

CEU | *Universidad
Cardenal Herrera*

MATEMÁTICAS

DE LA TITULACIÓN DE *Administración y Dirección
de Empresas.*

Profesor: Antonio Falcó Montesinos

22 de diciembre de 2008



Capítulo 1

Números reales

1.1. Introducción

Los números reales pueden ser representados por puntos en una línea. En dicha línea se elige un punto que será llamado cero y se elige una medida de distancia. A cada punto de la recta se le asocia su distancia al cero y un signo dependiendo de si está a la derecha (positivo) o a la izquierda (negativo) del cero.

Para los números reales se definen las operaciones suma (+) y producto ($\cdot$) con las siguientes propiedades:

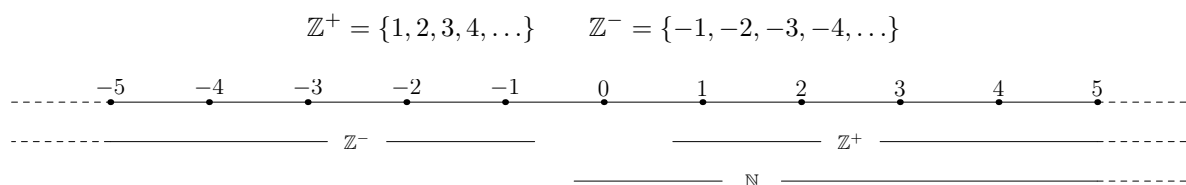
- | | |
|---|---|
| 1. Propiedad asociativa
$(x + y) + z = x + (y + z) \quad (x, y, z \in \mathbb{R})$ | 1. Propiedad Asociativa: $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ |
| 2. Propiedad conmutativa
$x + y = y + x \quad (x, y \in \mathbb{R})$ | 2. Propiedad Conmutativa: $x \cdot y = y \cdot x \quad (x, y \in \mathbb{R})$ |
| 3. Existencia de Elemento Neutro:
$x + 0 = x.$ | 3. Existencia de Elemento unidad $x \cdot 1 = x.$ |
| 4. Existencia de Elemento Simétrico.
$x + (-x) = 0.$ | 4. Existencia de Elemento inverso. Si $x \neq 0$, existe otro real $1/x$, tal que: $x \cdot (1/x) = 1.$ |
| | 5. Propiedad Distributiva
$(x + y) \cdot z = x \cdot z + y \cdot z \quad (x, y, z \in \mathbb{R})$ |

La propiedad asociativa significa que es innecesario el uso de paréntesis cuando sólo aparezca esta operación.

1.2. Números enteros

Los números naturales son: $0, 1, 2, 3, \dots$. Surgen de la idea del orden de una secuencia (ordinal) y de la necesidad de contar los elementos en un conjunto (cardinal). Por ejemplo, la población de un país es un número natural. Al conjunto de todos los números naturales se les representa por $\mathbb{N}$.

Un número entero es un número natural al que se le antepone un signo positivo (+) o negativo (-). Los números enteros se denotan por $\mathbb{Z}$. El conjunto $\mathbb{Z}^+$ es el conjunto de los enteros positivos no nulos y el conjunto $\mathbb{Z}^-$ es el conjunto de los enteros negativos no nulos.



La necesidad de incluir el signo viene de situaciones en las que una cantidad puede estar en el haber o ser debida. Podemos escribir $+m$ si la cantidad se encuentra en el haber y negativa $-m$ si debemos m unidades.

Si se tienen 5 monedas en bolsillo y debes 8 monedas, entonces en el balance debes 3 monedas, $(+5) + (-8) = (5 - 8) = -3$. Si solicitamos n veces un préstamo de m , lo que tenemos es una deuda de la cantidad $n \cdot m$, es decir, $n \cdot (-m) = -n \cdot m$. Multiplicar por -1 una cantidad es cambiar el haber a débito o viceversa, es decir, intercambiar los papeles de acreedor y el deudor.

Dado un entero m y n con n no nulo, existen enteros q y r tales que $0 \leq r < n$ y $m = q \cdot n + r$. A q se le llama *cociente* de la *división euclídea* de m entre n y a r se le llama resto de dicha división euclídea. La idea de la división euclídea es repartir m elementos entre n y ver los que nos sobran.

Cuando el resto de la división euclídea de m entre n es cero se dice que n es un *divisor* de m .

1.3. Números racionales

Aunque no podamos repartir 2 entre 3, si tiene sentido hablar de dicha proporción. Por ejemplo, antiguamente se podía efectuar un trueque de 2 vacas por 3 mulas. ¿Cuál es el precio de una mula si contamos en vacas? Una mula vale $2/3$ de una vaca, lo que significa que dos vacas equivalen a tres mulas. En una sociedad donde funcionase el trueque estarían utilizando continuamente proporciones de este estilo.

Los números racionales se escriben como m/n donde m y n son enteros, siendo n no nulo. El número m se denomina *numerador* y el número n se llama *denominador*. Ejemplos de números racionales son

$$\frac{5}{3}, \quad \frac{1}{2}, \quad -\frac{5}{8}, \quad \frac{1}{-5}, \quad \frac{0}{5}.$$

Lo que nos importa es la relación que hay entre las dos cantidades, por ejemplo, afirmar que el trueque de 4 vacas es por 6 mulas es lo mismo que 2 vacas se cambien por tres mulas. Por tanto, $4/6 = 2/3$. En consecuencia, si multiplicamos numerador y denominador por el mismo número, la fracción no cambia. Decimos que dos números racionales a/b y c/d son *iguales* o *equivalentes* si $a \cdot d = b \cdot c$.

La suma y el producto se definen como

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + cb}{bd} \quad \text{y} \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d}$$

Cuando a y b son no nulos y no tienen divisores comunes se dice que la fracción a/b es *irreducible*. Toda fracción con numerador no nulo se puede escribir como una fracción a/b irreducible.

1.4. Notación decimal.

Los números se escriben con un sistema de notación posicional que es denominado *representación decimal*. Se puede representar cualquier número con una cadena formada por los símbolos 0,1,2,3,4,5,6,7,8,9. A cada uno de los símbolos en la cadena se les llama dígitos. Los dígitos a la izquierda de la coma hay que multiplicarlos por $1/10^m$ siendo m el número de lugares a la izquierda de la coma

$$0.125 = 1 \cdot \frac{1}{10} + 2 \cdot \frac{1}{10^2} + 5 \cdot \frac{1}{10^3}.$$

Una simple operación nos permite escribir el número como fracción.

$$0.125 = \frac{125}{1000} = \frac{25}{200} = \frac{5}{40} = \frac{1}{8}.$$

Hay fracciones cuya representación decimal tendría infinitas cifras. Por ejemplo,

$$\frac{1}{3} = 0.33333333333333333333 \dots$$

Se puede probar que la representación decimal de una fracción tiene un número finito de dígitos o existe una secuencia de dígitos que se repiten. Para representar que los dígitos se repiten indefinidamente escribimos una barra sobre las cifras

$$1.\overline{142857} = 1.142857142857142857 \dots$$

Si dicho número ($q = 1.\overline{142857}$) lo multiplicamos por 10^m donde m es el número de cifras que se repiten, nos queda otro número en el que se repite la misma secuencia de dígitos. Al restar $10^m q - q$ nos queda un número con un número finito de cifras decimales.

$$10^6 q - q = 1142857.\overline{142857} - 1.\overline{142857} = 1142856$$

Despejando q y simplificando hasta encontrar la fracción irreducible, obtenemos:

$$q = \frac{1142856}{999999} = \dots = \frac{8}{7}$$

1.5. Incompletitud de los números racionales

A cualquier número real le corresponde al menos una representación decimal.¹ Éstas no se corresponden con números racionales cuando exista una secuencia infinita de dígitos en las que no es posible encontrar ninguna periodicidad, diciéndose en este caso que son números *irracionales*.

Hay proporciones físicas que se corresponden con números irracionales, por ejemplo, la relación entre el lado de un cuadrado y su diagonal es $\sqrt{2}$.

$$\sqrt{2} = 1.414213562373095048801688724209698078569671875376948073176679737990732 \dots$$

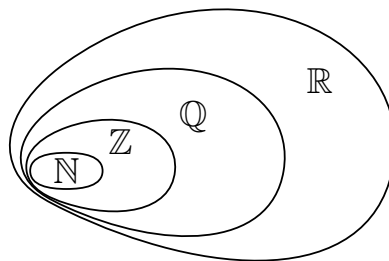
Un número irracional de importancia en economía es el número e

$$e = 2.718281828459045235360287471352662497757247093699959574966967627724076 \dots;$$

pero sin lugar a duda el número irracional más famoso es la relación entre la longitud de la circunferencia y su diámetro: el número π

$$\pi = 3.141592653589793238462643383279502884197169399375105820974944592307816 \dots$$

Todo número natural es un número entero, todo número entero es racional, todo racional es real. Más formalmente, se dice que $\mathbb{N}$ es un subconjunto de $\mathbb{Z}$, $\mathbb{Z}$ es subconjunto de $\mathbb{Q}$ y $\mathbb{Q}$ es un subconjunto de $\mathbb{R}$. Esta afirmación puede ser ilustrada gráficamente con el diagrama que aparece a la derecha de este párrafo.



1.6. Ecuaciones

Una ecuación es la afirmación de que dos expresiones son iguales. Las dos expresiones de la igualdad son llamados lados y son separadas por el símbolo de igualdad ($=$). Una *variable de la ecuación* es un símbolo que puede ser representado por un conjunto de valores. Por *resolver* una ecuación se entiende que es encontrar todos los valores de las variables para los que se verifica la ecuación.

¹Los números con un número finito de cifras decimales tienen dos representaciones. Por ejemplo, $12.23 = 12.22\overline{9}$

Ejemplo 1.1. Una persona invierte 1000€ en una cartera con dos tipos de bonos con un seis y un siete por ciento de rentabilidad. Si al año siguiente su cartera está valorada en 1064€, ¿cuál era la composición de su cartera?

Si la cantidad invertida en el primer bono es x (asignamos un símbolo arbitrario a la variable), entonces la cantidad invertida en el segundo bono es $1000 - x$. Por tanto, el dinero ganado por los bonos es $0.06x + 0.07 \cdot (1000 - x)$ y por otro lado es 64 €. Igualando ambas expresiones se obtienen los dos lados de la igualdad:

$$0.06x + 0.07(1000 - x) = 64.$$

Operamos y despejamos la variable x , con lo que se obtiene:

$$0.06x + 70 - 0.07x = 64; \quad 70 - 64 = 0.01x; \quad x = \frac{6}{0.01} = 600$$

Con lo que la cantidad que se invirtió en el primer bono es 600€ y 400€ en el segundo bono.

Ejemplo 1.2. El modelo IS *Investment/Saving* para una economía establece que el gasto total de la economía (consumo, inversión privada, gasto público, balance en la exportación,...) es igual al producto interior bruto.

En los modelos económicos una *variable* es cualquier magnitud que puede cambiar de valor. El modelo suele venir determinado por una ecuación que relacione las variables entre sí. El más simple de los modelos IS en una economía cerrada, es decir, con balance de exportación nulo, viene determinado por las siguientes ecuaciones:

$$Y = C + I + G; \quad C = a + b(Y - T)$$

donde la variable Y es el producto interior bruto, C es el consumo, T son los impuestos, I es la inversión y G es el gasto público. Los valores a y b son constantes estructurales de la economía estudiada. A las variable G y T son consideradas *exógenas* pues son establecidas por el gobierno y al resto de variables se les llama *endógenas*.

Modelos como el anterior han sido utilizados desde 1937 en el contexto de la economía keynesiana donde los gobiernos deben intentar mejorar los datos macroeconómicos por la correcta elección de los impuestos y el gasto público.

1.7. Desigualdades. Intervalos.

Una desigualdad es afirmar que una expresión es menor que otra. El símbolo $<$ significa “menor que”, el símbolo $>$ significa “mayor que”, $\leq$ significa “menor o igual que” y $\geq$ significa “mayor o igual que”. Se verifican las siguientes propiedades:

1. Dados dos números reales x e y , ocurre que, o bien $x \leq y$ o bien $y \leq x$.
2. Sean x, y, z tres números reales cualquiera. Si $x \leq y$ entonces $x + z \leq y + z$.
3. Sean x, y dos números reales cualquiera y $z \geq 0$. Si $x \leq y$ entonces $xz \leq yz$.

Antes de continuar, resaltamos las siguientes propiedades que pueden ser deducidas de la anteriores:

1. Si $a > b$ y $c \geq d$, entonces $a + c > b + d$.
2. Si $a > b \geq 0$ y $c \geq d > 0$, entonces $ac > bd$.
3. Si $k < 0$ y $a > b$, entonces $ka < kb$.
4. Si $0 < a < b$, entonces $a^2 < b^2$, y si $a < b < 0$, entonces $a^2 > b^2$.
5. Si $a < 0$ o $a > 0$, entonces $a^2 > 0$.

6. Si $a > b$, entonces $-a < -b$.
7. Si $a < 0$, $b > 0$, entonces $ab < 0$, mientras que si $a < 0$, $b < 0$, entonces $ab > 0$.
8. Si $a > 0$, entonces $1/a > 0$, mientras que si $a < 0$, entonces $1/a < 0$.
9. Si $a > b > 0$, entonces $1/b > 1/a > 0$, mientras que si $a < b < 0$, entonces $1/b < 1/a < 0$.

Cuando se escribe

$$\{x \in \mathbb{R} : x \text{ cumple la propiedad } P\}$$

significa “el conjunto de todos los números reales que verifican la propiedad P ”. La propiedad P puede ser, por ejemplo, verificar una ecuación o una inecuación.

Un intervalo es el conjunto de todos los números reales entre dos puntos de la recta real. Por tanto, los intervalos son conjuntos “sin huecos”. Dados dos números reales a y b tal que $a < b$, distinguimos cuatro tipos de intervalos:

- I. Intervalo abierto de extremos a y b . A los puntos a y b se les llama extremo superior y extremo inferior del intervalo respectivamente. El intervalo (a, b) son todos los puntos x de la recta real que verifican $a < x < b$, en notación de conjuntos se define como

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

A la hora de representarlo gráficamente sombreamos los puntos entre a y b . Los extremos a y b son señalados con una circunferencia para indicar que esos puntos no pertenecen al conjunto.



- II. Intervalo cerrado de extremos a y b . Son todos los puntos x que verifican $a \leq x \leq b$, en notación de conjunto viene definido por $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$. Para representarlo gráficamente sombreamos todos los puntos entre a y b . Los extremos a y b son señalados con punto negro para indicar que dichos puntos pertenecen al conjunto



- III. Intervalos semiabiertos de extremos a y b . Nos encontramos con dos casos. El primero es que pertenezca el extremo derecho, pero no el izquierdo. Con la notación de conjunto se escribe

$$(a, b] := \{x \in \mathbb{R} : a < x \leq b\}.$$

Como en los casos anteriores sombreamos los puntos entre a y b . El extremo inferior se señala con una circunferencia y el extremo superior es señalado con un círculo.



El caso opuesto es que pertenezca el extremo inferior, pero no el extremo superior del intervalo. Se define como sigue

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

y la representación gráfica es el extremo inferior señalado con un círculo, el superior señalado con una circunferencia y los puntos entre ambos sombreados.



IV. Intervalos semi-infinitos. Son los que no están limitados por uno de los lados. Se emplea el símbolo $+\infty$ para expresar que no está limitado a la derecha y el símbolo $-\infty$ para decir que no lo está por la izquierda. El primer caso viene dado por limitado y abierto por la izquierda e ilimitado por la derecha. Son todos los puntos x que verifican que $a < x$, empleando la notación de conjuntos escribimos:

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

Para representarlo señalamos el punto a con una circunferencia y sombreamos todos los puntos hacia la derecha



El siguiente caso es ilimitado a la derecha y cerrado por la izquierda. Lo que se define como

$$[a, \infty) := \{x \in \mathbb{R} : a \leq x\}$$

La representación gráfica es idéntica a la anterior, pero sustituimos la circunferencia por un círculo



Los dos siguientes casos son los ilimitados por la izquierda. Se definen por $(-\infty, a) := \{x \in \mathbb{R} : x < a\}$ cuya representación gráfica es



El último caso se define por $(-\infty, a] := \{x \in \mathbb{R} : x \leq a\}$ y se representa por



Un caso especial de intervalo es el ilimitado por los dos lados $(-\infty, \infty)$; en este caso se trata de toda la recta real. Por eso, escribir $\mathbb{R}$ es equivalente a $(-\infty, \infty)$.

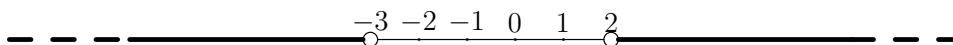
Comentar que los términos “abiertos” y “cerrados” pueden ser entendidos como que los extremos pertenezcan o no al conjunto. Existen definiciones más precisas, pero se salen de los objetivos del curso.

Ejemplo 1.3. Resolver la inecuación $(x - 2)(x + 3) > 0$.

El producto de dos números es positivo si los dos tienen el mismo signo. Hay dos posibilidades:

1. Los dos factores positivos y se verifica que $x - 2 > 0$ y $x + 3 > 0$. Despejando $x > 2$ y $x > -3$. Obsérvese que si se verifica la primera, entonces se verifica la segunda. La única manera de que se verifiquen las dos es que $x > 2$
2. Los dos factores son negativos. Con lo que se verifica $x - 2 < 0$ y $x + 3 < 0$. Ambas condiciones equivalen a $x < -3$.

En resumen, la variable x verifica que $x < -3$ ó $x > 2$. Podemos representar el conjunto solución gráficamente por



Obsérvese que en este caso la solución viene dada por dos intervalos: $(-\infty, -3)$ y $(2, +\infty)$.

1.8. Valor absoluto

Se llama *valor absoluto* a la distancia de un punto a al 0. Más formalmente, dado un número real x , se define su valor absoluto por

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

Algunas propiedades del valor absoluto son:

1. $|x| = |-x|$;
2. $|x \cdot y| = |x| \cdot |y|$;
3. $|x| \leq y$ es equivalente a $-y \leq x \leq y$;
4. $|x + y| \leq |x| + |y|$;
5. $||x| - |y|| \leq |x + y| \leq |x| + |y|$.

Ejemplo 1.4. Encuentra el valor de x para que $-1/2 < x < 4$ y $|2x + 1| = |x - 4|$.

Tenemos que $|2x + 1|^2 = (2x + 1)^2$ y $|x - 4|^2 = (x - 4)^2$, por lo tanto la ecuación para los valores absolutos es equivalente a: $(2x + 1)^2 = (x - 4)^2$ implica que $x^2 + 4x - 5 = 0$. Soluciones de esta ecuación son $x_1 = 1$, $x_2 = -5$, pero sólo x_1 verifica simultáneamente las condiciones del problema.

Capítulo 2

Funciones de una variable.

En la sección de economía de los periódicos leemos habitualmente la frase “en función de” en expresiones como “los tipos de interés suben en función de la inflación” o “el precio de los productos cárnicos van en función del precio de los cereales”. Todas estas frases expresan que las variables están interrelacionadas de forma que para cada precio de los cereales hay un precio de los productos cárnicos. En matemáticas la expresión “en función de” tiene una definición precisa para evitar ambigüedades.

2.1. Concepto de función. Dominio y recorrido

2.1.1. Función, Dominio y Recorrido

Dado un conjunto D de números reales, una *función* $f : D \rightarrow \mathbb{R}$ (o simplemente f) de una variable real con dominio D es una forma de asignar a cada número real x en D un único número $f(x)$.

Es decir, para definir una función se necesita:

- La ley de correspondencia, $x \mapsto f(x)$.
- El conjunto D donde está definida: **dominio**. Cuando se omite éste, entenderemos que es el mayor conjunto $D \subseteq \mathbb{R}$ donde esté definida.

La notación $f : D \rightarrow \mathbb{R}$ señala el nombre de la función (f), el nombre del conjunto del dominio (D) y el conjunto de llegada de la aplicación ($\mathbb{R}$). En el curso de *Matemáticas I* el conjunto de llegada es siempre $\mathbb{R}$, pero en *Matemáticas II* estudiaremos otras posibilidades.

Muchas veces las funciones se escriben como $y = f(x)$. A la variable x se le dice *independiente*, *exógena* o *pre-imagen* de y . A la variable y se le llama *dependiente*, *endógena* o *imagen* de x .

Un ejemplo de función es el costo de producir x kilogramos de un producto para una empresa. El dominio de esta función son los números reales no negativos $[0, +\infty)$, pues no se puede producir una cantidad negativa.

Otro ejemplo de función es la cotización de una empresa en bolsa en el mercado continuo. En este caso el dominio viene dado por un intervalo que representa el lapso de tiempo para el que tenemos datos.

En los ejercicios teóricos muchas funciones son definidas por una fórmula. Por ejemplo, la función $f : [0, 100] \rightarrow \mathbb{R}$ es definida por $f(Q) = Q(100 - Q)$. Con esta afirmación estamos diciendo que el dominio es $[0, 100]$ y que la variable independiente es Q . Desgraciadamente, en los modelos reales dichas fórmulas no están tan claras.

Llamamos *recorrido* de una función al conjunto de imágenes de la función f . Se denota por $\mathcal{R}(f)$ ó $f(D)$ con D el dominio de la función, escrito en notación de conjuntos

$$f(D) = \{f(x) : x \in D\}.$$

2.1.2. Álgebra de funciones

Se pueden obtener funciones a partir de otras mediante operaciones algebraicas suma y producto. Se pueden sumar dos funciones f y g con un mismo dominio D , la imagen de x es

$$(f + g)(x) = f(x) + g(x) \quad \forall x \in D.$$

La expresión “ $\forall x \in D$ ” significa “para todo x ”.

El producto de dos funciones f y g con un mismo dominio D , la imagen de x viene definida por

$$(f \cdot g)(x) = f(x) \cdot g(x) \quad \forall x \in D.$$

En el caso de dos funciones f y g con diferentes dominios, se supone que el dominio es la zona común a ambos dominios (los matemáticos le llamamos *intersección* a la zona común de dos dominios).

También se define la división de dos funciones de dominio D como $(f/g)(x) = f(x)/g(x)$ para todo x con $g(x) \neq 0$. Para la división hay que eliminar del dominio los puntos en los que se anule el denominador.

Decimos que dos funciones f y g son *iguales* cuando:

1. El dominio de f y de g coincide y
2. Para todo x en el dominio $f(x) = g(x)$.

2.2. Gráfica de una función

2.2.1. Pares ordenado

Un par ordenado se representa como (x, y) donde x e y son dos números reales. Al conjunto de todos los pares ordenados se denota $\mathbb{R} \times \mathbb{R}$ ó $\mathbb{R}^2$.

Decimos que (a, b) es igual a (c, d) si $a = c$ y $b = d$.

Se llama *gráfica* de una función $f : D \rightarrow \mathbb{R}$ al siguiente conjunto de pares ordenados

$$\text{Graf}(f) = \{(x, f(x)) : x \in D\}.$$

Lo interesante es que dicho conjunto puede ser “dibujado”, ya que los pares ordenados pueden ser representados como puntos en un plano.

2.2.2. Plano coordenado

En un plano se representan dos rectas perpendiculares que son llamadas *ejes de coordenadas*. El punto de intersección de ambas rectas es llamado *origen de coordenadas*. A la recta horizontal se le llama *eje x* o *eje de abscisa*. A la recta vertical se le llama *eje y* o *eje de ordenadas*. En cada uno de los ejes se establece una unidad de medida.

Dado un par ordenado (a, b) , trazamos una perpendicular al eje de abscisa en su punto a y una perpendicular al eje de ordenadas en el punto b . La intersección de ambas rectas nos da un punto que identificamos con el par ordenado (a, b) . Ver figura 2.1 para una representación gráfica de lo anterior. Los ejes de ordenadas dividen el plano en cuatro sectores. Dichos sectores se llaman *cuadrantes* y se enumeran en sentido contrario a las agujas del reloj.

$$\begin{aligned} \text{Cuadrante I} &:= \{(x, y) : x > 0, y > 0\}; & \text{Cuadrante II} &:= \{(x, y) : x < 0, y > 0\}; \\ \text{Cuadrante III} &:= \{(x, y) : x < 0, y < 0\}; & \text{Cuadrante IV} &:= \{(x, y) : x > 0, y < 0\}. \end{aligned}$$

En muchas magnitudes económicas sólo tienen sentido los valores no negativos. Por ello, en las ciencias económicas es de especial interés el primer cuadrante. En la figura 2.2 se muestran los diferentes cuadrantes en el plano coordenado.

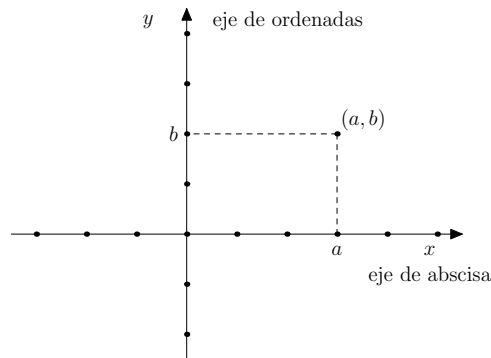


Figura 2.1: Identificación del par ordenado (a, b) con un punto del plano ordenado.

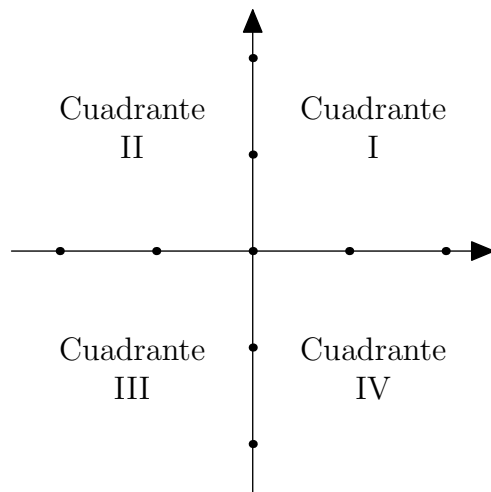


Figura 2.2: Los diferentes cuadrantes del plano de ejes coordenados.

La gráfica de una función es un conjunto de puntos que pueden ser representados en el plano coordenado. Por ejemplo, la gráfica de la función valor absoluto $f(x) = |x|$ aparece en la figura 2.4. Obsérvese que la gráfica de la función es la misma, pero hemos representado diferentes zonas del plano coordenado. Preste atención a que la representación es diferente a las gráficas anteriores. En este caso no hemos representado los ejes ordenados, hemos puesto las escalas en la parte inferior para el eje de abscisa y en la parte derecha para el eje de ordenadas. Se le asigna la correspondencia de cada punto de forma análoga a lo explicado anteriormente. La ventaja de esta representación es que muchas veces el origen de coordenadas está lejos de los valores que toma la función. Si intentásemos representar datos reales como los del siguiente ejemplo, escoger el trozo adecuado del plano coordenado es importante. En el ejemplo 2.1 se puede observar dicha situación.

Ejemplo 2.1. En la tabla 2.1 se encuentran los datos del IPC. Estos datos pueden ser considerados una función ya que a cada fecha se le asigna una cantidad. En inicio la función sólo tiene como dominio los 68 meses para los que tenemos datos, vemos la gráfica en el panel izquierdo de la figura 2.5. Muchas veces aparece dicha gráfica asignando una imagen a todos los puntos del intervalo $[2002, 2007.6]$, para ello unimos dos datos consecutivos por una recta. Esta gráfica puede observarse en el panel derecho de la figura 2.5.

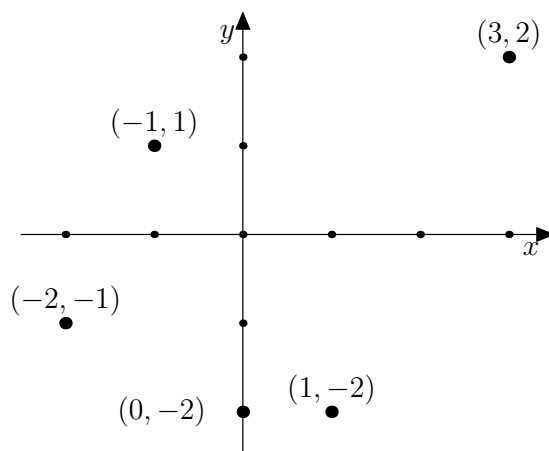


Figura 2.3: Ejes de coordenadas cartesianas y la representación de algunos pares ordenados en el plano.

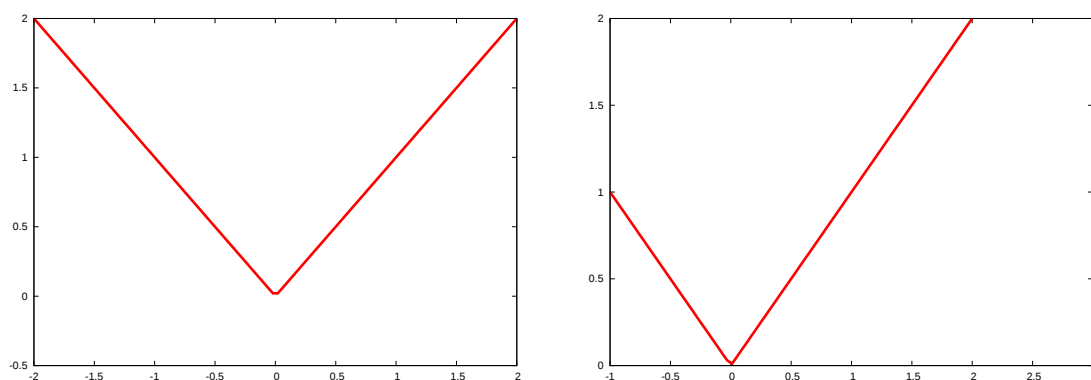


Figura 2.4: Representación de la función $f(x) = |x|$.

2.3. Funciones lineal y afín

2.3.1. Funciones constantes

La función más fácil de definir son las constantes. Una función *constante* toma el mismo valor para todos los puntos donde está definida.

2.3.2. Funciones lineales

En una función lineal el valor de salida es un múltiplo del valor de entrada. Por ejemplo, $f(x) = 3x$ y $f(x) = 7x$. Al número que multiplica a la variable dependiente se le llama *pendiente*.

Los intereses que aportan los bancos por depósitos a un año son una función lineal dependiente de la cantidad depositada. Si nos ofrecen un depósito al 6% de interés anual, entonces los intereses que obtenemos son $f(C) = 0.06C$ donde C es el capital ingresado.

Año	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Sep.	Oct.	Nov.	Dic.
2002	86,089	86,164	86,876	88,056	88,373	88,377	87,763	88,015	88,344	89,219	89,358	89,654
2003	89,285	89,475	90,096	90,829	90,744	90,804	90,236	90,649	90,919	91,533	91,831	91,989
2004	91,343	91,385	92,023	93,283	93,833	93,982	93,269	93,681	93,849	94,813	95,050	94,957
2005	94,157	94,401	95,148	96,518	96,703	96,929	96,336	96,759	97,353	98,145	98,293	98,504
2006	98,105	98,152	98,847	100,241	100,602	100,755	100,155	100,359	100,195	100,606	100,855	101,131
2007	100,451	100,520	101,282	102,681	102,963	103,152	102,402	102,542				

Cuadro 2.1: Datos del IPC(=Índice de Precios al Consumo) según aparece en la página del INE(=Instituto Nacional de Estadística) <http://www.ine.es/daco/daco42/daco421/1au.htm> son los datos del IPC tomando como base el año 2006.

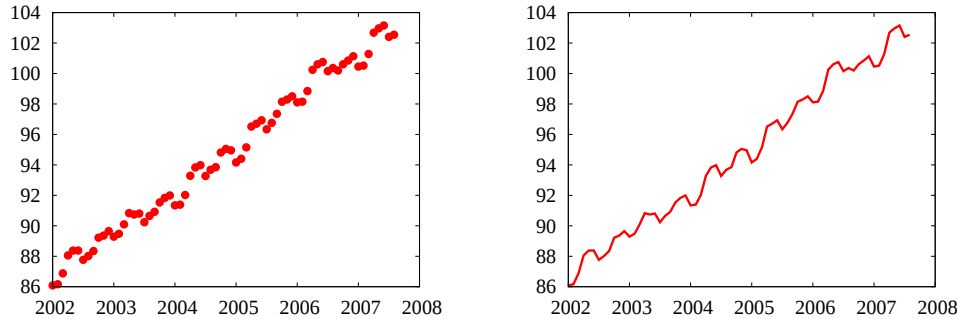


Figura 2.5: Gráficas con la evolución del IPC. En el panel izquierdo están señalados los 68 puntos correspondientes a los datos existentes y en el panel izquierdo esos puntos son unidos por rectas.

2.3.3. Funciones afines

Una función *afín* es la suma de una función lineal y una función constante. En algunos textos a las funciones afines se les llama también funciones lineales. Por tanto, para dar una función afín necesitamos la pendiente m y la ordenada al origen c para la la forma general de la $f(x) = mx + c$.

Una recta viene determinada por dos puntos. Veamos el cálculo de la pendiente dado dos puntos de la recta. Si los $P_1 = (x_1, y_1)$ y $P_2 = (x_2, y_2)$ pertenecen a una recta $y = mx + c$, entonces

$$y_1 = mx_1 + c \quad y_2 = mx_2 + c$$

Despejando c e igualando obtenemos

$$y_1 - mx_1 = y_2 - mx_2, \quad y_2 - y_1 = m(x_2 - x_1)$$

Con lo que la fórmula de la pendiente es el cociente del incremento en las ordenadas entre el incremento en las abscisas

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta y}{\Delta x} \quad (2.1)$$

donde $\Delta x = x_2 - x_1$ es el incremento en la variable x y $\Delta y = y_2 - y_1$ es el incremento en la variable y . La pendiente nos informa de las unidades que aumenta la variable dependiente cuando la independiente varía en una unidad. Obsérvese la figura 2.6 para una interpretación gráfica de dicho concepto.

Dado un punto $P = (x_1, y_1)$ por el que pasa la recta y la pendiente m . Se llama ecuación de la recta *punto-pendiente* a la expresión

$$y = m(x - x_1) + y_1 \quad (2.2)$$

Ejemplo 2.2. Calcule la ecuación de la recta que pasa por los puntos $(1, 1)$ y $(3, 2)$.

Primero calculamos el valor de la pendiente de la recta teniendo en cuenta que $x_1 = 1$, $y_1 = 1$, $x_2 = 3$ e $y_2 = 2$. Por la fórmula (2.1)

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{2 - 1}{3 - 1} = \frac{1}{2}$$

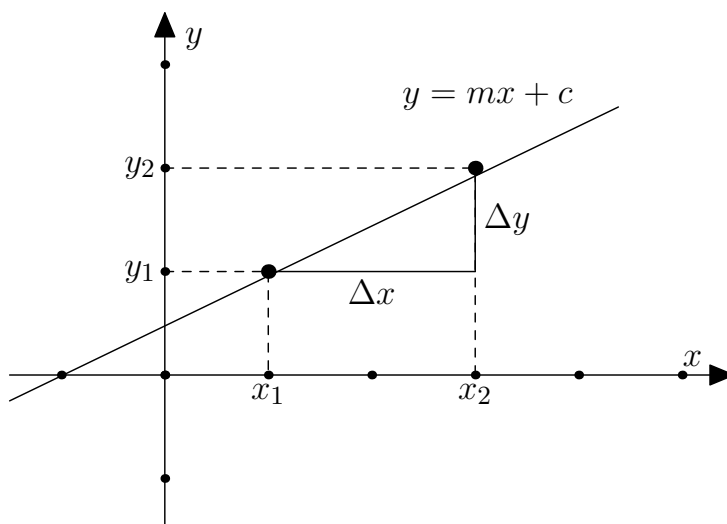


Figura 2.6: Interpretación gráfica del significado de la pendiente de una recta. La pendiente es el cociente Δy entre Δx .

Lo que significa que por cada incremento de unidad en la x incrementamos media unidad en la y . Vea la figura 2.6 donde se representa esta recta. Usando la fórmula (2.2) obtenemos que la ecuación de la recta es

$$y = \frac{1}{2}(x - 1) + 1$$

Ejemplo 2.3. El ministerio de trabajo de EE.UU. ofrece datos sobre el consumo frente a los ingresos. Esos datos para el año 2005 pueden observarse en la tabla 2.2. Esos datos se representan en la figura 2.7 donde se puede observar que los puntos “casi” pasan por la recta $C = 0.512292Y + 16331.7$ donde C es el consumo personal e Y son los ingresos personales.¹ La cantidad 16331.7\$ es el consumo mínimo que tiene que realizar cualquier americano (comida, ropa, ...) para satisfacer las necesidades básicas. Una vez satisfechas estas necesidades gastará cierta cantidad y otra la dedicará a ahorrar. El porcentaje con los que los americanos dedican al consumo está relacionado por el coeficiente 0.512292. Es decir, aproximadamente intentan consumir el 51 % de los ingresos. A dicho coeficiente se le llama *propensión marginal al consumo*.

