & & \ddots & \vdots & \vdots & \ddots & \vdots \\ & & & \lambda & * & \cdots & * \\ & \text{0's} & & & * & \cdots & * \\ & & & & \vdots & \vdots & \vdots \end{pmatrix} P^{-1} \right| \\
&= \left| P \left(xI_n - \begin{pmatrix} \lambda & 0 & \cdots & 0 & * & \cdots & * \\ 0 & \lambda & \cdots & 0 & * & \cdots & * \\ & & \ddots & \vdots & \vdots & \ddots & \vdots \\ & & & \lambda & * & \cdots & * \\ & \text{0's} & & & * & \cdots & * \\ & & & & \vdots & \vdots & \vdots \end{pmatrix} \right) P^{-1} \right| \\
&= \left| xI_n - \begin{pmatrix} \lambda & 0 & \cdots & 0 & * & \cdots & * \\ 0 & \lambda & \cdots & 0 & * & \cdots & * \\ & & \ddots & \vdots & \vdots & \ddots & \vdots \\ & & & \lambda & * & \cdots & * \\ & \text{0's} & & & * & \cdots & * \\ & & & & \vdots & \vdots & \vdots \end{pmatrix} \right| = (x-\lambda)^r q(x)
\end{aligned}
$$

para cierto polinomio $q(x)$. Por tanto, la dimensión r es, como máximo, la multiplicidad de λ en el polinomio característico. $\square$

Por tanto, la suma $\sum_{\lambda \text{ valor propio}} \dim S(A, \lambda)$ es menor o igual que la suma de las multiplicidades de las raíces, que es a lo sumo igual al grado del polinomio característico:

$$\sum_{\lambda \text{ valor propio}} \dim S(A, \lambda) \leq n.$$

- Las siguientes condiciones son equivalentes:
  1. Existe $P \in \mathbb{R}^{n \times n}$ regular tal que $P^{-1}AP$ es una matriz diagonal.
  2. Existe una base de $\mathbb{R}^n$ formada por vectores propios de A.
  3. Juntando las bases de los distintos subespacios fundamentales se obtiene una base de $\mathbb{R}^n$.
  4. El polinomio característico $|xI_n - A|$ tiene todas sus raíces reales, y para toda raíz λ se tiene que $\dim S(A, \lambda)$ coincide con la multiplicidad de λ.

Para probar esto basta darse cuenta de que si $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base formada por vectores propios: $A\mathbf{v}_i = \lambda_i \; \forall i$ (los λ_i's no necesariamente distintos), entonces

$$A\Big(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n\Big) = \Big(\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n\Big) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

esto es, $P^{-1}AP = \Lambda$, donde $P = \begin{pmatrix}\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_n\end{pmatrix}$ y Λ es la matriz diagonal con $\lambda_1, \ldots, \lambda_n$ en la diagonal.

Si se cumplen estas condiciones equivalentes, la matriz A se dice **diagonalizable**.

Ejemplo 6.2. Sabemos que los valores propios de $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ son -1 y 4. Además, $S(A,-1) = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle$ y $S(A,4) = \left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\rangle$. Esto nos da

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

esto es, $AP = P\Lambda$, con la matriz regular $P = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$ y la matriz diagonal $\Lambda = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}$.

5.2. Aplicaciones

En el siguiente capítulo veremos las aplicaciones de las matrices diagonalizables a la resolución de ciertos sistemas de ecuaciones diferenciales. La clave para estas aplicaciones es el hecho de que es muy fácil calcular las potencias de una matriz diagonalizable:

$$P^{-1}AP = \Lambda \implies A = P\Lambda P^{-1} \implies A^m = \left(P\Lambda P^{-1}\right)^m = P\Lambda^m P^{-1}.$$

Ejemplo. (Matrices de Leslie, 1945) Las *matrices de Leslie* se usan en ecología para medir la evolución de poblaciones estructuradas en edades. Son matrices cuadradas en las cuales todos los elementos son cero excepto, quizá, los situados en la primera fila y los situados en la primera subdiagonal debajo de la diagonal principal. La dinámica de evolución de la población viene dada por la siguiente ecuación:

$$\begin{pmatrix} N_0(k+1) \\ N_1(k+1) \\ N_2(k+1) \\ N_3(k+1) \\ \vdots \\ N_{n-1}(k+1) \end{pmatrix} = \begin{pmatrix} f_0 & f_1 & f_2 & \cdots & f_{n-2} & f_{n-1} \\ s_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{n-2} & 0 \end{pmatrix} \begin{pmatrix} N_0(k) \\ N_1(k) \\ N_2(k) \\ N_3(k) \\ \vdots \\ N_{n-1}(k) \end{pmatrix}$$

donde $N_i(t)$ indica el número de individuos de la clase i en el instante (discreto) t.

Por ejemplo, si una población de organismos está estructurada en tres grupos de edad y

- El número de hembras en el primer grupo de edad (edad 0) en el instante $k+1$ es el número de hembras nacidas en ese instante.
- El número de hembras en el segundo grupo de edad (edad 1) en el instante $k+1$ es la mitad de las hembras del primer grupo de edad en el instante k (un 50 % de supervivencia):

$$N_1(k+1) = \frac{1}{2}N_0(k).$$

- El número de hembras en el tercer grupo de edad (edad 2) en el instante $k+1$ es un tercio del número de hembras en el segundo grupo en el instante k:
$$N_2(k+1) = \frac{1}{3}N_1(k).$$
- Las hembras del primer grupo de edad se reproducen con una media de $\frac{14}{9}$ descendientes hembras y las del segundo grupo de edad con una media de $\frac{4}{3}$:
$$N_0(k+1) = \frac{14}{9}N_1(k) + \frac{4}{3}N_2(k).$$

Así, la evolución del número de hembras viene dada por la ecuación matricial:
$$\begin{pmatrix} N_0(k+1) \\ N_1(k+1) \\ N_2(k+1) \end{pmatrix} = \begin{pmatrix} 0 & \frac{14}{9} & \frac{4}{3} \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{pmatrix} \begin{pmatrix} N_0(k) \\ N_1(k) \\ N_2(k) \end{pmatrix}$$
Si denotamos por N_i el valor inicial $N_i(0)$, $i = 0, 1, 2$. La población de hembras en el instante k es
$$\begin{pmatrix} N_0(k) \\ N_1(k) \\ N_2(k) \end{pmatrix} = \begin{pmatrix} 0 & \frac{14}{9} & \frac{4}{3} \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{pmatrix}^k \begin{pmatrix} N_0 \\ N_1 \\ N_2 \end{pmatrix}$$
Para determinar la evolución calculamos los valores propios de la matriz de Leslie $A = \begin{pmatrix} 0 & \frac{14}{9} & \frac{4}{3} \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \end{pmatrix}$:
$$\begin{vmatrix} x & -\frac{14}{9} & -\frac{4}{3} \\ -\frac{1}{2} & x & 0 \\ 0 & -\frac{1}{3} & x \end{vmatrix} = x^3 - \frac{7}{9}x - \frac{2}{9} = (x-1)\left(x+\frac{1}{3}\right)\left(x+\frac{2}{3}\right)$$
Los subespacios fundamentales son:
$$S(A,1) = \left\langle \begin{pmatrix} 6 \\ 3 \\ 1 \end{pmatrix} \right\rangle,$$
$$S\left(A,-\frac{1}{3}\right) = \left\langle \begin{pmatrix} 2 \\ -3 \\ 3 \end{pmatrix} \right\rangle,$$
$$S\left(A,-\frac{2}{3}\right) = \left\langle \begin{pmatrix} 8 \\ -6 \\ 3 \end{pmatrix} \right\rangle,$$

luego si $P = \begin{pmatrix} 6 & 2 & 8 \\ 3 & -3 & -6 \\ 1 & 3 & 3 \end{pmatrix}$, tenemos

$$AP = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{2}{3} \end{pmatrix}$$

y ahora podemos calcular fácilmente las potencias:

$$A^k = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{2}{3} \end{pmatrix}^k P^{-1} = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & \left(-\frac{1}{3}\right)^k & 0 \\ 0 & 0 & \left(-\frac{2}{3}\right)^k \end{pmatrix} P^{-1}$$

Al aumentar k, $\left(-\frac{1}{3}\right)^k$ y $\left(-\frac{2}{3}\right)^k$ tienden a 0, luego A^k tiende hacia $P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P^{-1}$ y

$$\begin{pmatrix} N_0(k) \\ N_1(k) \\ N_2(k) \end{pmatrix} \xrightarrow[k\to\infty]{} P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P^{-1} \begin{pmatrix} N_0 \\ N_1 \\ N_2 \end{pmatrix} = \begin{pmatrix} 6 & 0 & 0 \\ 3 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} P^{-1} \begin{pmatrix} N_0 \\ N_1 \\ N_2 \end{pmatrix} = N_0' \begin{pmatrix} 6 \\ 3 \\ 1 \end{pmatrix}$$

donde $\begin{pmatrix} N_0' \\ N_1' \\ N_2' \end{pmatrix} = P^{-1} \begin{pmatrix} N_0 \\ N_1 \\ N_2 \end{pmatrix}$.

Sin necesidad de calcular P^{-1} observamos que la población tiende a un **equilibrio** con 60 % de hembras del primer grupo, 30 % del segundo y 10 % del tercero.

La situación del ejemplo anterior se repite a menudo.

Definición. Se dice que un valor propio λ de una matriz $A \in \mathbb{R}^{n\times n}$ es el *valor propio dominante* si $1 \leq \lambda \in \mathbb{R}$, la multiplicidad de λ es 1, y $\lambda > |\mu|$ para cualquier otro valor propio (posiblemente complejo) de A.

En esta situación, dado 'casi' cualquier vector $\mathbf{v} \in \mathbb{R}^n$, los vectores $A^k\mathbf{v}$ tienden a acercarse hacia el subespacio fundamental $S(A, \lambda)$.[5]

[5] Si A es diagonalizable y $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base de vectores propios: $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$, $i = 1, \ldots, n$, con $\lambda_1 = \lambda$ el valor propio dominante, entonces si $\mathbf{v} = t_1\mathbf{v}_1 + \cdots + t_n\mathbf{v}_n$, con $0 \neq t_1, \ldots, t_n \in \mathbb{R}$, y

$$A^k\mathbf{v} = t_1\lambda_1^k\mathbf{v}_1 + \cdots + t_n\lambda^k\mathbf{v}_n = \lambda_1^k \left(t_1\mathbf{v}_1 + t_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k \mathbf{v}_2 + \cdots + t_n \left(\frac{\lambda_n}{\lambda_1}\right)^k \mathbf{v}_n \right).$$

Como $\dfrac{\lambda_i}{\lambda_1}$ tiene valor absoluto < 1, sus potencias tienden a 0, luego los vectores $A^k\mathbf{v}$ tienden a acercarse hacia $S(A, \lambda_1)$.
El caso en que A no es diagonalizable requiere otra demostración.

