Álgebra lineal y Cálculo en varias variables

Un curso introductorio

Versión web

(link a versión para imprimir)

Gabriel Larotonda
FCEyN-UBA

2021

.

Introducción

Este texto pretende dar una introducción a los temas de cálculo en varias variables reales, y para ello no presupone conocimientos de álgebra lineal: las nociones y propiedades de vectores, subespacios y matrices necesarias para el desarrollo del cálculo están presentadas en los primeros capítulos. Esta presentación incluye las ideas de clasificación de superficies cuádricas, que creemos cumplen una doble función: mostrar la utilidad de las herramientas de matrices, y desarrollar una buena visión geométrica de los temas de extremos locales de funciones que visitaremos al final del texto.

Si bien el tratamiento es riguroso, hemos omitido la mayor parte de las demostraciones de los teoremas que enunciamos y utilizamos, concentrándonos más en sus interpretaciones, aplicaciones y usos. El lector interesado en las pruebas puede recurrir a textos clásicos de Cálculo como el de Stewart "Cálculo de varias variables" [6], o el texto del que suscribe "Cálculo y Análisis" [3].

Hemos incluido lecciones grabadas en video que recorren todos los temas de este texto de manera coherente, que pueden verse online o descargarse para ver luego. Los hipervínculos a esos videos pueden hallarse al comienzo de cada sección. No es necesario haber leído la sección correspondiente antes de ver la lección grabada en video; más bien el texto funciona como un acompañanate de aquellas clases, para leer y reforzar conceptos y ejemplos luego de la clase.

En el mismo espíritu, hemos usado el software GeoGebra para ilustrar muchos de los ejemplos que se presentan en el texto; el lector hallará que clickeando en el hipervínculo correspondiente se abre una página web de GeoGebra. En ocasiones el applet de Geogebra tiene sliders para modificar los gráficos presentados de manera dinámica; en todos

los casos las figuras espaciales pueden rotarse usando el mouse, y la escala del dibujo (zoom) se puede cambiar usando la rueda del mouse. En muchas ocasiones estas visualizaciones son de gran ayuda; en otros casos son simples curiosidades o juegos visuales. Dejamos en manos del lector experimentar y decidir cuál es el caso en cada una de ellas, ya que esto es en general una apreciación puramente subjetiva y personal.

Buenos Aires, Abril de 2021.

Capítulo 1

Vectores y sistemas lineales

El símbolo ☡ de camino sinuoso indica “precaución” o “atención”.

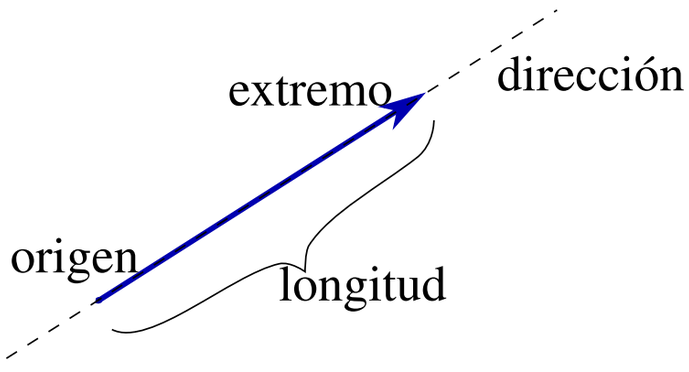
Para profundizar en los temas desarrollados en los primeros capítulos, recomendamos el libro de texto de Serge Lang “Introducción al Álgebra Lineal” [2].

1.1. Operaciones con vectores

• Corresponde a clases en video 1.1 y 1.2

DEFINICIÓN 1.1.1 (Vectores)**.** Un *vector* es una flecha que tiene

1. *dirección* (dada por la recta que la contiene),
2. *sentido* (dado por el lado hacia el que apunta la flecha),
3. *longitud* (dado por el largo desde el comienzo hasta el final.



- El punto del comienzo es el *origen* del vector y el punto del final el *extremo*.
- Para poder comparar vectores, es necesario que tengan un origen común.

DEFINICIÓN 1.1.2 (Pares ordenados)**.** Denotamos con $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ al conjunto de *pares ordenados* de números reales:

$$\mathbb{R}^2 = \{(a,b) : a,b \in \mathbb{R}\}.$$

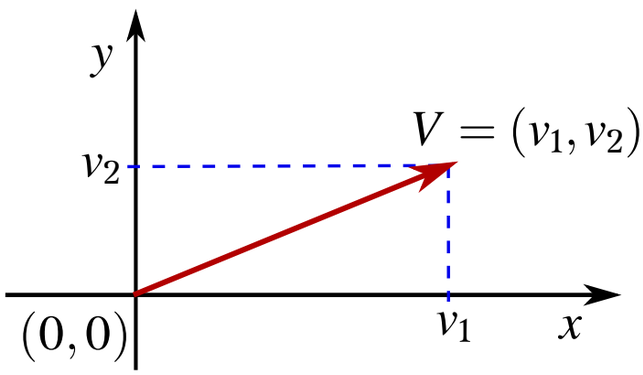


Figura 1.1: Plano cartesiano

Los pares ordenados de números reales se pueden representar en el plano con dos ejes perpendiculares, el eje horizontal donde van las x es el *eje de abcisas* y el eje vertical donde van las y es el *eje de ordenadas*. Este plano se denomina *plano euclideo* o *plano cartesiano*.

ƨ El orden es importante: el par $(1,3)$ es distinto del par $(3,1)$.

DEFINICIÓN 1.1.3 (Producto cartesiano)**.** Dados dos conjuntos A, B, el *producto cartesiano* de A con B es el nuevo conjunto que se consigue tomando pares ordenados de elementos de a y b respectivamente:

$$A \times B = \{(a,b) : a \in A, b \in B\}.$$

DEFINICIÓN 1.1.4 (Origen de coordendadas)**.** El par $(0,0)$ es el *origen de coordenadas* y lo anotamos con un cero grande $\mathbb{O} = (0,0)$.

DEFINICIÓN 1.1.5 (Vectores en coordenadas)**.** Dado un par ordenado $(a,b) \in \mathbb{R}^2$, lo podemos pensar como un punto del plano, pero también como un vector: dibujamos la flecha con origen en $(0,0)$ y extremo en (a,b) -ver la Figura 1.1-.

DEFINICIÓN 1.1.6 (Operaciones con vectores)**.** Sean $V = (v_1, v_2)$ $W = (w_1, w_2)$ y $\lambda \in \mathbb{R}$. Definimos

1. Producto de un vector por un número: $\lambda V = (\lambda v_1, \lambda v_2)$, se multiplican ambas coordenadas por el mismo número. Corresponde geométricamente a estirar, contraer o dar vuelta el vector (esto último con números negativos).

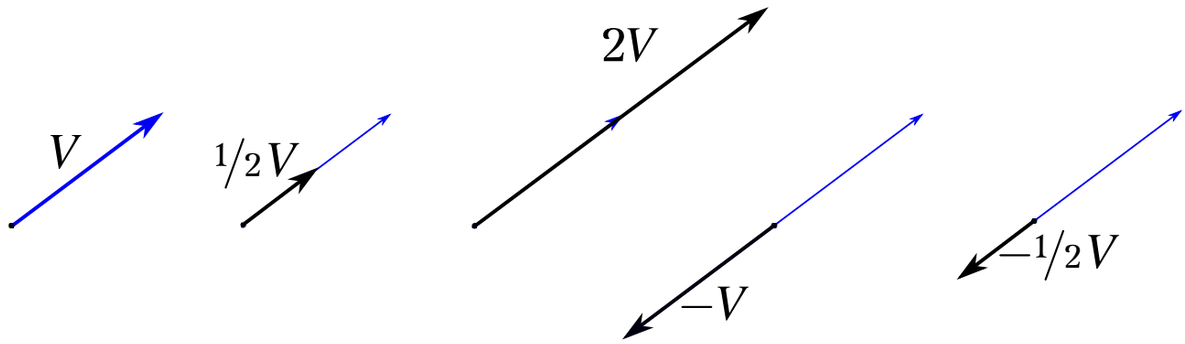


Figura 1.2: Producto de un número por el vector V.

2. Suma de vectores: $V + W = (v_1 + w_1, v_2 + w_2)$, se suman coordenada a coordenada. Corresponde geométricamente a tomar el vértico opuesto del paralelogramo formado por V y W:

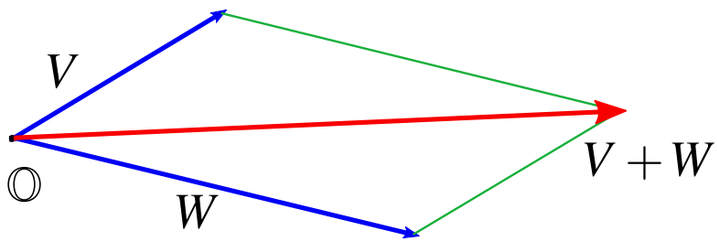


Figura 1.3: Suma de los vectores V, W.

PROPIEDADES 1.1.7 (suma de vectores)**.** Sean V, W, Z vectores y sean $\alpha, \beta \in \mathbb{R}$.

1. La suma es conmutativa y asociativa: $V + W = W + V$ y $(V + W) + Z = V + (W + Z)$.

2. El vector nulo (el origen) es el neutro de la suma: $V + \mathbb{O} = V$.

3. Multiplicar por el número 0 devuelve el origen: $0.V = \mathbb{O}$.

4. El producto por números distribuye con la suma: $\alpha(V + W) = \alpha V + \alpha W$ y también $(\alpha + \beta)V = \alpha V + \beta V$.

5. El vector *opuesto* se consigue multiplicando por -1: si $-V = (-1).V$ entonces $V + (-V) = V - V = \mathbb{O}$.

DEFINICIÓN 1.1.8 (Longitud de un vector)**.** Si $V = (v_1, v_2) \in \mathbb{R}^2$, su longitud se calcula usando el Teorema de Pitágoras que relaciona los lados de un triángulo rectángulo. Definimos

$$\|V\| = \sqrt{v_1^2 + v_2^2}$$

la *norma* de V, que no es otra cosa que su longitud (ver Figura 1.1).

PROPIEDADES 1.1.9 (Norma y operaciones con vectores)**.** Sean V, W vectores y $\lambda \in \mathbb{R}$.

1. $\|V\| \geq 0$ para todo V.
2. $\|V\| = 0$ si y sólo si $V = \mathbb{O}$.
3. $\|\lambda V\| = |\lambda| \, \|V\|$ donde $|\lambda|$ denota el módulo del número λ.
4. $\|V + W\| \leq \|V\| + \|W\|$ (desigualdad triangular).

DEFINICIÓN 1.1.10 (Producto interno)**.** Si $V = (v_1, v_2)$ y $W = (w_1, w_2)$ entonces el producto interno de V con W es un número real que puede ser cero, positivo o negativo:

$$V \cdot W = v_1 w_1 + v_2 w_2.$$

Otra notación: $V \cdot W = \langle V, W \rangle$.

PROPIEDADES 1.1.11 (Producto interno y operaciones con vectores)**.** Sean V, W, Z vectores y $\lambda \in \mathbb{R}$.

1. $V \cdot V = \|V\|^2$.
2. $(\lambda V) \cdot W = \lambda (V \cdot W) = V \cdot (\lambda W)$, *"saca escalares"*.
3. $(V + W) \cdot Z = V \cdot Z + W \cdot Z$, *"propiedad distributiva"*..

4. $V \cdot W = \|V\| \, \|W\| \cos\alpha$.

De la última propiedad, si $V, W \neq \mathbb{O}$, podemos escribir

$$\cos\alpha = \frac{V \cdot W}{\|V\| \, \|W\|}. \tag{1.1.1}$$

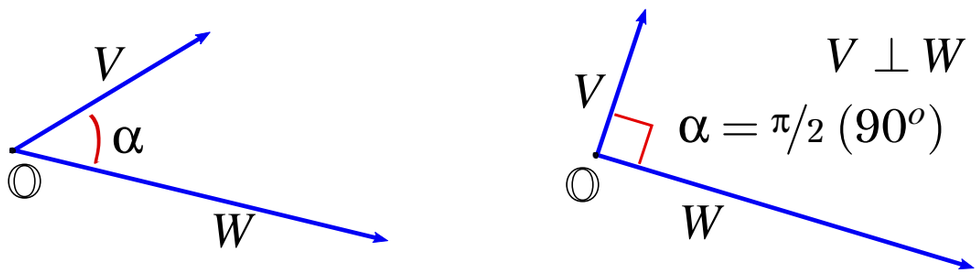


Figura 1.4: Ángulo entre V, W. En el segundo caso el ángulo es recto.

Como $|\cos\alpha| \leq 1$ para cualquier ángulo α, se tiene la *desigualdad de Cauchy-Schwarz*

$$-\|V\| \, \|W\| \leq V \cdot W \leq \|V\| \, \|W\|.$$

DEFINICIÓN 1.1.12 (Vectores ortogonales)**.** Decimos que V, W son *ortogonales* si son perpendiculares, y lo denotamos $V \perp W$.

OBSERVACIÓN 1.1.13 (Ortogonalidad y producto interno)**.** Dos vectores V, W son ortogonales $\Leftrightarrow$ el ángulo que forman es $\alpha = \pm\frac{\pi}{2}$ (90 grados) $\Leftrightarrow$ se tiene $\cos\alpha = 0$, y esto último por la ecuación (1.1.1) es equivalente a que $V \cdot W = 0$. Luego

$$V \perp W \Longleftrightarrow V \cdot W = 0. \tag{1.1.2}$$

OBSERVACIÓN 1.1.14 (Vectores en $\mathbb{R}^3$)**.** El espacio euclideo se representa con ternas ordenadas $V = (v_1, v_2, v_3) \in \mathbb{R}^3$, las operaciones son similares a las de $\mathbb{R}^2$. En general si $V = (v_1, v_2, \cdots, v_n) \in \mathbb{R}^n$ (y lo mismo para W, Z), definimos

1. $\lambda V = (\lambda v_1, \lambda v_2, \cdots, \lambda v_n)$

2. $V + W = (v_1 + w_1, v_2 + w_2, \cdots, v_n + w_n)$

3. $\|V\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$.

4. $V \cdot W = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n$.

Estas operaciones tienen las mismas propiedades ya enunciadas en 1.1.7, 1.1.9 y 1.1.11.

ƨ Dados $V, W \in \mathbb{R}^n$ podemos pensar que están en un plano bidimensional, luego aplican las reglas geométricas para producto por números y suma (usando la ley del paralelogramo). Es más, está bien definido el ángulo entre ellos, y se calcula usando la ecuación (1.1.1). También vale por este motivo $V \cdot W = 0 \Leftrightarrow V \perp W = 0$.

1.2. Rectas y planos

• Corresponde a clases en video 1.3, 1.4, 1.5a, 1.5b, 1.5c, 1.6, 1.7, 1.8

DEFINICIÓN 1.2.1 (Forma paramétrica de la recta)**.** Dados $V, P \in \mathbb{R}^n$ con $V \neq \mathbb{O}$, consideramos la *ecuación paramétrica* de la recta L dada por

$$L : tV + P, \quad t \in \mathbb{R}.$$

El vector V es el *vector director* de la recta, el punto P es el *punto de paso* de la recta y el número t es el *parámetro* de la recta.

DEFINICIÓN 1.2.2 (Recta y segmento entre dos puntos)**.** Dados $P \neq Q \in \mathbb{R}^n$, determinan una única recta con vector director $Q - P$ ya que $V = P - Q$ es una copia del segmento entre P y Q pero ahora con origen en $\mathbb{O}$. La ecuación de esta recta es entonces

$$L : t(Q - P) + P, \quad t \in \mathbb{R}.$$

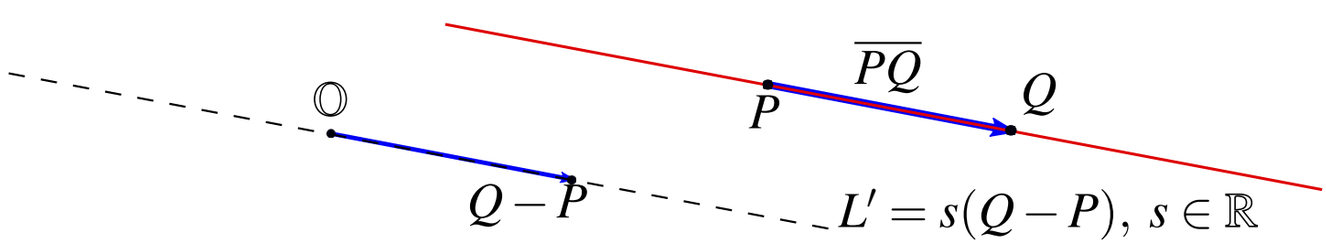


Figura 1.5: Recta que pasa por los puntos P, Q.

Cuando $t = 0$ obtenemos P y cuando $t = 1$ obtenemos Q. Si queremos únicamente los puntos entre P y Q (el *segmento* entre P y Q), restringimos el parámetro al intervalo pidiendo $t \in [0,1]$.

DEFINICIÓN 1.2.3 (Rectas paralelas)**.** Decimos que las rectas $L_1, L_2 \subset \mathbb{R}^n$ son *paralelas* si los vectores directores están alineados (tienen la misma dirección). Esto es, si

$$L_1 : tV + P \qquad L_2 : sW + Q$$

con los parámetros $s,t \in \mathbb{R}$, entonces L_1 es paralela a L_2 (denotado $L_1 // L_2$) si y sólo si $V = \lambda W$ para algún $\lambda \in \mathbb{R}$. Equivalentemente, como V, W son no nulos, son paralelas si y sólo si $W = \alpha V$ para algún $\alpha \in R$.

ƨ Las rectas L, L' de la Figura 1.5 son paralelas.

DEFINICIÓN 1.2.4 (Forma paramétrica del plano)**.** Dados $V, W, P \in \mathbb{R}^n$, el *plano* π *generado por* V, W *que pasa por* P tiene ecuación

$$\Pi : sV + tW + P, \quad s,t \in \mathbb{R}.$$

Los vectores V, W son los *generadores* de Π y s,t son los *parámetros*.

OBSERVACIÓN 1.2.5 (Plano que pasa por tres puntos)**.** Si $P, Q, R \in \mathbb{R}^n$ no están alineados, existe un único plano Π que pasa por los tres puntos. Su ecuación paramétrica es

$$\Pi : s(Q-P) + t(R-P) + P, \quad s,t \in \mathbb{R}.$$

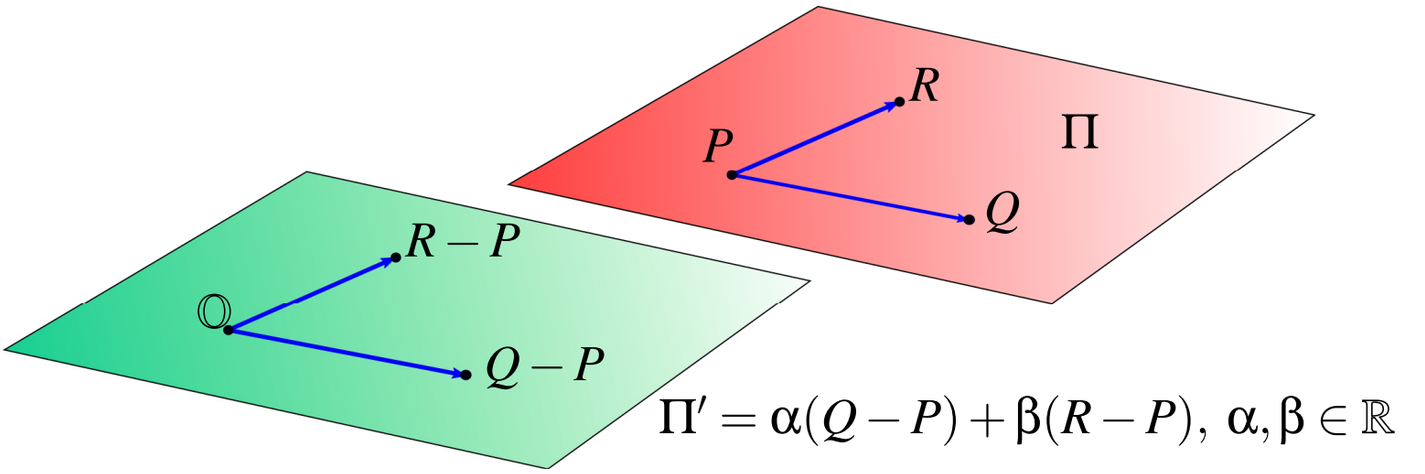


Figura 1.6: Plano Π *que contiene a los puntos* P, Q, R *(en rojo).*

DEFINICIÓN 1.2.6 (Producto vectorial)**.** Si $V, W \in \mathbb{R}^3$, el *producto vectorial* de V con W es un nuevo vector que es ortogonal a ambos, denotado $V \times W$. Se calcula armando una matriz 3×3 con V, W como segunda y tercera fila:

$$V \times W = \begin{vmatrix} i & j & k \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1).$$

El producto vectorial da el vector nulo si y solo si los vectores están alineados: $V \times W = \mathbb{O} \iff V = \lambda W$ ó $W = \lambda V$.

PROPIEDADES 1.2.7 (del producto vectorial)**.** Sean $V, W, Z \in \mathbb{R}^3$, $\lambda \in \mathbb{R}$. Entonces

1. $V \times W = -W \times V$ (*no* es conmutativo).
2. $(V + W) \times Z = V \times Z + W \times Z$.
3. $(\lambda V) \times W = \lambda(V \times W) = V \times (\lambda W)$.

OBSERVACIÓN 1.2.8. Si $\Pi : sV + tW$ es un plano por el origen ($s, t \in \mathbb{R}$), entonces

$$\begin{aligned}(V \times W) \cdot (sV + tW) &= V \times W \cdot (sV) + V \times W \cdot (tW) \\ &= t(V \times W) \cdot V + s(V \times W) \cdot W = t0 + s0 = 0\end{aligned}$$

por las propiedades recién enunciadas. Entonces $V \times W$ tiene producto interno nulo contra cualquier punto del plano Π. Luego $V \times W$ es perpendicular a *cualquier punto del plano* Π, por lo remarcado en (1.1.2).

DEFINICIÓN 1.2.9 (Vector normal a un plano)**.** Si $\pi : sV + tW + P$ es un plano en $\mathbb{R}^3$, un *vector normal* N_π del plano Π es cualquier vector no nulo que sea ortogonal a todos los puntos del plano.

ƨ Por la observación previa, siempre podemos tomar $N_\pi = V \times W$ si V, W son generadores del plano.

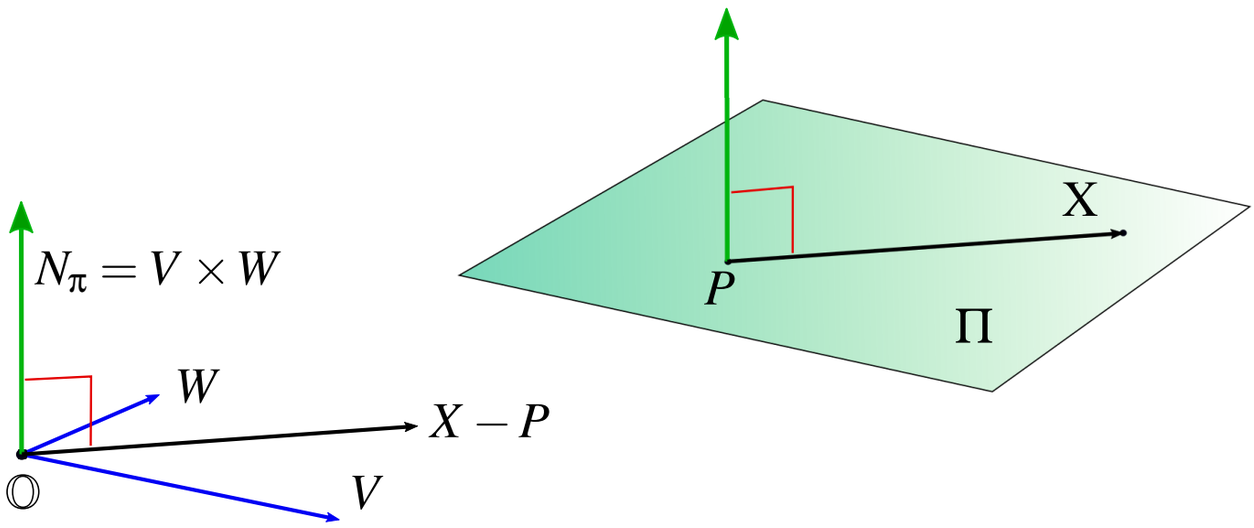


Figura 1.7: $X \in \Pi$ si y solo si $X - P \perp N_\pi$.

DEFINICIÓN 1.2.10 (Ecuación implícita del plano)**.** Sea $\Pi : sV + tW + P \subset \mathbb{R}^3$ un plano, $N_\pi = V \times W = (a,b,c)$ un vector normal al plano, entonce la ecuación implícita del plano es $\Pi : (X - P) \cdot N_\pi = 0$ (ver Figura 1.7). Equivalentemente

$$\Pi : \quad X \cdot N_\pi = P \cdot N_\pi.$$

En coordendas, si $X = (x,y,z)$ entonces la ecuación implícita es

$$\Pi : \quad ax + by + cz = d,$$

donde $d = P \cdot N_\pi \in \mathbb{R}$. Los números a,b,c,d son los *coeficientes* de la ecuación. Una tal ecucación es una *ecuación lineal*.

DEFINICIÓN 1.2.11 (Posiciones relativas)**.** Dos rectas $L_1, L_2 \subset \mathbb{R}^n$ son

1. *Paralelas* si tienen vectores directores alineados.
2. *Ortogonales* si se cortan y tiene vectores directores ortogonales.
3. *Alabeadas* si no se cortan pero tampoco son paralelas.

Dos planos $\Pi_1, \Pi_2 \subset \mathbb{R}^3$ son *paralelos* si tienen vectores normales paralelos; equivalentemente si $N_1 = \lambda N_2$ para algún $\lambda \in \mathbb{R}$. Los planos de la Figura 1.6 son paralelos.

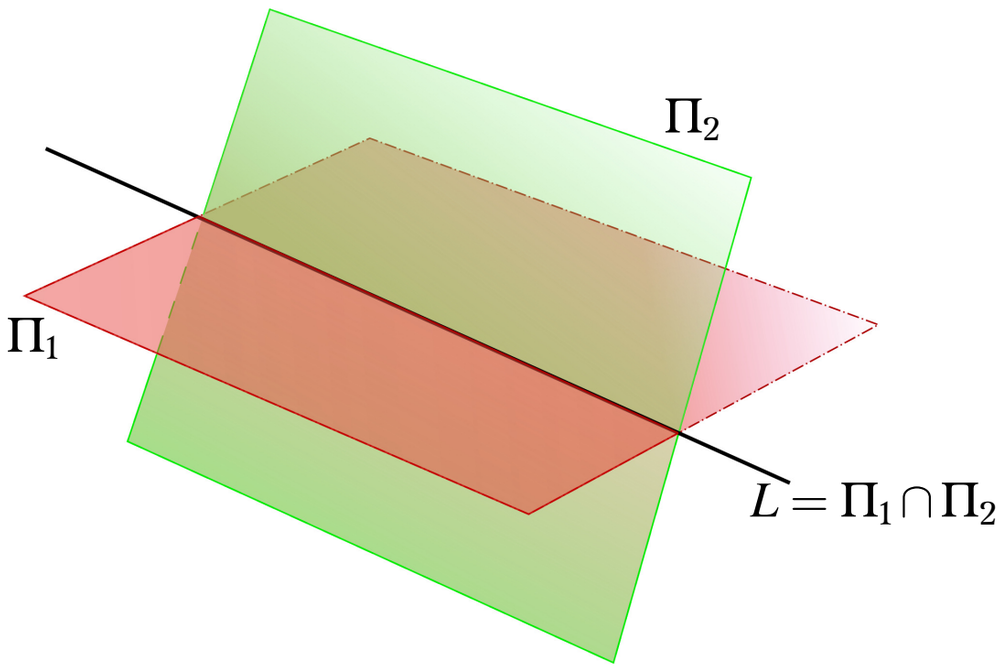


Figura 1.8: La intersección de dos planos (no paralelos) en $\mathbb{R}^3$ es una recta.

PROPOSICIÓN 1.2.12 (Dos planos en $\mathbb{R}^3$)**.** *Si $\Pi_1 \neq \Pi_2$ no son paralelos, entonces se cortan a lo largo de una recta en $\mathbb{R}^3$: $\Pi_1 \cap \Pi_2 = L$.*

1.3. Sistemas de ecuaciones lineales

• Corresponde a clases en video 1.9, 1.10

DEFINICIÓN 1.3.1 (Sistema)**.** Un sistema en n variables y con k ecuaciones se escribe como

$$S: \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots & = \ \vdots \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n &= b_k \end{cases} \tag{1.3.1}$$

- Los números a_{ij} con $i = 1,2,\ldots,k,\ j = 1,2,\ldots,n$ son los *coeficientes* de las ecuaciones del sistema.

- El sistema es *homogéneo* si todos los b_i (los términos independientes) son nulos.

- Una *solución* del sistema es una n-upla $X = (x_1, x_2, \ldots, x_n)$ que verifica *simultáneamente todas* las ecuaciones.

- Cuando se pide hallar las soluciones del sistema, siempre se entiende que queremos hallar *todas* las soluciones, todos los $X \in \mathbb{R}^n$ que verifican las ecuaciones.

ƨ Si el sistema es homogéneo, $X = \mathbb{O}$ es solución, decimos que es la *solución trivial* (puede haber otras).

DEFINICIÓN 1.3.2 (Sistemas equivalentes)**.** Dos sistemas de ecuaciones S_1, S_2 son *equivalentes* si tienen el mismo conjunto de soluciones.

ƨ Dos sistemas equivalentes no tienen por qué tener la misma cantidad de ecuaciones. Por ejemplo

$$S_1 : \begin{cases} 3x + y - z &= 4 \\ -6x - 2y + 2z &= -8 \end{cases} \qquad S_2 : \{3x + y - z = 4$$

son equivalentes porque la segunda ecuación de S_1 es en realidad, una copia de la primera ecuación (multiplicada por (-2)).

PROPOSICIÓN 1.3.3 (Operaciones con filas)**.** *Las operaciones permitidas para transformar un sistema en otro equivalente son dos:*

1. *Multiplicar una fila por un número $\lambda \neq 0$ (ambos lados del igual).*

2. *Sumar dos filas y guardar el resultado en el lugar de cualquiera de las dos que sumamos.*

2 Combinando estas dos operaciones podemos escribir

$$F_i + \lambda F_j \longrightarrow F_i \qquad \lambda \in \mathbb{R},$$

que se interpreta como: multiplicar la fila j por λ, sumarle la fila i, y guardar el resultado en el lugar de la fila i.

DEFINICIÓN 1.3.4 (Clasificación del conjunto de soluciones)**.** Los sistemas lineales se clasifican en *compatibles* e *incompatibles*. Los compatibles tienen al menos una solución, los incompatibles no tienen solución (equivalentemente, el conjunto de soluciones de un sistema incompatible es vacío).
Los sistemas compatibles se clasifican en *compatible determinado* (cuando tienen una única solución) y *compatible indeterminado* (cuando tiene más de una solución; en ese caso siempre tiene infinitas soluciones).

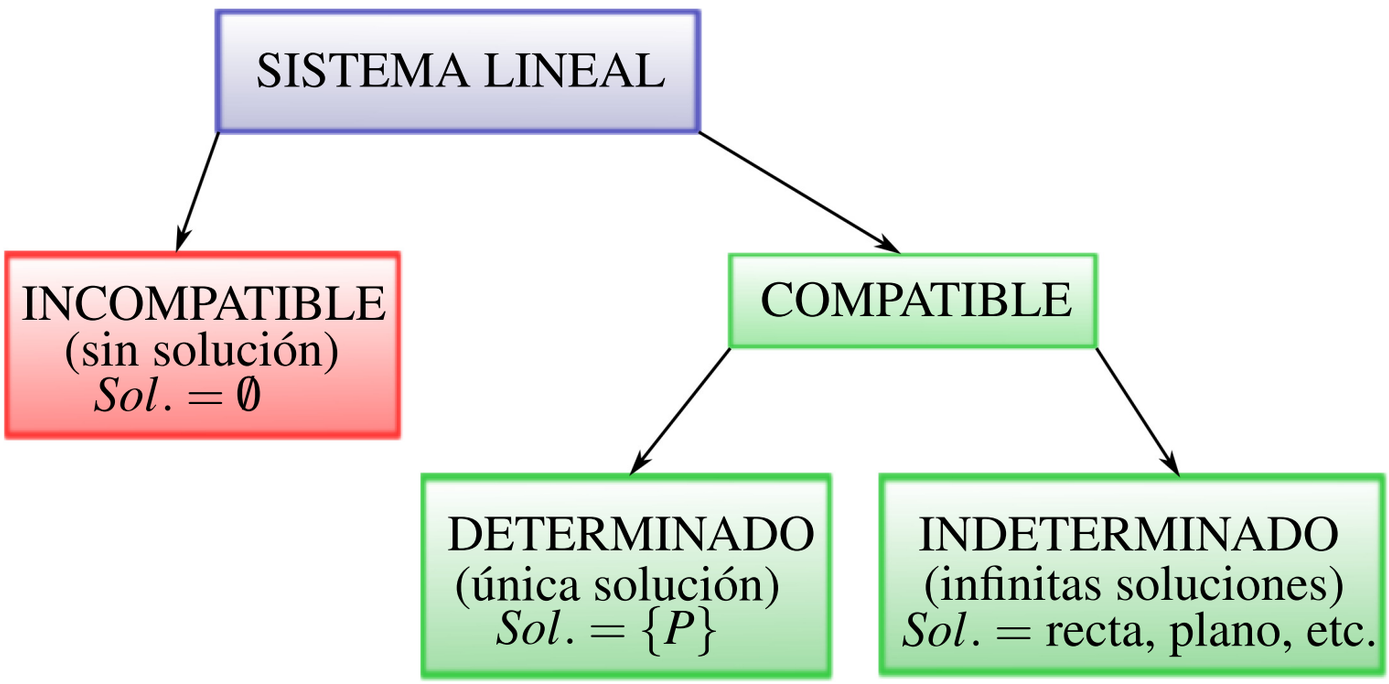


Figura 1.9: Clasificación de los sistemas de ecuaciones lineales.

DEFINICIÓN 1.3.5 (Matriz ampliada de un sistema lineal)**.** Dado un sistema lineal como (1.3.1), podemos escribir su *matriz ampliada* co-

mo

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} & b_k \end{array}\right).$$

El bloque que está a la izquierda de la barra vertical se denomina *matriz* del sistema; esta matriz tiene *n columnas* y *k filas*.

DEFINICIÓN 1.3.6 (Método de Gauss y matriz escalonada)**.** El *método de Gauss* para resolver un sistema lineal consiste en hacer operaciones con las filas de la matriz ampliada, hasta obtener una *matriz escalonada*: cada fila de una matriz escalonada tiene su primer coeficiente no nulo (comenzando desde la izquierda) *a la derecha* del primer coeficiente no nulo de la fila superior.

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & a_{kn} & b_k \end{array}\right).$$

Puede ocurrir que el primer coeficiente no nulo de la fila inferior esté *varios lugares* a la derecha, como en este ejemplo de matriz escalonada:

$$\left(\begin{array}{ccccccc|c} 3 & 0 & -2 & 7 & -4 & 0 & -2 & 2 \\ 0 & 0 & 0 & 3 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 7 & -1 & 2 \end{array}\right). \tag{1.3.2}$$

DEFINICIÓN 1.3.7 (Ecuaciones independientes)**.** En una matriz (o un sistema) escalonado, las ecuaciones son *independientes* (descartamos las filas de ceros). La cantidad de filas (ecucaciones) independientes es la cantidad de *restricciones* del sistema.

DEFINICIÓN 1.3.8 (Solución de un sistema lineal)**.** Se busca dar una expresión paramétrica del conjunto de soluciones del sistema, la *dimensión* del conjunto solución es la cantidad de parámetros indepen-

dientes. Se puede calcular la dimensión del conjunto como la dimensión del espacio total menos la cantidad de restricciones (cantidad de ecuaciones independientes).

EJEMPLO 1.3.9 (Soluciones del sistema (1.3.2))**.** Como la matriz del sistema (sin ampliar) tiene 7 columnas, es un sistema con siete variables y las soluciones son de la forma $X = (x_1, x_2, x_3, x_4, x_5, x_6, x_7)$. El sistema está escalonado, las tres ecuaciones son indpendientes. El conjunto solución tiene que tener dimensión $7-3=4$; la solución se tiene que poder expresar usando 4 parámetros independientes. Lo resolvemos: la última fila se reescribe como ecuación

$$7x_6 - x_7 = 2.$$

De aquí despejamos la variable x_6 en función de x_7

$$7x_6 = 2 + x_7 \Longrightarrow x_6 = {}^1\!/_7\, x_7 + {}^2\!/_7.$$

La segunda fila se corresponde con la ecuación $3x_4 + x_5 - 2x_6 + x_7 = 0$. De aquí despejamos $3x_4 = -x_5 + 2x_6 - x_7$. Reemplazando la variable x_6 en función de la x_7 según lo deducido de la tercer fila, y dividiendo luego por 3, se obtiene

$$x_4 = -{}^1\!/_3\, x_5 - {}^5\!/_{21}\, x_7 + {}^4\!/_{21}.$$

De la primer fila despejamos la variable x_1 en función de las demás variables, reemplazando además las variables x_6 y x_5 por los valores ya despejados (que estaban en función de las variables x_5 y x_7):

$$x_1 = {}^2\!/_3\, x_3 + {}^{19}\!/_9\, x_5 + {}^{11}\!/_9\, x_7 + {}^2\!/_9.$$

Las variables x_2, x_3, x_5, x_7 quedaron como variables libres: son los 4 parámetros independientes. La solución general del sistema es

$$(\frac{2}{3}\, x_3 + \frac{19}{9}\, x_5 + \frac{11}{9}\, x_7 + \frac{2}{9}, x_2, x_3, \frac{-1}{3}\, x_5 - \frac{5}{21}\, x_7 + \frac{4}{21}, x_5, \frac{1}{7}\, x_7 + \frac{2}{7}, x_7)$$

que se puede reescribir usando las propiedades de las operaciones con

vectores como

$$Sol. = x_2(0,1,0,0,0,0,0) + x_3(2/3,0,1,0,0,0,0) + x_5(19/9,0,0,-1/3,1,0,$$
$$+x_7(11/9,0,0,-5/21,0,1/7,1) + (2/9,0,0,4/21,0,2/7,0).$$

Podemos cambiar el nombre de los cuatro parámetros independientes y reescribir la solución

$$Sol. = \alpha(0,1,0,0,0,0,0) + \beta(2/3,0,1,0,0,0,0) + \gamma(19/9,0,0,-1/3,1,0,0)$$
$$+ \delta(11/9,0,0,-5/21,0,1/7,1) + (2/9,0,0,4/21,0,2/7,0) \quad \alpha,\beta,\gamma,\delta \in \mathbb{R}$$

de donde es claro que el conjunto solución tiene dimensión 4.

1.4. Matrices

• Corresponde a clases en video 1.11, 1.12, 1.13

DEFINICIÓN 1.4.1 ($\mathbb{R}^{k\times n}$)**.** Es el conjunto de los arreglos de números (o *coeficientes*) de k filas y n columnas ubicados en un rectángulo, encerrado por paréntesis $(\cdot)$ o corchetes $[\cdot]$.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{pmatrix} \in \mathbb{R}^{k\times n}.$$

Las entradas se indican con un subíndice, el primero indica la fila y el segundo la columna. Así por ejemplo a_{21} es el coeficiente de la segunda fila y la primer columna, mientras que a_{12} es el coeficiente de la primer fila y la segunda columna.

ƨ Dado un vector $V \in \mathbb{R}^d$, podemos pensarlo como un vector fila (matriz en $\mathbb{R}^{1\times d}$), o como un vector columna (matriz de $\mathbb{R}^{d\times 1}$), según se indique (o sea conveniente).

DEFINICIÓN 1.4.2 (Operaciones con matrices: suma y producto por

números). Sean $A, B \in \mathbb{R}^{k\times n}$. Entonces la suma $A+B \in \mathbb{R}^{k\times n}$ se calcula lugar a lugar. El producto de la matriz A por el número λ da también una matriz del mismo tamaño, cuyas entradas se obtienen multiplicando todas las entradas de A por λ. Valen las propiedades distributivas de suma con producto.

DEFINICIÓN 1.4.3 (Operaciones con matrices: producto). Si $A \in \mathbb{R}^{k\times n}$ y $B \in \mathbb{R}^{n\times d}$, se calcula el producto AB haciendo el producto interno de las filas de A por las columnas de B, obteniéndose una matriz de $k \times d$, de la siguiente manera:

$$AB = \begin{pmatrix} - & F_1 & - \\ - & F_2 & - \\ \vdots & & \vdots \\ - & F_k & - \end{pmatrix} \begin{pmatrix} | & | & \cdots & | \\ C_1 & C_2 & & C_d \\ | & | & \cdots & | \end{pmatrix} = \begin{pmatrix} F_1 \cdot C_1 & F_1 \cdot C_2 & \cdots & F_1 \cdot C_d \\ F_2 \cdot C_1 & F_2 \cdot C_2 & \cdots & F_2 \cdot C_d \\ \vdots & & & \vdots \\ F_k \cdot C_1 & F_k \cdot C_2 & \cdots & F_k \cdot C_d \end{pmatrix}.$$

- El producto se puede hacer si y solo si la cantidad de *columnas* de A (la matriz de la izquierda) es igual a la cantidad de *filas* de B (la matriz de la derecha; en el ejemplo de arriba este número es n. Así cada F_i tiene de largo n lugares, y cada C_j también. De esta manera, el producto interno de F_i con C_j está bien definido.
- Aunque se puedan calcular ambos productos, en general $AB \neq BA$.

PROPIEDADES 1.4.4 (Distributiva de la suma y el producto de matrices). Si A, B tienen el mismo tamaño, se pueden sumar. Si C tiene el tamaño adecuado para poder multiplicar por A (y por B) a la izquierda, entonces también se puede multiplicar por $A+B$ (a la izquierda) y vale

$$C \cdot (A+B) = C \cdot A + C \cdot B.$$

Lo mismo aplica si puede hacerse el producto por la derecha.

DEFINICIÓN 1.4.5 (Matriz traspuesta). Si $A \in \mathbb{R}^{k\times n}$, la matriz *traspuesta* de A (denotada A^t) es la matriz que se obtiene intercambiando

las filas de A por las columnas de A. Entonces $A^t \in \mathbb{R}^{n\times k}$. Por ejemplo

$$A = \begin{pmatrix} 2 & 3 \\ -1 & 0 \\ 4 & -2 \end{pmatrix} \in \mathbb{R}^{3\times 2} \quad \Longrightarrow \quad A^t = \begin{pmatrix} 2 & -1 & 4 \\ 3 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{2\times 3}.$$

Si $A \in \mathbb{R}^{n\times n}$ (es una matriz cuadrada, es decir, tiene la misma cantidad de filas que de columnas) entonces la matriz traspuesta de A se obtiene haciendo una reflexión de los coeficientes de A respecto de la diagonal. Decimos entonces que A es *simétrica* si $A^t = A$.

OBSERVACIÓN 1.4.6 (Escritura de un sistema lineal usando el producto)**.** Si $A \in \mathbb{R}^{k\times n}$ es la matriz de coeficientes de un sistema lineal y $X \in \mathbb{R}^n$ es el vector de las variables, el sistema (1.3.1) se puede escribir como

$$AX = b$$

donde $b \in \mathbb{R}^k$ es el vector columna de los términos independientes (hay tantos como filas). Esto es porque

$$AX = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n \end{pmatrix}.$$

Igualando este vector columna al vector columna b, recuperamos el sistema de ecuaciones lineales (1.3.1).

DEFINICIÓN 1.4.7 (Sistema homogéneo asociado)**.** Dado el sistema lineal $AX = b$, el *sistema homogéneo asociado* es el sistema lineal $AX = \mathbb{O}$, donde $\mathbb{O}$ denota el vector columna de ceros (del mismo largo que el vector columna b).

OBSERVACIÓN 1.4.8. Sea $AX = b$ sistema lineal. Entonces

1. Si V es solución del sistema homogéneo asociado y P es solución del sistema original, entonces $V+P$ también es solución del sistema original:

$$A(V+P) = AV + AP = \mathbb{O} + b = b.$$

2. Si P, Q son soluciones del sistema original, entonces $V = P - Q$ es solución del sistema homogéneo asociado:

$$AV = A(P-Q) = AP - AQ = b - b = \mathbb{O}.$$

3. Si S_0 es el conjunto de *todas* las soluciones del sistema homogéneo, y P es solución del sistema original, entonces

$$S_b = S_0 + P = \{V + P : V \in S_0\}$$

es el conjunto de *todas* las soluciones del sistema original.

DEFINICIÓN 1.4.9 (Combinación lineal)**.** Una *combinación lineal* de dos vectores (o dos matrices) V, W que están en el mismo espacio, es cualquier suma de dos múltiplos de V, W. Es decir, cualquier elemento de la forma

$$\alpha V + \beta W \qquad \alpha, \beta \in \mathbb{R}.$$

ƨ Si V, W son dos soluciones de un sistema homogéneo, cualquier combinación lineal de ellos también es solución del mismo sistema homogéneo.

PROPIEDADES 1.4.10 (de un sistema homogéneo)**.** Sea $AX = \mathbb{O}$ un sistema homogéneo. Entonces

1. Si V es solución, λV es solución para cualquier $\lambda \in \mathbb{R}$:

$$A(\lambda V) = \lambda(AV) = \lambda\mathbb{O} = \mathbb{O}.$$

2. Si V, W son soluciones, entonces la suma es solución:

$$A(V+W) = AV + AW = \mathbb{O} + \mathbb{O} = \mathbb{O}.$$

DEFINICIÓN 1.4.11 (Subespacio)**.** Un *subespacio* S de $\mathbb{R}^n$ es el conjunto de soluciones de un sistema homogéneo. Equivalentemente, es un conjunto $S \subset \mathbb{R}^n$ tal que

1. $\mathbb{O} \in S$
2. $V \in S \Rightarrow \lambda V \in S$ para todo $\lambda \in \mathbb{R}$.
3. $V, W \in S \Rightarrow V + W \in S$.

ƨ Cualquier combinación lineal de elementos de un subespacio es también un elemento del subespacio.

OBSERVACIÓN 1.4.12 (Rectas y planos como subespacios)**.** Una recta es un subespacio si y solo si contiene al origen. Lo mismo ocurre con un plano cualquiera: es un subespacio si y solo si contiene al origen.

DEFINICIÓN 1.4.13 (Generadores de un subespacio)**.** Los vectores $V_1, V_2, \cdots, V_r \in \mathbb{R}^n$ son *generadores* del subespacio $S \subset \mathbb{R}^n$ si todo elemento de S se puede escribir como combinación lineal de estos vectores: si $X \in S$, existen $\alpha_1, \alpha_2, \ldots, \alpha_r \in \mathbb{R}$ tales que

$$X = \alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_r V_r.$$

EJEMPLO 1.4.14 (Planos y rectas)**.** Dada una recta por el origen, es un subespacio y si tenemos su forma paramétrica $L : tV$, $t \in \mathbb{R}$, entonces el generador es V (alcanza con uno solo).
Dado un plano por el origen, si tenemos su forma paramétrica $\Pi : tV + sW$, $s, t \in \mathbb{R}$, entonces los generadores son V, W (necesitamos dos generadores).

EJEMPLO 1.4.15 (Comprobar si un vector está en un subespacio)**.** Si $V = (1,3,1), W = (0,2,-1)$ son generadores del subesapacio S, para saber si $P = (2,1,7)$ es un elemento de S, hay que plantear

$$(2,1,7) = \alpha(1,3,1) + \beta(0,2,-1).$$

Igualando las coordenadas obtenemos un sistema lineal con tres ecuaciones (hay tres coordenadas) y dos incógnitas (α y β). Su matriz ampliada es

$$\left(\begin{array}{cc|c} 1 & 0 & 2 \\ 3 & 2 & 1 \\ 1 & -1 & 7 \end{array}\right).$$

Si hacemos $F_3 - F_1 \to F_3$, luego $F_2 - 3F_1 \to F_2$, y finalmente $2F_3 + F_2 \to F_3$, llegamos a

$$\left(\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 2 & -5 \\ 0 & 0 & 5 \end{array}\right).$$

La última ecuación dice $0 = 5$ entonces el sistema es incompatible. Esto nos dice que P *no* pertenece a S, lo escribimos como $P \notin S$.

DEFINICIÓN 1.4.16 (Matriz nula y matriz identidad)**.** La *matriz nula* es aquella cuyas entradas son todas 0, la denotamos $\mathbb{O}_{k\times n}$ o $\mathbb{O}$ si se entiende su tamaño del contexto. Es el neutro de la suma: para toda matriz $A \in \mathbb{R}^{k\times n}$ se tiene

$$A + \mathbb{O} = A.$$

La *matriz identidad* es la matriz cuadrada de $n \times n$ cuyas entradas son todas nulas salvo las de la diagonal, que son 1, la denotamos $\mathbb{I}_n$. Por ejemplo:

$$\mathbb{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathbb{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \mathbb{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Es el neutro del producto de las matrices de $n \times n$: si $A \in \mathbb{R}^{n\times n}$ entonces

$$A\,\mathbb{I}_n = A = \mathbb{I}_n A.$$

DEFINICIÓN 1.4.17 (Matriz inversa)**.** Si $A \in \mathbb{R}^{n\times n}$, decimos que A tiene inversa (o que A es *inversible*) si existe una matriz $B \in \mathbb{R}^{n\times n}$ tal que

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

Si existe, la matriz B es única y en general se denota A^{-1} (se lee A *a la menos uno*), es la *matriz inversa* de A. Entonces si A es inversible, vale

$$AA^{-1} = \mathbb{I}_n = A^{-1}A.$$

ƨ Si $A, B \in \mathbb{R}^{n\times n}$ verifican $AB = \mathbb{I}_n$, entonces A es la (única) inversa de B y recíprocamente B es la (única) inversa de A. Se puede probar que en ese caso la condición $BA = \mathbb{I}_n$ se cumple automáticamente (no hace falta chequearla).

OBSERVACIÓN 1.4.18 (Cálculo de la matriz inversa)**.** Para ver si $A \in \mathbb{R}^{n\times n}$ tiene inversa (y calcularla en ese caso), planteamos un sistema con matriz ampliada que a la derecha lleva la matriz identidad: $(A|\mathbb{I}_n)$ o más explícitamente

$$\left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} & 0 & \cdots & 0 & 1 \end{array}\right).$$

Aplicamos el método de Gauss a la matriz de la izquierda, hay que escalonar debajo de la diagonal y luego también *arriba* de la diagonal.

- Supongamos que llegamos a $(\mathbb{I}_n|B)$. Entonces A es inversible y La matriz de la derecha es la inversa de A. Es decir, luego de escalonar obtenemos $(\mathbb{I}_n|A^{-1})$.

- Si al triangular, a la izquierda nos queda alguna fila de ceros, la matriz A *no tiene inversa*.

PROPIEDADES 1.4.19 (Matrices inversibles y sistemas)**.** Sea $A \in \mathbb{R}^{n\times n}$ la matriz asociada al sistema lineal $AX = b$ (aquí $X, b \in \mathbb{R}^n$) ya que A es cuadrada; el sistema tiene la misma cantidad de incógnitas que de ecuaciones). Entonces

1. El sistema es compatible determinado si y solo si A es inversible. Si tenemos la inversa A^{-1} entonces

$$AX = B \;\to\; A^{-1}(AX) = A^{-1}b \;\to\; \mathbb{I}_n X = A^{-1}b \;\to\; X = A^{-1}b.$$

   Luego la única solución del sistema es $X = A^{-1}b$.

2. Si $b \neq \mathbb{O}$, y A no es inversible, puede tratarse de un sistema incompatible o de un sistema compatible indeterminado.

3. Si $b = \mathbb{O}$ (sistema homogéneo) entonces

   a) La única solución es $X = \mathbb{O}$ si y sólo si A es inversible.

   b) Si A no es inversible el sistema es compatible indeterminado.

1.5. Determinante

• Corresponde a clase en video 1.14

DEFINICIÓN 1.5.1 (Determinante)**.** Es un número real que puede ser positivo, negativo o nulo, denotado $\det(A)$. Cuando tenemos la matriz podemos cambiar los paréntesis por barras verticales para indicar que se trata del determinante. Se calcula en $A \in \mathbb{R}^{2\times 2}$ de la siguiente manera:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Para matrices de 3×3 se elige una fila o una columna y se desarrolla siguiendo la regla de signos

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$$

y tomando los determinantes de las matrices que quedan al tachar esa fila y esa columna.