	20 % inferior	Segundo 20 %	Tercer 20 %	Cuarto 20 %	superior 20 %
Ingresos antes de impuestos	9676\$	25546\$	42622\$	67813\$	147737\$
Consumo	19120\$	28921\$	39098\$	54354\$	90469\$

Cuadro 2.2: Datos ofrecidos por el Ministerio de Trabajo de Estados Unidos en los que la población se divide en cinco grupos según su nivel de ingreso. Para cada grupo de aporta el ingreso medio y los gastos en consumo medios. Estos datos pueden consultarse en la web <http://www.bls.gov/cex/>.

¹ El cálculo de los coeficientes 0.512292 y 16331.7 será estudiado en la asignatura Matemáticas II.

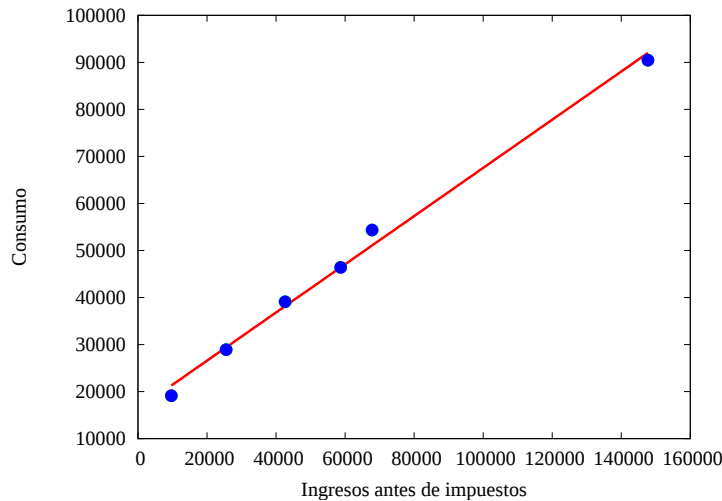


Figura 2.7: Gráfica del consumo frente al ingreso personal para los datos de la tabla 2.2. Se representa la recta $C = 0.512292Y + 16331.7$

2.4. Función potencia con exponente racional

2.4.1. Funciones potencia con exponente entero positivo

Se define como el producto de la variable independiente tantas veces como indique el exponente. Así, la función potencia con exponente n se define:

$$x^n = \underbrace{x \cdots x}_{n \text{ veces}}$$

Observese que si r es la tasa de interés anual y C es el capital inicial, la cantidad $C(1+r)^n$ es el capital acumulado tras n años. Por tanto, las funciones potencia tienen una interpretación económica clara.

2.4.2. Funciones potencia con exponente entero negativo

Si sabemos que el capital que tenemos en el momento actual en un depósito es C y dicho capital ha estado durante n años en dicho depósito con una tasa de interés anual r . ¿Cuál era el valor inicial ingresado? Obviamente, dicho valor es $C/(1+r)^n$. Lo que hemos hecho es recorrer el tiempo en sentido inverso. ¿Por qué no denotar esa cantidad por $C(1+r)^{-n}$? Por eso se define la función potencia con exponente negativo de la siguiente forma:

$$x^{-n} = \frac{1}{x^n} \quad \text{si } x \neq 0$$

El dominio en este caso es toda la recta real salvo el cero. Este conjunto se puede representar con $\mathbb{R} \setminus \{0\}$ o bien $(-\infty, 0) \cup (0, +\infty)$.

2.4.3. Funciones potencia con exponente racional

Por simplicidad consideraremos su dominio los reales no negativos, $\mathbb{R}^+ \cup \{0\}$.² Una entidad financiera nos ofrece una tasa de interés anual de r por un depósito de medio año de una cantidad C , es coherente

² Es posible definir en todo $\mathbb{R}$ aquellas potencias cuyo exponente tiene denominador impar

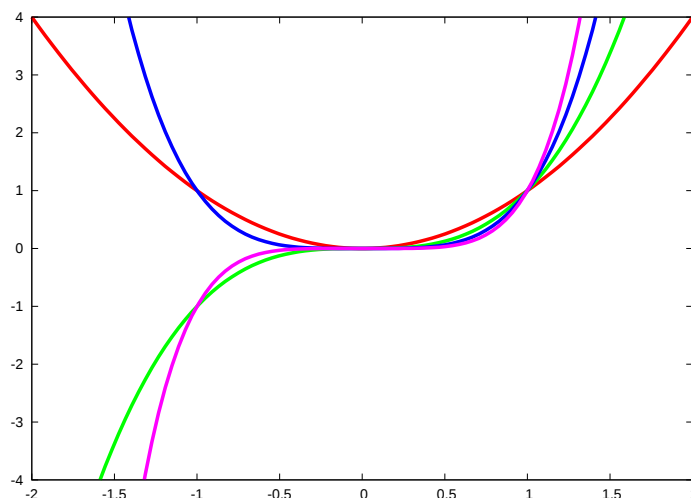


Figura 2.8: Gráficas de las funciones $f(x) = x^2$ en rojo, $f(x) = x^3$ en verde, $f(x) = x^4$ en azul y $f(x) = x^5$ en violeta. Obsérvese la diferente forma que tiene la función según el exponente de la función sea par o impar.

denotar la cantidad obtenida al final $C(1+r)^{1/2}$. Si volviéramos a poner el dinero en el depósito durante otro medio año tendríamos $C(1+r)^{1/2}(1+r)^{1/2}$ y esta cantidad debe ser equivalente a haber puesto todo el dinero durante un año, es decir, $C(1+r)$. Por tanto,

$$C(1+r)^{1/2}(1+r)^{1/2} = C(1+r)$$

Es decir, $(1+r)^{1/2}$ es la raíz cuadrada de $1+r$.

Razonamientos análogos nos llevan a determinar que si con la misma tasa de interés anual hubiésemos tenido un depósito mensual el valor final sería $C(1+r)^{1/12}$ y se deduce que $(1+r)^{1/12}$ es $\sqrt[12]{1+r}$.

Para todo x positivo se define la *potencia* de x elevado a m/n (con $n \neq 0$), $y = x^{m/n}$, como el único real positivo que verifica $y^n = x^m$.

La función $f(x) = x^{m/n}$ se define para los números reales no negativos si m/n es positivo y para los positivos si m/n es negativo. En la figura 2.10 se puede observar la forma típica de las funciones potencias. Para los exponentes positivos menores que uno la gráfica aparece “mirando” hacia la derecha, mientras que para exponentes positivos mayores que uno la forma es como la de exponentes enteros positivos. Algunas propiedades de la función potencia para $a, b \in \mathbb{Q}$ son:

1. $x^a x^b = x^{a+b}$;
2. $(x^a)^b = x^{ab}$;
3. $\frac{x^a}{x^b} = x^{a-b}$;
4. $x^{1/a} = \sqrt[a]{x}$ con $a \neq 0$.

2.5. Funciones pares e impares

La función $f(x)$ es *par* si $f(-x) = f(x)$, en consecuencia, el cambio de signo en la variable x no afecta a la función.

Las funciones potencia con exponente par son funciones pares. Otro ejemplo es la función valor absoluto cuya gráfica está en la figura 2.4. Para la gráfica de más funciones pares vea la figura 2.11

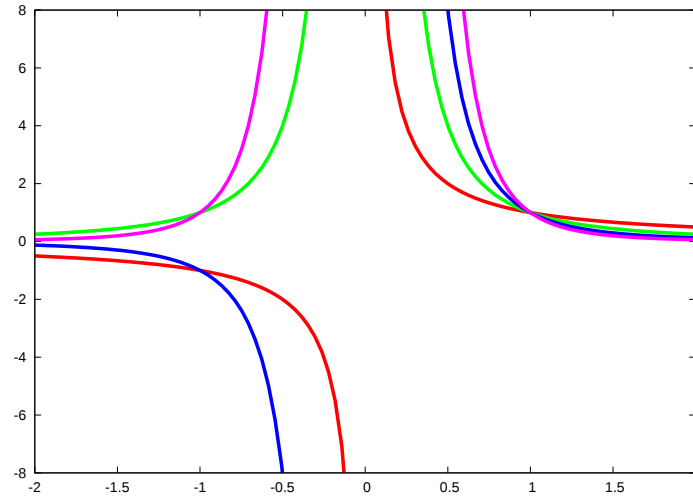


Figura 2.9: Gráficas de las potencias de exponente negativos. La función $f(x) = x^{-1}$ trazada en rojo, $f(x) = x^{-2}$ en verde, $f(x) = x^{-3}$ en azul y $f(x) = x^{-4}$ en violeta.

La función $f(x)$ es una función *impar* si $f(-x) = -f(x)$, en consecuencia, el cambio de signo en la variable x cambia el signo de la función.

Las funciones potencia con exponente impar son funciones impares. Para la gráfica de más funciones pares vea la figura 2.12.

Propiedades:

- El producto de dos funciones pares es par.
- El producto de dos funciones impares es par.
- El producto de una función par por otra impar es una función impar.
- Una función que no tiene paridad definida (no es par ni impar), puede ser descompuesta como la suma de una función par y otra impar.

2.6. Composición de funciones. Función recíproca de otra

2.6.1. Composición de Funciones

Dadas dos funciones $f : D_f \rightarrow \mathbb{R}$ y $g : D_g \rightarrow \mathbb{R}$ tales que el recorrido de g esté contenido en el dominio de f , D_f , se llama composición de f con g , $f \circ g : D_g \rightarrow \mathbb{R}$ a la función definida por $(f \circ g)(x) = f(g(x))$ para todo x en el dominio de g , D_g .

Componer funciones no es conmutativo: $f(g(x)) \neq g(f(x))$ como se observa en el ejemplo 2.4.

Ejemplo 2.4. Dados $f(x) = 3x - 1$ y $g(x) = x^2$, ambas definidas sobre toda la recta real, obtén $f(g(x))$ y $g(f(x))$

$$f(g(x)) = 3(x^2) - 1 = 3x^2 - 1, \quad g(f(x)) = (3x - 1)^2 = 9x^2 - 6x + 1$$

2.6.2. Funciones recíprocas

Una función se dice *inyectiva* si a cada valor en el recorrido $y = f(x)$ le corresponde una única pre-imagen x .

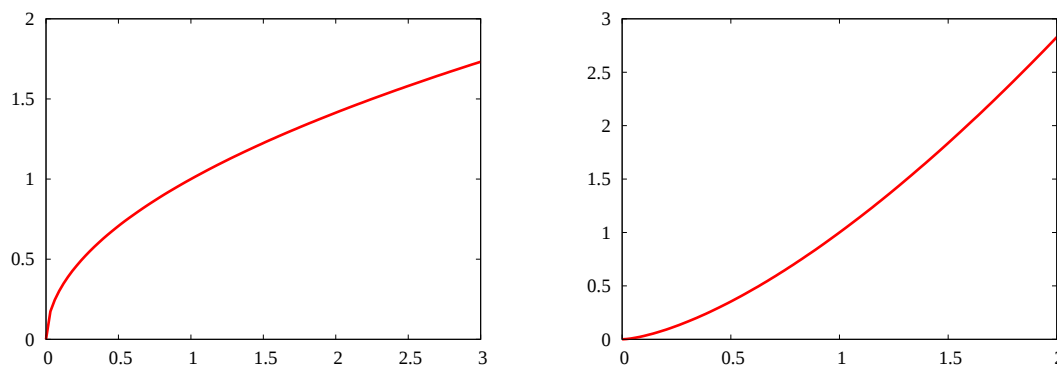


Figura 2.10: Gráfica de la función $f(x) = x^{1/2}$ en el panel de la izquierda y gráfica de la función $f(x) = x^{3/2}$ en el panel de la derecha.

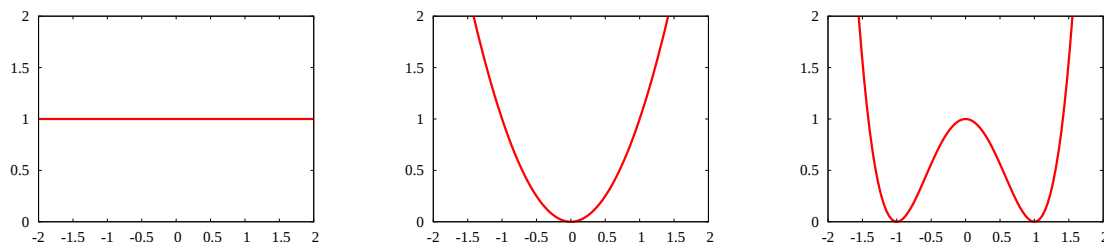


Figura 2.11: Ejemplos de funciones pares. De izquierda a derecha tenemos las funciones $f(x) = 1$, $f(x) = x^2$ y $f(x) = 1 - 2x^2 + x^4$.

La función g se dice que es la *inversa* o *recíproca* de f si $f(g(x)) = x$ para todo x en el dominio de g y $g(f(x)) = x$ para todo x en el dominio de f . La función recíproca de f se denota por f^{-1} .

La regla para obtenerla (si existe), consiste en invertir la relación algebraica existente originalmente: $y = f(x)$ y despejar la variable x para obtener $x = f^{-1}(y)$. Posteriormente (convención para cualquier función), se sustituye variable independiente por dependiente: $y = f^{-1}(x)$. En el ejemplo 2.5 se puede observar esta técnica.

Recuerda que $y = f^{-1}(x)$ significa $x = f(y)$ y que las funciones f y f^{-1} verifican $f(f^{-1}(x)) = x = f^{-1}(f(x))$.

Las gráficas de funciones recíprocas $f(x)$ y $f^{-1}(x)$ son simétricas con respecto a la recta bisectriz $y = x$.

Teorema 2.1. Una función tiene inversa si, y sólo si, es inyectiva.

La función $f : [0, +\infty) \rightarrow \mathbb{R}$ definida por $f(x) = x^{m/n}$ con m y n enteros positivos tiene por función inversa $f^{-1} : [0, +\infty) \rightarrow \mathbb{R}$ dada por $f^{-1}(x) = x^{n/m}$.

Una observación un tanto sutil: La función $f : \mathbb{R} \rightarrow \mathbb{R}$ definida por $f(x) = x^2$ no tiene inversa, ya que no es inyectiva. Sin embargo, la función $f : [0, +\infty) \rightarrow \mathbb{R}$ definida por $f(x) = x^2$ tiene por inversa $f^{-1}(x) = x^{1/2} = \sqrt{x}$. Vea la figura 2.13.

Ejemplo 2.5. Calcular la inversa de $y = f(x) = 2x - 3$ ($y = f(x)$)

Despejando la variable x llegamos a que

$$x = \frac{y + 3}{2} \quad (x = f^{-1}(y))$$

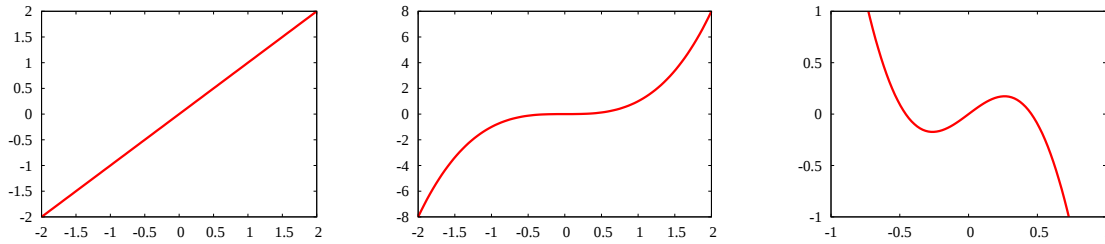


Figura 2.12: Gráfica de funciones impares. De izquierda a derecha tenemos las funciones $f(x) = x$, $f(x) = x^3$ y $f(x) = x^5 - 5x^3 + x$.

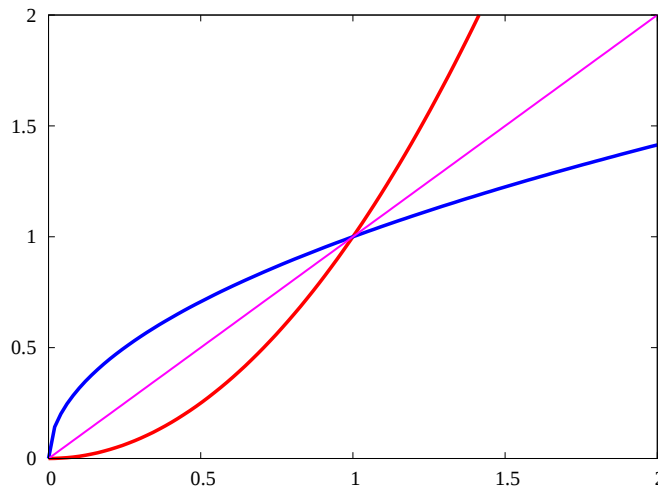


Figura 2.13: Comparación de las gráficas $g(x) = x^2$ y $g^{-1}(x) = \sqrt{x}$. Se representa también la bisectriz del primer cuadrante para que se compruebe la simetría respecto de dicho eje.

Intercambiando variables x e y , tenemos que

$$y = \frac{x+3}{2} \quad (y = f^{-1}(x))$$

2.7. Funciones definidas a trozos en economía

En economía hay magnitudes que no pueden ser negativas, pero muchas veces se dan fórmulas que en inicio pueden tomar valores negativos. Por ejemplo, encontramos que la demanda de un producto es $Q = 80 - P$ donde P es el precio. Queda implícito que si el precio es mayor que 80, entonces la demanda es nula. Dicha demanda Q se define por trozos de la siguiente forma:

$$Q(P) = \begin{cases} 80 - P & \text{si } 0 \leq P \leq 80; \\ 0 & \text{si } P > 80. \end{cases}$$

Ejemplo 2.6. Dados dos consumidores con consumos $Q_1 = 60 - 2P$ y $Q_2 = 100 - 8P$ de un producto. Calculemos la demanda total del producto.

Las dos funciones definidas a trozos son:

$$Q_1(P) = \begin{cases} 60 - 2P & \text{si } 0 \leq P \leq 30; \\ 0 & \text{si } P > 30; \end{cases} \quad \text{y} \quad Q_2(P) = \begin{cases} 100 - 8P & \text{si } 0 \leq P \leq 12.5; \\ 0 & \text{si } P > 12.5. \end{cases}$$

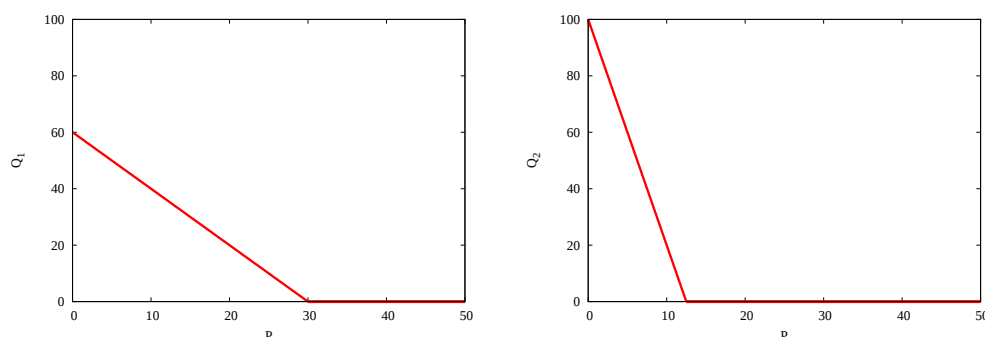


Figura 2.14: Gráficas del consumo de los dos consumidores frente al precio del producto.

Sus gráficas son mostradas en la figura 2.14. La demanda total, Q_T , es la suma de ambas funciones. Hay que sumar la expresiones de cada función separando en cada uno de los intervalos que hay un cambio de definición de las funciones Q_1 y Q_2 . Así, separaremos en tres trozos: $(0, 12.5]$, $(12.5, 30]$ y $(30, +\infty)$. En el primer trozo $(0, 12.5]$, la función Q_1 se define por $60 - 2P$ y Q_2 por $100 - 8P$ y su suma es $160 - 10P$. En el segundo trozo, $(12.5, 30]$, la función Q_1 se sigue definiendo por $60 - 2P$, la función Q_2 es nula y su suma vale $60 - 2P$. En el tercer trozo, $(30, +\infty)$, las funciones Q_1 y Q_2 son nulas, con lo que su suma también. La función consumo total queda definida:

$$Q_T(P) = \begin{cases} 160 - 10P & \text{si } 0 \leq P \leq 12.5; \\ 60 - 2P & \text{si } 12.5 < P \leq 30; \\ 0 & \text{si } P > 30. \end{cases}$$

y su gráfica aparece en la figura 2.15.

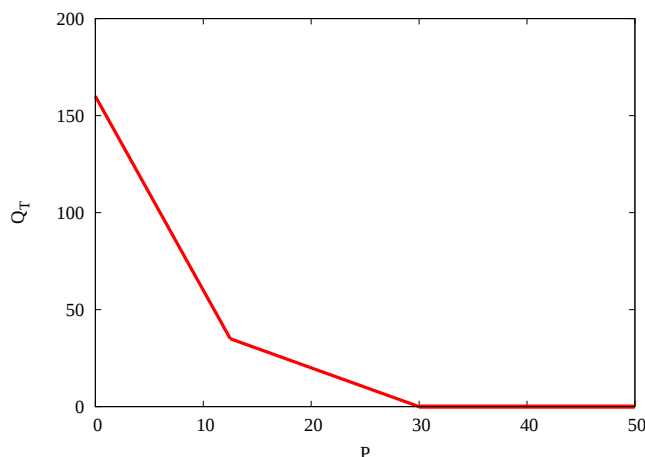


Figura 2.15: Gráfica del consumo total Q_T del producto frente a su precio.

Capítulo 3

Polinomios y funciones racionales.

3.1. Definición de polinomio. Operaciones.

Se llama *polinomio* a una función definida por

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

donde a_j para $j = 0, \dots, n$ son números reales y $a_n \neq 0$. Los números $a_n, a_{n-1}, \dots, a_1$ y a_0 son denominados *coeficientes del polinomio* y a n se le llama *grado* del polinomio. El coeficiente a_n del polinomio se llama *coeficiente principal* y el coeficiente a_0 se llama *término independiente*. Cada uno de los sumandos $a_j x^j$ se llaman *monomios*. El dominio de un polinomio es toda la recta real.

La suma de dos polinomios de grado m y n es otro polinomio de grado menor o igual que $\max\{m, n\}$ donde $\max\{m, n\}$ es el mayor de los dos valores: m y n .

El producto de dos polinomios de grado m y n es otro polinomio de grado $m \cdot n$.

Ejemplo 3.1. Una cadena televisiva puede retransmitir un partido si tiene los derechos de imagen de los dos equipos. Queremos medir la *utilidad* de tener los derechos de x equipos. En este caso el sentido más simple para utilidad es que la satisfacción del abonado a la cadena se mida con el número de partido que puede ver. Cada equipo juega en casa contra otros $x - 1$ equipos cuyos derechos son poseídos por la cadena. Por tanto, la función utilidad viene dada por $U(x) = x(x - 1) = x^2 - x$ y se trata de un polinomio.

Relacionado con este ejemplo está el valor de las redes de telecomunicación y un razonamiento similar al anterior es lo que se conoce como *ley de Metcalfe*. Evidentemente, el tema no es tan sencillo y existe bastante polémica sobre la forma de valorar las redes de telecomunicación.

Ejemplo 3.2. Si en un depósito cuya tasa de interés anual es r ingresamos a primero de los años 2005, 2006 y 2007 la cantidad de 2000€, la cantidad al final del año 2007 es

$$C(r) = 1000(1 + r) + 1000(1 + r)^2 + 1000(1 + r)^3 = 1000r^3 + 4000r^2 + 6000r + 3000.$$

Por tanto, el capital obtenido es un polinomio respecto de la tasa de interés.

3.2. División euclídea de polinomios

Teorema 3.1. Sean P y Q dos polinomios tales que el grado de P es mayor o igual que el de Q . Entonces, existen dos únicos polinomios q y r tales que

$$P(x) = q(x)Q(x) + r(x) \quad \text{para todo } x \in \mathbb{R} \quad (3.1)$$

y el grado de r es menor que el de Q .

La expresión (3.1) se puede escribir de la forma

$$\frac{P(x)}{Q(x)} = q(x) + \frac{r(x)}{Q(x)} \quad \text{si } Q(x) \neq 0.$$

A $q(x)$ se le llama *cociente* de la división $P(x)/Q(x)$ y a $r(x)$ se le llama *resto*.

Si $r(x) = 0$ en (3.1) decimos que $Q(x)$ es un *divisor* de $P(x)$ o que $P(x)$ es divisible por $Q(x)$.

3.2.1. Algoritmo general de división euclídea

Veamos sobre un ejemplo el funcionamiento del algoritmo. Sean $P(x) = 6x^3 + 2x - 3$ y $Q(x) = 2x^2 + x$. Dividimos los dos monomios con mayor potencia $(6x^3)/(2x^2) = 3x$ y escribimos dicha cantidad en el cociente. Multiplicamos el resultado de la división anterior por el divisor $Q(x)$, obteniendo $(3x)Q(x) = 6x^3 + 3x^2$, le restamos a $P(x)$ dicha cantidad y obtenemos $P_1(x) = -3x^2 + 2x - 3$. Obsérvese que tras estas operaciones se puede escribir

$$\frac{P(x)}{Q(x)} = 3x + \frac{P_1(x)}{Q(x)}$$

y el grado del polinomio P_1 es una unidad menor que el grado de P . A continuación se le aplica el mismo algoritmo a $P_1(x)$ y así sucesivamente hasta tener un polinomio de grado menor que $Q(x)$ en el numerador. Vea la figura 3.1 para una descripción completa del algoritmo. En este caso el cociente es $3x - \frac{3}{2}$ y el resto $\frac{7}{2}x - 3$. Esto significa que podemos escribir

$$\frac{P(x)}{Q(x)} = \frac{6x^3 + 2x - 3}{2x^2 + x} = 3x - \frac{3}{2} + \frac{\frac{7}{2}x - 3}{2x^2 + x}$$

Para comprobar que la división es correcta sólo necesitamos operar en el segundo lado de la igualdad para comprobar que efectivamente son iguales.

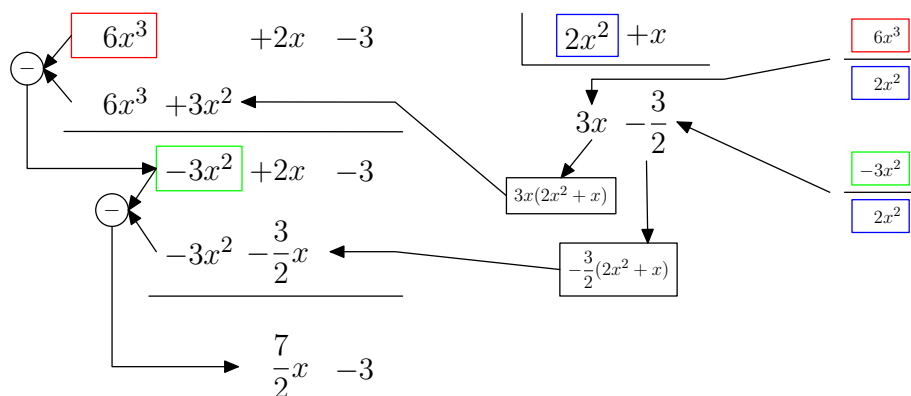


Figura 3.1: Algoritmo de división euclídea para los polinomios $P(x) = 6x^3 + 2x - 3$ y $Q(x) = 2x^2 + x$.

Ejemplo 3.3. Dividimos el polinomio $x^2 + 3x + 1$ entre $x + 1$.

$$\begin{array}{r} x^2 + 3x - 1 \mid x + 1 \\ x^2 + \\ \hline 2x - 1 \\ 2x + 2 \\ \hline -3 \end{array}$$

Por tanto, $\frac{x^2 + 3x - 1}{x + 1} = x + 2 - \frac{3}{x + 1}$

Ejemplo 3.4. Dividir el polinomio $x^4 - 4x^2 + x + 1$ entre $x^3 + x^2 - 4x - 4$.

$$\begin{array}{r} x^4 \quad -4x^2 + x + 1 \quad | \quad x^3 + x^2 - 4x - 4 \\ x^4 + x^3 - 4x^2 - 4x \\ \hline -x^3 \quad +5x + 1 \\ -x^3 - x^2 + 4x + 4 \\ \hline x^2 + x - 3 \end{array}$$

Por tanto, $\frac{x^4 - 4x^2 + x + 1}{x^3 + x^2 - 4x - 4} = x - 1 + \frac{x^2 + x - 3}{x^3 + x^2 - 4x - 4}$.

3.2.2. Algoritmo de Ruffini

Se puede efectuar la división euclídea cuando dividimos un polinomio $P(x)$ entre un binomio $Q(x) = x - x_0$ mediante un procedimiento más rápido que el algoritmo general que acabamos de describir. Sean

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0,$$

el resto es un número real r y el cociente

$$q(x) = b_{n-1} x^{n-1} + b_{n-2} x^{n-2} + \dots + b_1 x + b_0.$$

Realizamos el producto $q(x)(x - x_0)$

$$\begin{aligned} P(x) &= q(x)(x - x_0) + r = (b_{n-1} x^{n-1} + b_{n-2} x^{n-2} + \dots + b_1 x + b_0)(x - x_0) + r = \\ &= b_{n-1} x^n + (b_{n-2} - b_{n-1} x_0) x^{n-1} + (b_{n-3} - b_{n-2} x_0) x^{n-2} + \dots \\ &\dots + (b_2 - b_1 x_0) x^2 + (b_1 - b_0 x_0) x + (r - b_0 x_0) \end{aligned}$$

Igualando los coeficientes de los dos polinomios.

$$\begin{array}{lll} a_n = b_{n-1} & \implies & b_{n-1} = a_n \\ a_{n-1} = b_{n-2} - b_{n-1} x_0 & \implies & b_{n-2} = a_{n-1} + b_{n-1} x_0 \\ a_{n-2} = b_{n-3} - b_{n-2} x_0 & \implies & b_{n-3} = a_{n-2} + b_{n-2} x_0 \\ \vdots & \implies & \vdots \\ a_2 = b_1 - b_2 x_0 & \implies & b_1 = a_2 + b_2 x_0 \\ a_1 = b_0 - b_1 x_0 & \implies & b_0 = a_1 + b_1 x_0 \\ a_0 = r - b_0 x_0 & \implies & r = a_0 + b_0 x_0 \end{array}$$

La relación de fórmulas anteriores permite obtener los coeficientes de cociente y el resto sucesivamente. A la hora de efectuar el algoritmo con lápiz y papel es habitual escribirlo en el orden que se indica en la figura 3.2.

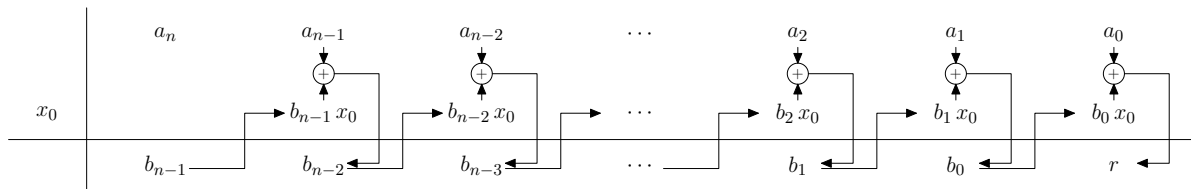


Figura 3.2: Algoritmo de Ruffini para la división por un binomio.

Ejemplo 3.5. Didivir el polinomio $x^3 + 3x^2 + 2x + 1$ y $x - 2$.

$$\begin{array}{r|rrrr} & 1 & 3 & 2 & 1 \\ 2 & & 2 & 10 & 24 \\ \hline & 1 & 5 & 12 & 25 \end{array}$$

Por tanto, podemos escribir

$$\frac{x^3 + 3x^2 + 2x + 1}{x - 2} = x^2 + 5x + 12 + \frac{25}{x - 2}$$

Ejemplo 3.6. Didivir el polinomio $x^3 + 2x^2 + 3x + 4$ y $x - 2$.

$$\begin{array}{r|rrrr} & 1 & 2 & 3 & 4 \\ 2 & & 2 & 8 & 22 \\ \hline & 1 & 4 & 11 & 26 \end{array}$$

lo que significa que

$$\frac{x^3 + 2x^2 + 3x + 4}{x - 2} = x^2 + 4x + 11 + \frac{26}{x - 2}$$

Ejemplo 3.7. Didivir el polinomio $x^2 + 3x + 1$ y $x + 1$. Obsérvese que se trata de repetir el ejemplo 3.3 por otro método

$$\begin{array}{r|rrr} & 1 & 3 & -1 \\ -1 & & -1 & -2 \\ \hline & 1 & 2 & -3 \end{array}$$

y queda la misma expresión que en el ejemplo 3.3

3.3. Raíces de un polinomio. Descomposición de un polinomio.

Se dice que x_0 es una *raíz* de un polinomio P si $P(x_0) = 0$.

Teorema 3.2. Si los coeficientes del polinomio P son enteros, el coeficiente principal de P es uno y P tiene una raíz entera, entonces dicha raíz es un divisor del término independiente.

Teorema 3.3. El polinomio $P(x)$ es divisible por $x - a$ si, y sólo si, a es una raíz del polinomio.

Demostración. Dividiendo el polinomio $P(x)$ entre $x - x_0$ tenemos que $P(x) = q(x)(x - x_0) + r$. Sustituyendo en $x = x_0$ se tiene que $P(x_0) = r$, con lo que x_0 es raíz si, y sólo si, $r = 0$. $\square$

Una raíz x_0 de un polinomio P se dice que su *multiplicidad* es mayor o igual que m si existe un polinomio $Q(x)$ tal que $P(x) = (x - x_0)^m Q(x)$, es decir, si $(x - x_0)^m$ es un divisor de $P(x)$.

Teorema 3.4. Un polinomio de grado n tiene a lo más n raíces contando cada raíz como indique su multiplicidad.

Teorema 3.5. Si un polinomio P de grado n tiene n raíces reales (contando cada raíz tantas veces como indique la multiplicidad), entonces se puede factorizar como producto de polinomios de primer grado, es decir,

$$P(x) = a_n(x - x_1)(x - x_2) \cdots (x - x_n)$$

donde a_n es el coeficiente principal del polinomio y $x_1, x_2, \dots, x_n$ son las raíces reales del polinomio repetidas tantas veces como indica su multiplicidad.

Un resultado fundamental de las matemáticas es el conocido como *teorema fundamental del álgebra* en el que se afirma que todo polinomio se puede escribir como producto de polinomios de primer y segundo grado.

Ejemplo 3.8. Factorizar el polinomio $P(x) = x^3 + x^2 - 4x - 4$.

Buscamos la raíces entre los divisores del término independiente que son ± 1 , ± 2 , y ± 4 . Tenemos que $P(1) = 1^3 + 1^2 - 4 \cdot 1 - 4 = -6$, con lo que $x = 1$ no es raíz. Probamos con $x = -1$, $P(-1) = (-1)^3 + (-1)^2 - 4(-1) - 4 = -1 + 1 + 4 - 4 = 0$; con que $x = -1$ es una raíz de P . Dividimos por $x - 1$ para extraer dicho factor. Apliquemos el algoritmo de Ruffini

$$\begin{array}{r|rrrr} & 1 & 1 & -4 & -4 \\ -1 & & -1 & 0 & 4 \\ \hline & 1 & 0 & -4 & 0 \end{array}$$

Podemos escribir $P(x) = (x - (-1))(x^2 - 4)$. Para factorizar $x^2 - 4$ resolvemos la ecuación $x^2 - 4 = 0$ que tiene por raíces $x = \pm 2$. Por tanto, la factorización es

$$P(x) = (x + 1)(x - 2)(x + 2).$$

Ejemplo 3.9. Factoriza el polinomio $P(x) = x^3 + 3x^2 - 4$.

Intentamos con los divisores del término independiente: ± 1 , ± 2 y ± 4 . Probamos si $x = 1$ es raíz, $P(1) = 1^3 + 3 \cdot 1^2 - 4 = 0$, con lo que $x = 1$ es raíz. Dividimos el polinomio P por $x - 1$ usando el algoritmo de Ruffini:

$$\begin{array}{r|rrrr} & 1 & 3 & 0 & -4 \\ 1 & & 1 & 4 & 4 \\ \hline & 1 & 4 & 4 & 0 \end{array}$$

Tenemos la factorización: $P(x) = (x - 1)(x^2 + 4x + 4)$. Resolvemos la ecuación de segundo grado $x^2 + 4x + 4 = 0$ y obtenemos que $x = -2$ es una raíz doble. Por consiguiente la factorización es:

$$P(x) = (x - 1)(x - (-2))^2 = (x - 1)(x + 2)^2.$$

3.4. Funciones racionales

Una función que es el cociente de polinomios se llama función *racional*. Dados dos polinomios $P(x)$ y $Q(x)$, la función definida por $f(x) = P(x)/Q(x)$ es una función racional. El dominio de las funciones racionales son todos los números reales salvo las raíces del denominador.

Si el grado del polinomio numerador es menor que el grado del polinomio denominador se dice que es una función racional *propia*. En caso contrario, se dice que es una función racional *impropia*. El teorema 3.1 (del resto) afirma que toda función racional impropia se puede escribir como la suma de un polinomio y una función racional propia.

La manipulación de las funciones racionales es básica para comprender los modelos más simples de dinámica de poblaciones.