Ejemplo. (Sucesión de Fibonacci) La sucesión de Fibonacci se define recursivamente así:

$$x_0 = 0, \quad x_1 = 1, \quad x_{n+2} = x_{n+1} + x_n \ \forall n \geq 0.$$

Sus primeros términos son $0, 1, 1, 2, 3, 5, 8, 13, 21, \ldots.$

Para calcular su término general consideramos $X_n = \begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix}$, $\forall n \geq 0$. Así tenemos:

$$X_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad X_{n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} X_n.$$

De aquí se concluye que $X_n = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n X_0$ para todo $n \geq 0$, y de nuevo tenemos que calcular las potencias de la matriz $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.

Calculamos los valores propios:

$$|xI_2 - A| = \begin{vmatrix} x & -1 \\ -1 & x-1 \end{vmatrix} = x(x-1) - 1 = x^2 - x - 1$$

y, por tanto, los valores propios son $\lambda_1 = \dfrac{1+\sqrt{5}}{2}$ y $\lambda_2 = \dfrac{1-\sqrt{5}}{2}$. Así pues, existe una matriz regular P tal que $P^{-1}AP = D$ con $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, luego $X_n = A^n X_0 = PD^nP^{-1}X_0$ y concluimos, sin necesidad de calcular P y P^{-1} que existen números reales a y b tales que $x_n = a\lambda_1^n + b\lambda_2^n$. Para $n = 0$ y $n = 1$ obtenemos $a + b = 0$ y $a\lambda_1 + b\lambda_2 = 1$, de donde concluimos $a = -b = \dfrac{1}{\lambda_1 - \lambda_2} = \dfrac{1}{\sqrt{5}}$. Por tanto,

$$x_n = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2} \right)^n$$

para todo $n \geq 0$.

Capítulo 7

Ecuaciones diferenciales

Una **ecuación diferencial ordinaria** (o EDO) es una ecuación que involucra una función $y = f(x)$ y sus derivadas.

§ 1. Definiciones básicas

El orden más alto de la derivada involucrada se dice *orden de la ecuación.*

Por ejemplo, la ecuación

$$3x^2y'' + xy' - y = 0 \tag{7.1}$$

es una ecuación diferencial ordinaria de orden 2.

Las ecuaciones diferenciales ordinarias de orden 1 serán el objeto principal de estudio en este capítulo.

Una *solución particular* en un intervalo es toda función $y = f(x)$ (escribiremos también $y = y(x)$) que satisface la ecuación en ese intervalo. Por ejemplo,

$$y = x^{\frac{-1}{3}}$$

es una solución de (7.1) en $(0, +\infty)$, pues $y' = -\frac{1}{3}x^{\frac{-4}{3}}$ e $y'' = \frac{4}{9}x^{\frac{-7}{3}}$, luego

$$3x^2y'' + xy' - y = \frac{4}{3}x^{\frac{-1}{3}} - \frac{1}{3}x^{\frac{-1}{3}} - x^{\frac{-1}{3}} = 0.$$

Pero existen más soluciones, como $y = x$, o $y = 5x - 7x^{\frac{-1}{3}}$.

La ecuación suele venir acompañada de unas *condiciones iniciales* que ayudan a determinar la solución que estamos buscando. Así, la solución $y = x + x^{\frac{-1}{3}}$ es la solución a (7.1) que verifica $y(8) = \frac{17}{2}$ e $y'(8) = \frac{47}{48}$.

Un *problema de valor inicial* es una ecuación diferencial junto a unas condiciones iniciales.

La *solución general* de una EDO es la forma más general que puede tomar la solución sin tener en cuenta las condiciones iniciales, mientras que una *solución particular* sí las tiene en cuenta.

Ejemplo 7.2. La solución general de la ecuación

$$y' = f(x)y$$

es

$$\boxed{y = Ke^{F(x)}}$$

donde $F(x)$ es una primitiva de $f(x)$ (esto es, $F'(x) = f(x)$).

En efecto, si y es una solución, entonces $(e^{-F(x)}y)' = -F'(x)e^{-F(x)}y + e^{-F(x)}y' = -f(x)e^{-F(x)}y + e^{-F(x)}y' = e^{-F(x)}(y' - f(x)y) = 0$, luego $e^{-F(x)}y$ es constante. Si K denota esta constante, entonces $y = Ke^{F(x)}$.

Alternativamente, podemos resolver esta ecuación como sigue. Suponiendo que y no es la solución trivial $y = 0$, tenemos $\dfrac{y'}{y} = f(x)$, luego $(\ln|y|)' = f(x)$ y $\ln|y| = \displaystyle\int f(x)dx = F(x) + C$ para alguna constante C. Así $y = \pm e^{F(x)+C} = Ke^{F(x)}$, siendo K la constante $\pm e^C$.

Ejemplo. Consideramos el problema de valor inicial:

$$2xy' + 4y = 3, \quad y(1) = -4.$$

La función $y = \frac{3}{4}$ es una solución de nuestra ecuación, si $z = z(x)$ es otra solución, entonces la función $w = z - \frac{3}{4}$ verifica $2xw' + 4w = 0$, luego $w' = -\dfrac{2}{x}w$. Por el Ejemplo 7.2, $w = Ke^{-2\ln|x|} = Ke^{-\ln(x^2)} = \dfrac{K}{x^2}$ para alguna constante K.

Por tanto la solución general es

$$y = \frac{3}{4} + \frac{K}{x^2}$$

Al imponer $y(1) = -4$ se tiene $K = -\dfrac{19}{4}$, y obtenemos la solución $y = \dfrac{3}{4} - \dfrac{19}{4x^2}$.

Ejemplo. En ocasiones, las soluciones se encuentran de modo *implícito*. Por ejemplo, el problema de valor inicial dado por

$$yy' = x, \quad y(2) = -1$$

es equivalente a $(y^2)' = 2x$, luego $y^2 = x^2 + c$ para alguna constante c. Al imponer $y(2) = -1$ obtenemos $c = -3$ y la solución de modo implícito $y^2 = x^2 - 3$.

Para obtener la solución explícita se despeja y obteniendo $y = \pm\sqrt{x^2 - 3}$. Las condiciones iniciales fuerzan $y = -\sqrt{x^2 - 3}$.

§ 2. Soluciones exactas

Vamos a considerar algunos ejemplos de ecuaciones diferenciales ordinarias que sabemos integrar.

2.1. Variables separadas

Una ecuación diferencial con variables separadas es una que puede escribirse en la forma

$$f(y)y' = g(x). \tag{7.3}$$

Estas ecuaciones se resuelven fácilmente en forma implícita encontrando integrales de las funciones $f(x)$ y $g(x)$: $F(x) = \int f(x)dx$, $G(x) = \int g(x)dx$. De este modo la ecuación la podemos escribir, usando la regla de la cadena, como

$$(F(y))' = G'(x)$$

que nos da la solución, de modo implícito,

$$F(y) = G(x) + c.$$

(c denotará siempre una constante.)

Receta. Escribimos la ecuación como

$$f(y)dy = g(x)dx$$

e integramos en ambos lados:

$$\int f(y)dy = \int g(x)dx\,.$$

En los ejercicios verás ejemplos muy interesantes relativos al crecimiento de poblaciones y a la logística.

Ejemplo.

$$y' = y^2x, \quad y(1) = -1.$$

Escribimos como

$$\frac{1}{y^2}dy = xdx$$

e integramos para obtener

$$-\frac{1}{y} = \frac{x^2}{2} + c$$

o, cambiando c por $\frac{c}{2}$,

$$y = \frac{-2}{x^2 + c}.$$

Notemos que haciendo el límite con $c \to \infty$ obtenemos la *solución singular* $y = 0$.

La condición inicial fuerza $c = 1$, luego la solución es $y = \dfrac{-2}{x^2 + 1}$.

2.2. Ecuaciones diferenciales lineales

Consideramos ahora las ecuaciones de la forma

$$y' + f(x)y = g(x). \tag{7.4}$$

Si tuviéramos una función $\mu(x)$, llamada *factor integrante*, tal que

$$\mu(x)\left(y' + f(x)y\right) = (\mu(x)y)'$$

esto es, si

$$\mu'(x) = \mu(x)f(x) \tag{7.5}$$

entonces podemos integrar fácilmente:

$$\mu(x)y = \int \mu(x)g(x)dx + c.$$

Pero (7.5) tiene como solución (recuerda el Ejemplo 7.2)

$$\mu(x) = e^{\int f(x)dx}.$$

Receta. La solución de (7.4) es

$$y = \frac{\displaystyle\int \mu(x)g(x)dx + c}{\mu(x)}$$

con

$$\mu(x) = e^{\int f(x)dx}.$$

Hay otro modo de proceder, llamado *método de variación de constantes*. Sea $F(x) = \displaystyle\int f(x)dx$ una primitiva de $f(x)$. Las funciones $y = Ke^{-F(x)}$ son las soluciones de la *ecuación homogénea*

$$y' + f(x)y = 0.$$

Cambiamos K por una función $h(x)$: $y = h(x)e^{-F(x)}$. Entonces

$$y' = h'(x)e^{-F(x)} - h(x)F'(x)e^{-F(x)} = h'(x)e^{-F(x)} - f(x)y,$$

luego necesitamos $z = h(x)$ tal que $z'e^{-F(x)} = g(x)$, esto es, $z = \displaystyle\int e^{F(x)}g(x)dx + c$, obteniendo así el resultado anterior.

2.3. Cambio de variables. Ecuaciones diferenciales de Bernouilli

Las ecuaciones diferenciales de la forma

$$y' + f(x)y = g(x)y^n \tag{7.6}$$

donde n es un número real, se dicen *ecuaciones de Bernouilli.* Si $n = 0$ o $n = 1$, estas ecuaciones son lineales. Si $n \geq 1$, $y = 0$ es una solución singular. Deseamos obtener el resto de soluciones.

Para ello dividimos por y^n:

$$y^{-n}y' + f(x)y^{1-n} = g(x),$$

y hacemos el *cambio de variable* $z = y^{1-n}$. Así $z' = (1-n)y^{-n}y'$ y (7.6) se transforma en

$$\frac{1}{1-n}z' + f(x)z = g(x)$$

que es lineal.

Ejemplo.

$$y' + \frac{4}{x}y = x^3y^2, \quad x > 0\,.$$

Tras dividir por y^2 obtenemos

$$y^{-2}y' + \frac{4}{x}y^{-1} = x^3.$$

Hacemos $z = y^{-1}$, luego $z' = -y^{-2}y'$ y la ecuación se convierte en

$$-z' + \frac{4}{x}z = x^3$$

que es lineal. Su factor integrante es

$$\mu(x) = e^{-\int \frac{4}{x}dx} = e^{-4\ln x} = x^{-4}$$

y tenemos la solución

$$z = \frac{-\int x^{-4}x^3dx + c}{x^{-4}} = \frac{-\ln x + c}{x^{-4}} = x^4(c - \ln x),$$

Por tanto, la solución general es

$$y = \frac{1}{x^4(c - \ln x)}.$$

§ 3. Soluciones aproximadas. Soluciones en series

Hasta ahora hemos visto varios tipos de ecuaciones para los que tenemos métodos exactos de resolución. Sin embargo, estos tipos son la excepción. La mayoría de EDOs no pueden resolverse de modo exacto.

3.1. Método de Euler

En estos casos hay que acudir a métodos numéricos que aproximan la solución con suficiente precisión.