EJEMPLO 1.5.2 (Determinante de una matriz 3×3)**.** Calculamos por la primer fila

$$\begin{aligned} \det(A) = \begin{vmatrix} 3 & 2 & -7 \\ 1 & 2 & 0 \\ 4 & -2 & 9 \end{vmatrix} &= 3 \begin{vmatrix} 2 & 0 \\ -2 & 9 \end{vmatrix} - 2 \begin{vmatrix} 1 & 0 \\ 4 & 9 \end{vmatrix} + (-7) \begin{vmatrix} 1 & 2 \\ 4 & -2 \end{vmatrix} \\ &= 3(2 \cdot 9 - 0(-2)) - 2(1 \cdot 9 - 0 \cdot 4) - 7(1(-2) - 2 \cdot 4) \\ &= 3 \cdot 18 - 2 \cdot 9 - 7(-10) = 54 - 18 + 70 = 106. \end{aligned}$$

Entonces $\det(A) = 106$. Si intentamos por otra fila (o columna), llegamos al mismo resultado. Por ejemplo, si lo calculamos por la segunda fila:

$$\begin{aligned} \begin{vmatrix} 3 & 2 & -7 \\ 1 & 2 & 0 \\ 4 & -2 & 9 \end{vmatrix} &= (-1) \begin{vmatrix} 2 & -7 \\ -2 & 9 \end{vmatrix} + 2 \begin{vmatrix} 3 & -7 \\ 4 & 9 \end{vmatrix} - 0 \begin{vmatrix} 3 & 2 \\ 4 & -2 \end{vmatrix} \\ &= (-1)(2 \cdot 9 - (-7)(-2)) + 2(3 \cdot 9 - (-7) \cdot 4) - 0 \\ &= (-1)(18 - 14) + 2(27 + 28) \\ &= (-1)4 + 2 \cdot 55 = -4 + 110 = 106. \end{aligned}$$

Es conveniente calcular el determinante usando la fila (o la columna) que tenga más ceros.

EJEMPLO 1.5.3 (Determinante de una matriz de 4×4)**.** La regla de los signos es similar a la de 3×3, se comienza con un $+$ en el lugar 11

(la esquina superior izquierda) y se van alternando los signos:

$$\begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}.$$

Damos un ejemplo, calculemos el determinante de la matriz

$$B = \begin{vmatrix} 2 & -1 & 0 & 5 \\ 3 & 0 & -2 & 0 \\ 1 & -1 & 4 & 3 \\ 2 & 3 & 0 & 1 \end{vmatrix}.$$

Desarrollamos por la segunda fila:

$$= (-1)\cdot 3 \begin{vmatrix} -1 & 0 & 5 \\ -1 & 4 & 3 \\ 3 & 0 & 1 \end{vmatrix} + 0 \begin{vmatrix} 2 & 0 & 5 \\ 1 & 4 & 3 \\ 2 & 0 & 1 \end{vmatrix} - (-2) \begin{vmatrix} 2 & -1 & 5 \\ 1 & -1 & 3 \\ 2 & 3 & 1 \end{vmatrix} + 0 \begin{vmatrix} 2 & -1 & 0 \\ 1 & -1 & 4 \\ 2 & 3 & 0 \end{vmatrix}.$$

Hay dos términos nulos (en la fila habia dos ceros). Calculamos los dos determinantes de 3×3 que quedaron (desarrollando por alguna fila o columna; por ejemplo el primero es conveniente desarrollarlo por la segunda *columna*, ya que tiene dos ceros). Obtenemos

$$\det(B) = -3 \cdot (-16) + 2 \cdot 0 = 48.$$

PROPIEDADES 1.5.4 (del determinante)**.** Sean $A, B \in \mathbb{R}^{n\times n}$, $\lambda \in \mathbb{R}$. Entonces

1. $\det(A^t) = \det(A)$.
2. $\det(A^k) = (\det(A))^k$ para todo $k \in \mathbb{N}$.
3. $\det(AB) = \det(BA) = \det(A)\det(B)$
4. $\det(\lambda A) = \lambda^n \det(A)$

ꙅ En general es FALSO que $\det(A+B) = \det(A) + \det(B)$, es decir el determinante no es una función lineal.

TEOREMA 1.5.5 (Determinante versus inversa)**.** *Sea $A \in \mathbb{R}^{n\times n}$. Entonces A tiene inversa si y sólo si* $\det(A) \neq 0$.

COROLARIO 1.5.6. *Si A, B son matrices inversibles del mismo tamaño, entonces AB también es inversible (y lo mismo vale para BA pero cuidado que $BA \neq AB$).*

COROLARIO 1.5.7. *Un sistema homogéneo cuadrado (misma cantidad de ecuaciones que de incógnitas) es compatible determinado (tiene única solución, la solución trivial $X = \mathbb{O}$) si y solo si el determinante de la matriz del sistema es no nulo.*

Capítulo 2

Subespacios, autovalores y diagonalización

2.1. Bases de subespacios

• Corresponde a clases grabadas 2.1, 2.2, 2.3a, 2.3b, 2.4, 2.5

DEFINICIÓN 2.1.1 (Independencia lineal)**.** Dados $V_1, V_2, \ldots, V_k \in \mathbb{R}^n$ son *linealmente independientes* si la única manera de obtener una combinación lineal de ellos que da cero, es tomando todos los coeficientes 0. Es decir, si $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$ son tales que

$$\alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_k V_k = \mathbb{O}$$

es porque $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$. Abreviamos independencia lineal con l.i. Si los vectores *no* son l.i., decimos que son *linealmente dependientes*, abreviado l.d.

Si los vectores son l.d. siempre es posible despejar *alguno* de ellos en función de todos los demás. En la Figura 2.1 el conjunto $\{V_1, V_2, V_3\}$ es l.d. porque V_3 se puede escribir como combinación lineal de V_1, V_2 (está en el mismo plano que generan ellos dos). También es cierto en esa figura que V_2 se puede escribir como combinación lineal de V_1, V_3 y que V_1 es combinación lineal de V_2, V_3.

El conjunto $\{V_1, V_2, V_4\}$ es l.i. porque V_1, V_2 lo son y además V_4 no está

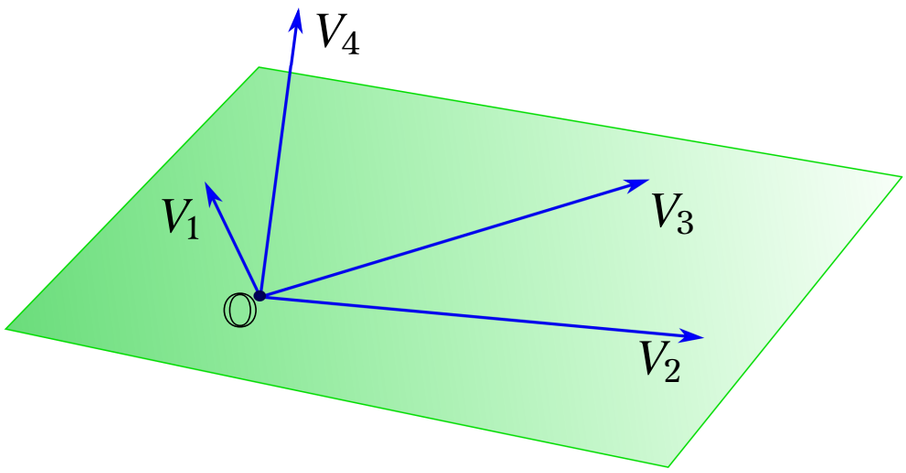


Figura 2.1: El conjunto $\{V_1, V_2, V_3\}$ es l.d., el conjunto $\{V_1, V_2, V_4\}$ es l.i.

en el plano generado por los dos primeros, luego no es combinación lineal de ellos.

OBSERVACIÓN 2.1.2. Los vectores $V_1, V_2, \ldots, V_r$ son l.d. si y sólo si existe una combinación lineal *no trivial* de ellos que da el vector nulo. Por no trivial queremos decir que alguno (o varios) de los coeficientes son distintos de cero.

ƨ Un vector $V \neq \mathbb{O}$ siempre es l.i. El vector nulo $V = \mathbb{O}$ es siempre l.d. Cualquier conjunto de vectores que incluya al vector nulo es l.d.

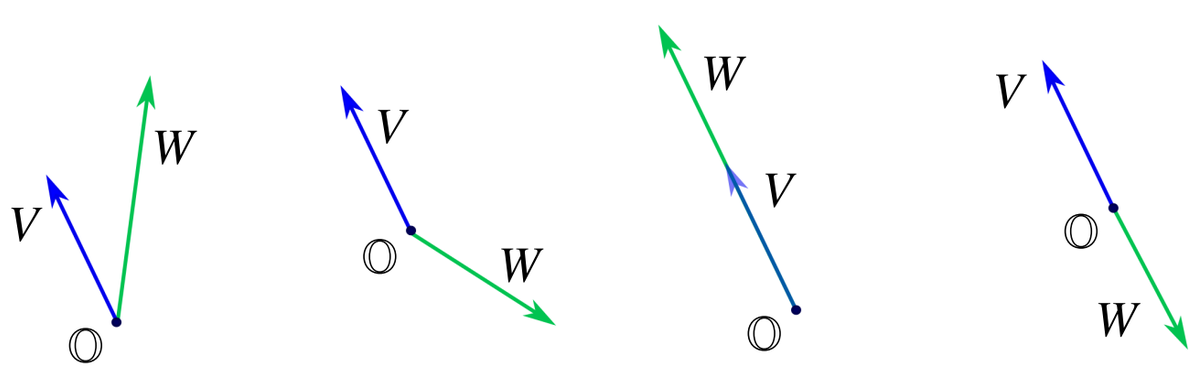


Figura 2.2: En los primeros dos casos $\{V, W\}$ es l.i., en los dos casos restantes es l.d.

ƨ Dados dos vectores V, W ambos no nulos, el conjunto es l.d. si y solo si los vectores están alineados, esto es, si $V = \alpha W$ (equivalentemente, si $W = \beta V$).

EJEMPLO 2.1.3. Si un conjunto de vectores es l.d., entonces se puede despejar *algún* vector en función de los demás, pero no es cierto que se

puedan despejar *todos* en función de los demás. Por ejemplo, consideremos los vectores $V_1 = (-1,2,0), V_2 = (2,-4,0), V_3 = (0,0,1)$. Este conjunto es l.d. porque tomando $\alpha_1 = -2, \alpha_2 = 1, \alpha_3 = 0$ tenemos

$$\begin{aligned} -2V_1 + V_2 + 0V_3 &= -2(-1,2,0) + (2,-4,0) + 0(0,0,1) \\ &= (2,-4,0) + (2,-4,0) + (0,0,0) = (0,0,0) = \mathbb{O}. \end{aligned}$$

Entonces podemos despejar V_1 en función de los demás:

$$V_1 = -\frac{1}{2}V_2 + 0V_3,$$

también se puede en este caso despejar V_2 en función de los demás de manera evidente. Pero *no* es posible despejar V_3 en función de V_1, V_2: ellos dos generan la misma recta por el origen $L : s(-1,2,0)$, $s \in \mathbb{R}$; el vector $V_3 = (0,0,1)$ *no* está en esa recta así que hallar a, b, s tales que

$$V_3 = aV_1 + bV_2 = s(-1,2,0)$$

es imposible.

DEFINICIÓN 2.1.4 (Base y dimensión)**.** Los $\{V_1, \ldots, V_r\}$ son *generadores* del subespacio S si $\forall\, W \in S$ existen $\alpha_1, \ldots, \alpha_r \in \mathbb{R}$ tales que

$$W = \alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_r V_r.$$

Una *base* de S es un conjunto de generadores linealmente independiente. En una base, el orden es importante: la base $\{V, W\}$ es distinta de la base $\{W, V\}$.
La *dimensión* del subespacio S es el número de vectores de una base cualquiera (este número *no* depende de la base elegida).

EJEMPLO 2.1.5 (Rectas y planos)**.** Una recta por el origen es un subespacio de dimensión 1: con un sólo vector no nulo de la recta lo generamos (un vector no nulo siempre es l.i.)

Un plano por el origen es un subespacio de dimensión 2: con dos vectores no alineados del plano podemos generarlo (si no están alineados son l.i.).

Ejemplo: en la Figura 2.1 podemos extraer varias bases del plano marcado: $B = \{V_1, V_2\}$, $B' = \{V_1, V_3\}$, $B'' = \{V_2, V_3\}$ (y también, por ejemplo, $B''' = \{V_2, V_1\}$ ya que el orden es relevante.

PROPIEDADES 2.1.6 (Ecuaciones *implícitas* y dimensión)**.** Sea $S \subset \mathbb{R}^n$ subespacio dado por ecucaciones implícitas: S es el conjunto de soluciones del sistema lineal homogéneo $AX = \mathbb{O}$. Con el método de Gauss, obtenemos un sistema equivalente *escalonado* $MX = \mathbb{O}$. Descartando las filas de ceros, supongamos que la cantidad de filas de M es j. Entonces:

• la dimensión del subespacio es $\dim(S) = n - j$, la dimensión del espacio total *menos* la cantidad de filas independientes (y no triviales) del sistema.

Atención que en este resultado se puede usar únicamente cuando tenemos las ecuaciones *impícitas* de S, que son entonces tantas como filas del sistema homogéneo que lo define. Es decir, si tenemos ecuaciones paramétricas, no tiene sentido usar este resultado.

EJEMPLO 2.1.7 (Planos y rectas)**.** Si $n = 2$, estamos en el plano $\mathbb{R}^2$

- 1 ecuación lineal homogénea no trivial

  $$ax + by = 0$$

  (que no sea la ecuación $0 = 0$) define una recta L:

  $$\dim(L) = 1 = n - j = 2 - 1$$

  donde 2 es la dimensión del espacio total y 1 es la cantidad de ecuaciones.

- 2 ecuaciones lineales homogéneas independientes (que no sea

una múltiplo de la otra), tiene única solución $S = \{\mathbb{O}\}$:

$$\dim(S) = 0 = n - j = 2 - 2.$$

Si $n = 3$, estamos en el espacio $\mathbb{R}^3$.

- 1 ecuación implícita (homogénea)

  $$ax + by + cz = 0$$

  no trivial define un plano Π por el origen:

  $$\dim(\Pi) = 2 = 3 - 1.$$

- 2 ecuaciones independientes: intersección de dos planos por el origen que no son paralelos, la solución es una recta L:

  $$\dim(L) = 1 = n - j = 3 - 2.$$

- 3 ecuaciones independientes: es un sistema compatible determinado, tiene única solución el origen (un punto). Su dimensión es $0 = n - j = 3 - 3$.

Si el sistema *no* está escalonado, el número de filas no dice nada: por ejemplo

$$S : \begin{cases} x + y - z &= 0 \\ 2x + 2y - 2z &= 0 \end{cases}$$

son dos ecuaciones implícitas en tres variables, pero puede el lector verificar que la solución es un plano (dimensión 2). Ocurre que la segunda fila es una múltiplo de la primera. Luego de escalonar, obtenemos *una sola ecuación* en $\mathbb{R}^3$, por ello la dimensión correcta es $2 = 3 - 1$.

PROPIEDADES 2.1.8 (Extraer una base)**.** Dado un conjunto de generadores $\{V_1, V_2, \ldots, V_r\}$ del subespacio S, para descartar los vectores redundantes y quedarnos con una base se puede escribir una matriz cuyas filas son los V_i, y escalonarla *sin intercambiar filas*. Las filas de

ceros que queden son los vectores que hay que descartar. Los que quedaron no nulos forman una base de S. También puede uno quedarse con los vectores *originales* de las filas que no se anularon y eso da (otra) base de S.

OBSERVACIÓN 2.1.9 (l.d. ó l.i.)**.** El método anterior también nos dice: si ninguna fila se anula, los vectores eran l.i., mientras que si se anulan una o más filas, eran l.d.

DEFINICIÓN 2.1.10 (Bases ortogonales y ortonormales)**.** Sea $S \subset \mathbb{R}^n$ subespacio (puede ser todo $\mathbb{R}^n$). Sea $B = \{V_1, \ldots, V_r\}$ base de S.
• Decimos que la base es *ortogonal* si tomados dos a dos, todos los vectores son perpendiculares entre si: $V_i \perp V_j$ para $i \neq j$.
• Si además todos los vectores tienen longitud 1, decimos que la base es *ortonormal*.

OBSERVACIÓN 2.1.11. Como el producto interno nulo describe la ortogonalidad, una base es ortogonal si y solo si

$$V_i \cdot V_j = 0 \qquad \forall i \neq j.$$

Por otro lado la base será ortonormal si además

$$1 = \|V_i\|^2 = V_i \cdot V_i \qquad \forall i.$$

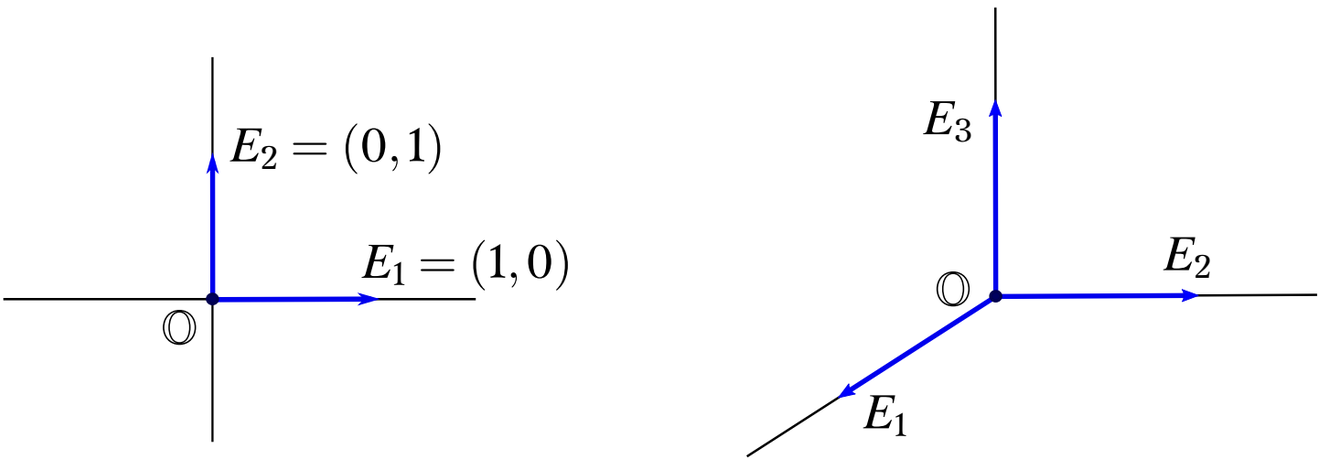


Figura 2.3: Bases canónicas de $\mathbb{R}^2$ y de $\mathbb{R}^3$

DEFINICIÓN 2.1.12 (Base canónica de $\mathbb{R}^n$)**.** Es la base ortonormal dada por los n vectores que tienen un 1 en algún lugar y todas las demás

entradas 0. Por ejemplo,

$$E = \{E_1, E_1\} = \{(1,0),(0,1)\}$$

es la base canónica de de $\mathbb{R}^2$. En el espacio $\mathbb{R}^3$ la base canónica es

$$E = \{E_1, E_2, E_3\} = \{(1,0,0),(0,1,0),(0,0,1)\}$$

2.2. Tranformaciones lineales y ortogonales

• Corresponde a clases en video 2.6, 2.7, 2.8

PROPIEDADES 2.2.1 (De la matriz traspuesta)**.** Si $A \in \mathbb{R}^{n\times n}$ entonces para todo par de vectores $V, W \in \mathbb{R}^n$ se verifica

$$(AV)\cdot W = V\cdot(A^tW) \qquad \text{ó equivalentemente} \qquad \langle AV, W\rangle = \langle V, A^tW\rangle$$

Esto puede deducirse de la propiedad de la traspuesta con el producto de matrices:

$$(AB)^t = B^tA^t$$

para todo par de matrices A, B para las que sea posible hacer el producto (esto es, $A \in \mathbb{R}^{k\times n}$, $B \in \mathbb{R}^{n\times d}$).

En general entonces $\langle AV, W\rangle \neq \langle V, AW\rangle$ salvo que A sea una matriz *simétrica* ($A^t = A$).

DEFINICIÓN 2.2.2 (Funciones, dominio, codominio, imagen, inyectividad)**.** Si X, Y son dos conjuntos cualesquiera, una función $f : X \to Y$ es una asignación que toma elementos de su *dominio* (que es un conjunto dentro de X denotado $Dom(f)$ en general) y les asigna un elemento $f(x)$ dentro del conjunto Y. El conjunto Y se denominia *codominio* de f, en general escribimos $y = f(x)$ o bien $x \mapsto f(x)$ para indicar esta asignación.
El conjunto de todos los $y = f(x)$ con $x \in Dom(f)$ se denomina *imagen* de f, lo denotamos $im(f)$. Cuando $im(f)$ es igual a todo Y decimos que

f es sobreyectiva (sinónimo=suryectiva). Entonces f no es sobreyectiva cuando la imagen es estrictamente más chica que el codominio.
Decimos que f es *inyectiva* si cada vez que $f(x_1) = f(x_2)$ es porque $x_1 = x_2$. Dicho de otra manera: en una función inyectiva, puntos distintos del dominio van a parar a puntos distintos del codominio.

EJEMPLO 2.2.3 (Función cuadrática)**.** Si $f : \mathbb{R} \to \mathbb{R}$ es la función elevar al cuadrado, la denotamos $f(x) = x^2$, su dominio y su codominio son todo $\mathbb{R}$.
Su imagen sin embargo, no es todo $\mathbb{R}$, ya que $x^2 \geq 0$ para todo $x \in \mathbb{R}$. Entonces $im(f) = [0, +\infty) \subset \mathbb{R}$ y la inclusión es estricta (o sea la imagen es más chica que el codominio). Luego f no es sobreyectiva.
Afirmamos que f tampoco es inyectiva: si tomamos $x_1 = 3, x_2 = -3$, es claro que son puntos distintos del dominio pero $f(x_1) = 3^2 = 9 = (-3)^2 = f(x_2)$, luego estos dos puntos van a parar al mismo $y = 9$.

DEFINICIÓN 2.2.4 (Matrices como funciones)**.** Dada una matriz $A \in \mathbb{R}^{k\times n}$, la función de $f : \mathbb{R}^n \to \mathbb{R}^k$ dada por $f(V) = AV$ (multiplicar contra A) se denomina *transformación lineal* asociada a la matriz A. Como función, tiene la propiedad

$$f(\alpha V + \beta W) = \alpha f(V) + \beta f(W) \qquad \forall V, W \in \mathbb{R}^n,\ \alpha, \beta \in \mathbb{R}$$

de linealidad, consecuencia de las propiedades del producto y la suma de matrices. En este caso $Dom(f) = \mathbb{R}^n$ mientras que el codominio de f es $\mathbb{R}^k$.

ƨ Aunque es un poco ambiguo, pensamos a veces a la matriz A como función y denotamos $A : \mathbb{R}^n \to \mathbb{R}^k$ a la misma. En ese caso decimos que el dominio de A es $\mathbb{R}^n$ y su codominio es $\mathbb{R}^k$.

DEFINICIÓN 2.2.5 (Núcleo y rango de una matriz)**.** Dada $A \in \mathbb{R}^{k\times n}$, el subconjunto del dominio dado por los ceros de A

$$\mathrm{Nu}(A) = \{V \in \mathbb{R}^k : AV = \mathbb{O}\}$$

se denomina *núcleo* de A. El vector nulo siempre está en $\mathrm{Nu}(A)$ porque $A\mathbb{O} = \mathbb{O}$. Decimos que el núcleo de A es *trivial* si $\mathrm{Nu}(A) = \{\mathbb{O}\}$.

El conjunto del codominio de A dado por

$$\mathrm{Ran}(A) = \{Z \in \mathbb{R}^n : Z = AX \text{ para algún } X \in \mathbb{R}^k\}$$

se denomina *rango* de A (a veces, *rango columna* de A) y es simplemente la imagen de A vista como función $A : \mathbb{R}^n \to \mathbb{R}^k$, es decir

$$\mathrm{Ran}(A) = im(A) = A(\mathbb{R}^k).$$

Las *columnas* de A son un conjunto de generadores de $\mathrm{Ran}(A)$ porque cada columna es AE_i para algún E_i de la base canónica.

OBSERVACIÓN 2.2.6 (Rango y núcleo son subespacios)**.** Si $V, W \in \mathrm{Nu}(A)$ entonces

$$A(sV + tW) = sAV + tAW = s\mathbb{O} + t\mathbb{O} = \mathbb{O}$$

luego $\mathrm{Nu}(A)$ es un subespacio de $\mathbb{R}^n$.
También $\mathrm{Ran}(A)$ es un subespacio de $\mathbb{R}^k$: si $Z_1, Z_2 \in \mathrm{Ran}(A)$ es porque existen X_1, X_2 tales que $Z_1 = AX_1$ y $Z_2 = AX_2$. Entonces

$$sZ_1 + tZ_2 = sAX_1 + tAX_2 = A(sX_1 + tX_2) = AX_0$$

así que llamando $X = sX_1 + tX_2$ se tiene $AX_0 = sZ_1 + tZ_2$ y así cualquier combinación lineal de Z_1, Z_2 está en $\mathrm{Ran}(A)$.

ƨ De hecho, habíamos definido subespacio como el conjunto de soluciones de un sistema homogéneo, y eso es lo mismo que el núcleo de la matriz asociada al sistema,

TEOREMA 2.2.7 (El rango y su dimensión)**.** *Para toda matriz A, la cantidad de filas l.i. de A es siempre igual a la cantidad de columnas l.i. de A. En particular la dimensión del rango por columnas (la imagen de A) es igual a la cantidad de filas l.i. o de columnas l.i. de A.*

Atención que este teorema sólo habla del tamaño (o sea la dimensión) del rango, y no del conjunto. Veamos en un ejemplo cómo interpretarlo bien y cómo uno puedo equivocarse, para evitar esa confusión: sea

$$A = \begin{pmatrix} 0 & 0 \\ 1 & 2 \end{pmatrix}.$$

En este caso vemos que A tiene una sola fila l.i. (también tiene una sola columna l.i porque la segunda columna es el doble de la primera), entonces la dimensión de la imagen de A (o lo que es lo mismo, la dimensión del rango de A) es exactamente 1 por el teorema anterior, es decir $\dim(Ran(A)) = 1$.

¿Es cierto que esa fila genera la imagen de A, o sea es cierto que $Ran(A) = \{(1,2)\}$? La respuesta categórica es *no*, es incorrecto, ese no es el subespacio imagen de A.

¿Cuál es el subespacio $Ran(A)$ entonces? Para hallarlo hay que mirar las columnas de A: como $AE_1 = (0,1)$ y $AE_2 = (0,2)$, estos dos son generadores del rango de A. Como se ve, son l.d., podemos tomar uno de ellos como base, entonces la respuesta correcta es

$$Ran(A) = \{(0,1)\}.$$

Veamos cómo se relacionan *los tamaños* de los subespacios rango y núcleo de una matriz:

TEOREMA 2.2.8 (de la dimensión para matrices)**.** *Si $A \in \mathbb{R}^{n\times n}$ entonces*

$$\dim(Ran(A)) + \dim(Nu(A)) = n.$$

OBSERVACIÓN 2.2.9 (Núcleo y determinante)**.** Si A es una matriz cuadrada, entonces por el teorema anterior A es sobreyectiva si y solo si el núcleo de A es trivial. Esto ocurre si y solo si A es inversible, y entonces podemos afirmar que para una matriz cuadrada

$$\mathrm{Nu}(A) = \{\mathbb{O}\} \quad \Longleftrightarrow \quad \det(A) \neq 0,$$

o equivalentemente

$$\mathrm{Nu}(A) \neq \{\mathbb{O}\} \iff \det(A) = 0.$$

DEFINICIÓN 2.2.10 (Matrices ortogonales)**.** Sea $U \in \mathbb{R}^{n\times n}$, decimos que U es una matriz *ortogonal* si $\langle UV, UW\rangle = \langle V, W\rangle$ para todo $V, W \in \mathbb{R}^n$. Son equivalentes

1. U es ortogonal
2. $\|UV\| = \|V\|$ para todo $V \in \mathbb{R}^n$
3. U es inversible y $U^{-1} = U^t$.

2 U es ortogonal si y solo si U^t es ortogonal.

OBSERVACIÓN 2.2.11 (Matrices y bases ortogonales)**.** Si

$$B = \{V_1, V_2, \ldots, V_n\}$$

es una base ortogonal de $\mathbb{R}^n$, y U es una matriz ortogonal, entonces

$$B' = \{UV_1, UV_2, \ldots, UV_n\}$$

también es ortogonal. Además si U es ortogonal los nuevos vectores tienen la misma longitud que los originales, luego: si B era base ortonormal entonces B' también lo es.

EJEMPLO 2.2.12 (rotaciones)**.** Si $n = 2$ tenemos la matriz con el parámetro $\theta \in [0, 2\pi]$,

$$U_\theta = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

Aplicada a cualquier vector lo transforma en uno rotado en un ángulo θ positivo (es decir, en contra del reloj).

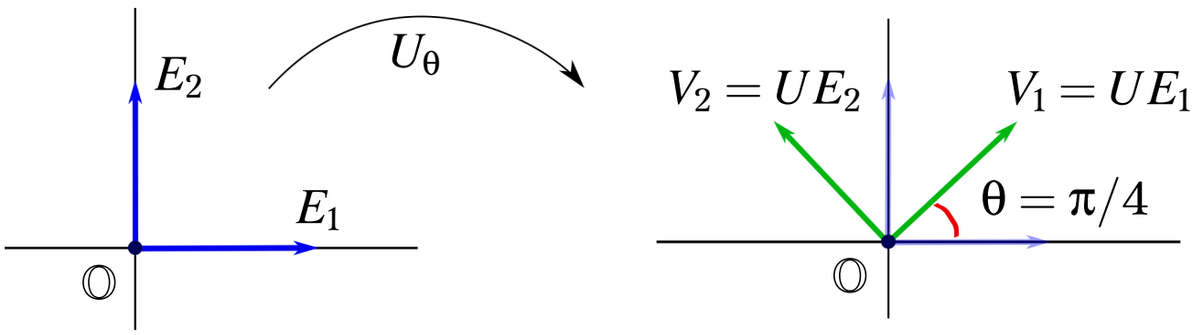


Figura 2.4: Rotación positiva (contra el reloj) de ángulo $\theta = \pi/4$

En particular tomando $\theta = \pi/4$ la base canónica se transforma en la base

$$B = \{V_1, V_2\} = \{(\sqrt{2}/2, \sqrt{2}/2)\,;\, (-\sqrt{2}/2, \sqrt{2}/2)\},$$

rotada en un ángulo recto en contra del reloj.

DEFINICIÓN 2.2.13 (Orientación en el plano)**.** Una base $\{V, W\}$ tiene *orientación positiva* si para ir de V hasta W (rotando hacia el lado más cercano) tenemos que movernos en contra del reloj (ver Figura 2.5). La base tiene *orientación negativa* si para llegar del primer vector al segundo hay que ir a favor del reloj (como la base $\{UV, UW\}$ de la misma figura).

PROPIEDADES 2.2.14 (Orientación)**.** Las matrices de rotación preservan la orientación (tienen determinante $\det(U) = 1$). Si queremos invertir la orientación podemos hacer una *reflexión* alrededor de una recta por el origen.

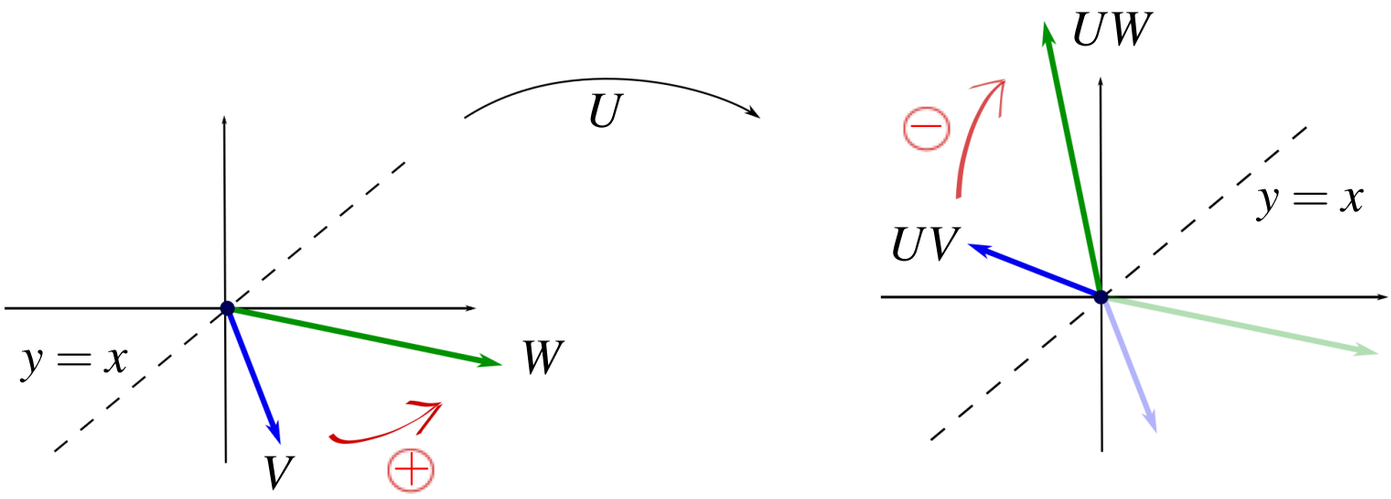


Figura 2.5: La reflexión U INVIERTE *la orientación de la base* $\{V, W\}$

Por ejemplo alrededor de la recta $y = x$ con la matriz

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

que es ortogonal (preserva longitudes y ángulos) pero no preserva la orientación (tiene $\det(U) = -1$): la base $\{V, W\}$ tiene orientación positiva (en contra del reloj) y la base $\{UV, UW\}$ tiene orientación negativa (a favor del reloj).

ƨ Puede probarse que toda matriz ortogonal es una rotación, o es una rotación seguida de una simetría como la del ejemplo anterior.

OBSERVACIÓN 2.2.15 (Rotaciones en $\mathbb{R}^3$)**.** Si fijamos un plano $\Pi \subset \mathbb{R}^3$ podemos hacer una rotación de ángulo θ alrededor de la normal N_π del plano. Por ejemplo si $\Pi : z = 0$ (cuya normal es $N = (0,0,1)$), la matriz

$$U_\theta^z = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

es una rotación alrededor del eje z. También

$$U_\theta^x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{pmatrix}, \qquad U_\theta^y = \begin{pmatrix} \cos\theta & 0 & -\operatorname{sen}\theta \\ 0 & 1 & 0 \\ \operatorname{sen}\theta & 0 & \cos\theta \end{pmatrix}$$

son matrices de rotación alrededor de los ejes x, y respectivamente.

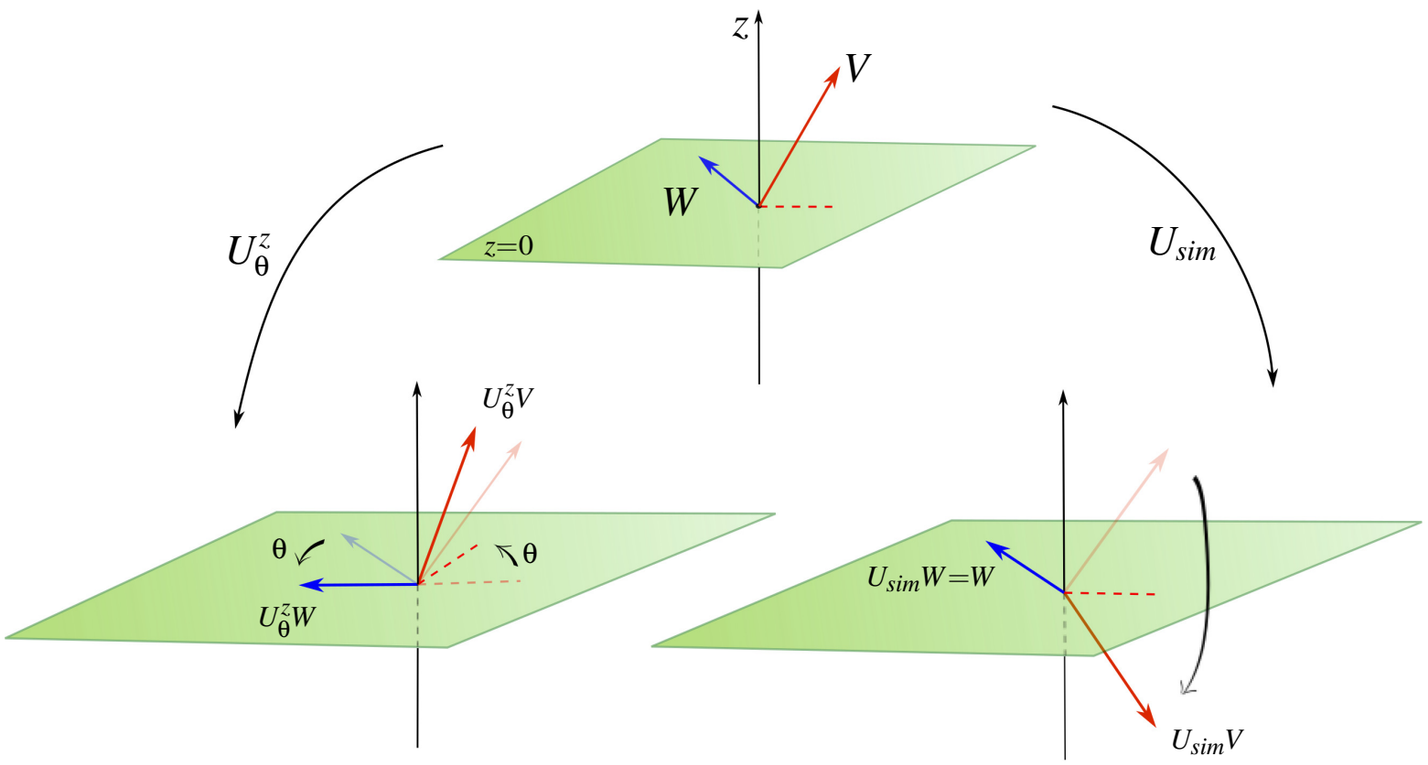


Figura 2.6: Una rotación y una reflexión

Si queremos invertir la orientación, podemos hacer una *reflexión* (o simetría) alrededor de un plano. Por ejemplo, alrededor del plano $\Pi : z = 0$, la matriz de simetría es

$$U_{sim} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad U_{sim} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

DEFINICIÓN 2.2.16 (Orientación en $\mathbb{R}^3$)**.** Sea $B = \{V_1, V_2, V_3\}$ base.

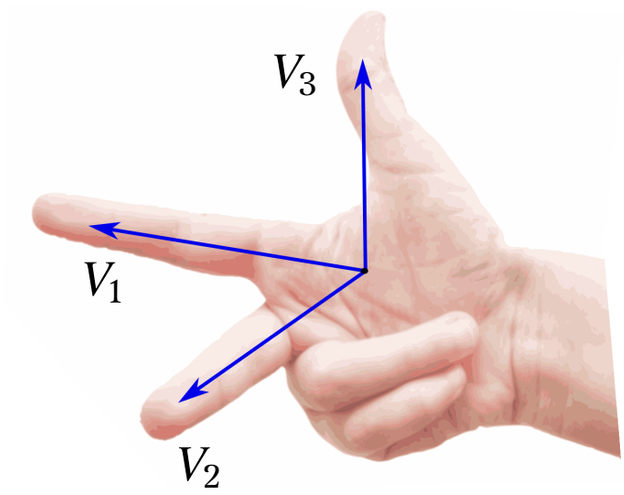


Figura 2.7: Regla de la mano derecha

Decimos que B está *orientada de forma positiva* si verifica la *regla de la mano derecha*: ubicando los vectores V_1, V_2 -en el orden dado- en los dedos índice y mayor de la mano derecha respectivamente, el vector V_3 debe apuntar en la dirección del dedo pulgar de la mano. La base canónica tiene orientación positiva.

ƨ Dados V, W l.i. en $\mathbb{R}^3$, la base $B = \{V, W, V \times W\}$ tiene orientación positiva, y la base $B' = \{V, W, -V \times W\}$ negativa.

ƨ En $\mathbb{R}^3$, puede probarse que toda matriz ortogonal es un producto de rotaciones y reflexiones como las de la Observación 2.2.15. Las rotaciones preservan la orientación (tienen $\det(U) = 1$), las reflexiones la invierten (tienen $\det(U) = -1$).

OBSERVACIÓN 2.2.17 (Bases y matrices ortogonales)**.** Si

$$B = \{V_1, V_2, \ldots, V_n\}$$

es una base ortonormal de $\mathbb{R}^n$, entonces la matriz

$$U = \begin{pmatrix} | & | & \ldots & | \\ V_1 & V_2 & \ldots & V_n \\ | & | & \ldots & | \end{pmatrix}$$

que tiene a los vectores como columnas es una matriz ortogonal. En

efecto

$$U^t U = \begin{pmatrix} - & V_1 & - \\ - & V_2 & - \\ \vdots & & \vdots \\ - & V_n & - \end{pmatrix} \begin{pmatrix} | & | & \dots & | \\ V_1 & V_2 & \dots & V_n \\ | & | & \dots & | \end{pmatrix} = \begin{pmatrix} V_1 \cdot V_1 & V_1 \cdot V_2 & \cdots & V_1 \cdot V_n \\ V_2 \cdot V_1 & V_2 \cdot V_2 & \cdots & V_2 \cdot V_n \\ \vdots & & & \vdots \\ V_n \cdot V_1 & V_n \cdot V_2 & \cdots & V_n \cdot V_n \end{pmatrix}$$

y como $V_i \cdot V_j = 0$ si $i \neq j$, mientras que $V_i \cdot V_i = 1$ para todo i (porque B es una base ortonormal), entonces

$$U^t U = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \mathbb{I}_n.$$

Esto dice que U^t es la matriz inversa de U. También U^t (que es la matriz con los V_i como filas) es una matriz ortogonal por el mismo motivo.

OBSERVACIÓN 2.2.18 (Matriz de cambio de base)**.** Si $B = \{V, W\}$ es una base de $\mathbb{R}^2$, escribimos $V = (v_1, v_2), W = (w_1, w_2)$ y tomamos C la matriz que tiene a sus vectores como *columna*, entonces C es inversible y

$$CE_1 = \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = V,$$

similarmente $CE_2 = W$.
Las mismas consideraciones aplican para base $B = \{V_1, V_2, \dots, V_n\}$ de $\mathbb{R}^n$, donde ahora la matriz $C \in \mathbb{R}^{n \times n}$: se tiene

$$CE_i = \begin{pmatrix} | & | & \dots & | \\ V_1 & V_2 & \dots & V_n \\ | & | & \dots & | \end{pmatrix} \begin{pmatrix} | \\ E_i \\ | \end{pmatrix} = V_i$$

para cada vector de la base canónica y cada V_i de la base B (en el orden dado de la base).

PROPIEDADES 2.2.19 (de la matriz como transformación lineal)**.** Dado un conjunto $\Omega \subset \mathbb{R}^n$ y una matriz inversible $C \in \mathbb{R}^{n\times n}$ el conjunto

$$\Omega' = C\Omega = \{CV : V \in \Omega\}$$

es el *transformado* por C. Por ejemplo

- Si $\Omega : \lambda V + P$ con $\lambda \in \mathbb{R}$ es una recta, entonces Ω' es otra recta:

$$C\Omega = \lambda CV + CP = \lambda V' + P'.$$

- Si $\Omega : sV + tW + P$ con $s,t \in \mathbb{R}$ es un plano, entonces

$$\Omega' = C\Omega = sCV + tCW + CP = sV' + tW' + P'$$

es también un plano.

- Si $C = U_\theta$ es una matriz de rotación, el objeto $\Omega' = C\Omega$ difiere de Ω en exactamente esa rotación (en otras palabras, obtenemos Ω' rotando Ω).

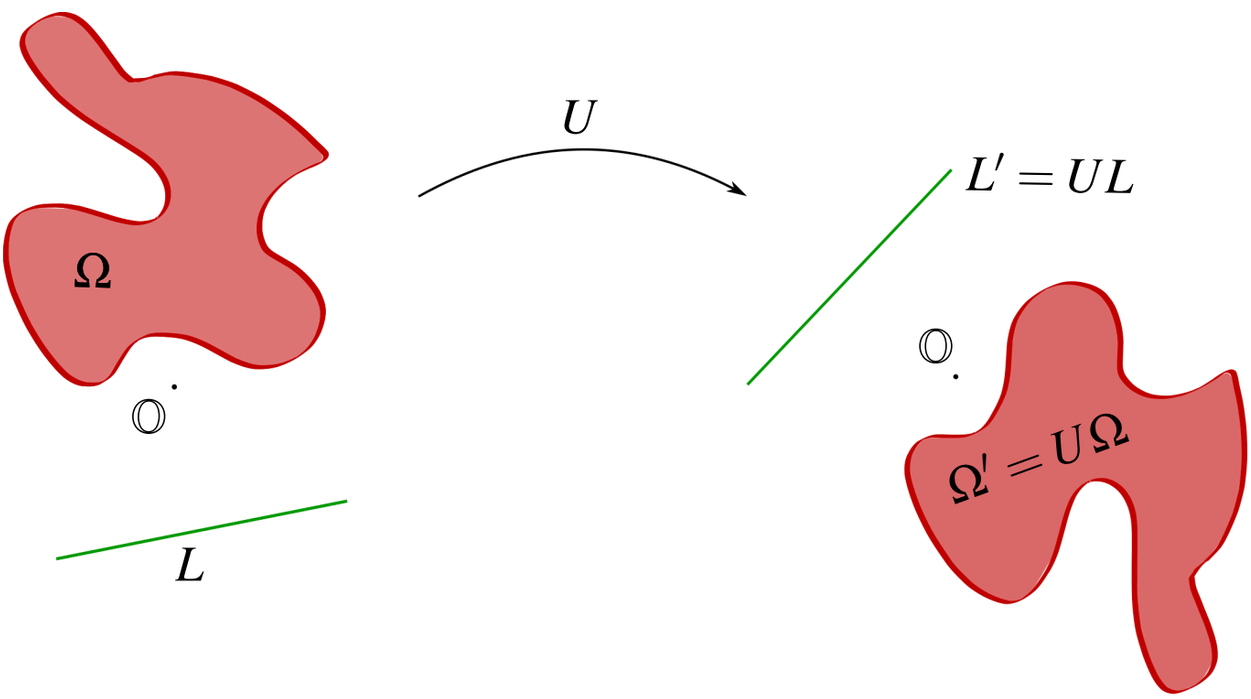


Figura 2.8: Movimiento rígido del objetos por una transformación ortogonal U

- Si U es una matriz ortogonal, el nuevo objeto $\Omega' = U\Omega$ sólo difiere del anterior en un *movimiento rígido* (el movimiento que hace la matriz ortogonal puede pensarse como una serie rotaciones y reflexiones).

2.3. Diagonalización

• Corresponde a clase en video 2.9

DEFINICIÓN 2.3.1 (Autovalores y autovectores)**.** Sea $A \in \mathbb{R}^{n\times n}$ decimos que $\lambda \in \mathbb{R}$ es un *autovalor* de A si existe $V \neq \mathbb{O}$ tal que $AV = \lambda V$. El vector V es un *autovector* de A.

Por ejemplo: si $A(1,-1) = (2,-2)$ entonces $V = (1,-1)$ es autovector de autovalor $\lambda = 2$ de la matriz A, porque $AV = 2V$.

OBSERVACIÓN 2.3.2 (Cálculo de autovalores y autovectores)**.** Se tiene $AV = \lambda V$ para algún $V \neq \mathbb{O}$ si y solo si $(A - \lambda\mathbb{I})V = \mathbb{O}$, luego $V \in \text{Nu}(A - \lambda\mathbb{I})$ es no nulo y el núcleo es no trivial. Como $A - \lambda\mathbb{I}$ es una matriz cuadrada, por la Observación 2.2.9 podemos concluir que

$$\lambda \in \mathbb{R} \text{ es autovalor de A} \iff \det(A - \lambda\mathbb{I}) = 0.$$

En ese caso, los autovectores son los elementos de $\text{Nu}(A - \lambda\mathbb{I})$, se obtienen resolviendo el sistema homogéneo $(A - \lambda\mathbb{I})X = \mathbb{O}$.

COROLARIO 2.3.3 (Autoespacios)**.** *Para cada autovalor $\lambda \in \mathbb{R}$ de la matriz A, el conjunto de autovectores es un subespacio no trivial de $\mathbb{R}^n$, denominado autoespacio del autovalor λ, lo denotamos*

$$E_\lambda = \{V \in \mathbb{R}^n : AV = \lambda V\} = Nu(A - \lambda\mathbb{I}).$$

DEFINICIÓN 2.3.4 (Polinomio característico)**.** Pensamos λ como incógnita real y notamos que $p(\lambda) = \det(A - \lambda\mathbb{I})$ es un polinomio en la variable λ, denominado *polinomio característico* de la matriz A.

A veces denotamos p_A para indicar que se trata del polinomio de la matriz A.

EJEMPLO 2.3.5 (de autovalores y autovectores)**.** Si $A = \begin{pmatrix} -10 & -6 \\ 18 & 11 \end{pmatrix}$

entonces su polinomio característico es

$$p(\lambda) = \begin{vmatrix} -10-\lambda & -6 \\ 18 & 11-\lambda \end{vmatrix} = (-10-\lambda)(11-\lambda)+18\cdot 6 = \lambda^2-\lambda-2.$$

Las raíces del polinomio son $\lambda_1 = -1$, $\lambda_2 = 2$. Para hallar los autovectores resolvemos primero el sistema homogéneo

$$(A-\lambda_1\mathbb{I})X = \mathbb{O}, \text{ es decir } (A+\mathbb{I})X = \mathbb{O},$$

cuya matriz ampliada es

$$\left(\begin{array}{cc|c} -10-\lambda_1 & -6 & 0 \\ 18 & 11-\lambda_1 & 0 \end{array}\right) = \left(\begin{array}{cc|c} -10-(-1) & -6 & 0 \\ 18 & 11-(-1) & 0 \end{array}\right) = \left(\begin{array}{cc|c} -9 & -6 & 0 \\ 18 & 12 & 0 \end{array}\right).$$

Como la segunda fila es (-2) por la primera, el sistema se reduce a $-9x-6y=0$. De aquí $6y=-9x$ o bien $y=-3/2x$. Entonces la solución es $(x,-3/2x)$ con x libre, o sea

$$E_{-1}: \{(1,-3/2)\} = \{(2,-3)\}.$$

El autoespacio E_{-1} correspondiente al autovalor $\lambda_1 = -1$ tiene dimensión 1. Cualquier vector no nulo de esta recta es autovector de este autovalor, por ejemplo $V_1 = (1,-3/2)$, o si uno prefiere no usar fracciones $V_1 = (2,-3)$.
Se resuelve de manera similar el sistema homogéneo $(A-\lambda_2\mathbb{I})X = \mathbb{O}$ (para $\lambda_2 = 2$) y se obtiene el autoespacio $E_2 = gen\{(1,-2)\}$, otro subespacio de dimensión 1. Podemos tomar $V_2 = (1,-2)$ como segundo autovector.

PROPIEDADES 2.3.6 (del polinomio característico)**.** Sea $A \in \mathbb{R}^{n\times n}$. Entonces:

1. p_A es un polinomio de grado n.

2. Las raíces de p_A son los autovalores de A.

En general un polinomio de grado n no tiene por qué tener n raíces reales distintas. Por ejemplo $p(\lambda) = \lambda^2 + 1$ no tiene ninguna raíz real, mientras que $p(\lambda) = (\lambda - 1)^2$ tiene una sola raíz real (doble).

DEFINICIÓN 2.3.7 (Matriz diagonal)**.** Una matriz cuadrada D es *diagonal* si todas las entradas fuera de la diagonal de D son nulas.

Cualquier múltiplo de la identidad es diagonal, pero como no es necesario que todas las entradas diagonales sean iguales, hay muchas más matrices diagonales.

DEFINICIÓN 2.3.8 (Matrices diagonalizables)**.** Una matriz cuadrada A es diagonalizable si existen una matriz diagonal del mismo tamaño D y una matriz inversible C tales que $A = CDC^{-1}$.

TEOREMA 2.3.9 (Bases de autovectores)**.** *$A \in \mathbb{R}^{n\times n}$ es diagonalizable si y sólo si existe una base $B = \{V_1, \ldots, V_n\}$ de $\mathbb{R}^n$ tal que todos los V_i son autovectores de A (y en ese caso la matriz D tiene a los autovalores de A en la diagonal). La matriz C que hay que tomar es la que tiene a los V_i como columnas.*

Una matriz diagonal D siempre es diagonalizable, basta tomar $C = \mathbb{I}$ (no es necesario que las entradas diagonales sean distintas).

EJEMPLO 2.3.10 (de matriz diagonalizable)**.** La matriz del Ejemplo 2.3.5 es diagonalizable. Como los autovalores son $\lambda = -1$ y $\lambda = 2$ tomamos

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

Como C hay que tomar la matriz que tiene a los autovectores como columnas, en el orden que pusimos los autovalores en D. Entonces, como pusimos primero el -1 hay que tomar como primer columna $V_1 = (2, -3)$, y como segunda columan $V_1 = (1, -2)$:

$$C = \begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix}.$$

Si calculamos la matriz inversa de C obtenemos $C^{-1} = \begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix}$. Entonces podemos verificar el teorema recién enunciado:

$$\begin{aligned} CDC^{-1} &= \begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 2 \\ 3 & -4 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -3 & -2 \end{pmatrix} = \begin{pmatrix} -10 & -6 \\ 18 & 11 \end{pmatrix} = A. \end{aligned}$$

2 La matriz $A = \begin{pmatrix} 3 & 0 \\ 1 & 3 \end{pmatrix}$ no es diagonalizable. Puede el lector verificar que A tiene solamente $\lambda = 3$ como autovalor, y que el autoespacio tiene dimensión 1, de hecho $E_3 = \{(0,1)\}$. Entonces no podemos armar una base de $\mathbb{R}^2$ con los autovectores de A, así que por el Teorema 2.3.9, A no es diagonalizable.