3.5. Descomposición en fracciones simples

Sea $P(x)/Q(x)$ una función racional propia y cuyo denominador Q sea posible factorizar como producto de polinomios de primer grado $Q(x) = (x - a_1)(x - a_2) \cdots (x - a_n)$ donde a_i son las raíces del denominador.

3.5.1. Descomposición con todas las raíces del denominador reales y distintas

Hacemos siguiente identificación

$$\frac{P(x)}{Q(x)} = \frac{A_1}{x - a_1} + \frac{A_2}{x - a_2} + \cdots + \frac{A_n}{x - a_n}$$

y determinamos los coeficientes indeterminados A_i efectuando en primer lugar la suma de las fracciones simples:

$$\frac{P(x)}{Q(x)} = \frac{A_1(x - a_2) \cdots (x - a_n) + A_2(x - a_1) \cdots (x - a_n) + \cdots + A_n(x - a_1) \cdots (x - a_{n-1})}{Q(x)}$$

Ejemplo 3.10. Descomponer en fracciones simples la función racional $\frac{x^2 + x - 3}{x^3 + x^2 - 4x - 4}$.
En el ejemplo 3.8 vimos que la factorización del denominador era $x^3 + x^2 - 4x - 4 = (x - 2)(x + 2)(x + 1)$, y, por lo tanto,

$$\frac{x^2 + x - 3}{x^3 + x^2 - 4x - 4} = \frac{A_1}{x - 2} + \frac{A_2}{x + 2} + \frac{A_3}{x + 1} = \frac{A_1(x + 2)(x + 1) + A_2(x - 2)(x + 1) + A_3(x - 2)(x + 2)}{(x - 2)(x + 2)(x + 1)}.$$

Identificando numeradores se tiene que

$$x^2 + x - 3 = A_1(x + 2)(x + 1) + A_2(x - 2)(x + 1) + A_3(x - 2)(x + 2)$$

Para $x = 2$, será $3 = 12A_1$, para $x = -2$, $-1 = 4A_2$, y para $x = -1$, $-3 = -3A_3$. Del sistema

$$\begin{cases} 3 = 12A_1, \\ -1 = 4A_2, \\ -3 = -3A_3, \end{cases}$$

obtenemos $A_1 = \frac{1}{4}$, $A_2 = -\frac{1}{4}$, $A_3 = 1$. De ese modo, la integral descompuesta en fracciones simples queda

$$\frac{x^2 + x - 3}{x^3 + x^2 - 4x - 4} = \frac{1}{4} \frac{1}{x - 2} - \frac{1}{4} \frac{1}{x + 2} + \frac{1}{x + 1}.$$

3.5.2. Raíces reales múltiples

Si el denominador tiene también raíces múltiples del tipo $(x - b)^k$, por cada una de ellas añadimos a la suma de fracciones simples del caso anterior las siguientes:

$$\frac{B_1}{(x - b)} + \frac{B_2}{(x - b)^2} + \cdots + \frac{B_k}{(x - b)^k}.$$

Ejemplo 3.11. Calcular $\frac{x^2 + x + 3}{x^3 + 3x^2 - 4}$.

En el ejemplo 3.9 vimos que la factorización del polinomio es $x^3 + 3x^2 - 4 = (x + 2)^2(x - 1)$.

La descomposición en fracciones simples con coeficientes indeterminados resulta

$$\frac{x^2 + x + 3}{x^3 + 3x^2 - 4} = \frac{A}{x - 1} + \frac{B_1}{(x + 2)} + \frac{B_2}{(x + 2)^2} = \frac{A(x + 2)^2 + B_1(x - 1)(x + 2) + B_2(x - 1)}{(x - 1)(x + 2)^2}$$

Identificando numeradores: $x^2 + x + 3 = A(x + 2)^2 + B_1(x - 1)(x + 2) + B_2(x - 1)$. Para $x = 1$, $5 = 9A$ para $x = -2$, $5 = -3B_2$, y (por ejemplo) para $x = 0$, $3 = 4A - 2B_1 - B_2$. El sistema resultante es

$$\begin{cases} 5 = 9A, \\ 5 = -3B_2, \\ 3 = 4A - 2B_1 - B_2, \end{cases}$$

de donde $A = \frac{5}{9}$, $B_1 = \frac{4}{9}$, y $B_2 = -\frac{5}{3}$. Finalmente,

$$\frac{x^2 + x + 3}{x^3 + 3x^2 - 4} = \frac{5}{9} \frac{1}{x - 1} + \frac{4}{9} \frac{1}{x + 2} - \frac{5}{3} \frac{1}{(x + 2)^2}.$$

Capítulo 4

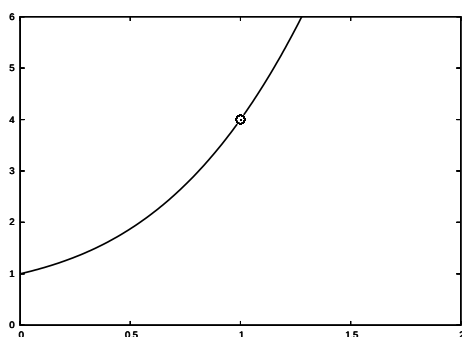
Límites. Continuidad. Sucesiones. Series

4.1. Concepto intuitivo de límite de una función en un punto

La definición rigurosa de límite que aparece en los libros de Análisis Matemático es

Se dice que el *límite* de la función f en $x = a$ es L si para todo $\epsilon > 0$ existe un $\delta > 0$ (dependiente de a y de ϵ) tal que $|f(x) - L| < \epsilon$ para todo $x \neq a$ y $|x - a| < \delta$.

Comprender y aplicar la definición es complicado. Afortunadamente, salvo que se quiera comprender la demostración de los teoremas, la idea intuitiva de límite es suficiente para comprender el concepto y las técnicas de cálculo de límites. Veamos la idea intuitiva con un ejemplo simple: la función $f(x) = \frac{x^4 - 1}{x - 1}$ no está definida en el uno. Sin embargo, viendo la gráfica y la tabla de valores en la figura 4.1 resulta obvio que al acercarse la variable x hacia valores cercanos a uno, las imágenes se aproximan a cuatro. En este caso decimos el límite de $f(x)$ en $x = 1$ vale 4 y lo denotamos por $\lim_{x \rightarrow 1} f(x) = 4$.



A la izquierda de $x = 1$		A la derecha de $x = 1$	
x	f(x)	x	f(x)
0.9	3.439	1.1	4.641
0.99	3.940399	1.01	4.060401
0.999	3.994003999	1.001	4.006004001
0.9999	3.99940004	1.0001	4.000600040001

Figura 4.1: Ilustración del concepto intuitivo de límite para la función $f(x) = \frac{x^4 - 1}{x - 1}$ en $x = 1$.

Es importante resaltar que el valor del límite no depende del valor de la imagen en dicho punto, sino de los valores que toma la función alrededor del punto. Sin embargo, como se ve en el próximo apartado, en muchos casos se obtienen los límites calculando el valor de la función en dicho punto.

4.1.1. Límites y sus propiedades

Los siguientes resultados se obtienen usando la definición de límite.

Teorema 4.1. Si los límites $\lim_{x \rightarrow a} f(x) = L$ y $\lim_{x \rightarrow a} g(x) = K$ son dos números reales, entonces

1. $\lim_{x \rightarrow a} k f(x) = k L$;
2. $\lim_{x \rightarrow a} f(x) + g(x) = L + K$;
3. $\lim_{x \rightarrow a} f(x)g(x) = L K$;
4. Si $K \neq 0$, entonces $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{L}{K}$.

Teorema 4.2. (Límite de función compuesta) Si f y g son funciones tales que $\lim_{x \rightarrow a} g(x) = L$ y $\lim_{x \rightarrow L} f(x) = K$ entonces $\lim_{x \rightarrow a} f(g(x)) = K$.

Teorema 4.3. Si P es un polinomio, entonces el límite del polinomio cuando en el punto a es $P(a)$, es decir,

$$\lim_{x \rightarrow a} P(x) = P(a)$$

Teorema 4.4. Sea $r > 0$, entonces $\lim_{x \rightarrow a} x^r = a^r$.

Una consecuencia importante de los resultados anteriores es que el límite de polinomios, funciones racionales y funciones radicales se obtienen por sustitución, allí donde se puedan evaluar.

Ejemplo 4.1. Calcular $\lim_{x \rightarrow 3} \left(\frac{x-2}{x+1} \right)^{1/2}$.

Por lo visto en los resultados anteriores, si en una función racional, o incluso con radicales, podemos sustituir el valor donde se toma límite “sin problemas” (el denominador no se anula), entonces dicho valor es el límite.

$$\lim_{x \rightarrow 3} \left(\frac{x-2}{x+1} \right)^{1/2} = \left(\frac{3-2}{3+1} \right)^{1/2} = \left(\frac{1}{4} \right)^{1/2} = \frac{1}{2}.$$

Si el límite del numerador es diferente de cero y el denominador es cero, entonces el límite no existe. Cuando el límite del numerador y del denominador es cero, tenemos una indeterminación del tipo 0/0 para calcular el límite se aplican técnicas especiales.

La primera técnica para resolver indeterminaciones que se muestra en los ejemplos y se aplica a las funciones racionales y consiste en cancelar factores comunes en numerador y denominador

Ejemplo 4.2. Calculemos el límite de $f(x) = \frac{x^2 - 1}{x + 1}$ cuando la variable x tiende a -1 .

Observemos que para $x = -1$ no está definido el valor. Sin embargo, para $x \neq -1$ simplificamos

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

Con lo que la gráfica de dicha función es $y = x - 1$ salvo el punto $(-1, -2)$ aparece en el panel izquierdo de la figura 4.2. Intuitivamente cuando nos acercamos a $x = -1$ el valor de y será -2 . Así escribimos

$$\lim_{x \rightarrow -1} f(x) = -2.$$

Ejemplo 4.3. Calcular el límite de $f(x) = \frac{x^4 + x^3 + x^2 + x - 4}{x - 1}$ cuando x tiende a uno.

Tenemos una indeterminación 0/0. Dividimos el polinomio $x^4 + x^3 + x^2 + x - 4$ entre $x - 1$ utilizando el método de Ruffini

$$\begin{array}{r|rrrrr} & 1 & 1 & 1 & 1 & -4 \\ 1 & & 1 & 2 & 3 & 4 \\ \hline & 1 & 2 & 3 & 4 & 0 \end{array}$$

Lo que significa que $x^4 + x^3 + x^2 + x - 4 = (x-1)(x^3 + 2x^2 + 3x + 4)$. Podemos resolver la indeterminación.

$$\lim_{x \rightarrow 1} \frac{x^4 + x^3 + x^2 + x - 4}{x - 1} = \lim_{x \rightarrow 1} \frac{(x-1)(x^3 + 2x^2 + 3x + 4)}{x - 1} = \lim_{x \rightarrow 1} x^3 + 2x^2 + 3x + 4 = 1^3 + 2 \cdot 1^2 + 3 \cdot 1 + 4 = 10.$$

Otra técnica para resolver indeterminaciones es intentar racionalizar el numerador o el denominador de la función. Si en el numerador o denominador tenemos una expresión de la forma $\sqrt{a} + b$, multiplicamos numerador y denominador por $\sqrt{a} - b$. Veamos un par de ejemplos de esta técnica.

Ejemplo 4.4. Calcule $\lim_{x \rightarrow 1} \frac{x-1}{\sqrt{x+3}-2}$.

Al sustituir el valor $x = 1$ en el numerador y en el denominador, comprobamos que se trata de una indeterminación del tipo $0/0$. Para resolverla multiplicamos numerador y denominador por $\sqrt{x+3} + 2$

$$\lim_{x \rightarrow 1} \frac{x-1}{\sqrt{x+3}-2} = \lim_{x \rightarrow 1} \frac{x-1}{\sqrt{x+3}-2} \frac{\sqrt{x+3}+2}{\sqrt{x+3}+2} = \lim_{x \rightarrow 1} \frac{(x-1)(\sqrt{x+3}+2)}{(\sqrt{x+3})^2 - 2^2} = \lim_{x \rightarrow 1} \frac{(x-1)(\sqrt{x+3}+2)}{x-1}$$

Cancelamos factores comunes en numerador y denominador

$$\lim_{x \rightarrow 1} \frac{(x-1)(\sqrt{x+3}+2)}{x-1} = \lim_{x \rightarrow 1} \sqrt{x+3} + 2 = 4.$$

Ejemplo 4.5. Calcular el límite de $f(x) = \frac{x-a^2}{\sqrt{x}-a}$ cuando x tiende a a^2 .

Al sustituir tenemos una indeterminación cero partido cero. Dividimos y multiplicamos por $\sqrt{x} + a$.

$$\lim_{x \rightarrow a^2} \frac{x-a^2}{\sqrt{x}-a} = \lim_{x \rightarrow a^2} \frac{x-a^2}{\sqrt{x}-a} \frac{\sqrt{x}+a}{\sqrt{x}+a} = \lim_{x \rightarrow a^2} \frac{(x-a^2)(\sqrt{x}+a)}{x-a^2} = \lim_{x \rightarrow a^2} \sqrt{x} + a = \sqrt{a^2} + a = |a| + a.$$

4.1.2. Límites laterales

Decimos que el límite por la izquierda de la función $f(x)$ cuando tiende hacia a es L si cuando nos aproximamos a a con valores de x menores que a , entonces las imágenes de a se acercan a L . Lo denotamos por $\lim_{x \rightarrow a^-} f(x) = L$.

La definición del límite por la derecha es análoga y se denota por $\lim_{x \rightarrow a^+} f(x) = L$.

Teorema 4.5. Sea f es una función definida en un intervalo abierto que contiene a un punto c . El límite de $f(x)$ cuando x tiende a c vale L si, y sólo si, los dos límites laterales coinciden, es decir,

$$\lim_{x \rightarrow c^+} f(x) = \lim_{x \rightarrow c^-} f(x) = L.$$

Ejemplo 4.6. La función $f(x) = \frac{x}{|x|}$ es una función definida por trozos que se puede expresar también como

$$f(x) = \begin{cases} -1 & \text{si } x < 0; \\ 1 & \text{si } x > 0. \end{cases}$$

Se observa en el panel derecho de la figura 4.2 que cuando nos acercamos al cero por la izquierda la función vale -1 , así el límite por la izquierda diremos que vale -1 y lo denotaremos por $(x \rightarrow 0^-)$

$$\lim_{x \rightarrow 0^-} \frac{x}{|x|} = -1$$

Análogamente el límite a la derecha, lo denotaremos por $x \rightarrow 0^+$

$$\lim_{x \rightarrow 0^+} \frac{x}{|x|} = 1.$$

Ejemplo 4.7. La función $f(x) = 1 + \sqrt{5-x}$ tiene un único límite lateral cuando x tiende a cinco, ya que la función no existe para los valores de $x > 5$. Se escribe $\lim_{x \rightarrow 5^-} 1 + \sqrt{5-x} = 1$

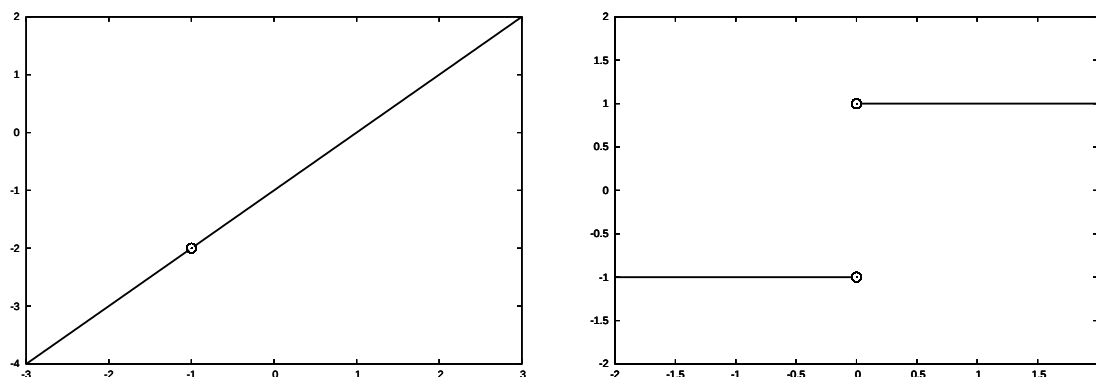


Figura 4.2: En el panel de la izquierda gráfica del ejemplo 4.2 y en el panel de la derecha gráfica de la función signo en ejemplo 4.6.

4.1.3. Límites infinitos

Diremos que el límite de una función $f(x)$ cuando x tiende hacia a es infinito positivo $(+\infty)$, si cuando nos acercamos hacia el valor a , entonces la función toma cada vez valores mayores de forma no acotada. De forma rigurosa se puede expresar:

Se dice que el límite de la función cuando x tiende hacia a es $+\infty$, si para todo $M > 0$ existe un $\delta > 0$ (dependiente de a y de M) tal que $f(x) \geq M$ para todo x tal que $|x - a| < \delta$ y $x \neq a$.

Es posible una definición análoga para límites con valor $-\infty$. Cuando se escribe $\lim_{x \rightarrow a} f(x) = \infty$ sin especificar el signo queremos decir que cuando nos acercamos hacia a el valor absoluto de la función toma cada vez valores mayores. Es decir, escribir $\lim_{x \rightarrow a} f(x) = \infty$ es lo mismo que $\lim_{x \rightarrow a} |f(x)| = +\infty$.

Es importante comentar que en estos casos cuando decimos que un límite vale infinito lo que queremos expresar es la causa de la no existencia del límite. Afortunadamente, no es necesario aplicar las definiciones para calcular los límites y basta con aplicar el teorema 4.6 y 4.7.

Teorema 4.6. Si el límite $\lim_{x \rightarrow a} f(x) \neq 0$ y $\lim_{x \rightarrow a} g(x) = 0$, entonces $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \infty$

La determinación de $+\infty$ ó $-\infty$ debe hacerse con el estudio del signo del cociente $f(x)/g(x)$ alrededor del punto.

Si algún límite lateral de $f(x)$ tiende a infinito cuando x tiende a a decimos que $x = a$ es una *asíntota vertical* de la gráfica de f .

Ejemplo 4.8. Calcular límite de $f(x) = \frac{4}{(x-2)^3}$ cuando x tiende a 2.

Como el numerador tiende a cuatro y el denominador a cero sabemos que el límite es infinito por teorema 4.6. Para ver el signo nos fijamos en los laterales del punto. Para x a la izquierda del 2, tenemos que $x - 2$ es negativo y $1/(x - 2)^3$ también. Por tanto, el límite por la izquierda es

$$\lim_{x \rightarrow 2^-} \frac{4}{(x-2)^3} = -\infty.$$

Para x a la derecha del dos, tenemos que $x - 2$ es positivo con lo que $1/(x - 2)^3$ es positivo. Así,

$$\lim_{x \rightarrow 2^+} \frac{4}{(x-2)^3} = +\infty.$$

Por tanto, la función f tiene asíntota vertical en $x = 2$. Vea el panel izquierdo de la figura 4.3.

Ejemplo 4.9. Calcular las asíntotas verticales de $f(x) = \frac{2+x}{x^2(1-x)}$ y aclarar su comportamiento por los laterales.

Las asíntotas de un función racional se encuentran en los puntos donde se anula el numerador y no el denominador. Por tanto, las asíntotas verticales son las rectas $x = 0$ y $x = 1$.

Cerca de $x = 0$ los tres factores (x^2 , $2+x$ y $1-x$) que aparecen son positivos, con lo que la función es positiva alrededor de $x = 0$. Podemos afirmar que $\lim_{x \rightarrow 0} f(x) = +\infty$.

A la izquierda de $x = 1$ tenemos que los tres factores son positivos, con lo que $f(x)$ es positiva a la izquierda del uno y $\lim_{x \rightarrow 1^-} f(x) = +\infty$. Sin embargo, a la derecha del uno los factores x^2 y $2+x$ son positivos, mientras que $1-x$ es negativo. Por tanto, $f(x)$ es negativo a la derecha del uno y $\lim_{x \rightarrow 1^+} f(x) = -\infty$. Vea el panel izquierdo de la figura 4.3.

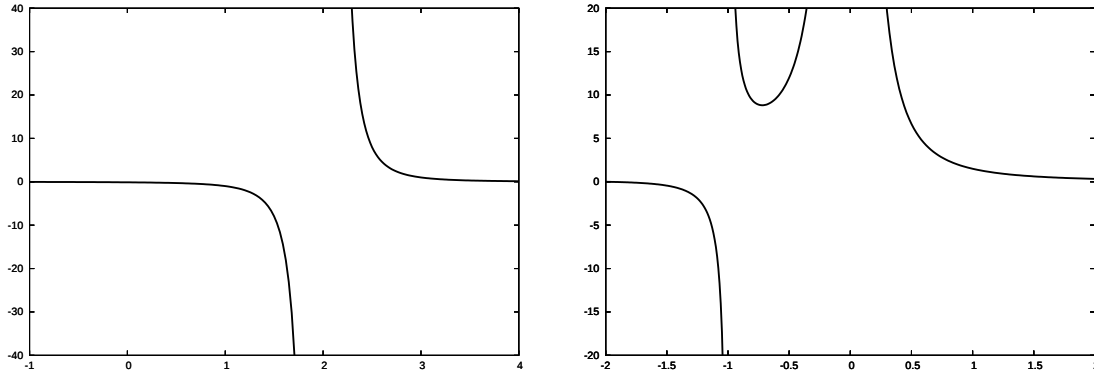


Figura 4.3: Gráficas de las funciones $f(x) = \frac{4}{(x-2)^3}$ y $f(x) = \frac{2+x}{x^2(1-x)}$

Teorema 4.7 (Álgebra del límite). Si $\lim_{x \rightarrow a} f(x) = +\infty$ y $\lim_{x \rightarrow a} g(x) = K$, entonces

1. $\lim_{x \rightarrow a} f(x) + g(x) = +\infty$
2. Si $K > 0$, entonces $\lim_{x \rightarrow a} f(x)g(x) = +\infty$.
3. Si $K < 0$, entonces $\lim_{x \rightarrow a} f(x)g(x) = -\infty$.
4. Si $K = 0$, entonces no sabemos el resultado del límite (indeterminación).

Si $\lim_{x \rightarrow a} f(x) = +\infty$ y $\lim_{x \rightarrow a} g(x) = +\infty$, entonces

1. $\lim_{x \rightarrow a} f(x) + g(x) = +\infty$
2. $\lim_{x \rightarrow a} f(x) - g(x)$ es una indeterminación.
3. $\lim_{x \rightarrow a} f(x)g(x) = +\infty$.

4.2. Concepto intuitivo de límite de una función en el infinito

Se dice que el límite de la función $f(x)$ en el infinito positivo $(+\infty)$ es L , si cuando tomamos cada vez mayores valores de la variable independiente x , las imágenes de la función se acercan a L . La formulación rigurosa de este concepto es:

El límite de la función $f(x)$ cuando x tiende a $+\infty$ es L , si para todo $\epsilon > 0$ existe un $M > 0$ (dependiente de ϵ) tal que $|f(x) - L| < \epsilon$ para todo $x > M$.

Se tiene una definición análoga para límites en el infinito negativo.

Para el cálculo de los límites en el infinito sólo necesitamos el teorema 4.8 y conocer que se verifica el mismo álgebra de límite que en el límite puntual (teoremas 4.1 y 4.7). Basta sustituir en estos $a = +\infty$ ó $a = -\infty$.

Se dice que f tiene una *asíntota horizontal* la recta $y = L$ si $\lim_{x \rightarrow +\infty} f(x) = L$ ó $\lim_{x \rightarrow -\infty} f(x) = L$.

Teorema 4.8. Si r es un número racional positivo, entonces $\lim_{x \rightarrow +\infty} 1/x^r = 0$.

Ejemplo 4.10. ¿A qué tiende $f(x) = \frac{2x-1}{x+1}$ cuando x tiende hacia infinito?

Si dividimos numerador y denominador por x (no podemos dividirlo si $x = 0$, pero este valor está lejos del límite), obtenemos

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

El numerador tiende a 2 y el denominador a 1. Así, la función tiende a 2

$$\lim_{x \rightarrow \infty} \frac{2 - \frac{1}{x}}{1 + \frac{1}{x}} = 2$$

Dada una función racional cuyo grado del denominador es una unidad mayor que el grado del denominador, al cociente de la división euclídea se le llama *asíntota oblicua*. La asíntota oblicua determina el comportamiento de la función en el infinito y se puede observar como la función se “parece” cada vez más a dicha recta cuando nos acercamos al infinito.

Ejemplo 4.11. Calcular la asíntota oblicua de $f(x) = \frac{x^4 - 4x^2 + x + 1}{x^3 + x^2 - 4x - 4}$.

En el ejemplo 3.4 vimos que el cociente era $x - 1$. En la figura 4.4 puede observarse como la función se parece cada vez más a la recta $y = x - 1$ que es la asíntota oblicua.

4.3. Continuidad

4.3.1. Continuidad

Una función $f(x)$ definida en un intervalo abierto que contiene a un punto c se dice *continua* en $x = c$ cuando

1. Existe el límite $\lim_{x \rightarrow c} f(x)$
2. El límite $\lim_{x \rightarrow c} f(x)$ vale $f(c)$.

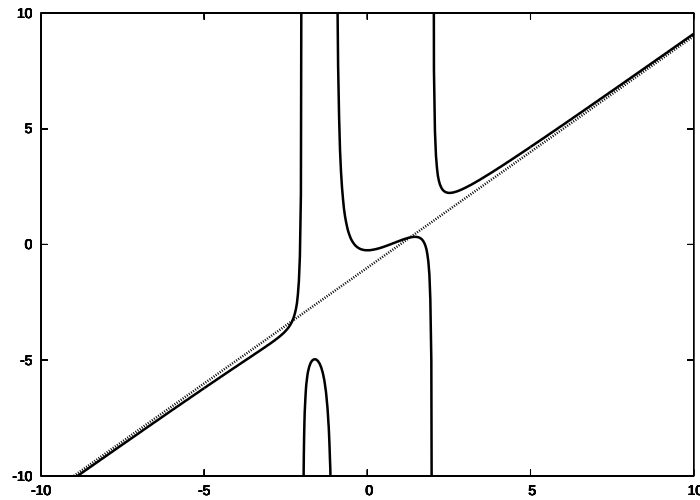


Figura 4.4:

Una función $f(x)$ definida en el intervalo $[a, b]$ se dice que es continua en $x = a$ si

$$\lim_{x \rightarrow a^+} f(x) = f(a)$$

Análoga definición se tiene para el extremo $x = b$.

Sea I un intervalo abierto de c y f una función definida en I salvo quizás el punto c . Si los límites laterales coinciden se dice que hay una *discontinuidad evitable*. Un ejemplo de discontinuidad evitable es la vista para $f(x) = (x^2 - 1)/(x + 1)$ en $x = -1$ vista en ejemplo 4.2. Si los límites laterales existen, pero no coinciden se dice que hay una *discontinuidad de salto finito*, por ejemplo, la función $f(x) = x/|x|$ vista en el ejemplo 4.6. En caso de que algún límite lateral valga infinito se dice que es una discontinuidad de *salto infinito* como ocurre para $f(x) = 4/(x - 2)^3$ para $x = 2$.

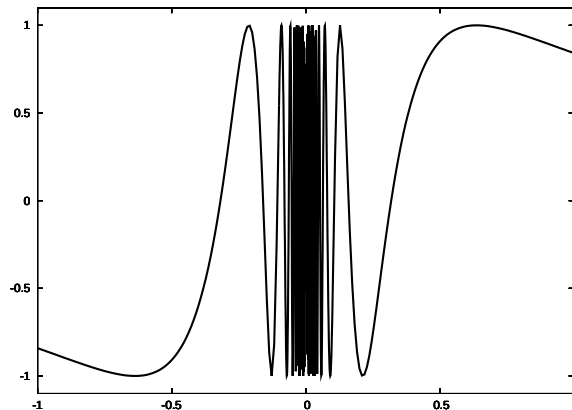


Figura 4.5: Ejemplo de discontinuidad que no es ni evitable ni de salto finito.

Teorema 4.9. Sea c un punto en un intervalo donde está definida la función f

1. Si $f(x)$ es continua en $x = c$, entonces $k f(x)$ es continua en $x = c$.

2. Si $f(x)$ y $g(x)$ es continua en $x = c$, entonces $f(x) + g(x)$ y $f(x)g(x)$ son continuas en $x = c$.
3. Si $f(x)$ y $g(x)$ es continua en $x = c$ con $g(c) \neq 0$, entonces $f(x)/g(x)$ es continua en $x = c$.
4. Si $f(x)$ es continua en $x = g(c)$ y $g(x)$ es continua en $x = c$, entonces $f(g(x))$ es continua en $x = c$.

Teorema 4.10 (Teorema del valor intermedio). Si $f(x)$ es una función continua definida en el intervalo $[a, b]$ y k un número intermedio de $f(a)$ y $f(b)$, entonces existe un c en $[a, b]$ tal que $f(c) = k$.

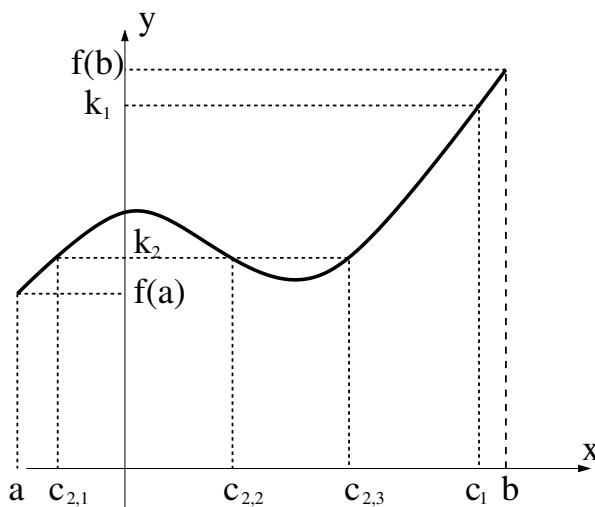


Figura 4.6: En la gráfica observamos que el valor k_1 tiene una única preimagen, c_1 , mientras que el valor k_2 tiene tres preimagenes $c_{2,1}$, $c_{2,2}$ y $c_{2,3}$.

4.4. Sucesiones. Límite.

En el lenguaje corriente, una sucesión es una ordenación de elementos. En matemáticas se define *sucesión* como una función cuyo dominio son los números enteros mayores que una cantidad n_0 . La sucesión se denota por $(a_n)_{n \geq n_0}$. Cuando no se expresa el n_0 se sobreentiende que el $n_0 = 0$. Por conveniencia a la hora de trabajar se emplea el término general o término n -ésimo de la sucesión que denotamos por a_n . Por ejemplo, para dar sucesiones simplemente escribimos $a_n = n^2$, $b_n = 2^n - 3^n$ y $c_n = 1/(1 + n^2)$.

En algunos casos los términos de una sucesión se definen utilizando el término precedente de la sucesión. En este caso, decimos que es una sucesión definida por *recurrencia*. En este caso es necesario dar el valores del primer término de la sucesión. Un ejemplo es $a_{n+1} = (n + 1)a_n$, con $n \geq 0$ y $a_0 = 1$. El término n -ésimo de la sucesión se llama *factorial* de n y se denota por $n!$. Para $n \geq 1$, $n!$ es el producto de los n primeros enteros positivos y $0! = 1$. Los primeros términos de la sucesión son

$$\begin{array}{cccccccc} 0! = 1, & 1! = 1, & 2! = 2, & 3! = 6, & 4! = 24, & 5! = 120, & 6! = 720, & 7! = 5040 \\ 8! = 40320, & 9! = 362880, & 10! = 3628800, & 11! = 39916800, & 12! = 479001600 \end{array}$$

Una sucesión $(a_n)_{n \geq 0}$ de números reales, diremos que tiene por límite un número real $-\infty < L < \infty$ si una vez que n ha excedido un determinado valor suficientemente grande, digamos N , todos los términos de la misma permanecen arbitrariamente cerca del número L . Cuando esta condición se cumple decimos que la sucesión es *convergente* y que su límite es L y escribimos de manera formal

$$\lim_{n \rightarrow \infty} a_n = L.$$

En el caso de que no exista el límite diremos que la sucesión es *divergente*.

La definición rigurosa y que en la práctica no emplearemos, afirma que $(a_n)_{n \geq 0}$ tiene por límite L , si para todo $\epsilon > 0$ es posible encontrar un número N (dependiente de ϵ) de forma que $|a_n - L| < \epsilon$ para todo $n \geq N$.

Teorema 4.11. Sea L un número real. Sea f una función de una variable real tal que $\lim_{x \rightarrow \infty} f(x) = L$. Si $a_n = f(n)$, entonces $\lim_{n \rightarrow \infty} a_n = L$.

Teorema 4.12. Si $\lim_{n \rightarrow \infty} a_n = L$ y $\lim_{n \rightarrow \infty} b_n = K$, entonces

1. $\lim_{n \rightarrow \infty} (a_n + b_n) = L + K$;
2. $\lim_{n \rightarrow \infty} c a_n = c L$;
3. $\lim_{n \rightarrow \infty} (a_n b_n) = LK$;
4. $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{L}{K}$ si $K \neq 0$.

Teorema 4.13. Si $\lim_{n \rightarrow \infty} a_n = L$, entonces $\lim_{n \rightarrow \infty} a_{n+N} = L$.

El resultado nos dice que para el valor del límite no importan los primeros términos de la sucesión. Por eso, en los resultados sobre límites sólo nos importa que esté definida para todo $n \geq N$ para algún N entero.

Teorema 4.14 (Encaje de sucesiones). Si $\lim_{n \rightarrow \infty} a_n = L = \lim_{n \rightarrow \infty} b_n$ y existe un entero $N > 0$ tal que $a_n \leq c_n \leq b_n$ para todo $n > N$, entonces

$$\lim_{n \rightarrow \infty} c_n = L.$$

Teorema 4.15. Si $\lim_{n \rightarrow \infty} |a_n| = 0$, entonces $\lim_{n \rightarrow \infty} a_n = 0$.

Una sucesión $(a_n)_{n \geq 0}$ es *creciente* si sus términos verifican $a_j \leq a_{j+1}$ para todo $j \geq 0$. Una sucesión $(a_n)_{n \geq 0}$ es estrictamente creciente si sus términos verifican $a_j < a_{j+1}$ para todo $j \geq 0$. Una sucesión $(a_n)_{n \geq 0}$ es *decreciente* si sus términos verifican $a_{j+1} \leq a_j$ para todo $j \geq 0$. Una sucesión $(a_n)_{n \geq 0}$ es estrictamente decreciente si sus términos verifican $a_{j+1} < a_j$ para todo $j \geq 0$. Una sucesión se dice que es monótona si es creciente o decreciente.

Una sucesión $(a_n)_{n \geq 0}$ se dice *acotada superiormente* si existe un número real M tal que $a_n \leq M$ para todo $n \geq 0$. El número M se llama *cota superior*. Una sucesión $(a_n)_{n \geq 0}$ se dice *acotada inferiormente* si existe un número real N tal que $N \leq a_n$ para todo $n \geq 0$. El número N se llama *cota inferior*. Se dice que es *acotada* si es acotada superior e inferiormente.

Teorema 4.16. Una sucesión monótona y acotada es convergente.

Ejemplo 4.12. Los límites los calculamos, básicamente empleando el argumento

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0.$$

Es del mismo modo evidente que $a_n = (-1)^n$ es divergente. Por otro lado se puede demostrar que $\lim_{n \rightarrow \infty} x^n = 0$ si $-1 < x < 1$, y que $(x^n)_{n \geq 0}$ es divergente para $x \leq -1$ y $x \geq 1$.

4.5. Progresiones aritméticas y geométricas.

4.5.1. Progresión aritmética

Se llama *progresión aritmética* de diferencia d y término inicial a_0 a una sucesión definida por la relación de recurrencia $a_{n+1} = a_n + d$ para $n \geq 0$. El término general de la sucesión es $a_n = a_0 + n d$.

Para determinar la suma de los $(n+1)$ -primeros términos, $S = a_0 + a_1 + \dots + a_{n-1} + a_n$, se observa que la suma de a_0 y a_n es lo mismo que la de a_1 y a_{n-1} y así sucesivamente

$$a_k + a_{n-k} = a_0 + kd + a_0 + (n-k)d = 2a_0 + nd = a_n + a_0.$$

Sumamos la cantidad S con ella misma, pero cambiando el orden de los sumandos, $a_n + a_{n-1} + \dots + a_1 + a_0$ y obtenemos que $2S = (a_n + a_0)(n+1)$. Con lo que la suma de los $(n+1)$ -primeros términos es

$$S = \frac{(a_n + a_0)(n+1)}{2}. \quad (4.1)$$

4.5.2. Progresión geométrica

Una sucesión geométrica de razón R y término inicial a_0 es una sucesión que se define de forma recurrente por $a_{n+1} = R a_n$. El término general de la sucesión es $a_n = a_0 R^n$.