Si tenemos el problema de valor inicial

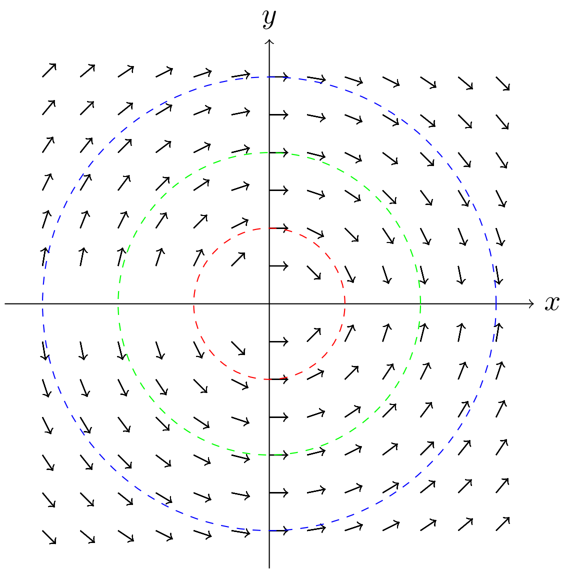
$$y' = f(x, y), \quad y(x_0) = y_0 \,. \tag{7.7}$$

En cada punto $(a, b) \in \mathbb{R}^2$, la tangente a una solución $y = y(x)$ de $y' = f(x, y)$ que pase por el punto (a, b): $b = y(a)$, es $f(a, b)$. Esto es, $\big(1, f(a, b)\big)$ es el vector tangente a la solución que pasa por (a, b). Estos vectores tangentes forman el *campo de direcciones* (o *campo de velocidades*) de las *curvas integrales* de la EDO: $t \mapsto \big(t, y(t)\big)$.

Ejemplo. Consideramos la EDO:

$$y' = -\frac{x}{y} \,.$$

Su campo de direcciones presenta la siguiente forma



(La longitud de los vectores del campo se ha normalizado.)

Dado el problema de valor inicial (7.7), para un $x_1 = x_0 + h$ tenemos, por el desarrollo de Taylor de la solución $y(x)$:

$$y(x_1) = y(x_0 + h) = y(x_0) + y'(x_0) \cdot h + \text{error}.$$

Ahora bien, sabemos que $y(x_0) = y_0$ y que $y'(x_0) = f(x_0, y_0)$. Entonces, para valores pequeños de h:

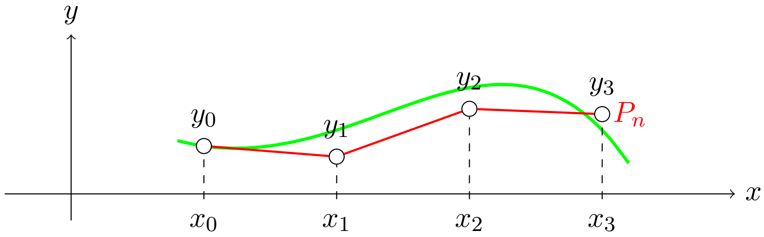
$$y(x_1) \simeq y_0 + f(x_0, y_0) \cdot h =: y_1$$

es una buena aproximación de la solución en x_1. Esto es, tomamos como aproximación al valor exacto de $y(x_1)$ el valor en x_1 de la recta tangente en x_0 a la solución $y(x)$.

Ahora consideramos $x_2 = x_1 + h$, y aproximamos el valor en x_2:

$$y(x_2) \simeq y_1 + f(x_1, y_1) \cdot h =: y_2\,.$$

Repitiendo el proceso, llegamos a una poligonal P_n (*poligonal de Euler*) que aproxima la curva solución:



La solución se obtiene como un límite:

$$y(x) \equiv \lim_{h\to 0} P_n(x).$$

3.2. Soluciones en series

En ocasiones, la ecuación diferencial nos permite calcular las derivadas de las funciones solución en un punto determinado, y esto nos permite recuperar la solución a partir de su polinomio o serie de Taylor.

Ejemplo. Consideramos la EDO

$$y'' + y = 0.$$

Toda solución verifica $y^{(n)} = -y^{(n-2)}$, luego si $y(0) = a$ e $y'(0) = b$, entonces $y^{(2n)}(0) = (-1)^n a$ e $y^{(2n+1)}(0) = (-1)^n b$. La función que tiene estas derivadas en 0 es $y = a\cos x + b \operatorname{sen} x$.

Alternativamente, buscamos soluciones de la forma

$$y(x) = \sum_{n=0}^{\infty} \frac{a_n}{n!} x^n,$$

luego

$$y'(x) = \sum_{n=1}^{\infty} \frac{a_n}{(n-1)!} x^{n-1}, \quad y''(x) = \sum_{n=2}^{\infty} \frac{a_n}{(n-2)!} x^{n-2}.$$

Sustituyendo en nuestra EDO obtenemos

$$\sum_{n=2}^{\infty} \frac{a_n}{(n-2)!} x^{n-2} + \sum_{n=0}^{\infty} \frac{a_n}{n!} x^n = 0,$$

de donde deducimos $a_{n+2} = -a_n \ \forall n$, y así

$$y(x) = a_0 \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{n!} + a_1 \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(n+1)!} = a_0 \cos x + a_1 \operatorname{sen} x.$$

§ 4. Sistemas de ecuaciones diferenciales lineales

Dado un número real a, la ecuación diferencial

$$x'(t) = ax(t)$$

tiene como soluciones $x(t) = ce^{at}$, $c \in \mathbb{R}$.

En general, vamos a considerar sistemas de ecuaciones diferenciales lineales:

$$\begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{nn} \\ \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

donde $a_{ij} \in \mathbb{R}\ \forall i, j = 1, \ldots, n$, que escribiremos, por simplicidad, como

$$\mathbf{x}' = A\mathbf{x}.$$

4.1. Matriz diagonalizable

Si la matriz de coeficientes A es diagonalizable, existe una matriz regular

$$P = \Big(\mathbf{u}_1 \mid \cdots \mid \mathbf{u}_n\Big)$$

cuyas columnas son vectores propios de A con

$$AP = P\Lambda$$

siendo

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

la matriz diagonal con los valores propios de A (repetidos según su multiplicidad).

Haciendo el cambio de variables $\mathbf{y} = P^{-1}\mathbf{x}$ tenemos

$$\mathbf{y}' = \left(P^{-1}\mathbf{x}\right)' = P^{-1}\mathbf{x}' = P^{-1}A\mathbf{x} = \Lambda P^{-1}\mathbf{x} = \Lambda\mathbf{y}.$$

Esto es, nuestro sistema se convierte en el sistema:

$$\begin{aligned} y_1'(t) &= \lambda_1 y_1(t), \\ y_2'(t) &= \lambda_2 y_2(t), \\ \vdots\ &\ \ \vdots \\ y_n'(t) &= \lambda_n y_n(t), \end{aligned}$$

cuya solución viene dada por

$$y_i(t) = c_i e^{\lambda_i t}, \quad i = 1, \ldots, n$$

que depende de los parámetros $c_1, \ldots, c_n \in \mathbb{R}$.

La solución al sistema inicial viene dada por $\mathbf{x} = P\mathbf{y}$, esto es:

$$\mathbf{x} = c_1 e^{\lambda_1 t}\mathbf{u_1} + c_2 e^{\lambda_2 t}\mathbf{u_2} + \cdots + c_n e^{\lambda_n t}\mathbf{u}_n.$$

Ejemplo. Consideramos el sistema

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}.$$

Por el Ejemplo 6.2 sabemos que sus valores propios son -1 y 4, y los correspondientes subespacios fundamentales están formados por los múltiplos escalares de $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ y $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$, respectivamente. Así pues, la solución general de nuestro sistema es

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = c_1 e^{-t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2 e^{4t} \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

esto es,

$$\begin{aligned} x_1(t) &= c_1 e^{-t} + 2c_2 e^{4t}, \\ x_2(t) &= -c_1 e^{-t} + 3c_2 e^{4t}. \end{aligned}$$

4.2. Valores propios complejos

Si se nos presenta un sistema $\mathbf{x}' = A\mathbf{x}$, con A matriz real 2×2 con valores propios complejos, no reales, estos valores propios vienen dados por un número complejo $\lambda = a + \mathrm{i}b$, $a, b \in \mathbb{R}$, y su conjugado $\bar{\lambda} = a - \mathrm{i}b$.

Sea $0 \neq \mathbf{u} \in \mathbb{C}^2$ un vector propio (con coeficientes complejos) de valor propio λ: $\mathbf{u} = \mathbf{p} + \mathrm{i}\mathbf{q}$, con $\mathbf{p}, \mathbf{q} \in \mathbb{R}^2$. Entonces $\mathbf{v} = \overline{\mathbf{u}} = \mathbf{p} - \mathrm{i}\mathbf{q}$ es un vector propio de valor propio $\bar{\lambda}$. La solución general *compleja* a nuestro sistema viene dada por las combinaciones lineales de $e^{\lambda t}\mathbf{u}$ y $e^{\bar{\lambda} t}\mathbf{v}$, esto es, por las combinaciones lineales de

$$e^{at}(\cos(bt) + \mathrm{i}\,\mathrm{sen}(bt))\mathbf{u} \quad \text{y} \quad e^{at}(\cos(bt) - \mathrm{i}\,\mathrm{sen}(bt))\mathbf{v},$$

o, equivalentemente, por las combinaciones lineales de las partes reales e imaginarias:

$$e^{at}(\cos(bt)\mathbf{p} - \mathrm{sen}(bt)\mathbf{q}) \quad \text{y} \quad e^{at}(\cos(bt)\mathbf{q} + \mathrm{sen}(bt)\mathbf{p}).$$

Las combinaciones lineales con coeficientes reales de estas últimas funciones nos dan la solución general (real).

Receta. La solución general del sistema

$$\mathbf{x}' = A\mathbf{x}$$

con A matriz real 2×2 con valor propio no real $\lambda = a + \mathrm{i}b$ y vector propio asociado $\mathbf{u} = \mathbf{p} + \mathrm{i}\mathbf{q}$ es

$$\begin{aligned}\mathbf{x} &= c_1 \operatorname{Re}\left(e^{\lambda t}\mathbf{u}\right) + c_2 \operatorname{Im}\left(e^{\lambda t}\mathbf{u}\right)\\ &= c_1 e^{at}(\cos(bt)\mathbf{p} - \operatorname{sen}(bt)\mathbf{q}) + c_2 e^{at}(\cos(bt)\mathbf{q} + \operatorname{sen}(bt)\mathbf{p})\end{aligned}$$

con $c_1, c_2 \in \mathbb{R}$.