De acuerdo a los últimos dos ejemplos, *algunas* matrices que no son simétricas ($A^t \neq A$) pueden ser diagonalizables, pero otras no. En cambio, *todas* las matrices simétricas *siempre* son diagonalizables, y eso es lo que dice el próximo teorema:

TEOREMA 2.3.11 (Diagonalización de matrices simétricas)**.** *Si $A^t = A \in \mathbb{R}^{n\times n}$ es una matriz simétrica, entonces*

1. *p_A tiene n raíces reales (contadas con multiplicidad).*

2. *Autovectores correspondientes a autovalores distintos son ortogonales.*

3. *Existe una base ortonormal $B = \{V_1, \ldots, V_n\}$ de $\mathbb{R}^n$, donde todos los V_i son autovectores de A.*

4. *Se puede factorizar la matriz A como*

$$A = UDU^t$$

donde D es una matriz diagonal que tiene a los autovalores de A y U es la matriz ortogonal que tiene a la base B de autovectores como vectores columna.

Como la base es ortonormal, la matriz que tiene a los vectores de la base como columnas (o como filas) es ortogonal y así en lugar de calcular la inversa de U, podemos trasponerla -y es lo mismo que invertirla, porque para matrices ortogonales, $U^{-1} = U^t$-.

DEFINICIÓN 2.3.12 (Matrices definidas positivas)**.** Decimos que la matriz simétrica $A^t = A \in \mathbb{R}^{n\times n}$ es

1. *semi-definida positiva* si $\langle AV, V\rangle \geqslant 0$ para todo $V \in \mathbb{R}^n$.
2. *definida positiva* si $\langle AV, V\rangle > 0$ para todo $V \in \mathbb{R}^n \setminus \{\mathbb{O}\}$.
3. *semi-definida negativa* si $\langle AV, V\rangle \leqslant 0$ para todo $V \in \mathbb{R}^n$.
4. *definida negativa* si $\langle AV, V\rangle < 0$ para todo $V \in \mathbb{R}^n \setminus \{\mathbb{O}\}$.
5. *indefinida* si existen $V, W \in \mathbb{R}^n$ tales que

$$\langle AV, V\rangle > 0 \quad \langle AW, W\rangle < 0.$$

PROPIEDADES 2.3.13 (Autovalores vs. definida positiva)**.** Escribiendo cualquier $V \in \mathbb{R}^n$ como combinación lineal de la base B que diagonaliza $A = A^t$,

$$V = \alpha_1 V_1 + \alpha_2 V_2 + \cdots + \alpha_n V_n,$$

obtenemos la forma diagonal

$$\langle AV, V\rangle = \lambda_1\alpha_1^2 + \lambda_2\alpha_2^2 + \cdots + \lambda_n\alpha_n^2$$

donde los λ_i son los autovalores de A. Como los coeficientes de V están al cuadrado, se tiene que A es

1. semi-definida positiva si y solo si $\lambda_i \geq 0$ para todo i.
2. definida positiva si y solo si $\lambda_i > 0$ para todo i.

3. semi-definida negativa si y solo si $\lambda_i \leq 0$ para todo i.

4. definida negativa si y solo si $\lambda_i < 0$ para todo i.

5. indefinida si existen $\lambda_i > 0$, $\lambda_j < 0$.

EJEMPLO 2.3.14 (definidas e indefinidas en 2 variables)**.** Aquí $V = (x, y)$, damos un ejemplo de cada caso

- definida positiva: si $\lambda_1 = \lambda_2 = 1$ entonces $\langle AV, V \rangle = x^2 + y^2$.
- definida negativa $\lambda_1 = \lambda_2 = -1$ entonces $\langle AV, V \rangle = -x^2 - y^2$.
- indefinida $\lambda_1 = 1$, $\lambda_2 = -1$ entonces $\langle AV, V \rangle = x^2 - y^2$.
- semi-definida positiva: si $\lambda_1 = 1$, $\lambda_2 = 0$ entonces $\langle AV, V \rangle = x^2$.
- semi-definida negativa: si $\lambda_1 = -1$, $\lambda_2 = 0$ entonces $\langle AV, V \rangle = -x^2$.

Capítulo 3

Conjuntos y funciones en la recta y el espacio

3.1. Conjuntos en la recta y el espacio

• Corresponde a clases en video 3.1, 3.2, 3.3

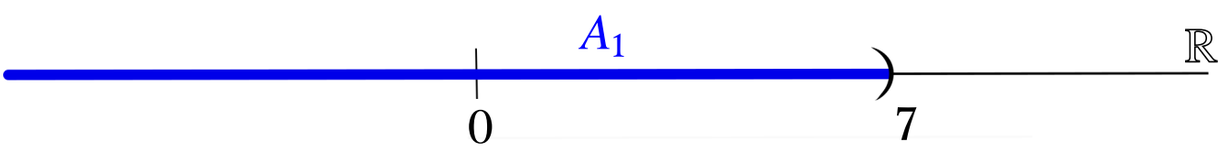
DEFINICIÓN 3.1.1 (Cotas de conjuntos de números reales)**.** Dado un conjunto $A \subset \mathbb{R}$, decimos que A está

• *acotado superiormente* si $\exists c \in \mathbb{R}$ tal que $x \leq c$ para todo $x \in A$. El número c es una *cota superior* de A.

• *acotado inferiormente* si $\exists d \in \mathbb{R}$ tal que $x \geq d$ para todo $x \in A$. El número d es una *cota inferior* de A.

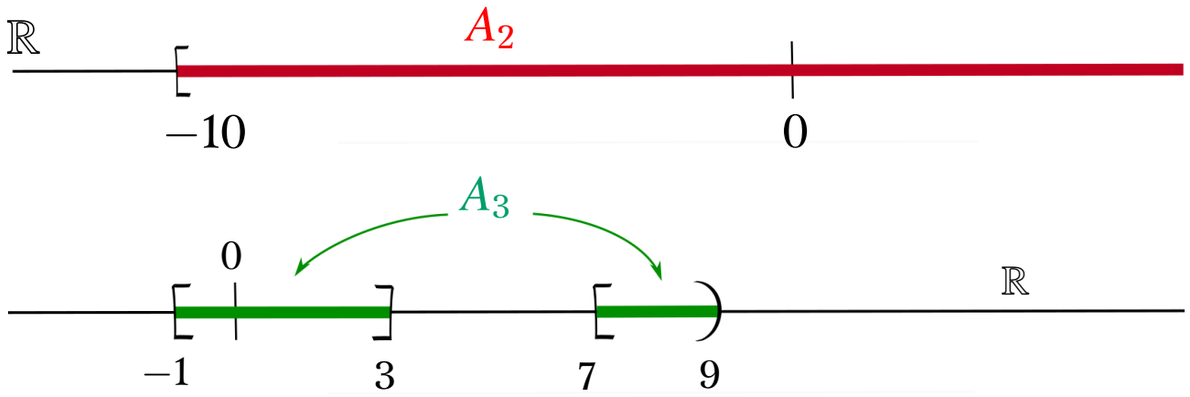
EJEMPLO 3.1.2 (conjuntos acotados y no acotados)**.** .

■ $A_1 = (-\infty, 7)$ está acotado superiormente por $c = 7$, pero también por $c = 7.1$, $c = 8$, $c = 125$, etc. Este conjunto no está acotado inferiormente.

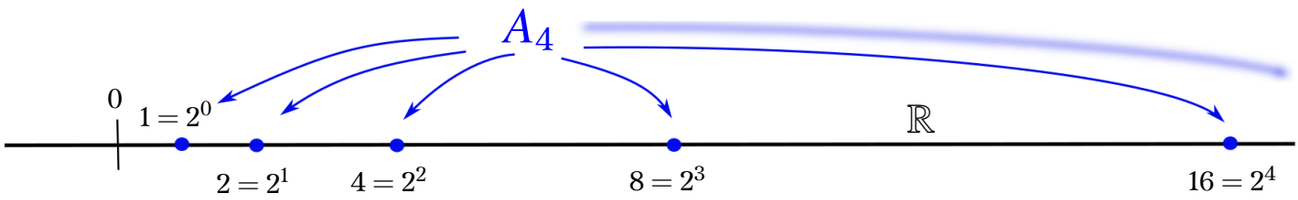


■ $A_2 = [-10, +\infty)$ no está acotado superiormente, pero si está acotado inferiormente. Cotas inferiores son $d = -10$ ó $d = -11$, etc.

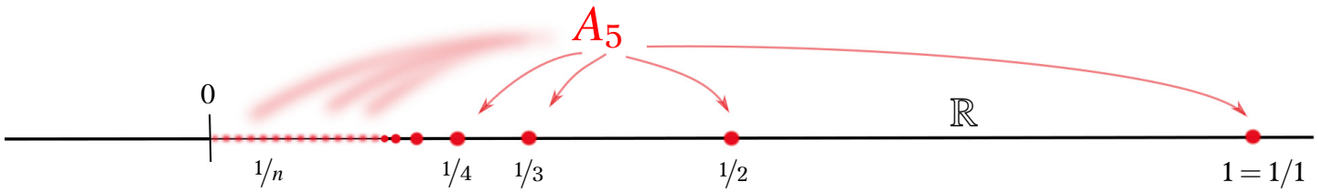
■ $A_3 = [-1, 3] \cup [7, 9)$ está acotado superiormente por $c = 9$ e inferiormente por $d = -1$. Los números entre intervalos (por ejemplo $x = 3$ ó $x = 4$) no son cotas del conjunto.



■ $A_4 = \{2^n : n \in \mathbb{N}_0\}$ no está acotado superiormente. Pero tiene cota inferior, por ejemplo $d = 0$ ó $d = 1$.



■ $A_5 = \{1/n : n \in \mathbb{N}\}$ está acotado tanto superior como inferiormente, $d = 0$ es cota inferior y $c = 1$ superior.



Como puede verse en los ejemplos, una vez hallada una cota superior, en realidad hay muchas (infinitas), porque cualquier número más grande sirve también de cota superior. Lo mismo vale para las cotas inferiores: una vez hallada una, cualquier número más *chico* también es cota inferior. Veamos como elegir una cota óptima

DEFINICIÓN 3.1.3 (Supremo e infimo)**.** Dado un conjunto $A \subset \mathbb{R}$, decimos que

- $s \in \mathbb{R}$ es el *supremo* de A si s es la mejor cota superior (la más pequeña). Esto es, s es el supremo de A (denotado $s = \sup(A)$) si s es cota superior de A y además $s \leq c$ para toda otra cota superior c de A.
- $i \in \mathbb{R}$ es el *infimo* de A (denotado $i = \inf(A)$) si es la mejor cota inferior (la más grande). Es decir, i es el ínfimo de A si es cota inferior y además $i \geq d$ para toda cota inferior d de A.

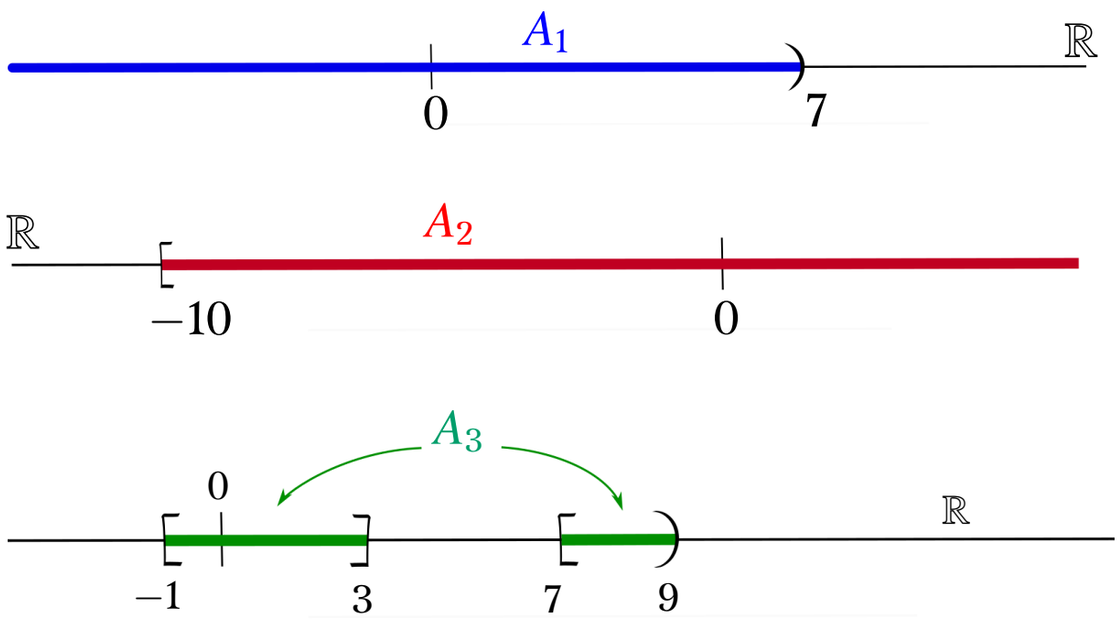
EJEMPLO 3.1.4 (supremos e infimos)**.** Conjuntos del Ejemplo 3.1.2:
- $\sup(A_1) = 7$, no tiene ínfimo porque no tiene cota inferior.
- $\inf(A_2) = -10$, no tiene supremo porque no tiene cota superior.
- $\inf(A_3) = -1$, $\sup(A_3) = 9$.
- $\inf(A_4) = 1$, no tiene supremo.
- $\inf(A_5) = 0$, $\sup(A_5) = 1$.

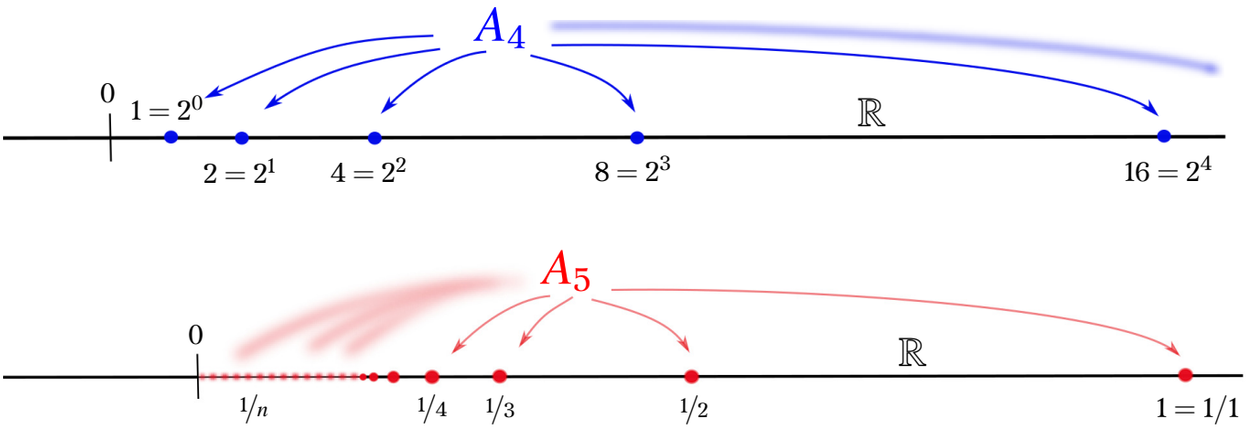
PROPIEDADES 3.1.5 (del supremo y el ínfimo)**.** Sea $A \subset \mathbb{R}$.

- Si A es acotado superiormente, tiene supremo.
- El supremo es único.
- El supremo no tiene por qué estar en el conjunto, cuando está, decimos que es el *máximo* del conjunto A, denotado $M = \text{máx}(A)$.
- Si A es acotado inferiormente, tiene ínfimo.
- El ínfimo es único.
- Si el ínfimo está en el conjunto A, decimos que es el *mínimo* de A, denotado $m = \min(A)$.

EJEMPLO 3.1.6 (máximos y mínimos de conjuntos del Ejemplo 3.1.2)**.**

.





- A_1 *no tiene* máximo porque $\sup(A_1) = 7$ no es un elemento de $A_1 = (-\infty, 7)$. Tampoco tiene mínimo porque no tiene ínfimo.
- $\min(A_2) = -10$ porque -10 es ínfimo y es un elemento del conjunto $A_2 = [-10, +\infty)$. El conjunto no tiene máximo porque no tiene supremo.
- $\min(A_3) = -1$. El conjunto no tiene máximo, porque $9 = \sup(A_3)$ no es un elemento de
- $\min(A_4) = 1$ (tomando $n = 0$ vemos que $1 = 2^0 \in A_4$). No tiene máximo porque no está acotado superiormente.
- $\max(A_5) = 1$ (tomando $n = 1$ vemos que $1 = 1/1 \in A_5$). No tiene mínimo porque $0 = \inf(A) \notin A_5$ (no hay ningún $n \in \mathbb{N}$ tal que $1/n = 0$).

DEFINICIÓN 3.1.7 (Módulo)**.** La función módulo toma un número y devuelve el mismo número (si era positivo ó 0) o devuelve el opuesto si el número era negativo. En cualquier caso, devuelve siempre un número mayor o igual a 0. La definición formal es como función partida

$$|x| = \begin{cases} x & si\ x \geq 0 \\ -x & si\ x < 0 \end{cases}$$

La función módulo es el análogo de la norma en dimensión 1, de hecho podemos decir sin equivocarnos que el módulo es la norma del espacio $\mathbb{R}$. Eso se ve bien si recordamos algunas de sus propiedades

PROPIEDADES 3.1.8 (de la función módulo)**.** Sean $x, y \in \mathbb{R}$. Entonces

- $|x| \geq 0$ y se tiene $|x| = 0$ únicamente en el caso que $x = 0$.
- $|yx| = |y||x|$
- $|x+y| \leq |x| + |y|$
- $|x| = \text{dist}(x, 0)$
- $|x-y| = \text{dist}(x, y)$.

OBSERVACIÓN 3.1.9 (Módulo como distancia)**.** Si tomamos $r \geq 0$, al escribir $|x| = r$ estamos indicando que x debe estar a distancia r del origen. Por ejemplo $|x| = 3$ indica que $x = 3$ o que $x = -3$. Entonces el conjunto

$$A = \{x \in \mathbb{R} : |x| \leq 3\} = \{x \in \mathbb{R} : \text{dist}(x, 0) \leq 3\}$$

consta de todos los puntos tales que su distancia al 0 es menor o igual que 3. Vemos que entonces se trata de un intervalo, $A = [-3, 3]$.

■ Esto a veces se menciona como la propiedad de "desdoblar" el módulo: pedir $|x| \leq 3$ es equivalente a pedir $-3 \leq x \leq 3$ y también es equivalente a pedir

$$\textit{simultáneamente} \begin{cases} x \geq -3 \\ x \leq 3 \end{cases}$$

donde la llave indica que pedimos que se cumplan *simultáneamente* las dos condiciones.

■ En general entonces, para $r \geq 0$ se tiene $|x| \leq r$ es equivalente a $-r \leq x \leq r$ y entonces es equivalente a pedir que se cumplan las dos desigualdades $x \geq -r$, $x \leq r$.

■ Por las mismas consideraciones, $|x| < r$ es equivalente a $-r < x < r$ (siempre que $r > 0$).

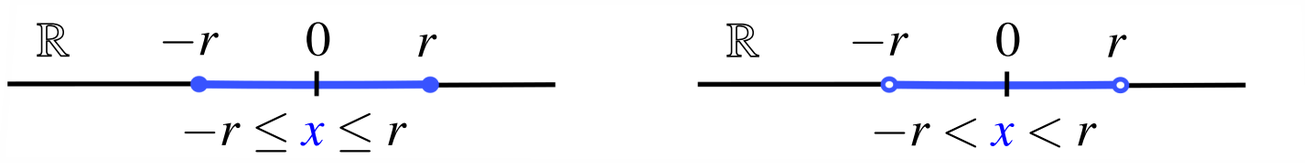


Figura 3.1: En azul, el conjunto $|x| \leq r$, y el conjunto $|x| < r$

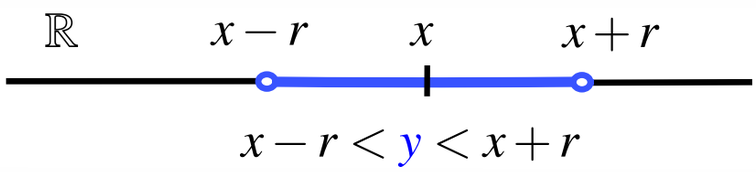
DEFINICIÓN 3.1.10 (Intervalos abiertos)**.** Un intervalo abierto es un conjunto de la forma

$$(a,b) = \{x \in \mathbb{R} : a < x < b\}$$

que consta de un segmento *sin* los extremos.
◇ Dado $x \in \mathbb{R}$, un intervalo abierto *alrededor* de x es de la forma

$$I = (x-r, x+r)$$

para algún $r > 0$. El número r se conoce como *radio* del intervalo, porque si tomamos un compás y lo pinchamos en x con apertura r, la intersección de la recta real con el interior del disco es el intervalo que nos interesa.



Observemos que este intervalo consta de los puntos de la recta real que están a distancia menor a r (del punto x):

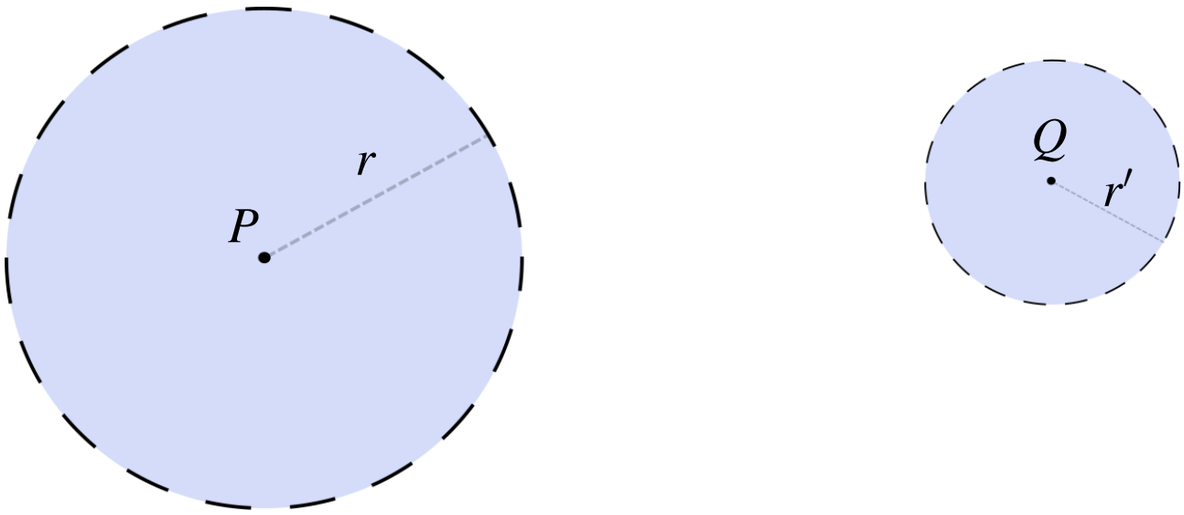
$$(x-r, x+r) = \{y \in \mathbb{R} : dist(y,x) < r\} = \{y \in \mathbb{R} : |y-x| < r\}.$$

Para tener en cuenta: en muchos textos se utilizan otras letras para el radio, por ejemplo la letra griega **d** que es δ y se lee "delta", o también la letra griega **e** que es ε y se lee "épsilon". Entonces por ejemplo con esta notación

$$I = \{x \in \mathbb{R} : |x+5| < \delta\} = (-5-\delta, -5+\delta)$$

es el intervalo abierto de centro $x = -5$ y radio $\delta > 0$.

Vamos a generalizar esta idea al espacio $\mathbb{R}^n$:

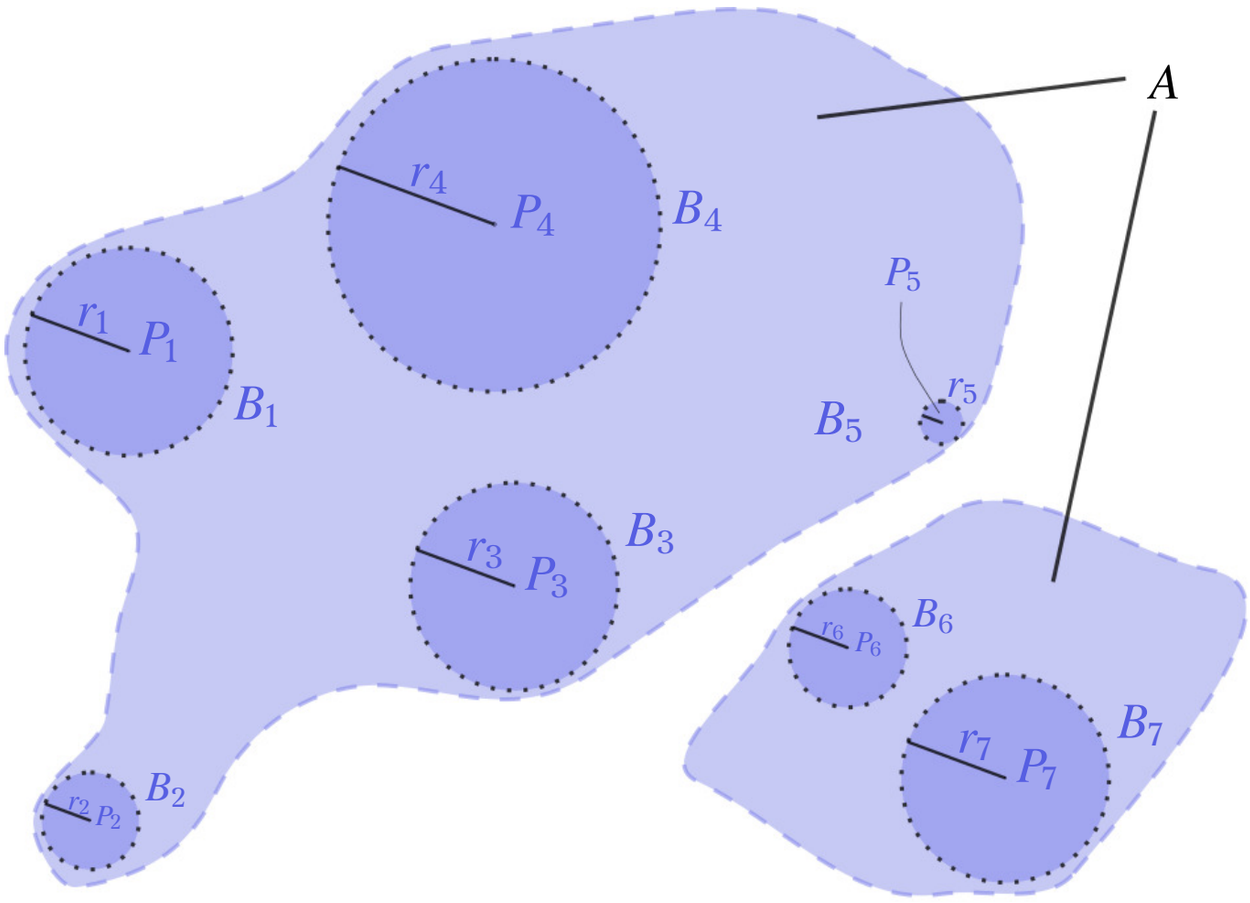


DEFINICIÓN 3.1.11 (Bolas y discos abiertos en $\mathbb{R}^n$)**.** Decimos que $B \subset \mathbb{R}^n$ es una *bola abierta de centro P y radio $r > 0$* si

$$B = B_r(P) = \{X \in \mathbb{R}^n : \|X - P\| < r\} = \{X \in \mathbb{R}^n : \text{dist}(X, P) < r\}.$$

Notemos que es una generalización de la noción de intervalo abierto alrededor de un punto: en la figura de arriba vemos bolas en el plano (son discos sin el borde).
En el caso de $\mathbb{R}^3$, la bola abierta es la esfera maciza, pero sin la cáscara.

DEFINICIÓN 3.1.12 (Puntos interiores - Conjuntos abiertos)**.** Sea $P \in A \subset \mathbb{R}^n$. Decimos que P es un *punto interior* del conjunto A si todos los puntos que lo rodean son también de A.
Más formalmente: existe $r > 0$ tal que la bola $B_r(P)$ centrada en P consta únicamente de elementos de A. Aún más formalmente: P es punto interior de A si existe $r > 0$ tal que $B_r(P) \subset A$.
El conjunto de los puntos interiores de A, se denota A°.
Siempre se tiene la inclusión $A^\circ \subset A$, porque por definición un punto interior es un punto (especial) de A.
Cuando $A = A^\circ$ decimos que A es un *conjunto abierto*.
Equivalentemente, *un conjunto es abierto si todos sus puntos son interiores*.
ƨ El radio depende del punto, para puntos muy cerca del borde del conjunto, hay que tomar un radio más pequeño.



DEFINICIÓN 3.1.13 (Complemento de un conjunto)**.** Dado $A \subset X$ (con A, X conjuntos), se denomina *complemento* de A al conjunto de todos los puntos de X que no están en A. El conjunto X a veces se denomina *conjunto universal*, quiere decir que es el contexto que hay que mirar fueraa de A, pero sólo en X. Se denota

$$A^c = \{x \in X : x \notin A\}.$$

Observemos que esto es lo mismo que decir: *para conseguir A^c tomamos el conjunto X y retiramos todos los puntos de A*. Por este motivo se puede usar la notación

$$A^c = X \setminus A,$$

donde el símbolo $\setminus$ se interpreta como un "menos", pero no en el sentido de restar números, sino en el sentido de restar=sacar.
Nuevamente: $A^c = X \setminus A$ se lee "X menos A" y se interpreta como "todos lo elementos de X, menos los elementos de A".

Algunos ejemplos simples:

■ $A_6 = \mathbb{R} \setminus \{2\}$ se interpreta como "todos lo números reales menos el número 2", se puede escribir alternativamente este conjunto como $A_6 = (-\infty, 2) \cup (2, +\infty)$.
■ $A_7 = \mathbb{R} \setminus \{-1, 4\}$ es "cualquier número real que no sea ni el -1 ni el 4". Se puede escribir como unión de intervalos: $A_7 = (-\infty, -1) \cup (-1, 4) \cup (4, +\infty)$.
■ $A_8 = \mathbb{R} \setminus [3, +\infty)$ consta de los números reales sin una semirrecta (del 3 en adelante). Podemos reescribirlo como $A_8 = (-\infty, 3)$. Como la semirrecta original contenía al 3, el conjunto A_8 *no* contiene al 3.
■ $A_9 = \mathbb{R} \setminus (-6, 2)$, vemos que $A_9 = (-\infty, -6] \cup [2, +\infty)$ con ambos números $-6, 2$ incluidos, puesto que lo que sacamos no los contenía.

ƨ Si el conjunto universal X (donde está A) no se explicita, hay que deducirlo del contexo. Así por ejemplo si nos dan un intervalo $A = [-2, 6)$ de números reales, hay que sobreentender que $X = \mathbb{R}$ y entonces

$$A^c = \mathbb{R} \setminus [-2, 6) = (-\infty, -2) \cup [6, +\infty)$$

En cambio si $A = \{\lambda V : \lambda \in \mathbb{R}\}$ con $V \in \mathbb{R}^n$, entendemos que A es una recta en $\mathbb{R}^n$ y entonces $X = \mathbb{R}^n$. Luego A^c será todo $\mathbb{R}^n$ menos esa recta.

DEFINICIÓN 3.1.14 (Frontera de un conjunto)**.** Dado un conjunto $A \subset \mathbb{R}^n$, decimos que $X \in \mathbb{R}$ es un punto de la *frontera* de A (o del *borde* de A) si *para toda bola abierta* $B = B_r(X)$ alrededor de X, podemos hallar

1. (al menos) un punto $P \in A$ (es decir $P \in A \cap B$)
2. (al menos) un punto $Q \notin A$ (es decir $Q \in A^c \cap B$).

El conjunto de puntos del borde de A se denota con $bd(A)$, aunque en algunos textos se denota como ∂A.

ƨ En la definición recién dada, es crucial que haya puntos de A y de A^c en *cualquier bola* alrededor del punto del borde X (no importa que tan pequeña sea la bola, siempre hay puntos de A y de A^c en la bola). Por ejemplo en la figura de arriba, $Q_1 \notin bd(A)$ porque la bola más pequeña no toca A^c.

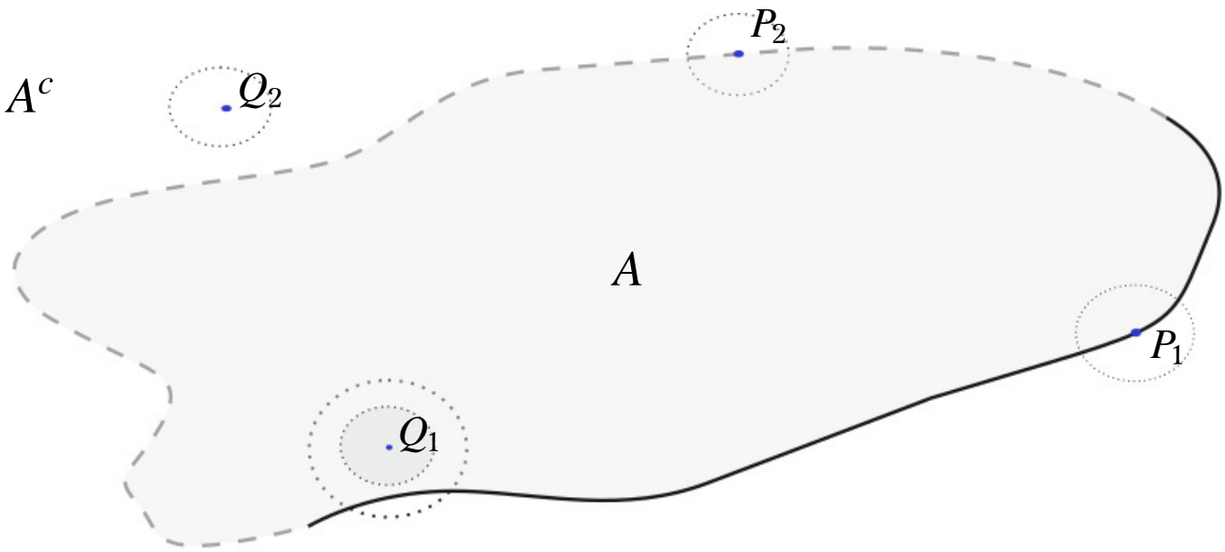


Figura 3.2: $P_1, P_2 \in bd(A)$, mientras que $Q_1, Q_2 \notin bd(A)$

Notemos que los puntos de la frontera de A pueden ser puntos de A, o puntos fuera de A. Por ejemplo en la figura de arriba, $P_1 \in A$, pero $P_2 \notin A$.

DEFINICIÓN 3.1.15 (Clausura de un conjunto)**.** Dado $A \subset \mathbb{R}^n$, el conjunto *clausura de* A (denotado $\overline{A}$) se obtiene tomando A y agregando todos los puntos del borde de A, es decir

$$\overline{A} = A \cup bd(A).$$

En general tenemos $A \subset \overline{A}$ por su misma definición, la clausura contiene al conjunto original.
Cuando vale $A = \overline{A}$ decimos que A es un *conjunto cerrado*.

Por ejemplo:

- Si $A = [-2,3)$, entonces $bd(A) = \{-2,3\}$ y entonces $\overline{A} = [-2,3]$.
- Si $A = (-\infty,1) \cup [3,10)$ entonces $\overline{A} = (\infty,1] \cup [3,10]$.
- Si $B = B_1(0)$ (la bola unitaria abierta), entonces $\overline{B} = \{X : \|X\| \leq 1\}$ (la bola unitaria cerrada).
- Si $C = \{1/3^n : n \in \mathbb{N}\}$, entonces $\overline{C} = C \cup \{0\}$.
- La clausura del conjunto de la Figura 3.2 se consigue agregando los puntos del borde faltante (como P_2):

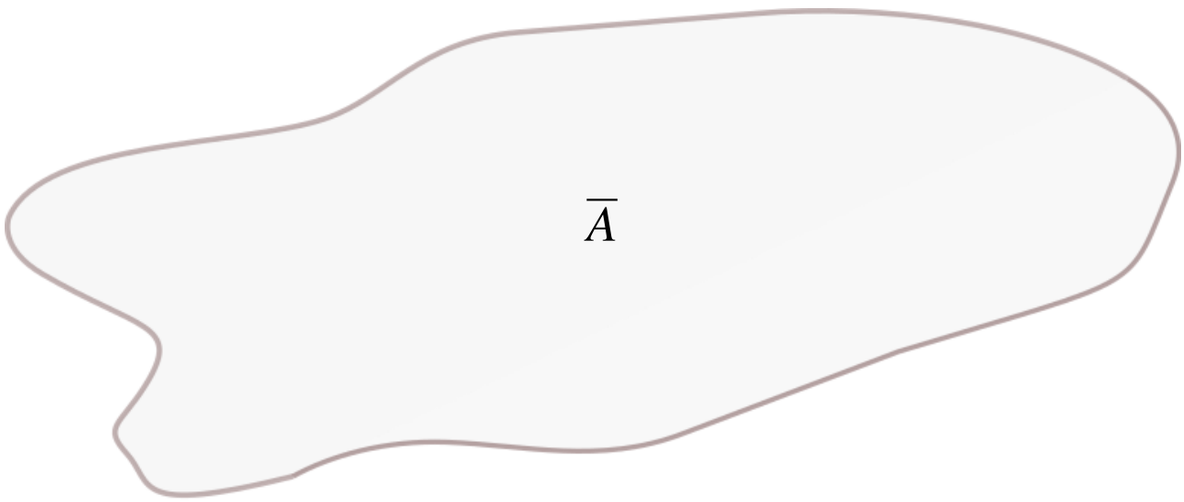


Figura 3.3: Clausura del conjunto A de la Figura 3.2

- Si $D = \{(x,y) : x > 0\}$ entonces $\overline{D} = \{(x,y) : x \geq 0\}$.

Algunas equivalencias útiles que relacionan estos conceptos (pensarlo sobre los dibujos que tenemos):

1. Un conjunto A es cerrado si y solo si $bd(A) \subset A$, es decir todos los puntos de la frontera son puntos de A.

2. Un conjunto es abierto si y solo si $bd(A) \subset A^c$, es decir todos los puntos de la frontera están fuera de A.

3. Un conjunto A es abierto si y solo si A^c es cerrado.

4. Un conjunto A es cerrado si y solo si A^c es abierto.

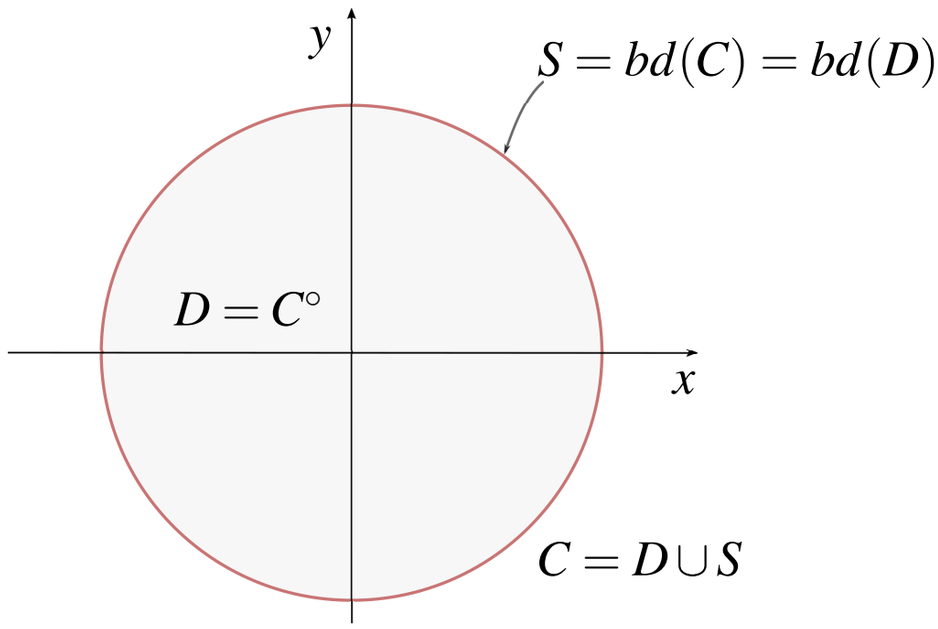
5. Hay conjuntos que no son ni abiertos ni cerrados, por ejemplo el intervalo $A = [3,7) \subset \mathbb{R}$, o el conjunto A de la Figura 3.2 de más arriba.

EJEMPLO 3.1.16 (Conjuntos dados por des/igualdades)**.** Veamos algunos ejemplos importantes:

1. $D = \{(x,y); x^2 + y^2 < 9\} \subset \mathbb{R}^2$ es un disco (de radio 3), es un conjunto abierto.

2. $C = \{(x,y) : x^2 + y^2 \leq 9\} \subset \mathbb{R}^2$ es un disco (de radio 3), en este caso es un conjunto cerrado. El interior de C es el disco abierto, o sea $C^\circ = D$.

3. Recíprocamente, la clausura del disco abierto D es el disco cerrado C, o sea $\overline{D} = C$.



4. Esto último es porque el borde del disco abierto es la circunferencia (de radio 3)

$$bd(C) = \{(x,y) : x^2 + y^2 = 9\} = S.$$

Esta circunferencia S también es un conjunto cerrado de $\mathbb{R}^2$.

5. La circunferencia S (la del item anterior o cualquier otra) no tiene interior, porque pegado a cualquier punto de S tenemos puntos que no están en S. Entonces $S^\circ = \emptyset$, tiene interior vacío.

6. $A = \{x : \ln(x) \leq 1\} \subset \mathbb{R}$ es un conjunto que no es abierto ni cerrado, esto es porque $A = (0, e]$ (por el dominio del logaritmo natural).
7. En muchos casos, los conjuntos definidos con $<$ son abiertos y los definidos con $\leq$ o con $=$ son cerrados, pero hay que tener cuidado porque el dominio de las fórmulas involucradas puede hacer que eso sea falso (como en el ejemplo anterior).

8. Si tomamos $B = \{(x,y,z) : x^2 + y^2 + z^2 < 16\}$ entonces B es una bola abierta (de radio 4) en $\mathbb{R}^3$. Si cambiamos el $<$ por $\leq$ tendremos la bola cerrada (se agrega la cáscara).

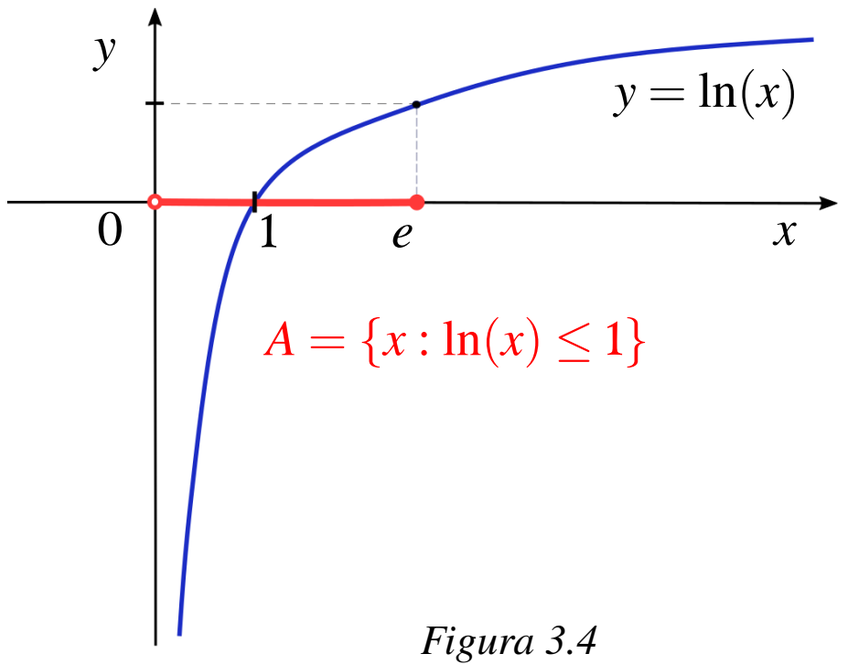


Figura 3.4

9. El conjunto $S^2 = \{x^2 + y^2 + z^2 = R^2\} \subset \mathbb{R}^3$ se denomina *esfera* de radio $R > 0$, es un conjunto cerrado. Esta cáscara S^2 no tiene interior pues es muy delgada, como no hay puntos interiores entonces tiene interior vacío.

10. Toda recta $L \subset \mathbb{R}^n$ es un subconjunto cerrado. Si $n \geq 2$, toda recta tiene interior vacío.

11. Todo plano $\Pi \subset \mathbb{R}^n$ es un subconjunto cerrado. Si $n \geq 3$, todo plano tiene interior vacío.

12. El conjunto $A = \{(x,y) : x \geq 0\}$ se denomina *semiplano*. Es un conjunto cerrado, y su interior es el semiplano abierto, es decir $A^\circ = \{(x,y) : x > 0\}$.

DEFINICIÓN 3.1.17 (Conjuntos acotados)**.** Un conjunto $A \subset \mathbb{R}^n$ es *acotado* si está acotado en norma, es decir si existe una constante $M > 0$ tal que $\|X\| \leq M$ para todo $X \in A$. En otras palabras, la distancia al origen de los puntos de A está acotada por alguna constante.

DEFINICIÓN 3.1.18 (Conjuntos compactos)**.** Un conjunto $A \subset \mathbb{R}^n$ es *compacto* si es cerrado y acotado.

EJEMPLO 3.1.19 (Conjuntos compactos y no compactos)**.** Veamos algunos casos típicos

1. Un intervalo $I \subset \mathbb{R}$ es compacto si y solo si tiene la forma $I = [a,b]$ con $a < b$ números reales.

2. El intervalo $(-\infty, 3]$ no es compacto pues no es acotado.

3. El intervalo $[4,7)$ no es compacto pues no es cerrado.

4. Una bola cerrada centrada en algún punto

$$B = \{X \in \mathbb{R}^n : \|X - P\| \leq R\}$$

es un conjunto compacto.

5. Si la bola es abierta no es compacta.

6. Un semiplano no es compacto (no es acotado).

7. Una recta $L \subset \mathbb{R}^n$ no es compacta porque no es acotada.

8. Un segmento de una recta

$$\vec{PQ} = \{tQ + (1-t)P : t \in [0,1]\} \subset \mathbb{R}^n$$

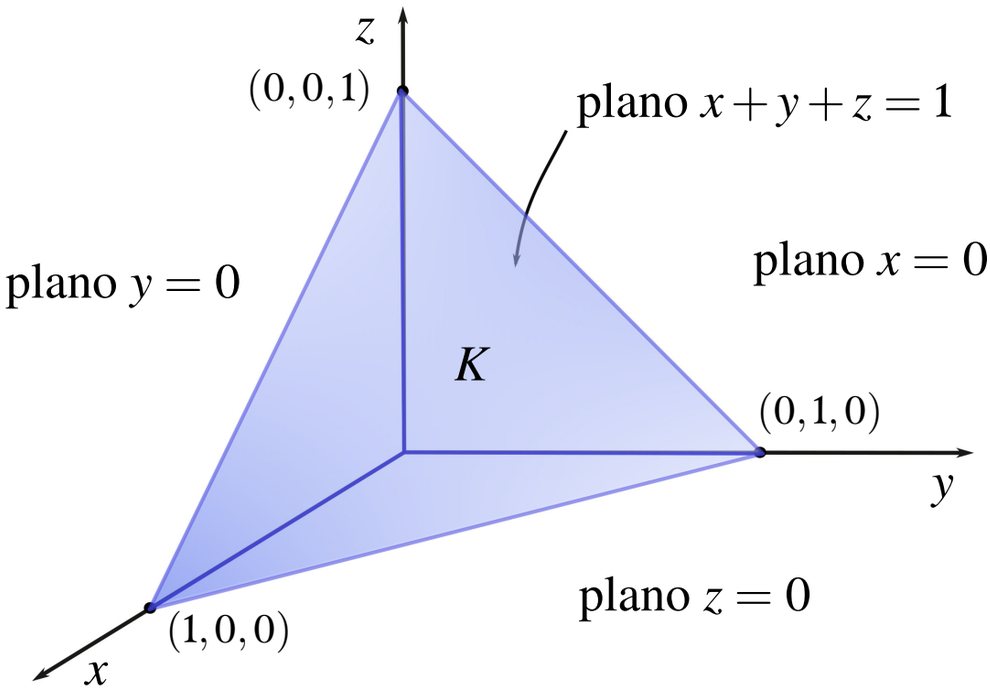
es un conjunto compacto.

9. Un plano nunca es compacto porque no es acotado.

10. El conjunto

$$K = \{(x,y,z) : x \geq 0, y \geq 0, z \geq 0, x+y+z \leq 1\}$$

es compacto, como puede verse al ilustralo (es un prisma triangular sólido):



3.2. Funciones, gráficos

• Corresponde a clases en video 3.4, 3.5a, 3.5b

DEFINICIÓN 3.2.1. Dada $f : X \to Y$ una función, su *gráfico* es el conjunto abstracto

$$Gr(f) = \{(x, f(x)) : x \in Dom(f)\} \subset X \times Y.$$

Es un subconjunto del producto cartesiano $X \times Y$.

OBSERVACIÓN 3.2.2 (Dibujos de gráficos)**.** Cuando $f : \mathbb{R} \to \mathbb{R}$, estamos acostumbrados a dibujar el conjunto $Gr(f)$ y llamar a ese dibujo "el gráfico de f".

Por ejemplo: si $f(x) = e^x$, dibujamos los pares ordenados (x, e^x) en $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, nos queda la curva exponencial. Notemos que los puntos que dibujamos son pares (x, y) pero con la condición $y = e^x = f(x)$.

Cuando hay restricciones de dominio, el dibujo sólo está "sobre" el dominio (o debajo si la imagen es negativa), por ejemplo en $f(x) = \ln(x)$, dibujamos los pares $(x, \ln(x))$ con $x > 0$, el dibujo es como en la Figura 3.4. Para $x \leq 0$ no hay dibujo (ni arriba ni debajo del eje x), porque esos x no están en el dominio del logaritmo.

OBSERVACIÓN 3.2.3 (Funciones escalares)**.** Decimos que f es una función escalar si $f:\mathbb{R}^n \to \mathbb{R}$, es decir cuando el codominio es $\mathbb{R}$. Para estas funciones, su gráfico es un subconjunto de $\mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$. Cuando en el dominio tenememos los número reales, los dibujos son los que dicutimos en el caso anterior.

Cuando $n = 2$ tenemos una función escalar de dos variables. Su gráfico es un subconjunto de $\mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3$, podemos intentar dibujarlo para darnos una idea de su forma. En general será una superficie con valles y montañas. Son puntos del espacio (x,y,z) que *obedecen a la ecuación implícita* $z = f(x,y)$.

1. El caso más sencillo es cuando nos queda una superficie plana horizontal. Este es el caso del gráfico de una función $f:\mathbb{R}^2 \to \mathbb{R}$ *constante*. Por ejemplo, si tomamos $f(x,y) = 5$, el gráfico son los puntos de la forma $(x,y,f(x,y)) = (x,y,5)$. O dicho de otra manera, aquellos (x,y,z) que verifican la condición $z = 5$. Estos son todos los puntos del plano horizontal $z = 5$.

2. En general si $f(x,y) = k$, con $k = cte$, su gráfico es el plano horizontal de altura k. El plano es exactamente el plano del piso cuando $k = 0$. Si $k > 0$ el plano está sobre el piso, si $k < 0$ el plano está debajo del piso.

3. El siguiente caso a considerar es el de un plano que no sea horizontal (pero que no sea vertical tampoco, porque sino no puede ser el gráfico de una función, como desarrollamos en el siguiente item). Por ejemplo si $f(x,y) = x + y$, los puntos del gráfico de f deben verificar $z = f(x,y)$ es decir $z = x + y$. Se trata del plano por el origen $x + y - z = 0$.

4. Si una de las variables está ausente en la fórmula de f, es que esa variable está libre. Por ejemplo, si $f(x,y) = y$, tenemos que mirar la ecuación $z = f(x,y)$ para hacer su gráfica, o sea $z = y$. Se trata de un plano a 45º respecto del piso, de ecuación $y - z = 0$ (dibujarlo).

5. Como mencionamos antes, un plano vertical -como por ejemplo la pared xz (de ecuación $\Pi : y = 0$)- no puede ser el gráfico de ninguna función $f = f(x,y)$.

   Esto es porque por ejemplo, sobre $(x,y) = (1,0)$ hay infinitos valores de z. Entonces $f(1,0)$ no estaría bien definido (recordemos que para ser función, tenemos que decir cuánto vale f en cada punto del dominio, esa imagen tiene que ser un solo punto).

6. Toda función lineal afín como $f(x,y) = ax + by + c$ tiene como gráfica un plano porque hay que mirar el conjunto de ecuación $z = ax + by + c$, que equivalentemente es la ecuación del plano $\Pi : ax + by - z = -c$.