Sea $S_n = a_0 + a_1 + \dots + a_{n-1} + a_n$ la suma de los $(n+1)$ -primeros términos de la progresión geométrica. Si multiplicamos la cantidad por R , obtenemos

$$R S_n = R a_0 + R a_1 + \dots + R a_{n-1} + R a_n = a_1 + a_2 + \dots + a_n + a_{n+1}.$$

Si consideramos la diferencia de S_n y $R S_n$, se tiene que

$$S_n - R S_n = (a_0 + a_1 + \dots + a_{n-1} + a_n) - (a_1 + a_2 + \dots + a_n + a_{n+1}) = a_0 - a_{n+1}.$$

Como $a_{n+1} = a_0 R^{n+1}$, al despejar S_n se consigue que

$$S_n = \frac{a_0 - a_0 R^{n+1}}{1 - R} = a_0 \frac{1 - R^{n+1}}{1 - R}. \quad (4.2)$$

La progresión geométrica en finanzas

La fórmula del interés compuesto sirve para un capital donde los intereses son reinvertidos es

$$C_f = C_i(1+r)^n \quad (4.3)$$

donde C_i es la cantidad de capital inicial, C_f es la cantidad compuesta al final de n periodos de interés con tasa periódica de interés r . Obsérvese que se trata de una sucesión geométrica de razón $R = 1+r$ y término inicial C_i .

Hay que distinguir los términos interés anual nominal e interés anual efectivo. Si el año se compone de m periodos donde la tasa periódica es r , el *interés nominal* es $m r$. El *interés efectivo* es obtener la tasa de crecimiento obtenida por nuestro capital al final del año cuando los intereses de cada periodo son reinvertidos. Dicha cantidad se calcula con la fórmula

$$r_e = \left(1 + \frac{r}{m}\right)^m - 1.$$

donde r es la tasa de interés nominal. En la asignatura de matemáticas salvo que se diga lo contrario cuando se habla de interés nos referimos al interés efectivo.

En finanzas el valor del dinero depende del momento en el que nos referimos, ya que el dinero puede ser invertido a una tasa de interés anual¹, r . Se dice *valor futuro* de una cantidad C a n años vista es $C(1+r)^n$. Se llama *valor presente* de un capital C dentro n años a $C(1+r)^{-n}$ que es el capital que se necesita invertir para obtener dentro de n años la cantidad C .

¹ Por simplicidad suponemos la hipótesis (no muy realista) de que la tasa de interés es constante

Ejemplo 4.13. Supongamos que el valor del tipo de interés es fijo a un 5 % anual. Para dentro de diez años deseamos asegurar la compra de un bien que valdrá 10000€ y dividimos el pago en tres plazos. El primero se efectúa en este momento por un valor de 3000€, el segundo por 2000€ se efectúa a los cinco años y el último plazo lo realizamos un año antes de la compra. ¿Cuál es la cantidad a pagar en el último plazo?

Los 3000€ de hoy tienen un valor futuro de $3000(1.05)^{10}$ € a los diez años y los 2000€ del segundo pago tienen un valor futuro a los diez años de $2000(1.05)^5$ €, ya que el dinero está invertido durante cinco años. El último plazo se paga una cantidad desconocida, P , que tiene por valor futuro $1.05 P$ € al año siguiente. La suma de los tres valores futuros debe ser igual a 10000€. En consecuencia, se tiene la ecuación

$$3000(1.05)^{10} + 2000(1.05)^5 + P 1.05 = 10000$$

despejando el valor de P se llega a

$$P = \frac{10000 - 3000(1.05)^{10} - 2000(1.05)^5}{1.05} = 10000(1.05)^{-1} - 3000(1.05)^9 - 2000(1.05)^4 \approx 2438.81$$

Observando la expresión final es un ejercicio interesante calcular dicho valor pensando en el valor del dinero cuando se ejecuta el último plazo.

Una *anualidad* es una sucesión finita de pagos hecha en periodos fijos de tiempo durante determinado lapso de tiempo. El intervalo de tiempo de la anualidad es el *término* de la anualidad. Un ejemplo de anualidad es depositar en una cuenta 100€ cada tres meses.

El *valor presente de la anualidad* es la suma del valor presente de todos los pagos a realizar. Suponemos que todos los pagos son realizados al final del periodo, lo que se conoce como *anualidad corriente*. Si la tasa de interés del periodo de pago es r y se efectúan n pagos por un importe de P €, entonces el valor presente de la anualidad, A , es

$$A = P(1+r)^{-1} + P(1+r)^{-2} + \dots + P(1+r)^{-n+1} + P(1+r)^{-n}$$

Si consideramos $R = (1+r)^{-1}$, la expresión anterior se escribe como

$$A = PR + PR^2 + \dots + PR^{n-1} + PR^n = P + PR + PR^2 + \dots + PR^{n-1} + PR^n - P$$

Los n primeros sumandos son la suma de los n primeros término de una progresión geométrica de razón R y término inicial P . Podemos aplicar la fórmula (4.2), obteniendo que el valor presente es

$$A = P \frac{1 - R^{n+1}}{1 - R} - P = P \frac{R - R^{n+1}}{1 - R}$$

Como $R = (1+r)^{-1}$, se tiene que

$$A = P \frac{1 - (1+r)^{-n}}{r}. \quad (4.4)$$

La cantidad $\frac{1 - (1+r)^{-n}}{r}$ se denota $a_{\overline{n}|r}$. Por $a_{\overline{n}|r}$ denotamos el valor presente de la anualidad de una unidad monetaria de duración n periodos a una tasa de interés periódica r . Esta cantidad es usada para el cálculo de las amortizaciones de los préstamos.

Ejemplo 4.14. Si se pide un préstamo de 1000€ al 6 % de interés y se desea pagar mediante una anualidad durante diez años. ¿Cuál es el precio de la anualidad?

La tasa de interés es $r = 0.06$. Si la cantidad a pagar anualmente es A , el valor presente de la anualidad es $Aa_{\overline{10}|0.06}$ debe ser igual al préstamo pedido 1000. Tenemos la ecuación $1000 = Aa_{\overline{10}|0.06}$, con lo que

$$A = \frac{1000}{a_{\overline{10}|0.06}} = \frac{1000}{\frac{1 - (1+r)^{-10}}{r}} = \frac{1000}{\frac{1 - (1.06)^{-10}}{0.06}} = 135.8679582 \dots$$

4.6. Series.

Cada vez los préstamos hipotecarios duran más tiempo, con ello se busca menor cuota, pero que pasaría si se tiene un préstamo a tiempo perpetuo donde la amortización estuviese en el final de los tiempos. Si la cuota del préstamo es A , P el préstamo pedido y r la tasa de interés, entonces igualando el valor presente de la cantidad presente y la anualidad se tiene

$$P = A \frac{1}{1+r} + A \frac{1}{(1+r)^2} + A \frac{1}{(1+r)^3} + A \frac{1}{(1+r)^4} + A \frac{1}{(1+r)^5} + A \frac{1}{(1+r)^6} + A \frac{1}{(1+r)^7} + A \frac{1}{(1+r)^8} + \dots$$

Observamos que se trata de una suma de infinitos términos. ¿Qué sentido tienen la suma de infinitos términos? En este apartado daremos un sentido a esta expresión.

4.6.1. Series y convergencia

Una *serie* de números reales es una expresión del tipo

$$a_0 + a_1 + a_2 + a_3 + \dots + a_n + \dots$$

en la que se suman todos los términos de una sucesión de números reales. Para representar la suma infinita en forma más compacta, emplearemos la notación: $\sum_{n=0}^{\infty} a_n$. El índice del sumatorio n es lo que se llama una variable muda, en el sentido que puede ser reemplazada por cualquier otro símbolo sin que por ello cambie el sentido de la suma, y en consecuencia

$$\sum_{n=0}^{\infty} a_n = \sum_{r=0}^{\infty} a_r = \dots = \sum_{s=0}^{\infty} a_s.$$

Consideremos la suma finita de los $(n+1)$ -primeros términos

$$S_n = a_0 + a_1 + a_2 + \dots + a_n = \sum_{r=0}^n a_r$$

a la que llamaremos la n -ésima *suma parcial* de la serie $\sum_{n=0}^{\infty} a_n$. Las sumas parciales forman una sucesión de números reales $S_0, S_1, S_2, \dots, S_n, \dots$ con término general S_n . Diremos que la serie infinita $\sum_{n=0}^{\infty} a_n$ *converge* a L si $\lim_{n \rightarrow \infty} S_n = L$. Si el límite no existe diremos que la serie es *divergente*.

Teorema 4.17. Consideremos dos series convergentes $\sum_{n=0}^{\infty} a_n = A$ y $\sum_{n=0}^{\infty} b_n = B$. Entonces se cumple para cualquier número real k

$$\sum_{n=0}^{\infty} k a_n = k A, \quad \sum_{n=0}^{\infty} (a_n + b_n) = A + B.$$

Ejemplo 4.15. Calcular la suma de la serie geométrica $\sum_{k=0}^{\infty} \left(-\frac{2}{3}\right)^k$

La suma de una serie geométrica de razón r viene dada por la fórmula

$$S_n = 1 + \left(-\frac{2}{3}\right) + \left(-\frac{2}{3}\right)^2 + \dots + \left(-\frac{2}{3}\right)^{n-1} = \sum_{k=0}^n \left(-\frac{2}{3}\right)^k.$$

es la suma de una serie geométrica de razón $r = -2/3$, luego

$$S_n = \frac{1 - (-2/3)^{n+1}}{1 - (-2/3)},$$

como $\lim_{n \rightarrow \infty} (-2/3)^n = 0$, tenemos que

$$\lim_{n \rightarrow \infty} S_n = \frac{1}{1 - (-2/3)} = \frac{3}{5},$$

en consecuencia la serie converge a $3/5$

$$\sum_{n=0}^{\infty} \left(-\frac{2}{3}\right)^n = \frac{3}{5}.$$

Ejemplo 4.16. Calcular $\sum_{n=10}^{\infty} \left(\frac{1}{2}\right)^n$.

Del teorema 4.17 tenemos que $\sum_{n=10}^{\infty} \left(\frac{1}{2}\right)^n = \frac{1}{2^{10}} \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n$. Por la fórmula de la suma de una progresión geométrica (4.2) tenemos que

$$S_n = \sum_{k=0}^n \left(\frac{1}{2}\right)^k = \frac{1 - (1/2)^{n+1}}{1 - 1/2} = 2 - \left(\frac{1}{2}\right)^n$$

Como $\lim_{n \rightarrow \infty} S_n = 2$, tenemos que

$$\sum_{n=10}^{\infty} \left(\frac{1}{2}\right)^n = \frac{1}{2^{10}} \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = \frac{1}{2^{10}} \cdot 2 = \frac{1}{2^9}.$$

Las series de la forma $\sum_{n=n_0}^{\infty} r^n$ donde $|r| < 1$ se pueden sumar por el procedimiento anterior y se conocen como *series geométricas*.

Hay otro caso de serie que es posible sumar. Se llaman *telescopicas* a aquellas series en cuya suma se cancelan términos de forma que es fácil expresar fácilmente su suma parcial. Para entender como se producen las cancelaciones cuando es el sumatorio de una función racional es necesario efectuar la descomposición en fracciones simples vista en el tema anterior.

Ejemplo 4.17. Calcular $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$

Observemos que $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$. La suma parcial en este caso es

$$S_n = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n \cdot (n+1)} = \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n+1}\right),$$

observando que el último término de un paréntesis se cancela con el primero del siguiente paréntesis, escribimos que $S_n = 1 - \frac{1}{n+1}$, y en consecuencia se tiene $\lim_{n \rightarrow \infty} S_n = 1$.

Ejemplo 4.18. Calcular $\sum_{n=0}^{\infty} \frac{1}{n^2 + 4n + 3}$

Hacemos la descomposición en fracciones simples de la función racional que aparece en el sumatorio.

La factorización de $n^2 + 4n + 3$ es $(n + 1)(n + 3)$. La descomposición en fracciones simples dejando los coeficientes indeterminados es

$$\frac{1}{n^2 + 4n + 3} = \frac{A}{n + 1} + \frac{B}{n + 3}$$

De la expresión $1 = A(n + 3) + B(n + 1)$, se obtiene $A = 1/2$ y $B = -1/2$. Así, tenemos que

$$\sum_{n=0}^{\infty} \frac{1}{n^2 + 4n + 3} = \sum_{n=0}^{\infty} \left(\frac{1}{2} \frac{1}{n + 1} - \frac{1}{2} \frac{1}{n + 3} \right) = \frac{1}{2} \sum_{n=0}^{\infty} \left(\frac{1}{n + 1} - \frac{1}{n + 3} \right)$$

Las sumas parciales, S_n de $\sum_{n=0}^{\infty} \left(\frac{1}{n + 1} - \frac{1}{n + 3} \right)$ son

$$\left(\frac{1}{1} - \frac{1}{3} \right) + \left(\frac{1}{2} - \frac{1}{4} \right) + \left(\frac{1}{3} - \frac{1}{5} \right) + \left(\frac{1}{4} - \frac{1}{6} \right) + \dots + \left(\frac{1}{n-1} - \frac{1}{n+1} \right) + \left(\frac{1}{n} - \frac{1}{n+2} \right) + \left(\frac{1}{n+1} - \frac{1}{n+3} \right)$$

Observando que el último sustraendo de cada paréntesis se cancela con el primer término de dos paréntesis a la derecha, tenemos que $S_n = 1 + \frac{1}{2} - \frac{1}{n+2} - \frac{1}{n+3}$ y tomando límite $\sum_{n=0}^{\infty} \left(\frac{1}{n+1} - \frac{1}{n+3} \right) = \frac{3}{2}$. Por tanto,

$$\sum_{n=0}^{\infty} \frac{1}{n^2 + 4n + 3} = \frac{1}{2} \sum_{n=0}^{\infty} \left(\frac{1}{n+1} - \frac{1}{n+3} \right) = \frac{3}{4}.$$

Ejemplo 4.19. Calcular $\sum_{n=1}^{\infty} \frac{1}{n}$

Consideremos la suma parcial $S_{2^n} : S_{2^n} = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{2^n}$. Entonces, agrupamos de la siguiente forma

$$S_{2^n} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4} \right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \right) + \dots + \left(\frac{1}{2^{n-1}+1} + \dots + \frac{1}{2^n} \right).$$

Nótese que:

$$\left(\frac{1}{3} + \frac{1}{4} \right) > \left(\frac{1}{4} + \frac{1}{4} \right) = \frac{1}{2} \quad \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \right) > \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} \right) = \frac{1}{2},$$

así hasta el final donde

$$\left(\frac{1}{2^{n-1}+1} + \dots + \frac{1}{2^n} \right) > \left(\frac{1}{2^n} + \dots + \frac{1}{2^n} \right) = \frac{1}{2}.$$

En consecuencia obtenemos que

$$S_{2^n} > 1 + \frac{1}{2} + \frac{1}{2} + \dots + \frac{1}{2} = 1 + \frac{n}{2}$$

por lo tanto

$$\lim_{n \rightarrow \infty} S_{2^n} = +\infty$$

y la serie es divergente. A esta serie se le conoce con el nombre de *serie armónica*.

Teorema 4.18. Si la serie $\sum_{n=0}^{\infty} a_n$ es convergente, entonces $\lim_{n \rightarrow \infty} a_n = 0$.

Demostración. Nótese que $a_n = S_n - S_{n-1}$, como la serie es convergente

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} S_n - \lim_{n \rightarrow \infty} S_{n-1} = L - L = 0. \quad \square$$

Consecuencia del anterior resultado es el siguiente teorema

Teorema 4.19 (Criterio de divergencia). Si $\lim_{n \rightarrow \infty} a_n \neq 0$, entonces la serie $\sum_{n=0}^{\infty} a_n$ es divergente.

Ejemplo 4.20. Determinar la convergencia de la serie $\sum_{n=1}^{\infty} \frac{2n}{5n+3}$

Como $\lim_{n \rightarrow \infty} \frac{2n}{5n+3} = \frac{2}{5} \neq 0$ la serie es divergente.

4.6.2. Comparación de series

Teorema 4.20 (Criterio de comparación). Si la serie de términos positivos $\sum_{n=0}^{\infty} a_n$ es convergente y la serie $\sum_{n=0}^{\infty} b_n$ cumple que $0 \leq b_n \leq a_n$ entonces la serie $\sum_{n=0}^{\infty} b_n$ es convergente.

Ejemplo 4.21. La serie $\sum_{n=0}^{\infty} \frac{1}{2^n + n}$ es convergente, ya que $0 \leq \frac{1}{2^n + n} \leq \frac{1}{2^n}$ y $\sum_{n=0}^{\infty} \frac{1}{2^n}$ es convergente.

4.6.3. Convergencia absoluta y condicional

Si una serie $\sum_{n=0}^{\infty} a_n$ contiene términos positivos y negativos es convergente y la suma de los valores absolutos $\sum_{n=0}^{\infty} |a_n|$ es convergente, la serie $\sum_{n=0}^{\infty} a_n$ se dice que es *absolutamente convergente*.

Si una serie $\sum_{n=0}^{\infty} a_n$ contiene términos positivos y negativos es convergente y la suma de los valores absolutos $\sum_{n=0}^{\infty} |a_n|$ es divergente, la serie $\sum_{n=0}^{\infty} a_n$ se dice que es *condicionalmente convergente*.

Teorema 4.21. Si la serie $\sum_{n=0}^{\infty} a_n$ es absolutamente convergente, entonces

1. la serie $\sum_{n=0}^{\infty} a_n$ es convergente
2. Se pueden reagrupar términos de la serie sin que se altere el valor de la suma.
3. Se pueden reordenar términos sin que se altere el valor de la suma.

Si la serie es condicionalmente convergente este tipo de operaciones *no son válidas*. De hecho, se prueba que reordenando la serie se puede obtener cualquier valor.

Ejemplo 4.22. Emplea el criterio de comparación para determinar la convergencia o divergencia de la serie $\sum_{n=1}^{\infty} \frac{(-1)^n}{3^n + 1}$.

El valor absoluto del término general verifica la acotación

$$\left| \frac{(-1)^n}{3^n + 1} \right| = \frac{1}{3^n + 1} \leq \frac{1}{3^n}$$

Como $\sum_{n=0}^{\infty} \frac{1}{3^n}$ es convergente, $\sum_{n=0}^{\infty} \left| \frac{(-1)^n}{3^n + 1} \right|$ es convergente por el criterio de comparación (teorema 4.20) y la serie $\sum_{n=1}^{\infty} \frac{(-1)^n}{3^n + 1}$ es absolutamente convergente.

4.6.4. El criterio del cociente

Teorema 4.22 (Criterio del cociente). Si la serie $\sum_{n=0}^{\infty} a_n$ cumple que $a_n \neq 0$ y que

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = L,$$

entonces

1. Si $L < 1$ la serie $\sum_{n=0}^{\infty} a_n$ es absolutamente convergente.
2. Si $L > 1$ la serie es divergente.

Comentario al teorema: Cuando $L = 1$ el límite no nos proporciona ninguna información sobre la convergencia de la serie. Por ejemplo, $\sum_{n=0}^{\infty} 1$ es divergente, $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ es convergente como se vio en el ejemplo 4.17, y el límite L del teorema es uno.

Demostración. Supongamos que el límite es $L < 1$. Entonces podemos fijar un número r de forma que $L < r < 1$. De la existencia del límite podemos deducir que para $n > N$ se cumple que $|a_{n+1}| < r|a_n|$. En consecuencia se tiene

$$|a_{N+2}| < r|a_{N+1}|, |a_{N+3}| < r|a_{N+2}| < r^2|a_{N+1}|, \dots$$

en general $|a_{N+m+1}| < r^m|a_{N+1}|$. Entonces

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^N a_n + \sum_{n=N+1}^{\infty} a_n = \sum_{n=1}^N a_n + R_N$$

donde R_N es el resto de la suma de la serie. Este resto cumple

$$|R_N| = \left| \sum_{n=N+1}^{\infty} a_n \right| \leq \sum_{n=N+1}^{\infty} |a_n| < |a_{N+1}|(1 + r + r^2 + r^3 + \dots)$$

como la suma en el paréntesis es una serie geométrica de razón $r < 1$ el resto es una serie convergente, y en consecuencia la serie original es absolutamente convergente. Si $r > 1$, entonces $\lim_{n \rightarrow \infty} |a_n| \neq 0$ y la serie es entonces divergente. $\square$

Ejemplo 4.23. Determinar por el criterio del cociente si es convergente $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n!}$. Consideremos

$$a_n = \frac{(-1)^{n+1}}{n!}, \text{ entonces}$$

$$\frac{a_{n+1}}{a_n} = \frac{-n!}{(n+1)!} = \frac{-1}{n+1}$$

por lo tanto

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 0 < 1$$

por lo que la serie es absolutamente convergente. Con lo que se ha demostrado que $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n!}$ y $\sum_{n=1}^{\infty} \frac{1}{n!}$ son convergentes.

Ejemplo 4.24. Determinar por el criterio del cociente si es convergente $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2}$.

Consideremos $a_n = \frac{(-1)^{n+1}}{n^2}$, entonces $\frac{a_{n+1}}{a_n} = \frac{-n^2}{(n+1)^2}$ luego

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 1$$

con lo que en este caso no podemos decidir nada.

4.7. Series de potencias. Radio de convergencia.

4.7.1. Series de potencias

Dado un número real x_0 , se llama *serie de potencias* centrada en $x = x_0$, a una expresión del tipo

$$a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n + \cdots = \sum_{n=0}^{\infty} a_n(x - x_0)^n$$

siendo los coeficientes de la serie a_n son números reales dados. Si empleamos el criterio del cociente, podemos afirmar que esta serie es absolutamente convergente si

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}(x - x_0)^{n+1}}{a_n(x - x_0)^n} \right| < 1,$$

o de forma equivalente

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| |x - x_0| < 1.$$

Si asumimos que $a_n \neq 0$ obtenemos que

$$|x - x_0| < \lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right|.$$

En consecuencia, si existe el límite

$$\lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right| = R$$

entonces la serie es absolutamente convergente para los valores de x que cumplen $|x - x_0| < R$. Esta desigualdad define el intervalo abierto $(x_0 - R, x_0 + R)$ en la cual la serie es siempre absolutamente convergente. Al número R se le conoce con el nombre de *radio de convergencia* de la serie de potencias y al intervalo abierto $(x_0 - R, x_0 + R)$ *intervalo de convergencia*.

Ejemplo 4.25. Encontrar el radio de convergencia de la serie

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}.$$

Esta serie de potencias está expandida alrededor del punto $x_0 = 0$ y el n -ésimo coeficiente se escribe como $a_n = (-1)^{n+1} \frac{1}{n}$. Por lo tanto el radio de convergencia de la serie es

$$\lim_{n \rightarrow \infty} \left| \frac{a_n}{a_{n+1}} \right| = \lim_{n \rightarrow \infty} \left| \frac{n+1}{n} \right| = 1.$$

Consecuentemente el intervalo de convergencia es $(-1, 1)$.

Es importante observar que una serie de potencia es una nueva forma de definir funciones. Así, $f(x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}$ es una función cuyo dominio es el intervalo $(-1, 1)$.

De lo visto para la progresión geométrica se deduce que

$$f(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad \text{para } -1 < x < 1$$

Por tanto, $\sum_{n=0}^{\infty} x^n$ y $\frac{1}{1-x}$ coinciden en el intervalo $(-1, 1)$. En la figura 4.7 se puede observar como al ir sumando cada vez más sumando se parece cada vez más a $1/(1-x)$.

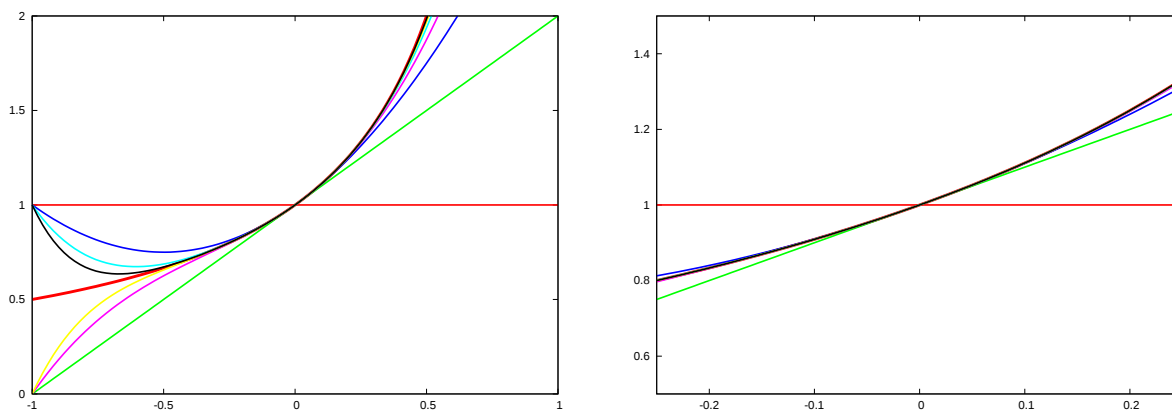


Figura 4.7: Sumas parciales de la serie de potencia $\sum_{n=0}^{\infty} x^n$ y su suma $\frac{1}{1-x}$ representadas en el intervalo $(-1, 1)$ en el panel de la izquierda y el $(-1/4, 1/4)$ en el panel de la derecha.

Ejemplo 4.26. Calcular el radio de convergencia de $\sum_{n=0}^{\infty} \frac{1}{2^n + 3^n} x^{2n+1}$.

Podemos sacar factor común x de la serie y obtenemos $x \sum_{n=0}^{\infty} \frac{1}{2^n + 3^n} x^{2n}$, el radio de la serie de potencia es el mismo que el radio de potencia de $\sum_{n=0}^{\infty} \frac{1}{2^n + 3^n} x^{2n}$, ya que el radio de potencia determina el intervalo donde la expresión de la serie tiene sentido. Llamando $y = x^2$, tenemos que respecto la variable y la serie de potencia es $\sum_{n=0}^{\infty} \frac{1}{2^n + 3^n} y^n$. A esta expresión le aplicamos la fórmula del radio de convergencia:

$$R = \lim_{n \rightarrow \infty} \frac{\frac{1}{2^n} + \frac{1}{3^n}}{\frac{1}{2^{n+1}} + \frac{1}{3^{n+1}}} = \lim_{n \rightarrow \infty} \frac{2 + \frac{2^{n+1}}{3^n}}{1 + \frac{2^{n+1}}{3^{n+1}}} = \lim_{n \rightarrow \infty} \frac{2 + 2 \left(\frac{2}{3}\right)^n}{1 + \left(\frac{2}{3}\right)^{n+1}} = 2.$$

La serie tiene sentido cuando $-2 < y < 2$. Como $y = x^2$, la serie tiene sentido en el intervalo $(-\sqrt{2}, \sqrt{2})$ y el radio de convergencia de la serie original es $\sqrt{2}$.

Capítulo 5

Derivada. Reglas básicas de derivación

5.1. Concepto de derivada. Marginal

Dada una magnitud económica y en función de otra magnitud x de forma que $y = f(x)$ se llama *tasa media de variación* de f en el intervalo $[x_1, x_2]$ al cociente de la variación de la magnitud y entre la variación de x , es decir,

$$\frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

En la figura 5.1 se puede observar que la tasa de variación es la pendiente de la recta que une los puntos (x_1, y_1) y (x_2, y_2) .

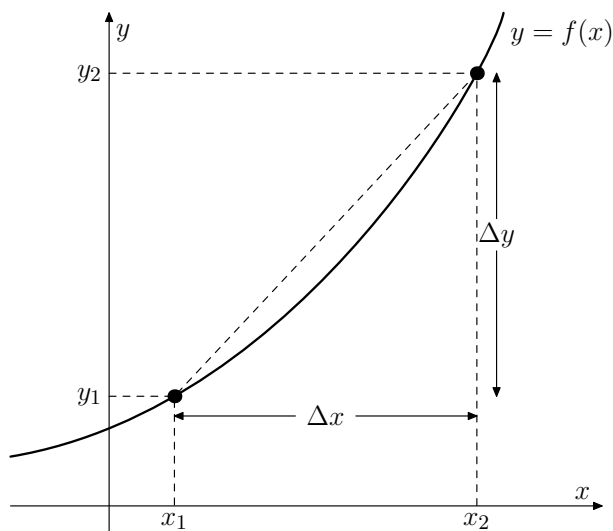


Figura 5.1: Interpretación gráfica de la tasa de variación.

Si la magnitud y fuera la distancia recorrida por un móvil y x fuese el tiempo, entonces la tasa de variación es de la velocidad media en un lapso de tiempo.

Para conocer la “velocidad” instantánea (y no la media) con la que cambia una magnitud económica y al variar x para un valor $x = x_0$, se trata de tomar un Δx tan pequeño como sea posible, lo que nos

conduce a emplear límites y que dicha “velocidad” instantánea sea definida por

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

A esta velocidad es a lo que llamamos *tasa instantánea de variación* de la magnitud y respecto x en $x = x_0$. Si consideramos una función f sin sentido “físico” o “económico” esta tasa de variación instantánea es lo que llamamos derivada.

Una función $f(x)$ definida en un intervalo abierto que contiene al punto $x = a$ se dice que es *derivable* en $x = a$ si existe el límite

$$\lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h} \quad (5.1)$$

y a dicho límite se le llama *derivada* de $f(x)$ en $x = a$ y lo denotamos por $f'(a)$, $\frac{df}{dx}(a)$ o si $y = f(x)$ como $\frac{dy}{dx}(a)$. Si la función f es derivable en un intervalo, hablamos de la función derivada denotada por f' ó $\frac{df}{dx}$. Muchas veces escribimos $\frac{d}{dx}$ seguido de la expresión que vamos a derivar entre corchetes, por eso, “calcula $\frac{d}{dx}[(x^3 - 1)^2]$ ” es equivalente a “dada $f(x) = (x^3 - 1)^2$, calcula $f'(x)$ ”.

Teorema 5.1. Si la función f es diferenciable en $x = a$, entonces f es continua en $x = a$.

Ejemplo 5.1. La función constante $f(x) = k$ es derivable en todo punto y $f'(x) = 0$, ya que

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{k - k}{h} = \lim_{x \rightarrow a} \frac{0}{h} = 0.$$

Ejemplo 5.2. La función $f(x) = x$ es derivable en todo punto y $f'(x) = 1$, ya que

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{x + h - x}{h} = \lim_{h \rightarrow 0} \frac{h}{h} = \lim_{h \rightarrow 0} 1 = 1.$$

Ejemplo 5.3. La función $f(x) = x^2$ es derivable en todo punto y $f'(x) = 2x$, ya que

$$\begin{aligned} f'(a) &= \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{(x + h)^2 - x^2}{h} = \lim_{h \rightarrow 0} \frac{x^2 + 2xh + h^2 - x^2}{h} = \\ &= \lim_{h \rightarrow 0} \frac{2xh + h^2}{h} = \lim_{h \rightarrow 0} 2x + h = 2x. \end{aligned}$$

Ejemplo 5.4. La función $f(x) = \frac{1}{x} = x^{-1}$ es derivable en $x \neq 0$ y su derivada es $f'(x) = -x^{-2}$.

$$\begin{aligned} f'(x) &= \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{\frac{1}{x + h} - \frac{1}{x}}{h} = \lim_{h \rightarrow 0} \frac{\frac{x - (x + h)}{(x + h)x}}{h} = \\ &= \lim_{h \rightarrow 0} -\frac{h}{hx(x + h)} = \lim_{h \rightarrow 0} -\frac{1}{x(x + h)} = -\frac{1}{x^2}. \end{aligned}$$

Ejemplo 5.5. La función $f(x) = |x|$ es continua en $x = 0$, pero no es diferenciable en $x = 0$, ya que no coinciden los límites laterales del límite que define la derivada:

$$\lim_{h \rightarrow 0^+} \frac{f(0 + h) - f(0)}{h} = \lim_{h \rightarrow 0^+} \frac{h}{h} = 1, \quad \neq \quad \lim_{h \rightarrow 0^-} \frac{f(0 + h) - f(0)}{h} = \lim_{h \rightarrow 0^-} \frac{-h}{h} = -1.$$

A dichos límites laterales se les llama derivadas laterales. Cuando no coinciden, la función no es derivable

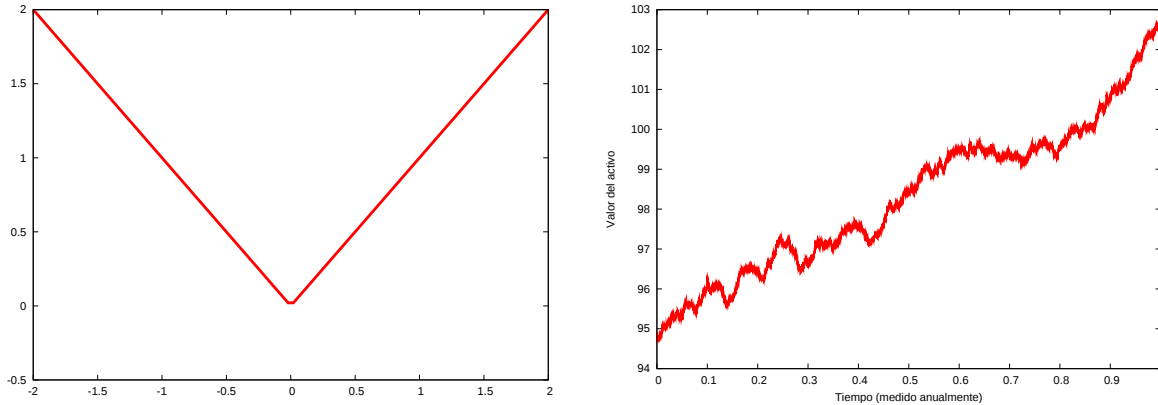


Figura 5.2: Ejemplos de funciones no derivables. La no derivabilidad se manifiesta en la presencia de un “pico” en el punto. En el panel de la izquierda aparece la gráfica del valor absoluto donde se observa la presencia de un “pico” en $x = 0$. En el panel de la derecha, la simulación del precio de un activo en la bolsa donde si se hacen sucesivas ampliaciones se descubre que para cada punto existe un “pico”, con lo que dicha función no es derivable en ninguna de sus puntos.

Si $C(x)$ es el coste de producción de x unidades se le llama coste marginal de producción y nos informa de la variación en el coste de producción ante cambio en la producción. Análogamente, se habla de ingreso por ventas marginal, beneficio de producción marginal y otros casos.

En economía muchas veces se estudia la variación en tantos por cientos y no en términos absolutos, por ejemplo, que en un año el producto interior bruto de Gibraltar aumente 100 millones de dolares puede considerarse un éxito para el país, pero esa cantidad para los Estados Unidos significa estar en un periodo de recesión. Con ese propósito se define que la *tasa proporcional de variación* de f en $x = a$ como el cociente $f'(a)/f(a)$.

Sean y y x dos magnitudes económicas relacionadas por una función $y = f(x)$ y un valor $x = x_0$, la derivada se emplea para aproximar el cambio en y para pequeños cambios en x . De esta forma Δy se estima como $f'(x_0)\Delta x$.

5.2. Recta tangente a una función

Una interpretación geométrica fundamental: la derivada es la pendiente de la recta tangente. Dada una función $y = f(x)$, la recta tangente en el punto $(x_0, f(x_0))$ es la recta a la que se aproximan las rectas secantes que unen $(x_0, f(x_0))$ con $(x, f(x))$ cuando el punto x se acerca a x_0 . Un ejemplo de cómo las secantes aproximan a una recta aparece en la figura 5.3.

La pendiente de la recta secante que pasa por $(x_0, f(x_0))$ con $(x, f(x))$ es el cociente $\frac{f(x) - f(x_0)}{x - x_0}$. La recta tangente se obtiene cuando el valor x se acerca a x_0 . Si llamamos $h = x - x_0$ dicho límite se escribe como en (5.1)

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Si f es una función derivable en x_0 e $y_0 = f(x_0)$ se llama *recta tangente* a la curva $y = f(x)$ en el punto (x_0, y_0) a la recta que pasa por (x_0, y_0) y tiene por pendiente $f'(x_0)$. Por tanto, la ecuación de la recta es

$$y - y_0 = f'(x_0)(x - x_0) \quad (5.2)$$

Si f es continua en x_0 y $\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \infty$ se llama *recta tangente* a la recta vertical $x = x_0$.

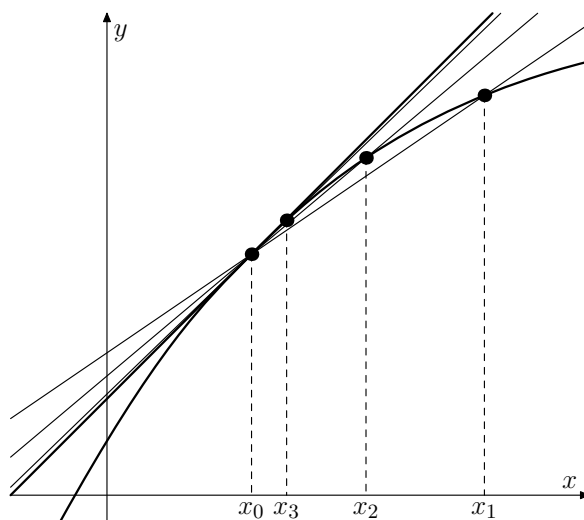


Figura 5.3: Dado un punto x_0 y una función f , si trazamos la secante que pasa por $(x_1, f(x_1))$ y $(x_0, f(x_0))$, la que pasa por $(x_2, f(x_2))$ y $(x_0, f(x_0))$ y la que pasa por $(x_3, f(x_3))$ y $(x_0, f(x_0))$, observamos como dichas rectas se van aproximando a la recta tangente de $y = f(x)$ en $x = x_0$ que aparece con trazo grueso.