Ejemplo. Consideramos el sistema

$$\begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}.$$

El polinomio característico es $\begin{vmatrix} x-1 & -1 \\ 1 & x-1 \end{vmatrix} = (x-1)^2 + 1$, cuyas raíces son $1 \pm \mathrm{i}$. El subespacio fundamental (complejo) de valor propio $1+\mathrm{i}$ es $S(A, 1+\mathrm{i}) = \mathbb{C}\begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix}$. Las partes reales e imaginarias de $e^{(1+\mathrm{i})t}\begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} = e^t(\cos t + \mathrm{i}\operatorname{sen} t)\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathrm{i}\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$ son

$$e^t\left(\cos t\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \operatorname{sen} t\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = e^t\begin{pmatrix} \cos t \\ -\operatorname{sen} t \end{pmatrix} \qquad \text{y} \qquad e^t\left(\cos t\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \operatorname{sen} t\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = e^t\begin{pmatrix} \operatorname{sen} t \\ \cos t \end{pmatrix},$$

luego la solución general es

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = c_1 e^t\begin{pmatrix} \cos t \\ -\operatorname{sen} t \end{pmatrix} + c_2 e^t\begin{pmatrix} \operatorname{sen} t \\ \cos t \end{pmatrix},$$

esto es,

$$\begin{aligned} x_1(t) &= e^t(c_1 \cos t + c_2 \operatorname{sen} t),\\ x_2(t) &= e^t(-c_1 \operatorname{sen} t + c_2 \cos t).\end{aligned}$$

4.3. Valor propio repetido

Consideremos ahora los sistemas

$$\mathbf{x}' = A\mathbf{x},$$

donde A es una matriz real 2×2 con un único valor propio real, de multiplicidad 2:

$$|xI_2 - A| = (x - \lambda)^2$$

de manera que A no sea diagonalizable.

Veamos primero un caso particular:

$$\begin{cases} x_1' = \lambda x_1, \\ x_2' = x_1 + \lambda x_2. \end{cases}$$

La matriz de coeficientes $A = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}$ tiene un único valor propio: λ, con multiplicidad 2, pero la dimensión de $S(A, \lambda)$ es 1, luego no es diagonalizable. Sin embargo, sabemos resolver el sistema, pues $x_1 = c_2 e^{\lambda t}$ para alguna constante c_2, y tenemos $x_2' = c_2 e^{\lambda t} + \lambda x_2$, que es una ecuación diferencial lineal con solución

$$x_2 = e^{\lambda t}(c_2 t + c_1)$$

para alguna constante c_1. Así pues, la solución al sistema es

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = e^{\lambda t} \begin{pmatrix} c_2 \\ c_2 t + c_1 \end{pmatrix} = c_1 e^{\lambda t} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + c_2 e^{\lambda t} \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right).$$

Notemos que el vector $\mathbf{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ no es vector propio de valor propio λ, pero sí lo es $\mathbf{v} = (A - \lambda I_2)\, \mathbf{u} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. La solución general nos queda de la forma

$$\mathbf{x}(t) = c_1 e^{\lambda t} \mathbf{v} + c_2 e^{\lambda t} \left(\mathbf{u} + t \mathbf{v} \right).$$

En general, si $0 \neq \mathbf{p}$ es un vector propio, entonces $A\mathbf{p} = \lambda \mathbf{p}$ y $\mathbf{x}(t) = e^{\lambda t} \mathbf{p}$ es una solución del sistema, puesto que

$$\mathbf{x}' = \lambda e^{\lambda t} \mathbf{p} = A(e^{\lambda t} \mathbf{p}) = A\mathbf{x}.$$

Buscamos soluciones de la forma

$$\mathbf{x}(t) = e^{\lambda t} \mathbf{u} + t e^{\lambda t} \mathbf{v} = e^{\lambda t}(\mathbf{u} + t\mathbf{v})$$

con $\mathbf{v} \neq 0$. De

$$\begin{aligned} \mathbf{x}'(t) &= \lambda e^{\lambda t}(\mathbf{u} + t\mathbf{v}) + e^{\lambda t} \mathbf{v} \\ &= e^{\lambda t}\Big((\lambda \mathbf{u} + \mathbf{v}) + t \lambda \mathbf{v} \Big) \end{aligned}$$

$$\mathbf{x}'(t) = A\mathbf{x}(t) = e^{\lambda t}(A\mathbf{u} + tA\mathbf{v})$$

obtenemos

$$A\mathbf{u} + tA\mathbf{v} = (\lambda \mathbf{u} + \mathbf{v}) + t\lambda \mathbf{v} \quad \forall t \in \mathbb{R},$$

de donde deducimos

$$A\mathbf{v} = \lambda \mathbf{v}, \quad A\mathbf{u} = \lambda \mathbf{u} + \mathbf{v},$$

o equivalentemente

$$(A - \lambda I_2)\mathbf{v} = 0, \quad (A - \lambda I_2)\mathbf{u} = \mathbf{v}.$$

Receta. En las condiciones anteriores, sea $\mathbf{u} \in \mathbb{R}^2 \setminus S(A, \lambda)$, entonces $\mathbf{v} = (A - \lambda I_2)\mathbf{u}$ es un vector propio (de valor propio λ) y la solución general del sistema $\mathbf{x}' = A\mathbf{x}$ es

$$c_1 e^{\lambda t}\mathbf{v} + c_2 e^{\lambda t}(\mathbf{u} + t\mathbf{v}).$$

Ejemplo. Dada la ecuación diferencial:

$$y'' - 2y' + y = 0, \tag{7.8}$$

la podemos convertir en un sistema como sigue. Sea $x_1 = y$ y $x_2 = y'$, entonces nos queda

$$\begin{cases} x_1' = x_2, \\ x_2' = -x_1 + 2x_2, \end{cases}$$

esto es,

$$\mathbf{x}' = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \mathbf{x}. \tag{7.9}$$

Calculamos los valores propios de la matriz A en el sistema:

$$\begin{vmatrix} x & -1 \\ 1 & x-2 \end{vmatrix} = x(x-2) + 1 = (x-1)^2.$$

Obtenemos fácilmente $S(A, 1) = \mathbb{R}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. El vector $\mathbf{u} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ no es pues vector propio, pero

$$(A - I_2)\mathbf{u} = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} =: \mathbf{v}$$

sí es vector propio.

La solución general a nuestro sistema (7.9) es

$$\mathbf{x}(t) = c_1 e^t \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^t \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = e^t \begin{pmatrix} c_1 + c_2 t \\ c_1 + c_2 + c_2 t \end{pmatrix}$$

luego la solución general a la ecuación (7.8) es

$$y = e^t(c_1 + c_2 t),$$

con $c_1, c_2 \in \mathbb{R}$.

4.4. Resumen

Para sistemas de ecuaciones diferenciales lineales de grado dos:

$$\mathbf{x}' = A\mathbf{x} \tag{7.10}$$

podemos resumir lo anterior como sigue:

- Si A es diagonalizable y $\{\mathbf{u_1}, \mathbf{u_2}\}$ es una base de vectores propios, con valores propios asociados λ_1 y λ_2 (pueden ser iguales), entonces la solución general de (7.10) es

  $$\mathbf{x} = c_1 e^{\lambda_1 t}\mathbf{u_1} + c_2 e^{\lambda_2 t}\mathbf{u_2}$$

  con $c_1, c_2 \in \mathbb{R}$ constantes arbitrarias.

- Si A posee un valor propio $\lambda \in \mathbb{C} \setminus \mathbb{R}$ y $\mathbf{u} \in \mathbb{C}^2$ es un vector propio (no nulo) de valor propio λ, entonces la solución general de (7.10) es

  $$\mathbf{x} = c_1 \operatorname{Re}(e^{\lambda t}\mathbf{u}) + c_2 \operatorname{Im}(e^{\lambda t}\mathbf{u})$$

  con $c_1, c_2 \in \mathbb{R}$ constantes arbitrarias, donde Re e Im denotan la parte real e imaginaria: $\operatorname{Re}(a + \mathrm{i}b) = a$, $\operatorname{Im}(a + \mathrm{i}b) = b$.

  Debemos recordar la exponencial de números complejos: $e^{a+b\mathrm{i}} = e^a(\cos b + \mathrm{i}\operatorname{sen} b)$.

- Si A no es diagonalizable y $\lambda \in \mathbb{R}$ es su único valor propio (con multiplicidad 2), entonces elegimos un vector $\mathbf{u} \in \mathbb{R}^2$ que no sea vector propio de A, y consideramos el vector $\mathbf{v} = (A - \lambda I_2)\mathbf{u}$. La solución general de (7.10) es

  $$c_1 e^{\lambda t}\mathbf{v} + c_2 e^{\lambda t}(\mathbf{u} + t\mathbf{v}).$$

  con $c_1, c_2 \in \mathbb{R}$ constantes arbitrarias.

Capítulo 8

Funciones de varias variables. Diferenciación

En este capítulo estudiaremos funciones

$$f : D \longrightarrow \mathbb{R}$$

donde el dominio D es un subconjunto de $\mathbb{R}^n$ para algún n, concentrándonos especialmente en el caso $n = 2$.

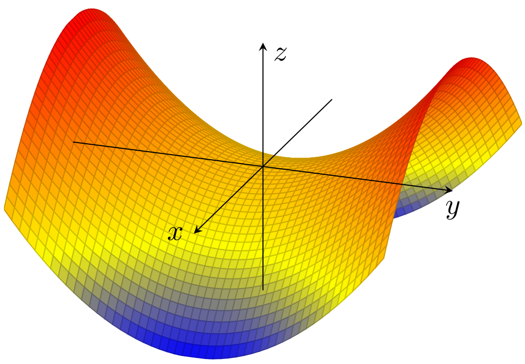
Ejemplo. $f(x, y) = x \ln(x^2 - y)$. El dominio de f es el conjunto

$$D = \{(x, y) \in \mathbb{R}^2 \mid y < x^2\}$$

Dada una función de dos variables $f : D \longrightarrow \mathbb{R}$, $D \subseteq \mathbb{R}^2$:

- La **gráfica** de f es la superficie de $\mathbb{R}^3$:

$$\{(x, y, z) \in \mathbb{R}^3 \mid z = f(x, y)\}.$$

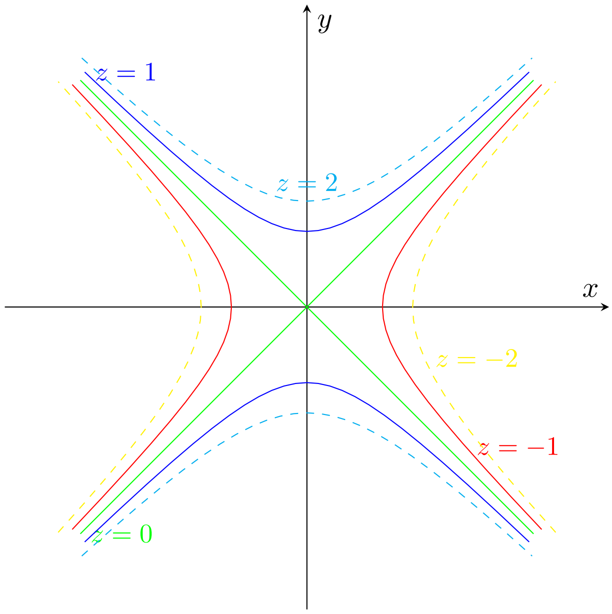


(8.1)

Gráfica de $f(x, y) = y^2 - x^2$

- Las **curvas de nivel** son las curvas en $\mathbb{R}^2$ de ecuación $f(x,y)=k$, con k en el rango de f.