DEFINICIÓN 3.2.4 (Funciones cuadráticas)**.** Una función de la forma

$$f(x,y) = ax^2 + by^2 + cxy + dx + ey + k$$

se conoce como *función cuadrática* en dos variables. Cuando $d = e = k = 0$ nos queda

$$f(x,y) = ax^2 + by^2 + cxy$$

y en este caso decimos que es una *forma cuadrática homogénea*. Para la definición vamos a pedir que f no sea nula, o sea que a, b ó c sean no nulos.
También se dice que f es *homogénea de grado* 2. Esto es porque

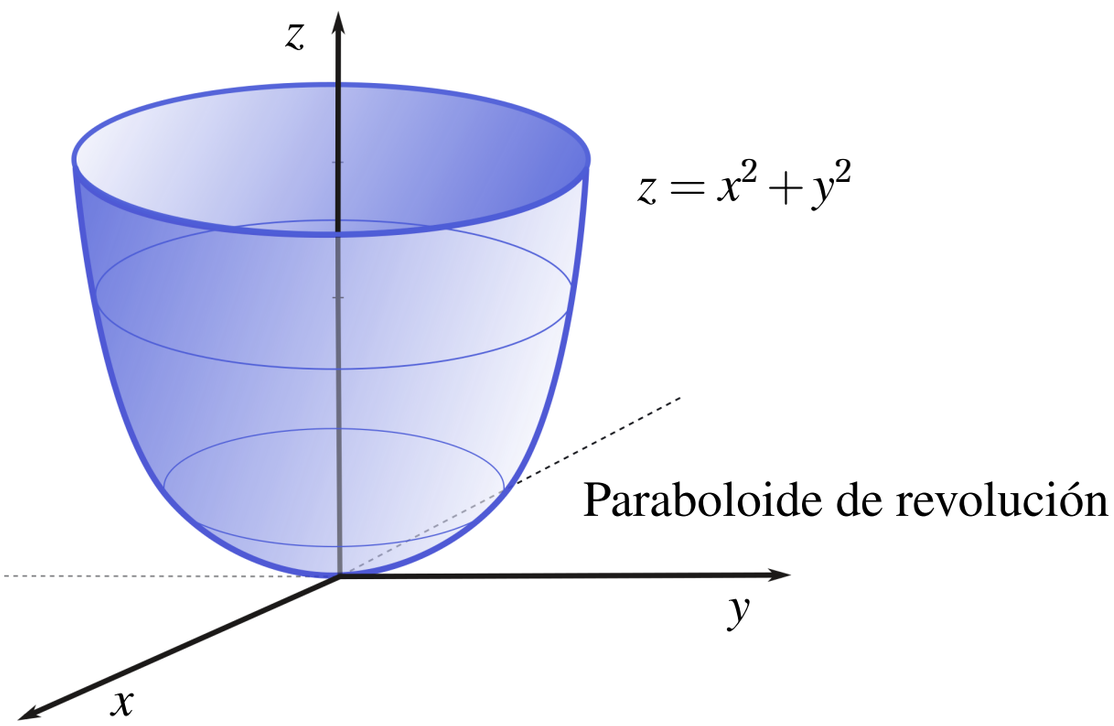
$$f(tx,ty) = a(tx)^2 + b(ty)^2 + c(txty) = at^2x^2 + bt^2y^2 + ct^2xy = t^2 f(x,y)$$

es decir, f saca escalares al cuadrado. Vectorialmente, $f(tV) = t^2 f(V)$ para todo $V \in \mathbb{R}^2$ y todo $t \in \mathbb{R}$.

Vamos a graficar (hacer el dibujo del gráfico de) funciones cuadráticas homogéneas. Los ejemplos más importantes son los que siguen: *paraboloide* y *silla de montar*. Luego tenemos el *cilindro parabólico*, que es un caso "degenerado" porque una de las variables está libre.

Podemos verificar que estas que mostramos a continuación son las gráficas usando un programa graficador como GeoGebra, como puede verse en este link.
Una explicación más detallada de por qué los gráficos tienen esta forma particular está más abajo en la Sección 3.3.

EJEMPLO 3.2.5 (Paraboloide de revolución)**.** $f(x,y) = x^2 + y^2$. Su gráfico es el conjunto en $\mathbb{R}^3$ dado por la ecuación implícita $z = x^2 + y^2$. Se conoce como *paraboloide de revolución*:

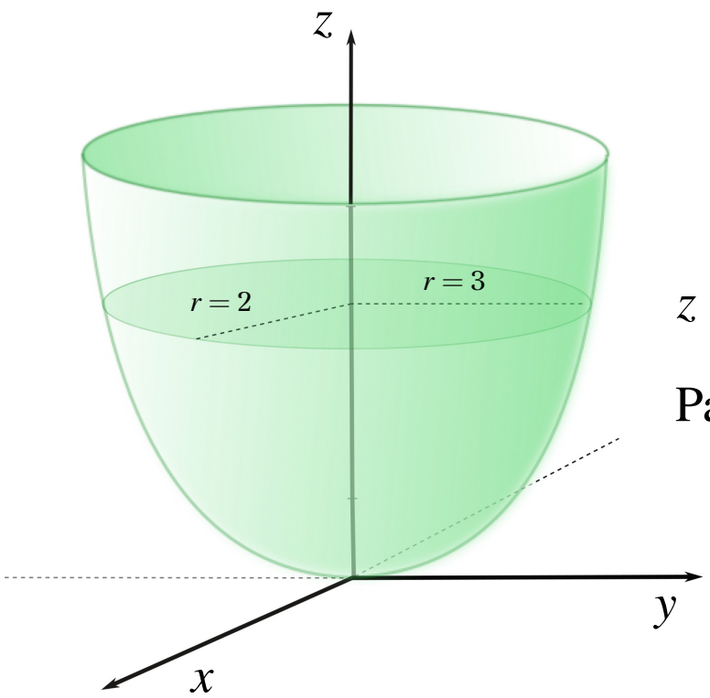


■ Los cortes horizontales de esta figura son circunferencias de ecuación $x^2 + y^2 = z_0$, y los cortes verticales devuelven una parábola como por ejemplo $z = y^2$ (tomando $x = 0$). Podemos obtener la gráfica de f girando la parábola alrededor del eje z (eso explica el nombre).

■ Si invertimos el signo y consideramos $f(x,y) = -x^2 - y^2$ la gráfica es un paraboloide pero invertido. Esto es porque este cambio obedece a fijar x, y y cambiar z por $-z$, que es una reflexión respecto del plano del piso (el plano $z = 0$).

■ Repetimos la observación del ítem previo en otros términos: aplicamos la transformación $U(x,y,z) = (x,y,-z)$ para pasar de la figura $z = x^2 + y^2$ a la figura $z = -x^2 - y^2$. La transformación U es una transformación ortogonal porque es una reflexión, podemos escribir su matriz si observamos que $UE_1 = E_1$, $UE_2 = E_2$, $UE_3 = -E_3$, luego

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$



$z = 4x^2 + 9y^2$

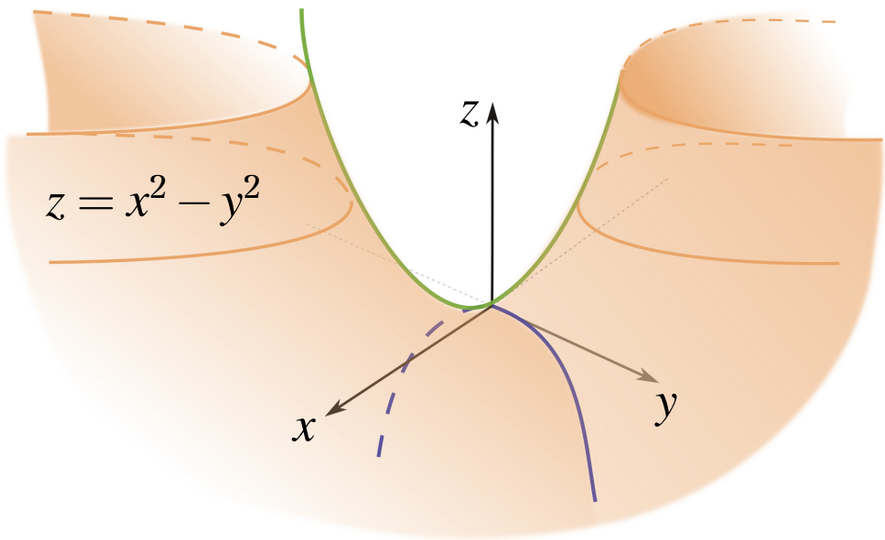
Paraboloide elíptico

■ Si x^2, y^2 están acompañados por constantes positivas el gráfico es similar al del paraboloide, pero está "achatado" en los ejes x, y. Por ejemplo $f(x,y) = 4x^2 + 9y^2$ tiene un gráfico similar al del paraboloide pero los cortes horizontales no son circunferencias, sino elipses (por ese motivo se conoce como *paraboloide elíptico*. En este caso el corte con el plano horizontal $z = 1$ nos da una elipse de radio $r = 2$ en la dirección del eje x, y de radio $r = 3$ en la dirección del eje y:

■ La misma observación pero para el paraboloide invertido, por ejemplo la gráfica de $f(x,y) = -4x^2 - 9y^2$ es similar al paraboloide invertido, pero achatado en las direcciones de los ejes x, y.

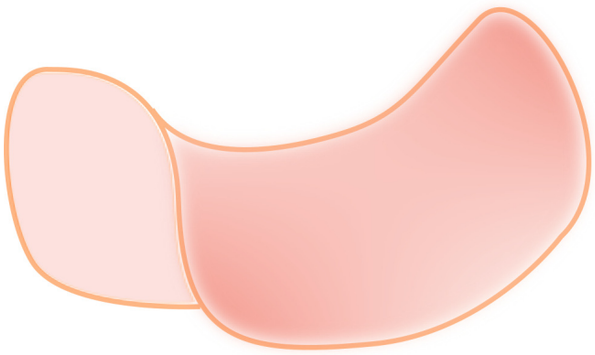
Ahora vamos a presentar un ejemplo sustancialmente distinto del paraboloide o sus variantes comprimida y/o invertida: una figura conocida como la silla de montar.

EJEMPLO 3.2.6 (Silla de montar)**.** $f(x,y) = x^2 - y^2$. Su gráfico es el conjunto en $\mathbb{R}^3$ dado por la ecuación implícita $z = x^2 - y^2$. Se conoce como *silla de montar*:
■ Si hacemos cortes horizontales, vemos hipérbolas de ecuación implícita $x^2 - y^2 = z_0$ (siempre que $z_0 \neq 0$, ver la Sección 3.3 debajo para más detalles).



■ Si hacemos el corte con el plano vertical $y = 0$ (la pared xz), vemos $z = x^2 - 0^2$, que es la parábola hacia arriba $z = x^2$.

■ Si hacemos el corte con el plano vertical $x = 0$ (la pared yz), vemos $z = 0^2 - y^2$, que es la parábola invertida $z = -y^2$.



$z = x^2 - y^2$

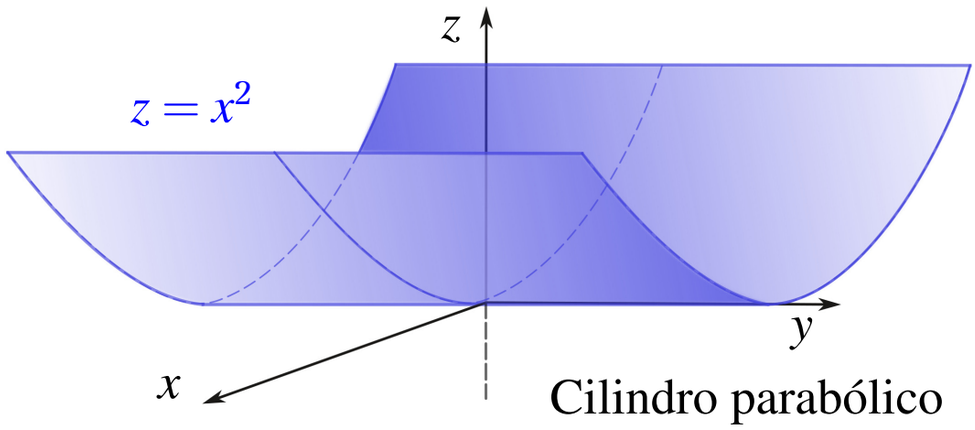
Silla de montar

Figura 3.5: En selectos ámbitos académicos se la conoce como la papa frita.

■ Si invertimos el signo en la silla de montar, $f(x,y) = -x^2 + y^2$, el dibujo es simétrico respecto del plano del piso. Esto es porque como ya explicamos, este cambio obedece a fijar x, y y cambiar z por $-z$, que es una reflexión respecto del plano del piso (el plano $z = 0$).

EJEMPLO 3.2.7 (Cilindro parabólico)**.** Consideremos $f(x,y) = x^2$. Notamos que la función no depende de y; eso quiere decir que la figura tiene la misma forma para todo y (sólo depende de x). El gráfico de f es la superficie $z = x^2$. Cortando esta superfice con planos verticales $y = cte$ (planos paralelos a la pared xz) vemos siempre la misma curva: la parábola $z = x^2$. La superficie se denomina *cilindro parabólico*, y se obtiene deslizando esta parábola hacia la izquierda y la derecha

indefinidamente:



Cilindro parabólico

Si invertimos el signo de la función f y tomamos $f(x,y) = -x^2$, su gráfico también será un cilindro parabólico, pero en este caso estará invertido respecto del plano $z = 0$, con el vértice de la parábola apuntando hacia arriba.

EJEMPLO 3.2.8. Si consideramos $f(x,y) = xy$, su gráfico también es un silla de montar. Esto es porque si hacemos el cambio de variables $x = u+v$, $y = u-v$, $z = z$, la ecuación implícita del gráfico $z = xy$ se transforma en

$$z = (u+v)(u-v) = u^2 - v^2$$

que es la misma ecuación del ejemplo anterior (con otras letras). Una explicación más detallada de este cambio de variables se da a continuación:

TEOREMA 3.2.9 (Forma canónica del gráfico de una función cuadrática)**.** *Sea $f : \mathbb{R}^2 \to \mathbb{R}$ una función cuadrática homogénea no nula. Luego de un cambio de coordenadas ortogonal $U \in \mathbb{R}^{2\times 2}$ en el dominio de f, y una dilatación/compresión en las direcciones de los ejes, el gráfico de f es alguno de estos cuatro:*

- $z = u^2 + v^2$ *(paraboloide)*
- $z = -u^2 - v^2$ *(paraboloide invertido)*
- $z = u^2 - v^2$ *(silla de montar)*

- $z = \pm u^2$ *(cilindro parabólico).*

En lugar de probar el teorema, vamos a mostrar cómo elegir el cambio de variables en dos ejemplos. La idea es escribir una matriz simétrica A que nos ayuda a encontra el cambio de variables, por medio de sus autovectores.

EJEMPLO 3.2.10. Sea $f(x,y) = 6x^2 + 9y^2 - 4xy$. Lo primero que vamos a hacer es encontrar una matriz simétrica $A^t = A$, es decir

$$A = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix},$$

de manera tal que si hacemos $\langle AX, X\rangle$ recuperamos $f(x,y)$. Afirmamos que la matriz que necesitamos es exactamente

$$A = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix},$$

vamos a verificarlo. Calculamos primero

$$AX = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 6x - 2y \\ -2x + 9y \end{pmatrix}.$$

Entonces, recordando que pensamos A como función de $\mathbb{R}^2$ en $\mathbb{R}^2$, tenemos $A(x,y) = (6x - 2y, -2x + 9y)$. Ahora calculamos el producto interno $\langle AX, X\rangle$:

$$\begin{aligned} AX \cdot X &= (6x - 2y, -2x + 9y) \cdot (x,y) = (6x - 2y)x + (-2x + 9y)y \\ &= 6x^2 - 2yx - 2xy + 9y^2 = 6x^2 + 9y^2 - 4xy = f(x,y) \end{aligned}$$

como habíamos afirmado (notar que los lugares de la diagonal son los coeficientes de x^2, y^2 respectivamente, pero que el lugar fuera de la diagonal es *la mitad* del coeficiente de xy, porque aparece dos veces). Ahora buscamos los autovalores y autovectores de A, por el teorema para matrices simétricas sabemos que hay una base ortonormal $B =$

$\{V_1, V_2\}$ de todo $\mathbb{R}^2$ tal que V_1, V_2 son autovectores de A. Esa base nos da el cambio de variables, o en otras palabras, poniendo esa base como columnas armamos una matriz ortogonal U que es la transformación (el cambio de variables) ortogonal que lleva nuestra gráfica de f a una de las 3 formas indicadas por el teorema.

Afirmamos que los autovalores de A son $\lambda = 10$, $\lambda = 5$ y que los correspondientes autoespacios son $E_{10} = \{(-1,2)\}$, $E_5 = \{(2,1)\}$, estas cuentas las puede verificar el lector. Los autoespacios son perpendiculares, pero falta normalizar los autovectores: tomamos

$$V_1 = (^{-1}/\sqrt{5}\,, ^{2}/\sqrt{5}), \quad V_2 = (^{2}/\sqrt{5}\,, ^{1}/\sqrt{5})$$

para obtener la base ortonormal $B = \{V_1, V_2\}$ y una transformación ortogonal que los tiene como columnas

$$U = \begin{pmatrix} | & | \\ V_1 & V_2 \\ | & | \end{pmatrix} = \begin{pmatrix} ^{-1}/\sqrt{5} & ^{2}/\sqrt{5} \\ ^{2}/\sqrt{5} & ^{1}/\sqrt{5} \end{pmatrix}.$$

Tomamos las nuevas variables $\tilde{X} = (u, v)$ de la siguiente manera: definimos $(u, v) = \tilde{X} = U^t X$, es decir

$$u = V_1 \cdot X = \frac{-1}{\sqrt{5}}x + \frac{2}{\sqrt{5}}y, \qquad v = V_2 \cdot X = \frac{2}{\sqrt{5}}x + \frac{1}{\sqrt{5}}y.$$

Escribimos la ecuación implícita del gráfico de f, que es $z = f(x,y)$ y recordamos que $A = UDU^t$ donde D es la matriz diagonal de los autovalores. Ahora observamos que, por la propiedad de la trasposición respecto del producto interno, se tiene:

$$\begin{aligned} z = f(x,y) &= AX \cdot X = (UDU^t)X \cdot X = (DU^t)X \cdot U^t X \\ &= D(U^t X) \cdot U^t X = D\tilde{X} \cdot \tilde{X}. \end{aligned}$$

Entonces, en las nuevas variables (u, v), la ecuación implícita de la superficie se escribe con una matriz diagonal D. Por eso, cuando desarrollemos esta última expresión $D\tilde{X} \cdot \tilde{X}$, veremos que no hay término

mixto, es decir no hay término con el producto uv:

$$\begin{aligned} z &= \begin{pmatrix} 10 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \cdot (u,v) \\ z &= (10u, 5v) \cdot (u,v) \\ z &= 10u^2 + 5v^2. \end{aligned}$$

Esto quiere decir que luego de la transformación ortogonal U (que es una rotación y/o una simetría, o una composición de ellas), la gráfica de $f(x,y) = 6x^2 + 9y^2 - 4xy$ coincide con el conjunto $z = 10u^2 + 5v^2$. Ahora vamos a usar que los coeficientes los podemos escribir como

$$10 = \sqrt{10}^2, \qquad 5 = \sqrt{5}^2,$$

y entonces si volvemos a cambiar las variables por $x_1 = \sqrt{10}u$, $x_2 = \sqrt{5}v$ vemos que

$$z = 10u^2 + 5v^2 = (\sqrt{10}u)^2 + (\sqrt{5}v)^2 = x_1^2 + x_2^2.$$

Es decir, luego de esta última transformación (que se puede pensar como una dilatatación o compresión en la dirección de cada uno de los dos ejes), vemos que la gráfica de f coincide con el conjunto

$$z = x_1^2 + x_2^2$$

que es un paraboloide de revolución.

EJEMPLO 3.2.11. Sea $f(x,y) = xy$, veamos que luego de una transformación ortogonal su gráfico coincide con el de la silla de montar $z = u^2 - v^2$ (como ya comentamos más arriba en el Ejemplo 3.2.8). Afirmamos que la matriz que necesitamos para obtener $f(x) = AX \cdot X$ es

$$A = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}$$

ya que no hay términos con x^2 ó y^2. Lo comprobamos: $A(x,y) = 1/2(y,x)$

luego

$$\langle A(x,y);(x,y)\rangle = 1/2\langle (y,x);(x,y)\rangle = 1/2(yx+xy) = xy = f(x,y).$$

Los autovalores de A son $\lambda = 1/2$, $\lambda = \lambda = -1/2$ y los correspondientes autoespacios son $E_{1/2} = \{(1,1)\}$, $E_{-1/2} = \{(1,-1)\}$, estas cuentas las puede verificar el lector. Los autoespacios son perpendiculares, pero falta normalizar los autovectores: tomamos

$$V_1 = (1/\sqrt{2}\,, 1/\sqrt{2}), \quad V_2 = (1/\sqrt{2}\,, -1/\sqrt{2})$$

para obtener la base ortonormal $B = \{V_1, V_2\}$ y una transformación ortogonal que los tiene como columnas

$$U = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix}.$$

Tomamos las nuevas variables $\tilde{X} = (u,v)$ de la siguiente manera: definimos $(u,v) = \tilde{X} = U^t X$, es decir

$$u = V_1 \cdot X = \frac{1}{\sqrt{2}}x + \frac{1}{\sqrt{2}}y, \qquad v = V_2 \cdot X = \frac{1}{\sqrt{2}}x + \frac{-1}{\sqrt{2}}y.$$

Comentario: notemos que salvo por el factor constante $1/\sqrt{2}$ *(que está puesto para normalizar), este cambio de variables es esencialmente* $u = x+y$, $v = x-y$. *Si uno despeja* x,y *en función de* u,v *obtiene* $x = \sqrt{2}/2(u+v)$ *y también* $u = \sqrt{2}/2(u-v)$, *que salvo el factor constante es el cambio de variable que propusimos en el Ejemplo 3.2.8. La diferencia es que aquel, más simple, no es ortogonal porque "aplasta" un poco y en cambio este si lo es.*

Siguiendo con el razonamiento para $f(x,y) = AX \cdot X$, tenemos nuevamente que $A = UDU^t$ donde D es la matriz diagonal de los autovalores de A, y U la matriz ortogonal que tiene a los autovectores como colum-

na. Nuevamente tenemos

$$\begin{aligned} z = f(x,y) = AX \cdot X = (UDU^t)X \cdot X = (DU^t)X \cdot U^tX \\ = D(U^tX) \cdot U^tX = D\tilde{X} \cdot \tilde{X} \end{aligned}$$

Es decir, en las nuevas variables, la ecuación implícita se escribe usando una matriz diagonal. Por eso no hay términos cruzados (los términos con el producto uv no aparecen), como se ve a continuación si escribimos explícitamente $D\tilde{X} \cdot \tilde{X}$:

$$\begin{aligned} z &= \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \cdot (u,v) \\ z &= (1/2u, -1/2v) \cdot (u,v) \\ z &= 1/2u^2 - 1/2v^2. \end{aligned}$$

Esta superficie ya es una silla de montar, pero algo comprimida en las direcciones de los ejes. Para convencernos, podemos hacer un segundo cambio de variables, llamando $x_1 = u/\sqrt{2}$, $x_2 = v/\sqrt{2}$ y obtenemos

$$z = \frac{1}{2}u^2 - \frac{1}{2}v^2 = (\frac{1}{\sqrt{2}}u)^2 - (\frac{1}{\sqrt{2}}v)^2 = x_1^2 - x_2^2.$$

Vemos que luego de comprimir un poco en la dirección de los ejes u, v, la figura que obtuvimos al rotar el gráfico de $f(x,y) = xy$ es la superficie dada de forma implícita por la ecuación $z = x_1^2 - x_2^2$, que es la silla de montar.

A la vista de los ejemplos, podemos resumir aún más el resultado del Teorema 3.2.9, pues se ve que los coeficientes de u^2, v^2, en la expresión nueva para el gráfico de f, son los autovalores de la matriz A con la que construimos f:

TEOREMA 3.2.12. *Sea $f : \mathbb{R}^2 \to \mathbb{R}$ cuadrática homogénea, sea $A^t = A \in \mathbb{R}^{2\times 2}$ tal que $f(X) = AX \cdot X$, donde $X = (x,y)$ son las variables. Si $\lambda_1, \lambda_2 \in \mathbb{R}$ son los autovalores de A, entonces (luego de una transformación ortogonal en el dominio) el gráfico de f es*

- *Un paraboloide (cuando* $\lambda_1, \lambda_2 > 0$*)*
- *Un paraboloide invertido (cuando* $\lambda_1, \lambda_2 < 0$*)*
- *Una silla de montar (cuando* $\lambda_1 > 0$ *y* $\lambda_2 < 0$*, o al revés).*
- *Un cilindro parabólico (cuando* λ_1 *ó* $\lambda_2 = 0$*).*

Si recordamos los nombres que dimos para estos casos en el final del resumen de la Guía 2, vemos que el gráfico de f es

- Un paraboloide si A es definida positiva.
- Un paraboloide invertido si A es definida negativa.
- Una silla de montar si A es indefinida.
- Un cilindro parabólico cuando A es semi-definida.

OBSERVACIÓN 3.2.13 (Cuádricas)**.** Estas superficies (paraboloide, silla de montar, cilindro parabólico) son ejemplos de lo que se conoce como *cuádricas*: superficies de $\mathbb{R}^3$ dadas (de manera implícita) por un polinomio de grado 2. En la sección 3.3 veremos otros ejemplos de cuádricas y haremos una lista de todos los posibles casos.

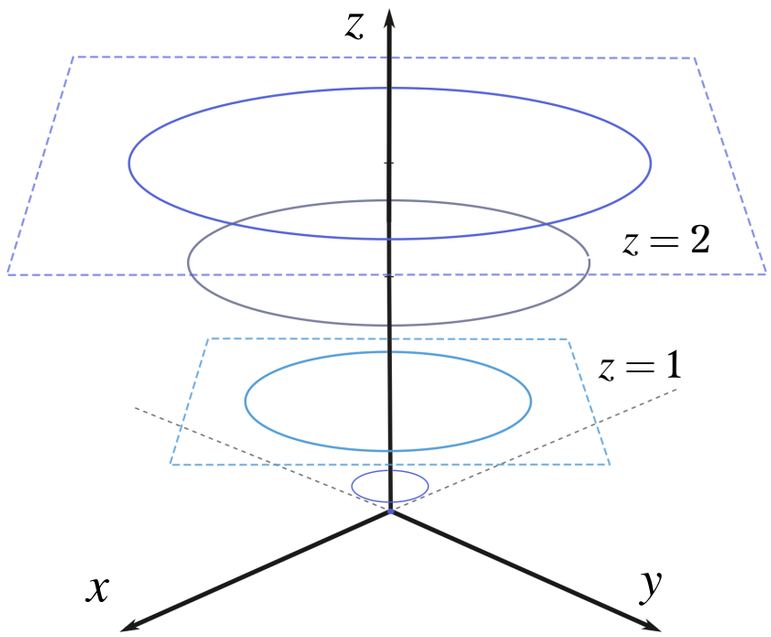
3.3. Curvas y superficies de nivel

• Corresponde a clases en video 3.6, 3.7

Como ya mencionamos, cuando tenemos una función $f : \mathbb{R}^2 \to \mathbb{R}$, a veces es útil mirar las denominadas *curvas de nivel* de la función f, que consisten en igualar f a una constante y ver qué curva obtenemos en el plano (x, y). Como estamos cortando con $z = cte$, podemos pensar que estamos haciendo cortes de la superficie del gráfico de f (que estaba dada por la ecuación implícita $z = f(x, y)$) con planos horizontales (porque esos planos tienen ecuación $z = cte$).

EJEMPLO 3.3.1 (Curvas de nivel del paraboloide)**.** Consideremos la función escalar $f(x,y) = x^2 + y^2$. Veamos sus curvas de nivel, y cómo estas ayudan a describir la superficie dada por el gráfico de f.

■ Si fijamos $z = R^2 > 0$, y miramos la curva de nivel $f(x,y) = R^2$, nos queda la ecuación $x^2 + y^2 = R^2$, que es una circunferencia centrada en el origen de radio $R > 0$.



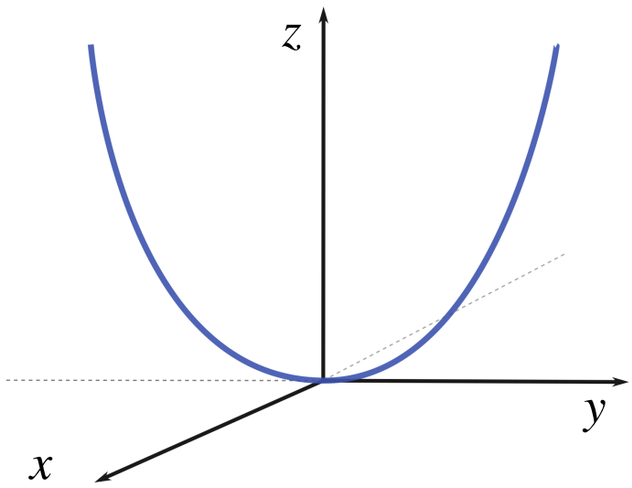
■ Si miramos la curva de nivel de $z = 0$, obtenemos $x^2 + y^2 = 0$, esto sólo es posible si $(x,y) = (0,0)$ entonces en este caso no obtenemos una curva sino un punto del plano (el origen).

■ Si miramos curvas de nivel con $z_0 < 0$, no hay solución y nos da el conjunto vacío; por ejemplo con $z = -1$ tenemos que ver el conjunto $x^2 + y^2 = -1$ que es vacío.

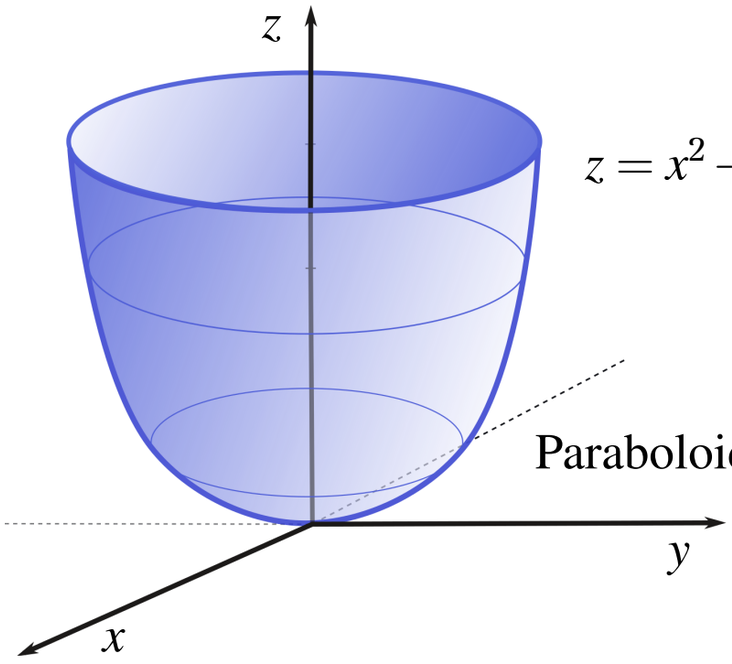
■ Resumiendo, los cortes con planos horizontales nos devuelven: nada para z negativo, el origen para $z = 0$ y circunferencias cada vez más grandes (de mayor radio) a medida que cortamos con $z > 0$ más grande.

■ Lo que queda por decidir es cómo unir estas circunferencias, para eso hacemos un corte con el plano vertical $x = 0$ (que es el plano yz, la pared del fondo). Lo que vemos es la ecuación $z = f(0,y)$ que en este caso es $z = 0^2 + y^2$, es decir $z = y^2$. La graficamos, es una parábola hacia arriba:

■ Eso quiere decir que hay que unir las circunferencias apiladas con una parábola, y por eso la gráfica de $f(x,y) = x^2 + y^2$ tiene la forma que



mencionamos en la sección anterior: se obtiene haciendo revolucionar la parábola $z = y^2$ alrededor del eje z.



EJEMPLO 3.3.2 (Curvas de nivel de una silla de montar)**.** Tomamos $f(x,y) = x^2 - y^2$, veamos sus curvas de nivel.
■ Si fijamos $z_0 = R^2 > 0$, y miramos la curva de nivel $f(x,y) = R^2$, nos queda la ecuación $x^2 - y^2 = R^2$. Dividimos por R^2 y vemos que queda

$$\frac{x^2}{R^2} - \frac{y^2}{R^2} = 1 \Longrightarrow \left(\frac{x}{R}\right)^2 - \left(\frac{y}{R}\right)^2 = 1.$$

Haciendo el cambio de variables $u = x/R$, $v = y/R$ (que consiste en comprimir un poco en la dirección de los ejes), vemos la ecuación $u^2 - v^2 = 1$. Entonces estos cortes devuelven hipérbolas (con dos ramas). ■ Si miramos la curva de nivel de $z = 0$, obtenemos $x^2 - y^2 = 0$, que equivale a $|x| = |y|$ y entonces tenemos dos posibilidades, $y = x$ ó

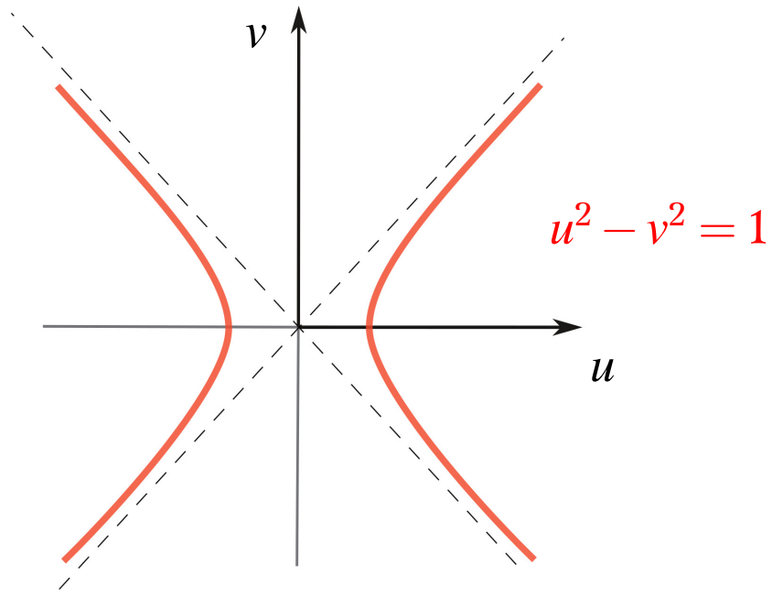


Figura 3.6: Hipérbola

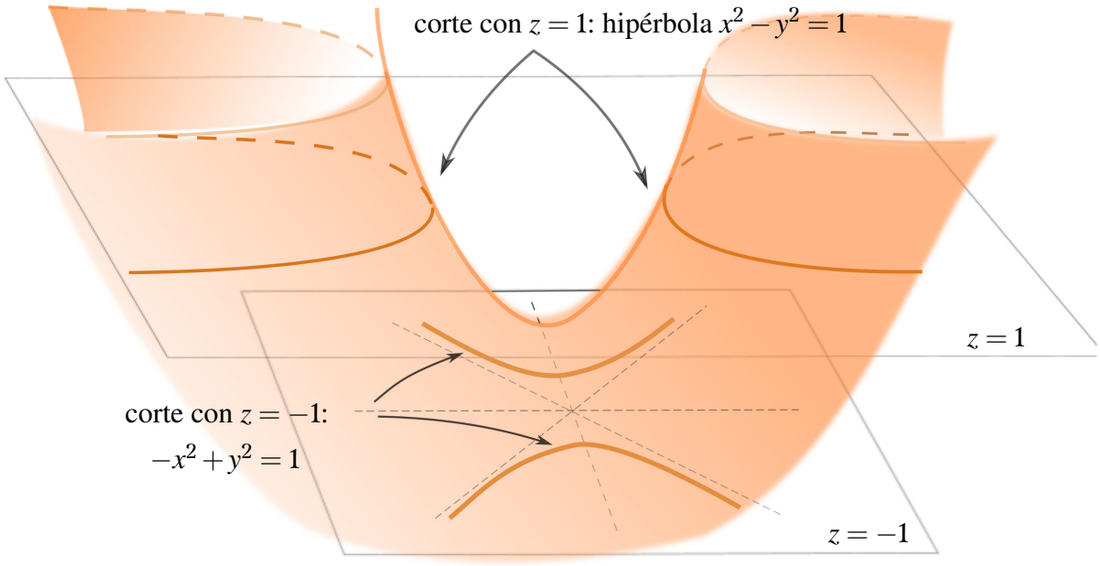


Figura 3.7: Curvas de nivel de la silla de montar $z = x^2 - y^2$

$y = -x$ que son dos rectas por el origen. ¡Si, en efecto, la silla de montar tiene dos rectas horizontales dentro, aunque no sea tan evidente a simple vista!

■ Si miramos curvas de nivel con $z_0 = -R^2 < 0$, el razonamiento es el mismo que para $z_0 > 0$, la única diferencia es que en lugar de ver la ecuación $u^2 - v^2 = 1$ luego del cambio de variables, ahora vemos la ecuación $-u^2 + v^2 = 1$, que equivale a intercambiar u con v en la anterior. Esta es una simetría, y entonces estas curvas de nivel también dan hipérbolas (con sus dos ramas) pero ahora tienen otra orientacion que las de arriba.

DEFINICIÓN 3.3.3 (Superficies de nivel)**.** Cuando tenemos una fun-

ción escalar de 3 variables $f:\mathbb{R}^3\to\mathbb{R}$, su gráfico es el subconjunto de $\mathbb{R}^4$ dado por $(x,y,z,f(x,y,z))$, o si se quiere son los (x,y,z,w) tales que $w=f(x,y,z)$. Como es un subconjunto de $\mathbb{R}^4$, no podemos esperar hacer un dibujo. Pero podemos estudiar los cortes con algunas de las variables fijas, por ejemplo tomando $w=w_0=cte$. Estos cortes se conocen como *superficies de nivel* de f, son superficies dadas por la ecuación implícita

$$f(x,y,z)=cte.$$

■ Si f es lineal sus superficies de nivel serán planos. Por ejemplo, $f(x,y,z)=2x+y-z$. Si hacemos $f(x,y,z)=0$ vemos el plano por el origen $\Pi:2x+y-z=0$. Mientras que $2x+y-z=5$ es un plano, pero que no pasa por el origen. Notemos que todas las superficies de nivel de f son planos paralelos, todos con normal $N=(2,1-1)$.

Ahora vamos a estudiar las superficies conocidas como *cuádricas*. Esta es una familia de ejemplos relativamente simples (y relevantes) dados por polinomios de grado 2 en las tres variables x,y,z. Ya vimos los casos particulares del paraboloide, la silla de montar y el cilindro parabólico, cuando teníamos superficies de la forma $z=f(x,y)$, es decir cuando la superficie era el gráfico de una función de dos variables (Teorema 3.2.9).

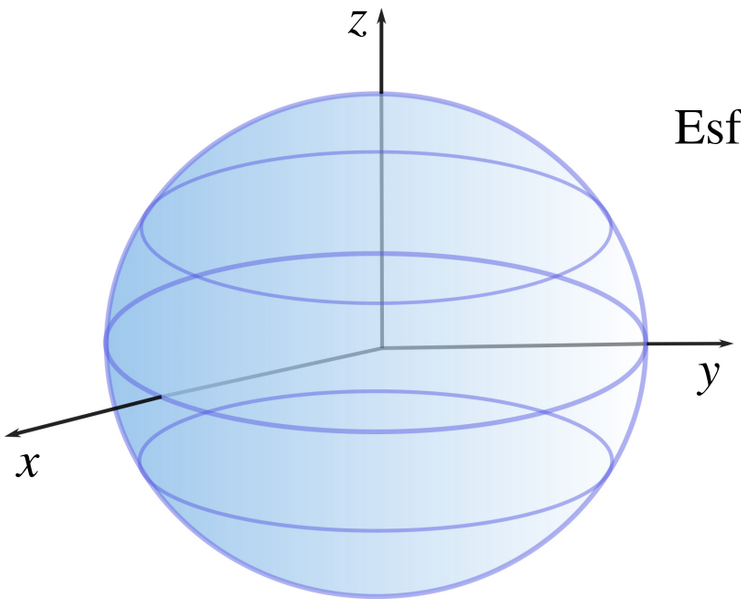
EJEMPLO 3.3.4 (Cuádricas)**.** Miramos superficies de nivel $f(x,y,z)=w_0$, con f polinomio de grado 2.

■ $f(x,y,z)=x^2+y^2+z^2$. Sea $R^2>0$, consideramos la superficie de nivel $f(x,y,z)=R^2$ (estamos tomando $w=R^2$). Notamos que si $X=(x,y,z)$, la ecuación se reescribe como

$$\text{Esfera de radio } R: \quad R^2=x^2+y^2+z^2=\|X\|^2, \ \ luego\ \|X\|=R.$$

Son los puntos de $\mathbb{R}^3$ que distan del origen exactamente R, se trata de una *superficie esférica de radio R*, o más brevemente una *esfera de radio R*.

■ Si tomamos $w=0$ y miramos $f(x,y,z)=0$, vemos $x^2+y^2+z^2=0$, así que en este caso el conjunto de soluciones no es una superficie sino



Esfera

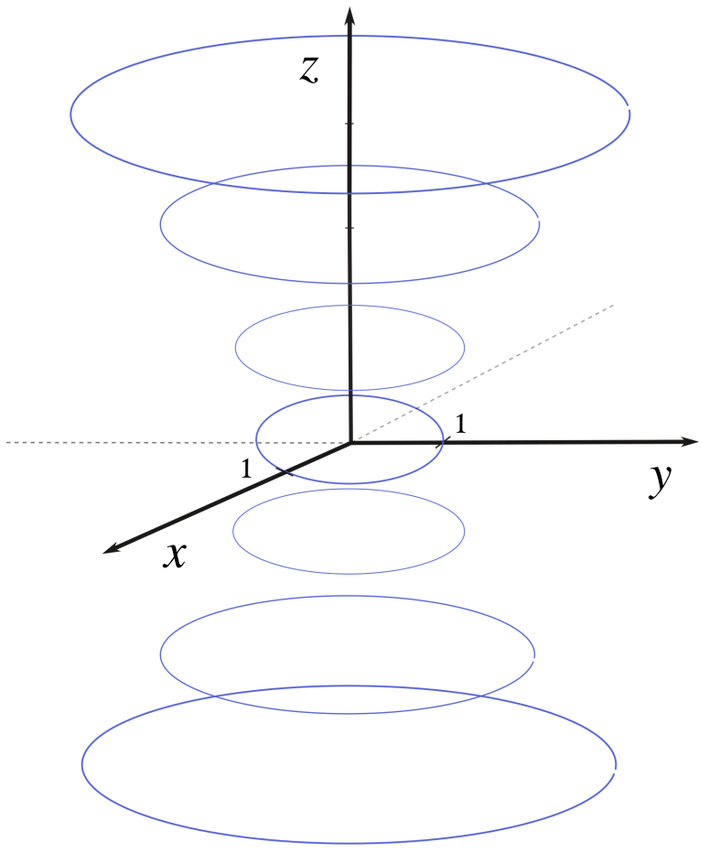
$x^2+y^2+z^2=1$

únicamente un punto, el punto $\mathbb{O}=(0,0,0)$.
■ Si miramos el conjunto $f(x,y,z)=z_0$ con $z_0<0$, resulta vacío: por ejemplo $x^2+y^2+z^2=-1$ no tiene soluciones.

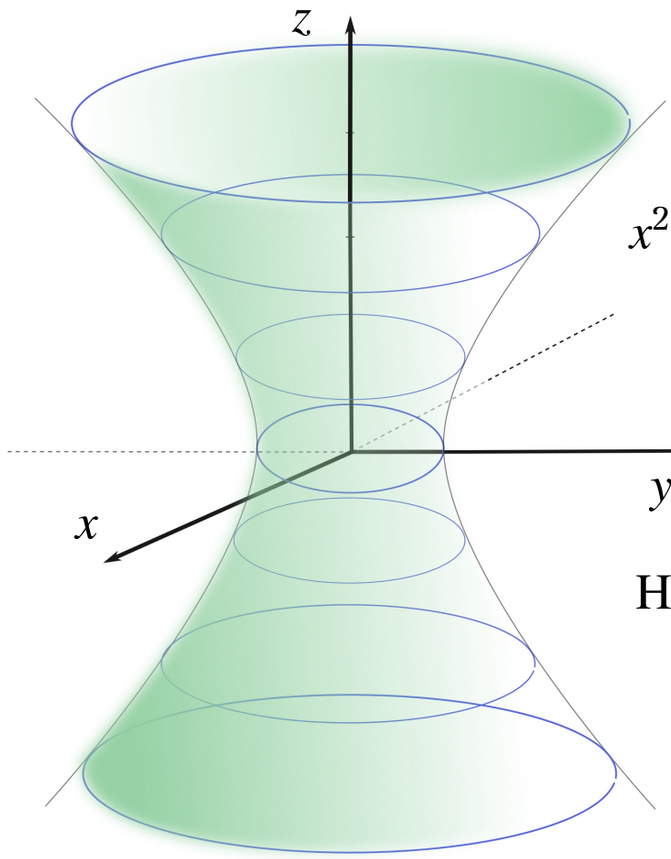
■ Cambiemos de función, sea $g(x,y,z)=x^2+y^2-z^2$. Miramos primero $x^2+y^2-z^2=R^2$, es decir tomamos $w>0$. Por ejemplo, si $R=1$ vemos $x^2+y^2-z^2=1$. Esta figura se conoce como *hiperboloide de una hoja*, su ecuación se reescribe como

$$\text{Hiperboloide de una hoja:}\quad x^2+y^2=1+z^2.$$

Para describirla podemos hacer cortes con planos horizontales (mirar las curvas de nivel tomando $z=z_0=cte$), notamos que $r^2=1+z_0^2\geq 1>0$ es positivo no importa el signo de z_0. Entonces todos los cortes horizontales son de la forma $x^2+y^2=r^2$, con $r\geq 1$ y se trata de circunferencias, todas de radio $r\geq 1$. El radio mínimo $r=1$ se consigue cuando $z_0=0$, y hay simetría respecto de z, entonces vemos circunferencias apiladas que crecen en radio a medida que nos alejamos del piso (en ambas direcciones).
El nombre de la figura se explica por su perfil: para ver cómo se unen estas circunferencias hacemos ahora un corte de $x^2+y^2=1+z^2$ con el plano vertical $x=0$. Vemos la ecuación $y^2=1+z^2$, que se reeescribe como $y^2-z^2=1$, y como ya sabemos, es una hipérbola (con sus dos ramas, como en la Figura 3.6 de más arriba). Entonces el hiperboloide



de una hoja se puede pensar como la superficie que se obtiene girando esta hipérbola alrededor del eje z.



$x^2+y^2-z^2=1$

Hiperboloide de de una hoja

■ Volviendo al caso general de la superficie de nivel $x^2+y^2-z^2=R^2$,

si dividimos por R^2 vemos

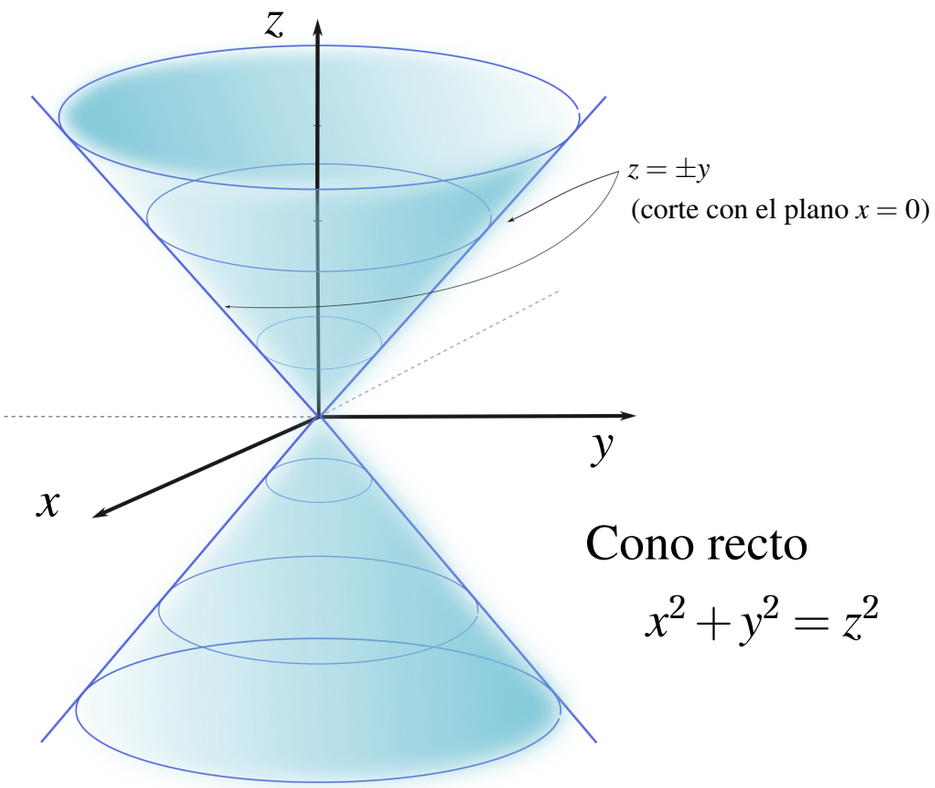
$$\frac{x^2}{R^2}+\frac{y^2}{R^2}-\frac{z^2}{R^2}=1 \Longrightarrow (\frac{x}{R})^2+(\frac{y}{R})^2-(\frac{z}{R})^2=1.$$

Con el cambio de variables $(x,y,z)\mapsto(u,v,w)=(\frac{1}{R}x,\frac{1}{R}y,\frac{1}{R}z)$, tenemos la superficie $u^2+v^2-w^2=1$ que como acabamos de discutir es un hiperboloide de una hoja. Entonces todas las superficies de nivel de $g(x,y,z)=x^2+y^2-z^2=R^2$ (para $R\neq 0$) son similares: son todos hiperboloides de una hoja, y lo que cambia es el radio mínimo (el de la circunferencia que se obtiene al cortar con el plano $z=0$).

■ Si tomamos $w=0$ vemos la superficie de nivel $x^2+y^2-z^2=0$ que ya no es un hiperboloide. Se trata de la superficie conocida como

$$\text{Cono recto:}\quad x^2+y^2=z^2.$$

o simplemente *cono*. En realidad se trata de dos conos opuestos por el vértice, como indica el dibujo: Tomando curvas de nivel, vemos que se



trata de circunferencias apiladas (salvo en $z=0$ donde es un punto). Luego para ver como unirlas, notamos que el perfil (el corte con el plano vertical $x=0$) nos da la ecuación $y^2=z^2$, que se trata de las

rectas $y = z$, $y = -z$.

■ Ahora veamos qué pasa si miramos $g(x,y,z) = w_0$ con $w_0 < 0$, es decir $x^2 + y^2 - z^2 = -R^2$ con $R \neq 0$. Dividiendo por R^2 y haciendo el mismo cambio de variables de antes, notamos que son todas (compresiones de) la superficie de nivel con $R = 1$, que es $x^2 + y^2 - z^2 = -1$. Esta es distina que el hiperboloide de una hoja por el signo, es la superficie conocida como

$$\text{Hiperboloide de dos hojas:} \quad x^2 + y^2 = z^2 - 1.$$

Para describirla podemos tomar curvas de nivel con $z = z_0 = cte$. Notemos que $z_0^2 - 1$ puede ser positivo, negativo o nulo, dependiendo de z_0. Separamos en casos: primero veamos donde se anula, esto es cuando $z_0^2 = 1$, y esto ocurre cuando $z_0 = 1$ ó bien $z_0 = -1$. Entonces los cortes con estos dos planos horizontales nos dan (en ambos casos) la ecuación $x^2 + y^2 = 0$, que sólo tiene solución al $(x,y) = (0,0)$.
Si tomamos $-1 < z_0 < 1$, vemos que $z_0^2 < 1$ y entonces $z_0^2 - 1 < 0$. En este caso la ecuación $x^2 + y^2 = z_0^2 - 1$ no tiene solución. Esto quiere decir que los cortes con planos $z = z_0$, con z_0 entre -1 y 1, no cortan la superficie $x^2 + y^2 = z^2 - 1$.
Por último, consideramos $|z_0| > 1$ (o sea debajo del -1 y por encima del 1), con esta condición $z_0^2 - 1 > 0$ y entonces la curva $x^2 + y^2 = z_0^2 - 1$ es una circunferencia (de radio cada vez mayor a medida que z se aleja del piso (ver Figura 3.9 debajo).
Para unir estas circunferencias apiladas, miramos el perfil de la superficie dado por el corte con el plano $x = 0$, vemos la ecuación $0^2 + y^2 = z^2 - 1$ o equivalentemente $z^2 - y^2 = 1$. Se trata de una hipérbola, ahora simétrica respecto del piso.
Entonces el hiperboloide de dos hojas es la superficie que puede obtenerse haciendo girar esta hipérbola (las dos ramas) alrededor del eje z.