Ejemplo 5.6. Encuentre la recta tangente a la curva $y = x^2$ en el punto $(2, 4)$.

Tenemos que $f(x) = x^2$ y en el ejemplo 5.3 vimos que $f'(x) = 2x$. Aplicando la fórmula (5.2) se obtiene que $y - 4 = f'(2)(x - 2)$. Como $f'(2) = 2 \cdot 2 = 4$, tenemos que la recta tangente es $y - 4 = 4(x - 2)$. Despejando la variable y se expresa como $y = 4x - 4$.

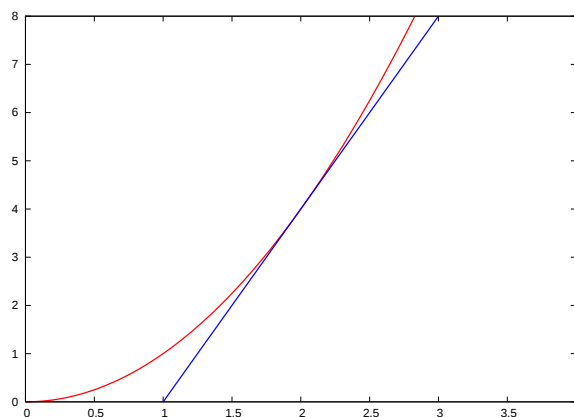


Figura 5.4: La gráfica de $y = x^2$ y la recta tangente a dicha curva en el punto $(2, 4)$

5.3. Derivada de la suma y el producto

5.3.1. Regla de la suma

Teorema 5.2. Si f es derivable en a , entonces kf es derivable y

$$\frac{d}{dx}[kf](a) = k \frac{df}{dx}(a).$$

Si f y g son derivables en a , entonces $f + g$ es derivable en a y

$$\frac{d}{dx}[f + g](a) = \frac{df}{dx}(a) + \frac{dg}{dx}(a) = f'(a) + g'(a).$$

Ejemplo 5.7. Derivar $f(x) = 3x^2 - x + 3$.

Usando el teorema 5.2 y los ejemplos 5.1, 5.2 y 5.3, se obtiene que

$$\frac{d}{dx}[f(x)] = \frac{d}{dx}[3x^2 - x + 3] = \frac{d}{dx}[3x^2] - \frac{d}{dx}[x] + \frac{d}{dx}[3] = 3 \frac{d}{dx}[x^2] - 1 + 0 = 3(2x) - 1 = 6x - 1.$$

5.3.2. Regla del producto

Teorema 5.3. Si f y g son derivable en $x = a$, entonces el producto fg es derivable en $x = a$

$$\frac{d}{dx}[fg](a) = f'(a)g(a) + f(a)g'(a).$$

Demostración. Es simplemente calcular un límite

$$(fg)'(a) = \lim_{h \rightarrow 0} \frac{(fg)(a+h) - (fg)(a)}{h} = \lim_{h \rightarrow 0} \frac{f(a+h)g(a+h) - f(a)g(a)}{h}.$$

La idea es sumar y resta $f(a)g(a+h)$ en el numerador y, posteriormente agrupar términos de forma conveniente

$$\begin{aligned} (fg)'(a) &= \lim_{h \rightarrow 0} \frac{f(a+h)g(a+h) - f(a)g(a+h) + f(a)g(a+h) - f(a)g(a)}{h} \\ &= \lim_{h \rightarrow 0} g(a+h) \frac{f(a+h) - f(a)}{h} + f(a) \lim_{h \rightarrow 0} \frac{g(a+h) - g(a)}{h} = g(a)f'(a) + f(a)g'(a). \quad \square \end{aligned}$$

Ejemplo 5.8. La función $f(x) = x^3$ es derivable en todo punto y $f'(x) = 3x^2$. La función x^3 se escribe como el producto de x^2 por x y se aplican el teorema 5.3 más los ejemplos 5.2 y 5.3 para derivar

$$\frac{d}{dx}[f(x)] = \frac{d}{dx}[x^3] = \frac{d}{dx}[x^2x] = \frac{d}{dx}[x^2]x + x^2 \frac{d}{dx}[x] = (2x)x + x^2(1) = 2x^2 + x^2 = 3x^2.$$

5.4. Regla de la cadena

Si el marginal de las ganancias (G) respecto de la producción (Q) (es decir, la “velocidad” de crecimiento en las ganancias respecto de las unidades producidas) es de 3€/unidad y la tasa de variación de la producción respecto del tiempo es de 100unidad/año, se tiene que la tasa de variación de las ganancias respecto del tiempo vale

$$\frac{\Delta G}{\Delta t} = \frac{\Delta G}{\Delta Q} \frac{\Delta Q}{\Delta t} = 3 \frac{\text{€}}{\text{unidad}} 100 \frac{\text{unidad}}{\text{año}} = 300 \frac{\text{€}}{\text{año}}.$$

En este ejemplo hemos manejado magnitudes con “velocidad” constante. ¿Qué sucede cuando la velocidad no es constante? Tomando límite y haciendo el tiempo “muy pequeño”, se obtiene intuitivamente que

$$\frac{dG}{dt} = \frac{dG}{dQ} \frac{dQ}{dt}.$$

Este resultado es conocido como regla de la cadena y su demostración es excesivamente técnica para el ámbito de estos apuntes.¹ La regla de la cadena se enuncia como sigue

Teorema 5.4. Si g es derivable en a y f es derivable en $g(a)$, entonces la función composición $(f \circ g)(x) = f(g(x))$ es derivable en $x = a$ y su derivada es

$$\frac{d}{dx}[f \circ g](a) = f'(g(a))g'(a).$$

Ejemplo 5.9. Derivar $f(x) = (x^2 + 1)^3$ usando la regla de la cadena. Llamemos $y = (x^2 + 1)^3 = z^3$ con $z = x^2 + 1$ y usando la regla de la cadena

$$\frac{dy}{dx} = \frac{dy}{dz} \frac{dz}{dx} = (3z^2)(2x) = 6(x^2 + 1)^2 x.$$

5.4.1. Regla del cociente

Teorema 5.5. Si f es derivable en $x = a$ y $f(a) \neq 0$, entonces

$$\frac{d}{dx}\left[\frac{1}{f}\right](a) = -\frac{f'(a)}{f(a)^2}.$$

Demostración. Con el ejemplo 5.4 y la regla de la cadena (teorema 5.4). □

Teorema 5.6. Si f y g son derivables en $x = a$ y $g(a) \neq 0$, entonces

$$\frac{d}{dx}\left[\frac{f}{g}\right](a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Demostración. Se puede deducir de los teoremas 5.3 y 5.5.

$$\frac{d}{dx}\left[f\left(\frac{1}{g}\right)\right](a) = f'\left(\frac{1}{g}\right)(a) + f(a)\frac{d}{dx}\left[\frac{1}{g}\right](a) = \frac{f'(a)}{g(a)} + f(a)\frac{-g'(a)}{g(a)^2} = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}. \quad \square$$

5.4.2. Derivada de la función recíproca

Supongamos que f tiene inversa f^{-1} y que ambas son derivables. Derivando ambos miembros de la igualdad $f(f^{-1}(x)) = x$, obtenemos que

$$f'(f^{-1}(x))\frac{d}{dx}[f^{-1}(x)] = 1.$$

Despejando la derivada de la función recíproca, se consigue que

$$\frac{d}{dx}[f^{-1}(x)] = \frac{1}{f'(f^{-1}(x))}.$$

Se puede ser más preciso y lo único que se necesita para que exista la inversa y sea verdad la fórmula anterior es que la derivada en el punto no sea nula

Teorema 5.7. Sea f una función que tiene una función recíproca f^{-1} . Si f es continua en su dominio, entonces f^{-1} es continua.

Si f es creciente en su dominio, entonces f^{-1} es creciente.

Si f es decreciente en su dominio, entonces f^{-1} es decreciente.

Si f es derivable en c y $f'(c) \neq 0$, entonces f^{-1} es derivable en $f(c)$ y

$$\frac{d}{dx}[f^{-1}](a) = \frac{1}{f'(f^{-1}(a))}.$$

¹Ver apéndices del libro Cálculo I con autores Larson, Hostetler & Edwards

Ejemplo 5.10. La función $g(x) = \sqrt{x} = x^{1/2}$ es derivable en $(0, +\infty)$ y su derivada es $g'(x) = \frac{1}{2}x^{-1/2}$. La raíz cuadrada es la función recíproca de $f(x) = x^2$ en $(0, +\infty)$, como $f'(x) = 2x$ no se anula en $(0, +\infty)$ la función es derivable en esos puntos y

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

5.5. Derivada de la función potencia

5.5.1. Números binómicos y binomio de Newton

Si desarrollamos la expresión $(a+b)^n$ y denotamos al coeficiente del término $a^j b^{n-j}$ por $\binom{n}{j}$, tenemos que

$$(a + b)^n = \sum_{j=0}^n \binom{n}{j} a^j b^{n-j}.$$

Los coeficientes $\binom{n}{j}$ son las formas de escoger j elementos entre n y se pueden calcular con la fórmula

$$\binom{n}{j} = \frac{n!}{j!(n-j)!} = \frac{n(n-1)\cdots(n-m+1)}{m!}.$$

Se verifican las siguientes propiedades:

Teorema 5.8. Sean $n = 1, 2, \dots$ y $m = 0, 1, 2, \dots, n$. Se verifica:

1. $\binom{n}{0} = 1 = \binom{n}{n}$ y $\binom{n}{1} = n$.
2. $\binom{n}{m} = \binom{n}{n-m}$.
3. $\binom{n}{m} = \binom{n-1}{m-1} + \binom{n-1}{m}$ para $m = 1, \dots, n-1$.
4. $\binom{n}{m} = \frac{n}{m} \binom{n-1}{m-1}$.

La tercera propiedad es importante porque permite construir lo que se conoce como triángulo de Tartaglia mediante sumas de los dos elementos más próximos de la fila superior.

The figure consists of two triangular diagrams. The left triangle represents the expansion of $(a+b)^6$. It has 7 rows. The top row has one entry: $\binom{6}{0}$. The second row has two entries: $\binom{6}{0}$ and $\binom{6}{1}$. The third row has three entries: $\binom{6}{0}$, $\binom{6}{1}$, and $\binom{6}{2}$. The fourth row has four entries: $\binom{6}{0}$, $\binom{6}{1}$, $\binom{6}{2}$, and $\binom{6}{3}$. The fifth row has five entries: $\binom{6}{0}$, $\binom{6}{1}$, $\binom{6}{2}$, $\binom{6}{3}$, and $\binom{6}{4}$. The sixth row has six entries: $\binom{6}{0}$, $\binom{6}{1}$, $\binom{6}{2}$, $\binom{6}{3}$, $\binom{6}{4}$, and $\binom{6}{5}$. The seventh row has seven entries: $\binom{6}{0}$, $\binom{6}{1}$, $\binom{6}{2}$, $\binom{6}{3}$, $\binom{6}{4}$, $\binom{6}{5}$, and $\binom{6}{6}$. The right triangle represents the expansion of $(x+y)^6$. It also has 7 rows. The top row has one entry: 1. The second row has two entries: 1 and 1. The third row has three entries: 1, 2, and 1. The fourth row has four entries: 1, 3, 3, and 1. The fifth row has five entries: 1, 4, 6, 4, and 1. The sixth row has six entries: 1, 5, 10, 10, 5, and 1. The seventh row has seven entries: 1, 6, 15, 20, 15, 6, and 1. Below the right triangle is a dotted line.

Usaremos la siguiente generalización de número binómico: se define el *número binómico* de un número real α sobre un natural n como

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-(n-1))}{n!} \quad \text{para } n \geq 1 \quad \text{y} \quad \binom{\alpha}{0} = 1.$$

5.5.2. derivada de la potencia

Teorema 5.9. La función $f(x) = x^q$ con q racional es derivable en su dominio (salvo en $x = 0$ si $q < 1$) y su derivada es $f'(x) = q x^{q-1}$

Demostración. Primero veamos que la función $f(x) = x^n$ con $n \in \mathbb{N}$ es derivable en todo punto y $f'(x) = n x^{n-1}$, ya que

$$f'(a) = \lim_{h \rightarrow 0} \frac{(a+h)^n - a^n}{h} = \lim_{h \rightarrow 0} \frac{\sum_{j=0}^n \binom{n}{j} a^j h^{n-j} - a^n}{h} = \lim_{h \rightarrow 0} \frac{a^n + n h a^{n-1} + \dots + h^n - a^n}{h} = n a^{n-1}.$$

Por el teorema 5.7 la derivada de $f(x) = x^{1/n}$ (que es la recíproca de x^n) es $f'(x) = \frac{1}{n} x^{1/n-1}$ y, por último, la derivada de $x^{n/m}$ se calcula empleando la regla de la cadena (teorema 5.4). $\square$

Teorema 5.10. Si la función f definida por la serie $f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ con radio de convergencia $R > 0$, entonces f es derivable en el intervalo $(x_0 - R, x_0 + R)$. Se verifica que:

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1}.$$

Ejemplo 5.11. El precio de una obligación a 3 años donde la entidad emisora tiene que pagar 1€ está relacionado con el tipo de interés a ese mismo plazo, por la expresión:

$$P = \frac{1}{(1+r)^3}. \quad (5.3)$$

Encuentra el porcentaje de cambio en P para pequeñas variaciones de r y usa el resultado para estimar el cambio de precio porcentual que experimenta este título financiero si el tipo de interés desciende 25 puntos básicos cuando el tipo actual es de $r = 2\%$ (de $r = 2\%$ a $r = 1.75\%$).

Podemos basar nuestro razonamiento aproximando la diferencial de la expresión (5.3) a pequeños incrementos.

Notemos que $\Delta P \approx P'(r_0) \Delta r = -3(1+r_0)^{-4} \Delta r$ y por lo que en la proximidad del precio y del tipo de interés inicial (antes de la bajada brusca de tipos de interés) se puede hacer la estimación del cambio porcentual que se va a producir en precio:

$$\Delta P \approx P'(r_0) \Delta r = -3P(r_0) \frac{r - r_0}{1 + r_0},$$

de donde se obtiene que

$$\% \text{cambio en } P = 100 \frac{\Delta P}{P(r_0)} = -300 \frac{\Delta r}{1 + r_0}.$$

Sustituyendo para nuestro caso concreto $\Delta r = -0.0025$ y $1 + r_0 = 1.02$, se tiene que

$$\% \text{cambio en } P = -300 \frac{-0.0025}{1.02} \approx 0.75\%.$$

5.6. Derivadas de ordenes superiores

Las derivadas de orden superior son definidas por la aplicación sucesiva de la diferenciación con respecto a x . La derivada segunda de f respecto de x es definida por

$$\frac{d^2 f}{dx^2} = \frac{d}{dx} \left[\frac{df}{dx} \right].$$

La derivada segunda se representa también por $f''(x)$.

De forma recursiva, se define la derivada n -ésima como

$$\frac{d^n f}{dx^n} = \frac{d}{dx} \left[\frac{d^{n-1} f}{dx^{n-1}} \right].$$

También usaremos la notación $f^{(n)}(x)$ para representar la función $\frac{d^n f}{dx^n}$.

Por $f^{(0)}$ o $\frac{d^0 f}{dx^0}$ se representa a la función f sin derivar. Esta notación será útil en diversas fórmulas.

Ejemplo 5.12. Dada $f(x) = 3x^2 - \frac{2}{x}$. Calcula $f^{(1)}(2) = f'(2)$ y $f^{(2)}(3) = f''(3)$.
Tenemos que $f'(x) = 6x + 2x^{-2}$ y $f''(x) = 6 - 4x^{-3}$. Sustituyendo tenemos que

$$f'(2) = 12 + \frac{2}{4} = \frac{25}{2} \quad \text{y} \quad f''(3) = 6 - 4\left(\frac{1}{27}\right) = \frac{158}{27}.$$

Ejemplo 5.13. Derivada n -ésima de la potencia $f(x) = (x+a)^\tau$ con τ un número racional. Calculamos varias derivadas de f ,

$$f'(x) = \tau(x+a)^{\tau-1}, \quad f''(x) = \tau(\tau-1)(x+a)^{\tau-2}, \quad f'''(x) = \tau(\tau-1)(\tau-2)(x+a)^{\tau-3}$$

y así, sucesivamente, observamos que la derivada n -ésima, $f^{(n)}(x)$, es la potencia $(x+a)^{\tau-n}$ multiplicado por el producto $\tau(\tau-1) \cdots (\tau-(n-1))$. Con lo que, por la definición de número binómico, se tiene que

$$\frac{d^n}{dx^n} [(x+a)^\tau] = n! \binom{\tau}{n} (x+a)^{\tau-n}. \quad (5.4)$$

Ejemplo 5.14. Calculemos la derivada n -ésima de $f(x) = \frac{x^2 + x - 3}{x^3 + x^2 - 4x - 4}$.

En el ejemplo 3.10 encontramos que la descomposición en fracciones simples es

$$f(x) = \frac{1}{4} \frac{1}{x-2} - \frac{1}{4} \frac{1}{x+2} + \frac{1}{x+1}.$$

Aplicando la fórmula (5.4), obtenemos

$$f^{(n)}(x) = \frac{n!}{4} \binom{-1}{n} (x-2)^{-1-n} - \frac{n!}{4} \binom{-1}{n} (x+2)^{-1-n} + n! \binom{-1}{n} (x+1)^{-1-n}.$$

Ejemplo 5.15. Calcular la derivada n -ésima de $f(x) = \frac{x^2 + x + 3}{x^3 + 3x^2 - 4}$.

Por el ejemplo 3.11 tenemos que la descomposición en fracciones simples es

$$f(x) = \frac{5}{9} \frac{1}{x-1} + \frac{4}{9} \frac{1}{x+2} - \frac{5}{3} \frac{1}{(x+2)^2}.$$

De nuevo aplicando la fórmula (5.4), obtenemos

$$f^{(n)}(x) = n! \frac{5}{9} \binom{-1}{n} (x-1)^{-1-n} + n! \frac{4}{9} \binom{-1}{n} (x+2)^{-1-n} - n! \frac{5}{3} \binom{-2}{n} (x+2)^{-2-n}.$$

5.6.1. Representación de funciones por series de potencias

A la serie de potencia $\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x-x_0)^n$ se la llama *expansión en serie de Taylor* de $f(x)$ alrededor del punto x_0 . Cuando $x_0 = 0$ se le llama *expansión en serie de MacLaurin* y se escribe como

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} x^n.$$

Aunque en los ejemplos más habituales la expansión en serie de MacLaurin coincide con la función, existen casos donde esto no sucede.

Ejemplo 5.16. Calcular la expansión en serie de MacLaurin para $f(x) = (1+x)^\tau$ con τ un número racional.

Por la fórmula (5.4) tenemos que $f^{(n)}(x) = n! \binom{\tau}{n} (1+x)^{\tau-n}$. Por tanto, $f^{(n)}(0) = n! \binom{\tau}{n}$ y la expansión en serie de MacLaurin es

$$\sum_{n=0}^{\infty} \frac{n! \binom{\tau}{n}}{n!} x^n = \sum_{n=0}^{\infty} \binom{\tau}{n} x^n \quad (5.5)$$

En este ejemplo es posible probar (aunque se trata de un resultado técnico) que el desarrollo de MacLaurin coincide con la función

Teorema 5.11. Si x verifica que $-1 < x < 1$, entonces $(1+x)^\tau = \sum_{n=0}^{\infty} \binom{\tau}{n} x^n$.

5.7. Elasticidad de la demanda

Supongamos dos magnitudes económicas x e y relacionadas por una función f , $y = f(x)$. Con el concepto de elasticidad se informa de la forma que una variación porcentual de la variable x produce una variación porcentual en la variable y . Por eso *elasticidad* de f respecto de x (se denota $El_x f(x)$ ó σ_{yx}) se define como

$$El_x f(x) = \frac{x}{f(x)} f'(x).$$

Ejemplo 5.17. Sea D la demanda y Q la producción de un bien. Si están relacionados por $D = 5000Q^{-1.7}$. Calcular la elasticidad.

Derivemos D respecto de Q ,

$$\frac{dD}{dQ} = 5000(-1.7)Q^{-2.7} = -8500Q^{-2.7}$$

con lo que la elasticidad es

$$\sigma_{DQ} = \frac{Q}{D} \frac{dD}{dQ} = \frac{Q}{5000Q^{-1.7}} 8500Q^{-2.7} = -1.7$$

Que la elasticidad sea constante a -1.7 significa que si la producción aumenta un 1 %, entonces la demanda disminuye un 1.7 %.

En un tema posterior veremos que las únicas funciones con elasticidad constante son las potencias multiplicadas por un número.

5.8. Regla de Leibnitz

Es una generalización de la regla del producto.

Si calculamos la derivada segunda del producto:

$$\frac{d^2}{dx^2}[fg] = f \frac{d^2}{dx^2}[g] + 2 \frac{d}{dx}[f] \frac{d}{dx}[g] + \frac{d^2}{dx^2}[f]g$$

y la derivada tercera es

$$\frac{d^3}{dx^3}[fg] = f \frac{d^3}{dx^3}[g] + 3 \frac{d}{dx}[f] \frac{d^2}{dx^2}[g] + 3 \frac{d^2}{dx^2}[f] \frac{d}{dx}[g] + \frac{d^3}{dx^3}[f]g.$$

Continuando el proceso descubrimos que los coeficientes que aparecen son los números binómicos y se puede probar por inducción la siguiente fórmula:

$$\frac{d^n}{dx^n}[fg] = \binom{n}{0} f \frac{d^n g}{dx^n} + \binom{n}{1} \frac{df}{dx} \frac{d^{n-1} g}{dx^{n-1}} + \binom{n}{2} \frac{d^2 f}{dx^2} \frac{d^{n-2} g}{dx^{n-2}} + \cdots + \binom{n}{n-1} \frac{d^{n-1} f}{dx^{n-1}} \frac{dg}{dx} + \binom{n}{n} \frac{d^n f}{dx^n} g.$$

Teorema 5.12. Si f y g son n veces derivables, el producto fg es también n veces derivable y viene dado por la fórmula:

$$\frac{d^n}{dx^n}[fg] = \sum_{k=0}^n \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x). \quad (5.6)$$

Ejemplo 5.18. Relaciona la derivada n -ésima de $f(x) = \frac{1}{1+x^2}$ con las derivadas sucesivas anteriores y calcule la derivada segunda con dicha fórmula

Tenemos la expresión $f(x)(1+x^2) = 1$. La función $g(x) = 1+x^2$ verifica $g'(x) = 2x$, $g''(x) = 2$ y $g^{(n)}(x) = 0$ con $n \geq 3$. Aplicando la fórmula de Leibnitz al producto $f(x)g(x)$

$$\begin{aligned} \frac{d^n}{dx^n}(f(x)g(x)) &= \binom{n}{0} f^{(n)}(x)g(x) + \binom{n}{1} f^{(n-1)}(x)g'(x) + \cdots + \binom{n}{2} f^{(n-2)}(x)g''(x) = \\ &= (1+x^2)f^{(n)}(x) + n2xf^{(n-1)}(x) + \frac{n(n-1)}{2} 2f^{(n-2)}(x) = \\ &= (1+x^2)f^{(n)}(x) + n2xf^{(n-1)}(x) + n(n-1)f^{(n-2)}(x). \end{aligned}$$

Como $f(x)(1+x^2) = 1$ y la derivada n -ésima del producto $f(x)g(x)$ es cero, tenemos que

$$(1+x^2)f^{(n)}(x) + n2xf^{(n-1)}(x) + n(n-1)f^{(n-2)}(x) = 0.$$

Despejando tenemos $f^{(n)}(x)$ en función de $f^{(n-1)}(x)$ y $f^{(n-2)}(x)$

$$f^{(n)}(x) = -\frac{n2x}{1+x^2} f^{(n-1)}(x) - \frac{n(n-1)}{1+x^2} f^{(n-2)}(x).$$

Como $f(x) = \frac{1}{1+x^2}$ y $f'(x) = -\frac{2x}{(1+x^2)^2}$. La derivada segunda, aplicando la fórmula anterior, es

$$f''(x) = -\frac{4x}{1+x^2} f'(x) - \frac{2}{1+x^2} f(x) = -\frac{4x}{1+x^2} \left(-\frac{2x}{(1+x^2)^2} \right) - \frac{2}{1+x^2} \frac{1}{1+x^2} = \frac{6x^2 - 2}{(1+x^2)^3}.$$

Capítulo 6

Consecuencias de la continuidad y la derivabilidad.

6.1. Extremos relativos y absolutos. Teorema de Weierstrass.

La función f se dice que alcanza un *máximo absoluto* en $x = c$ si $f(c) \geq f(x)$ para todo x en el dominio de f y al valor $f(c)$ se le llama *máximo absoluto* de la función f . La función f se dice que alcanza un *mínimo absoluto* en $x = c$ si $f(c) \leq f(x)$ para todo x en el dominio de f y al valor $f(c)$ se le llama *mínimo absoluto* de la función f . Frecuentemente omitiremos el adjetivo absoluto, es decir, simplemente se dice máximo o mínimo.

Los máximos y mínimos de una función se llaman *extremos* de la función. Cuando decimos que en $x = c$ la función tiene un extremo queremos decir que en ese punto se alcanza un máximo o un mínimo, pero no especificamos de cuál de los dos casos se trata.

Teorema 6.1 (Teorema de Weierstrass). Una función f continua en un intervalo cerrado $[a, b]$ alcanza su máximo y su mínimo en algún punto del intervalo.

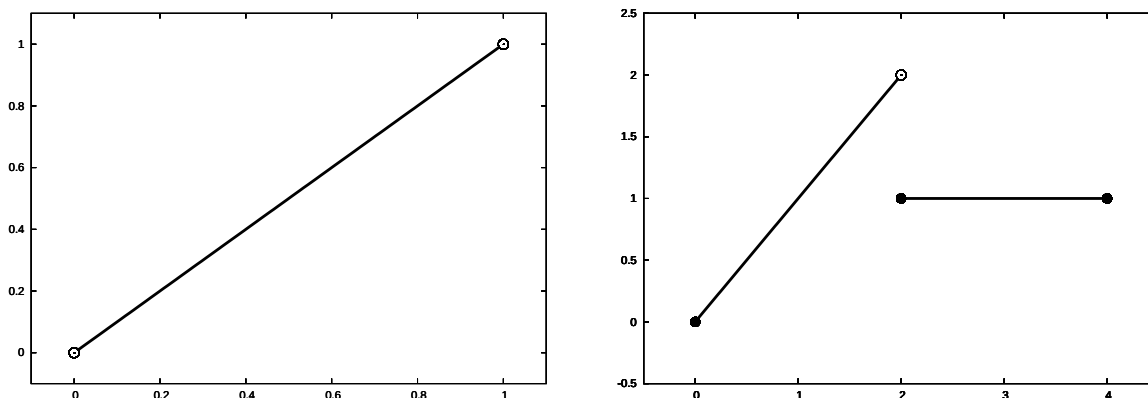


Figura 6.1: La función $f(x) = x$ en $(0, 1)$ no alcanza ni máximo ni mínimo como se observa en el panel izquierdo. A dicha función f no se le puede aplicar el teorema de Weierstrass porque no está definida en un intervalo cerrado. La función definida por trozos $g(x) = x$ si $0 \leq x < 2$ y $g(x) = 1$ si $2 \leq x \leq 4$ no se le puede aplicar el teorema de Weierstrass porque no es continua y se observa en el panel derecho que dicha función no alcanza un máximo absoluto en ninguno de sus puntos.

La función f se dice que tiene un *máximo local* o *relativo* en $x = c$ si existe un intervalo abierto I al cual pertenece c tal que $f(x) \leq f(c)$ para todo x en I . Se dice que tiene un *mínimo relativo* en $x = c$ si existe un intervalo abierto I al cual pertenece c tal que $f(x) \geq f(c)$ para todo x en I .

Decimos que la función f tiene un *punto crítico* en $x = c$ si $f'(c) = 0$ o f no es diferenciable en $x = c$. Es decir, los puntos no críticos son aquellos que tienen una recta tangente no horizontal.

Teorema 6.2. Si la función f definida en el intervalo abierto (a, b) es diferenciable en $x = c$ con $a < c < b$ y tiene un extremo relativo en $x = c$, entonces $f'(c) = 0$.

Demostración. Supongamos que en $x = c$ se alcanza un mínimo. Si $h < 0$, entonces $(f(c+h) - f(c))/h$ es negativo, luego $f'(c) \leq 0$. Por el otro lado, si $h > 0$, entonces el cociente $(f(c+h) - f(c))/h$ es positivo, luego $f'(c) \geq 0$. De $f'(c) \geq 0$ y $f'(c) \leq 0$, se deduce que $f'(c) = 0$. $\square$

Una consecuencia del anterior teorema es:

Teorema 6.3. Si la función f definida en un intervalo (a, b) tiene un máximo relativo o un mínimo relativo en $x = c$ (con $a < c < b$), entonces c es un punto crítico de la función f .

El teorema 6.3 junto al teorema de Weierstrass (teorema 6.1) implica que los extremos absolutos de una función definida en un intervalo cerrado $[a, b]$ sólo pueden ser alcanzados en:

- Los extremos del intervalo cerrado, $x = a$ ó $x = b$;
- Los puntos críticos.

Por ello al estudiar los extremos absolutos de una función en un intervalo cerrado $[a, b]$ se empieza localizando los puntos críticos y, luego, el máximo absoluto es la mayor de las evaluaciones de la función en la lista anterior (formada por puntos críticos y extremos del intervalo) y el mínimo la menor de las evaluaciones de dicha lista.

Ejemplo 6.1. Calcula los extremos de la función $f(x) = x^3 - 6x^2 + 9x + 7$ en el intervalo $[0, 5]$.

Señalamos que como el enunciado pide localizar extremos y no especifica si son absolutos o relativos, debemos entender que nos piden localizar los extremos absolutos.

Como la función es diferenciable en el intervalo $(0, 5)$. Los posibles máximos y mínimos se tienen que alcanzar en los extremos del intervalo o donde se anule la derivada. Al ser $f'(x) = 3x^2 - 12x + 9$, los puntos donde se anula son los que verifican $x^2 - 4x + 3 = 0$ son $x = 1$ y $x = 3$. Por tanto, el máximo y el mínimo se encuentran en la siguiente lista: $x = 0$, $x = 1$, $x = 3$ y $x = 5$. Evaluemos la función en estos valores:

$$\begin{aligned} f(0) &= 7, & f(1) &= 1 - 6 \cdot 1^2 + 9 \cdot 1 + 7 = 11, \\ f(3) &= 3^3 - 6 \cdot 3^2 + 9 \cdot 3 + 7 = 7, & f(5) &= 5^3 - 6 \cdot 5^2 + 9 \cdot 5 + 7 = 27. \end{aligned}$$

Por tanto, el mínimo se alcanza simultáneamente en $x = 0$ y $x = 3$, mientras que el máximo se alcanza en $x = 5$.

Ejemplo 6.2. Calcular los extremos absolutos de $f(x) = \sqrt{x} + |x - 1|$ en el intervalo $[0, 2]$.

La función se escribe por trozos como

$$f(x) = \begin{cases} \sqrt{x} + 1 - x & \text{si } 0 \leq x < 1, \\ \sqrt{x} + x - 1 & \text{si } 1 \leq x \leq 2. \end{cases}$$

En los intervalos $(0, 1)$ y $(1, 2)$ está definida por funciones que son derivables, así tenemos que

$$f'(x) = \begin{cases} \frac{1}{2}x^{-1/2} - 1 & \text{si } 0 < x < 1, \\ \frac{1}{2}x^{-1/2} + 1 & \text{si } 1 < x \leq 2. \end{cases}$$

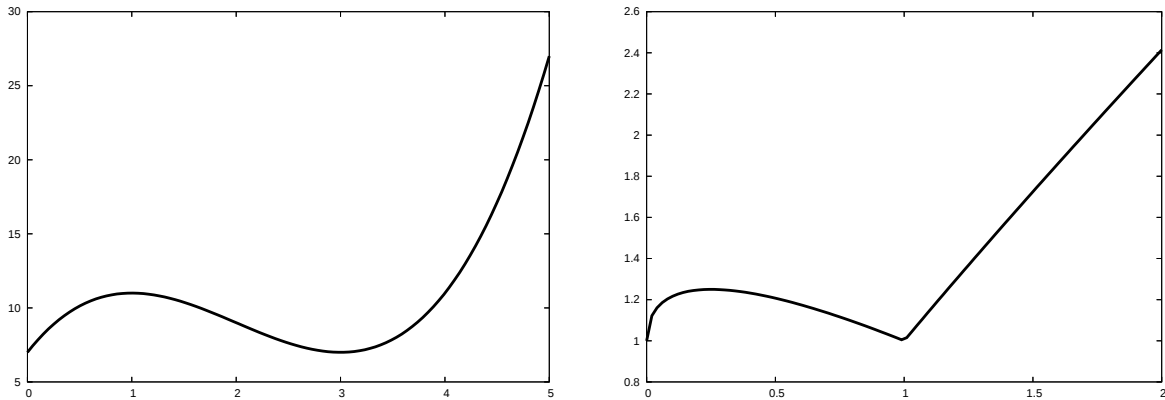


Figura 6.2: En el panel de la izquierda gráfica de la función $f(x) = x^3 - 6x^2 + 9x + 7$ definida en el intervalo $[0, 5]$ y el panel de la derecha $f(x) = \sqrt{x} + |x - 1|$ definida en el intervalo $[0, 2]$.

En $x = 1$ las derivadas laterales son diferentes y, por tanto, tenemos que la función no es derivable en dicho punto. Veamos los puntos con derivada nula, la expresión $\frac{1}{2}x^{-1/2} - 1$ sólo se anula en $x = \frac{1}{4}$; mientras que $\frac{1}{2}x^{-1/2} + 1$ no se anula en ningún punto.

Tenemos que los posibles puntos donde se alcanzan los extremos absolutos son: puntos donde la función no es derivable ($x = 1$), puntos con derivada nula ($x = 1/4$) y extremos del intervalo de definición ($x = 0$ y $x = 2$). Evaluemos la función en dichos puntos:

$$\begin{aligned} f(0) &= \sqrt{0} + |0 - 1| = 1 & f\left(\frac{1}{4}\right) &= \sqrt{\frac{1}{4}} + \left|\frac{1}{4} - 1\right| = \frac{5}{4} = 1.25 \\ f(1) &= \sqrt{1} + |1 - 1| = 1 & f(2) &= \sqrt{2} + |2 - 1| = \sqrt{2} + 1 \approx 2.41 \end{aligned}$$

Luego el máximo se alcanza en $x = 2$ y el mínimo en $x = 0$ y $x = 1$.

6.2. Teorema de Rolle. Teorema del valor medio. Monotonía de una función.

6.2.1. Teorema de Rolle

Teorema 6.4 (Teorema de Rolle). Sea f una función continua en un intervalo cerrado $[a, b]$, diferenciable en el intervalo abierto (a, b) y tal que $f(a) = f(b)$, entonces existe al menos un número c con $a < c < b$ tal que $f'(c) = 0$.

Demostración. Una función continua en un intervalo cerrado $[a, b]$ alcanza un máximo y un mínimo por el teorema de Weierstrass (teorema 6.1).

Si el máximo y el mínimo coinciden, entonces la función es constante $f(x) = f(a)$ para todo x en el intervalo $[a, b]$ y la derivada de f se anula en todos los puntos del intervalo.

Si el máximo y el mínimo no coinciden, entonces alguno de ellos es diferente del valor $f(a) = f(b)$. Si este extremo se alcanza en $x = c$ con $a < c < b$, al ser la función f diferenciable en $x = c$, tenemos que en $x = c$ la derivada es nula, $f'(c) = 0$. $\square$

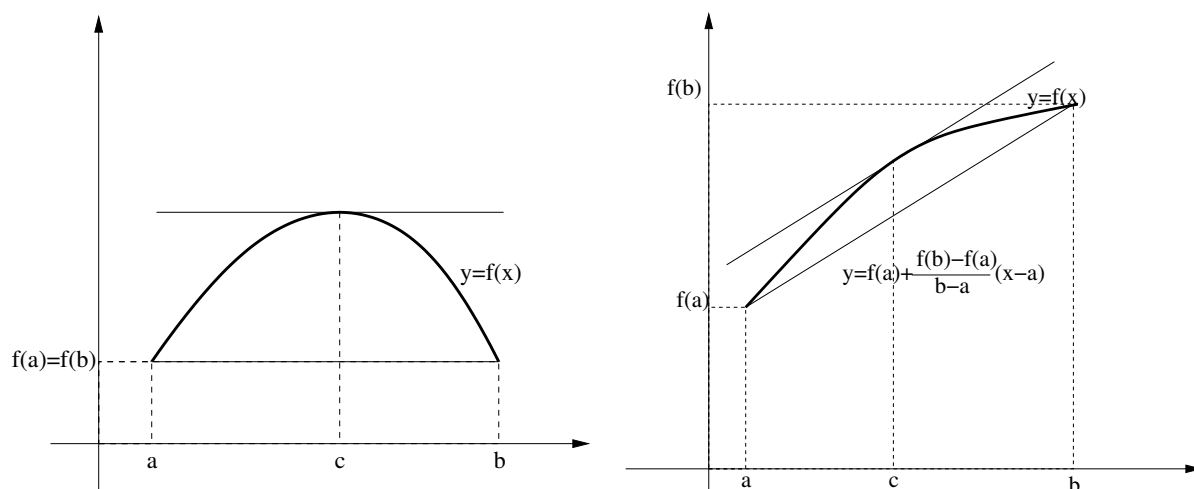


Figura 6.3: Interpretación geométrica de los teoremas de Rolle y del valor medio. En el panel izquierdo: La interpretación geométrica del teorema de Rolle es simple porque si tenemos una función continua en $[a, b]$ y diferenciable en (a, b) (es decir, no se forman “picos”) con $f(a) = f(b)$, entonces hay un punto en el que la tangente $y = f'(x)$ debe ser horizontal. En el panel derecho: La interpretación geométrica del teorema del valor medio es que en algún punto de la gráfica $y = f(x)$ su recta tangente es paralela a la recta que une a los puntos $(a, f(a))$ y $(b, f(b))$.