Curvas de nivel de $z=y^2-x^2$



§ 1. Límites. Continuidad

- Decimos que el **límite** de f cuando (x,y) tiende hacia (a,b) es L y escribimos

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

si la función se acerca a L cuando (x,y) se acerca a (a,b), con $(x,y)\neq(a,b)$. Formalmente $\lim_{(x,y)\to(a,b)} f(x,y)=L$ si para todo $\epsilon>0$, existe $\delta>0$ tal que

$$(x,y)\in D \text{ y } 0<\sqrt{(x-a)^2+(y-b)^2}<\delta \implies |f(x,y)-L|<\epsilon.$$

Como para funciones de una variable, se tiene $\lim_{(x,y)\to(a,b)} f(x,y)=L$ si para toda sucesión (x_n,y_n) en $D\setminus\{(a,b)\}$ con $\lim x_n=a$ y $\lim y_n=b$, se tiene $\lim f(x_n,y_n)=L$.

Si $\lim_{(x,y)\to(a,b)} f(x,y)=L$, entonces para cada curva $\sigma: I\longrightarrow D$ con $\sigma(t)=\big(x(t),y(t)\big)$ tal que $\lim_{t\to t_0} x(t)=a$ y $\lim_{t\to t_0} y(t)=b$, se tiene

$$\lim_{t\to t_0} f\big(x(t),y(t)\big)=L.$$

(Notemos que este último límite es un límite de una sola variable.) Esto quiere decir que al acercarnos hacia (a,b) a través de cualquier curva, el límite siempre ha de ser L.

Ejemplo. No existe $\lim_{(x,y)\to(0,0)} \dfrac{xy}{x^2+y^2}$.

En efecto, con $\sigma(t) = (t, 0)$ se tiene $\lim_{t\to 0} t = 0 = \lim_{t\to 0} 0$ y

$$\lim_{t\to 0} \frac{t\cdot 0}{t^2+0} = \lim_{t\to 0} \frac{0}{t^2} = \lim_{t\to 0} 0 = 0,$$

pero con $\sigma(t) = (t, t)$ se tiene

$$\lim_{t\to 0} \frac{t^2}{t^2+t^2} = \lim_{t\to 0} \frac{1}{2} = \frac{1}{2},$$

luego acercándonos al punto $(0, 0)$ por dos curvas distintas obtenemos distintos límites.

Todo lo anterior se extiende sin dificultad a funciones de tres o más variables.

Debido a que los límites de funciones quedan determinados por límites de sucesiones, se tienen las propiedades usuales de los límites:

- Si existen los límites $\lim_{(x,y)\to(a,b)} f(x, y)$ y $\lim_{(x,y)\to(a,b)} g(x, y)$ y $r \in \mathbb{R}$, entonces:
  - $\lim_{(x,y)\to(a,b)} rf(x, y) = r \lim_{(x,y)\to(a,b)} f(x, y)$,
  - $\lim_{(x,y)\to(a,b)} \big(f(x, y) + g(x, y)\big) = \lim_{(x,y)\to(a,b)} f(x, y) + \lim_{(x,y)\to(a,b)} g(x, y)$ (salvo la indeterminación $\infty - \infty$),
  - $\lim_{(x,y)\to(a,b)} \Big(f(x, y)g(x, y)\Big) = \lim_{(x,y)\to(a,b)} f(x, y) \lim_{(x,y)\to(a,b)} g(x, y)$ (salvo $0 \cdot \infty$),
  - $\lim_{(x,y)\to(a,b)} \dfrac{f(x, y)}{g(x, y)} = \dfrac{\lim_{(x,y)\to(a,b)} f(x, y)}{\lim_{(x,y)\to(a,b)} g(x, y)}$ (salvo $\dfrac{0}{0}$ o $\dfrac{\infty}{\infty}$).
- **Teorema del Sandwich:** Si existe $\delta > 0$ tal que $f(x, y) \le h(x, y) \le g(x, y)$ $\forall (x, y) \in D$ con $0 < \sqrt{(x-a)^2 + (y-b)^2} < \delta$ y $\lim_{(x,y)\to(a,b)} f(x, y) = \lim_{(x,y)\to(a,b)} g(x, y) = L$, entonces $\lim_{(x,y)\to(a,b)} h(x, y) = L$.

En ocasiones es útil usar 'coordenadas polares' para calcular límites, especialmente en el origen, pues

$$\lim_{(x,y)\to(0,0)} f(x, y) = \lim_{r\to 0^+} f(r\cos\theta, r\operatorname{sen}\theta).$$

Ejemplo. $\lim_{(x,y)\to(0,0)} (x^2 + y^2) \ln(x^2 + y^2) =?$

$$\begin{aligned}\lim_{(x,y)\to(0,0)} (x^2 + y^2) \ln(x^2 + y^2) &= \lim_{r\to 0^+} r^2 \ln(r^2) \\ &= \lim_{r\to 0^+} 2r^2 \ln r = 0.\end{aligned}$$

Como para funciones de una variable, una función $f(x, y)$ se dice **continua** en $(a, b) \in D$, si

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b),$$

y análogamente para funciones de más variables.

§ 2. Derivadas parciales

Dada una función de dos variables $f(x, y)$, la **derivada parcial** de f con respecto a x en $(a, b) \in D$ es el límite (si existe):

$$\frac{\partial f}{\partial x}(a,b) := \lim_{h \to 0} \frac{f(a+h,b) - f(a,b)}{h}\,,$$

que es un límite en una sola variable. Análogamente la **derivada parcial** de f con respecto a y en $(a, b) \in D$ es el límite (si existe):

$$\frac{\partial f}{\partial y}(a,b) := \lim_{h \to 0} \frac{f(a,b+h) - f(a,b)}{h}\,.$$

Las definiciones son análogas para funciones de tres o más variables.

Notación: Las derivadas parciales de una función $z = f(x, y)$ se denotan de varias maneras:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x} f = f_x = z_x = \frac{\partial z}{\partial x}\,.$$

Geométricamente, $f_x(a, b)$ mide la razón de cambio de f en la dirección del eje OX y, por tanto, es la pendiente de la recta tangente a la curva $t \to f(a + t, b)$ en el punto $t = 0$. Esta curva está contenida en la gráfica de f.

Del mismo modo, $f_y(a, b)$ mide la razón de cambio de f en la dirección del eje OY, esto es, la pendiente de la recta tangente a la curva $t \to f(a, b + t)$ en el punto $t = 0$.

Ejemplo. Si $f(x, y) = x^2 y - \ln x$, se tiene:

$$f_x = 2xy - \frac{1}{x}, \quad f_y = x^2, \quad f_x(1,0) = -1, \quad f_y(1,0) = 1.$$

Ejemplo. La función $f(x, y) = \dfrac{xy}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$, $f(0, 0) = 0$, está definida en todo $\mathbb{R}^2$. Sabemos que no existe $\lim_{(x,y)\to(0,0)} f(x, y)$, luego no es continua en $(0, 0)$. Sin embargo,

$$f_x(0,0) = \lim_{h \to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h \to 0} 0 = 0\,,$$

ya que $f(h, 0) = \dfrac{h \cdot 0}{h^2 + 0^2} = 0$ para todo $h \neq 0$. Del mismo modo, $f_y(0, 0) = 0$.

Vemos, por tanto, que la existencia de derivadas parciales no implica continuidad.

De manera más general, la **derivada direccional** de f en el punto $P = (a, b)$ y en la dirección $\mathbf{u}$ (vector de longitud 1) se define como,

$$D_{\mathbf{u}} f(P) := \lim_{t \to 0} \frac{f(P + t\mathbf{u}) - f(P)}{t} = \left.\frac{d}{dt}\right|_{t=0} f(P + t\mathbf{u}).$$

§ 3. Diferenciabilidad

Dada una función de dos variables (por simplicidad, los resultados son válidos para un número arbitrario de variables con las modificaciones obvias) $f : D \to \mathbb{R}$, f se dice **diferenciable** en el punto $P = (a, b) \in D$, si existen las derivadas parciales $f_x(a, b)$ y $f_y(a, b)$ y además

$$\lim_{(x,y)\to(a,b)} \frac{f(x, y) - \Big(f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)\Big)}{\sqrt{(x - a)^2 + (y - b)^2}} = 0.$$

Esto lo podemos escribir como

$$\lim_{\mathbf{u}\to\mathbf{0}} \frac{f(P + \mathbf{u}) - \Big(f(P) + f_x(P)u_1 + f_y(P)u_2\Big)}{\|\mathbf{u}\|} = 0\,,$$

donde $\mathbf{u} = (u_1, u_2)$.

Si f es diferenciable en $P = (a, b) \in D$, entonces es continua en P y podemos aproximar f en las cercanías de P por

$$f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)\,.$$

Además,

- Para todo vector $\mathbf{u} = (u_1, u_2)$ de longitud 1 existe la derivada direccional $D_{\mathbf{u}}f(P)$ y

  $$D_{\mathbf{u}}f(P) = f_x(P)u_1 + f_y(P)u_2 = \nabla f(P) \bullet \mathbf{u}\,,$$

  donde $\nabla f(P) = \big(f_x(P), f_y(P)\big)$ se llama (**vector**) **gradiente** de f en P.

  En efecto, al ser f diferenciable en P y ser $\|t\mathbf{u}\| = |t|$, se tiene

  $$\lim_{t\to 0} \frac{f(P + t\mathbf{u}) - \Big(f(P) + f_x(P)tu_1 + f_y(P)tu_2\Big)}{t} = 0$$

  luego

  $$\lim_{t\to 0} \frac{f(P + t\mathbf{u}) - f(P)}{t} = f_x(P)u_1 + f_y(P)u_2.$$

- El valor máximo de $D_{\mathbf{u}}f(P)$ al variar $\mathbf{u}$ es $\|\nabla f(P)\|$ y se obtiene si $\mathbf{u}$ tiene la misma dirección que el vector gradiente $\nabla f(P)$.

  Por tanto

  $$\boxed{\nabla f(P) \text{ indica la dirección de mayor crecimiento de } f \text{ en } P.}$$

  Esto es debido a que

  $$D_{\mathbf{u}}f(P) = \nabla f(P) \bullet \mathbf{u} = \|\nabla f(P)\| \cos\alpha,$$

  con α el ángulo que forman el vector unitario $\mathbf{u}$ y $\nabla f(P)$.

- El plano tangente a la gráfica de f en $P = (a, b)$ es

$$z - f(a, b) = f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

Por otra parte,

- Si existen las derivadas parciales f_x y f_y en un disco alrededor de $P = (a, b)$, y son continuas en P, entonces f es diferenciable en P.

§ 4. Regla de la cadena

Podemos componer funciones de varias variables de muchos modos, y para cada uno de ellos hay una versión de la regla de la cadena. Veremos aquí un par de ellos:

- Si $z = f(x, y)$ es diferenciable, y $x = g(t)$ e $y = h(t)$ son derivables, entonces z es una función derivable de t y
$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

- Si $z = f(x, y)$ es una función diferenciable, y a su vez $x = g(s, t)$ e $y = h(s, t)$ son diferenciables, entonces
$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s},$$
$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}.$$

§ 5. Derivadas de orden superior. Polinomio de Taylor

Como para funciones de una variable, si f es una función de dos variables, x e y, así lo son sus derivadas parciales, f_x y f_y. Por tanto podemos considerar las derivadas parciales $(f_x)_x$, $(f_x)_y$, $(f_y)_x$ y $(f_y)_y$. Utilizaremos las siguientes notaciones:

$$f_{xx} = (f_x)_x = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2},$$
$$f_{xy} = (f_x)_y = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x},$$
$$f_{yx} = (f_y)_x = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y},$$
$$f_{yy} = (f_y)_y = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}.$$

Teorema. ***Igualdad de las derivadas cruzadas:*** *Si f_{xy} y f_{yx} son continuas, entonces $f_{xy} = f_{yx}$.*

Si f es suficientemente derivable en un entorno de un punto $P = (a, b) \in D$, con derivadas continuas, y tomamos un punto (x, y) cercano a P, sea $u_1 = x - a$, $u_2 = y - b$.