■ Ahora cambiamos de función. Sea $f(x,y,z) = x^2 + y^2$, notemos que en la superficie de nivel $x^2 + y^2 = w_0$ la variable z no aparece. Eso quiere

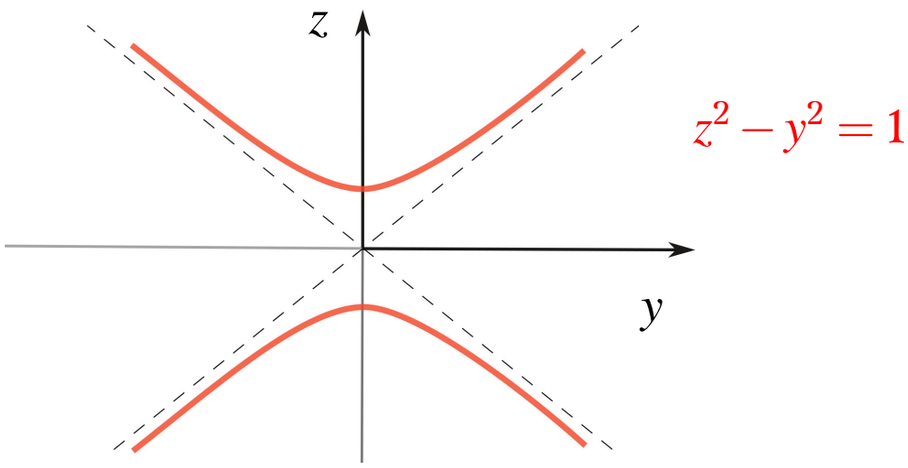


Figura 3.8: Las dos ramas de una hipérbola en el plano yz

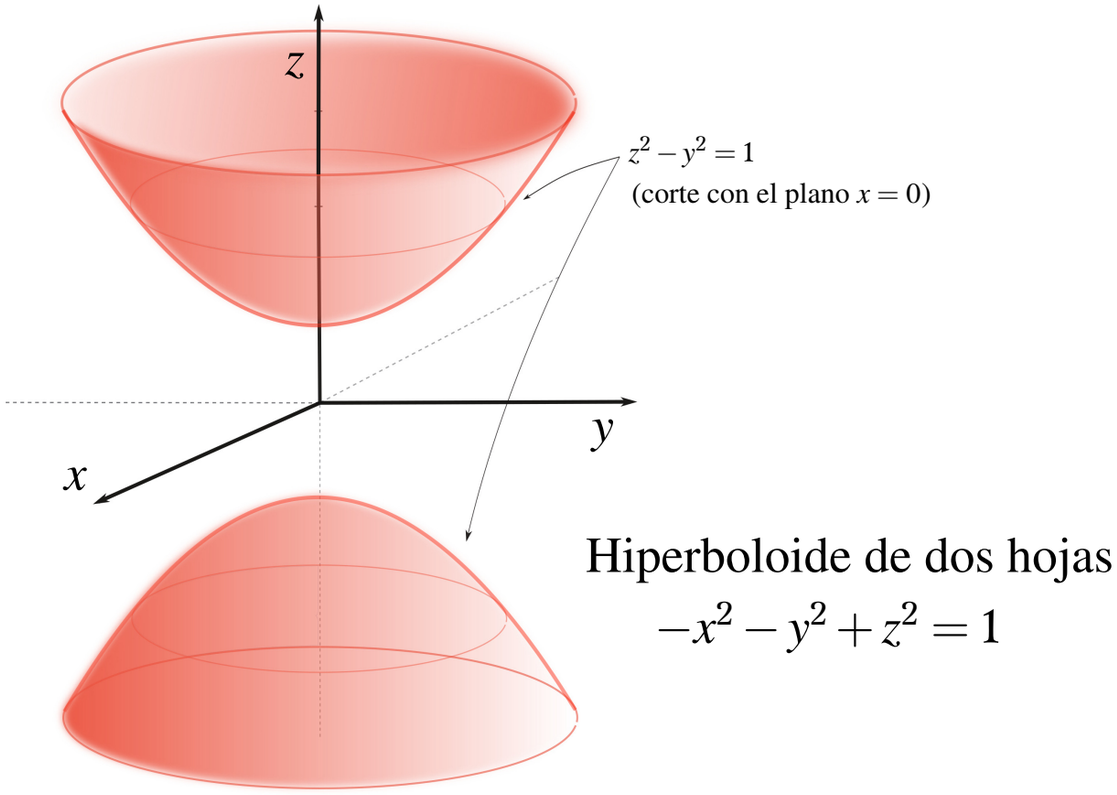


Figura 3.9: Hiperboloide de 2 hojas

decir que z *está libre*, e implica que una vez hecho el dibujo para algún z (por ejemplo, para $z = 0$) la superficie se obtiene dejando z libre por encima y por debajo de ese dibujo.

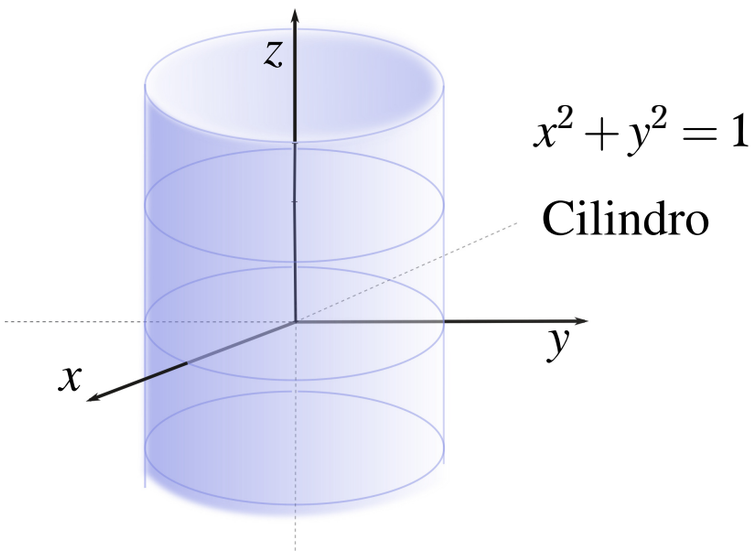
■ Si $w_0 = 0$ obtenemos $x^2 + y^2 = 0$, que en principio es el punto $x = 0, y = 0$. Pero si recordamos que estamos en $\mathbb{R}^3$ y que z está libre, lo que obtenemos es el conjunto $(0, 0, z)$, que es una recta vertical (es el eje z).

■ Si $w_0 < 0$ la superficie de nivel $x^2 + y^2 = w_0$ es vacía, porque el lado izquierdo es siempre no negativo. Notemos que no importa el valor de z (por más libre que esté), la ecuación nunca se va a verificar.

■ Si $w_0 = R^2 > 0$, tenemos $x^2 + y^2 = R^2$. Dividiendo por R^2 y cambiando las variables por x/R e y/R respectivamente obtenemos la superfice llamada

$$\text{Cilindro recto:} \quad x^2 + y^2 = 1$$

o simplemente *cilindro*. Vemos que todos los cortes de esta superficie con $z = cte$ son idénticos, porque la ecuación implícita no tiene z. Estos cortes son todos $x^2 + y^2 = 1$, que es una circunferencia de radio 1. Entonces la superficie se obtiene apilando circunferencias de radio 1, y es exactamente un cilindro vertical.



■ Nuevamente cambiamos de función. Consideramos en este ejemplo la función $f(x,y,z) = z^2 - y^2$. Ahora la variable faltante es x, y si miramos la superficie de nivel para $w_0 = 1$ vemos la ecuación $z^2 - y^2 = 1$. Esta superficie, al cortarla con planos verticales $x = cte$ (planos paralelos a la pared yz), presenta siempre la hipérbola plana $z^2 - y^2 = 1$ (como la de la Figura 3.8). La superficie se denomina *cilindro hiperbólico*, y se obtiene deslizando esta hipérbola hacia adelante y hacia atrás en la dirección del eje x:

OBSERVACIÓN 3.3.5 (Variantes de cuádricas)**.** Intercambiando las variables, obtenemos versiones de estas superficies que apuntan en las direcciones de los distintos ejes.

Por ejemplo, consideremos las siguientes ecuaciones:

$$a)\ x^2 + z^2 = y^2 \qquad b)\ -x^2 + y^2 - z^2 = 1 \qquad c)\ x^2 + z^2 = 1.$$

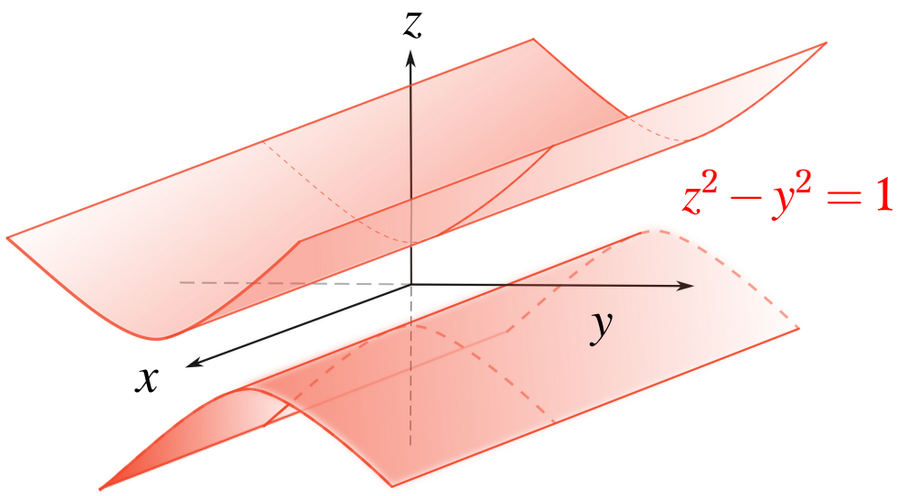


Figura 3.10: Cilindro hiperbólico

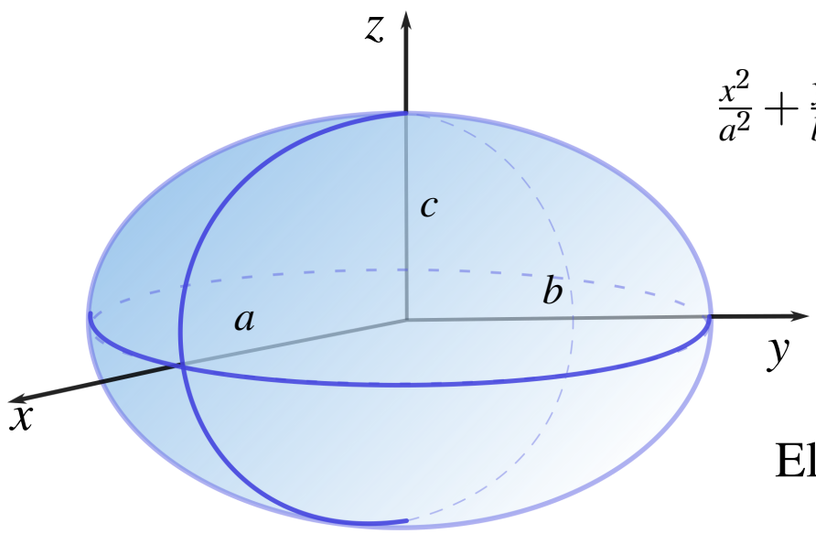
Para hacer un dibujo de estas superficies, no es necesario volver a calcular sus curvas de nivel, perfil, etc. Porque podemos notar (en cada caso) que la superficie dada tiene una ecuación similar a alguna de las que ya estudiamos, con la salvedad de que en estas de aquí, los roles de las variables están intercambiados. Eso quiere decir que estas superficies de aquí se pueden obtener a partir de alguna de las que ya estudiamos, haciendo una simetría en $\mathbb{R}^3$. ¿Qué simetría? La que se obtiene al intercambiar esas variables.

Por ejemplo: la ecuación $a)$ se obtiene a partir de la ecuación del cono recto $x^2+y^2=z^2$, intercambiando la variable y con la variable z. Luego la superficie $a)$ de aquí se obtiene del cono recto haciendo una simetría alrededor del plano $\Pi : y-z=0$ (ya que los puntos del plano, que verifican $y=z$, no se modifican al intercambiar las variables). Si queremos dibujar $a)$, notamos que en el cono recto original, el eje de rotación es el eje z. Luego $a)$ es un cono con eje de rotación alrededor del eje y.

Con los mismos argumentos, $b)$ es un hiperboloide de dos hojas con eje de rotació alrededor del eje y, mientras que $c)$ es un cilindro alrededor del eje y.

También podemos multiplicar las variables por constantes positivas, en ese caso obtenemos superficies cuádricas que resultan variantes "aplastadas" o "estiradas" en las direcciones de los ejes, tal como vimos en los casos del paraboloide y la silla de montar, en la Sección 3.2.

Por ejemplo la superficie dada por la ecuación $z = 3x^2$ es un cilindro parabólico, mientras que $4x^2 + 9y^2 = z^2$ es un cono que no es recto, puesto que los cortes horizonales son elipses.



$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

Elipsoide

Un último ejemplo interesante (y con nombre propio) es el de la superficie dada por la ecuación implícita

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

que puede obtenerse a partir de la esfera, modificando (x, y, z) por x/a, y/b, z/c respectivamente. Esta superficie cuádrica se conoce como *elipsoide* de radios a, b, c (cuando $a = b = c$ obtenemos una esfera).

TEOREMA 3.3.6 (Cuádricas en $\mathbb{R}^3$)**.** *Toda superficie cuádrica de $\mathbb{R}^3$ puede transformarse (mediante un cambio escalas en los ejes y un cambio de coordenadas ortogonal), en una de las superficies siguientes:*

- *Paraboloide:* $z = x^2 + y^2$
- *Silla de montar:* $z = x^2 - y^2$.
- *Cilindro parabólico:* $z = x^2$.
- *Esfera:* $x^2 + y^2 + z^2 = 1$
- *Hiperboloide de una hoja:* $x^2 + y^2 - z^2 = 1$

- *Hiperboloide de dos hojas:* $x^2+y^2-z^2=-1$
- *Cono:* $z^2=x^2+y^2$
- *Cilindro:* $x^2+y^2=1$
- *Cilindro hiperbólico:* $z^2-y^2=1$

Eso quiere decir que los ejemplos que vimos en este resumen, agotan todas las posibles superficies cuádricas en el espacio $\mathbb{R}^3$. Las funciones de la lista se conocen como las *formas canónicas* de los polinomios cuadráticos en dos variables. Puede verse una representación dinámica de estas figuras, así como sus intersecciones con distintos planos (sus curvas de nivel) en este link de Geogebra.

El método de prueba es similar al de aquel teorema, pero con algunas adiciones. No lo demostraremos, pero podemos ver muy claramente cómo funciona esta clasificación en un ejemplo:

EJEMPLO 3.3.7. Sea $g(x,y,z)=2x^2+2xy+2y^2-z^2$. Llevar g a su forma canónica con un cambio de variables adecuado y describir sus superficies de nivel.
La matriz para escribir $g(X)=AX\cdot X$ se consigue poniendo en la diagonal los coeficientes cuadráticos, y fuera de ella *la mitad* de los coeficientes cruzados, por ejemplo en el lugar $1-2$ va la mitad del coeficiente de xz, que es un 1, como no hay xz ni yz hay ceros en los lugares $1-3$ y $2-3$ respectivamente. La matriz tiene que ser simétrica así que es:

$$A=\begin{pmatrix}2&1&0\\1&2&0\\0&0&-1\end{pmatrix},$$

invitamos al lector a verificar que $g(X)=AX\cdot X$. Los autovalores de A son $3,1,-1$ y los autovectores (ya normalizados) son

$$V_1=(1/\sqrt{2},1/\sqrt{2},0)\quad V_2=(-1/\sqrt{2},1/\sqrt{2},0)\quad V_3=(0,0,1)$$

Las nuevas variables $\tilde{X} = (x_1, x_2, x_3)$ se obtienen aplicando la tranformación ortogonal U^t a las variables originales $X = (x, y, z)$, entonces tenemos

$$\begin{aligned} x_1 &= V_1 \cdot X = {}^{x}/\sqrt{2} + {}^{y}/\sqrt{2}y \\ x_2 &= V_2 \cdot X = {}^{-x}/\sqrt{2} + {}^{y}/\sqrt{2}y. \\ x_3 &= V_3 \cdot X = z \end{aligned}$$

La forma canónica de f en las nuevas variables es

$$D\tilde{X} \cdot \tilde{X} = 3x_1^2 + x_2^2 - x_3^2.$$

Entonces luego de una compresión de factor $1/\sqrt{3}$ en la primer variable la función original es equivalente a

$$f(x, y, z) = x^2 + y^2 - z^2.$$

Inspeccionando la lista del teorema, vemos que las superficies de nivel de esta última función son hiperboloides o un cono, entonces las de la superficie original también (salvo una transformación ortogonal y un dilatación de factor $\sqrt{3}$ en la dirección del eje x).

ƨ Si miramos ecuaciones implícitas generales $f(x, y, z) = cte$ donde f no es un pollinomio de grado 2, obtenemos muchos más ejemplos de superficies que no hemos presentado aquí. Es decir que en este resumen hemos hecho énfasis en las cuádricas pero hay muchas otras superficies en $\mathbb{R}^3$ que no lo son. Su dibujo puede ser mucho más complicado, y no tenemos una buena clasificación, pero de ser necesario se puede usar un programa graficador para obtenerlo. Lo que queremos destacar para finalizar, es que entender cuáles son las posibles cuádricas nos servirá sobre el final de este texto para comprender cómo se hallan los extremos de funciones de varias variables, y así poder resolver problemas de máximos y mínimos que involucren más de una variable.

Capítulo 4

Límite y derivadas de funciones

4.1. Límites y continuidad

• Corresponde a clases en video 4.1, 4.2a, 4.2b, 4.3, 4.4

DEFINICIÓN 4.1.1 (Limite en una variable)**.** Dada $f : A \to \mathbb{R}$, donde $A \subset \mathbb{R}$ es un intervalo o una unión de intervalos, decimos que *el límite de f cuando x tiende a x_0 es ℓ* si podemos conseguir que todos los $f(x)$ estén cercanos a ℓ, siempre que x esté cercano a x_0.
No nos interesa el valor de f en el punto $y_0 = f(x_0)$, ni siquiera nos interesa si $x_0 \in A$ (el dominio de f); necesitamos si poder aproximarnos a x_0 con puntos $x \in A$.
Formalmente $\lim_{x\to x_0} f(x) = \ell$ si

$$|f(x) - \ell| \text{ es chico, siempre que } |x - x_0| \text{ es chico } (x \neq x_0).$$

La idea de "es chico" requiere alguna formalización mayor. Porque no es aceptable que esté sujeta a interpretación del que hace la afirmación (o del que la lee).

Para ello vamos a decir que $|f(x) - \ell|$ tiene que ser *tan chico como uno quiera*, siempre que acercemos lo suficiente x a x_0. Esto se puede expresar mejor así: para cada eror posible $E > 0$, queremos encontrar

una $D = D(E) > 0$ (distancia) de manera tal que si $dist(x, x_0) < D$, entonces $|f(x) - \ell| < E$. Entonces $\lim_{x \to x_0} f(x) = \ell$ si para todo $E > 0$, existe $D > 0$ de manera tal que

$$|x - x_0| < D \Longrightarrow |f(x) - \ell| < E.$$

No hay que olvidar que hay que excluir la evaluación en x_0, para ello podemos escribir $x \neq x_0$ o para ser más breves escribimos $0 < |x - x_0| < D$ -puesto que $|x - x_0| \geq 0$ siempre, y $|x - x_0| = 0$ únicamente cuando $x = x_0$.

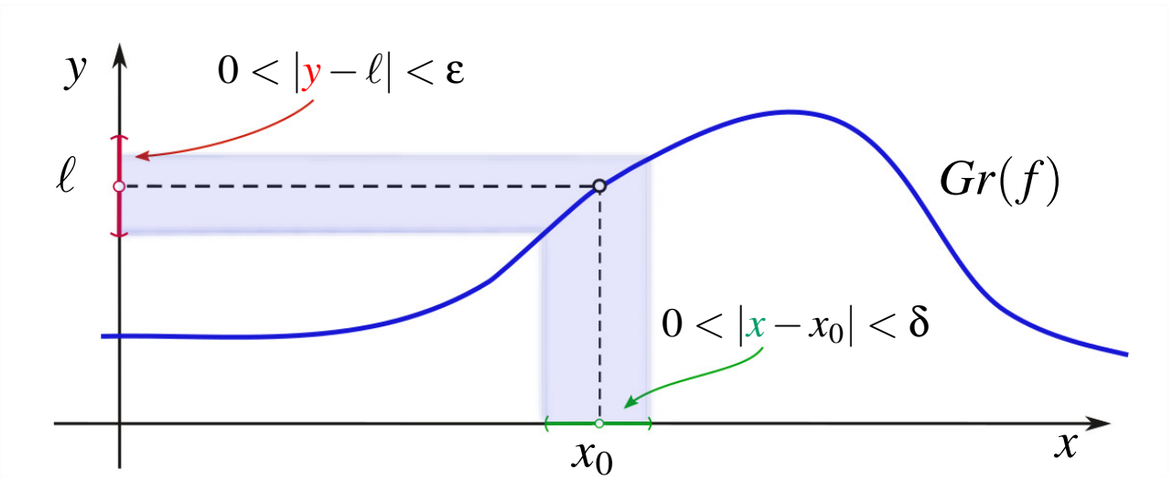


Figura 4.1: La banda horizontal celeste indica los $f(x)$ con $0 < |x - x_0| < \delta$

En la bibliografía es usual usar letras griegas para el error y la distancia, así que la definición de límite suele escribirse entonces como sigue:

$\lim_{x \to x_0} f(x) = \ell$ si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que

$$0 < |x - x_0| < \delta \Longrightarrow |f(x) - \ell| < \varepsilon. \qquad (4.1.1)$$

2 Como estamos aproximando x a x_0, es necesario que x_0 esté en A o pegado a él. Luego debe ser $x_0 \in A \cup bd(A) = \overline{A}$ (la clausura de A).

OBSERVACIÓN 4.1.2 (Límites en infinito)**.** Escribimos $\lim_{x \to +\infty} f(x) = \ell$ si $f(x)$ se aproxima a ℓ cuando x es cada vez mayor. Formalmente, el límite en $+\infty$ es ℓ si para todo $\varepsilon > 0$ existe $K > 0$ tal que

$$x > K \Longrightarrow |f(x) - \ell| < \varepsilon.$$

La misma idea se usa para $\lim_{x\to-\infty} f(x) = \ell$, pero ahora queremos que $f(x)$ esté cerca de ℓ cuando x sea *negativo* y muy grande en valor absoluto (gráficamente, x está muy a la izquierda de 0 y $f(x)$ muy cerca de la altura ℓ).

A veces nos interesa ver qué ocurre si nos aproximamos a x_0 por un sólo lado (ya sea porque f no está definida a un lado, o porque el comportamiento es distinto a ambos lados de x_0):

DEFINICIÓN 4.1.3 (Límites laterales)**.** Si $Dom(f) \subset R$, decimos que $\lim_{x\to x_0^+} f(x) = \ell$ si $\forall\, \varepsilon > 0,\ \exists \delta > 0$ tal que

$$x_0 < x < x_0 + \delta \Longrightarrow |f(x) - \ell| < \varepsilon.$$

Este se llama límite lateral por la derecha, y requerimos entonces x cerca de x_0 con la condición adicional $x > z_0$.
Similarmente se define el límite lateral por la izquierda, en ese caso es $x < x_0$.

ƨ Si hay puntos de $A = Dom(f)$ a ambos lados de x_0, entonces: existen los dos límites laterales *y son iguales* si y solo si existe el límite en x_0.

Ahora extendemos la definición de límite en un punto a funciones de varias variables.

DEFINICIÓN 4.1.4. [Limite en $\mathbb{R}^n$] Sea $A \subset \mathbb{R}^n$, sea $f : A \to \mathbb{R}$, sea $P \in \overline{A}$. Decimos que $\lim_{X\to P} f(X) = \ell$ si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que

$$0 < \|X - P\| < \delta \Longrightarrow |f(X) - \ell| < \varepsilon.$$

La idea es exactamente la misma, queremos que $f(x)$ esté cerca de ℓ si X está cerca del punto P (y además pedimos $X \neq P$ como antes). Decimos "el límite de f cuando X tiende a P es ℓ".

OBSERVACIÓN 4.1.5 (Notación)**.** Escribimos $\lim_{X\to X_0} f(X) = \ell$ a veces de la siguiente manera

$$f(X) \xrightarrow[X\to X_0]{} \ell.$$

La idea es clara: los valores de $f(X)$ se aproximan a ℓ cuando los valores de X se aproximan a X_0.

Algunas propiedades útiles se listan a continuación

PROPIEDADES 4.1.6 (del límite de funciones)**.** Sea $A \subset \mathbb{R}^n$, sean $f, g : A \to \mathbb{R}$, sea $X_0 \in \overline{A}$. Entonces si existen los límites de f y de g en X_0, se tiene

1. $\lim_{X\to X_0}(f(X)+g(X)) = \lim_{X\to X_0} f(X) + \lim_{X\to X_0} g(X)$.
2. $\lim_{X\to X_0} f(X)g(X) = \lim_{X\to X_0} f(X) \cdot \lim_{X\to X_0} g(X)$.
3. Si $g \neq 0$ en A y$\lim_{X\to X_0} g(X) \neq 0$, entonces

$$\lim_{X\to X_0} \frac{f(X)}{g(X)} = \frac{\lim_{X\to X_0} f(X)}{\lim_{X\to X_0} g(X)}.$$

Recordemos que -si puede hacerse- la composición de dos funciones $(g \circ f)(X) = g(f(X))$ se calcula viendo primero el valor de f en x, y luego viendo el valor de g en $f(X)$.

PROPIEDADES 4.1.7 (de la composición y el límite)**.** Si $f(X) \xrightarrow[X\to X_0]{} \ell$ y además $g(X) \xrightarrow[X\to\ell]{} L$, entonces existe el límite de la composición y es igual a L:

$$(g \circ f)(X) \xrightarrow[X\to X_0]{} L.$$

Esto es porque si comenzamos cerca de X_0, los valores de $Y = f(X)$ están cerca de ℓ, luego los valores de $Z = g(Y)$ están cerca de L.

OBSERVACIÓN 4.1.8 (Límites que dan infinito)**.** Escribimos

$$\lim_{X\to X_0} f(X) = +\infty$$

si $f(X)$ es cada vez mayor (sin cota superior) a medida que X se aproxima a X_0. Formalmente, $f(X) \xrightarrow[X\to X_0]{} +\infty$ si para todo $M > 0$ existe

$\delta > 0$ tal que

$$0 < \|X - X_0\| < \delta \Longrightarrow f(X) > M.$$

Vemos que el gráfico se hace cada vez más alto a medida que X se aproxima a X_0. Aquí f puede ser una función de una o de varias variables.
La misma idea se usa para $\lim_{X \to X_0} f(x) = -\infty$: decimos que

$$f(X) \xrightarrow[X \to X_0]{} -\infty$$

si para todo $m < 0$ existe $\delta > 0$ tal que

$$0 < \|X - X_0\| < \delta \Longrightarrow f(X) < m.$$

En este caso vemos que el gráfico se hace más bajo (hacia $-\infty$) a medida que $X \to X_0$ (no hay cota inferior para las imágenes de f).

PROPIEDADES 4.1.9 (del sandwich y "cero por acotada")**.** Cuando queremos probar que un límite es 0, podemos hacer uso una desigualdad del tipo

$$|f(t)| \leq |g(t)|.$$

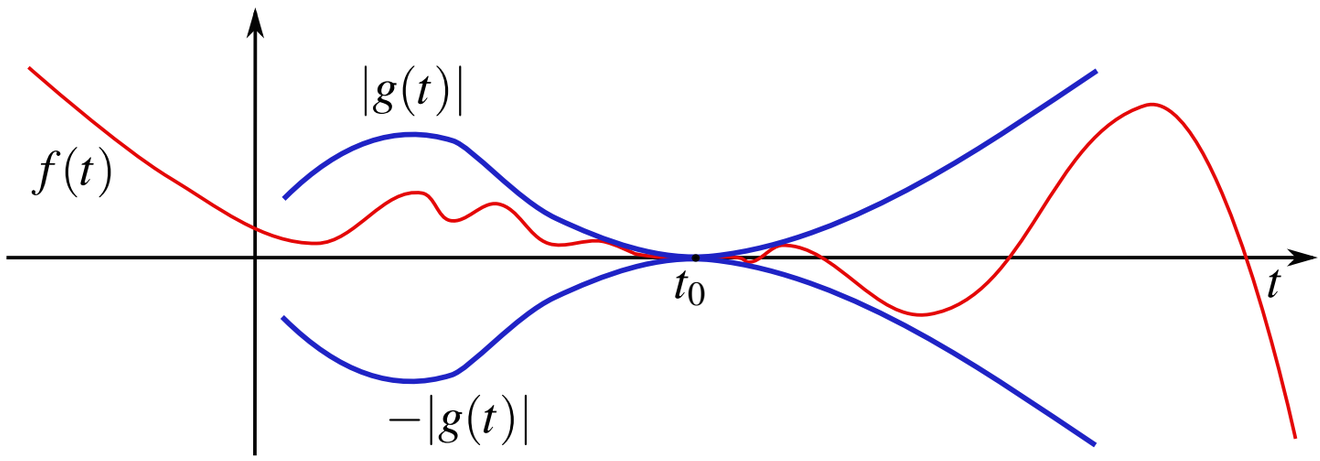
Si sabemos que $g(t) \to 0$ cuando $t \to t_0$, entonces podemos concluir que también $f(t) \to 0$ cuando $t \to t_0$. Esto es porque

$$-|g(t)| \leq f(t) \leq |g(t)|$$

y entonces f no puede escapar de tener límite nulo en $t = t_0$, como se ve en la figura:
Por ejemplo:

■ Calcular $\lim_{x \to 0} x \operatorname{sen}(1/x)$. Como $|\operatorname{sen}(1/x)| \leq 1$ entonces

$$|f(x)| = |x \operatorname{sen}(1/x)| \leq |x| \to 0$$



cuando $x \to 0$. Luego $\lim_{x\to 0} x\,\mathrm{sen}(1/x) = 0$. Aquí la función que podemos controlar fácilmente, que es más grande que f, es la función $g(x) = x$.
Este es el tipo de límite donde justificamos que da 0 porque es del tipo "0 por acotada". En este caso la función que tiende a cero es $g(x) = x$, mientras que el otro factor está acotado.

■ Calcular $\lim_{x\to 1^+} \ln(x)\cos(1/\ln(x))$. Afirmamos que $\ell = 0$ porque $\ln(1) = 0$. Sin embargo el segundo factor no tiene límite, por eso necesitamos justificarlo así:

$$|\ln(x)\cos(1/\ln(x))| \leq |\ln(x)| \to 0$$

cuando $x \to 1$. Entonces por la propiedad del sandwich el límite original es 0. Nuevamente es de la forma "0 por acotada", en este caso el factor acotado es $\cos(1/\ln(x))$, que no tiene límite pero sus valores están acotados entre -1 y 1.

PROPOSICIÓN 4.1.10 (Cálculo de límites usando la composición)**.** *Como la composición de límites es el límite de la composición, podemos aprovechar eso para calcular límites donde reconozcamos expresiones conocidas.*

Veamos algunos ejemplos importantes de esta aplicación. El lector curioso puede ver la forma que tiene las superficies dadas por los gráficos de las siguientes funciones, en este link. Allí notarán que el resultado que obtenemos en las cuentas debajo, no siempre es evidente a partir del dibujo.

■ Como $\lim_{t\to 0}(1+t)^{1/t} = e$, entonces si $f(x,y) \to 0$ cuando $(x,y) \to (x_0,y_0)$, al hacer la composición con f podemos calcluar el límite:

1. $\ell = \lim_{(x,y)\to(0,0)} \left(1+x^2+y^4\right)^{\frac{-7}{x^2+y^4}}$. Como $(r^b)^a = r^{ab}$, tenemos que

$$\ell = \lim_{(x,y)\to(0,0)} \left(\left(1+x^2+y^4\right)^{\frac{1}{x^2+y^4}} \right)^{-7},$$

y entonces

$$\ell = \left(\lim_{(x,y)\to(0,0)} \left(1+x^2+y^4\right)^{\frac{1}{x^2+y^4}} \right)^{-7} = e^{-7}.$$

Aplicamos la idea de la composición a $f(x,y) = x^2+y^4$.

2. $\lim_{(x,y)\to(1,2)} \left(1+y\ln(x)^2\right)^{\frac{y}{3\ln(x)^2}}$. Como $f(x,y) = y\ln(x)^2 \to 0$ cuando $(x,y) \to (1,2)$, tenemos

$$\left(1+y\ln(x)^2\right)^{\frac{y}{3\ln(x)^2}} = \left(1+y\ln(x)^2\right)^{\frac{y^2}{3\cdot y\ln(x)^2}}$$

$$= \left(\left(1+y\ln(x)^2\right)^{\frac{1}{y\ln(x)^2}} \right)^{y^2/3} \longrightarrow e^{2^2/3} = e^{4/3}.$$

■ Como $\lim_{t\to 0} \frac{\operatorname{sen}(t)}{t} = 1$, tenemos

1. $\lim_{(x,y)\to(0,0)} \dfrac{\operatorname{sen}(5xy)}{xy} = 5\lim_{(x,y)\to(0,0)} \dfrac{\operatorname{sen}(5xy)}{5xy} = 5\cdot 1 = 5.$

Esta vez compusimos con $f(x,y) = 5xy$.

2. $\lim_{(x,y)\to(3,1)} \dfrac{\operatorname{sen}(xy-x)}{y-1} = \lim_{(x,y)\to(3,1)} \dfrac{\operatorname{sen}(xy-x)}{xy-x}\dfrac{xy-x}{y-1}$

$$= \lim_{(x,y)\to(3,1)} \frac{\operatorname{sen}(xy-x)}{xy-x}\frac{x(y-1)}{y-1}$$

$$= \lim_{(x,y)\to(3,1)} \frac{\operatorname{sen}(xy-x)}{xy-x}\cdot x = 1\cdot 3 = 3.$$

En este caso compusimos con $f(x,y) = xy - x = x(y-1)$. Notemos que para esta función, $f(x,y) \to 0$ cuando $(x,y) \to (3,1)$ (si no fuera así, el límite estaría mal calculado).

ƨ Un caso muy importante de $f(x,y) \to 0$ que vamos a usar (cuando $(x,y) \to (0,0)$) es mirando la función $f(x,y) = \|(x,y)\|$, o sus potencias como

$$\|(x,y)\|^2 = x^2 + y^2.$$

PROPIEDADES 4.1.11 (de la norma)**.** Como $x^2 + y^2 \geq x^2$, se tiene tomando raíz que $|x| \leq \sqrt{x^2+y^2} = \|(x,y)\|$. Entonces

$$-\|(x,y)\| \leq x \leq \|(x,y)\|$$

y también

$$-\|(x,y)\| \leq y \leq \|(x,y)\|.$$

Entonces, si $(x,y) \to (0,0)$ es claro que $x \to 0$ y también $y \to 0$, pero esta cuenta nos dice que la norma lo hace más rápido que las coordenadas.
Por otro lado es claro que si $(x,y) \to (0,0)$, entonces también se tiene que $\|(x,y)\| \to 0$.

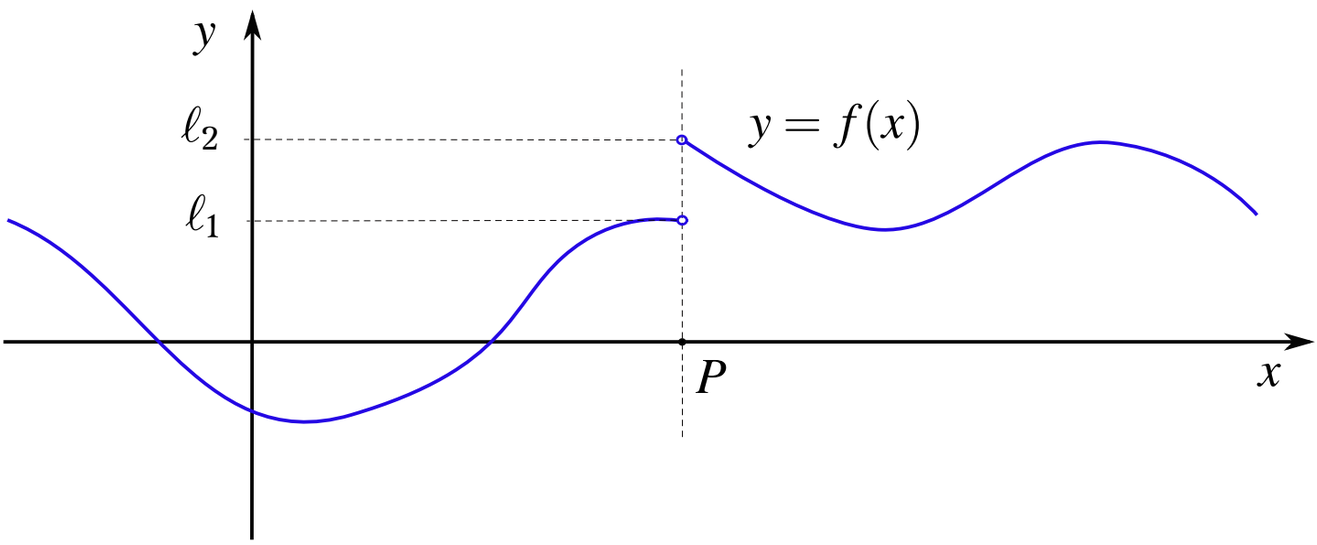


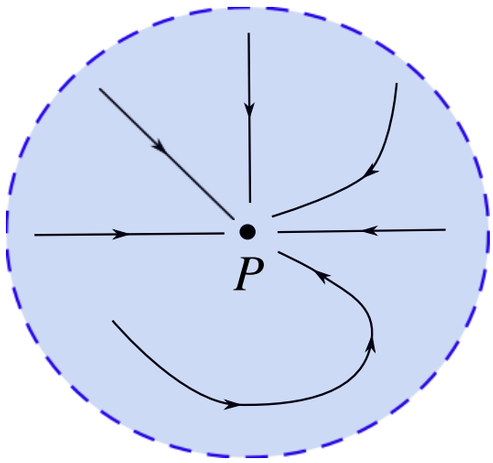
Figura 4.2: Límites laterales distintos, el límite en P no existe

OBSERVACIÓN 4.1.12 (Limite a lo largo de curvas)**.** Cuando el límite en P existe, no importa por qué camino nos acercemos a P, siempre

debe ser $|f(X)-\ell| \to 0$. Entonces si encontramos dos maneras distintas de acercarnos, que den límites distintos, es porque el límite en P *no existe*.

El primer ejemplo de esto es para funciones de una variable (Figura 4.2: cuando tenemos dos límites laterales que no coinciden. Sabemos que el límite no existe en ese caso.

Si miramos funciones de dos variables, un punto de su dominio estará en $\mathbb{R}^2$. Hay muchas maneras de acercarse a un punto $P \in \mathbb{R}^2$.



Por ejemplo, por cualquier recta que pase por P. Esto no agota todos los caminos, pero es la primer cosa con la que uno puede probar. Si encontramos que el limite da distinto a lo largo de dos rectas, es porque no existe el límite doble en P.

EJEMPLO 4.1.13. Vamos resolver límites en algunos ejemplos. Nuevamente podemos ver las superficies dadas por $Gr(f)$ en Geogebra, si clickeamos en este link.

Como ya mencionamos, no siempre es de mucha utilidad ver el gráfico, y que es necesario saber algunas técnicas como las que explicamos a continuación para ver si un límite existe (y cuánto da). En el applet de Geogebra también están algunas de las funciones de la Guia de TP 4.

■ Sea $f(x,y) = \dfrac{x^2-y^2}{x^2+y^2}$. Queremos ver si existe $\lim_{(x,y)\to(0,0)} f(x,y)$.

Comenzamos probando por los ejes, que pasan por el punto $P=(0,0)$. Probamos primer por el eje x. Eso quiere decir que tomamos $y=0$ y

vemos que queda

$$f(x,0) = \frac{x^2 - 0^2}{x^2 + 0^2} = \frac{x^2}{x^2} = 1.$$

Entonces f es constante en el eje x y el límite a lo largo del eje x da 1. Ahora probamos por el eje y. Eso quiere decir que tomamos $x = 0$ y vemos que queda

$$f(0,y) = \frac{0^2 - y^2}{0^2 + y^2} = \frac{-y^2}{y^2} = -1.$$

Entonces f es constante en el eje y y el límite a lo largo da -1.
Como los límites por los dos ejes son distintos, no existe el límite de f en $P = (0,0)$.

■ Sea $f(x,y) = \dfrac{xy}{x^2+y^2}$, en $P = (0,0)$. Probamos primer por el eje x. Eso quiere decir que tomamos $y = 0$ y vemos que queda

$$f(x,0) = \frac{0}{x^2} = 0.$$

Entonces f es constante en el eje x y el límite a lo largo del eje x da 0. Ahora probamos por el eje y. Eso quiere decir que tomamos $x = 0$ y vemos que queda

$$f(0,y) = \frac{0}{y^2} = 0.$$

Entonces f es constante en el eje y y el límite a lo largo da 0.
Como los límites por los dos ejes son iguales, todavía no podemos llegar a ninguna conclusión. Probamos con una recta por el origen genérica, $y = mx$. Tenemos

$$f(x,mx) = \frac{x \cdot mx}{x^2 + (mx)^2} = \frac{m \cdot x^2}{x^2 + m^2x^2} = \frac{m \cdot x^2}{x^2(1+m^2)}$$

Entonces el límite a lo largo de la recta $y = mx$ es

$$\lim_{x\to 0} f(x,mx) = \frac{m}{1+m^2}.$$

Si $m = 0$ tenemos el eje x, que ya sabíamos que daba 0. Pero si tomamos cualquier otra pendiente, por ejemplo $m = 1$, vemos que el límite da $1/2$. Eso quiere decir que el límite de f en $(0,0)$ acercándonos por la recta diagonal $y = x$ da $1/2$. Como por los ejes daba 0, vemos que no existe el límite de f en $P = (0,0)$.

■ Sea $f(x,y) = \dfrac{x^2 y}{x^4 + y^2}$ en $P = (0,0)$. Por los ejes, el límite da 0 (ver la cuenta del ejemplo anterior). Por las rectas $y = mx$ tenemos

$$\begin{aligned}\lim_{x\to 0} f(x,mx) &= \lim_{x\to 0} \frac{mx^3}{x^4 + m^2x^2} = \lim_{x\to 0} \frac{mx^3}{x^4 + m^2x^2} \\ &= \lim_{x\to 0} \frac{mx^3}{x^2(x^2 + m^2)} = \lim_{x\to 0} \frac{mx}{(x^2 + m^2)} = \frac{0}{m^2} = 0,\end{aligned}$$

puesto que podemos suponer que $m \neq 0$ (el caso $m = 0$ ya lo estudiamos cuando nos acercamos por el eje x). Entonces por cualquier recta por el origen da 0.

¿Quiere decir que el límite doble es 0? No necesariamente, ya que hay otros caminos para acercarse al origen que no son rectas.

Por ejemplo, la parábola $y = x^2$. Calculamos f a lo largo de esa curva,

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

Entonces f es constante a lo largo de esa parábola (que pasa por el origen) y así podemos afirmar que el límite de f a lo largo de esa parábola es $1/2$. Como este límite es distinto de 0 (que era lo que daba a lo largo de, por ejemplo, el eje x), podemos afirmar que no existe el límite de f en $P = (0,0)$.

■ Sea $f(x,y) = \dfrac{4x^2 y}{x^2 + y^2}$ en $P = (0,0)$. Por los ejes, el límite da 0 (ver

la cuenta dos ejemplos atrás). Por las rectas $y = mx$ tenemos

$$\lim_{x\to 0} f(x,mx) = \lim_{x\to 0} \frac{4mx^3}{x^2+m^2x^2} = \lim_{x\to 0} \frac{4mx^3}{x^2+m^2x^2}$$
$$= \lim_{x\to 0} \frac{4mx^3}{x^2(1+m^2)} = \lim_{x\to 0} \frac{4mx}{(1+m^2)} = \frac{0}{1+m^2} = 0.$$

Entonces por cualquier recta por el origen da 0. ¿Quiere decir que el límite doble es 0? Podemos probar por parábolas y otras curvas. Invitamos al lector a hacerlo. Vamos a ver que a lo largo de todas ellas el límite es 0.

¿Cómo podemos probar que el límite doble es cero? No tiene sentido probar por infinitos caminos. Lo que vamos a hacer es probar que $f(x,y) - \ell$ (en este caso $f(x,y)$ pues el candidato es $\ell = 0$) es tan pequeño como querramos, a medida que $(x,y) \to 0$. O sea vamos a probar que el límite doble es 0 usando su definición. Para eso usamos la Propiedad 4.1.11 de la norma , y escribimos

$$|f(x,y) - 0| = |f(x,y)| = \frac{|4x^2y|}{|x^2+y^2|} = \frac{4x^2|y|}{x^2+y^2} \leq \frac{4\|(x,y)\|^2 \cdot \|(x,y)\|}{\|(x,y)\|^2}.$$

Aquí usamos que $x^2 \leq \|(x,y)\|^2$ y también que $|y| \leq \|(x,y)\|$. El denominador, por otro lado, era exactamente igual a $\|(x,y)\|^2$.

Entonces cancelando las normas (recordemos que no nos interesa que pasa exactamente en $(x,y) = (0,0)$), tenemos

$$|f(x,y) - 0| = |f(x,y)| \leq 4\|(x,y)\|.$$

Pero sabemos que $(x,y) \to (0,0)$, luego $4\|(x,y)\| \to 0$. Por la propiedad del sandwich, debe ser $f(x,y) \to 0$ cuando $(x,y) \to (0,0)$, luego el límite si existe y es nulo, esto es

$$\lim_{(x,y)\to(0,0)} \frac{4x^2y}{x^2+y^2} = 0.$$

2 Idea: cuando tenemos un cociente de polinomios, y el límite es de la

forma $0/0$, es útil pensar qué grado tiene el numerador y qué grado tiene el denominador. Si tienen el mismo grado, es bastante probable que el límite no exista. En cambio si el grado del numerador (en el ejemplo previo, grado 3) es mayor que el del denominador (en el ejemplo previo, grado 2) entonces el numerador tiende a cero más rápido que el denominador y eso nos dice que es bastante probable que el líimite exista y sea nulo. Por supuesto que esta idea es sólo un indicador, que luego hay revisar en los hechos con las técnicas que indicamos aquí.

Continuidad

DEFINICIÓN 4.1.14 (Continuidad)**.** Cuando $X_0 \in A$ podemos comparar el límite con el valor al evaluar, si coinciden decimos que f es continua en X_0. Resumiendo f es continua en X_0 si

$$\lim_{X \to X_0} f(X) = f(X_0).$$

Entonces necesitamos que exista el límite y que coincida con $f(X_0)$.

2 También interesa ver si se puede extender la continuidad de f a otros puntos que no estén en el dominio. Luego el estudio de continuidad se hace en $\overline{A} = A \cup bd(A)$.

DEFINICIÓN 4.1.15 (Discontinuidades evitables)**.** Si $X_0 \in A = Dom(f)$ y existe el límite

$$\lim_{X \to X_0} f(X) = \ell,$$

pero $f(X_0) \neq \ell$, decimos que X_0 es una *discontinuidad evitable* de f. Esto es porque si cambiamos el valor de f en X_0 por el número ℓ, la función queda continua en X_0.
También decimos que $X_0 \in bd(A) \setminus A$ es una discontinuidad evitable si existe el límite

$$\lim_{X \to X_0} f(X) = \ell,$$

en este caso *definimos* f en X_0 como ℓ y de esta manera la extensión queda continua en X_0.

DEFINICIÓN 4.1.16 (Discontinuidades esenciales)**.** Si $X_0 \in \overline{A}$ pero no existe el límite cuando $X \to X_0$ de f, decimos que la discontinuidad es *esencial*.

Para funciones de una variable, un caso particular de discontinuidad esencial es el salto: cuando existen ambos límites laterales pero son *distintos*.
Una discontuinidad esencial que no es de salto la tiene $f(x) = \text{sen}(1/x)$ en $x_0 = 0$. Aquí no existe ninguno de los dos límites laterales.

DEFINICIÓN 4.1.17 (Continuidad en conjuntos)**.** Dada $f : A \to \mathbb{R}$ decimos que f es continua en el conjunto $A \subset \mathbb{R}^n$ si es continua en todos y cada uno de los puntos $P \in A$.

PROPIEDADES 4.1.18 (de las funciones continuas)**.** Sean $f, g : A \to \mathbb{R}$ funciones continuas. Entonces

1. La suma $f + g$ es continua en A.
2. El producto $f \cdot g$ es continuo en A.
3. Si g no se anula en A, entonces el cociente f/g es continuo en A.
4. Si componemos dos funciones continuas (siempre que se pueda hacer la composición), el resultado es una nueva función continua.

Esto nos dice que hay que identificar dos tipos de problemas: cuando hay un cociente, ver donde se anula el denominador, y cuando hay una composición, ver dónde hay problemas del dominio. Allí estarán las posibles discontinuidades.

Curvas

Una curva es una función de una variable $\alpha : I \to \mathbb{R}^n$, donde $n \geq 2$. En general nos referimos a su imagen que es un conjunto de $\mathbb{R}^n$ que se parametriza con una sola variable. Por ejemplo

$$\alpha(t) = (\cos(t), \text{sen}(t))$$

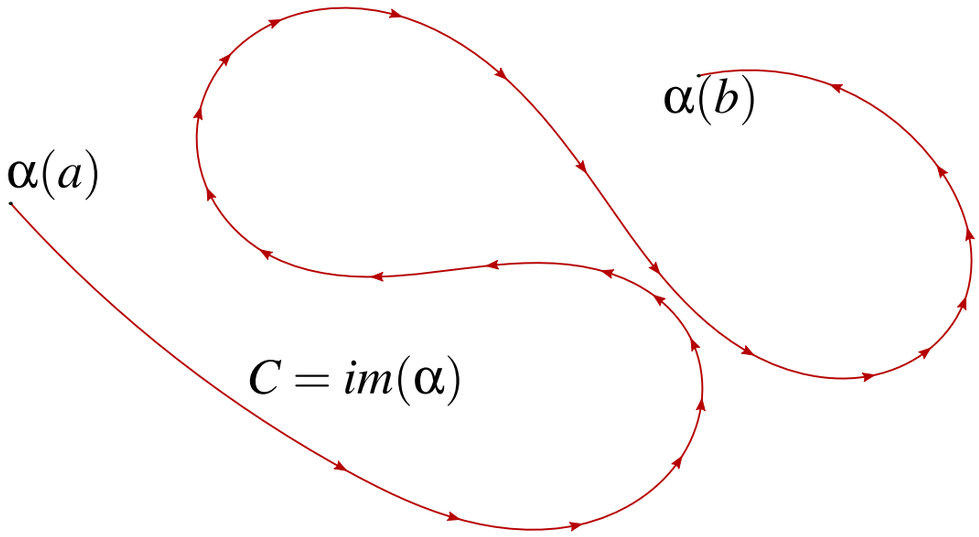


Figura 4.3: Curva $C \subset \mathbb{R}^2$

es una curva en $\mathbb{R}^2$, su imagen es una circunferencia. Mientras que

$$\alpha(t) = (2t, 3t, -t)$$

es una curva en $\mathbb{R}^3$, su imagen es una recta.
En general escribimos las curvas en $\mathbb{R}^2$ como

$$\alpha(t) = (x(t), y(t))$$

donde $x : I \to \mathbb{R}$ y también $y : I \to \mathbb{R}$ son dos funciones de una variable. Por ejemplo en el caso anterior $x(t) = \cos(t)$, mientras que $y(t) = \text{sen}(t)$.
Para las curvas en $\mathbb{R}^3$ podemos usar $\alpha(t) = (x(t), y(t), z(t))$ donde cada coordenada es una función de $\mathbb{R}$ en $\mathbb{R}$, por ejemplo en el segundo caso de arriba $x(t) = 2t$, $y(t) = 3t$, $z(t) = -t$.

DEFINICIÓN 4.1.19 (Continuidad de curvas)**.** Una curva $\alpha : I \to \mathbb{R}^n$ dada por

$$\alpha(t) = (x_1(t), x_2(t), \ldots, x_n(t))$$

es continua si todas las funciones $x_i : I \to \mathbb{R}$ son continuas.

ƨ Podemos notar que una curva es continua en $t_0 \in \bar{I}$ cuando el límite de cada coordenada (en el punto t_0) coincide con evaluar cada coordenada en t_0.

DEFINICIÓN 4.1.20 (Conjuntos conexos)**.** Diremos que $A \subset \mathbb{R}^n$ es

conexo (más precisamente conexo por arcos o arco-conexo) si dados $P, Q \in A$ los podemos unir con una curva continua $\alpha : I \to \mathbb{R}^n$ tal que $im(\alpha) \subset A$. Si $I = [a,b]$ entonces debe ser $\alpha(a) = P$, $\alpha(b) = Q$ (o alrevés).