6.2.2. Teorema del valor medio para la derivada

Teorema 6.5 (Teorema del valor medio). Sea f una función continua en el intervalo cerrado $[a, b]$ y diferenciable en el intervalo abierto (a, b) , entonces existe c tal que $a < c < b$ tal que

$$f(b) - f(a) = f'(c)(b - a)$$

Demostración. Definimos la función

$$F(x) = (f(x) - f(a))(b - a) - (f(b) - f(a))(x - a).$$

Por las propiedades de las funciones diferenciables y continuas tenemos que $F(x)$ es continua en el intervalo $[a, b]$ y es diferenciable en el intervalo (a, b) . Como $F(a) = F(b) = 0$, el teorema de Rolle afirma que su derivada

$$F'(x) = f'(x)(b - a) - (f(b) - f(a))$$

debe anularse para al menos un c con $a < c < b$. Para dicho c se verifica $f(b) - f(a) = f'(c)(b - a)$. $\square$

6.2.3. Funciones crecientes y decrecientes

Se dice que f es *creciente* en un intervalo I si para todo x_1 y x_2 con $x_1 < x_2$ se implica que $f(x_1) \leq f(x_2)$. Se dice que f es *estrictamente creciente* en un intervalo I si para todo x_1 y x_2 con $x_1 < x_2$ se implica que $f(x_1) < f(x_2)$. Definiciones análogas para decreciente y estrictamente decreciente.

Teorema 6.6. Sea $f(x)$ una función continua en $[a, b]$ y diferenciable en (a, b) .

Si la derivada de f es positiva ($f'(x) > 0$) en todo punto del intervalo (a, b) , entonces f es estrictamente creciente en $[a, b]$.

Si la derivada de f es negativa ($f'(x) < 0$) en todo punto del intervalo (a, b) , entonces f es estrictamente decreciente en $[a, b]$.

Si la derivada de f es nula ($f'(x) = 0$) en todo punto del intervalo (a, b) , entonces f es constante en $[a, b]$.

Demostración. Consideremos $a \leq x < y \leq b$. Por el teorema del valor medio existe un c tal que $x < c < y$ tal que

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

Si la derivada de f es positiva, entonces $f'(c) > 0$ y $f(y) - f(x)$ es positivo, por lo que f es estrictamente creciente en $[a, b]$. Para derivada de f negativa razonamiento análogo. $\square$

Para ver si en un punto crítico, $x = c$, se alcanza un máximo o un mínimo relativo se utiliza el *criterio de la primera derivada* donde hay que estudiar el signo de la derivada en sus puntos cercanos. Así,

- Si $f'(x) > 0$ en la inmediata izquierda de $x = c$ y $f'(x) < 0$ en la inmediata derecha de $x = c$, entonces la función f alcanza un máximo relativo en $x = c$.
- Si $f'(x) < 0$ en la inmediata izquierda de $x = c$ y $f'(x) > 0$ en la inmediata derecha de $x = c$, entonces la función f alcanza un mínimo relativo en $x = c$.
- Si $f'(x)$ tiene el mismo signo a la inmediata izquierda y a la inmediata derecha, entonces f no alcanza un extremo relativo en $x = c$.

Ejemplo 6.3. Estudiar los extremos relativos de la función $f(x) = x(x+1)^{2/3}$.

La función $x^{2/3}$ no es derivable en $x = 0$. Así, la función $f(x)$ dada no es derivable en $x = -1$. Salvo en dicho punto, la expresión de la derivada es:

$$f'(x) = (x+1)^{2/3} + x \frac{2}{3}(x+1)^{-1/3} = (x+1)^{2/3} \left(1 + \frac{2}{3} \frac{x}{x+1}\right) = (x+1)^{2/3} \frac{5x+3}{3x+3}.$$

Como $(x+1)^{2/3}$ es positiva, el signo depende del otro factor. El estudio del signo nos indicaría que la derivada es positiva (función estrictamente creciente) en los intervalos $(-\infty, -1)$ y $(-3/5, +\infty)$ y la función tiene derivada negativa (función estrictamente decreciente) en el intervalo $(-1, -3/5)$. Así, en $x = -1$ la función pasa de creciente a decreciente, luego se trata de un máximo relativo y en $x = -3/5$ la función pasa de decreciente a creciente, con lo que se trata de un mínimo relativo.

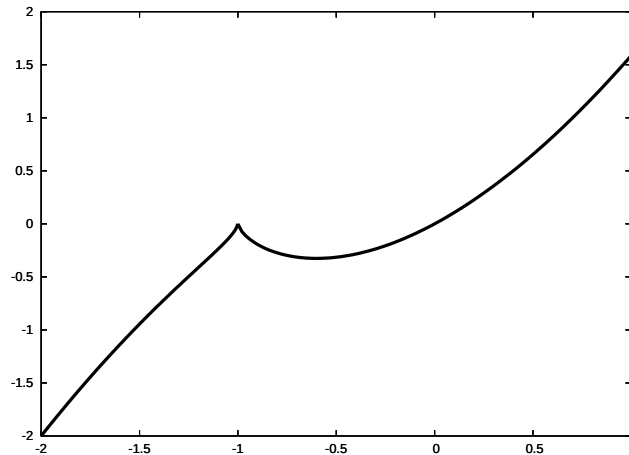


Figura 6.4: Obsérvese que en $x = -1$ la función no es derivable, pero en ese punto tiene un máximo relativo.

6.3. Regla de L'Hôpital.

Teorema 6.7 (Regla de L'Hôpital). Sean $f(x)$ y $g(x)$ funciones diferenciables en un intervalo que contiene el punto a tales que $f(a) = g(a) = 0$, hay un entorno de $x = a$ donde la función g no se anula y existe el límite cuando x tiende a a para el cociente $f'(x)/g'(x)$. Entonces,

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

Demostración. Sean $f(x)$ y $g(x)$ funciones diferenciables en un entorno de $x = a$ tales que $f(a) = g(a) = 0$. Si aplicamos el teorema del valor medio en el intervalo $[a, x]$ para f y g tenemos que existe dos puntos c_f y c_g tales que

$$f(x) = f(a) + f'(c_f)(x - a) = f'(c_f)(x - a) \quad g(x) = g(a) + g'(c_g)(x - a) = g'(c_g)(x - a)$$

Tenemos la siguiente igualdad entre los cocientes

$$f(x)/g(x) = f'(c_f)/g'(c_g).$$

Como $a < c_f, c_g < x$, se tiene que cuando x tiende a a , los puntos c_f y c_g también tienden a a . Consiguientemente, los límites de los cocientes $f(x)/g(x)$ y $f'(x)/g'(x)$ son los mismos. $\square$

Ejemplo 6.4. Evalúa $\lim_{x \rightarrow 1} \frac{x^3 + 2x^2 - x - 2}{x^2 - 1}$

Consideremos las funciones $f(x) = x^3 + 2x^2 - x - 2$ y $g(x) = x^2 - 1$ son diferenciables en todo valor real, $f(1) = 1^3 + 2 \cdot 1^2 - 1 - 2 = 0$ y $g(1) = 1^2 - 1 = 0$. Así, tenemos una indeterminación $0/0$. Veamos el límite del cociente $f'(x)/g'(x)$. Como $f'(x) = 3x^2 + 4x - 1$ y $g'(x) = 2x$, tenemos que

$$\lim_{x \rightarrow 1} \frac{x^3 + 2x^2 - x - 2}{x^2 - 1} = \lim_{x \rightarrow 1} \frac{3x^2 + 4x - 1}{2x} = \frac{6}{2} = 3.$$

Ejemplo 6.5. Evalúa $\lim_{x \rightarrow 0} \frac{\sqrt{1+4x} - 1}{3x}$.

Consideremos la función $f(x) = \sqrt{1+4x} - 1 = (1+4x)^{1/2} - 1$ y la función $g(x) = 3x$. Ambas funciones son diferenciables en todo $\mathbb{R}$ y $f(0) = g(0) = 0$. Así, tenemos una indeterminación $0/0$. Calculamos la derivada

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

y la derivada de $g(x)$ es $g'(x) = 3$, con lo que el límite del cociente es

$$\lim_{x \rightarrow 0} \frac{\sqrt{1+4x} - 1}{3x} = \lim_{x \rightarrow 0} \frac{2(1+4x)^{-1/2}}{3} = \frac{2}{3}.$$

Ejemplo 6.6. Calcular $\lim_{x \rightarrow 1} \frac{x^{1/3} - 1}{x^2 - 3x + 2}$.

Lamemos $f(x) = x^{1/3} - 1$ y $g(x) = x^2 - 3x + 2$. Como $f(1) = g(1) = 0$, tenemos una indeterminación del tipo $0/0$ y al ser derivables en un entorno de $x = 1$ podemos aplicar la regla de L'Hôpital

$$\lim_{x \rightarrow 1} \frac{f'(x)}{g'(x)} = \lim_{x \rightarrow 1} \frac{(1/3)x^{-2/3}}{2x - 3} = \frac{1/3}{2 - 3} = -\frac{1}{3}$$

Por tanto, $\lim_{x \rightarrow 1} \frac{f(x)}{g(x)} = -\frac{1}{3}$.

6.4. Representación gráfica de funciones. Estudio cualitativo.

Seguidamente recomendamos las etapas a seguir para determinar el comportamiento cualitativo de una función derivable cualquiera.

1. Obtención del *dominio de la función* $f(x)$.

2. *Cálculo de los puntos de corte con los ejes.*

Corte con el eje OX: se hace $y = 0$ en la ecuación de definición de la función $y = f(x)$, equivale a obtener las raíces de la ecuación $f(x) = 0$.

Corte con el eje OY: se hace $x = 0$ y se calcula el correspondiente valor $y = f(0)$

3. *Signo de la función.* Analizar el signo de la función en los intervalos que delimitan los puntos de corte sobre el eje de abscisas y las discontinuidades de la función. Si la función f es positiva en un intervalo $I \subseteq \mathbb{R}$ dibujamos la gráfica por encima del eje de abscisa. Mientras que si f es negativa en un intervalo $I \subseteq \mathbb{R}$ dibujamos la gráfica por debajo del eje abscisa.

4. Análisis de la existencia (o no) de simetrías

Simetría par. Si $f(x) = f(-x)$ para todo x en el dominio

Simetría impar. Si $f(x) = -f(-x)$ para todo x en el dominio.

5. Búsqueda de asíntotas:

Verticales. Si $\lim_{x \rightarrow a} f(x) = \infty$, entonces la recta $x = a$ es una asíntota.

Horizontales. Si $\lim_{x \rightarrow \pm\infty} f(x) = b$, entonces la recta $y = b$ es una asíntota.

Oblicuas. Si la función $f(x)$ verifica:

$$m = \lim_{x \rightarrow \pm\infty} \frac{f(x)}{x}, \quad n = \lim_{x \rightarrow \pm\infty} (f(x) - mx),$$

entonces $y = mx + n$ es una asíntota oblicua.

6. Cálculo de la primera derivada.

Puntos críticos Valores de x tales que la derivada $f'(x) = 0$ o que la función no es derivable.

Este punto es un

Mínimo relativo: Los puntos críticos $(x, f(x))$ cuando se cumpla que $f'(x - h) < 0$ y, simultáneamente, $f'(x + h) > 0$ ($x \pm h$ en la cercanía del punto estacionario).

Máximo relativo: Los puntos críticos $(x, f(x))$ cuando $f'(x - h) > 0$ y, simultáneamente, $f'(x + h) < 0$.

Monotonía Los puntos críticos separan la recta real en intervalos $(I_1, I_2, \dots)$ en los que el signo de la derivada no cambia. Si en un intervalo concreto (I_i) se cumple que:

Si $f'(x) > 0$, entonces la función $f(x)$ es creciente en I_i .

Si $f'(x) < 0$, entonces la función $f(x)$ es decreciente en I_i .

Ejemplo 6.7. Representa cualitativamente la función $f(x) = \frac{x}{1+x^2}$.

1. El dominio de f es $\mathbb{R}$, pues el denominador no se anula.

2. Cortes con los ejes. Con el eje OX, la ecuación $f(x) = 0$ sólo tiene la solución $x = 0$.

Corte con el eje OY. Como $f(0) = 0$, se trata del mismo punto $(0, 0)$, es decir, la función pasa por el origen.

3. Signo de la función. El denominador es siempre positivo, por lo que el signo sólo depende del numerador. Por tanto, la función es negativa cuando $x < 0$ y la función es positiva cuando $x > 0$.
4. La función tiene simetría impar: $f(-x) = -f(x)$
5. Asíntotas. Dado que el denominador nunca se anula no pueden existir asíntotas verticales. Por otro lado, $\lim_{x \rightarrow \pm\infty} \frac{x}{1+x^2} = 0$ existe una asíntota horizontal en $y = 0$ y no pueden haber asíntotas oblicuas.
6. Cálculo de la derivada. La derivada resulta $f'(x) = \frac{1-x^2}{1+x^2}$.

Puntos críticos Si planteamos $\frac{1-x^2}{1+x^2} = 0$ las soluciones son $x = \pm 1$.

Intervalos de monotonía Tenemos tres regiones delimitadas por los dos puntos donde se anula la derivada:

$$I_1 = (-\infty, -1), \quad I_2 = (-1, 1), \quad I_3 = (1, +\infty)$$

Como el signo del derivada es constante en cada uno de los intervalos, para saber el signo evaluamos en un punto de cada intervalo $f'(-2) = -3/5 < 0$, con lo que $f'(x) < 0$ en I_1 y en este intervalo la función es decreciente. Evaluamos en $x = 0$ y $f'(0) = 1 > 0$, con lo que $f'(x) > 0$ en I_2 , por tanto f es creciente en I_2 . Por último, $f'(2) = -3/5 < 0$ y la función es de nuevo decreciente en el intervalo I_3 .

Extremos relativos: De los intervalos de crecimiento deducimos que para $x = -1$ tenemos un mínimo relativo, y que para $x = 1$ tenemos un máximo relativo. Obtengamos el valor de las ordenadas para clarificar donde situar estos puntos singulares de la función:

$$\text{Mínimo en: } x_m = -1 \rightarrow y_m = f(x_m) = \frac{-1}{1+(-1)^2} = -\frac{1}{2}. \text{ En conclusión } m(-1, -\frac{1}{2}).$$

$$\text{Máximo en: } x_m = 1 \rightarrow y_m = f(x_m) = \frac{1}{1+1^2} = \frac{1}{2}. \text{ En conclusión } M(1, \frac{1}{2}).$$

Con estos elementos ya podemos representar la función que aparece en el panel izquierdo de la figura 6.5.

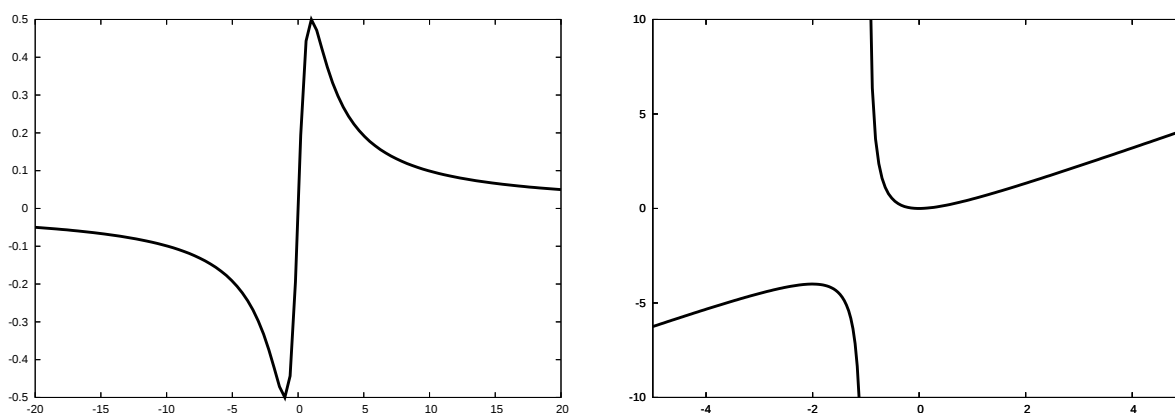


Figura 6.5: Gráfica de $f(x) = \frac{x}{1+x^2}$ en el panel izquierdo y gráfica de $f(x) = \frac{x^2}{x+1}$ en el panel derecho.

Ejemplo 6.8. Representa la función $f(x) = \frac{x^2}{x+1}$.

1. El dominio es $\mathbb{R} \setminus \{-1\}$, pues es una función racional cuyo denominador se anula en $x = -1$.
2. El único corte con los ejes ordenados es $(0, 0)$.
3. Signo de la función. El numerador se anula en $x = 0$ y el denominador en $x = -1$. En los intervalos en que estos puntos separa la recta real el signo de la función f es constante. Como $f(-2) = -4 < 0$, $f(-0.5) = 0.5 > 0$ y $f(1) = 0.5 > 0$. Tenemos que la función f es negativa en el intervalo $(-\infty, -1)$ y positiva en el intervalo $(-1, +\infty)$.
4. No hay simetría, pues $f(-x) = \frac{(-x)^2}{1+(-x)} = \frac{x^2}{1-x}$ que no es $-f(x)$ ni $f(x)$.
5. Asíntotas. La recta $x = -1$ es una asíntota vertical porque para $x = -1$ se anula el denominador y no el numerador. Como el grado del numerador es una unidad mayor que el del denominador hay una asíntota oblicua. Para encontrar dicha asíntota oblicua dividimos los polinomios por el método de Ruffini.

$$\begin{array}{r|rrr} & 1 & 0 & 0 \\ -1 & & -1 & 1 \\ \hline & 1 & -1 & 1 \end{array} \quad \frac{x^2}{x+1} = x - 1 + \frac{1}{x+1}$$

y la asíntota es el cociente del división $y = x - 1$.

6. La primera derivada es $f'(x) = \frac{x^2 + 2x}{(x+1)^2}$. Su denominador se anula en $x = -1$ y el numerador en $x = -2$ y $x = 0$. La función f' tiene signo constante en los intervalos en que estos puntos separa la recta real. Como $f'(-3) > 0$, $f'(-1.5) < 0$, $f'(-0.5) < 0$ y $f'(1) > 0$. Tenemos que f es decreciente en el intervalo $(-2, 0)$; mientras que es creciente en los intervalos $(-\infty, -2)$ y $(0, +\infty)$.

Ejemplo 6.9. Representa la gráfica $f(x) = \frac{x^2 - 3x - 4}{x^2 - x - 2}$.

El dominio de una función racional son aquellos puntos donde no se anula el denominador. Tenemos que resolver $x^2 - x - 2 = 0$ que por la fórmula de segundo grado es

$$x = \frac{1 \pm \sqrt{1 - 4(-2)}}{2} = \frac{1 \pm 3}{2}, \quad x = -1 \text{ ó } x = 2.$$

Con lo que el dominio es $\mathbb{R} \setminus \{-1, 2\} = (-\infty, -1) \cup (-1, 2) \cup (2, +\infty)$.

Para ver el signo de la función hay que dividir la recta real en las discontinuidades ($x = -1$ y $x = 2$) y aquellos puntos donde se anula la función. Un cociente se anula donde se anule el denominador. Por tanto, f se anula donde $x^2 - 3x - 4 = 0$ y resolvemos la ecuación de segundo grado

$$x = \frac{3 \pm \sqrt{9 - 4(-4)}}{2} = \frac{3 \pm 5}{2}, \quad x = 4 \text{ ó } x = -1.$$

Dividimos la recta por los intervalos que separan los puntos: $x = -1$, $x = 2$ y $x = 4$. En cada uno de los subintervalos el signo es constante, para ver dicho signo evaluamos la función en algún punto del subintervalo: $f(-10) = 126/108 > 0$, $f(0) = -4/(-2) > 0$, $f(3) = -4/4 < 0$ y $f(10) = 66/88 > 0$. Por tanto, es positiva en $(-\infty, -1)$, $(-1, 2)$ y $(4, +\infty)$; mientras que es negativa en $(2, 4)$.

Cálculo de asíntotas. Hay asíntota vertical donde se anule el denominador y no el numerador. En nuestro caso en $x = -2$.

En el punto $x = -1$ se anula numerador y denominador, con lo que debemos estudiar el límite para saber si hay asíntota vertical. Al sustituir los valores tenemos una indeterminación $0/0$ con lo que factorizamos, ya que sabemos la raíces $x^2 - 3x - 4 = (x+1)(x-4)$ y $x^2 - x - 2 = (x+1)(x-2)$

$$\lim_{x \rightarrow -1} \frac{x^2 - 3x - 4}{x^2 - x - 2} = \lim_{x \rightarrow -1} \frac{(x+1)(x-4)}{(x+1)(x-2)} = \lim_{x \rightarrow -1} \frac{x-4}{x-2} = \frac{-5}{-3} = \frac{5}{3}.$$

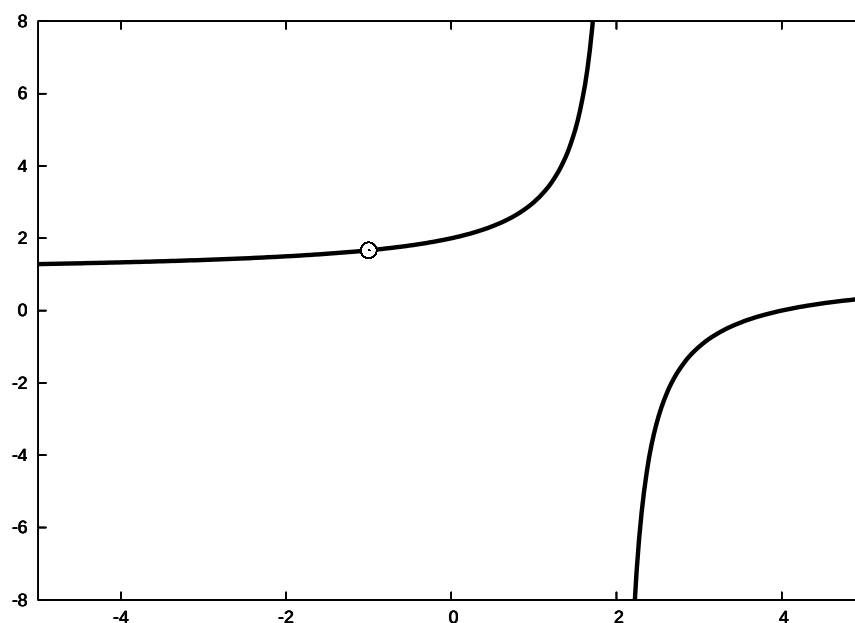


Figura 6.6: Gráfica de la función $f(x) = \frac{x^2 - 3x - 4}{x^2 - x - 2}$.

Con lo que es $x = -1$ no hay asíntota vertical, sino discontinuidad evitable. En realidad,

$$f(x) = \frac{x - 4}{x - 2} \quad \text{si } x \neq -1. \quad (6.1)$$

Para ver las asíntotas horizontales, calculamos el límite en el infinito que vale 1. Por tanto, $y = 1$ es una asíntota horizontal.

Asíntota oblicua no hay, puesto que hay horizontal.

Calculamos la primera derivada

$$f'(x) = \frac{d}{dx} \left[\frac{x - 4}{x - 2} \right] = \frac{(1)(x - 2) - (x - 4)(1)}{(x - 2)^2} = \frac{2}{(x - 2)^2}$$

que obviamente es siempre positiva y la función es creciente de $(-\infty, -1)$, de $(-1, 2)$ y de $(2, +\infty)$.

Capítulo 7

Integración

7.1. Primitivas e integral definida

Una función $F(x)$ que verifica $F'(x) = f(x)$ para todo x en un intervalo se dice que es una *primitiva* de f en dicho intervalo.

Ejemplo 7.1. Consideremos la función $f(x) = 4x^3$, es posible obtenerla por derivación como: $\frac{d}{dx}[x^4] = 4x^3$; por lo tanto, una primitiva de f es $F(x) = x^4$. Como la derivada de cualquier constante es cero, se sigue que las siguientes funciones también son primitivas de $f(x) = 4x^3$: $F_1(x) = x^4 - 7$, $F_2(x) = x^4 + 1$, $F_3(x) = x^4 + 91$.

Teorema 7.1. Dadas dos primitivas de una misma función f , $F_1(x)$ y $F_2(x)$, en un intervalo (a, b) , éstas se diferencian en una constante, es decir, $F_1(x) = F_2(x) + C$ para todo $x \in (a, b)$.

Al conjunto de todas las funciones primitivas de una función f ($F(x) + C$) se le llama *integral indefinida* o *anti-derivada* de f y se denota como $\int f(x)dx$.

Teorema 7.2. Se verifican las siguientes propiedades:

1. La integración es la operación inversa de la derivación $\int \frac{d}{dx}F(x)dx = F(x) + C$;
2. $\int [f(x) + g(x)]dx = \int f(x)dx + \int g(x)dx$;
3. $\int kf(x)dx = k \int f(x)dx$.

Teorema 7.3. Si r es un racional distinto de -1 y a es un número real cualquiera

$$\int (x + a)^r dx = \frac{(x + a)^{r+1}}{r + 1} + C.$$

Demostración. La derivada de $\frac{(x + a)^{r+1}}{r + 1}$ es $\frac{1}{r + 1}(r + 1)(x + a)^r = (x + a)^r$. □

Ejemplo 7.2. Calcula $\int (5x^7 + 2x^2 + 7x^{1/3}) dx$.

Combinando la linealidad (teorema 7.2) y integral indefinida de la potencia (teorema 7.3), se tiene que

$$\int (5x^7 + 2x^2 + 7x^{1/3}) dx = 5 \frac{x^{7+1}}{7+1} + 2 \frac{x^{2+1}}{2+1} + 7 \frac{x^{1/3+1}}{1/3+1} + C = \frac{5}{8}x^8 + \frac{2}{3}x^3 + \frac{21}{4}x^{4/3} + C$$

Ejemplo 7.3. Calcular $\int (\sqrt{x-1} - 1) (\sqrt{x-1} + 1) dx$

Realizamos el producto del integrando con anterioridad a la integración: $(\sqrt{x-1} - 1) (\sqrt{x-1} + 1) = (\sqrt{x-1})^2 - 1^2 = x - 2$. La integración es ahora inmediata: $\int x dx - 2 \int dx = \frac{x^2}{2} - 2x + C$.

Ejemplo 7.4. Calcula $\int \frac{\sqrt[3]{x-5} + 2}{3} dx$.

$$\begin{aligned} \int \frac{\sqrt[3]{x-5} + 2}{3} dx &= \int \left(\frac{1}{3}(x-5)^{1/3} + \frac{2}{3} \right) dx = \frac{1}{3} \int (x-5)^{1/3} dx + \frac{2}{3} \int dx = \\ &= \frac{1}{3} \frac{(x-5)^{1/3+1}}{1/3+1} + \frac{2}{3}x + C = \frac{1}{4}(x-5)^{4/3} + \frac{2}{3}x + C. \end{aligned}$$

Teorema 7.4. Una serie de potencia $f(x) = \sum_{n=0}^{\infty} a_n x^n$ con radio de convergencia R , tiene como primitiva

$F(x) = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}$ en el intervalo $(-R, R)$, es decir,

$$\int \left(\sum_{n=0}^{\infty} a_n x^n \right) dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} + C.$$

Ejemplo 7.5. Encuentra la función F que verifica que $F'(x) = x^2 - 4x$ y $F(0) = 2$.

La función F es una primitiva de $f(x) = x^2 - 4x$, luego existe un C tal que

$$F(x) = \frac{x^3}{3} - 4\frac{x^2}{2} + C = \frac{1}{3}x^3 - 2x^2 + C.$$

Como $F(0) = C$, podemos determinar el valor de C , ya que en la expresión anterior $F(0) = C$, con lo que $C = 2$ y la función buscada es $F(x) = \frac{1}{3}x^3 - 2x^2 + 2$.

Ejemplo 7.6. En una empresa que manufactura harina el coste fijo de producción es 1000€ por año donde el coste fijo de producción son todos los gastos que permanecen constantes independientemente de la producción. Si Q es la producción en kilo y C es el coste de producción medido en euros, el coste en materias primas (trigo) es de 0.2€ por kilo producido y el marginal del coste de la mano de obra varía como $0.1Q^{-0.9}$. Por tanto, el marginal del coste de producción viene dado por

$$\frac{dC}{dQ} = 0.2 + 0.1Q^{-0.9}$$

Determina con estos datos el coste según la producción.

La función de coste es una primitiva de $0.2 + 0.1Q^{-0.9}$, con lo que hay una constante A tal que

$$C(Q) = 0.2Q + 0.1 \frac{Q^{0.1}}{0.1} + A.$$

En este caso hemos llamado a la constante de integración A para no confundirla con el coste. Como los costes en ausencia de producción valen 1000€ tenemos que $C(0) = 1000$, luego $A = 1000$ y la función de costo es

$$C(Q) = 0.2Q + Q^{0.1} + 1000.$$

7.2. Integral definida

Dada una función f acotada no negativa y dos números $a \leq b$, la integral definida denotada por $\int_a^b f(x) dx$ es el área de un trapecio curvilíneo definido por

$$\{(x, y) : a \leq x \leq b \quad 0 \leq y \leq f(x)\}$$

cuya superficie aparece sombreada en la figura 7.1. Si la función no negativa acotada en $[a, b]$ posee sólo un número finito de discontinuidades, es seguro que tiene sentido hablar de área de la anterior superficie. Sin embargo, existen funciones muy complicadas de definir (no nos aparecerán en ningún ejemplo práctico) para las que no tiene sentido hablar del área de dicha superficie. En consecuencia, $\int_a^b f(x) dx$ consiste en un número siempre que tenga sentido.

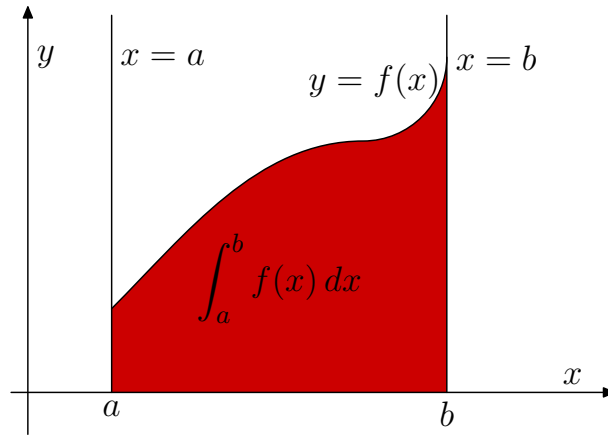


Figura 7.1: El área de la superficie sombreada define la integral de función $f(x)$ en intervalo $[a, b]$.

Ejemplo 7.7. El área de un rectángulo de base $b = 3$ y altura $h = 2$, que es $b \cdot h = 2 \cdot 3 = 6$, se puede identificar con la integral definida $\int_0^3 2 dx$. En el primer panel de la figura 7.2 aparece dicha superficie sombreada.

Ejemplo 7.8. La integral definida $\int_0^4 x dx$ es el área de un triángulo cuya base y altura valen 4, por lo que

$$\int_0^4 x dx = \frac{1}{2} 4 \cdot 4 = \frac{4^2}{2} = 8.$$

En el segundo panel de la figura 7.2 aparece la superficie correspondiente sombreada.

Ejemplo 7.9. La función $f(x) = 1 - |x|$ en el intervalo $[-1, 1]$ está definida por trozos como

$$f(x) = \begin{cases} 1 + x & \text{si } -1 \leq x < 0; \\ 1 - x & \text{si } 0 \leq x \leq 1. \end{cases}$$

La integral definida $\int_{-1}^1 (1 - |x|) dx$ es el área del triángulo sombreado en el tercer panel de la figura 7.2. Como se trata de un triángulo cuya base mide dos unidades y su altura una, el área de dicho triángulo vale uno, así $\int_{-1}^1 (1 - |x|) dx = 1$.

Ejemplo 7.10. La integral definida $\int_{-1}^1 \sqrt{1-x^2} dx$ es $\frac{\pi}{2}$, pues la superficie que aparece sombreada en el cuarto panel de la figura 7.2 es la mitad de un círculo de radio uno.

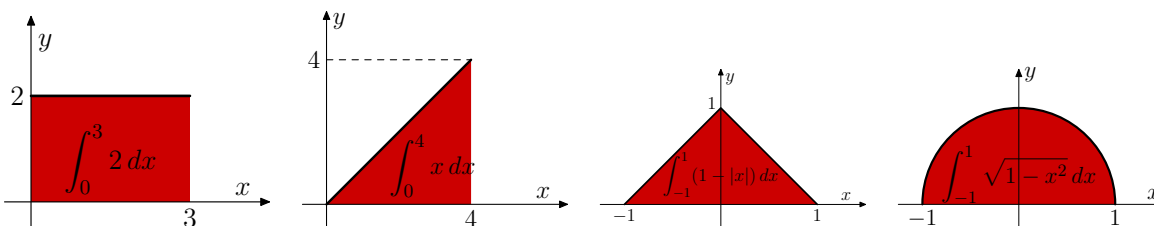


Figura 7.2: Gráficas de las funciones definidas en los ejemplos 7.7, 7.8, 7.9 y 7.10

La integral definida $\int_a^b f(x) dx$ de una función acotada cualquiera (puede ser negativa) se define por

$$\int_a^b f(x) dx = \int_a^b \max\{f(x), 0\} dx - \int_a^b \max\{-f(x), 0\} dx$$

donde ambos términos tienen sentido porque las funciones $\max\{f(x), 0\}$ y $\max\{-f(x), 0\}$ son funciones no negativas. Vea la figura 7.3 donde aparece un ejemplo gráfico de estas funciones.

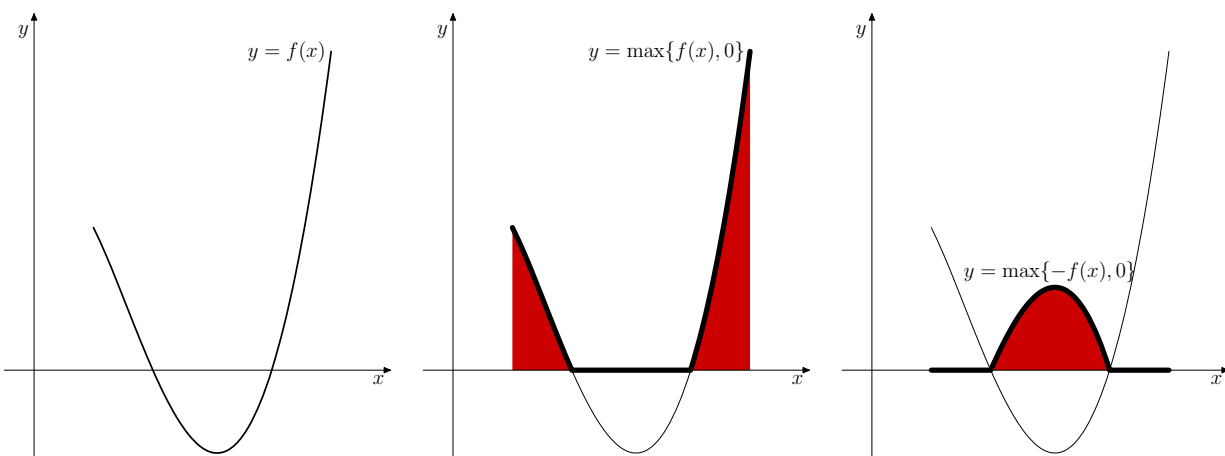


Figura 7.3: La definición de la integral para una función no necesariamente positiva separa el intervalo de definición de la integral en el trozo donde la función está por encima del eje de abscisa y en el trozo que está por debajo del eje de abscisa.

Ejemplo 7.11. Calcular $\int_{-2}^3 x dx$.