Consideremos la función de una variable $F(t) := f(a + tu_1, b + tu_2)$. Su polinomio de Taylor de orden 2 en $t = 0$ es

$$F(0) + F'(0)t + \frac{F''(0)}{2}t^2\,, \tag{8.2}$$

que es una buena aproximación de $F(t)$.

Por la regla de la cadena tenemos:

$$\begin{aligned}
F'(t) &= f_x(a + tu_1, b + tu_2)u_1 + f_y(a + tu_1, b + tu_2)u_2, \\
F''(t) &= f_{xx}(a + tu_1, b + tu_2)u_1^2 + f_{xy}(a + tu_1, b + tu_2)u_1u_2 \\
&\qquad + f_{yx}(a + tu_1, b + tu_2)u_1u_2 + f_{yy}(a + tu_1, b + tu_2)u_2^2 \\
&= f_{xx}(a + tu_1, b + tu_2)u_1^2 + 2f_{xy}(a + tu_1, b + tu_2)u_1u_2 + f_{yy}(a + tu_1, b + tu_2)u_2^2.
\end{aligned}$$

y, en particular,

$$\begin{aligned}
F'(0) &= f_x(a, b)u_1 + f_y(a, b)u_2, \\
F''(0) &= f_{xx}(a, b)u_1^2 + 2f_{xy}(a, b)u_1u_2 + f_{yy}(a, b)u_2^2.
\end{aligned}$$

Así pues (8.2) con $t = 1$ nos queda, teniendo en cuenta que $u_1 = x - a$ y $u_2 = y - b$:

$$\boxed{\begin{aligned}
&f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) \\
&\qquad + \frac{1}{2}f_{xx}(a, b)(x - a)^2 + f_{xy}(a, b)(x - a)(y - b) + \frac{1}{2}f_{yy}(a, b)(y - b)^2\,,
\end{aligned}}$$

que es un polinomio de grado 2 en x, y llamado **polinomio de Taylor de grado** 2 **de** f **en el punto** (a, b).

Este polinomio es una buena aproximación al valor de $f(x, y)$ en las cercanías de (a, b).

Ejemplo. El polinomio de Taylor de grado 2 de la función $\cos(x^2y)$ en el punto $\left(\frac{\pi}{2}, 0\right)$ es

$$1 - \frac{\pi^4}{32}y^2,$$

pues $f\left(\frac{\pi}{2}, 0\right) = 1$, $f_x\left(\frac{\pi}{2}, 0\right) = f_y\left(\frac{\pi}{2}, 0\right) = f_{xx}\left(\frac{\pi}{2}, 0\right) = f_{xy}\left(\frac{\pi}{2}, 0\right) = 0$ y $f_{yy}\left(\frac{\pi}{2}, 0\right) = -\frac{\pi^4}{16}$.

El polinomio de Taylor de grado 3 es:

$$\begin{aligned}
&f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) \\
&+ \frac{1}{2}f_{xx}(a, b)(x - a)^2 + f_{xy}(a, b)(x - a)(y - b) + \frac{1}{2}f_{yy}(a, b)(y - b)^2 \\
&+ \frac{1}{6}f_{xxx}(a, b)(x - a)^3 + \frac{1}{2}f_{xxy}(a, b)(x - a)^2(y - b) + \frac{1}{2}f_{xyy}(a, b)(x - a)(y - b)^2 + \frac{1}{6}f_{yyy}(a, b)(y - b)^3.
\end{aligned}$$

§ 6. Extremos

Se dice que la función $f(x, y)$ tiene un **máximo relativo** en el punto (a, b) si $f(x, y) \leq f(a, b)$ para todos los puntos (x, y) de un disco centrado en (a, b).

Análogamente para **mínimo relativo**.

Proposición. *Si f tiene un extremo (máximo o mínimo) relativo en (a, b), y tiene sus derivadas parciales en (a, b), entonces $f_x(a, b) = 0 = f_y(a, b)$. Esto es, $\nabla f(a, b) = 0$.*

En efecto, si $f(x, y)$ tiene un extremo (máximo o mínimo) relativo en (a, b), la función de una variable $f(x, b)$ tiene un extremo relativo en a, luego si es derivable, la derivada en $x = a$ vale 0, esto es: $f_x(a, b) = 0$. Igualmente $f_y(a, b) = 0$.

Los puntos en los que $\nabla f(x, y) = 0$ se dicen *puntos críticos*. Los puntos críticos no son necesariamente extremos.

Ejemplo. $f(x, y) = y^2 - x^2$ (su gráfica está en (8.1)).

$\nabla f(x, y) = (-2x, 2y)$, luego el único punto crítico es $(0, 0)$. En la dirección del eje OX, la función toma valores $f(x, 0) = -x^2$, que tiene un máximo para $x = 0$. Sin embargo, en la dirección del eje OY, la función toma valores $f(0, y) = y^2$, que tiene un mínimo para $y = 0$. Por tanto $(0, 0)$ no es extremo relativo de f.

Los puntos como en el ejemplo anterior, que son mínimos relativos de f en alguna dirección, y máximos relativos en otras direcciones, se llaman **puntos silla**.

Dada una función con derivadas parciales de segundo orden en (a, b), la matriz

$$H = \begin{pmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{yx}(a, b) & f_{yy}(a, b) \end{pmatrix}$$

se llama **matriz Hessiana** de f en el punto (a, b). Su determinante $|H|$ se llama simplemente **Hessiano** de f. El siguiente criterio se basa en el desarrollo de Taylor de grado 2 alrededor de los puntos críticos.

Teorema. *Si las derivadas parciales segundas de f son continuas en un entorno de un punto crítico (a, b), entonces:*

- *si $|H| > 0$ y $\operatorname{traza}(H) > 0$, (a, b) es un mínimo relativo.*
- *si $|H| > 0$ y $\operatorname{traza}(H) < 0$, (a, b) es un máximo relativo.*
- *si $|H| < 0$, (a, b) es un punto silla.*

Si $|H| = 0$ el criterio anterior no nos da información, es lo análogo al caso de puntos críticos para funciones de una variable con derivada segunda nula.

Ejemplos.

- $f(x, y) = x^2 + y^2 + xy$.

  $\nabla f(x, y) = (2x + y, 2y + x)$, luego su único punto crítico es $(0, 0)$. La matriz Hessiana es constante:

  $$H = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

  con $|H| = 3 > 0$ y $\text{traza}(H) = 4 > 0$, luego $(0, 0)$ es un mínimo relativo.

- $f(x, y) = x^3 - 4xy + y$.

  $\nabla f(x, y) = (3x^2 - 4y, -4x + 1)$, luego su único punto crítico es $\left(\frac{1}{4}, \frac{3}{64}\right)$. La matriz Hessiana es

  $$H(x, y) = \begin{pmatrix} 6x & -4 \\ -4 & 0 \end{pmatrix}, \qquad H\left(\frac{1}{4}, \frac{3}{64}\right) = \begin{pmatrix} \frac{3}{2} & -4 \\ -4 & 0 \end{pmatrix}$$

  Así, el Hessiano en el punto crítico vale -16, luego el punto crítico es un punto silla.

En el caso de una función f de n variables, $n \geq 3$, debemos de calcular también la matriz Hessiana H. Si las derivadas parciales segundas son continuas, esta matriz es simétrica y tiene todos sus valores propios reales. El criterio anterior es ahora un poco más complicado:

Teorema. *Si las derivadas parciales segundas de f son continuas en un entorno de un punto crítico $P = (a_1, \ldots, a_n)$, entonces:*

- *si los valores propios de H son todos positivos, P es un mínimo relativo.*
- *si los valores propios de H son todos negativos, P es un máximo relativo.*
- *si hay valores propios positivos y negativos, P es un punto silla.*

Nota. El determinante de la matriz Hessiana es el producto de sus valores propios, contados según su multiplicidad, y la traza es la suma de los mismos. Esto hace que este criterio se reduzca al anterior en el caso de dos variables.

§ 7. Campos vectoriales

Un **campo vectorial** de dos variables es una aplicación $\mathbf{F} : D \longrightarrow \mathbb{R}^2$, $(D \subseteq \mathbb{R}^2)$. Así

$$\mathbf{F}(x, y) = \Big(P(x, y), Q(x, y)\Big) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$$

donde las *componentes* de $\mathbf{F}$ son funciones $D \longrightarrow \mathbb{R}$.

$\mathbf{F}$ asigna un vector del plano a cada punto del dominio D. Análogamente se definen los campos vectoriales en dimensiones superiores.

Ejemplo. Campo gradiente: Dada una función diferenciable $f : D \longrightarrow \mathbb{R}$, f define su campo gradiente: $\nabla f : D \longrightarrow \mathbb{R}^2$.

Definición. Un campo vectorial $\mathbf{F}$ se dice **conservativo** si es el campo gradiente de una función diferenciable f: $\mathbf{F} = \nabla f$.

En este caso, la función f se dice **potencial** de $\mathbf{F}$.

Proposición. *Sea $\mathbf{F}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ un campo vectorial tal que sus componentes P y Q tienen derivadas parciales continuas.*

- *Si $\mathbf{F}$ es conservativo, entonces*
$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}\,.$$
- *Si D es* simplemente conexo[1]*, el recíproco es cierto. Esto es, si $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}$, entonces $\mathbf{F}$ es conservativo.*

Para tres variables, el resultado correspondiente es un poco más complicado. Dado un campo vectorial en tres variables $\mathbf{F}(x, y, z) = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$ cuyas componentes tienen derivadas parciales continuas, se llama **rotacional** de $\mathbf{F}$ al campo vectorial

$$\operatorname{rot}\mathbf{F} := \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}\,.$$

Nos podemos acordar mejor de su definición mediante la siguiente regla nemotécnica:

$$\operatorname{rot}\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}$$

Proposición. *Sea $\mathbf{F}(x, y, z) = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$ un campo vectorial tal que sus componentes P, Q y R tienen derivadas parciales continuas.*

- *Si $\mathbf{F}$ es conservativo, entonces* $\operatorname{rot}\mathbf{F} = 0$.
- *Si D es simplemente conexo, el recíproco es cierto. Esto es, si* $\operatorname{rot}\mathbf{F} = 0$*, entonces $\mathbf{F}$ es conservativo.*

Recuerda el Teorema Fundamental del Cálculo Integral:

$$\int_a^b f'(x)dx = f(b) - f(a),$$

que nos dice que el valor de la integral de $f'(x)$ depende solo del valor de f en los extremos.