TEOREMA 4.1.21 (Bolzano)**.** *Sea $A \subset \mathbb{R}^n$ conjunto conexo, sea $f : A \to \mathbb{R}^n$ una función continua. Sean $P, Q \in A$. Entonces*

1. *Si $f(P) > 0$, $f(Q) < 0$ (o alrevés), entonces existe $R \in A$ tal que $f(R) = 0$.*

2. *Si $f(P) < f(Q)$ entonces la imagen de f toma todos los valores intermedios entre $f(P)$ y $f(Q)$. Explícitamente, $[f(P), f(Q)] \subset Im(f)$.*

ƨ No necesariamente el intervalo y la imagen de f son exactamente iguales, la imagen puede ser más grande.

TEOREMA 4.1.22 (Teorema de Weierstrass)**.** *Sea $K \subset \mathbb{R}^n$, sea $f : K \to \mathbb{R}$. Entonces f es acotada y alcanza máximo y mínimo en K.*

Este teorema requiere algunas explicaciones. Recordemos para empezar que un conjunto es compacto cuando es cerrado y acotado.

■ Como $f : \mathbb{R}^n \to \mathbb{R}$, la imagen es un subconjunto de números reales. Decimos que f es acotada en K si el conjunto imagen de f con dominio K es un conjunto acotado.

■ Usamos $f|_K$ para indicar la restricción de f al conjunto K. Entonces f es acotada si $A = im(f|_K) \subset \mathbb{R}$ es acotado.

■ Decimos que f alcanza máximo y mínimo en $K \subset \mathbb{R}$ cuando la imagen de f es un conjunto compacto. Porque en ese caso, si tomamos

$$M = \text{máx}\{f(x) : x \in K\} \in \mathbb{R},$$

-que existe porque la imagen es compacta- existe un punto $P_M \in K$ tal que $f(P_M) = M$.

■ El *valor máximo* de f en K es $M = f(P_M)$, mientras que el máximo se puede *alcanzar* en uno o más puntos de K -el punto P_M no tiene por qué ser único, ver la Figura 4.4-. Entonces el valor máximo es

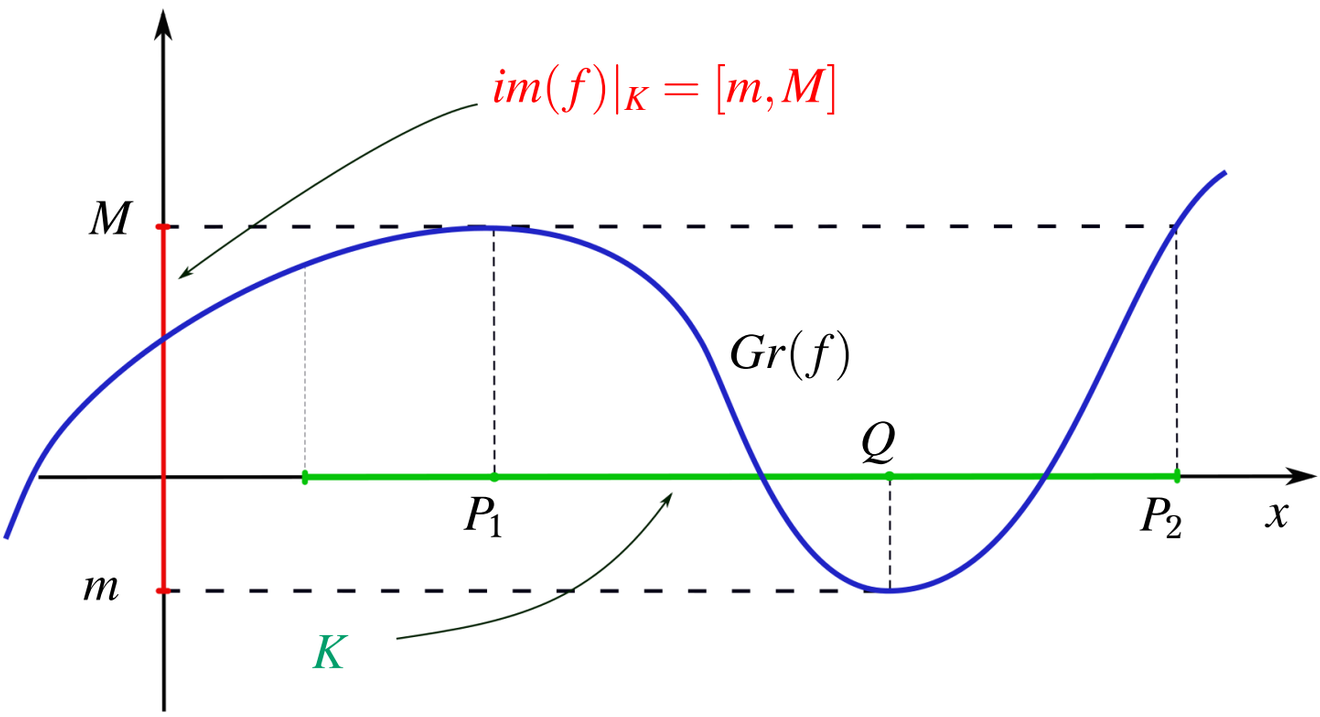


Figura 4.4: Q es el único mínimo de $f|_K$, mientras que P_1, P_2 son máximos de $f|_K$

único, pero los "máximos" de f (los puntos donde hay que evaluar para conseguirlo) pueden ser varios.

■ Similarmente, el *valor mínimo* de f en K que menciona el teorema, es $m = f(Q_m) \in \mathbb{R}$ para algún $X_m \in K$, ya que el conjunto imagen de $f|_K$ tiene mínimo m. Entonces el valor mínimo es único, pero los "mínimos" de f (los puntos donde hay que evaluar para conseguirlo) pueden ser varios.

■ Notemos que si bien $[m,M] \subset im(f)$, si restringimos f a K, la imagen es exactamente ese intervalo: $[m,M] = im(f|_K)$.

4.2. Derivadas parciales

• Corresponde a clases en video 4.5a, 4.5b, 4.6a, 4.6b, 4.7a, 4.7b

Comenzamos repasando la noción de derivada en una variable.

DEFINICIÓN 4.2.1 (Cociente incremental y derivada)**.** Sea $I \subset \mathbb{R}$, sea $f : I \to \mathbb{R}$ y sea $x \in I^o$. La derivada de f respecto de x se obtiene calcu-

lando el límite de los cocientes incrementales cuando $h \to 0$:

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

(siempre que el límite exista). Equivalentemente, haciendo el cambio de variable $y = x + h$

$$f'(x) = \lim_{y\to x} \frac{f(y) - f(x)}{y - x}.$$

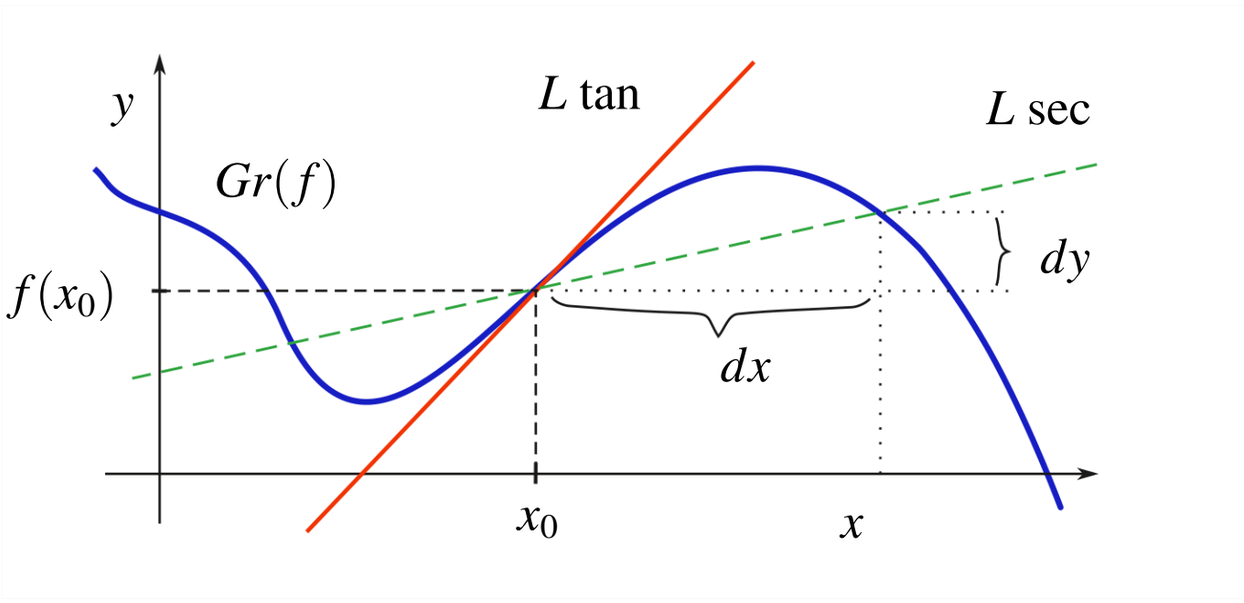


Figura 4.5: La recta tangente se obtiene como límite de rectas secantes

Cuando el límite existe y da un número decimos que f *es derivable en* x. Para que el límite exista el numerador debe tender a cero, luego es condición necesaria para que f sea derivable que

$$\lim_{y\to x} f(y) = f(x),$$

y entonces el límite coincide con evaluar. Enunciamos esto como teorema:

TEOREMA 4.2.2 (Derivable implica continua)**.** *Si f es derivable en x entonces f es continua en x.*

ƨ La afirmación recíproca no es cierta, por ejemplo la función módulo $f(x) = |x|$ es continua en $x = 0$ pero no es derivable allí, ya que

$$\lim_{h\to 0^+} \frac{f(0+h) - f(0)}{h} = \lim_{h\to 0^+} \frac{h - 0}{h} = 1$$

mientras que

$$\lim_{h\to 0^-} \frac{f(0+h)-f(0)}{h} = \lim_{h\to 0^-} \frac{-h-0}{h} = -1.$$

Esto dice que los límites laterales *del cociente incremental* de f en $x = 0$ son distintos, y entonces no existe el límite *del cociente incremental*. Por eso $f(x) = |x|$ no es derivable en $x = 0$.

2 Como se aprecia en la Figura 4.5, el cociente incremental

$$\frac{dy}{dx} = \frac{f(x)-f(x_0)}{x-x_0}$$

representa la pendiente de la recta *secante* que pasa por $P = (x_0, f(x_0))$ y por $Q = (x, f(x))$. Lo que queremos es ver qué pendiente tiene la recta en el límite cuando $x \to x_0$, y ese número (el límite de los cocientes incrementales) es la derivada.

OBSERVACIÓN 4.2.3 (Recta tangente)**.** Cuando f es derivable en x_0 la derivada representa la pendiente de la recta tangente al gráfico de f en ese punto. La ecuación de la recta tangente es

$$y = f'(x_0)(x-x_0) + f(x_0).$$

La recta tangente es la recta que pasa por el punto $(x_0, f(x_0))$ que mejor aproxima al gráfico de f cerca de ese punto.

PROPOSICIÓN 4.2.4 (Derivadas de funciones conocidas)**.** *De la definición de derivada y propiedades algebraicas de cada función, es posible deducir las siguientes fórmulas:*

Tabla de derivadas $(d1)$

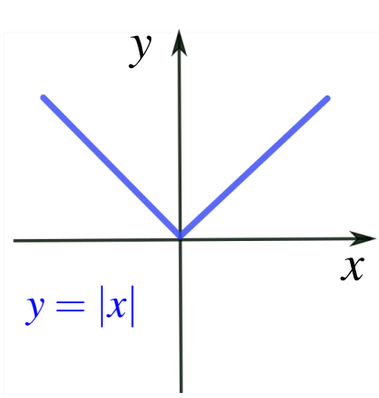
	$Dom(f)$	$f(x)$	$f'(x)$
$n \in \mathbb{N}_0$	$\mathbb{R}$	x^n	nx^{n-1}
$k \in \mathbb{Z}_{<0}$	$\mathbb{R}_{\neq 0}$	x^k	kx^{k-1}
$\alpha \in \mathbb{R}$	$x > 0$	x^α	$\alpha x^{\alpha-1}$
	$x > 0$	$\ln(x)$	$1/x$
	$\mathbb{R}$	e^x	e^x
$a \neq 0$	$\mathbb{R}$	a^x	$a^x \ln(a)$
	$\mathbb{R}$	$\operatorname{sen}(x)$	$\cos(x)$
	$\mathbb{R}$	$\cos(x)$	$-\operatorname{sen}(x)$

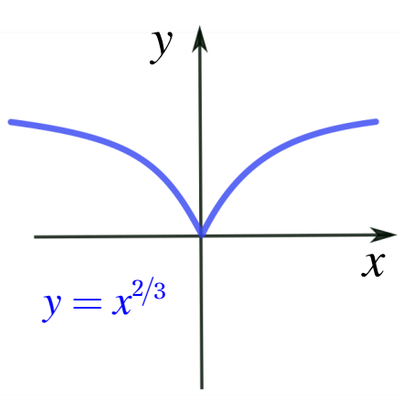
2 De la tabla (y algo de sentido común) notamos que el dominio de la derivada coincide con el de la función, ya que para poder derivar tenemos que tener función. Por ejemplo la fórmula $1/x$ para la derivada del logaritmo a priori tiene dominio todos los números no nulos, pero a posteriori (sabiendo que la estamos pensando como la derivada de ln) tiene dominio sólo $x > 0$.

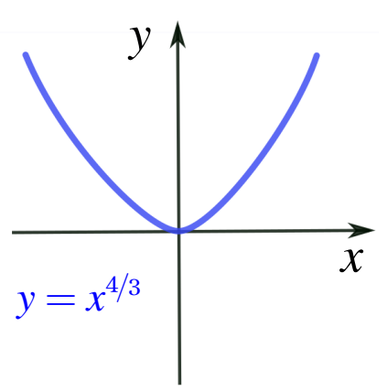
2 En algunos casos el dominio de la derivada es estrictamente más chico que el de la función original, por ejemplo

1. $f(x) = |x|$. Entonces $Dom(f) = \mathbb{R}$ pero $Dom(f') = \mathbb{R} \setminus \{0\}$.

2. $f(x) = x^{2/3}$. Entonces $Dom(f) = \mathbb{R}$ pero $Dom(f') = \mathbb{R} \setminus \{0\}$, puesto que
$$f'(x) = {}^2\!/_3 x^{-1/3} = {}^2\!/_3 \frac{1}{x^{1/3}}.$$







3. En general el dominio de $f(x) = x^{p/q}$ es todo $\mathbb{R}$ si la fracción es no negativa y q es impar, pero el dominio de la derivada no tiene al $x = 0$ cuando $p/q < 1$.

OBSERVACIÓN 4.2.5 (Funciones derivables)**.** De los ejemplos y la definición vemos que para calcular la derivada, tiene sentido hacerlo cuando el punto x_0 es interior. Entonces de aquí en más nos concentraremos en derivar funciones f en dominios abiertos. Dado $A \subset \mathbb{R}$ abierto, diremos que f es derivable en A si existe su derivada en todos y cada uno de los puntos $x \in A$.

PROPIEDADES 4.2.6 (Reglas de derivación-una variable)**.** Reglas para derivar productos sumas y composiciones. Sea $A \subset \mathbb{R}$ abierto, sean f, g derivables en $A \subset \mathbb{R}$. Entonces

1. $f + g$ es derivable en A y vale $(f+g)' = f' + g'$.
2. $f \cdot g$ es derivable en A y vale $(fg)' = f'g + fg'$.
3. Si $\lambda \in \mathbb{R}$ entonces $(\lambda f)' = \lambda f'$.
4. Si g no se anula en A entonces f/g es derivable en A y vale
$$\left(\frac{f}{g}\right)' = \frac{f' \cdot g - f \cdot g'}{g^2}.$$
5. *Regla de la cadena:* si se puede hacer la composición $g \circ f$, entonces es derivable y
$$(g \circ f)'(x) = g'(f(x)) \cdot f'(x).$$
6. *Derivada de la función inversa:* si $f : A \to B$ es biyectiva y con derivada no nula en A entonces $f^{-1} : B \to A$ es derivable y vale
$$\left(f^{-1}\right)'(y) = \frac{1}{f'(f^{-1}(y))}$$

para todo $y \in B$.

2 Simplemente cambiando las letras, si puede hacerse la otra composición (en el otro orden) $f \circ g$ y ambas son derivables entonces la composición es derivable y

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x).$$

También puede reescribirse la fórmula de la derivada de la inversa notando que $y = f(x)$, entonces

$$\left(f^{-1}\right)'(f(x)) = \frac{1}{f'(x)}$$

PROPIEDADES 4.2.7 (Derivadas)**.** Combinando las propiedades mencionadas con las derivadas de la Tabla (d1), obtenemos las derivadas de otras funciones que aparecen repetidamente en los ejemplos y aplicaciones. La tabla no indica los dominios ni de f ni de f', eso queda a cargo del lector:

Tabla de derivadas $(d2)$

$f(x)$	$f'(x)$
$\tan(x)$	$\frac{1}{\cos^2(x)}$
$e^{g(x)}$	$e^{g(x)}g'(x)$
$\ln(g(x))$	$\frac{g'(x)}{g(x)}$
$\operatorname{arc\,sen}(x)$	$\frac{1}{\sqrt{1-x^2}}$
$\operatorname{arc\,cos}(x)$	$\frac{-1}{\sqrt{1-x^2}}$
$\operatorname{arctan}(x)$	$\frac{1}{1+x^2}$
$\cosh(x)$	$\operatorname{senh}(x)$
$\operatorname{senh}(x)$	$\cosh(x)$

Sobre el final están las *funciones trigonométricas hiperbólicas*:

$$\cosh(x) = \frac{e^x + e^{-x}}{2}, \qquad \operatorname{senh}(x) = \frac{e^x - e^{-x}}{2}.$$

Notar que a diferencia de las usuales, no están acotadas y tampoco cambia el signo al derivar. También cumplen una relación similar a la pitagórica (pero con otro signo):

$$\cosh^2(x) - \operatorname{senh}^2(x) = 1.$$

Notemos que $\operatorname{senh}(0) = 0$; como $\cosh^2(x) = 1 + \operatorname{senh}^2(x) \geq 1$ el coseno hiperbólico no se anula nunca.

TEOREMA 4.2.8 (Teoremas del cálculo diferencial en una variable)**.** *Sea $f : [a,b] \to \mathbb{R}$ continua. Supongamos que f es derivable en (a,b). Entonces*

■ *Teorema de Rolle: Si $f(a) = f(b)$, existe $c \in (a,b)$ tal que $f'(c) = 0$.*

■ *Teorema de Lagrange: En general existe $c \in (a,b)$ tal que*

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

COROLARIO 4.2.9 (Crecimiento y decrecimiento)**.** *Sea $f : I \to \mathbb{R}$ derivable con I un intervalo abierto, entonces*

1. *Si $f' \geq 0$ en I, f es creciente en I. Esto es*

$$x \leq y \Longrightarrow f(x) \leq f(y).$$

   Una función constante es creciente (admitimos derivada nula).

2. *Si $f' > 0$ en I, f es estrictamente creciente en I. Esto es*

$$x < y \Longrightarrow f(x) < f(y),$$

   en este caso no se admiten funciones constantes.

3. *Si $f' \leq 0$ en I, f es decreciente en I. Esto es*

$$x \leq y \Longrightarrow f(x) \geq f(y).$$

Una función constante es decreciente (admitimos derivada nula).

4. *Si $f' < 0$ en I, f es estrictamente decreciente en I. Esto es*

$$x < y \Longrightarrow f(x) > f(y),$$

en este caso no se admiten funciones constantes.

Las funciones con derivada no nula en un intervalo son inyectivas, ya que son estrictamente monótonas (creciente o decreciente).

Las funciones crecientes preservan desigualdades, las decrecientes las invierten. Las que no son ni crecientes ni decrecientes (como por ejemplo $f(x) = x^2$) se llevan mal con las desigualdes.

DEFINICIÓN 4.2.10 (Derivadas sucesivas)**.** Si $f : A \to \mathbb{R}$ es derivable en A, podemos hallar una fórmula $f'(x)$ para todo $x \in A$. Obtuvimos así una nueva función $f' : A \to \mathbb{R}$. Nos podemos preguntar si es derivable en A. Si lo es, denotamos

$$f'' = (f')', \text{ es decir } f''(x) = (f'(x))'.$$

Así sucesivamente, la derivada n-ésima la denotamos $f^{(n)}(x)$ (usamos el paréntesis en el exponente para que no se confunda con la potencia usual que está relacionada con el producto).

DEFINICIÓN 4.2.11 (Funciones de clase C^k)**.** Si f es derivable k veces *y además su derivada k-ésima es continua*, decimos que f es de clase C^k.
Si queremos hacer explícito el dominio, denotamos $C^k(A)$ al conjunto de todas las funciones que son de clase C^k en el conjunto $A \subset \mathbb{R}^n$.
Decimos que f es de clase C^∞ si existen todas las derivadas sucesivas de todos los órdenes.

Todos los polinomios, las funciones seno y coseno, la función exponencial son de clase C^∞. La función $f(x) = x^{7/3}$ es de clase C^2 pero no es de clase C^3: calculamos

$$f'(x) = \frac{7}{3}x^{4/3}, \quad f''(x) = \frac{7}{3}\cdot\frac{4}{3}x^{1/3}, \quad f'''(x) = \frac{7}{3}\cdot\frac{4}{3}\cdot\frac{1}{3}x^{-1/3},$$

vemos que f''' ni siquiera está definida en $x = 0$.

Una f es de clase C^1 si existe f' y además f' es una función continua. Si f no es derivable no puede ser C^1, pero hay funciones derivables que no son C^1. Por ejemplo:

EJEMPLO 4.2.12 (Una función derivable que no es C^1)**.** Sea $f : \mathbb{R} \to \mathbb{R}$

$$f(x) = \begin{cases} x^2 \operatorname{sen}(1/x) & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Calculamos su derivada fuera de $x = 0$ y obtenemos

$$x^2 \operatorname{sen}(1/x)' = 2x \operatorname{sen}(1/x) + x^2 \cos(1/x)\frac{-1}{x^2} = 2x \operatorname{sen}(1/x) - \cos(1/x).$$

Calculamos su derivada en $x = 0$ usando la definición

$$f'(0) = \lim_{x\to 0} \frac{x^2 \operatorname{sen}(1/x) - 0}{x - 0} = \lim_{x\to 0} x \operatorname{sen}(1/x) = 0$$

donde usamos la propiedad "0 por acotada=0". Entonces f es derivable en todo $\mathbb{R}$, con derivada

$$f'(x) = \begin{cases} 2x \operatorname{sen}(1/x) - \cos(1/x) & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Para ver que $f \notin C^1(\mathbb{R})$, basta ver que la función derivada no es continua en algún punto. Afirmamos que no es continua en $x = 0$. Sabemos que $f'(0) = 0$ por lo recién calculado. Por otro lado vemos que

$$\lim_{x\to 0} f'(x) = \lim_{x\to 0} 2x \operatorname{sen}(1/x) - \cos(1/x) = 0 - \lim_{x\to 0} \cos(1/x) = 0 -$$

entonces como no existe el límite de f' cuando x tiende a 0, f' no es continua y así f no es una función de clase C^1.

ƨ Como el único problema está en $x = 0$, podemos decir que f es C^1 en todo $\mathbb{R}$ sin el 0. Esto es, $f \in C^1(\mathbb{R} \setminus \{0\})$, pero $f \notin C^1(\mathbb{R})$.

Funciones de varias variables

DEFINICIÓN 4.2.13 (Derivadas parciales)**.** La derivada parcial de $\frac{\partial}{\partial x}$ respecto de x de una función $f(x,y)$ se obtiene pensando que x es variable mientras que y está fijo (es constantes).
Por ejemplo si $f(x,y) = x^2 y \cos(y^2+3x)$, tenemos

$$\frac{\partial f}{\partial x}(x,y) = 2xy\cos(y^2+3x) - x^2 y\,\text{sen}(y^2+3x)\cdot 3.$$

La derivada parcial respecto de y se obtiene en cambio fijando la variable x y pensando a la función como únicamente de la variable y. Entonces en el mismo ejemplo

$$\frac{\partial f}{\partial x}(x,y) = x^2\cos(y^2+3x) - x^2 y\,\text{sen}(y^2+3x)\cdot 2y.$$

DEFINICIÓN 4.2.14 (Vector gradiente)**.** Si existen ambas derivadas parciales en el punto $P = (x_0, y_0)$, el vector

$$\nabla f(P) = (\frac{\partial f}{\partial x}(P), \frac{\partial f}{\partial y}(P))$$

se denomina *vector gradiente a f en el punto P*. El símbolo ∇ es la letra griega *nabla*.

ƨ Cuando la función tiene 3 (o más variables) la definición es la misma: se fijan todas las demás variables -se las piensa como constantes- y se deriva respecto de ella. Así por ejemplo

$$\frac{\partial}{\partial z}\left(x^3z^2 + \ln(z^2+y^3)\right) = 2x^3z + \frac{1}{(z^2+y^3)}\cdot 2z.$$

En este caso como la función tiene tres variables el vector gradiente tendrá 3 coordenadas,

$$\nabla f(x_0,y_0,z_0) = (\frac{\partial f}{\partial x}(x_0,y_0,z_0), \frac{\partial f}{\partial y}(x_0,y_0,z_0), \frac{\partial f}{\partial z}(x_0,y_0,z_0))$$

Si pensamos la derivada por definición, vemos que

$$\frac{\partial f}{\partial x}(x_0,y_0) = \lim_{x\to x_0} \frac{f(x,y_0)-f(x_0,y_0)}{x-x_0}$$

(fijamos $y = y_0$, movemos x), mientras que

$$\frac{\partial f}{\partial y}(x_0,y_0) = \lim_{y\to y_0} \frac{f(x_0,y)-f(x_0,y_0)}{y-y_0}.$$

(fijamos $x = x_0$, movemos y). Mismas consideraciones para funciones de más variables. Entonces las derivadas parciales en el punto pueden existir o no, puede existir una si y la otra no, etc.

Una vez calculada una derivada parcial de f, obtenemos una nueva función con la misma cantidad de variables. En realidad, si existen las derivadas parciales, tenemos una función para cada variable, las denotamos

$$\frac{\partial f}{\partial x}, \qquad \frac{\partial f}{\partial y}, \qquad \frac{\partial f}{\partial z}, \quad etc.$$

Para abreviar podemos escribir

$$\frac{\partial f}{\partial x}(P) = f_x(P), \qquad \frac{\partial f}{\partial y}(P) = f_y(P), \quad etc.$$

DEFINICIÓN 4.2.15 (Funciones de clase C^1)**.** Sea $A \subset \mathbb{R}^n$ abierto y $f : A \to \mathbb{R}$. Una función es de clase C^1 en A -denotado $f \in C^1(A)$ cuando

1. Existen todas las derivadas parciales en todos los puntos de A.

2. Todas las funciones derivadas parciales $\frac{\partial f}{\partial x^i}$ son funciones continuas en A.

Se extiende esta definición a las derivadas sucesivas, decimos que $f \in C^k(A)$ si existen todas las derivadas parciales sucesivas de f hasta orden k y son todas funciones continuas.

TEOREMA 4.2.16 ($C^1 \Rightarrow$ continua)**.** *Sea $f : A \to \mathbb{R}$ con $A \subset \mathbb{R}^n$ abierto. Si $f \in C^1(A)$, entonces f es continua en A.*

No probaremos este teorema, pero hacemos una observación importante.

Que existan las derivadas parciales no es suficiente para garantizar la continuida de una función de 2 (o más) variables. Para poder asegurarlo, es necesario que además de existir, sean funciones continuas.

OBSERVACIÓN 4.2.17 (Derivadas parciales y crecimiento)**.** Las derivadas parciales de f indican cuánto crece la función f en la dirección del respectivo eje.

DEFINICIÓN 4.2.18 (Plano tangente)**.** Si $f : A \to \mathbb{R}$ es de clase C^1 (con $A \in \mathbb{R}^2$), el *plano tangente* a f en el punto $P = (x_0, y_0) \in A$ es el plano de ecuación

$$z = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + f(x_0, y_0).$$

Este es el plano (Figura 4.6) que pasa por el punto $(x_0, y_0, f(x_0, y_0))$ del gráfico de f y tiene la propiedad de ser el que mejor aproxima a dicho gráfico (cerca del punto mencionado).

OBSERVACIÓN 4.2.19 (Plano tangente como aproximación lineal)**.** Sea

$$f(x, y) = 1 + e^{x+y-7x^2-9y^2}.$$

Las derivadas parciales de f son

$$f_x(x, y) = e^{x+y-7x^2-9y^2}(1 - 14x), \qquad f_y(x, y) = e^{x+y-7x^2-9y^2}(1 - 18y).$$

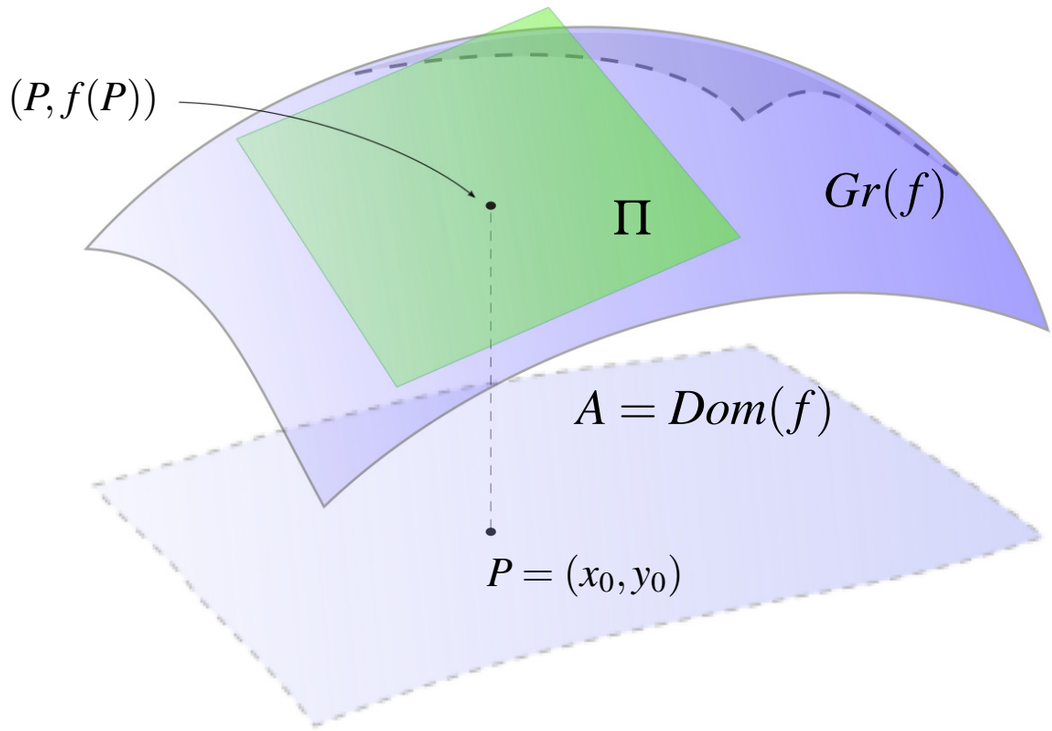


Figura 4.6: Plano Π *tangente al gráfico de* f *en* $(P, f(P))$

Como $f(0,0) = 1+1 = 2$, $f_x(0,0) = 1$, $f_y(0,0) = 1$, entonces el plano tangente a f en $P = (0,0)$ es

$$z = x + y + 2 = \pi(x,y).$$

Si pensamos π como función, su gráfica es el plano, y esta función lineal aproxima a f cerca de $P = (0,0)$, denotamos esto como:

$$\pi(x,y) \simeq f(x,y) \qquad si \quad (x,y) \simeq (0,0).$$

La diferencia es lo que se conoce como *error* de la aproximación. Podemos ver una representación gráfica de f y su plano tangente en $P = (0,0)$ en el applet de Geogebra (click).

2 Más variables: cuando la función tiene más de dos variables, no esperamos representarla gráficamente. Pero las ideas de derivadas parciales son idénticas, y lo mismo ocurre con la idea de aproximación lineal. Entonces por ejemlo si $f : \mathbb{R}^3 \to \mathbb{R}$, su gráfico es el conjunto de $\mathbb{R}^4$ dado por

$$Gr(f) = \{(x,y,z,f(x,y,z)) : (x,y,z) \in Dom(f)\} \subset \mathbb{R}^4,$$

que se puede describir por la ecuación implícita $w = f(x,y,z)$. Vamos a definir su *plano tangente* en el punto $(x_0,y_0,z_0) \in Dom(f)$ como el subespacio de $\mathbb{R}^4$ dado por la ecuación implícita

$$w = f_x(x_0,y_0,z_0)(x-x_0) + f_y(x_0,y_0,z_0)(y-y_0) + f_z(x_0,y_0,z_0)(z-z_0) + f(x_0,y_0,z_0)$$

o abreviadamente $w = \Pi(x,y,z)$. El nombre *plano* en realidad es formalmente incorrecto ya que la dimensión de este subespacio es 3 y no 2. Pero la idea de aproximación numérica sigue siendo válida: para (x,y,z) cercano a (x_0,y_0,z_0), se tiene

$$f(x,y,z) \simeq \Pi(x,y,z).$$

Volvamos ahora a las derivadas y su interpretación. Nuevamente trabajamos con funciones de dos variables, pero estas ideas se pueden extender a 3 o más de ellas.

Si queremos encontrar cuánto crece f en alguna otra dirección que no sea de un eje, tomamos $V = (v_1,v_2)$ de norma unitaria (porque queremos normalizar, ya que los vectores canónicos miden 1) y definimos

DEFINICIÓN 4.2.20 (Derivadas direccionales)**.** Sea $A \subset \mathbb{R}^2$ abierto, sea $f : A \to R$. Decimos (si existe) que el límite

$$\frac{\partial f}{\partial V}(x_0,y_0) = \lim_{t\to 0} \frac{f(x_0+tv_1, y_0+tv_2) - f(x_0,y_0)}{t}$$

es la *derivada direccional* de f en $X_0 = (x_0,y_0)$ en la dirección de $V = (v_1,v_2)$ -con $\|V\| = 1$-. Podemos escribirlo sin coordenadas como

$$\frac{\partial f}{\partial V}(X_0) = \lim_{t\to 0} \frac{f(X_0+tV) - f(X_0)}{t} = \lim_{X_t \to X_0} \frac{f(X_t) - f(X_0)}{\|X_t - X_0\|},$$

donde $X_t = X_0 + tV$.

OBSERVACIÓN 4.2.21 (Interpretación geométrica como pendiente)**.** La variación de f en el cociente de la derivada direccional es el numerador. La variación en el dominio es

$$dist(X_0, X_0+tV) = \|X_0+tV-X_0\| = |t|\,\|V\| = |t|$$

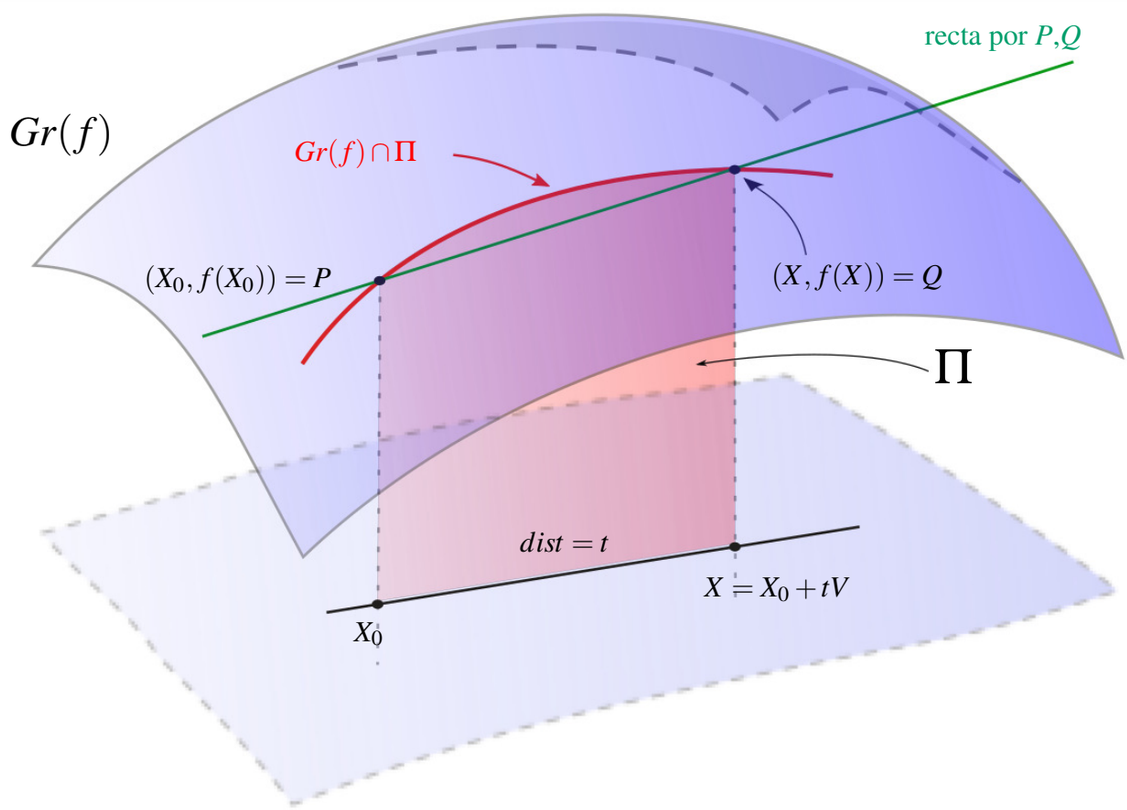


Figura 4.7: Corte de $Gr(f)$ con un plano vertical Π sobre la recta $L : X_0 + tV$

porque $\|V\| = 1$. Entonces el cociente incremental de la derivada direccional representa la variación de las alturas (imágenes de f) dividido por la variación en el dominio (el parámetro t).

Si cortamos $Gr(f)$ con un plano vertical Π sobre la recta $L : X_0 + tV$ (Figura 4.7), y miramos la figura de perfil, vemos el dibujo para una función de una variable (Figura 4.5): la recta que une P con Q en el dibujo de arriba es la recta secante a la intersección de $Gr(f)$ con Π.

Su pendiente indica una aproximación a la inclinación de $Gr(f)$ en esa dirección, y al tomar límite obtenemos el número exacto de la inclinación, que es la derivada direccional $\frac{\partial f}{\partial V}(X_0)$. Para cada dirección, este número puede ser distinto (es como pararse en una montaña y mirar en todas las direcciones, dependiendo de hacia donde miramos se sube más o menos, se baja, se permanece horizontal, etc.)

ƨ En particular si tomamos $V = E_1 = (1,0)$, tiene norma 1 y se tiene

$$\frac{\partial f}{\partial V}(x_0, y_0) = \lim_{t\to 0} \frac{1}{h} f(x_0 + h, y_0) - f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0),$$

la derivada parcial respecto de x es un caso particular de derivada direccional. Similarmente, la derivada parcial respecto de y es otro caso particular de derivada direccional (ahora tomando $V = E_2 = (0,1)$).

ꙅ Como ya mencionamos, no alcanza que existan las derivadas parciales para que una función de dos variables sea continua, y de hecho, tampoco alcanza con que existan todas las derivadas direccionales.
Un ejemplo de una función que no es continua, pero para la cual existen todas las derivadas direccionales en el origen, está dado por

$$f(x,y) = \begin{cases} \dfrac{x^3 y}{x^6 + y^2} & \text{si } (x,y) \neq (0,0), \\ \\ 0 & \text{si } (x,y) = (0,0). \end{cases}$$

Esta función también tiene su gráfico de Geogebra en este link.

Veamos que ahora que para una función "buena" (de clase C^1), podemos calcular todas sus derivadas direccionales de manera bastante simple (sin calcular límites):

TEOREMA 4.2.22 (Gradiente y derivadas direccionales)**.** *Sea $A \subset \mathbb{R}^n$ abierto y sea $f \in C^1(A)$. Entonces para todo $V \in \mathbb{R}^n$ con $\|V\| = 1$ y todo $P \in A$ se tiene*

$$\frac{\partial f}{\partial V}(P) = \langle \nabla f(P), V \rangle = \nabla f(P) \cdot V.$$

El teorema anterior dice que podemos evitar calcular los límites para las demas direcciones, siempre que f sea de clase C^1. En particular notemos que si f tiene dos variables y escribimos $V = (V_1, V_2)$, entonces

$$\frac{\partial f}{\partial V}(x_0, y_0) = v_1 f_x(P) + v_2 f_y(P).$$

ꙅ Todas las derivadas direccionales (en el punto P) son combinación lineal de las dos derivadas parciales (en el punto P). Si esto no es cierto para algún punto o para algún V, entonces podremos afirmar que f no es de clase C^1.

TEOREMA 4.2.23 (Dirección de máximo crecimiento)**.** *Si $f : A \to \mathbb{R}$ es de clase C^1, la dirección de máximo crecimiento en el punto $P \in A$, es la dirección del gradiente en ese punto. Además para todo $V \in \mathbb{R}^n$, se tiene*

$$-\|\nabla f(P)\| \leq \frac{\partial f}{\partial V}(P) \leq \|\nabla f(P)\|.$$

Luego cuánto crece f (en todas las direcciones posibles, mirando desde P) es a lo sumo la magnitud dada por la longitud del gradiente de f en P.

ƨ El valor máximo de las derivadas parciales en un punto P dado es exactamente $\|\nabla f(P)\|$. Todas las demás derivadas direccionales son menores o iguales a este número.

ƨ Cuando decimos que la dirección de máximo crecimiento es la del gradiente, queremos decir que (suponiendo que $\nabla f(P) \neq 0$) hay que tomar como dirección

$$V = \frac{1}{\|\nabla f(P)\|} \nabla f(P).$$

Si $\nabla f(P) = 0$, todas las derivadas direccionales son nulas, entonces en cualquier dirección se crece a la misma velocidad.

4.3. Diferencial y regla de la cadena

• Corresponde a clases en video 4.8, 4.9

Comenzamos listando algunas propiedaes del gradiente que son fáciles de deducir del caso de una variable.

PROPIEDADES 4.3.1 (Reglas de derivación-varias variables)**.** Reglas para derivar productos sumas y composiciones. Sea $A \subset \mathbb{R}^n$ abierto, sean $f, g : A \to \mathbb{R}$ de clase C^1. Sea $P \in A$, entonces

1. $f + g \in C^1(A)$ y vale $\nabla(f+g)(P) = \nabla f(P) + \nabla g(P)$.
2. $f \cdot g \in C^1(A)$ y vale $\nabla(fg)(P) = g(p)\nabla f(P) + f(p)\nabla g(P)$.

3. Si $\lambda \in \mathbb{R}$ entonces $\nabla(\lambda f)(P) = \lambda \nabla f(P)$.

4. Si g no se anula en A entonces $f/g \in C^1(A)$ y vale

$$\nabla\left(f/g\right)(P) = \frac{g(P)\nabla f(P) - f(P)\nabla g(P)}{g(P)^2}.$$

Para la composición de funciones (y la regla de la cadena), es necesario considerar funciones a valores en $\mathbb{R}^k$, para poder luego aplicar otra función con dominio en $\mathbb{R}^k$. Para eso discutimos a continuación las nociones de campo vectorial, y su diferencial.

Dada $F : \mathbb{R}^2 \to \mathbb{R}^2$, podemos escribir $F(x,y) = (f_1(x,y), f_2(x,y))$ donde $f_i : \mathbb{R}^2 \to \mathbb{R}$ son funciones escalares. Decimos que F es un *campo vectorial* o más brevemente que F es un *campo*. Otras notaciones comunes son

$$F(u,v) = (x(u,v), y(u,v))$$

donde ahora $x : \mathbb{R}^2 \to R$ es una función escalar y lo mismo ocurre con $y : \mathbb{R}^2 \to \mathbb{R}$.
Si ahora $F : \mathbb{R}^3 \to \mathbb{R}^3$ podemos escribir

$$F(x,y,z) = (f_1(x,y,z), f_2(x,y,z), f_3(x,y,z))$$

con las f_i funciones escalares, es decir $f_i : \mathbb{R}^3 \to \mathbb{R}$.

La definición de campo se generaliza a $F : \mathbb{R}^n \to \mathbb{R}^k$, aunque el nombre campo se suele reservar para $n, k \geq 2$.
Cuando $k = 1$ tenemos una *función escalar*, $f : \mathbb{R}^n \to \mathbb{R}$.
Cuando $n = 1$, $k \geq 2$ tenemos una *curva*, $\alpha : \mathbb{R} \to \mathbb{R}^k$.

DEFINICIÓN 4.3.2 (Continuidad, C^k)**.** Decimos que el campo F es continuo si todas las funciones f_i son continuas. Decimos que el campo es de clase C^k si todas las funciones f_i son de clase C^k.

DEFINICIÓN 4.3.3 (Matriz Diferencial)**.** Si $F = (f_1, f_2)$ es un campo con dominio $A \subset \mathbb{R}^2$, y tanto f_1 como f_2 tiene derivadas parciales en A,

definimos

$$DF(P) = \begin{pmatrix} - & \nabla f_1(P) & - \\ - & \nabla f_2(P) & - \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x}(P) & \frac{\partial f_1}{\partial y}(P) \\ \frac{\partial f_2}{\partial x}(P) & \frac{\partial f_2}{\partial y}(P) \end{pmatrix},$$

que es una matriz 2×2 para cada $P \in A$.
Con la notación $F(u,v) = (x(u,v), y(u,v))$ se escribe (omitiendo el punto de evaluación (u,v)) como

$$DF = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

Cuando $F : \mathbb{R}^3 \to \mathbb{R}^3$ obtenemos una matriz 3×3.
Cuando $F : \mathbb{R} \to \mathbb{R}$, obtenemos un número (matriz de 1×1), la derivada usual.

¿Qué pasa si la cantidad de variables de salida es distinta de la de llegada? Obtenemos una matriz que no es cuadrada. Por ejemplo

1. Si $f : \mathbb{R}^3 \to \mathbb{R}$, hay una sola función, luego hay una sola fila y entonces

$$Df(P) = \nabla f(P) = (\frac{\partial f}{\partial x}(P), \frac{\partial f}{\partial y}(P), \frac{\partial x}{\partial z}(P)).$$

   Lo mismo ocurre para funciones de dos variables: el gradiente y la diferencial coinciden.

2. Si $\alpha : \mathbb{R} \to \mathbb{R}^k$ (es una curva), diferenciando obtenemos su *vector tangente*. Por ejemplo si $\alpha(t) = (x(t), y(t))$, se tiene

$$D\alpha(t) = \alpha'(t) = (x'(t), y'(t)).$$

   En rigor, habría que presentar este vector como *matriz columna*, ya que hay dos funciones y una sola variable.

3. Si $F : \mathbb{R}^2 \to \mathbb{R}^3$, entonces $F(u,v) = (x(u,v), y(u,v), z(u,v))$ luego al poner cada función como fila tenemos tres filas. Y como hay

dos variables tenemos dos columnas:

$$DF = \begin{pmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \\ \partial z/\partial u & \partial z/\partial v \end{pmatrix} \in \mathbb{R}^{3\times 2}.$$

4. En general entonces, si $F : \mathbb{R}^n \to \mathbb{R}^k$ tiene todas sus derivadas parciales, su diferencial en cada punto del dominio $P \in \mathbb{R}^n$ será una matriz de k filas (variables de llegada) y n columnas (variables de salida).

5. La matriz de la diferencial de F en P se suele denominar *matriz Jacobiana* de F, o simplemente *Jacobiano* de F.

TEOREMA 4.3.4 (Regla de la cadena)**.** *Sean F, G campos de clase C^k tales que se pueda hacer la composición $F \circ G$. Entonces $F \circ G$ es de clase C^k y además para cada P en su dominio, vale*

$$D(F \circ G)(P) = DF(G(P)) \cdot DG(P)$$

donde el producto es el de matrices.

OBSERVACIÓN 4.3.5 (Composición y producto de matrices)**.** Veamos algunos ejemplos para terminar este resumen. Como antes, en rojo las variables de salida, y en azul las de llegada.

1. Si $G : \mathbb{R}^3 \to \mathbb{R}^2$ mientras que $F : \mathbb{R}^2 \to \mathbb{R}^4$, entonces podemos hacer la composición $F \circ G : \mathbb{R}^3 \to \mathbb{R}^4$. Mediante la regla de la cadena vemos que

$$\underbrace{D(F \circ G)(P)}_{4\times 3} = \underbrace{DF(G(P))}_{4\times 2} \cdot \underbrace{DG(P)}_{2\times 3},$$

puesto que $DF(Q) \in \mathbb{R}^{4\times 2}$ mientras que $DG(P) \in \mathbb{R}^{2\times 3}$.

2. Ahora consideramos $f:\mathbb{R}^2\to\mathbb{R}$, con $G:\mathbb{R}^2\to\mathbb{R}^2$. Entonces podemos hacer la composición $f\circ G:\mathbb{R}^2\to\mathbb{R}$ y así

$$\nabla(f\circ G)(x,y)=\underbrace{D(f\circ G)(x,y)}_{1\times 2}=\underbrace{\nabla f(G(x,y))}_{1\times 2}\cdot\underbrace{DG(x,y)}_{2\times 2}.$$

3. En cambio si $\alpha:\mathbb{R}\to\mathbb{R}^2$ es una curva, y $f:\mathbb{R}^2\to\mathbb{R}$ es una función escalar, entonces $g=f\circ\alpha:\mathbb{R}\to\mathbb{R}$, y así

$$g'(t)=\underbrace{D(f\circ\alpha)(t)}_{1\times 1}=\underbrace{\nabla f(\alpha(t))}_{1\times 2}\cdot\underbrace{\alpha'(t)}_{2\times 1}=\langle\nabla f(\alpha(t)),\alpha'(t)\rangle.$$

4. Si escribimos $\alpha(t)=(x(t),y(t))$ podemos escribir $f(u,v)$ - cambiamos el nombre de las variables de f para que no se confundan con las de α -. Entonces la última ecuación se reescribe así:

$$(f\circ\alpha)'=\langle(\frac{\partial f}{\partial u},\frac{\partial f}{\partial v});(x',y')\rangle,$$

luego

$$(f\circ\alpha)'=\frac{\partial f}{\partial u}\cdot x'+\frac{\partial f}{\partial v}\cdot y'.$$

En esta expresión hay que sobreentender que x',y' están evaluados en t, mientras que las derivadas parciales de f están evaluadas en $(x(t),y(t))$.

5. Es común en la literatura ver abusos de notación donde esta última ecuación se escribe así:

$$(f\circ\alpha)'=\frac{\partial f}{\partial x}\cdot x'+\frac{\partial f}{\partial y}\cdot y'.$$

Esto no es un problema si uno entiende del contexto que primero hay que derivar f respecto de sus variables y luego componer con las que dependen del parámetro t.

EJEMPLO 4.3.6 (del uso de la regla de la cadena)**.** Sin multiplicar matrices, y teniendo en cuenta los ejemplos recién vistos, si tenemos

una expresión del tipo

$$g(x,y) = f(x^2+3y, 2x^3-y^2),$$

se trata de la composición de un campo $G : \mathbb{R}^2 \to \mathbb{R}^2$ (dentro de f), con la función $f : \mathbb{R}^2 \to \mathbb{R}$.

Podemos calcular las derivadas parciales de g respecto de x e y sin necesidad de calcular la matriz diferencial de G, de la siguiente manera:

$$\begin{aligned}\frac{\partial g}{\partial x} &= \frac{\partial f}{\partial u}(x^2+3y, 2x^3-y^2)\cdot\frac{\partial}{\partial x}(x^2+3y) + \frac{\partial f}{\partial v}(x^2+3y, 2x^3-y^2)\cdot\frac{\partial}{\partial x}(2x^3-y^2)\\ &= \frac{\partial f}{\partial u}(x^2+3y, 2x^3-y^2)\cdot 2x + \frac{\partial f}{\partial v}(x^2+3y, 2x^3-y^2)\cdot 6x^2\end{aligned}$$

De manera similar, calculamos la otra derivada parcial

$$\begin{aligned}\frac{\partial g}{\partial y} &= \frac{\partial f}{\partial u}(x^2+3y, 2x^3-y^2)\cdot\frac{\partial}{\partial y}(x^2+3y) + \frac{\partial f}{\partial v}(x^2+3y, 2x^3-y^2)\cdot\frac{\partial}{\partial y}(2x^3-y^2)\\ &= \frac{\partial f}{\partial u}(x^2+3y, 2x^3-y^2)\cdot 3 + \frac{\partial f}{\partial v}(x^2+3y, 2x^3-y^2)\cdot(-2y)\end{aligned}$$

Es usual también ver esta cuenta descripta de la siguiente manera, sin las variables auxiliares u, v (y requiere interpretarla como lo hicimos recién): las derivadas parciales de $f(x^2+3y, 2x^3-y^2)$ son

$$\frac{\partial f}{\partial x}(x^2+3y, 2x^3-y^2)\cdot 2x + \frac{\partial f}{\partial y}(x^2+3y, 2x^3-y^2)\cdot 6x^2$$

$$\frac{\partial f}{\partial x}(x^2+3y, 2x^3-y^2)\cdot 3 + \frac{\partial f}{\partial y}(x^2+3y, 2x^3-y^2)\cdot(-2y).$$

Capítulo 5

Taylor, integración y ecuaciones diferenciales

5.1. Polinomio de Taylor

• Corresponde a clases en videos 5.1a, 5.1b, 5.2a, 5.2b

Para una función derivable $f : \mathbb{R} \to \mathbb{R}$, la derivada primera en un punto nos permitía saber qué pendiente tiene la gráfica de f (en ese punto).