La superficie por encima del eje de abscisa es un triángulo cuya base y altura valen 3 y, por consiguiente, el área de dicho triángulo es $9/2$. Por debajo del eje de abscisa la superficie es un triángulo cuyo base y altura son 2 y, por tanto, su área es 2. El valor de la integral

$$\int_{-2}^3 x dx = \frac{9}{2} - 2 = \frac{5}{2}.$$

Si $b < a$ se define la integral definida por $\int_a^b f(x) dx = -\int_b^a f(x) dx$.

Teorema 7.5. La integral definida verifica las propiedades:

1. $\int_a^b [f(x) + g(x)] dx = \int_a^b f(x) dx + \int_a^b g(x) dx$;
2. $\int_a^b k f(x) dx = k \int_a^b f(x) dx$.
3. $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$
4. $\int_a^a f(x) dx = 0$.
5. Si f es integrable en $[a, b]$, y $f(x) \leq g(x)$ para x con $a \leq x \leq b$ y $\int_a^b f(x) dx \leq \int_a^b g(x) dx$
6. La variable de integración es muda, $\int_a^b f(x) dx = \int_a^b f(t) dt$.

Todas estas propiedades son intuitivas dado que la integral no es más que el área comprendida entre la gráfica de la función y el eje de abscisas.

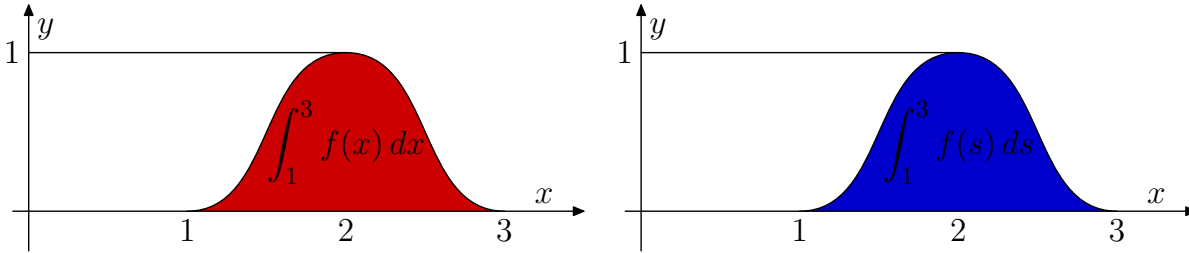


Figura 7.4: El valor de $\int_a^b f(x) dx$ y de $\int_a^b f(t) dt$ es el mismo porque las dos expresiones representan el área de la misma superficie.

Se denomina *partición* del intervalo $[a, b]$ a un conjunto $\Delta = \{x_0, x_1, x_2, \dots, x_{N-1}, x_N\}$ tales que

$$a = x_0 < x_1 < x_2 < \dots < x_{N-1} < x_N = b.$$

Se denomina *diámetro* de la partición $\Delta = \{x_0, x_1, x_2, \dots, x_N\}$ al mayor de los valores $|x_i - x_{i-1}|$ con $i = 1, \dots, N$, es decir,

$$\text{diam}(\Delta) = \max_{i=1, \dots, N} |x_i - x_{i-1}| = \max\{x_1 - x_0, x_2 - x_1, \dots, x_N - x_{N-1}\}.$$

Dada una función f definida en $[a, b]$ y una partición $\Delta = \{x_0, x_1, x_2, \dots, x_N\}$ en $[a, b]$ se dice que un valor S es una *suma integral* de f y Δ si existen c_i en el intervalo $[x_{i-1}, x_i]$ para $i = 1, \dots, N$ tales que

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

Como se puede observar en la figura 7.5 la suma integral consiste en la suma del área de los rectángulos con base $x_i - x_{i-1}$ y altura $f(c_i)$. Resulta bastante intuitivo que dada una partición con diámetro pequeño cualesquiera de sus sumas integrales se aproximan bastante al verdadero valor de la integral definida. Esta idea se enuncia de forma rigurosa en el teorema 7.6

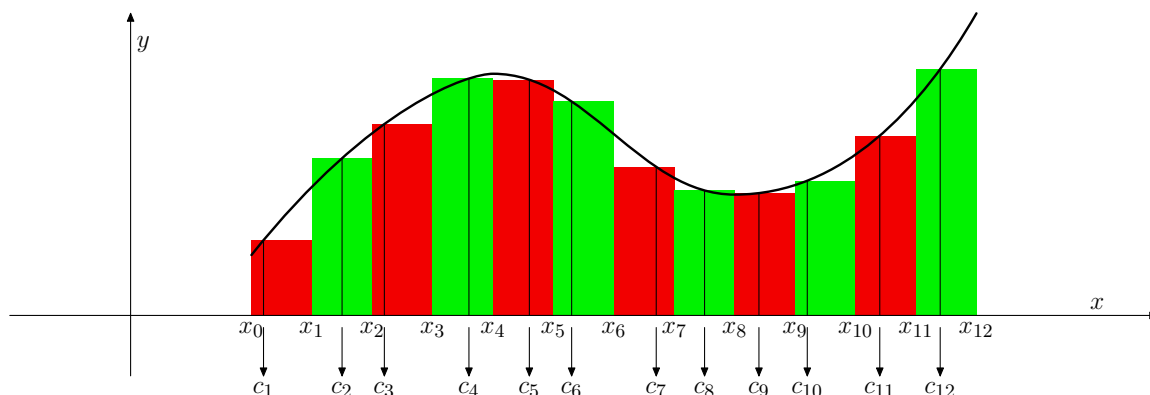


Figura 7.5: La suma del área de los rectángulos sombreados es una suma integral para la partición $\Delta = \{x_0, x_1, x_2, \dots, x_N\}$ y la función f .

Teorema 7.6. Si f es una función acotada con un número finito de discontinuidades, entonces para todo $\epsilon > 0$ existe un $\delta > 0$ tal que para toda suma integral, S , de f y partición Δ de $[a, b]$ con $\text{diam}(\Delta) < \delta$ se verifica que

$$|S - \int_a^b f(x) dx| < \epsilon.$$

7.3. La integral como media

Dada una magnitud económica que es función del tiempo, $f(t)$ (por ejemplo, la cotización de una acción en la bolsa) deseamos medir el valor medio de la acción en un lapso o intervalo de tiempo $[0, T]$. La primera idea para aproximar dicho valor es tomar una partición $\Delta = \{t_0, t_1, \dots, t_N\}$ con $t_j = jT/N$ del intervalo $[0, T]$ y en cada subintervalo $[t_{i-1}, t_i]$ medimos el valor de la acción en un momento arbitrario c_i . La aproximación al valor medio viene dada por la media de los valores de la acción en dichos momentos, $\frac{1}{N} \sum_{i=1}^N f(c_i)$, pero como $\frac{T}{N}$ es la distancia entre dos puntos consecutivos de la partición tenemos que

$$\frac{1}{N} \sum_{i=1}^N f(c_i) = \frac{1}{T} \sum_{i=1}^N f(c_i) \frac{T}{N} = \frac{1}{T} \sum_{i=1}^N f(c_i)(t_i - t_{i-1})$$

Por consiguiente, dicha aproximación al valor medio es una suma integral de f para la partición Δ dividida por T . La aproximación obtenida mejora conforme aumentemos el número de puntos de la partición Δ . Por tanto, el valor medio en el límite es la integral definida en $[0, T]$ dividida por la longitud del intervalo, T .

Por todas las consideraciones anteriores se define el *valor medio* de una función en el intervalo $[a, b]$ como

$$\langle f \rangle = \frac{1}{b-a} \int_a^b f(x) dx.$$

Ejemplo 7.12. Calcular el valor medio de la función $f(x) = 1 - |x|$ en el intervalo $[-1, 1]$.

En el ejemplo 7.9 vimos que $\int_{-1}^1 (1 - |x|) dx = 1$. Por tanto,

$$\langle f \rangle = \frac{1}{1 - (-1)} \int_{-1}^1 (1 - |x|) dx = \frac{1}{2}.$$

Teorema 7.7. Si f es continua en el intervalo $[a, b]$, existe $c \in (a, b)$ tal que $\langle f \rangle = f(c)$ o lo que es equivalente

$$\int_a^b f(x) dx = f(c)(b - a)$$

Demostración. La función f continua en $[a, b]$ alcanza su máximo y su mínimo por el teorema de Weierstrass. Sean c_m y c_M los puntos donde se alcanzan el mínimo y el máximo respectivamente. Por tanto, $f(c_m) \leq f(x) \leq f(c_M)$ para todo x en el intervalo $[a, b]$. Utilizando la propiedad de monotonía de la integración (teorema 7.5 apartado 5) se tiene que

$$f(c_m)(b - a) = \int_a^b f(c_m) dx \leq \int_a^b f(x) dx \leq \int_a^b f(c_M) dx = f(c_M)(b - a).$$

Dividiendo por $b - a$ la anterior desigualdad se obtiene que

$$f(c_m) \leq \frac{1}{b - a} \int_a^b f(x) dx \leq f(c_M).$$

Por el teorema del valor intermedio al ser f continua, existe c tal que $f(c) = \frac{1}{b - a} \int_a^b f(x) dx$. □

7.4. Teorema fundamental del cálculo integral

El siguiente resultado es conocido como teorema fundamental del cálculo y muestra la íntima conexión que existe entre las integrales indefinidas y las definidas.

Teorema 7.8. Si f es continua en $[a, b]$, la función $A(x) = \int_a^x f(t) dt$ es derivable en $[a, b]$ y $A'(x) = f(x)$

Demostración. La derivada de la función A se define por

$$A'(x) = \lim_{h \rightarrow 0} \frac{A(x + h) - A(x)}{h} = \lim_{h \rightarrow 0} \frac{1}{h} \left(\int_a^{x+h} f(t) dt - \int_a^x f(t) dt \right) = \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} f(t) dt$$

Por el teorema 7.7 para cada h existe un c_h entre x y $x + h$ tal que $\frac{1}{h} \int_x^{x+h} f(t) dt = f(c_h)$. Cuando h tiende a cero, $f(c_h)$ tiende a $f(x)$ por continuidad y, por tanto, $A'(x) = f(x)$. □

Ejemplo 7.13. Derivar la función $F(x) = \int_1^x \frac{1}{t} dt$.

Por el teorema 7.8 tenemos que $F'(x) = \frac{1}{x}$.

Ejemplo 7.14. Derivar la función $F(x) = \int_{x^2}^{x^2+1} \frac{1}{t+1} dt$.

Si llamamos $G(y) = \int_0^y \frac{1}{t+1} dt$ con el teorema 7.8 sabemos que $G'(y) = \frac{1}{y+1}$.

$$\begin{aligned} F(x) &= \int_{x^2}^{x^2+1} \frac{1}{t+1} dt = \int_0^{x^2+1} \frac{1}{t+1} dt - \int_0^{x^2} \frac{1}{t+1} dt \\ &= \int_0^{x^2+1} \frac{1}{t+1} dt - \int_0^{x^2} \frac{1}{t+1} dt = G(x^2 + 1) - G(x^2). \end{aligned}$$

Derivamos la función F usando la regla de la cadena (teorema 5.4)

$$\frac{d}{dx}[F(x)] = \frac{d}{dx}[G(x^2 + 1)] - \frac{d}{dx}[G(x^2)] = G'(x^2 + 1)2x - G'(x^2)2x = 2x \left(\frac{1}{x^2 + 1} - \frac{1}{x^2} \right) = \frac{-2x}{x^2(x^2 + 1)}$$

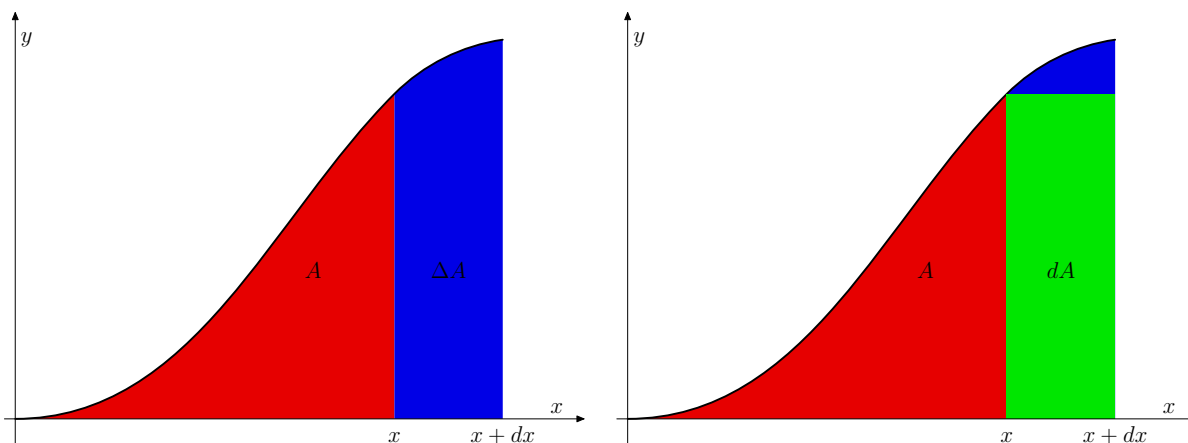


Figura 7.6: Interpretación geométrica del teorema fundamental del cálculo.

Regla de Barrow

Una consecuencia del teorema fundamental del cálculo integral es la conocida regla de Barrow. Si $F(x)$ es una primitiva de una función continua $f(x)$, se tiene que existe una constante C tal que

$$F(x) = \int_a^x f(t) dt + C$$

y, finalmente,

$$F(b) - F(a) = \left(\int_a^b f(t) dt + C \right) - \left(\int_a^a f(t) dt + C \right) = \int_a^b f(t) dt.$$

Dicha fórmula sigue siendo correcta aunque f no sea continua en todos sus puntos.

Teorema 7.9. Si F es una primitiva de f entonces:

$$\int_a^b f(x) dx = [F(x)]_a^b = F(b) - F(a)$$

Ejemplo 7.15. Calcular $\int_1^3 (x-2)^4 dx$.

Como $\frac{(x-2)^5}{5}$ es una primitiva de $(x-2)^4$, tenemos que

$$\int_1^3 (x-2)^4 dx = \left[\frac{(x-2)^5}{5} \right]_1^3 = \frac{(3-2)^5}{5} - \frac{(1-2)^5}{5} = \frac{1}{5} - \frac{-1}{5} = \frac{2}{5}$$

Ejemplo 7.16. Calcula el valor medio de $f(x) = \sqrt{2x+1}$ en el intervalo $[0, 1]$.

$$\begin{aligned} \langle f \rangle &= \frac{1}{1-0} \int_0^1 (2x+1)^{1/2} = \int_0^1 2^{1/2} \left(x + \frac{1}{2}\right)^{1/2} dx = \sqrt{2} \left[\frac{(x + \frac{1}{2})^{3/2}}{3/2} \right]_0^1 = \\ &= \sqrt{2} \left(\frac{(3/2)^{3/2}}{3/2} - \frac{(1/2)^{3/2}}{3/2} \right) = \sqrt{3} - \frac{1}{3}. \end{aligned}$$

7.5. Integrales impropias

Sea f una función continua en el intervalo $[a, +\infty)$ con a un número real, se define la *integral impropia* de f en el intervalo $[a, +\infty)$ como el límite

$$\int_a^{+\infty} f(x) dx = \lim_{b \rightarrow +\infty} \int_a^b f(x) dx$$

Si el límite anterior existe se dice que la integral impropia $\int_a^{+\infty} f(x) dx$ es convergente.

Sea f una función continua en el intervalo $(-\infty, b]$ con b un número real, se define la integral impropia de f en el intervalo $(-\infty, b]$ como el límite

$$\int_{-\infty}^b f(x) dx = \lim_{a \rightarrow -\infty} \int_a^b f(x) dx$$

Esta definición es básica para la estadística, ya que las funciones de distribución acumulada se definen mediante ellas.

Ejemplo 7.17. Calcular $\int_2^{+\infty} \frac{1}{x^2} dx$.

$$\int_2^{+\infty} \frac{1}{x^2} dx = \lim_{b \rightarrow +\infty} \int_2^b x^{-2} dx = \lim_{b \rightarrow +\infty} \left[\frac{x^{-2+1}}{-2+1} \right]_2^b = \lim_{b \rightarrow +\infty} \frac{b^{-1}}{-1} - \frac{2^{-1}}{-1} = \lim_{b \rightarrow +\infty} \frac{1}{2} - \frac{1}{b} = \frac{1}{2}.$$

Ejemplo 7.18. Calcular $\int_1^{+\infty} \frac{1}{\sqrt[3]{x+1}} dx$.

$$\int_1^{+\infty} \frac{1}{\sqrt[3]{x+1}} dx = \lim_{a \rightarrow +\infty} \int_1^a (x+1)^{-1/3} dx = \lim_{a \rightarrow +\infty} \left[\frac{(x+1)^{2/3}}{2/3} \right]_1^a = \lim_{a \rightarrow +\infty} \frac{(a+1)^{2/3}}{2/3} - \frac{(1+1)^{2/3}}{2/3} = +\infty$$

7.6. Criterio integral para convergencia de series

Teorema 7.10. Sea f una función continua, positiva y decreciente para $x \geq 1$, y sea $a_n = f(n)$. Entonces la serie $\sum_{n=1}^{\infty} a_n$ converge o diverge conforme la integral impropia $\int_1^{\infty} f(x) dx$ converge o diverge.

Demostración. Como la función es decreciente para $x \geq 1$, se verifica la siguiente acotación

$$\sum_{r=2}^n f(r) \leq \int_1^n f(x) dx \leq \sum_{r=1}^{n-1} f(r) \leq \sum_{r=1}^n f(r).$$

Como $a_r = f(r)$ obtenemos

$$\sum_{r=2}^n a_r \leq \int_1^n f(x) dx \leq \sum_{r=1}^n a_r$$

Entonces las sumas y la integral convergen o divergen respectivamente cuando consideramos los límites cuando n tiende a ∞ . $\square$

Teorema 7.11. La serie $\sum_{n=1}^{\infty} \frac{1}{n^p}$ es divergente para $0 < p \leq 1$, y convergente para $p > 1$.

Demostración. El término general de la serie es $a_n = \frac{1}{n^p}$, entonces consideraremos la función

$$f(x) = \frac{1}{x^p} = x^{-p},$$

para que con $a_n = f(n)$ podamos emplear el criterio de la integral. Dicha función está definida para $[1, +\infty)$ es continua y decreciente. Para emplear el criterio de la integral calculamos

$$\int_1^n x^{-p} dx = \left[\frac{x^{-p+1}}{1-p} \right]_1^n = \frac{1}{1-p} \left(\frac{1}{n^{p-1}} - 1 \right),$$

por lo que estamos considerando que $p \neq 1$. Si $0 < p < 1$, se tiene que

$$\lim_{n \rightarrow \infty} \frac{1}{n^{p-1}} = \lim_{n \rightarrow \infty} n^{1-p} = \infty,$$

y en consecuencia la integral impropia es divergente. Si $p > 1$, entonces $\lim_{n \rightarrow \infty} \frac{1}{n^{p-1}} = 0$. y en consecuencia la integral impropia es convergente. Si $p = 1$ en el ejemplo 4.19 ya se demostró que la serie $\sum_{n=1}^{\infty} \frac{1}{n}$ es divergente. $\square$

Capítulo 8

Funciones exponenciales y logarítmicas

En el capítulo anterior vimos que $\int x^r dx = \frac{x^{r+1}}{r+1} + C$ si $r \neq -1$. Las primitivas para el caso $r = -1$ permiten definir la función logaritmo que fueron fundamentales para agilizar el cálculo de operaciones elementales lo que fue descubierto por John Napier. La recíproca del logaritmo es la función exponencial que permite expresar el capital respecto del tipo de interés de forma continua. También juega un papel decisivo la exponencial en los modelos que estudian las poblaciones.

8.1. Definición de logaritmo natural. Propiedades.

El *logaritmo natural* o *neperiano* se define como $\ln(x) := \int_1^x \frac{1}{t} dt$. En algunos libros se denota como $\log(x)$.

Teorema 8.1. La función logaritmo natural verifica las siguientes propiedades:

1. su dominio es $(0, \infty)$;
2. es derivable en todo su dominio y su derivada es $\frac{1}{x}$, $\frac{d}{dx}[\ln(x)] = \frac{1}{x}$;
3. es estrictamente creciente en $(0, +\infty)$;
4. $\ln(1) = 0$;
5. $\ln(ab) = \ln(a) + \ln(b)$ con $a > 0$ y $b > 0$;
6. $\ln(a/b) = \ln(a) - \ln(b)$ con $a > 0$ y $b > 0$;
7. $\ln(a^r) = r \ln(a)$ con $a > 0$ y r racional;
8. verifica los siguientes límites:

$$\lim_{x \rightarrow 0^+} \ln(x) = -\infty \quad \text{y} \quad \lim_{x \rightarrow +\infty} \ln(x) = +\infty.$$

Demostración. La función $f(t) = 1/t$ es continua en $[x, 1]$ si $0 < x < 1$ y en $[1, x]$ para $x > 1$, con lo que tiene sentido la integral que define el logaritmo para cualquier valor positivo. Por el teorema fundamental del cálculo integral (teorema 7.8) el logaritmo es derivable y su derivada es $1/x$. Como la derivada es

positiva en cualquier valor positivo se tienen que la función es estrictamente creciente por el teorema 6.6. De las propiedades para integrales definidas se tiene que

$$\ln(1) = \int_1^1 \frac{1}{t} dt = 0.$$

La función $g(x) = \ln(ax) - \ln(x)$ es derivable y su derivada es $g'(x) = \frac{1}{ax} a - \frac{1}{x} = 0$. Por el teorema 6.6 la función $g(x)$ es constante. Como $g(1) = \ln(a)$, tenemos que $g(b) = g(1) = \ln(a)$, con lo que $\ln(ab) = \ln(a) + \ln(b)$.

Aplicando la anterior propiedad tenemos que

$$\ln(a) = \ln\left(b \frac{a}{b}\right) = \ln(b) + \ln\left(\frac{a}{b}\right)$$

y, despejando, $\ln\left(\frac{a}{b}\right) = \ln(a) - \ln(b)$.

Si n es un natural, tenemos que

$$\ln(a^n) = \ln(a \cdot a \cdots a) = \ln(a) + \ln(a) + \cdots + \ln(a) = n \ln(a).$$

Por otro lado,

$$\ln(a^{-n}) = \ln\left(\frac{1}{a^n}\right) = -\ln(a^n) = -n \ln(a).$$

Si n es un entero y m es un natural, entonces

$$n \ln(a) = \ln(a^n) = \ln((a^{n/m})^m) = m \ln(a^{n/m}),$$

y, despejando, $\ln(a^{n/m}) = \frac{n}{m} \ln(a)$.

Para probar el resultado con los límites se tiene en cuenta que $\ln(2)$ es positivo, con lo que

$$\lim_{n \rightarrow \infty} \ln(2^n) = \lim_{n \rightarrow \infty} n \ln(2) = +\infty \quad \lim_{n \rightarrow \infty} \ln\left(\left(\frac{1}{2}\right)^n\right) = \lim_{n \rightarrow \infty} -n \ln(2) = -\infty.$$

Lo anterior junto a que el logaritmo natural sea estrictamente creciente implica el resultado buscado sobre límites $\square$

Ejemplo 8.1. Calcula la derivada de $f(x) = \ln(4 + x^2)$.

Llamando $g(x) = 4 + x^2$, tenemos que $f(x) = \ln(g(x))$. Por la regla de la cadena se tiene que

$$f'(x) = \frac{g'(x)}{g(x)} = \frac{2x}{4 + x^2}.$$

Ejemplo 8.2. Calcular $\sum_{k=1}^{\infty} \ln\left(\frac{k+1}{k}\right)$.

Con las propiedades de la función logaritmo,

$$S_n = \ln 2 + \ln \frac{3}{2} + \cdots + \ln \frac{n+1}{n} = \ln \left(\frac{2 \cdot 3 \cdot 4 \cdots (n+1)}{1 \cdot 2 \cdot 4 \cdots n} \right) = \ln(n+1).$$

En consecuencia $\lim_{n \rightarrow \infty} S_n = \infty$ y la serie es divergente.

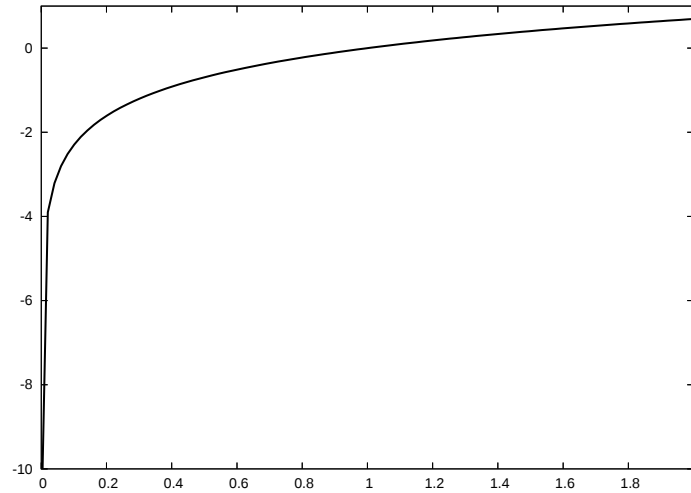


Figura 8.1: Gráfica de la función logaritmo.

Ejemplo 8.3. Encontrar la expansión de Maclaurin de la función $\ln(1+x)$.
Calculamos la derivada n -ésima

$$\begin{aligned}\frac{d^n}{dx^n}[\ln(1+x)] &= \frac{d^{n-1}}{dx^{n-1}}\left[\frac{d}{dx}[\ln(1+x)]\right] = \frac{d^{n-1}}{dx^{n-1}}[(1+x)^{-1}] = \\ &= \binom{-1}{n-1}(n-1)!(1+x)^{-1-(n-1)} = \binom{-1}{n-1}(n-1)!(1+x)^{-n} =\end{aligned}$$

Luego, $f^n(0) = (-1)(-2)(-3)\cdots(-(n-2))(-(n-1)) = (-1)^{n-1}(n-1)!$. Por tanto, la serie de MacLaurin es

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

Como se comentó en un tema anterior para muchas funciones su serie de MacLaurin coincide con la función, la función $\ln(1+x)$ es uno de esos casos.

Teorema 8.2. Para $-1 < x < 1$ se verifica que $\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n$

Ejemplo 8.4. Encuentra los tres primeros términos de la serie de McLaurin para $x \ln(1+x^2)$.
En el ejemplo 8.2 vimos que

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

por lo tanto, sustituyendo x por x^2 obtenemos que

$$\ln(1+x^2) = x^2 - \frac{x^4}{2} + \frac{x^6}{3} - \frac{x^8}{4} + \dots,$$

y multiplicando conseguimos el desarrollo de $x \ln(1+x^2)$

$$x \ln(1+x^2) = x^3 - \frac{x^5}{2} + \frac{x^7}{3} + \dots,$$

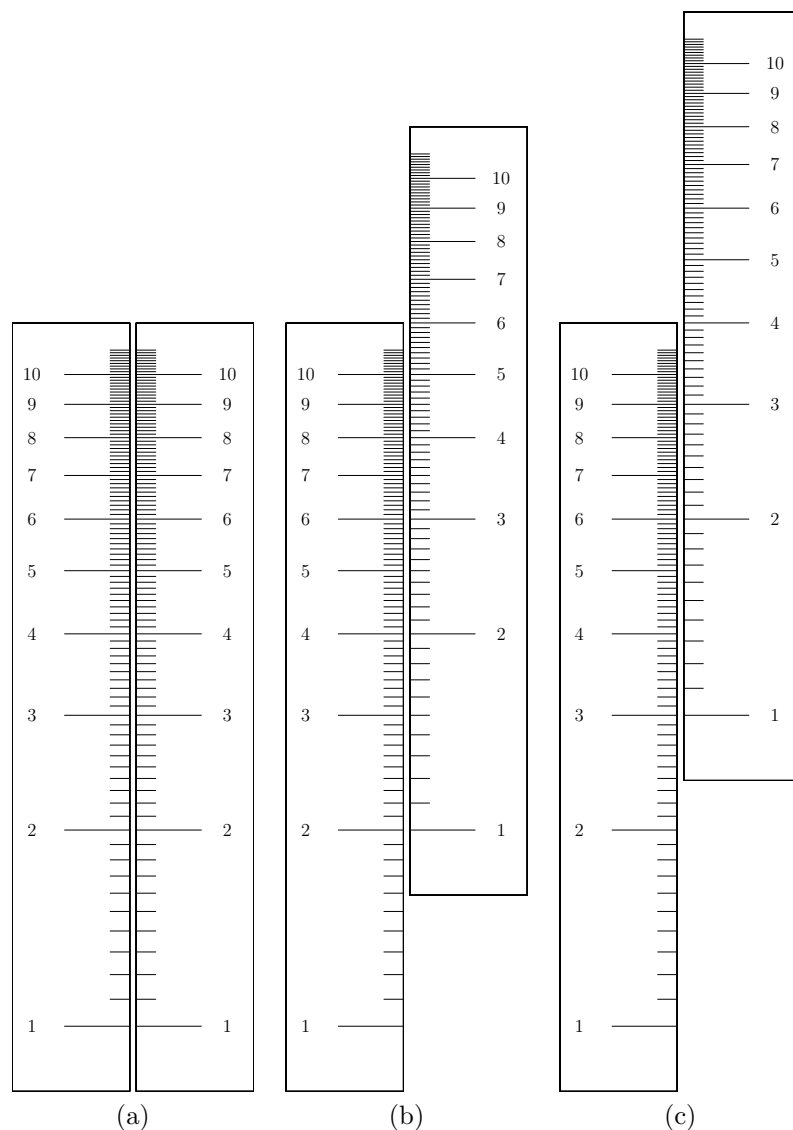


Figura 8.2: Ejemplo de uso de la regla de cálculo. La regla de cálculo consiste en dos reglas simétricas (a). Observe como si desplazamos la regla de cálculo de la derecha hasta hacer coincidir el uno con el dos, tenemos a la izquierda el doble de la cantidad de la derecha (panel (b)). En el panel (c), desplazamos la regla de cálculo hasta que coincida el uno con el tres y se obtiene que las cantidades de la izquierda se corresponden con el triple de la derecha. La idea consiste en que la graduación está en escala logarítmica y la propiedad $\ln(ab) = \ln(a) + \ln(b)$ permite que las multiplicaciones se obtengan por simple desplazamiento de la regla de la derecha. La regla de cálculo fue ampliamente utilizada en la primera mitad del siglo XX hasta que la calculadora electrónica la reemplazó. La calculadora electrónica sigue fundamentando muchos de sus cálculos en la función logaritmo.

8.2. Integrales de funciones racionales.

Teorema 8.3. Para la función logaritmo natural se verifica que

$$\frac{d}{dx}[\ln |x + a|] = \frac{1}{x + a}$$

para todo $x \neq -a$ y, en consecuencia,

$$\int \frac{1}{x + a} dx = \ln |x + a| + C.$$

Si el grado del numerador es mayor o igual que el grado del denominador se hace la división euclídea de la función racional. Cuando tenemos una función racional con grado del numerador menor que el del denominador se escribe la descomposición en fracciones simples. Dicha descomposición es fácil de integrar con el resultado anterior.

Ejemplo 8.5. Calcular $\int \frac{x^2 + 3x + 1}{x + 1} dx$.

Realizamos la división euclídea de los dos polinomios de la función racional

$$\begin{array}{r} x^2 + 3x - 1 \quad | \quad x + 1 \\ x^2 + \quad x \quad \quad \quad \\ \hline 2x - 1 \\ 2x + 2 \\ \hline -3 \end{array}$$

Por tanto, $\frac{x^2 + 3x - 1}{x + 1} = x + 2 - \frac{3}{x + 1}$ expresión que es fácil de integrar

$$\int \frac{x^2 + 3x - 1}{x + 1} dx = \int (x + 2 - \frac{3}{x + 1}) dx = \frac{x^2}{2} + 2x - 3 \ln |x + 1| + C$$

Ejemplo 8.6. Calcular $\int \frac{1}{x^2 - 1} dx$.

Las raíces del denominador $p(x) = x^2 - 1$ son $x = \pm 1$. Por tanto, la descomposición en fracciones simples es

$$\frac{1}{x^2 - 1} = \frac{A}{x - 1} + \frac{B}{x + 1}$$

se tiene que $1 = A(x + 1) + B(x - 1)$ y dando los valores $x = \pm 1$ se obtiene que $A = 1/2$ y $B = -1/2$. Tras la descomposición en fracciones simples podemos calcular la integral

$$\int \frac{1}{x^2 - 1} dx = \int \frac{1}{2} \left(\frac{1}{x - 1} - \frac{1}{x + 1} \right) dx = \frac{1}{2} \ln |x - 1| - \frac{1}{2} \ln |x + 1| + C.$$

Ejemplo 8.7. Calcular $\int \frac{x^2 + x - 3}{x^3 + x^2 - 4x - 4} dx$.

En el ejemplo 3.10 vimos que la descomposición en fracciones simples es

$$\frac{x^2 + x - 3}{x^3 + x^2 - 4x - 4} = \frac{1}{4} \frac{1}{x - 2} - \frac{1}{4} \frac{1}{x + 2} + \frac{1}{x + 1}.$$

Por tanto, su integral es

$$\int \frac{x^2 + x - 3}{x^3 + x^2 - 4x - 4} dx = \frac{1}{4} \ln |x - 2| - \frac{1}{4} \ln |x + 2| + \ln |x + 1| + C.$$

Ejemplo 8.8. Calcular $\int \frac{x^2 + x + 3}{x^3 + 3x^2 - 4} dx$.

En el ejemplo 3.11 vimos que la descomposición en fracciones simples es

$$\frac{x^2 + x + 3}{x^3 + 3x^2 - 4} = \frac{5}{9} \frac{1}{x-1} + \frac{4}{9} \frac{1}{x+2} - \frac{5}{3} \frac{1}{(x+2)^2}.$$

con lo que la integral resulta

$$\int \frac{x^2 + x + 3}{x^3 + 3x^2 - 4} dx = \frac{5}{9} \ln|x-1| + \frac{4}{9} \ln|x+2| - \frac{5}{3} \frac{(x+2)^{-1}}{-1} + C.$$

8.3. Definición de la función exponencial. Propiedades.

La inversa de la función logaritmo natural se llama función *exponencial* natural y se denota por $\exp(x)$.

Teorema 8.4. La exponencial verifica que

1. su dominio es el conjunto de todos los números reales, $\mathbb{R}$;
2. es derivable en todos los puntos y su derivada es ella misma $\frac{d}{dx}[\exp(x)] = \exp(x)$;
3. es estrictamente creciente en su dominio;
4. $\exp(0) = 1$;
5. $\exp(a+b) = \exp(a)\exp(b)$ para todo a y b real;
6. $\frac{\exp(a)}{\exp(b)} = \exp(a-b)$ para todo a y b real.
7. $\lim_{x \rightarrow +\infty} \exp(x) = +\infty$ y $\lim_{x \rightarrow -\infty} \exp(x) = 0$;
8. su integral es $\int \exp(x) dx = \exp(x) + C$.

Ejemplo 8.9. Calcular la serie de MacLaurin de la exponencial $f(x) = \exp(x)$. Es obvio que $f^{(n)}(x) = \exp(x)$, con lo que $f^{(n)}(0) = \exp(0) = 1$. Por tanto, la serie de MacLaurin de la exponencial es

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$$

De nuevo, la serie de MacLaurin de la exponencial coincide con la función.

Teorema 8.5. Para todo x real, se tiene que $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.

8.3.1. El número e

El número e se define como $e = \exp(1)$ o, lo que es lo mismo, el número real positivo tal que $\ln(e) = 1$. Por el teorema 8.5 tenemos que $e = \sum_{n=0}^{\infty} \frac{1}{n!}$ es posible encontrar una aproximación a dicho valor:

$$e = 2.718281828459045235360287471...$$

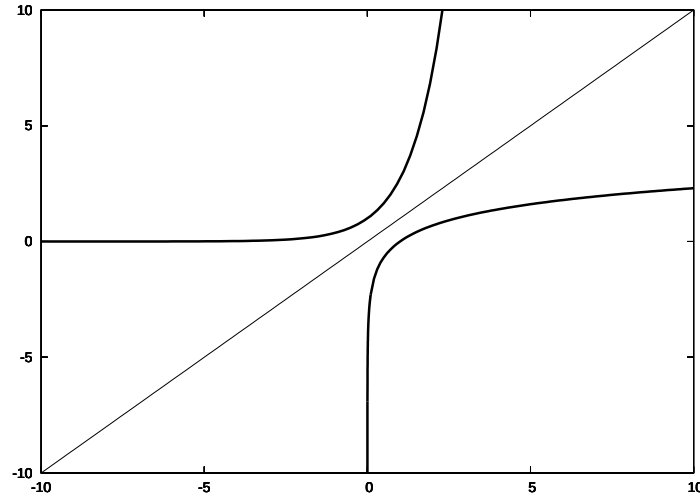


Figura 8.3: Obsérvese la simetría respecto del eje $y = x$ de la gráficas $y = e^x$ e $y = \ln(x)$.

Teorema 8.6. Si r es un número racional, $\exp(r) = e^r$.

Demostración. Para n natural tenemos que $\exp(n) = \exp(1 + \dots + 1) = \exp(1) \cdots \exp(1) = e^n$. Por otro lado, $1 = \exp(0) = \exp(n + (-n)) = \exp(n) \exp(-n)$ y despejando $\exp(-n) = e^{-n}$.