[1] D es simplemente conexo si no tiene agujeros que lo atraviesan.

Si ahora $f : D \longrightarrow \mathbb{R}$ es una función de dos variables (el argumento para más variables es el mismo) cuyo gradiente ∇f es continuo, y $\sigma : [a, b] \longrightarrow D$, $t \to (x(t), y(t))$ es una curva continua y con derivada continua en (a, b) (diremos *curva suave*), entonces

$$\begin{aligned}\int_\sigma \nabla f \bullet d\sigma &= \int_a^b \Big(f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) \Big) dt \\ &= \int_a^b \Big(f(x(t), y(t)) \Big)' dt \qquad \text{por la regla de la cadena} \\ &= f(\sigma(b)) - f(\sigma(a)) \qquad \text{por el Teorema Fundamental del Cálculo}\end{aligned}$$

Así pues la integral de línea $\int_\sigma \nabla f \bullet d\sigma$ *¡solo depende del valor de f en los extremos de la curva σ!*

Decimos que la integral de línea $\int_\sigma \mathbf{F} \bullet d\sigma$ es *independiente del camino* si $\int_\sigma \mathbf{F} \bullet d\sigma = \int_\gamma \mathbf{F} \bullet d\gamma$ para toda curva γ con los mismos puntos inicial y final que σ.

Teorema. *Sea $\mathbf{F} : D \longrightarrow \mathbb{R}^2$ un campo continuo. Entonces*

$$\boxed{\begin{array}{c}\mathbf{F} \textit{ es conservativo si, y solo si, para toda curva suave en } D, \\ \int_\sigma \mathbf{F} \bullet d\sigma \textit{ es independiente del camino.}\end{array}}$$

Ejemplo. El dominio de definición es importante. Consideremos el campo vectorial

$$\begin{aligned}\mathbf{F} : \mathbb{R}^2 \setminus \{(0, 0)\} &\longrightarrow \mathbb{R}^2 \\ (x, y) &\mapsto \frac{-y}{x^2 + y^2}\mathbf{i} + \frac{x}{x^2 + y^2}\mathbf{j}\,,\end{aligned}$$

que está definido en un dominio con un agujero (el origen). Se tiene

$$\begin{aligned}\frac{\partial}{\partial y}\left(\frac{-y}{x^2 + y^2}\right) &= \frac{-(x^2 + y^2) + 2y^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2} \\ \frac{\partial}{\partial x}\left(\frac{x}{x^2 + y^2}\right) &= \frac{x^2 + y^2 - 2x^2}{(x^2 + y^2)^2} = \frac{y^2 - x^2}{(x^2 + y^2)^2}\,,\end{aligned}$$

luego $\dfrac{\partial P}{\partial y} = \dfrac{\partial Q}{\partial x}$.

Si $\mathbf{F}$ fuera conservativo, la integral de línea sobre la circunferencia unidad $\sigma : [0, 2\pi] \to \mathbb{R}^2$, $\sigma(t) = (\cos t, \operatorname{sen} t)$ sería trivial pues $\sigma(0) = \sigma(2\pi)$. Pero

$$\int_\sigma \mathbf{F} \bullet d\sigma = \int_0^{2\pi} \mathbf{F}(\sigma(t)) \bullet \sigma'(t)\, dt = \int_0^{2\pi} \left(\frac{\operatorname{sen}^2 t}{\cos^2 t + \operatorname{sen}^2 t} + \frac{\cos^2 t}{\cos^2 t + \operatorname{sen}^2 t} \right) dt = \int_0^{2\pi} dt = 2\pi \neq 0,$$

luego $\mathbf{F}$ no es conservativo.

Capítulo 9

Integrales múltiples

Este capítulo está dedicado al estudio de integrales de funciones de varias variables.

§ 1. Integrales dobles

1.1. Sumas de Riemann

Al igual que para integrales de funciones de una variable, dada una función continua de dos variables

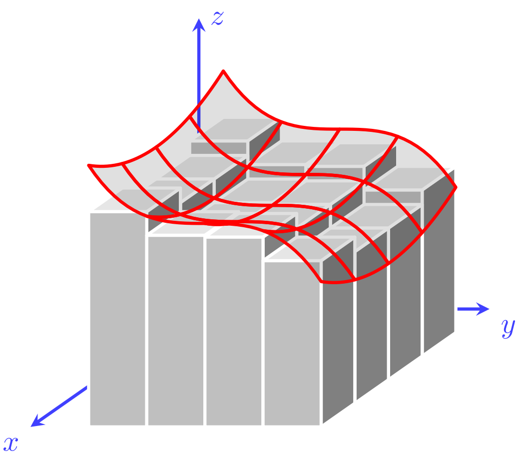
$$f : D \longrightarrow \mathbb{R},$$

definida en un rectángulo $D = [a, b] \times [c, d]$, con $f(x, y) \geq 0 \ \forall (x, y) \in D$, podemos aproximar el volumen bajo la gráfica de f tomando particiones

$$P_x = \{a = x_0 < x_1 < \cdots < x_n = b\}, \qquad P_y = \{c = y_0 < y_1 < \cdots < y_m = d\}$$

y puntos $(x^*_{ij}, y^*_{ij}) \in [x_{i-1}, x_i] \times [y_{j-1}, y_j]$, y considerar la *suma de Riemann*

$$\sum_{i=1}^{n} \sum_{j=1}^{m} f(x^*_{ij}, y^*_{ij})(x_i - x_{i-1})(y_j - y_{j-1}).$$



Al tomar particiones cada vez más finas, estas suman convergen hacia el volumen bajo la gráfica de f.

Una función arbitraria $f : D = [a, b] \times [c, d] \longrightarrow \mathbb{R}$ se dice *integrable* si existe el límite de sus sumas de Riemann. Dicho límite se dice **integral** de f y se denota:

$$\iint_D f(x, y)\, dxdy$$

1.2. Integrales iteradas

Por otra parte, dada $f : D = [a, b] \times [c, d] \longrightarrow \mathbb{R}$, para cada $y \in [c, d]$ podemos considerar la función de una variable

$$\begin{aligned} f(*, y) : [a, b] &\longrightarrow \mathbb{R} \\ x &\mapsto f(x, y), \end{aligned}$$

y su integral $F(y) = \int_a^b f(x, y)\, dx$, obteniendo así una función $F : [c, d] \longrightarrow \mathbb{R}$ que, a su vez, podemos integrar

$$\int_c^d \left(\int_a^b f(x, y)\, dx \right) dy\,. \tag{9.1}$$

También podemos proceder en sentido inverso:

$$\int_a^b \left(\int_c^d f(x, y)\, dy \right) dx\,. \tag{9.2}$$

Las integrales en (9.1) y (9.2) se dicen **integrales iteradas**.

Teorema. ***(Fubini)*** *Si $f : D = [a, b] \times [c, d] \longrightarrow \mathbb{R}$ es una función continua, entonces f es integrable y se tiene*

$$\boxed{\iint_D f(x, y)\, dxdy = \int_c^d \left(\int_a^b f(x, y)\, dx \right) dy = \int_a^b \left(\int_c^d f(x, y)\, dy \right) dx\,.}$$

Más generalmente, el resultado es cierto si f es acotada en D y discontinua solo en un número finito de curvas suaves.

Ejemplo. El volumen bajo la gráfica de

$$\begin{aligned} f : [0, 2] \times [0, 2] &\longrightarrow \mathbb{R} \\ (x, y) &\mapsto 16 - x^2 - 2y^2 \end{aligned}$$

es

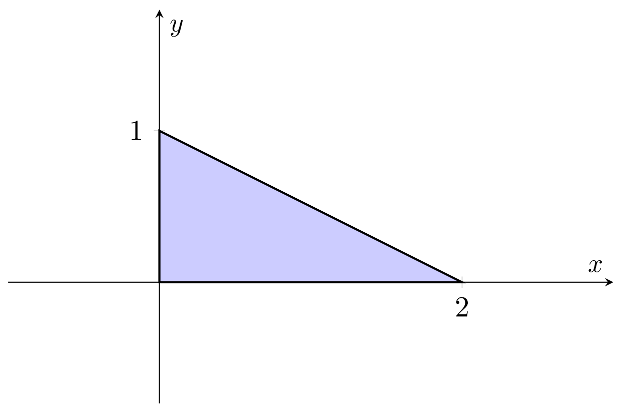
$$\iint_D f(x,y)\,dxdy = \int_0^2 \left(\int_0^2 (16-x^2-2y^2)\,dx\right) dy$$
$$= \int_0^2 \left[(16-2y^2)x - \frac{x^3}{3}\right]_{x=0}^{x=2} dy$$
$$= \int_0^2 \Big(32-4y^2-\frac{8}{3}\Big)dy = \left[\frac{88}{3}y - 4\frac{y^3}{3}\right]_0^2 = \frac{176-32}{3} = 48.$$

1.3. Integrales dobles sobre dominios generales

Si $f : D \longrightarrow \mathbb{R}$ es una función continua sobre un dominio acotado de $\mathbb{R}^2$, la integral doble $\iint_D f(x,y)\,dxdy$ se define también como límite de sumas de Rieman y se calcula mediante integrales iteradas.

Ejemplos:

- $\iint_D x\,dxdy$ siendo D el triángulo de vértices $(0,0)$, $(2,0)$ y $(0,1)$.



La recta que pasa por $(2,0)$ y $(0,1)$ es $x+2y=2$. Así tenemos

$$\iint_D f(x,y)\,dxdy = \int_0^2 \left(\int_0^{\frac{2-x}{2}} x\,dy\right) dx$$
$$= \int_0^2 [xy]_0^{\frac{2-x}{2}}\,dx = \frac{1}{2}\int_0^2 (2x-x^2)\,dx$$
$$= \frac{1}{2}\left[x^2-\frac{x^3}{3}\right]_0^2 = \frac{1}{2}\left(4-\frac{8}{3}\right) = \frac{2}{3}\,.$$

También podemos hacer

$$\begin{aligned}\iint_D f(x,y)\,dxdy &= \int_0^1 \left(\int_0^{2-2y} x\,dx\right) dy \\ &= \int_0^1 \left[\frac{x^2}{2}\right]_0^{2(1-y)} dy = 2\int_0^1 (1-y)^2\,dy \\ &= -2\left[\frac{(1-y)^3}{3}\right]_0^1 = \frac{2}{3}\,.\end{aligned}$$

- $\iint_D y\,dxdy$ siendo $D = \{(x,y)\in\mathbb{R}^2 \mid x^2+y^2 \le 1, y \ge 0\}$:

$$\begin{aligned}\iint_D y\,dxdy &= \int_0^1 \left(\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} y\,dx\right) dy = \int_0^1 [yx]_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\,dy \\ &= \int_0^1 2y\sqrt{1-y^2}\,dy = \left[-\frac{2}{3}(1-y^2)^{\frac{3}{2}}\right]_0^1 = \frac{2}{3}\,.\end{aligned}$$

Propiedades de las integrales dobles:

- $\iint_D \Big(f(x,y)+g(x,y)\Big)dxdy = \iint_D f(x,y)\,dxdy + \iint_D g(x,y)\,dxdy$,
- $\iint_D cf(x,y)\,dxdy = c\iint_D f(x,y)\,dxdy$ (c constante).
- Si $f(x,y) \ge g(x,y)$ $\forall (x,y)\in D$, entonces $\iint_D f(x,y)\,dxdy \ge \iint_D g(x,y)\,dxdy$.
- $\iint_D dxdy$ es el área de D.
- $\iint_D f(x,y)\,dxdy = \iint_{D_1} f(x,y)\,dxdy + \iint_{D_2} f(x,y)\,dxdy$, si $D = D_1 \cup D_2$ y D_1 y D_2 no se solapan excepto, quizás, en sus fronteras.