Podíamos también aproximar el valor de $f(x)$ -para x cercano a x_0- con el valor de la recta tangente. Para ello recordemos que la ecuación de la recta tangente en x_0 es

$$L_{tg} : y = f(x_0) + f'(x_0)(x - x_0),$$

que también podemos escribir como $\ell_{tg}(x) = f(x_0) + f'(x_0)(x - x_0)$.

Esto nos permite escribir

$$f(x) \simeq f'(x_0)(x - x_0) + f(x_0) \tag{5.1.1}$$

donde el símbolo $\simeq$ se lee "aproximadamente igual". Esta aproximación es una igualdad exacta cuando $x = x_0$.

Pero también es exacta *si f inicialmente era una función lineal*. Ya que si f es una recta, su derivada es constante, y entonces su recta tangente (en cualquier punto) es *exactamente igual* a la gráfica de f.

Cuando f no es lineal ¿cuál es la diferencia entre la recta tangente y la gráfica de f? En otros términos ¿cuál es el error cometido cuando usamos la aproximación (5.1.1)? Este error depende de cuánto nos alejemos de x_0. Mientras más cercanos estemos a x_0, más pequeño será el error, vamos a llamarlo $R(x)$. Es claro que

$$R(x) = f(x) - \ell_{tg}(x) = f(x) - f(x_0) - f'(x_0)(x - x_0).$$

Cabe aclarar que esta la fórmula del error (también llamada *resto de orden* 1) depende de f.

También, por supuesto, depende del punto x_0. Ya que si cambiamos el punto, cambia el valor de f y el de f' en x_0. Uno podría escribir algo como $R(f, x_0)(x)$ para indicar esto, pero en general no será necesario ya que una vez elegida la función, fijaremos el punto y estaremos trabajando con el resto para esa f y ese punto concreto.

Dados x, x_0 en un intervalo decimos que c está entre x, x_0 cuando c es alguno de los puntos intermedios entre x y x_0.
Notemos que si $x < x_0$, entonces $c \in (x, x_0)$, mientras que si $x > x_0$, tendremos $c \in (x_0, x)$. Podemos abreviar esto escribiendo $c \in \overline{xx_0}$, que podemos leerlo como que c está en el segmento entre x y x_0.

TEOREMA 5.1.1 (Fórmula del resto de orden 1)**.** *Sea f de clase C^2 en un intervalo I. Entonces existe $c \in \overline{xx_0}$ tal que $R(x) = 1/2 f''(c)(x - x_0)^2$. Luego fijados $x, x_0 \in I$ podemos escribir*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + 1/2 f''(c)(x - x_0)^2.$$

En otros términos

$$f(x) = \ell_{tg}(x) + R(x) = \ell_{tg}(x) + 1/2 f''(c)(x - x_0)^2.$$

EJEMPLO 5.1.2 (Ejemplos del error de orden 1)**.** Veamos cómo es el resto (o error) de orden 1 para algunos casos sencillos.

■ Sea $f(x) = \operatorname{sen}(x)$, sea $x_0 = 0$. Como $f'(x) = \cos(x)$, tenemos $f(0) = 0$, mientras que $f'(0) = 1$. Luego la recta tangente al seno por el origen

es

$$y = x.$$

Como $f''(x) = -\operatorname{sen}(x)$, el teorema nos dice que para cada $x \in \mathbb{R}$ existe c entre 0 y x tal que

$$\operatorname{sen}(x) = x - 1/2 \cdot \operatorname{sen}(c)x^2.$$

El resto o error está acotado por una parábola, como muestra esta cuenta:

$$|R(x)| = |f(x) - \ell_{tg}(x)| = \frac{1}{2}| \cdot \operatorname{sen}(c)| \cdot x^2 \leq \frac{1}{2}x^2.$$

Entonces por ejemplo si $|x| \leq 1/3$, entonces $x^2 \leq 1/9$ y entonces el error será menor o igual a $1/18$. Con esto podemos afirmar que para $|x| \leq 1/3$,

$$\operatorname{sen}(x) \simeq x,$$

con un error de *a lo sumo* $1/18 = 0,055 < 0,1$. Por supuesto, para x más pequeño el error será menor.

La utilidad de esta aproximación es que mientras mantengamos x en el rango mencionado (entre $-0,33$ y $0,33$) entonces podemos afirmar que $\operatorname{sen}(x) = x$ con una precisión de un decimal.

■ Sea $f(x) = e^x$, sea $x_0 = 0$. Calculamos $f'(x) = f''(x) = e^x$, entonces $f(0) = f'(0) = 1$ y así la recta tangente a la función exponencial por el origen es

$$y = 1 + x.$$

El teorema nos dice que para cada $x \in \mathbb{R}$ existe c entre 0 y x tal que

$$e^x = 1 + x + 1/2 \cdot e^c \cdot x^2.$$

Supongamos que $|x| \leq 1/4$, recordemos que c está entre 0 y x. Si $x < 0$ podemos afirmar que $e^c < e^0 = 1$, pero qué pasa si $x > 0$? Observamos que como la exponencial es una función creciente el valor de e^c es a los sumo $e^{0,25} < e^1 = e < 3$ -esta cota es muy burda pero será suficiente

para este ejemplo-. Entonces el resto o error se acota de la siguiente manera:

$$|R(x)| = 1/2 \cdot e^c \cdot x^2 \leq 1/2 \cdot 3 \cdot x^2 \leq 1/2 \cdot 3 \cdot 1/16 \simeq 0,094 < 0,1$$

ya que $|x| \leq 1/4$ implica que $x^2 \leq 1/4^2 = 1/16$. Con esto podemos afirmar que si $|x| < 1/4$ entonces

$$e^x \simeq 1 + x$$

con una precisión de un decimal.

Lo que queremos hacer ahora es conseguir una aproximación mejor de f, aproximación *cuadrática* (también llamada de segundo orden). Si f es una función cuadrática queremos que sea exacta esta aproximación. Y cuando f no sea una función cuadrática, queremos poder estimar, e incluso dar una fórmula para el resto (o error).

Fijada f de clase C^2 y un punto x_0 en su dominio, buscamos entonces una función cuadrática P que coincida a orden 2 con f en el punto x_0 ¿Qué quiere decir esto? Que pedimos

$$f(x_0) = P(x_0) \qquad f'(x_0) = P'(x_0) \qquad f''(x_0) = P''(x_0).$$

No es difícil ver que el único polinomio cuadrático que cumple esto es

$$P(x) = f(x_0) + f'(x_0)(x - x_0) + 1/2 f''(x_0)(x - x_0)^2$$

el *polinomio de Taylor de orden 2* de f en x_0.

El factor $1/2$ que antecede a la derivada segunda está puesto para que al derivar ese término, el 2 que baja del cuadrado se cancele y recuperemos $f''(x_0)$.
Ahora queremos controlar el error, llamado resto de orden 2, que es simplemente la diferencia entre la función y el polinomio:

$$R(x) = f(x) - P(x) = f(x) - f(x_0) - f'(x_0)(x - x_0) - 1/2 f''(x_0)(x - x_0)^2.$$

Para hacerlo, vamos a pedir que f tenga una derivada más:

TEOREMA 5.1.3 (Polinomio de orden 2 con resto de Taylor)**.** *Sea f de clase C^3, sea x_0 en el dominio de f. Entonces para cada x existe c entre x, p tal que*

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + 1/2 f''(x_0)(x-x_0)^2 + 1/6 f'''(c)(x-x_0)^3.$$

El resto de Taylor de orden 2 es la diferencia entre f y su polinomio de orden 2, luego

$$R(x) = 1/6 f'''(c)(x-x_0)^3.$$

Una vez más, R depende de f y de x_0, y además hay que aclarar que se trata del resto de orden 2, ya que no es lo mismo este resto que el resto de orden 1.

■ Vemos que en este caso el resto es pequeño porque si x está cerca de x_0, entonces $|x-x_0| < 1$ y entonces $|x-x_0|^3$ es aún más pequeño. Este resto es menor que el de orden 1, y por eso el polinomio de orden 2 es una mejor aproximación que la recta tangente.

■ También vemos que si f es una función cuadrática, $f''' \equiv 0$ y entonces $f = P$: el polinomio de Taylor de f (en cualquier punto) es exactamente igual a f.

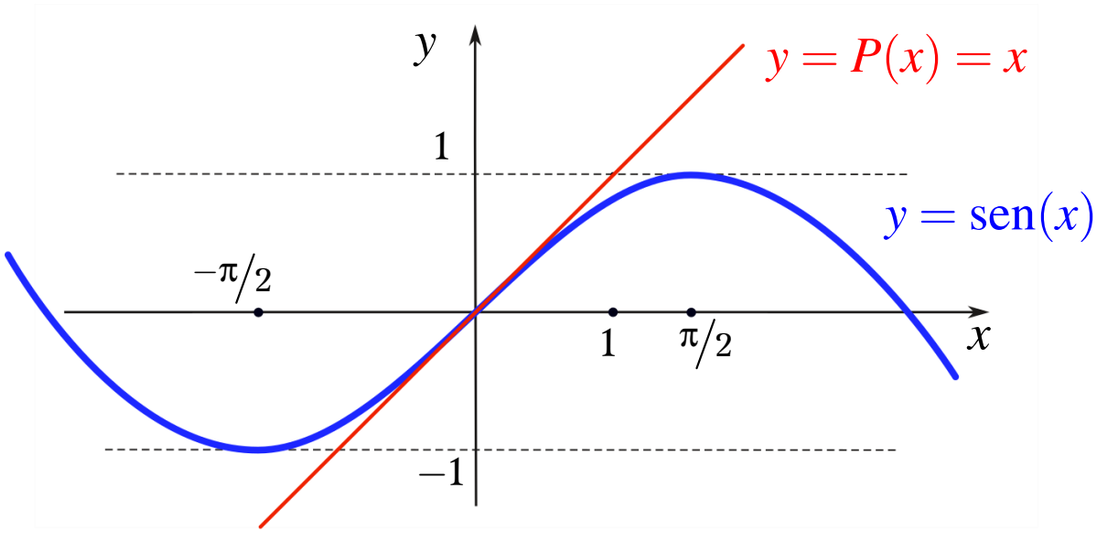
Veamos ahora los ejemplos de las funciones seno y exponencial en $x_0 = 0$, pero busquemos su polinomio de Taylor de orden 2 y su resto.

■ Sea $f(x) = \operatorname{sen}(x)$, sea $x_0 = 0$. Calculamos las derivadas, tenemos $f'(x) = \cos(x)$, $f''(x) = -\operatorname{sen}(x)$, $f'''(x) = -\cos(x)$. Entonces, como $\cos(0) = 1$ mientras que $\operatorname{sen}(0) = 0$, el teorema nos dice que para cada x, existe c entre 0 y x tal que

$$\operatorname{sen}(x) = x - 1/6 \cdot \cos(c) x^3.$$

Vemos que el polinomio de orden 2 es $P(x) = x$ ya que la derivada segunda es nula, así que coincide con la recta tangente en $x_0 = 0$.
Sin embargo, ahora la fórmula del resto es la de orden 2, y esto es mucho mejor porque aparece x^3. Supongamos nuevamente (como cuando estudiamos la aproximación de orden 1) que $|x| \leq 1/3$. Entonces

$$|R(x)| = \frac{1}{6} \cdot |\cos(c)| \cdot |x|^3 \leq \frac{1}{6} \cdot |x|^3 \leq \frac{1}{6} \cdot \frac{1}{3^3} \simeq 0,0061 < 0,01$$



puesto que $|\cos(c)| \leq 1$ y $|x| \leq 1/3$. Pero entonces en realidad la aproximación

$$\operatorname{sen}(x) \simeq x$$

tiene una precisión de *dos* decimales, siempre que $|x| \leq 1/3$.

■ Invitamos al lector a calcular el polinomio de Taylor de orden 2 de la función seno, pero ahora en $x_0 = -\pi/2$, y ver que allí es una parábola. Luego calcular el polinomio en $x_0 = \pi/2$ y ver que allí es una parábola invertida.

■ Sea $f(x) = \cos(x)$, sea $x_0 = 0$. Calculamos las derivadas, tenemos $f'(x) = -\operatorname{sen}(x)$, $f''(x) = -\cos(x)$, $f'''(x) = \operatorname{sen}(x)$. Entonces, como $\cos(0) = 1$ mientras que $\operatorname{sen}(0) = 0$, el teorema nos dice que para cada x, existe c entre 0 y x tal que

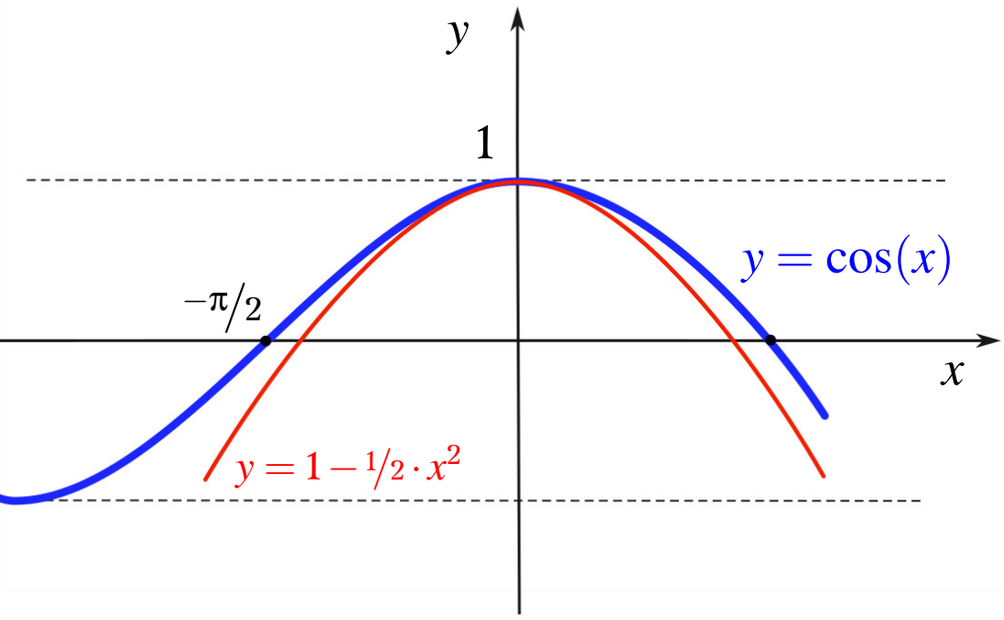
$$\cos(x) = 1 - 1/2 \cdot x^2 + 1/6 \cdot \operatorname{sen}(c) x^3.$$

Vemos que el polinomio de orden 2 en el origen del coseno es

$$P(x) = 1 - 1/2 \cdot x^2,$$

una parábola invertida con vértice en el punto de paso $(0,1)$.
Supongamos (como hicimos para el seno) que $|x| \leq 1/3$. Entonces

$$|R(x)| = \frac{1}{6} \cdot |\operatorname{sen}(c)| \cdot |x|^3 \leq \frac{1}{6} \cdot |x|^3 \leq \frac{1}{6} \cdot \frac{1}{3^3} \simeq 0,0061 < 0,01$$



puesto que $|\operatorname{sen}(c)| \leq 1$ y $|x| \leq 1/3$. Luego la aproximación

$$\cos(x) \simeq 1 - 1/2 \cdot x^2$$

tiene una precisión de *dos* decimales, siempre que $|x| \leq 1/3$.

■ Sea $f(x) = e^x$, tomemos $x_0 = 0$. Como todas las derivadas son e^x, tenemos por el teorema de Taylor de orden 2 que

$$e^x = 1 + x + 1/2 \cdot x^2 + 1/6 \cdot e^c \cdot x^3$$

para algún c entre $0, x$. En este caso el polinomio de orden 2 de la exponencial en el origen es

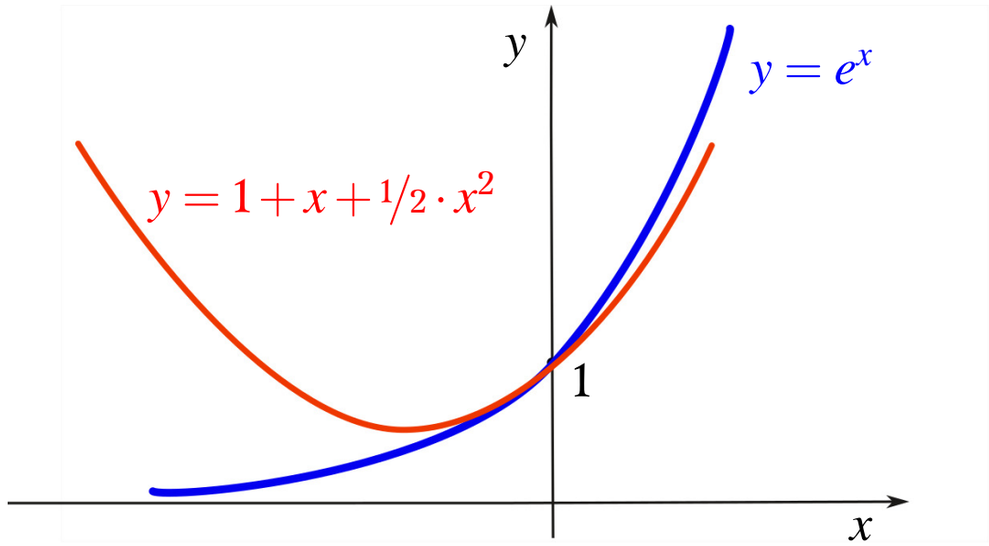
$$P(x) = 1 + x + 1/2 \cdot x^2.$$

Se trata de una parábola hacia arriba con vértice en $(-1, 1/2)$.
Supogamos nuevamente que $|x| \leq 1/4$: nuevamente como c está entre 0 y x tenemos $e^c < 3$ y así

$$|R(x)| = 1/6 \cdot e^c \cdot |x|^3 \leq 3/6 \cdot 1/4^3 \simeq 0,0078 < 0,01$$

ya que $|x| \leq 1/4$ implica $|x| \leq 1/4^3$. Esta acotación del error nos permite afirmar que si $|x| \leq 1/4$, entonces

$$e^x \simeq 1 + x + 1/2 \cdot x^2$$



con una precisión de dos decimales.

Polinomio de Taylor de orden 2 (en dos variables)

La idea es exactamente la misma en dos variables que en una.
Si $f : \mathbb{R}^3 \to \mathbb{R}$ es de clase C^3, buscamos un polinomio cuadrático que sea el que mejor aproxima a f cerca de un punto dado $X_0 = (x_0, y_0)$.

Para hallar este polinomio de grado 2 en 2 variables, pedimos que coincida con f en P y que coincidan todas las derivadas parciales hasta orden 2 también, es decir que $P(x,y)$ cumpla las seis condiciones

$$P(x_0,y_0) = f(x_0,y_0), \qquad \frac{\partial P}{\partial x}(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0)$$

$$\frac{\partial P}{\partial y}(x_0,y_0) = \frac{\partial f}{\partial y}(x_0,y_0), \qquad \frac{\partial^2 P}{\partial x^2}(x_0,y_0) = \frac{\partial^2 f}{\partial x^2}(x_0,y_0)$$

$$\frac{\partial^2 P}{\partial y^2}(x_0,y_0) = \frac{\partial^2 f}{\partial y^2}(x_0,y_0), \quad \frac{\partial^2 P}{\partial x \partial y}(x_0,y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)$$

ƨ Cuando f es un polinomio cuadrático, será *exactamente igual* a su polinomio de Taylor de orden 2, y cuando no lo sea, tendremos un error llamado resto de Taylor

$$R(x,y) = f(x,y) - P(x,y)$$

que ahora es una función de dos variables (y nuevamente depende de f y del punto P).

Para lidiar con las derivadas parciales de f de manera ordenada, será útil introducir la matriz Hessiana de f (también llamada matriz de la diferencial segunda de f):

DEFINICIÓN 5.1.4 (Matriz Hessiana)**.** Sea $A \subset \mathbb{R}^2$ abierto, sea $X_0 = (x_0, y_0)$ y sea $f \in C^2(A)$. El *Hessiano* de f en el punto X_0 es la matriz de las derivadas segundas, ordenadas de la siguiente manera:

$$Hf(x_0,y_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x_0,y_0) & \frac{\partial^2 f}{\partial y \partial x}(x_0,y_0) \\ \\ \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) & \frac{\partial^2 f}{\partial y^2}(x_0,y_0) \end{pmatrix}.$$

La matriz es fácil de recordar, especialmente por el siguiente teorema:

TEOREMA 5.1.5 (Derivadas cruzadas)**.** *Si f es de clase C^2 entonces*

$$\frac{\partial^2 f}{\partial x\, \partial y}(x_0,y_0) = \frac{\partial^2 f}{\partial y\, \partial x}(x_0,y_0).$$

Notemos que las seis condiciones que cumple el polinomio P de orden 2 de la función f se pueden resumir ahora como

$$f(X_0) = P(X_0), \qquad \nabla f(X_0) = \nabla P(X_0), \qquad Hf(X_0) = HP(X_0).$$

TEOREMA 5.1.6 (Fórmula de Taylor con resto de orden 2)**.** *Sea $A \subset \mathbb{R}^2$ abierto, sea $f \in C^3(A)$, sea $(x_0,y_0) \in A$.*

Dado $(x,y) \in B_R(P) \subset A$ entonces

$$\begin{aligned} f(x,y) =& f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)\cdot(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)\cdot(y-y_0) \\ &+ 1/2 \cdot \frac{\partial^2 f}{\partial x^2}(x_0,y_0)\cdot(x-x_0)^2 + 1/2 \cdot \frac{\partial^2 f}{\partial y^2}(x_0,y_0)\cdot(y-y_0)^2 \\ &+ \frac{\partial^2 f}{\partial y \partial x}(x_0,y_0)\cdot(x-x_0)\cdot(y-y_0) + R(x,y) \end{aligned}$$

donde R es el resto, y existe un punto $C = (c_1,c_2) \in A$, más precisamente

$C \in \overline{PX}$, tal que

$$\begin{aligned} R(x,y) = \quad & 1/6 \cdot \frac{\partial^3 f}{\partial x^3}(x_0,y_0)\cdot(x-x_0)^3 + 1/6 \cdot \frac{\partial^3 f}{\partial y^3}(x_0,y_0)\cdot(y-y_0)^3 \\ & + 1/2 \cdot \frac{\partial^3 f}{\partial x^2 \partial y}(x_0,y_0)\cdot(x-x_0)^2\cdot(y-y_0) \\ & + 1/2 \cdot \frac{\partial^3 f}{\partial y^2 \partial x}(x_0,y_0)\cdot(x-x_0)\cdot(y-y_0)^2 \end{aligned}$$

2 La diferencia en los coeficientes de las derivadas segundas y terceras obedece a que como f es de clase C^3, entonces hay varias derivadas sucesivas que son idénticas

$$f_{xy} = f_{yx}, \qquad f_{xxy} = f_{xyx} = f_{yxx}, \qquad f_{yyx} = f_{yxy} = f_{xyy}. \tag{5.1.2}$$

Todas las derivas de orden 2 van divididas por 2, y todas las derivadas del resto de orden tres van divididas por $3! = 6$, pero como agrupamos las idénticas, quedan esos coeficientes.

OBSERVACIÓN 5.1.7 (Notación vectorial/matricial)**.** Si denotamos al punto donde hacemos el desarrollo como $X_0 = (x_0, y_0)$, y similarmente denotamos $X = (x, y)$, podemos escribir la fómula de Taylor de manera compacta como

$$f(X) = f(X_0) + \langle \nabla f(X_0), X - X_0 \rangle + 1/2 \langle Hf(X_0)(X - X_0), X - X_0 \rangle + R(X),$$

que es muy similar a la fórmula de orden 2 para funciones de una variable.

EJEMPLO 5.1.8 (Polinomio de orden 2 en dos variables)**.** En el resumen de la Unidad 4, habíamos considerado

$$f(x,y) = 1 + e^{x+y-7x^2-9y^2},$$

cuyas derivadas parciales son

$$f_x(x,y) = e^{x+y-7x^2-9y^2}(1-14x), \qquad f_y(x,y) = e^{x+y-7x^2-9y^2}(1-18y).$$

Con esto, obtuvimos que el plano tangente a f en $P = (0,0)$ era

$$z = x + y + 2 = \pi(x,y),$$

luego $\pi(x,y) \simeq f(x,y)$ para (x,y) cerca del origen. Tenemos una representación gráfica de esto en el applet de Geogebra (click). Calculemos ahora el polinomio de orden 2 en el origen, con su resto de Taylor. Para eso necesitamos las derivadas sucesivas de f, que son

$$f_{xx}(x,y) = e^{x+y-7x^2-9y^2}(196x^2 - 28x - 13)$$

$$f_{yy}(x,y) = e^{x+y-7x^2-9y^2}(324y^2 - 36y - 17)$$

$$f_{xy}(x,y) = f_{yx}(x,y) = e^{x+y-7x^2-9y^2}(1-14x)(1-18y)$$

$$f_{xxx}(x,y) = e^{x+y-7x^2-9y^2}(-2744x^3 + 588x^2 + 546x - 41)$$

$$f_{xxy}(x,y) = e^{x+y-7x^2-9y^2}(1-18y)(196x^2 - 28x - 13)$$

$$f_{yyy}(x,y) = e^{x+y-7x^2-9y^2}(-5832y^3 + 872y^2 + 918y - 53)$$

$$f_{yyx}(x,y) = e^{x+y-7x^2-9y^2}(1-14x)(324y^2 - 36y - 17).$$

Las derivadas que faltan son iguales a algunas de estas, por las relaciones (5.1.2) dadas porque f es de clase C^3 en todo $\mathbb{R}^2$. Escribamos primero la expresión del resto, para eso fijado $X = (x,y)$ sabemos que existe $C = (c_1, c_2)$ en el segmento que une $X_0 = \mathbb{O}$ con X tal que

$$\begin{aligned} R(x,y) = \quad & 1/6 \cdot f_{xxx}(c_1,c_2) \cdot x^3 + 1/6 \cdot f_{yyy}(c_1,c_2) \cdot y^3 \\ & + 1/2 \cdot f_{xxy}(c_1,c_2) \cdot x^2 y + 1/2 \cdot f_{yyx}(c_1,c_2) \cdot x y^2 \\ = \quad & 1/6 \cdot e^{c_1+c_2-7c_1^2-9c_2^2}(-2744c_1^3 + 588c_1^2 + 546c_1 - 41) \cdot x^3 \\ & + 1/6 \cdot e^{c_1+c_2-7c_1^2-9c_2^2}(-5832c_2^3 + 872c_2^2 + 918c_2 - 53) \cdot y^3 \\ & + 1/2 \cdot e^{c_1+c_2-7c_1^2-9c_2^2}(1-18c_2)(196c_1^2 - 28c_1 - 13) \cdot x^2 y \\ & + 1/2 \cdot e^{c_1+c_2-7c_1^2-9c_2^2}(1-14c_1)(324c_2^2 - 36c_2 - 17) \cdot x y^2. \end{aligned}$$

No vamos a acotar este resto en dos variables, pero es posible hacerlo

sabiendo qué tan lejos vamos del origen con $X = (x, y)$ (ya que entonces $|c_1| \leq |x|$ y también $|c_2| \leq |y|$. Sabemos que si X está cerca del origen $\mathbb{O}$, entonces $f(x, y) \simeq P(x, y)$ con el error controlado por este resto.

Pero ¿quién es este polinomio cuadrático P en este caso?

Podemos escribirlo con las derivadas que ya calculamos, evaluadas en $(x_0, y_0) = (0, 0)$ como indica el teorema de arriba. Tenemos $f(0,0) = 2$, $f_x(0,0) = 1$, $f_y(0,0) = 1$, $f_{xx}(0,0) = -13$, $f_{yy}(0,0) = -17$ y por último $f_{xy}(0,0) = f_{yx}(0,0) = 1$. Entonces

$$\begin{aligned} P(x,y) = & f(0,0) + f_x(0,0) \cdot x + f_y(0,0) \cdot y + f_{xy}(0,0) \cdot xy \\ & + 1/2 \cdot f_{xx}(0,0) \cdot x^2 + 1/2 \cdot f_{yy}(0,0) \cdot y^2, \end{aligned}$$

luego

$$P(x,y) = 2 + x + y + xy - 13/2 \cdot x^2 - 17/2 \cdot y^2.$$

Podemos ver que el término lineal afín del polinomio es $2 + x + y$, que era exactamente el plano tangente a f en el origen. También puede el lector verificar que el gráfico de P se trata de un paraboloide invertido. Por supuesto, la aproximación del polinomio de grado 2 es mucho mejor que la del plano, y podemos ver esto representado gráficamente en el applet de Geogebra (click).

5.2. Integración

• Corresponde a clases en videos 5.3a, 5.3b, 5.4

Recordemos algunas definiciones y propiedades:

DEFINICIÓN 5.2.1 (Integrales)**.** Sea $f : I \to \mathbb{R}$ continua, con $I \subset \mathbb{R}$ un intervalo.

■ Una *primitiva (o antiderivada)* F de la función f es una función derivable en I tal que $F'(x) = f(x)$ para todo $x \in I$.

■ La función nula $f \equiv 0$ tiene como primitiva a cualquier función constante, y esas son las únicas primitivas de la función nula.

■ Si F, G son dos primitivas de f entonces existe una constante $c \in \mathbb{R}$ tal que $G(x) = F(x) + c$ para todo $x \in I$.

■ La *integral indefinida* $\int f$ es la colección de todas las primitivas de f en I. Entonces si F es alguna primitiva, tenemos que

$$\int f = F + c, \qquad c \in \mathbb{R}$$

■ Para toda F derivable se tiene $\int F' = F + c$.

■ La integral definida de f se calcula usando una primitiva cualquiera mediante la *Regla de Barrow*

$$\int_a^b f = F(x)\Big|_a^b = F(b) - F(a)$$

y esta cuenta no depende de cuál primitiva usemos.

■ En particular si F es derivable entonces $\int_a^b F' = F(b) - F(a)$.

■ La *función integral* de $f : [a,b] \to \mathbb{R}$ es la dada por $I(x) = \int_a^x f$, definida para $x \in [a,b]$. Notemos que Eso nos permite escribir la fórmula

$$I'(x) = \left(\int_a^x f\right)' = f(x).$$

■ Si F es una primitiva de f entonces la función integral no es otra cosa que $I(x) = F(x) - F(a)$. Luego la función integral de f es una primitiva específica de f, es la única primitiva de f que se anula en $x = a$.

■ Cuando es necesario, podemos aclarar la variable de integración escribiendo

$$\int f = \int f(x)dx, \qquad \int_a^b f = \int_a^b f(x)dx = F(x)\Big|_{x=a}^{x=b}$$

■ La *regla de integración por partes* nos permite calcular tanto integrales indefinidas

$$\int f'g = fg - \int fg'$$

como definidas

$$\int_a^b f'(x)g(x)dx = f(x)g(x)\Big|_a^b - \int_a^b f(x)g'(x)dx.$$

Esta regla es una consecuencia inmediata de la regla de derivación del producto.

■ Si F es una primitiva de f, el *método de sustitución* nos permite escribir la igualdad

$$\int f(g(x))g'(x)dx = \int f(u)du = F(u)+c = F(g(x))+c,$$

mediante la sustitución $u = g(x)$.

■ Para integrales definidas obtenemos entonces

$$\int_a^b (f\circ g)\cdot g' = F(g(b)) - F(g(a)),$$

siempre que $F'(x) = f(x)$ en el intervalo dado.

PROPIEDADES 5.2.2 (de las integrales)**.** Sean f, g continuas en el intervalo I, entonces

■ $\int f+g = \int f + \int g.$

■ Si $\lambda \in \mathbb{R}$ entonces $\int \lambda f = \lambda \int f.$

■ Si $a,b,c \in I$ entonces $\int_a^c f = \int_a^b f + \int_b^c f.$

■ Si $a,b \in I$ entonces $\int_a^b f = -\int_b^a f.$

■ Si $f \geq 0$ en $[a,b]$ entonces $\int_a^b f \geq 0$. Es importante aquí que $a < b$.

■ Si $f \geq g$ en $[a,b]$ entonces $\int_a^b f \geq \int_a^b g$.

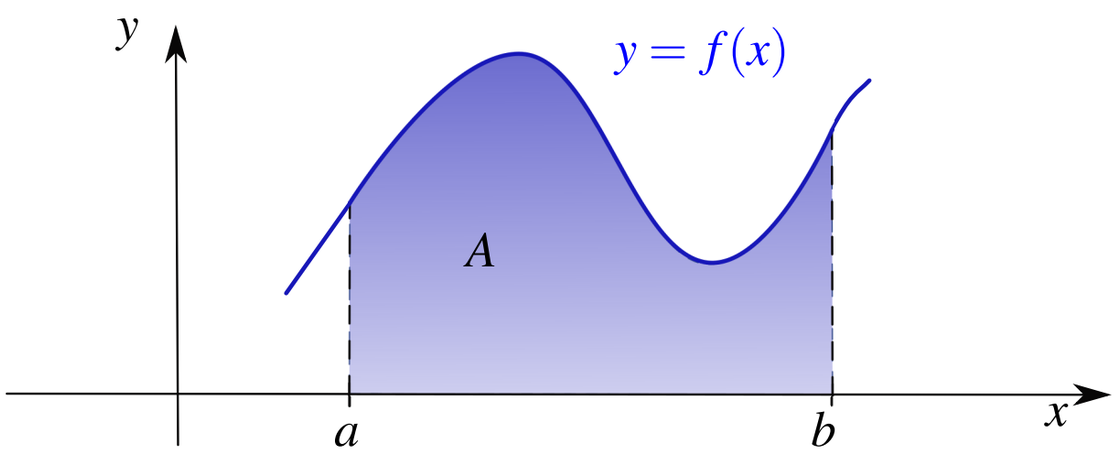
ƨ En general $\int_a^b f \geq 0$ no implica que $f \geq 0$ en todo el intervalo, porque puede haber cancelaciones.

ƨ Por el mismo motivo $\int_a^b f = 0$ no implica que $f(x) = 0$ para todo $x \in [a,b]$.

TEOREMA 5.2.3 (Teorema fundamental del cálculo integral)**.** .
Si $f : [a,b] \to \mathbb{R}$ es continua y $f \geq 0$, entonces

$$A = \int_a^b f(x)dx$$

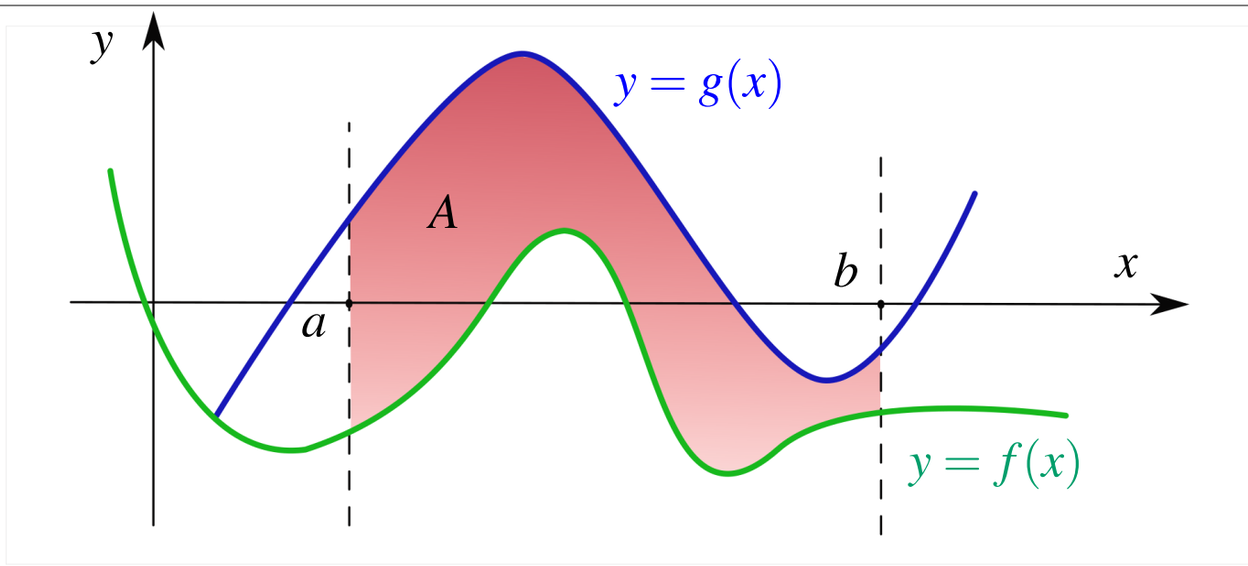
donde $A \geq 0$ es el área bajo el gráfico de f y sobre el eje de las x (encerrada por las rectas verticales $x = a, x = b$).



TEOREMA 5.2.4 (Área entre curvas)**.** *Si $g \geq f$ en el intervalo $[a,b]$ y ambas son integrables, el área encerrada entre los gráficos de g y de f se calcula con la integral definida*

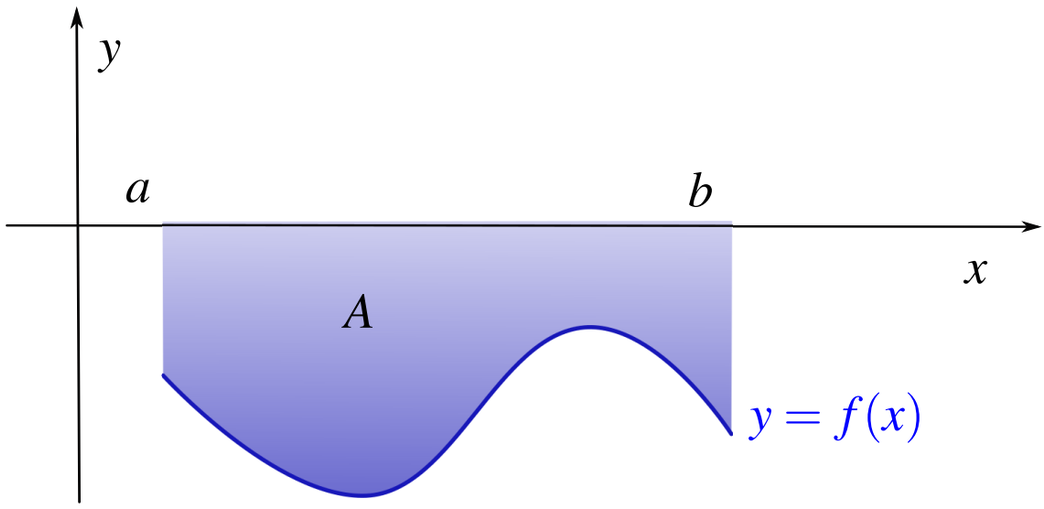
$$\int_a^b (g-f) = \int_a^b g - \int_a^b f.$$

ƨ No es relevante si f ó g cambian de signo, lo único importante es que g (el techo) permanezca siempre por encima de f (el piso) mientras x recorre el intervalo $[a,b]$.



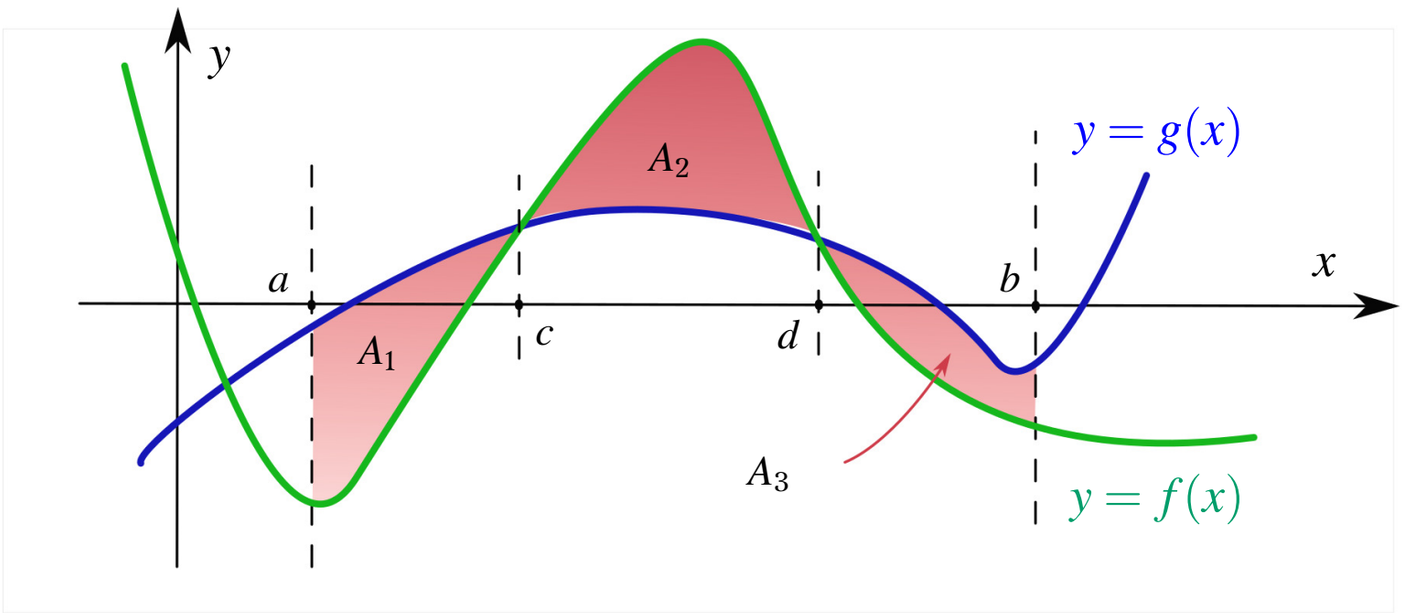
2 Si $f \leq 0$ en el intervalo $[a,b]$ entonces el área A atrapada *debajo* del eje x y *encima* del gráfico de f se calcula como

$$A = \int_a^b (-f) = -\int_a^b f.$$



Podemos pensar aquí que $g = 0$; entonces es un caso especial del teorema anterior, ya que $f \leq g = 0$ y así $g - f = 0 - f = -f$. También podemos pensar que si $f \leq 0$, entonces la función $-f$ es positiva (su gráfico es el simétrico respecto del eje x); luego el área atrapada bajo $-f$ y sobre el eje es la misma que esta.

OBSERVACIÓN 5.2.5 (Curvas que se cortan)**.** Si los gráficos de f, g se entrecruzan, primero tenemos que encontrar donde, y luego separar el intervalo en intervalos más pequeños donde una está encima de la otra. Para ello se requiere hallar donde se cruzan, y estos son los $x \in [a,b]$ tales que $f(x) = g(x)$.



En ejemplo gráfico que tenemos debajo, los puntos de corte en el intervalo $[a,b]$ son $x = c, x = d$. Entonces el área total es

$$A = A_1 + A_2 + A_3 = \int_a^c (g-f) + \int_c^d (f-g) + \int_d^b (g-f).$$

5.3. Ecuaciones diferenciales

• Corresponde a clases en video 5.5a, 5.5b

Si $f : I \to \mathbb{R}$ representa una cantidad (por ejemplo, una distancia que varía con el tiempo, que es la variable), entonces el cociente incremental

$$\frac{f(t) - f(t_0)}{t - t_0} = \frac{f(t_0 + \Delta t) - f(t_0)}{\Delta t} = \frac{\Delta f}{\Delta t}$$

Representa la variación de esta cantidad respecto de la variable (en el ejemplo, la velocidad media, que es el cociente de la distancia recorrida por el tiempo transcurrido).

Cuando el intervalo de tiempo tiende a cero ($\Delta t \to 0$) obtenemos la velocidad instantánea. Entonces si f representa la distancia de un móvil en función del tiempo, f' representa la velocidad instantánea de ese móvil.

Hay mucha otras aplicaciones de este modelo. Por ejemplo si f representa una población de bacterias y t es el tiempo, la derivada representa

la velocidad con la que crece (o decrece) el número de bacterias). O si f representa la cantidad de litros de una sustancia que pasa por una compuerta a través del tiempo, entonces f' es la velocidad de ese líquido por la compuerta. También hay modelos económicos donde f representa la cantidad de capital, y la variable puede ser el tiempo, o la cantidad de insumo de un cierto producto que produce una ganancia de capital; en todos los casos la derivada f' es la velocidad de cambio.

Hay modelos donde uno puede inferir la relación entre la cantidad que uno tiene (por ejemplo la posición del movil) y su velocidad de cambio (por ejemplo basándose en el conocimiento de la superficie por la que se mueve el móvil, o por el tamaño de las compuertas, etc. Entonces uno obtiene una relación entre f y su derivada f'. La pregunta es si a partir de esta relación podemos hallar alguna (o algunas) funciones que cumplan esta relación; esta f hallada es el *modelo* matemático para el problema planteado.

EJEMPLO 5.3.1 (Bacterias)**.** Un primer ejemplo simple es el siguiente: supongamos que sabemos que la velocidad es siempre proporcional a la cantidad. Por ejemplo en el crecimiento de bacterias, mientras más hay más rápido crecen. Digamos que sabemos que $f'(x) = 3f(x)$ para todo x en el intervalo de tiempo que nos interesa. ¿Podemos hallar la función f que nos dice la cantidad de bacterias en cada instante?

En este caso la poblacion es un número positivo, o sea $f(x) > 0$. De la relación que tenemos, podemos escribir

$$\frac{f'(x)}{f(x)} = 3 \qquad \forall x$$

Pero entonces podemos probar integrar. Como tenemos una igualdad de funciones (a la derecha está la función constante 3, vemos que

$$\int \frac{f'(x)}{f(x)} dx = \int 3\, dx = 3x + c.$$

Ahora hacemos una sustitución, $u = f(x)$. Luego $du = f'(x)dx$ y así

nos queda

$$3x + c = \int \frac{f'(x)}{f(x)} dx = \int \frac{1}{u} du = \ln|u| = \ln|f(x)| = \ln f(x)$$

pues $f(x) > 0$. Exponenciando ambos lados vemos que debe ser

$$f(x) = e^{3x+c} = e^{3x} e^{c} = k \cdot e^{3x}$$

donde $k > 0$ es una constante a determinar. ¿Cómo la hallamos? Para eso necesitamos saber la cantidad de bacterias en algún instante del tiempo, digamos 10^5 en el instante $x = 1$: este dato es simplemente $f(1) = 10^5$. Pero entonces reemplazando en la expresión que obtuvimos para f vemos que

$$10^5 = f(1) = k \cdot e^3,$$

luego $k = \frac{10^5}{e^3}$, y así

$$f(x) = \frac{10^5}{e^3} \cdot e^{3x}$$

es la función que modela la población de bacterias.

OBSERVACIÓN 5.3.2 (Notación)**.** En general se suele usar la letra x para denotar la variable, y la cantidad se denota con la variable y. En realidad, se trata de una función $y = y(x)$. Entonces la relación del ejemplo $f'(x) = 3f(x)$ se reescribe como

$$y'(x) = 3y(x) \qquad \text{o simplemente} \qquad y' = 3y.$$

Esto es lo que se conoce como una *ecuación diferencial* de orden 1 (porque sólo involucra la derivada primera). Como vimos, puede tener muchas soluciones, pero una vez fijada una condición inicial (en el ejemplo, $y(1) = 10^5$) hay una única solución de la ecuación diferencial, que es una función que verifica la relación.

OBSERVACIÓN 5.3.3 (Método de separación de variables)**.** Respecto del método que usamos para resolver la ecuación del ejemplo, este

cambio de notación sugiere una estrategia que funciona en muchos casos. Escribimos $y'(x) = \frac{dy}{dx}$, luego la ecuación diferencial se reescribe como

$$\frac{dy}{dx} = 3y.$$

Podemos operar algebraicamente sobre esta ecuación para dejar los diferenciales en el numerador, y todo lo que tenga x de un solo lado (y todo lo que tenga y del otro lado):

$$\frac{dy}{y} = 3\,dx.$$

Este método se conoce como *separación de variables*. La justificación concreta está en cómo lo resolvimos más arriba, cuando hicimos una sustitución. Siguiendo desde aquí, podemos integrar a ambos lados

$$\int \frac{1}{y}\,dy = \int \frac{dy}{y} = 3\int dx.$$

Notamos que a cada lado, se integra respecto del diferencial de la variable que sólo figura a ese lado. Luego integrando vemos que

$$\ln|y| = 3x + c, \qquad \text{esto es} \qquad \ln|y(x)| = 3x + c.$$

Exponenciando vemos que $|y(x)| = e^{3x+c} = k \cdot e^3 c$ como antes. Si suponemos que $y(x) > 0$, arribamos a la misma solución, y la condición inicial nos permite hallar k.

2 Si no suponemos que $y > 0$ (por ejemplo si la ecuación diferencial $y' = 3y$ representa una cantidad que puede cambiar de signo, como la cantidad de capital, que es negativo cuando hay deudas), entonces en principio la solución es de forma $y(x) = \pm ke^{3x}$, y esto es lo mismo que decir que $y(x) = Ae^{3x}$ donde ahora A puede ser una constante positiva o negativa.
Esta constante A se determina también con las condiciones iniciales; por ejemplo si en tiempo $x = 0$ el capital era -10 (deuda) entonces $y(x) = -10e^{3x}$, y vemos que siempre estaremos en deuda, y la deuda

será cada vez mayor.

EJEMPLO 5.3.4 (Segunda Ley de Newton)**.** La misma postula que la aceleración de un móvil (que es la derivada de la velocidad, o sea la derivada segunda de la posición) es propocional a la fuerza F que se ejerce sobre el móvil para moverlo.
Por ejemplo si $P(t)$ denota la posición en función del tiempo, y m es la masa del objeto, la segunda ley de Newton se escribo como la ecuación diferencial

$$mv'(t) = F(v(t))$$

donde $v(t) = P'(t)$ es la velocidad.
Supongamos que la fuerza es constante, $F = 1$ y para simplificar que la masa m también es 1. Tenemos la ecuación diferencial $v' = 1$, que con nuestra notación se escribe como $\frac{dv}{dt} = 1$. Luego $dv = dt$, así que integrando a ambos lados vemos que

$$v = \int dv = \int dt = t + c, \qquad \text{o sea} \quad v(t) = t + c.$$

La constante c está determinada por la velocidad inicial del móvil, porque $v(t_0) = t_0 + c$ así que $c = v_0 - t_0$, donde usamos la notación $v_0 = v(t_0)$. Entonces la velocidad es $v(t) = t + v_0 - t_0$. Para fijar ideas, supongamos que $t_0 = 1$ y que $v_0 = 3$, luego para este móvil la ecuación de su velocidad es

$$v(t) = t + 2.$$

Ahora queremos hallar la posición, para ello recordamos que $p'(t) = v(t)$ así que ahora tenemos la ecuación diferencial

$$\frac{dp}{dt} = p' = v = t + 2,$$

que se reescribe como $dp = (t+2)dt$. Integrando ambos lados tenemos

$$p = \int 1 \cdot dp = \int (t+2)dt = \frac{t^2}{2} + 2t + k,$$

es decir $p(t) = 1/2 \cdot t^2 + 2t + k$. Nuevamente hay que determinar la constante k, para eso necesitamos el dato de alguna posición, por ejemplo $p(t_0) = p_0$; recordemos que teníamos $t_0 = 1$, y ahora supongamos que $p_0 = 10$. Entonces

$$10 = p_0 = p(t_0) = 1/2 \cdot t_0^2 + 2t_0 + k = 1/2 + 2 + k.$$

así que de aquí podemos despejar $k = 10 - 3/2 = 17/2$. Así que la ecuación de movimiento de nuestro objeto es

$$p(t) = 1/2 \cdot t^2 + 2t + 17/2.$$

Dejamos como ejercicio: resolver este problema del movimiento para una masa genérica m y una fuerza constante genérica F. Hay que resolver es la ecuación diferencial $mv'(t) = F$, para m, F constantes genéricas. Se obtiene una función cudrática con parámetros, y es interesante pensar qué ocurre cuando los parámetros m, F varían.

EJEMPLO 5.3.5 (Un problema de mezcla)**.** Un tanque mezclador tiene 100 litros de una solución compuesta por 80 % de agua y 20 % de acetona. Se le comienza a agregar agua a una velocidad de 5 litros por minuto. Mientras se mezcla dentro del tanque, este deja salir por debajo líquido a una velocidad también de 5 litros por minuto, para usar la mezcla en otro proceso. Después de 40 minutos ¿cuánta acetona habrá en el tanque?

Denotamos $y = y(x)$ la cantidad de acetona en el tanque respecto del tiempo, sabemos que $y_0 = y(0) = 20$ litros. La cantidad de líquido total es constante, pero no así la de acetona, que va disminuyendo. La velocidad a la que se pierde es $y' = -$LO QUE SALE (en litros por minuto).

Lo que sale por debajo es una proporción del total, el factor de proporcionalidad es

$$\frac{\text{SALE}}{\text{TOTAL}} y = \frac{5}{100} y = \frac{1}{20} y.$$

Entonces la ecuación diferencial que modela el problema es

$$y' = -\frac{1}{20}y.$$

La reescribimos como $\frac{dy}{dx} = -\frac{1}{20}y$ lo que nos lleva a

$$\frac{1}{y}dy = -\frac{1}{20}dx.$$

Integrando ambos lados, obtenemos $\ln|y| = -\frac{1}{20}x + c$. Sabemos que $y(x) > 0$ para todo x, y también que $y_0 = 20$, luego

$$c = c - \frac{1}{20}\cdot 0 = \ln y_0 = \ln(20).$$

Entonces exponenciando obtenemos $y(x) = e^c e^{-x/20} = 20e^{-x/20}$. Luego después de 40 minutos, la cantidad de acetona en el tanque es

$$y(40) = 20e^{-40/20} = 20e^{-2} \simeq \frac{20}{7,4} \simeq 2,7 \textit{ litros },$$

o equivalentemente un $2,7\%$ de acetona y un $97,3\%$ de agua.

Capítulo 6

Extremos de funciones, optimización

6.1. Extremos locales

• Corresponde a clases en video 6.1a, 6.1b, 6.2a, 6.2b, 6.3a, 6.3b

Un *extremo* de una función f es un máximo o mínimo de f.

Recordemos que dado $\delta > 0$, la bola $B_\delta(X_0)$ es el conjunto de puntos que distan menos que δ de X_0. Podemos usarlo en este caso para indicar cercanía con X_0. Los *extremos locales* de una función son extremos con respecto a los puntos cercanos. También se usa el la palabra *relativo* como sinónimo de local. Más precisamente:

DEFINICIÓN 6.1.1 (Máximos y mínimos locales)**.** Sea $f : A \to \mathbb{R}$ (aquí $A \subset \mathbb{R}^n$). Dado $X_0 \in A$, decimos que X_0 es

1. mínimo local de f si existe $\delta > 0$ tal que

$$f(X) \geq f(X_0) \qquad \forall X \in A \cap B_\delta(X_0).$$

2. máximo local de f si existe $\delta > 0$ tal que

$$f(X) \leq f(X_0) \qquad \forall X \in A \cap B_\delta(X_0).$$

3. mínimo local estricto de f si existe $\delta > 0$ tal que

$$f(X) > f(X_0) \qquad \forall X \in A \cap B_\delta(X_0), \quad X \neq X_0.$$

4. máximo local estricto de f si existe $\delta > 0$ tal que

$$f(X) < f(X_0) \qquad \forall X \in A \cap B_\delta(X_0). \quad X \neq X_0.$$

5. Un máximo o mínimo local es un *extremo* local de f.

Si removemos la condición local, es decir si la desigualdad vale para todo $X \in A$ (y no sólo para X cercano a X_0) decimos que el extremo es *global*, o *absoluto*.

Para algunas funciones sencillas, podemos identificar extremos sin mayor preámbulo que comparar $f(X_0)$ con $f(X)$ para X cercano a X_0. Veamos algunos ejemplos importantes:

■ Sea $f(x) = x^2$, como $f(0) = 0$ entonces $f(x) = x^2 > 0 = f(0)$ si $x \neq 0$, luego $x_0 = 0$ es mínimo local estricto. De hecho, es un mínimo *global* porque la desigualdad vale en todo el dominio $A = \mathbb{R}$ de la función.

■ Sea $f(x) = x^2 - x^3 = x^2(1-x)$, sea $x_0 = 0$. Si $x \in B_1(0)$, en particular $x < 1$ y así $1 - x > 0$. Multiplicando esta última desigualdad por x^2 vemos que

$$f(x) = x^2(1-x) \geq 0 = f(0).$$

Entonces $x_0 = 0$ es mínimo local de f, de hecho es mínimo local estricto porque si $x \neq 0$ entonces $x^2 > 0$ y así

$$f(x) = x^2(1-x) > 0 = f(0).$$

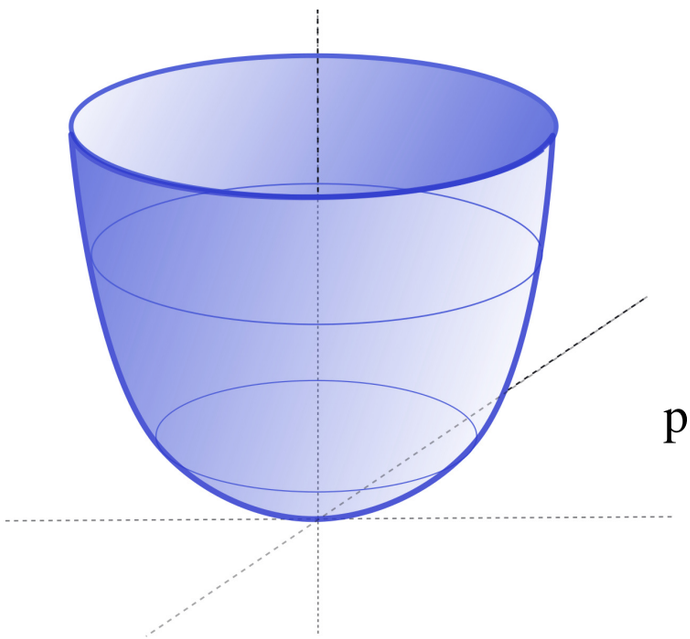
Es importante notar que si nos alejamos de $x_0 = 0$ deja de ser mínimimo, por ejemplo $f(5) = 25 \cdot (-4) < 0$ y es claro que 0 no es mínimo absoluto de f.

■ Sea $f(x) = x^4$, con $x_0 = 0$. Entonces $f(x) = x^4 > 0 = f(0)$ si $x \neq 0$ luego $x_0 = 0$ es mínimo estricto absoluto de f.

■ Si invertimos los signos obtendremos máximos, por ejemplo $f(x) = -x^2$ o bien $f(x) = -x^4$ tienen máximo estricto en $x_0 = 0$.

■ La función $f(x) = x^3$ no tiene un extremo (ni local ni global) en $x_0 = 0$. Esto es porque x^3 cambia de signo alrededor de $f(0) = 0$.

Pasemos ahora a funciones de dos variables.



paraboloide

■ Sea $f(x,y) = x^2 + y^2$, entonces $f(0,0) = 0$ y como

$$f(x,y) = x^2 + y^2 \geq 0 = f(0,0)$$

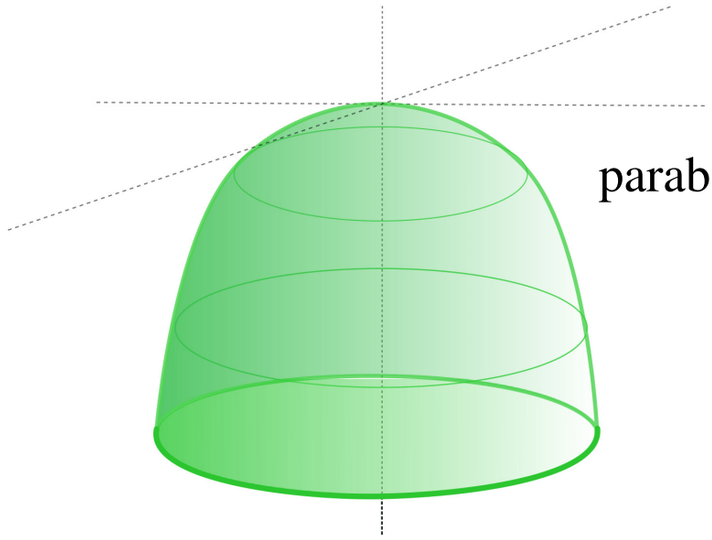
entonces $X_0 = (0,0)$ es mínimo de f. Al igual que en el caso de una variable, la desigualdad es estricta si $X \neq (0,0)$, así que es un mínimo estricto global de f.
■ En general, si el gráfico de f es un paraboloide, su vértice será mínimo estricto.

■ Lo mismo ocurre con $f(x,y) = x^4 + y^4$ o con $f(x,y) = x^2 + y^4$ o con $f(x,y) = 3x^4 + 5y^2$, etc.

■ Si invertimos los signos obtendremos un máximo, por ejemplo $X_0 = (0,0)$ es máximo extricto de $f(x,y) = -2x^2 - 3y^2$. En particular los paraboloides invertidos tienen máximo extricto en su vértice.

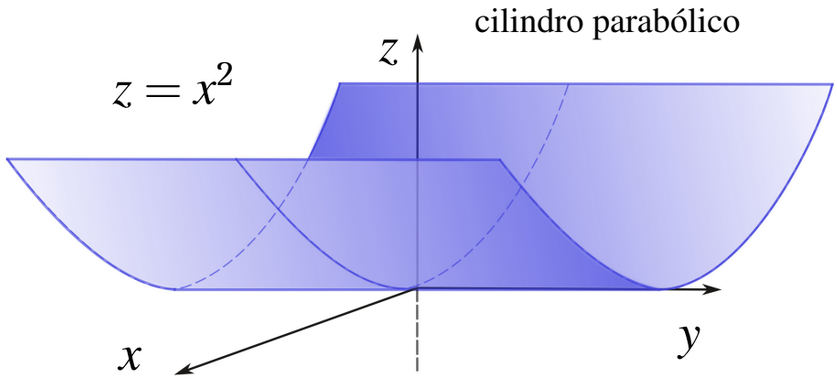
■ Veamos los casos llamados "degenerados" donde algún coeficiente es nulo. Por ejemplo sea $f(x,y) = x^2$, pensemos qué ocurre en el punto $X_0 = (0,0)$. Es claro que

$$f(x,y) = x^2 \geq 0 = f(0,0)$$



paraboloide invertido

entonces podemos afirmar que $(0,0)$ es mínimo global de f. Pero no es estricto porque por ejemplo $f(0,1) = 0^2 = 0$. De hecho $f(0,y) = 0$ para todo $y \in \mathbb{R}$, la función se anula sobre el eje de las y. Podemos afirmar que todos los puntos del eje y son mínimos (no estrictos) de f. Eso es más fácil de visualizar si recordamos que el gráfico de $f(x,y) = x^2$ es el cilindro parabólico:



■ El otro caso degenerado es cuando el coeficiente no nulo es negativo, por ejemplo $f(x,y) = -y^2$. Aquí vemos que $X_0 = (0,0)$ es máximo (no estricto), y de hecho cualquier punto del eje x lo es porque

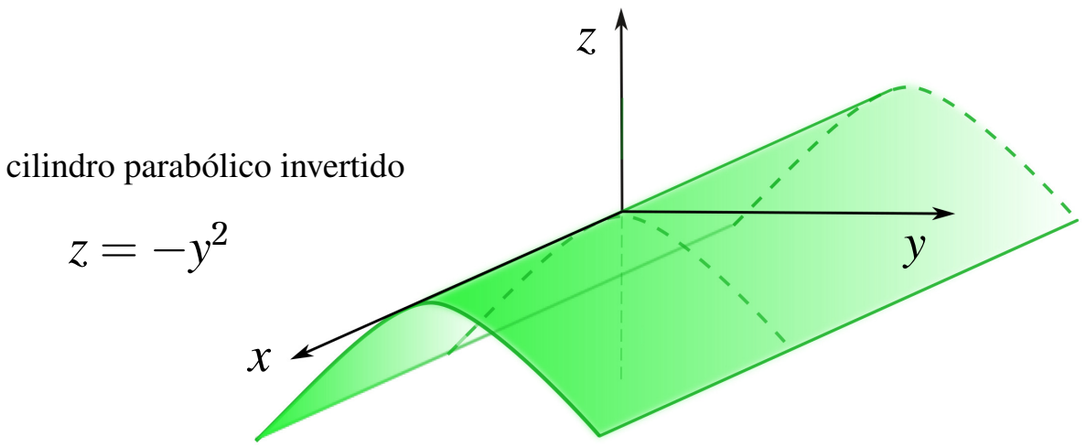
$$f(x,y) = -y^2 \leq 0 = f(x,0)$$

En este caso el gráfico es f es un cilindro parabólico invertido (Figura **??**.

■ Un extremo local estricto (que no es global) lo tiene

$$f(x,y) = (x-1)^2(11+x-y^2) + (y-2)^2$$

en $X_0 = (1,2)$. Para verlo, supongamos que $(x,y) \in B_1(X_0)$, entonces en



particular $|x-1|<1$ y también $|y-2|<1$. Respecto de x, vemos que $0<x<2$ y entonces $11+x>11$. Respecto de y, tenemos que $1<y<3$, luego $1<y^2<9$, entonces $-1>-y^2>-9$. En particular $-y^2>-9$ en esta bola. Combinando las dos desigualdas para x e y que obtuvimos, vemos que

$$11+x-y^2=(11+x)-y^2>11-9=2.$$

Es nos dice que el segundo factor de f (para $X\in B_1(X_0)$) es estrictamente positivo. Como el primer factor también lo es, vemos que si $(x,y)\in B_1(X_0)$ sin tocar $X_0=(1,2)$, entonces

$$f(x,y)=(x-1)^2(11+x-y^2)+(y-2)^2>(x-1)^2\cdot 2>0=f(1,2).$$

Por eso $X_0=(1,2)$ es mínimo local estricto de f.

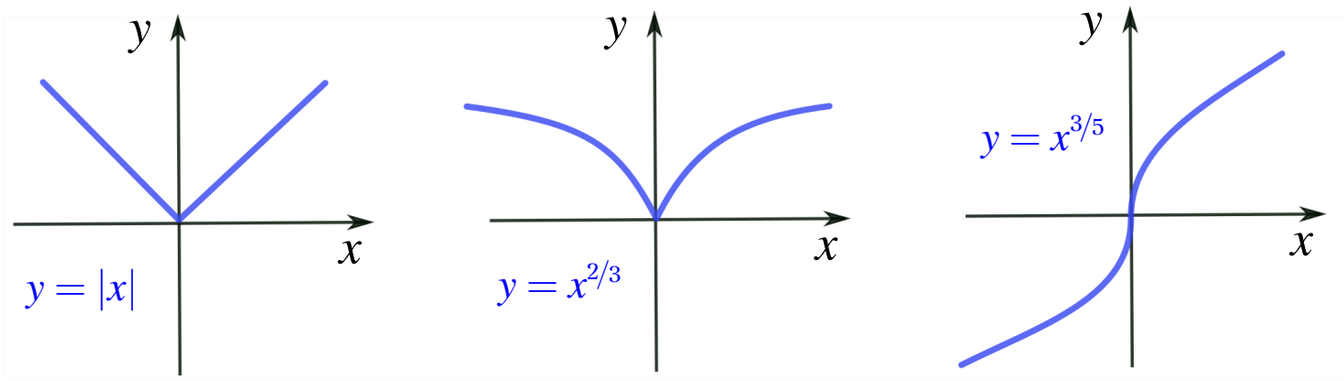
Hasta aquí, hemos observado o señalado los extremos "a mano". Lo que queremos es indicar un criterio que nos indique cómo hallarlos. Para eso comenzamo con una definición útil:

DEFINICIÓN 6.1.2 (Punto crítico)**.** Sea $A\subset\mathbb{R}^n$ abierto, sea $f:A\to\mathbb{R}$. Decimos que $X_0\in A$ es *punto crítico* de f si alguna de las dos siguientes condiciones se cumple

1. No existe alguna de las derivadas parciales de f en X_0.
2. Existen todas las derivadas parciales y son nulas. Equivalentemente $\nabla f(X_0)=\mathbb{O}$.

Los puntos críticos de primera clase son aquellos donde la función no es suave. Por ejemplo

■ La función módulo $f(x) = |x|$ tiene un punto crítico en $x_0 = 0$ porque no es derivable allí. Es un mínimo estricto.



■ $f(x) = x^{2/3}$ tiene un punto crítico en $x_0 = 0$ porque no es derivable allí, la derivada para $x \neq 0$ es $f'(x) = {}^2\!/_3 x^{-1/3}$. También es un mínimo estricto.

■ $f(x) = x^{3/5}$ tiene un punto crítico en $x_0 = 0$ porque no es derivable allí, la derivada para $x \neq 0$ es $f'(x) = {}^3\!/_5 x^{-2/5}$. Sin embargo, $x_0 = 0$ no es extremo de la función.

■ $f(x,y) = \sqrt{x^2+y^2}$ no tiene derivadas parciales en $(0,0)$,

$$f_x(x,y) = \frac{x}{\sqrt{x^2+y^2}}, \qquad f_y(x,y) = \frac{y}{\sqrt{x^2+y^2}},$$

son sus derivadas parciales pero no están definidas en el origen. El origen es un mínimo estricto ya que el gráfico de f es (la parte superior) del cono.

Vemos que no necesariamente en un punto crítico hay un extremo. Ahora miremos los puntos críticos de segunda clase, donde se anulan las derivadas:

■ $f(x) = x^2$ es derivable en $\mathbb{R}$ y su derivada es $f(x) = 2x$ luego su único punto crítico es x_0, y es un mínimo estricto.

■ $f(x) = x^3$, tiene derivada $f'(x) = 3x^2$ que se anula en $x_0 = 0$, pero no es un extremo de f.

■ $f(x,y) = x^2 + y^2$ tiene gradiente $\nabla f(x,y) = (2x, 2y)$, y este se anula únicamente en $X_0 = (0,0)$. Como ya discutimos, es un mínimo estricto.

■ $f(x,y) = x^2 - y^2$ tiene gradiente $\nabla f(x,y) = (2x, -2y)$, y este se anula en $X_0 = (0,0)$. Sin embargo no este punto no es un extremo de f (el gráfico de f es la silla de montar y el origen es el punto de ensilladura).

Nuevamente vemos que no necesariamente en un punto crítico hay un extremo.

DEFINICIÓN 6.1.3 (Punto silla)**.** También llamado punto de ensilladura, es aquel punto crítico X_0 de f donde f no tiene un extremo. Es decir, un punto silla es un punto crítico de f que no es máximo ni mínimo.

¿Por qué el énfasis en buscar los puntos críticos entonces? Por el siguiente teorema:

TEOREMA 6.1.4 (Fermat)**.** *Sea $A \subset \mathbb{R}^n$, sea $f : A \to \mathbb{R}$ tal que f tiene un extremo en $X_0 \in A$. Entonces X_0 es un punto crítico de f.*

ƨ Entonces si buscamos extremos, sabemos que necesariamente estarán en los puntos críticos de la función. Si primero buscamos el conjunto de todos los puntos críticos de f, los extremos (si tiene) estarán entre algunos de estos puntos. Es una manera de reducir el problema: en lugar de buscar extremos en todos los puntos de $A = Dom(f)$, buscamos extremos entre los puntos críticos. Los puntos críticos que veamos que no son extremos, son puntos silla.

OBSERVACIÓN 6.1.5 (Taylor en un punto crítico)**.** Si f es C^3 cerca del punto X_0, teníamos la fórmula de Taylor para f dada por

$$f(X) = f(X_0) + \langle \nabla f(X_0), X - X_0 \rangle + 1/2 \langle Hf(X_0)(X - X_0), X - X_0 \rangle + R(X),$$

Si X_0 es un punto crítico, al ser f suave, debe ser por el Teorema de Fermat $\nabla f(X_0) = \mathbb{O}$. Luego el término lineal se anula y tenemos

$$f(X) = f(X_0) + 1/2 \langle Hf(X_0)(X - X_0), X - X_0 \rangle + R(X).$$

Suponiendo que el resto es pequeño (porque estamos con X cerca de X_0) vemos que cerca de un punto crítico

$$f(X) \simeq f(X_0) + 1/2 \langle Hf(X_0)(X - X_0), X - X_0 \rangle.$$

Supongamos que la matriz Hessiana de f es definida positiva. Esto era que todos sus autovalores sean estrictamente positivos; luego de un cambio de variable adecuado vemos que el término de orden 2 es un paraboloide (hacia arriba)

$$1/2\langle Hf(X_0)(X-X_0), X-X_0\rangle = u^2+v^2 > 0$$

Entonces

$$f(X_0)+1/2\langle Hf(X_0)(X-X_0), X-X_0\rangle = f(X_0)+u^2+v^2 > f(X_0),$$

así que

$$f(X) \simeq f(X_0)+1/2\langle Hf(X_0)(X-X_0), X-X_0\rangle > f(X_0).$$

Esto nos permite concluir que si el Hessiano en el punto crítico X_0 es definido positivo, entonces

$$f(X) > f(X_0),$$

para X cercano a X_0, y entonces X_0 es un mínimo local estricto de f.

Recordemos que si $M = M^t \in \mathbb{R}^{2\times 2}$ es una matriz simétrica, y λ_i son los autovalores entonces

■ M es definida positiva si $\lambda_i > 0$ para todo i.

■ M es definida negativa si $\lambda_i < 0$ para todo i.

■ M es indefinida existen i, j con $\lambda_i \cdot \lambda_j < 0$ (tienen signos opuestos).

■ M es semi-definida si (al menos) algun λ_i es nulo y todos los demás tienen el mismo signo.

Con la idea de la observación anterior (y algunas cosas técnicas más que no mencionaremos) se puede probar el siguiente teorema, que nos da un criterio para decidir qué tipo de punto crítico tenemos.

TEOREMA 6.1.6 (Criterio del Hessiano)**.** *Sea $X_0 \in A \subset \mathbb{R}^n$ punto crítico de $f : A \to \mathbb{R}$ (de clase C^2). Entonces*

1. *Si $Hf(X_0)$ es definido positivo, entonces X_0 es un mínimo local estricto de f.*

2. *Si $Hf(X_0)$ es definido negativo, entonces X_0 es un máximo local estricto de f.*

3. *Si $Hf(X_0)$ es indefinido, entonces X_0 es un punto silla (no es extremo).*

4. *Si $Hf(X_0)$ es semi-definida, el criterio no decide.*

Este criterio también se conoce (en una variable) como *criterio de la derivada segunda*, y se puede enunciar en 1 variable de manera idéntica, sólo que ahora el rol del Hessiano lo cumple la derivada segunda en el punto -se puede pensar que es un caso particular del anterior, aunque aquí hay un sólo número y no dos, por eso el caso indefinido no figura-

TEOREMA 6.1.7 (Criterio en una variable)**.** *Sea $x_0 \in I \subset \mathbb{R}$ punto crítico de $f : I \to \mathbb{R}$ (de clase C^2). Entonces*

1. *Si $f''(x_0) > 0$, entonces x_0 es un mínimo local estricto de f.*

2. *Si $f''(x_0) < 0$, entonces X_0 es un máximo local estricto de f.*

3. *Si $f''(x_0) = 0$, el criterio no decide.*

Veamos en nuestros ejemplos de una variable como funciona esto:

■ $f(x) = x^2$, $f'(x) = 2x$, $f''(x) = 2$. El punto crítico es $x_0 = 0$, y como la derivada segunda es constante y positiva, es un mínimo local estricto.

■ $f(x) = -x^2$, $f'(x) = -2x$, $f''(x) = -2$. El punto crítico es $x_0 = 0$, y como la derivada segunda es constante y negativa, es un máximo local estricto.

■ $f(x) = x^3$, $f'(x) = 3x^2$, $f''(x) = 6x$. El punto crítico es $x_0 = 0$, pero no es un extremo de f como observamos anteriormente (notemos que aquí $f''(0) = 0$).

■ $f(x) = x^4$, $f'(x) = 4x^3$, $f''(x) = 12x$. El punto crítico es $x_0 = 0$, pero $f''(0) = 0$ así que el criterio no nos ayuda. Ya discutimos al comienzo de esta sección que se trata de un mínimo.

■ $f(x) = -x^4$, $f'(x) = -4x^3$, $f''(x) = -12x$. El punto crítico es $x_0 = 0$, pero $f''(0) = 0$ así que el criterio no nos ayuda. Ya discutimos al comienzo de esta sección que se trata de un máximo.

ƨ Regla mnemotécnica: Si es positiva la derivada segunda es como en x^2 luego vemos $\cup$ que es un mínimo. Si es negativa es como en $-x^2$ luego vemos $\cap$ que es un máximo.

Veamos ahora cómo funciona el criterio en dos variables. Los casos de estudio son las superficies cuádricas:

■ $f(x,y) = x^2 + y^2$, luego $\nabla f(x,y) = (2x, 2y)$ con punto crítico $X_0 = (0,0)$. Calculamos las derivadas segundas y tenemos

$$f_{xx}(x,y) = 2, f_{yy}(x,y) = 2, f_{xy}(x,y) = f_{yx}(x,x) = 0$$

luego la matriz Hessiana es constante e igual a

$$Hf(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

En particular así es el Hessiano en el punto crítico $(0,0)$. Esta matriz ya está diagonalizada, sus autovalores son $\lambda_1 = \lambda_2 = 2$, ambos estrictamente positivos. Entonces $Hf(0,0)$ es definido positivo, así que por el criterio del Hessiano, $(0,0)$ es un mínimo local estricto de f (cosa que ya sabíamos por otros medios).

■ $f(x,y) = -x^2 - y^2$, luego $\nabla f(x,y) = (-2x, -2y)$ con punto crítico $X_0 = (0,0)$, calculamos las derivadas segundas y tenemos

$$f_{xx}(x,y) = -2, f_{yy}(x,y) = -2, f_{xy}(x,y) = f_{yx}(x,y) = 0$$

luego la matriz Hessiana es constante e igual a

$$Hf(x,y) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

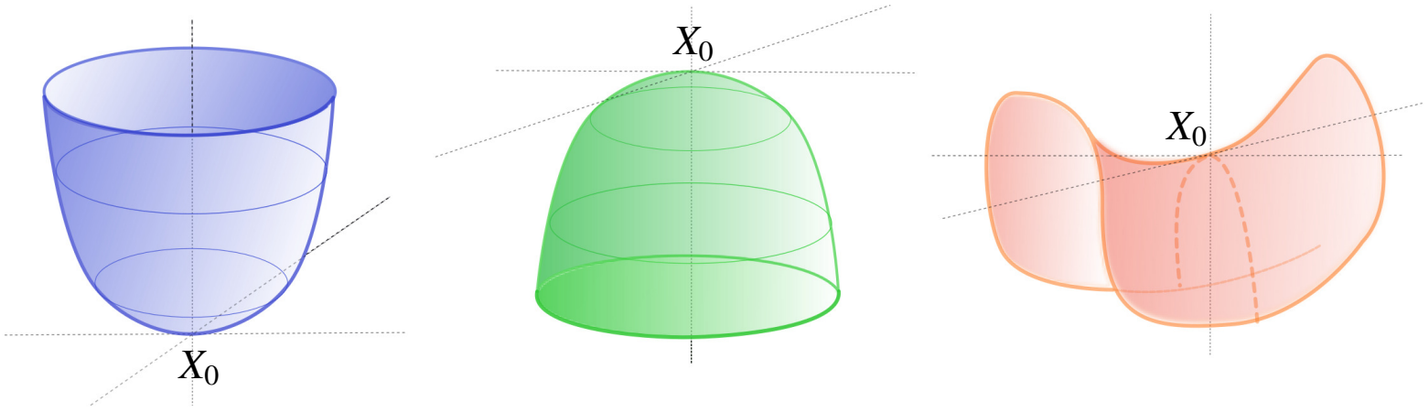


Figura 6.1: Puntos críticos X_0 con Hessiano definido positivo, negativo e indefinido respectivamente. Cerca del punto X_0 la superficie tiene la forma de una de estas.

En particular así es el Hessiano en el punto crítico $(0,0)$. Esta matriz ya está diagonalizada, sus autovalores son $\lambda_1 = \lambda_2 = -2$, ambos estrictamente negativos. Entonces $Hf(0,0)$ es definido negativo, así que por el criterio del Hessiano, $(0,0)$ es un máximo local estricto de f.

■ $f(x,y) = x^2 - y^2$, luego $\nabla f(x,y) = (2x, -2y)$ con punto crítico $X_0 = (0,0)$. Calculamos las derivadas segundas y tenemos

$$f_{xx}(x,y) = 2, f_{yy}(x,y) = -2, f_{xy}(x,y) = f_{yx}(x,y) = 0$$

luego la matriz Hessiana es constante e igual a

$$Hf(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

En particular así es el Hessiano en el punto crítico $(0,0)$. Esta matriz ya está diagonalizada, sus autovalores tienen signos opuestos. Entonces $Hf(0,0)$ es indefinido, así que por el criterio del Hessiano, $(0,0)$ es un punto silla de f.

ƨ Veamos ahora por qué el criterio no decide cuando hay un autovalor nulo (y todos los demás tienen el mismo signo). El problema es que con un autovalor nulo, el resto de Taylor R (que despreciamos en la Observación 6.1.5) si juega un papel en cada caso particular.

■ $f(x,y) = x^2 - y^4$, luego $\nabla f(x,y) = (2x, -4y^3)$ con punto crítico $X_0 = (0,0)$. Calculamos las derivadas segundas y tenemos

$$f_{xx}(x,y) = 2, f_{yy}(x,y) = -12y^2, f_{xy}(x,y) = f_{yx}(x,y) = 0$$

luego la matriz Hessiana es

$$Hf(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Los autovalores son $\lambda_1 = 2$, $\lambda_2 = 0$, es semi-definida. Pero esto no nos ayuda porque la función *no tiene un mínimo* en $(0,0)$. Para verlo nos acercamos al origen por el eje y, notamos que $f(0,y) = -y^4 < 0 = f(0,0)$ así que 0 no es mínimo. El dibujo de la superficie $z = x^2 - y^4$ (que es el gráfico de f) es muy parecido al de la silla de montar. Entonces $(0,0)$ es punto silla de f.

En general, como sólo queremos saber qué signo tienen los autovalores (y no exactamente cuánto valen) podemos usar el siguiente criterio

TEOREMA 6.1.8 (Criterio del determinante 2×2)**.** *Sea*

$$M = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

matriz simétrica 2×2*. Entonces*

1. *M es semi-definida* $\iff$ $\det(M) = 0$.
2. *M es definida positiva* $\iff$ $(\det(M) > 0 \wedge a_{11} > 0)$.
3. *M es definida negativa* $\iff$ $(\det(M) > 0 \wedge a_{11} < 0)$.
4. *M es indefinida* $\iff$ $\det(M) < 0$.

Aunque no lo probaremos, la idea de la prueba se basa en diagonalizar la matriz M y luego observar que la regal funciona bien cuando tenemos una matriz diagonal. Esto es porque

$$ac - b^2 = \det(M) = \lambda_1 \lambda_2$$

(el producto de los autovalores). Y además $a_{11} = \langle ME_1, E_1 \rangle$. Entonces es claro que el signo del determinante depende de los signos de los autovalores, y luego la entrada $1-1$ de la matriz es un caso particular

de la cuenta $\langle MV, V\rangle$ que nos permite decidir si M es definida negativa o positiva (ver el final del Resumen 2).

ƨ Podemos recordar el criterio del determinante en 2×2 con estos tres diagramas, que corresponden respectivamente a una matriz definida positiva, una definida negativa y una indefinida:

$$\begin{pmatrix} + & 0 \\ 0 & + \end{pmatrix} \qquad \begin{pmatrix} - & 0 \\ 0 & - \end{pmatrix} \qquad \begin{pmatrix} + & 0 \\ 0 & - \end{pmatrix}$$

El criterio dice que podemos encontrar los signos de los autovalores sin necesidad de diagonalizar la matriz ni de hallar los autovalores, y la manera de recordar el criterio es "imaginar" que la matriz está diagonalizada, entonces estamos en alguno de estos tres casos (siempre que el determinante sea no nulo).

Veamos cómo usar los teoremas y criterios en un ejemplo concreto:

EJEMPLO 6.1.9. Sea $f(x,y) = e^{x+y}\cdot(x^2-2y^2)$, hallar los extremos relativos de f. Calculamos su gradiente y lo igualamos a cero para hallar los puntos críticos

$$f_x(x,y) = e^{x+y}\cdot 1\cdot(x^2-2y^2) + e^{x+y}\cdot(2x) = e^{x+y}\cdot(x^2-2y^2+2x) = 0$$

$$f_y(x,y) = e^{x+y}\cdot 1\cdot(x^2-2y^2) + e^{x+y}\cdot(-4x) = e^{x+y}\cdot(x^2-2y^2-4y) = 0.$$

Obtenemos dos ecuaciones que se tienen que cumplir simultáneamente. El factor con la exponencial es nunca nulo así que lo podemos cancelar en ambas. Nos quedan las ecuaciones

$$x^2-y^2+2x=0 \qquad x^2-2y^2-4y=0.$$

Las ecuaciones son no lineales, pero son parecidas. Entonces restando la primera de la segunda nos queda $2x+4y=0$ de donde deducimos que $x=-2y$. Luego $x^2=4y^2$ y reemplazando esto en la segunda ecuación nos queda

$$4y^2-2y^2-4y=0$$

es decir $2y^2-4y=0$. Se factoriza como $2y(y-2)=0$, así que tiene

que ser $y = 0$ o bien $y = 2$. Cuando $y = 0$, debe ser también $x = 0$ (puesto que era $x = -2y$). Obtenemos $P_1 = (0,0)$. Cuando $y = 2$ debe ser $x = -4$, obtenemos el punto $P_2 = (-4,2)$. Luego f tiene dos puntos críticos.

Veamos de qué tipo son, para eso necesitamos calcular las derivadas segundas de f. Son

$$f_{xx}(x,y) = e^{x+y} \cdot (x^2 - 2y^2 + 4x + 2),\ f_{yy}(x,y) = e^{x+y} \cdot (x^2 - 2y^2 - 8y - 5),$$

$$f_{xy}(x,y) = f_{yx}(x,y) = e^{x+y} \cdot (x^2 - 2y^2 + 2x - 4y).$$

Tenemos

$$Hf(P_1) = \begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix}$$

y es claro que este Hessiano es indefinido. Entonces $P_1 = (0,0)$ es punto silla de f.

Por otro lado tenemos el otro punto crítico $P_2 = (-4,2)$ donde

$$Hf(P_2) = \begin{pmatrix} e^{-2} \cdot (-6) & e^{-2} \cdot (-8) \\ e^{-2} \cdot (-8) & e^{-2} \cdot (-12) \end{pmatrix}.$$

Es claro que $a_{11} = -6e^{-2} < 0$, y por otro lado el cálculo del determinante arroja

$$\det(Hf(P_1)) = (e^{-2})^2(6 \cdot 12 - 8 \cdot 8) = e^{-4} \cdot 8 > 0.$$

El criterio del determinante nos dice que este Hessiano es definido negativo, y entonces f tiene un máximo local estricto en $P_2 = (-4,2)$.

2 Esta función no tiene ningún mínimo local, y tiene un sólo máximo local. Puede verse su gráfico en el applet de Geogebra.

Para estudiar extremos de funciones de 3 variables, también hay un criterio para decidir los signos de los autovalores usando el determinante, que es útil aunque algo más técnico:

TEOREMA 6.1.10 (Criterio del determinante 3×3)**.** *Sea*

$$M = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

matriz simétrica 3×3*, a la sub-matriz de* 2×2

$$MP = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

se la denomina menor principal de M. Entonces M es

- *semi-definida* $\iff \det(M) = 0.$
- *definida positiva* $\iff (\det(M) > 0 \wedge \det(MP) > 0 \wedge a_{11} > 0).$
- *definida negativa* $\iff (\det(M) < 0 \wedge \det(MP) > 0 \wedge a_{11} < 0).$
- *M es indefinida en todos los demás casos (siempre que* $\det(M) \neq 0$*).*

Nuevamente, el criterio nos dice qué signos tienen los autovalores de M sin necesidad de hallarlos ni de diagonalizar M; para recordar el criterio es útil recordar en primer lugar que

$$\det(M) = \lambda_1\lambda_2\lambda_3,$$

el determinante siempre es el producto de los autovalores.

Descartado el caso de algún autovalor nulo (que equivale a determinante nulo) para saber si es positiva, negativa o indefinida son de ayuda los siguientes dos diagramas que corresponden al caso definido positivo y definido negativo respectivamente:

$$\begin{pmatrix} + & 0 & 0 \\ 0 & + & 0 \\ 0 & 0 & + \end{pmatrix} \qquad \begin{pmatrix} - & 0 & 0 \\ 0 & - & 0 \\ 0 & 0 & - \end{pmatrix}$$

6.2. Optimización

• Corresponde a clases en video 6.4a, 6.4b

Recordemos aquí el

TEOREMA 6.2.1 (Weierstrass)**.** *Sea $K \subset \mathbb{R}^n$, sea $f : K \to \mathbb{R}$. Entonces f es acotada y alcanza máximo y mínimo en K.*

El mismo nos será de utilidad para justificar por qué los extremos que hallaremos serán absolutos en los ejemplos que siguen. En esta sección discutiremos cómo modelar un problema con una función y hallar el óptimo (la solución del problema) buscando los extremos de la función.

EJEMPLO 6.2.2 (Optimización)**.** Queremos construir una caja con un volumen de $1000cm^3$, de manera tal que la superficie total de la caja sea mínima. ¿Qué dimensiones tienen que tener la caja?

Solución: Si la caja tiene dimensiones x, y, z (es un prisma rectangular, o sea un ladrillo), entonces la restricción es que $V = xyz = 1000$. Por otro lado, la superficie de las tapas sumadas es

$$S = 2xy + 2xz + 2yz.$$

Podemos usar la restricción $z = 1000/(xy)$ para escribir la función a minimizar

$$S(x,y) = 2xy + 2x\frac{1000}{xy} + 2y\frac{1000}{xy} = 2xy + 2000\frac{1}{y} + 2000\frac{1}{x},$$

sabiendo que debe ser $x, y, z > 0$ puesto que son las longitudes de los lados.
Antes de seguir, observamos que si x ó y tienden a 0, la superficie tiende a infinito. Y también que ocurre lo mismo cuando x ó y tienden a infinito. Entonces la función S no tiene máximo en la región

$$A = \{(x,y) : x > 0,\ y > 0\}.$$

Veamos que tiene un mínimo local, que es además mínimo global. Para ello calculamos las derivadas parciales de S y las igualamos a cero para hallar los puntos críticos:

$$S_x(x,y) = 2y - \frac{2000}{x^2} = 0, \qquad S_y(x,y) = 2x - \frac{2000}{y^2} = 0.$$

La primer ecuación nos dice que $x^2y = 1000$, mientras la segunda nos dice que $xy^2 = 1000$. Entonces debe ser $x^2y = xy^2$, y como $x > 0, y > 0$ podemos cancelar y debe ser $x = y$, pero de $x^2y = 1000$ deducimos que $x^3 = 1000$, luego $x = 10$. Como $y = x$ también debe ser $y = 10$. El único punto crítico es $P = (10,10)$, allí la función vale

$$S(10,10) = 2 \cdot 10^2 + 200 + 200 = 600.$$

El punto P es mínimo global de S en la región A indicada más arriba, nuestra intuición nos dice que tiene que haber un mínimo, y como hay un único punto crítico tiene que ser en este punto.
Entonces el mínimo absoluto se alcanza cuando $x = 10, y = 10, z = 10$ (notar que es cuando la caja es cúbica).

OBSERVACIÓN 6.2.3. El número del volumen $V = 1000$ es anecdótico. Revisando el razonamiento y las cuentas, vemos que la caja rectangular con menor superficie (para un volumen dado) siempre tiene todos sus lados iguales, es un cubo.

ƨ El argumento que dimos para concluir que el mínimo es absoluto es un poco impreciso, ya que por ejemplo, la función podría tomar valores menores que 600 sin tener un mínimo local, ya que el dominio A no es acotado (o sea que nuestra justificación no es tan buena como decíamos).

Un argumento más preciso sería: como S tiende a infinito tanto en los bordes de la región como en infinito, podemos hallar un rectángulo R dentro de A (que contenga al punto P) de manera tal que fuera del rectángulo (y en el borde) la función sea estrictamente mayor que 600.

Restringimos la función S al rectángulo R y buscamos el mínimo absoluto de $S|_R$ (que existe por el Teorema de Weirstrass). El mínimo no puede estar en el borde porque allí es estrictamente mayor que en P, donde S vale 600. Entonces el mínimo se alcanza en un punto del interior de R, y en ese punto el gradiente tiene que ser nulo pues en particular es mínimo local. Pero como el único punto donde se anula el gradiente de S es P, el mínimo absoluto tiene que ser P.

Ahora bien, esto de que la función a minimizar tiende a infinito en los ejes y en infinito, también es un tanto impreciso, y deberíamos poder mostar el recinto R donde estamos seguros que debe estar el óptimo. No vamos a hacer esto en general, pero para este ejemplo particular vamos a mostrar cómo se puede hacer.

Será útil el siguiente cálculo auxiliar: para $t > 0$, la función $g(t) = t + \frac{40}{t}$ tiene un mínimo absoluto en $t = 2\sqrt{10}$, y el valor mínimo es $4\sqrt{10}$. Esto es fácil de ver haciendo un estudio del crecimiento y decrecimiento de la función: como $g'(t) = 1 - \frac{40}{t^2}$, la derivada es negativa para $t \in (0, 2\sqrt{10})$ y positiva para $t > \sqrt{10}$.

Supongamos que $y \geq 25$, entonces

$$\begin{aligned} S(x,y) &= 2xy + \frac{2000}{y} + \frac{2000}{x} \geq 50x + \frac{2000}{x} \\ &= 50(x + \frac{40}{x}) \geq 50 \cdot 4\sqrt{10} \simeq 632 > 600 = S(P). \end{aligned}$$

Similarmente, cuando $x \geq 25$ entonces $S(x,y) > 600$. Ahora supongamos que $0 < y \leq 1$, entonces $1/y \geq 1$ y así

$$S(x,y) = 2xy + \frac{2000}{y} + \frac{2000}{x} > \frac{2000}{y} \geq 2000 > 600 = S(P).$$

Con el mismo argumento $S(x,y) > 600 = S(P)$ si $0 < x \leq 1$.
Ya tenemos nuestra región, ahora podemos precisar el argumento para terminar de probar que el mínimo absoluto de S está en el punto $P = (10,10)$. Sea

$$R = \{(x,y) : 1 \leq x \leq 25,\ 1 \leq y \leq 25\},$$

las cotas que encontramos recién nos dicen que en el borde y por fuera de R tenemos $S > 600$. Buscamos el mínimo absoluto de $S|_R$ (que existe ya que S es continua y R es compacto). Tiene que estar en el interior, porque $P \in R^\circ$ y alli la función vale 600 (mientras que en el borde es estrictamente mayor). Como el mínimo tiene que estar en el interior, que es abierto, el gradiente tiene que ser nulo allí. Como el único punto crítico que hallamos fue P, ese tiene que ser el mínimo absoluto de $S|_R$, y por ende es el mínimo absoluto de $S|_A$.

EJEMPLO 6.2.4. Queremos construir de nuevo una caja pero ahora tenemos $12m^2$ de cartón para las 6 tapas, entonces ¿Cuáles son las dimensiones de la caja para que el volumen contenido sea máximo?
El volumen de la caja es nuevamente $V = xyz$ donde x, y, z son las variables positivas que representan las longitudes de los lados de la caja. La restricción para la superficie es que

$$S = 2xy + 2xz + 2yz = 12.$$

De esta restricción podemos despejar una de las variables, nuevamente elegimos despejar z, como $2xy + 2(x+y)z = 12$ obtenemos

$$z = \frac{12 - 2xy}{2(x+y)} = \frac{6 - xy}{x+y}.$$

Antes de seguir vamos a hacer una observación: como debe ser $z > 0$ tenemos una restricción adicional que es

$$\frac{6 - xy}{x+y} > 0,$$

como $x + y > 0$ (pues tanto x como y lo son) debe ser $6 - xy > 0$, o equivalentemente $xy < 6$.
Entonces la función volumen a optimizar es

$$V(x,y) = xy \cdot \frac{6 - xy}{x+y} = \frac{6xy - x^2y^2}{x+y}$$

en el dominio

$$A = \{(x,y) : x > 0, y > 0, xy < 6\}.$$

Vamos a hallar los puntos críticos de V en la región A, para eso vemos dónde se anula el gradiente

$$V_x(x,y) = \frac{-y^2(x^2+2xy-6)}{(x+y)^2} = 0,$$

$$V_y(x,y) = \frac{-x^2(y^2+2xy-6)}{(x+y)^2} = 0,$$

Descartada la solución nula, tenemos dos ecuaciones

$$x^2+2xy-6=0, \qquad y^2+2xy-6=0.$$

Si las restamos obtenemos $x^2 = y^2$ y como ambos son positivos debe ser $x = y$. Reemplazamos en la primer ecuación y nos queda $x^2+2x^2-6 = 0$ lo que nos dice que $3x^2 = 6$, luego $x^2 = 2$ y así debe ser $x = y = \sqrt{2}$. El único punto crítico de V es $P = (\sqrt{2}, \sqrt{2})$ y allí el volumen vale

$$V(\sqrt{2}, \sqrt{2}) = 2 \cdot \frac{6-2}{2\sqrt{2}} = \frac{4}{\sqrt{2}} = 2\sqrt{2}$$

(en el último paso racionalizamos la fracción). Afirmamos que este es el máximo absoluto de la función V en la región A. El argumento es similar al del ejemplo anterior:

Podemos calcular z pues lo teniamos en función de x, y. Entonces el volumen máximo se alcanza cuando $x = y = z = \sqrt{2}$, y este volumen es $V = 2\sqrt{2}$.

OBSERVACIÓN 6.2.5. Nuevamente podemos observar que el óptimo se alcanza cuando los tres lados son iguales, y la caja es un cubo. Esto es válido para cualquier restricción de superficie: la caja que tiene mayor volumen (para una superficie dada) es cúbica.

6.3. Extremos absolutos

• Corresponde a clase en video 6.5

Nuevamente enunciamos el Teorema de Weierstarass, ya que en esta sección estudiaremos extremos de funciones en regiones compactas.

TEOREMA 6.3.1 (Weierstrass)**.** *Sea $K \subset \mathbb{R}^n$, sea $f : K \to \mathbb{R}$. Entonces f es acotada y alcanza máximo y mínimo en K.*

Entonces podemos garantizar que restrigiendo adecuadamente el dominio, una función continua siempre tendrá máximo y mínimo (que se puede alcanzar en uno o más puntos del dominio, como discutimos anteriormente). Lo que hay que tener presente, es que los extremos se pueden alcanzar tanto dentro del dominio como en los bordes. Comencemos con una variable.

EJEMPLO 6.3.2 (Extremos absolutos en un intervalo)**.** Hallar los extremos absolutos de

$$f(x) = 10x \cdot e^{-2x^2+3x}$$

en el intervalo $[-1, 0]$. Calculamos la derivada

$$\begin{aligned} f'(x) &= 10 \cdot e^{-2x^2+3x} + 10x \cdot e^{-2x^2+3x} \cdot (-4x+3) \\ &= 10 \cdot e^{-2x^2+3x} \cdot (-4x^2+3x+1) \end{aligned}$$

y buscamos los puntos críticos igualandola a cero. Tenemos que buscar las raíces de

$$-4x^2+3x+1 = 0$$

que son $x_1 = -1/4$, $x_2 = 1$ (pero el segundo no está en el intervalo que nos interesa). Agregando los bordes, los puntos críticos en nuestro intervalo son

$$p.c. = \{-1, -1/4, 0\}.$$

No es necesario hacer un estudio del tipo de punto crítico, porque buscamos extremos absolutos. Hacemos una lista del valor de f en cada

punto crítico y los comparamos. El valor máximo y el valor mínimo es lo que buscamos:

x	$f(x)$				
-1	$10e^{-2-3}\cdot(-1)$	$-10e^{-5}$	$\simeq$	$-0,066$	
$-1/4$	$10e^{-1/8-3/4}\cdot(-1/4)$	$-5/2\cdot e^{-7/8}$	$\simeq$	$-1,04$	mín.
0	$10e^{0}\cdot 0$	0	$\simeq$	0	máx.

Notamos que el mínimo se alcanza en el interior del intervalo $[-1,0]$, pero el máximo se alcanza en un borde.

Vamos a terminar este texto con un ejemplo de extremos absolutos para una función de dos variables.

EJEMPLO 6.3.3. Hallar los extremos absolutos de

$$f(x,y) = x^2 + y^2 - 2x - 2y$$

en el conjunto $K = \{(x.y) : x^2 + y^2 \leq 4\}$. Claramente f es continua y el conjunto es compacto, así que existen máximo y mínimo absoluto de $f|_K$.
Para hallarlos busquemos primero los puntos críticos en el interior del conjunto, para eso buscamos donde se anula el gradiente de f

$$f_x(x,y) = 2x - 2 = 0, \qquad f_y(x,y) = 2y - 2 = 0.$$

Encontramos el punto $P = (1,1)$ como único punto crítico, y vemos que está en el interior de K (que es un disco de radio 2). Ahora queremos estudiar $f|_{bd(K)}$ pero a diferencia del caso de una variable, el borde es toda una curva (son infinitos puntos).
Lo que vamos a hacer es estudiar f con esa restricción, para eso notamos que la curva

$$\alpha(t) = (2\cos(t), 2\,\mathrm{sen}(t))$$

recorre el borde de K cuando $t \in [0, 2\pi]$. Entonces le damos a (x,y)

esos valores y estaremos evaluando f únicamente en el borde,

$$t \mapsto f(2\cos(t), 2\operatorname{sen}(t)).$$

Llamemos $g(t)$ a esta composición (g es una función auxiliar). Tenemos que estudiar entonces los extremos en el intervalo $I = [0, 2\pi]$ de la función

$$\begin{aligned} g(t) &= f(2\cos(t), 2\operatorname{sen}(t)) \\ g(t) &= (2\cos(t))^2 + (2\operatorname{sen}(t))^2 - 2 \cdot 2\cos(t) - 2 \cdot 2\operatorname{sen}(t) \\ g(t) &= 4 - 4\cos(t) - 4\operatorname{sen}(t). \end{aligned}$$

En este caso entonces, redujimos el problema de estudiar f en el borde de K a estudiar una función de un variable en un intervalo compacto. Es un problema auxiliar, lo resolvemos como hicimos con el primer ejemplo: tenemos $t = 0, t = 2\pi$ como puntos críticos por estar en el borde del intervalo y ahora buscamos también los ceros de la derivada de g en el interior del intervalo

$$g'(t) = 4\operatorname{sen}(t) - 4\cos(t) = 0,$$

de donde vemos que debe ser $\cos(t) = \operatorname{sen}(t)$. Notemos que si el coseno se anula (en $\pi/2$ y también en $3\pi/2$), la ecuación no se verifica porque el seno no se anula allí. Entonces podemos suponer que t no es ninguno de estos dos números y dividir por $\cos(t)$ para ver que debe ser $tg(t) = 1$. Las soluciones de esta ecuación son 2: tenemos $t_1 = \pi/4$ y también $t_2 = 5\pi/4$ (también puede verse que en esos dos puntos de la circunferencia trigonométrica es donde el seno es igual al coseno).

Resumiendo los puntos críticos de g en el intervalo $[0, 2\pi]$ son

$$p.c. = \{0, \pi/4, 5\pi/4, 2\pi\}.$$

Pero ¿a qué puntos del borde de la circunferencia corresponden estos valores de t? Es fácil de calcular, ya que estamos recorriendo la cir-

cunferencia de radio 2 con el parámetro del arco. Reemplazamos en la parametrización $\alpha(t)$ y vemos que corresponden a los puntos

$$p.c. = \{(2,0)\,;\,(\sqrt{2},\sqrt{2})\,;\,(-\sqrt{2},-\sqrt{2})\,;\,(2,0)\}$$

Estos son los puntos críticos de $f|_{bd(K)}$, el punto $(2,0)$ aparece repetido porque la curva que usamos para parametrizar el borde comienza y termina en el mismo lugar. Entonces tenemos que sumar estos puntos críticos al del interior $P = (1,1)$ y esos son todos los puntos críticos de $f|_K$, los extremos absolutos tienen que estar entre ellos. Hacemos la tabla de valores y buscamos máximo y mínimo:

(x,y)	$f(x,y)$				
$(1,1)$	$1+1-2-2$	-2	$\simeq$	-2	mín.
$(2,0)$	$4+0-4-0$	0	$\simeq$	0	
$(\sqrt{2},\sqrt{2})$	$2+2-2\sqrt{2}-2\sqrt{2}$	$4-4\sqrt{2}$	$\simeq$	$-1,65$	
$(-\sqrt{2},-\sqrt{2})$	$2+2+2\sqrt{2}+2\sqrt{2}$	$4+4\sqrt{2}$	$\simeq$	$9,65$	máx.

Esta función alcanza máximo absoluto en el borde del disco de radio 2, y el mínimo absoluto en el interior. Puede verse su gráfico en el applet de Geogebra.

Estos ejemplos concluyen nuestra presentación de los temas de extremos absolutos de funciones, con los que cerramos el texto y nos despedimos; ojalá les haya sido útil y a la vez entretenido.

Índice alfabético

E

F

G

H

I

L

M

N

O

P

R

S

T

V

Bibliografía

[1] R. Courant, J. Fritz, *Introducción al cálculo y el análisis matemático*. Vol. 1 y 2, Ed. Limusa-Wiley , Méjico, 1998.

[2] S. Lang, *Introducción al álgebra lineal*. Ed. Addison-Wesley Iberoamericana, Argentina, 1990.

[3] G. Larotonda, *Cálculo y Análisis*. Publicaciones del Departamento de Matemática de la FCEyN-UBA (2010).

[4] J.E. Marsden, A.J. Tromba, *Cálculo vectorial*. Ed. Addison-Wesley Iberoamericana, Argentina, 1991.

[5] R. J. Noriega, *Cálculo diferencial e integral*. Ed. Docencia, Buenos Aires, 1987.

[6] J. Stewart *Cálculo en varias variables*. 7ma edición, Cengage Learning Editores (2012).