1.4. Cambio de variable en integrales dobles

En una variable, si deseamos calcular $\int_I f(x)\,dx$ y hacemos el cambio $x = \varphi(u)$, entonces

$$\int_I f(x)\,dx = \int_{\varphi^{-1}(I)} f(\varphi(u))|\varphi'(u)|\,du\,.$$

El valor absoluto se debe a que si φ es decreciente (derivada negativa) e $I = [a,b]$, $\varphi^{-1}(I) = [c,d]$, entonces $\varphi(d) = a$ y $\varphi(c) = b$, luego

$$\int_a^b f(x)\,dx = \int_d^c f(\varphi(u))\varphi'(u)\,du = \int_c^d f(\varphi(u))|\varphi'(u)|\,du.$$

En dos variables los cambios son de la forma $x = \varphi(u, v)$, $y = \psi(u, v)$, o

$$(x, y) = \Phi(u, v) := (\varphi(u, v), \psi(u, v)).$$

Debemos ahora sustituir *dxdy* por *dudv* multiplicado por el valor absoluto del *Jacobiano* de Φ: $\det \begin{pmatrix} \frac{\partial \varphi}{\partial u} & \frac{\partial \varphi}{\partial v} \\ \frac{\partial \psi}{\partial u} & \frac{\partial \psi}{\partial v} \end{pmatrix}$, de modo que se tiene

$$\boxed{\iint_D f(x, y)\, dxdy = \iint_{\Phi^{-1}(D)} f(\varphi(u, v), \psi(u, v)) \left| \det \begin{pmatrix} \frac{\partial \varphi}{\partial u} & \frac{\partial \varphi}{\partial v} \\ \frac{\partial \psi}{\partial u} & \frac{\partial \psi}{\partial v} \end{pmatrix} \right| dudv\,.}$$

Un cambio de variables muy útil es el paso a *coordenadas polares*:

$$x = r\cos\theta, \qquad y = r \operatorname{sen}\theta\,.$$

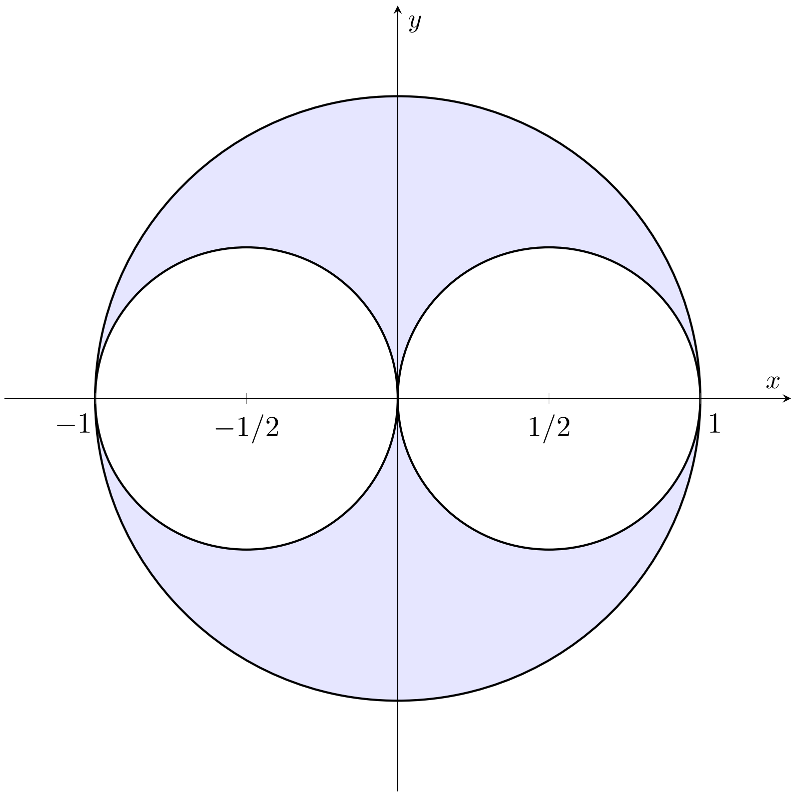
Ejemplos:

- Volumen de la *Bóveda de Viviani*:

$$V = \iint_D \sqrt{1 - x^2 - y^2}\, dxdy$$

donde el dominio es

$$D = \left\{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1,\ \left(x - \frac{1}{2}\right)^2 + y^2 \geq \frac{1}{4},\ \left(x + \frac{1}{2}\right)^2 + y^2 \geq \frac{1}{4} \right\}.$$



Por simetría $V = 4\iint_{D^+} \sqrt{1-x^2-y^2}\,dxdy$ con

$$\begin{aligned} D^+ &= \left\{(x,y)\in\mathbb{R}^2 \mid x,y\geq 0,\ x^2+y^2\leq 1,\ \left(x-\frac{1}{2}\right)^2+y^2\geq\frac{1}{4}\right\} \\ &= \{(x,y)\in\mathbb{R}^2 \mid x,y\geq 0,\ x^2+y^2\leq 1,\ x^2+y^2\geq x\} \end{aligned}$$

Cambiando a polares: $x = r\cos\theta$, $y = r\operatorname{sen}\theta$, el Jacobiano es

$$\begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\operatorname{sen}\theta \\ \operatorname{sen}\theta & r\cos\theta \end{vmatrix} = r(\cos^2\theta+\operatorname{sen}^2\theta) = r,$$

y D^+ se corresponde con

$$\tilde{D}^+ = \left\{(r,\theta)\in[0,+\infty)\times[0,2\pi] \mid 0\leq\theta\leq\frac{\pi}{2},\ r\leq 1,\ r\geq\cos\theta\right\},$$

luego tenemos

$$\begin{aligned} \iint_{D^+}\sqrt{1-x^2-y^2}\,dxdy &= \int_0^{\frac{\pi}{2}}\left(\int_{\cos\theta}^1 \sqrt{1-r^2}\,r\,dr\right)d\theta \\ &= \int_0^{\frac{\pi}{2}}\left[-\frac{1}{3}(1-r^2)^{\frac{3}{2}}\right]_{\cos\theta}^1 d\theta \\ &= \frac{1}{3}\int_0^{\frac{\pi}{2}}\operatorname{sen}^3\theta\ d\theta = \frac{1}{3}\int_0^{\frac{\pi}{2}}(1-\cos^2\theta)\operatorname{sen}\theta\ d\theta \\ &= -\frac{1}{3}\left[\cos\theta-\frac{\cos^3\theta}{3}\right]_0^{\frac{\pi}{2}} = \frac{1}{3}\left(1-\frac{1}{3}\right) = \frac{2}{9}. \end{aligned}$$

Así, el volumen de la bóveda de Viviani es $\dfrac{8}{9}$. (¡No depende de π!)

- La *campana de Gauss* es la gráfica de la función

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

que es un múltiplo escalar de la función de densidad de la distribución normal en Estadística.

Calcular directamente $A = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx$ tiene la dificultad de que no conocemos una primitiva. Pero, por el Teorema de Fubini[1]:

$$A^2 = \left(\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx\right)\left(\int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}}\,dy\right) = \iint_{\mathbb{R}^2} e^{-\frac{x^2+y^2}{2}}\,dxdy\,,$$

[1] El Teorema de Fubini se refiere a integrales sobre rectángulos acotados, pero tomando límites se prueba su validez también para rectángulos no acotados.

y cambiando a coordenadas polares, nos queda

$$\begin{aligned}\iint_{\mathbb{R}^2} e^{-\frac{x^2+y^2}{2}}\,dxdy &= \int_0^{+\infty}\int_0^{2\pi} e^{-\frac{r^2}{2}}\,r\,drd\theta\\ &= 2\pi\int_0^{+\infty} e^{-\frac{r^2}{2}}\,r\,dr\\ &= 2\pi\left[-e^{-\frac{r^2}{2}}\right]_0^{+\infty} = 2\pi\,.\end{aligned}$$

Por tanto:

$$\boxed{\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx = \sqrt{2\pi}\,.}$$

1.5. Superficies en $\mathbb{R}^3$

De modo análogo a la fórmula para la longitud de una curva en $\mathbb{R}^2$, el área de la gráfica de una función diferenciable $z = f(x, y)$ definida en un dominio $D \subseteq \mathbb{R}^2$ es

$$\boxed{\text{Área} = \iint_D \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\,dxdy\,.}$$

§ 2. Integrales triples

Dada una función de tres variables $f : D \longrightarrow \mathbb{R}$, $D \subseteq \mathbb{R}^3$, su integral

$$\iiint_D f(x, y, z)\,dxdydz$$

es el límite, como para dos variables, de las sumas de Riemann correspondientes. El Teorema de Fubini y las propiedades de las integrales dobles son válidas también, con los cambios naturales, para las integrales triples.

En particular, el volumen de una región $D \subseteq \mathbb{R}^3$ es

$$\iiint_D dxdydz\,.$$

Ejemplos:

- $D = [0, 1] \times [-1, 2] \times [0, 3]$,

$$\begin{aligned}\iiint_D x\,dxdydz &= \int_0^3 \left(\int_{-1}^2 \left(\int_0^1 x\,dx\right) dy\right) dz\\ &= \int_0^3 \left(\int_{-1}^2 \frac{1}{2}\,dy\right) dz = \int_0^3 \frac{3}{2}\,dz = \frac{9}{2}\,.\end{aligned}$$

- $\iiint_D \sqrt{x^2+y^2}\,dxdydz$, donde D es la región acotada por el paraboloide $z = x^2+y^2$ y el plano $z = 4$:

$$D = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2 \le z \le 4\}\,.$$

Con $\bar{D} = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 \le 4\}$, tenemos

$$\begin{aligned}\iiint_D \sqrt{x^2+y^2}\,dxdydz &= \iint_{\bar{D}} \left(\int_{x^2+y^2}^{4} \sqrt{x^2+y^2}\,dz \right) dxdy \\ &= \iint_{\bar{D}} \sqrt{x^2+y^2}(4-x^2-y^2)\,dxdy\ .\end{aligned}$$

Ahora pasamos a polares: $x = r\cos\theta$, $y = r\operatorname{sen}\theta$:

$$\begin{aligned}\iint_{\bar{D}} \sqrt{x^2+y^2}(4-x^2-y^2)\,dxdy &= \int_0^{2\pi}\int_0^2 r(4-r^2)r\,drd\theta \\ &= 2\pi \int_0^2 (4r^2-r^4)\,dr = 2\pi \left[4\frac{r^3}{3} - \frac{r^5}{5} \right]_0^2 \\ &= 2\pi \left(\frac{32}{3} - \frac{32}{5} \right) = \frac{128\pi}{15}\,.\end{aligned}$$