Si n es un entero y m es un natural

$$e^n = \exp(n) = \exp\left(n \frac{m}{m}\right) = \exp\left(m \frac{n}{m}\right) = \exp\left(\frac{n}{m}\right)^m$$

con lo que $\exp\left(\frac{n}{m}\right) = e^{n/m}$. □

Por extensión, para cualquier x real denotaremos por e^x a $\exp(x)$.

Ejemplo 8.10. Resolver la ecuación $\ln(x^2 - 9) = 2$.

Tomando exponenciales resulta $x^2 - 9 = e^2$ y despejando $x = \pm\sqrt{9 + e^2}$, con lo que tenemos dos soluciones de la ecuación.

8.4. Definición de funciones exponenciales con base arbitraria.

Se define la exponencial en base a de x como $a^x := \exp(x \ln(a)) = e^{x \ln(a)}$ para todo $a > 0$.

Teorema 8.7. Para a y b positivos y x e y cualesquiera la función exponencial verifica las siguientes propiedades

1. $a^0 = 1$
2. $a^x a^y = a^{x+y}$
3. $\frac{a^x}{a^y} = a^{x-y}$
4. $(a^x)^y = a^{xy}$
5. $a^x b^x = (ab)^x$

6. $\left(\frac{a}{b}\right)^x = \frac{a^x}{b^x}$

7. es derivable y su derivada es $\frac{d}{dx}[a^x] = \ln(a)a^x$

8. si $a > 1$ es una función estrictamente creciente y si $a < 1$ es una función estrictamente decreciente.

Demostración. Tenemos que $a^0 = \exp(0 \ln(a)) = \exp(0) = 1$ y la exponencial de una suma es

$$a^{x+y} = \exp((x+y) \ln(a)) = \exp(x \ln(a) + y \ln(a)) = \exp(x \ln(a)) \exp(y \ln(a)) = a^x a^y$$

de forma similar se obtiene el resultado para a^{x-y} y, usando las propiedades de la exponencial y el logaritmo se obtienen el resto de las propiedades

$$(a^x)^y = \exp(x \ln(a))^y = \exp(y \ln(x \ln(a))) = \exp(y x \ln(a)) = a^{xy}$$

$$a^x b^x = \exp(x \ln(a)) \exp(x \ln(b)) = \exp(x (\ln(a) + \ln(b))) = \exp(x \ln(ab)) = (ab)^x.$$

La derivada se obtiene usando la regla de la cadena

$$\frac{d}{dx}[a^x] = \frac{d}{dx}[\exp(x \ln(a))] = \exp(x \ln(a)) \frac{d}{dx}[x \ln(a)] = a^x \ln(a)$$

Como a^x es positiva al ser la imagen de la exponencial, el signo de la primera derivada es el de $\ln(a)$ que es positivo si $a > 1$ y negativo si $a < 1$. Esto último implica las propiedades de crecimiento y decrecimiento vistas. $\square$

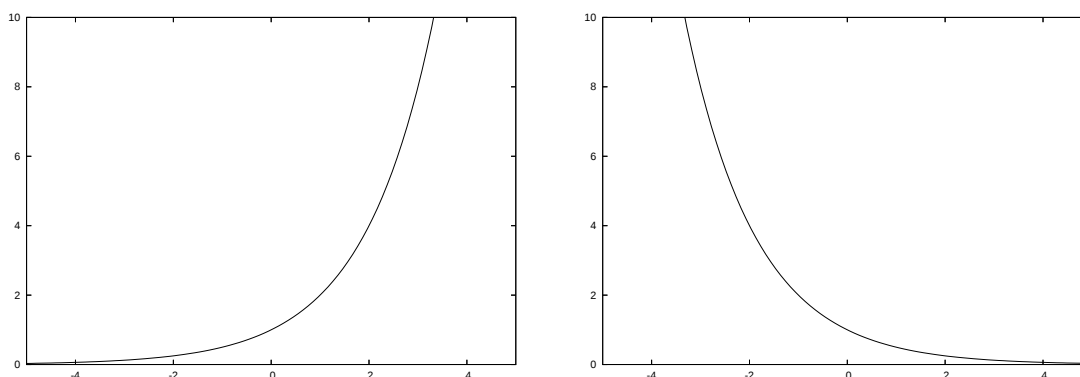


Figura 8.4: Gráficas de las funciones 2^x y $\left(\frac{1}{2}\right)^x$. Obsérvese como la primera es creciente, ya que la base es mayor que uno y la segunda es decreciente por ser en este caso la base es menor que uno.

Ejemplo 8.11. Calcular $\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x$.

Por la definición $\left(1 + \frac{1}{x}\right)^x = \exp\left(x \ln\left(1 + \frac{1}{x}\right)\right)$. Calculamos el límite de a lo que se aplica la exponencial $x \ln\left(1 + \frac{1}{x}\right)$ que cuando $x \rightarrow \infty$ tiene una indeterminación $0 \cdot \infty$, que se puede escribir como una indeterminación $0/0$ pasando la x al denominador

$$\lim_{x \rightarrow \infty} x \ln\left(1 + \frac{1}{x}\right) = \lim_{x \rightarrow \infty} \frac{\ln\left(1 + \frac{1}{x}\right)}{\frac{1}{x}} = \lim_{x \rightarrow +\infty} \frac{\frac{1}{1 + \frac{1}{x}} \left(-\frac{1}{x^2}\right)}{-\frac{1}{x^2}} = \lim_{x \rightarrow \infty} \frac{1}{1 + \frac{1}{x}} = 1.$$

Luego, $\lim_{x \rightarrow \infty} (1 + \frac{1}{x})^x = \exp(1) = e$.

El número e surgió por motivaciones económicas. Imaginemos un banquero ingenuo que anuncia un 5 % de interés anual, pero que aplica un interés compuesto en plazos tan pequeños como desee el cliente, es decir, si dividimos el año en n partes se aplica a cada una de las partes una tasa de interés de $0.05/n$. Por ejemplo, si divide en meses y se reinvierte cada mes el capital la tasa de interés obtenida al final es

$$(1 + \frac{0.05}{12})^{12} - 1 \approx 0.05116189788173318980487389096$$

que es mayor que el 5 % anunciado. En el caso de querer aprovecharnos de un banquero tan ingenuo debemos dividir en partes tan pequeñas como sea posible. En el límite de lo pequeño obtenemos

$$\lim_{n \rightarrow \infty} (1 + \frac{0.05}{n})^n - 1 = e^{0.05} - 1 \approx 0.05127109637602403969751763633$$

es decir, nunca obtenemos una tasa de interés mayor por pequeña que sea el plazo, pero con plazos muy pequeños nos aproximamos a dicho valor.

Ejemplo 8.12. Si un depósito ofrece un 5 % de interés anual. Calcula la tasa de variación proporcional para un capital inicial de 1000€ en cualquier instante.

El capital en el instante t (medido en años) es función del tiempo $C(t) = (1.05)^t 1000$. Por tanto, $\frac{dC}{dt} = 1000(1.05)^t \ln(1.05) = \ln(1.05)$ y, por tanto, la tasa de variación proporcional es

$$\frac{1}{C} \frac{dC}{dt} = \frac{1}{1000(1.05)^t} \ln(1.05) = \ln(1.05)$$

Por tanto, la tasa de variación proporcional del capital es constante y vale $\ln(1.05)$.

Ejemplo 8.13. Calcule el límite de $\lim_{x \rightarrow \infty} (\frac{x+1}{x+2})^x$.

Podemos escribir $(\frac{x+1}{x+2})^x = \exp(x \ln(\frac{x+1}{x+2}))$. El exponente $x \ln((x+1)/(x+2))$ tiene una indeterminación $\infty \cdot 0$. Si lo escribimos como un cociente podemos aplicar L'Hôpital

$$\begin{aligned} \lim_{x \rightarrow +\infty} x \ln(\frac{x+1}{x+2}) &= \lim_{x \rightarrow +\infty} \frac{\ln(\frac{x+1}{x+2})}{1/x} = \lim_{x \rightarrow +\infty} \frac{\ln(1 - \frac{1}{x+2})}{1/x} = \\ &= \lim_{x \rightarrow +\infty} \frac{1}{\frac{(x+2)^2}{-1/x^2}} \frac{x+2}{x+1} = \lim_{x \rightarrow +\infty} \frac{-x^2}{(x+2)(x+1)} = -1 \end{aligned}$$

Con lo que el límite vale e^{-1} .

8.5. Derivación logarítmica.

Para una función que no involucra logaritmos pero aparecen potencias, la derivada suele obtenerse tomando logaritmos naturales a ambos lados y derivando.

Ejemplo 8.14. Derivar $f(x) = \frac{(x+1)^3}{(x^2+4)^{3/2}}$ mediante derivación logarítmica.

Tomando logaritmos, obtenemos que

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

Derivando, llegamos a que

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

Con lo que la derivada resulta ser

$$f'(x) = 3 \frac{(x+1)^3}{(x^2+4)^{3/2}} \left(\frac{1}{x+1} - \frac{x}{x^2+4} \right) = 3 \frac{(x+1)^2(4-x)}{(x^2+4)^{5/2}}.$$

Ejemplo 8.15. Derivar $f(x) = x^x$

Si tomamos logaritmo obtenemos,

$$\ln(f(x)) = \ln(x^x) = x \ln(x).$$

Derivando, se obtiene que

$$\frac{f'(x)}{f(x)} = \ln(x) + x \frac{1}{x} = 1 + \ln(x).$$

Por tanto, $f'(x) = x^x(1 + \ln(x))$.

Ejemplo 8.16. Deriva la función $f(x) = \ln(x)^{1/x}$.

Aplicamos la función logaritmo a ambos miembros de la igualdad

$$\ln(f(x)) = \ln(\ln(x)^{1/x}) = \frac{1}{x} \ln(\ln(x))$$

y derivando respecto de la variable x ambos miembros de la igualdad

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

y, por último, despejando obtenemos que

$$f'(x) = \ln(x)^{1/x} \frac{1}{x^2} \left(\frac{1}{\ln(x)} - \ln(\ln(x)) \right).$$

8.6. Ecuaciones diferenciales: variables separadas. Modelo logístico de crecimiento de poblaciones.

Una *ecuación diferencial* es una ecuación donde la incógnita es una función de la que se conoce alguna relación que liga la derivada, la variable independiente y la propia función (variable dependiente).

El ejemplo más sencillo de ecuación diferencial fue estudiado en el tema anterior y se trata de la derivada de la función incógnita dependiendo de la variable independiente, $\frac{dy}{dx} = f(x)$, que se resuelve integrando

función f respecto la variable x , $y(x) = \int f(x) dx$.

Otro caso de ecuación diferencial “resoluble” es el conocido como *variables separables* donde la derivada de la variable dependiente es igual al producto de una función de la variable independiente por una función de la variable dependiente, es decir, $\frac{dy}{dx} = f(x)g(y)$. Para resolverla se pasan todos los términos

con la variable dependiente a un lado de la igualdad y los de la independiente al otro, $\frac{dy}{g(y)} = f(x) dx$.

Seguidamente, se integra dicha expresión

$$\int \frac{dy}{g(y)} = \int f(x) dx.$$

Despejando y en función de x se obtiene la solución buscada.

Ejemplo 8.17. Resuelve la ecuación diferencial $\frac{dy}{dx} = \alpha y$ con α cualquier real.

Aplicamos el método descrito anteriormente. Para la ecuación $\frac{dy}{dx} = \alpha y$ despejamos “todo” lo relacionado con “ y ” a un miembro y lo relacionado con “ x ” al otro, incluyendo en el “todo” a los diferenciales. (No tiene mucho “sentido” separar el cociente $\frac{dy}{dx}$ es dos términos, pues en realidad no es un cociente. De todas formas es posible justificar este paso de una forma rigurosa.)

$$\frac{dy}{y} = \alpha dx$$

Integramos y obtenemos $\ln(y) = \alpha x + C_0$. De nuevo volvemos a ser poco “riguroso” en este paso, ya que la integral de $1/x$ es $\ln|x|$ y no $\ln(x)$. Recomendamos saltarse dicho detalle, pues en este tipo de método es un engorro innecesario. Por último, despejamos la función y en función x

$$y(x) = e^{\alpha x + C_0} = e^{C_0} e^{\alpha x} = C_1 e^{\alpha x}$$

donde llamamos $C_1 = e^{C_0}$.

En economía muchas de sus leyes se establecen mediante ecuaciones diferenciales. Uno de los ejemplos más conocidos es el del crecimiento malthusiano de una población formulado por el economista Thomas Robert Malthus (1766-1837) sostiene que la velocidad con la que crece una población es proporcional a la cantidad de la población, lo que se puede escribir mediante una ecuación diferencial $\frac{dP}{dt} = kP$ donde P es número de individuos de la población, t es el tiempo (medido en años) y k es la tasa proporcional de variación de la población por años.

Por el ejercicio 8.17 tenemos que $P(t) = Ce^{kt}$ siendo $P(0) = Ce^0 = C$, luego $P(t) = P(0)(e^k)^t$. Obsérvese que se trata de una fórmula análoga a la del cálculo del capital de un depósito, ya que en ambas situaciones la tasa proporcional de variación es constante.

Ejemplo 8.18. La población de España a uno de enero de 2000 era de 40499791 habitantes y al cabo de cinco años (el uno de enero de 2005) la población era de 44108530. Si la población de España tuviese un crecimiento malthusiano, ¿cuándo alcanzaríamos los 50 millones de habitantes?

Sea t el tiempo medido en años desde el uno de enero de 2000 y $P(t)$ la población en el instante t . Tenemos que $P(0) = 40499791$ y $P(5) = 44108530$ son conocidos y se pide encontrar τ tal que $P(\tau) = 50000000$. Por lo visto en el ejemplo 8.17 $P(t) = Ce^{kt}$, tenemos que $C = P(0)$ y, por consiguiente, $P(5) = P(0)e^{5k}$, con lo que al despejar se obtiene que

$$k = \frac{1}{5}(\ln(P(5)) - \ln(P(0)))$$

Al ser $P(\tau) = 50000000 = P(0)e^{k\tau}$, despejando se obtiene que

$$\tau = \frac{1}{k}(\ln(50000000) - \ln(P(0))) = 5 \frac{\ln(50000000) - \ln(P(0))}{\ln(P(5)) - \ln(P(0))} = 12.34390482538,$$

es decir, a mediados de abril de 2012 la población alcanzaría los 50 millones con este modelo de crecimiento.

Ejemplo 8.19. Encuentre las funciones que tienen elasticidad constante σ .

Recordemos que la definición de elasticidad es $\sigma = \sigma_{xy} = \frac{x}{y} \frac{dy}{dx}$. Despejamos de forma conveniente

$\frac{dy}{y} = \sigma \frac{dx}{x}$ e integrando resulta $\ln(y) = \sigma \ln(x) + C$. Por último, despejamos y en función de x

$$y(x) = \exp(\sigma \ln(x) + C) = \gamma x^\sigma$$

donde la constante γ es $\exp(C)$.

Ejemplo 8.20. Encontrar la solución de la ecuación $\frac{dy}{dx} = y - y^2$.

Despejamos lo relacionado con x a un miembro y lo relacionado con la y al otro.

$$\frac{dy}{y(1-y)} = dx.$$

Integramos, para ello descomponemos en fracciones simples $\frac{1}{y(1-y)} = \frac{A}{y} + \frac{B}{y-1}$ y encontramos que $A = 1$ y $B = -1$, con lo que al integrar obtenemos

$$\ln(y) - \ln(y-1) = x + C.$$

Despejamos la variable y en función de x

$$\ln\left(\frac{y}{1-y}\right) = x + C; \quad \frac{y}{1-y} = e^{x+C}; \quad \frac{1-y}{y} = Ke^{-x}; \quad 1-x = Ke^x x; \quad (Ke^{-x} + 1)y = 1.$$

La solución general viene dada por

$$y(x) = \frac{1}{1 + Ke^{-x}}.$$

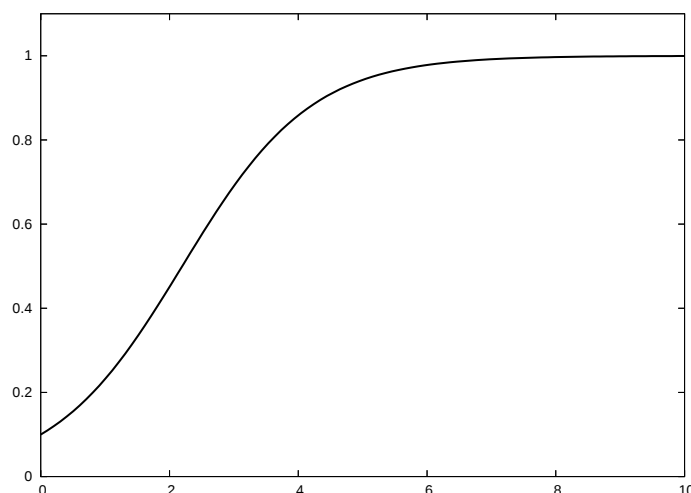


Figura 8.5: Gráfica la función $y(x) = \frac{1}{1 + e^{-x}}$, una de las soluciones de la ecuación diferencial $\frac{dy}{dx} = y - y^2$. Si consideramos el modelo logístico, el máximo sostenible de población es 1, es decir, la variable y no da el número de habitantes, sino la razón entre el número de habitantes actual y el máximo sostenible.

Según la teoría malthusiana la población crece a ritmo exponencial lo que ocasionaba el agotamiento de los recursos, lo que conllevaba que surgiesen epidemias, guerras y otros tipos de calamidades. El matemático Pierre François Verhulst propuso un modelo diferente para el crecimiento de las poblaciones donde la tasa proporcional de variación depende de los recursos existentes. El modelo más simple supone que la tasa proporcional de variación es lineal respecto a la razón entre el margen de la población para crecer y la población máxima sostenible. En términos de ecuaciones,

$$\frac{dP}{dt} = kP\left(1 - \frac{P}{M}\right)$$

donde k es la tasa de crecimiento intrínseca y M es el máximo de población sostenible. La ecuación anterior se trata de una ecuación de variables separables que es totalmente análoga al ejemplo 8.20. Como se observa en la figura 8.5 la población tiende a valores cercanos al máximo sostenible.

8.7. Definición de funciones logaritmos una base cualquiera.

Se define el logaritmo en base a positiva y distinto de uno por $\log_a(x) = \frac{\ln(x)}{\ln(a)}$.

Como señalamos al inicio del tema cuando no se escribe ninguna base se sobreentiende que es el número e y es lo que se llama logaritmo natural o neperiano. Algunas bases muy utilizadas tienen nombres: para la base 10 se llama logaritmo *decimal* o *vulgar* y para la base 2 se llama logaritmo *binario*.

Teorema 8.8. Propiedades de la función logarítmica:

1. es la función recíproca de a^x $a^{\log_a(x)} = x$ para $x > 0$ y $\log_a(a^x) = x$ para todo x real;
2. $\log_a a = 1$;
3. $\log_a(xy) = \log_a(x) + \log_a(y)$;
4. $\log_a(x/y) = \log_a(x) - \log_a(y)$;
5. $\log_a(x^y) = y \log_a(x)$;
6. $\log_a(x) = \frac{\log_b(x)}{\log_b(a)}$ tal que $b \neq 1$;
7. $\frac{d}{dx}[\log_a(x)] = \frac{1}{x \ln(a)}$.

Demostración. Es directa para todos los apartados. Por ejemplo, para demostrar las tres primeras:

$$\begin{aligned}a^{\log_a(x)} &= \exp\left(\frac{\ln(x)}{\ln(a)} \ln(a)\right) = \exp(\ln(x)) = x, \\ \log_a(a^x) &= \frac{\ln(\exp(x \ln(a)))}{\ln(a)} = \frac{x \ln(a)}{\ln(a)} = x, \\ \log_a(a) &= \frac{\ln(a)}{\ln(a)} = 1, \\ \log_a(xy) &= \frac{\ln(xy)}{\ln(a)} = \frac{\ln(x) + \ln(y)}{\ln(a)} = \frac{\ln(x)}{\ln(a)} + \frac{\ln(y)}{\ln(a)} = \log_a(x) + \log_a(y)\end{aligned}$$

Demostrar el resto de las propiedades es aplicar las propiedades del logaritmo y exponencial natural de forma conveniente. $\square$

Capítulo 9

Técnicas de integración

9.1. Integración por partes

- Conocemos que

$$d(f(x)g(x)) = g(x)f'(x)dx + f(x)g'(x)dx$$

es decir

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

luego

$$\int g(x)f'(x)dx = f(x)g(x) - \int f(x)g'(x)dx + C$$

- **Ejemplo:** Calcular

$$\int xe^x dx.$$

- Solución: Tomando

$$f(x) = e^x, g(x) = x$$

entonces

$$f'(x) = e^x, g'(x) = 1$$

y la anterior fórmula queda como

$$\begin{aligned}\int xe^x dx &= \int g(x)f'(x)dx = \\ &= f(x)g(x) - \int f(x)g'(x)dx \\ &= e^x x - \int e^x \times 1 dx \\ &= e^x x - \int e^x dx \\ &= e^x x - e^x + C\end{aligned}$$

- **Ejemplo:** Calcular

$$\int x \cos x dx.$$

- Solución: Tomando

$$g(x) = x, f'(x) = \cos x$$

tenemos

$$g'(x) = 1, f(x) = \sin x.$$

Entonces

$$\begin{aligned} \int x \cos x dx &= x \sin x - \int \sin x \times 1 dx \\ &= x \sin x + \int -\sin x dx \\ &= x \sin x + \cos x + C \end{aligned}$$

9.2. Cambio de variable

- Consideremos el cálculo de la primitiva

$$\int (\sqrt{1-x^2}) dx$$

si escribimos

$$x = \sin t \Rightarrow dx = d(\sin t) = \cos t dt,$$

luego

$$\int (\sqrt{1-x^2}) dx = \int (\sqrt{1-\sin^2 t}) \cos t dt.$$

como $\sin^2 t + \cos^2 t = 1$, despejando $\cos t$ obtenemos

$$\cos t = \sqrt{1-\sin^2 t}.$$

Finalmente,

$$\int (\sqrt{1-x^2}) dx = \int \cos^2 t dt.$$

- Vamos a calcular

$$\int \cos^2 t dt = \int \cos t \cos t dt$$

mediante integración por partes. Entonces

$$\begin{aligned} \int \cos^2 t dt &= \int \cos t \cos t dt \\ &= \sin t \cos t - \int \sin t (-\sin t) dt \\ &= \sin t \cos t + \int \sin^2 t \\ &= \sin t \cos t + \int (1 - \cos^2 t) dt \\ &= \sin t \cos t + \int 1 dt - \int \cos^2 t dt \\ &= \sin t \cos t + t - \int \cos^2 t dt. \end{aligned}$$

Pasando la integral de $\cos^2 t$ al otro extremo de la igualdad obtenemos

$$2 \int \cos^2 t \, dt = \sin t \cos t + t,$$

despejando

$$\int \cos^2 t \, dt = \frac{1}{2} (\sin t \cos t + t).$$

- Como $x = \sin t$, entonces $t = \arcsin x$, y podemos escribir

$$\begin{aligned} \int (\sqrt{1-x^2}) \, dx &= \int \cos^2 t \, dt \\ &= \frac{1}{2} (\sin t \cos t + t) \\ &= \frac{1}{2} (\sin t \sqrt{1-\sin^2 t} + t) \\ &= \frac{1}{2} (x \sqrt{1-x^2} + \arcsin x) \end{aligned}$$

9.3. Integración por sustitución

- Supongamos que la función que queremos integrar es del tipo

$$\int f' [g(x)] g'(x) \, dx$$

como por ejemplo

$$\int \frac{\sin [\log x]}{x} dx,$$

donde

$$f'(x) = \sin x, g(x) = \log x, g'(x) = \frac{1}{x}.$$

- Nótese que

$$d(f[g(x)]) = f'[g(x)] g'(x) \, dx$$

entonces integrando

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

- **Ejemplo:** Calcular

$$\int \frac{\sin [\log x]}{x} dx.$$

- Solución: Como

$$d(-\cos [\log x]) = \sin [\log x] \frac{1}{x} dx,$$

entonces

$$(-\cos [\log x]) + C = \int \sin [\log x] \frac{1}{x} dx.$$

- **Ejemplo:** Calcular

$$\int 3x^2 \cos x^3 dx.$$

- Solución: En este caso

$$d(\sin x^3) = 3x^2 \cos x^3 dx$$

entonces

$$\sin x^3 + C = \int 3x^2 \cos x^3 dx$$

- **Ejemplo:** Calcular

$$\int \cos^2 x \sin x dx.$$

- Solución: En este caso

$$d(\cos^3 x) = -3 \cos^2 x \sin x dx$$

entonces

$$(\cos^3 x) + C = -3 \int \cos^2 x \sin x dx$$

luego

$$\int \cos^2 x \sin x dx = -\left(\frac{1}{3} \cos^3 x + \frac{C}{3}\right).$$

- Se le llama por substitución por que se emplea el cambio de variable

$$u = g(x),$$

entonces

$$du = g'(x) dx$$

luego

$$\begin{aligned} \int f'[g(x)] g'(x) dx &= \int f'(u) du \\ &= f(u) + C \\ &= f(g(x)) + C. \end{aligned}$$

9.4. Integración de funciones racionales

- Consideremos A y p constantes reales y un número natural $k \geq 2$, entonces la primitiva siguiente:

$$\begin{aligned} \int \frac{A}{(x-p)^k} dx &= \int A(x-p)^{-k} dx \\ &= \left(\frac{A}{1-k}\right) (x-p)^{1-k} \\ &= \frac{A}{(1-k)(x-p)^{k-1}} \end{aligned}$$

- Consideremos A y p constantes reales y $k = 1$ entonces la primitiva siguiente:

$$\int \frac{A}{(x-p)} dx = A \log |x-p|.$$

- Consideremos C , D , q y r constantes reales y un número natural $k \geq 1$. Entonces la integral

$$\int \frac{Cx + D}{[(x - q)^2 + r^2]^k} dx$$

se reduce al cálculo de las integrales

$$\int \frac{dx}{[(x - q)^2 + r^2]^k}$$

y

$$\int \frac{(x - q)}{[(x - q)^2 + r^2]^k} dx$$

- La primera integral se reduce mediante $y = x - q$ a

$$\begin{aligned} \int \frac{(x - q)}{[(x - q)^2 + r^2]^k} dx &= \int \frac{dy}{(y^2 + r^2)^k} \\ &= \int \frac{r^{2k}}{\left(\left(\frac{y}{r}\right)^2 + 1\right)^k} dy \\ &= r^{2k+1} \int \frac{dz}{(z^2 + 1)^k}, \end{aligned}$$

tomando $z = y/r \Rightarrow (dz = r^{-1} dy \Leftrightarrow r dz = dy)$.

- Para resolver esta última integral, consideremos en primer lugar el caso $k = 1$

$$\int \frac{dz}{(z^2 + 1)} = \arctan z.$$

- El caso $k = 2$ se resuelve por integración por partes

$$\begin{aligned} \int \frac{dz}{(z^2 + 1)^2} &= \int \frac{1 + z^2 - z^2}{(z^2 + 1)^2} dz \\ &= \int \frac{z^2 + 1}{(z^2 + 1)^2} dz - \int \frac{z^2}{(1 + z^2)^2} dz \\ &= \int \frac{1}{(z^2 + 1)} dz - \int \frac{z^2}{(1 + z^2)^2} dz \\ &= \arctan z - \int \frac{z^2}{(1 + z^2)^2} dz. \end{aligned}$$

Finalmente, la integral que resta se resuelve por partes, teniendo en cuenta que

$$\frac{d}{dx} \left(\frac{1}{(1 + z^2)} \right) = -2 \left(\frac{z}{(1 + z^2)^2} \right),$$

de donde

$$\begin{aligned}
 \int \frac{z^2}{(1+z^2)^2} dz &= \int \left(\frac{z}{(1+z^2)^2} \right) z dz \\
 &= -\frac{1}{2} \left(\frac{1}{(1+z^2)} \right) z - \int \left(-\frac{1}{2} \frac{dz}{(1+z^2)} \right) \\
 &= -\frac{1}{2} \left(\frac{1}{(1+z^2)} \right) z + \frac{1}{2} \int \frac{dz}{(1+z^2)} \\
 &= -\frac{1}{2} \left(\frac{1}{(1+z^2)} \right) z + \frac{1}{2} \arctan z
 \end{aligned}$$

■ En resumen,

$$\begin{aligned}
 \int \frac{dz}{(z^2+1)^2} &= \arctan z - \left(-\frac{z}{2(1+z^2)} + \frac{1}{2} \arctan z \right) \\
 &= \frac{1}{2} \arctan z + \frac{z}{2(1+z^2)}.
 \end{aligned}$$

■ La segunda de las integrales, es decir,

$$\int \frac{(x-q)}{[(x-q)^2+r^2]^k} dx$$

la calculamos empleando el cambio de variable

$$y = (x-q)^2 + r^2 \Rightarrow dy = 2(x-q)dx$$

entonces

$$\begin{aligned}
 \int \frac{(x-q)dx}{[(x-q)^2+r^2]^k} &= \int \frac{dy}{2y^k} \\
 &= \frac{1}{2} \int \frac{dy}{y^k},
 \end{aligned}$$

entonces

$$\int \frac{dy}{y^k} = \begin{cases} \frac{1}{1-k} \frac{1}{y^{k-1}} & si \quad k \geq 2 \\ \log |y| & si \quad k = 1. \end{cases}$$

En resumen,

$$\int \frac{(x-q)dx}{[(x-q)^2+r^2]^k} = \begin{cases} \frac{1}{2} \left(\frac{1}{1-k} \right) \frac{1}{[(x-q)^2+r^2]^{k-1}} & si \quad k \geq 2 \\ \frac{1}{2} \log[(x-q)^2+r^2] & si \quad k = 1. \end{cases}$$

■ **Definición:** Llamaremos fracción parcial a una expresión de la forma

$$\frac{Cx+D}{[(x-q)^2+r^2]^k}$$

y

$$\frac{1}{[(x-q)^2 + r^2]^k}$$

o

$$\frac{(x-q)}{[(x-q)^2 + r^2]^k}$$

donde C , D , q y r son constantes reales y $k \geq 1$ es un número natural cualquiera.

Teorema 9.1. Sea $P(x)/Q(x)$, una función racional cualquiera, siendo $P(x)$ y $Q(x)$ dos polinomios cualesquiera y donde el grado del polinomio del denominador estrictamente mayor que el del numerador. Entonces esta función se puede escribir como una suma de fracciones parciales cuyos denominadores son los factores del polinomio $Q(x)$.

■ **Ejercicio:** Calcular

$$\int \frac{x-1}{(x-2)^2(x-3)} dx$$

■ **Solución:** Consideremos el escribir

$$\frac{x-1}{(x-2)^2(x-3)} = \frac{A}{(x-2)^2} + \frac{B}{x-2} + \frac{C}{x-3}.$$

Tenemos que calcular cuál es el valor de las constantes A , B y C . Para ello

$$\begin{aligned} \frac{x-1}{(x-2)^2(x-3)} &= \frac{A}{(x-2)^2} + \frac{B}{x-2} + \frac{C}{x-3} \\ &= \frac{A(x-3)}{(x-2)^2(x-3)} + \frac{B(x-2)(x-3)}{(x-2)^2(x-3)} + \frac{C(x-2)^2}{(x-2)^2(x-3)} \\ &= \frac{(B+C)x^2 + (A-5B-4C)x + (-3A+6B+4C)}{(x-2)^2(x-3)}. \end{aligned}$$

Igualando numeradores se tiene que cumplir

$$\begin{aligned} B+C &= 0 \\ A-5B-4C &= 1 \\ -3A+6B+4C &= -1. \end{aligned}$$

Resolviendo este sistema de ecuaciones obtenemos

$$A = -1, B = -2, C = 2.$$

En consecuencia

$$\begin{aligned} \int \frac{x-1}{(x-2)^2(x-3)} dx &= \int \frac{-1dx}{(x-2)^2} + \int \frac{-2dx}{x-2} + \int \frac{2dx}{x-3} \\ &= \frac{1}{(x-2)} - 2 \log |x-2| + 2 \log |x-3|. \end{aligned}$$

■ **Ejercicio:** Calcular

$$\int \frac{x-1}{(x^2+1)^2(x+1)} dx.$$

■ **Solución:** Consideremos

$$\frac{x-1}{(x^2+1)^2(x+1)} = \frac{Ax+B}{(x^2+1)^2} + \frac{Cx+D}{x^2+1} + \frac{E}{x+1}.$$

Reduciendo a común denominador obtenemos la igualdad entre numeradores

$$(Ax+B)(x+1) + (Cx+D)(x^2+1)(x+1) + E(x^2+1)^2 = x-1,$$

de donde obtenemos

$$A = 1, \quad B = 0, \quad C = 1/2, \quad D = -1/2, \quad E = -1/2.$$

En consecuencia

$$\begin{aligned} & \int \frac{(x-1)dx}{(x^2+1)^2(x+1)} \\ &= \int \frac{xdx}{(x^2+1)^2} + \int \frac{((1/2)x-1/2)dx}{x^2+1} + \int \frac{(-1/2)dx}{x+1}. \end{aligned}$$

Ahora bien

$$\int \frac{xdx}{(x^2+1)^2} = -\frac{1}{2} \cdot \frac{1}{x^2+1},$$

$$\begin{aligned} \int \frac{((1/2)x-1/2)dx}{x^2+1} &= \frac{1}{2} \int \frac{xdx}{x^2+1} - \frac{1}{2} \int \frac{dx}{x^2+1} \\ &= \frac{1}{2} \left(-\frac{1}{2} \cdot \frac{1}{x^2+1} \right) - \frac{1}{2} \arctan x. \end{aligned}$$

y

$$\int \frac{(-1/2)dx}{x+1} = -\frac{1}{2} \log |x+1|.$$

Capítulo 10

Sistemas de ecuaciones lineales

10.1. Introducción

Consideremos el siguiente sistema lineal de ecuaciones:

$$\begin{array}{rrcr} 2x_1 & -x_2 & +5x_3 & = & 8 \\ x_1 & & -4x_3 & = & -7 \end{array}$$

el objetivo es encontrar los valores x_1, x_2 y x_3 de forma que al sustituirlos en ambas ecuaciones se den las igualdades pertinentes.

Vamos a encontrar dos codificaciones de este problema que nos permitan no solo resolverlo sino poder interpretar las soluciones.

1. Codificación geométrica

$$x_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ -4 \end{bmatrix} = \begin{bmatrix} 8 \\ -7 \end{bmatrix}.$$

2. Codificación algebraica:

$$\begin{bmatrix} 2 & -1 & 5 \\ 1 & 0 & -4 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ -7 \end{bmatrix},$$

que también podemos escribir como

$$\left[\begin{array}{ccc|c} 2 & -1 & 5 & 8 \\ 1 & 0 & -4 & -7 \end{array} \right].$$

A los objetos que aparecen en las expresiones anteriores como

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 8 \\ -7 \end{bmatrix}$$

se les llama **vectores**, y a los objetos

$$\left[\begin{array}{ccc|c} 2 & -1 & 5 & 8 \\ 1 & 0 & -4 & -7 \end{array} \right], \left[\begin{array}{ccc|c} 2 & -1 & 5 & 8 \\ 1 & 0 & -4 & -7 \end{array} \right].$$

se le llama **matrices**.

10.1.1. Vectores y sus operaciones algebraicas

Una matriz con una única columna se le llama **vector columna**, o simplemente **vector**.

$$\mathbf{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

Al conjunto de vectores con dos entradas se denota por $\mathbb{R}^2$.

Dos vectores de $\mathbb{R}^2$ son **iguales** si sus correspondientes entradas coinciden:

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} \neq \begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix}$$

Dados dos vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^2$, el vector suma $\mathbf{u} + \mathbf{v}$ se obtiene sumando las correspondientes entradas de ambos, esto es,

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} + \begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix} = \begin{bmatrix} 3 + 0.2 \\ -1 + 0.3 \end{bmatrix}$$

Dado un vector $\mathbf{u}$ y un número real cualquiera x , el **múltiplo escalar de $\mathbf{u}$ por x** es el vector $x\mathbf{u}$ que se obtiene multiplicando cada entrada de $\mathbf{u}$ por x , esto es,

$$\sqrt{2} \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 3\sqrt{2} \\ -\sqrt{2} \end{bmatrix}$$

Al número x en este contexto se le llama **escalar**.

Al igual que el cero es el elemento neutro para la suma de números reales, el vector cero

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

es el elemento neutro para la suma de vectores

$$\mathbf{0} + \mathbf{v} = \mathbf{v} + \mathbf{0} = \mathbf{v}.$$

Ejemplo 10.1. Dados

$$\mathbf{u} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \text{ y } \mathbf{v} = \begin{bmatrix} 2 \\ -5 \end{bmatrix},$$

encuentra $4\mathbf{u}$, $(-3)\mathbf{v}$ y $4\mathbf{u} + (-3)\mathbf{v}$.

$$4\mathbf{u} = \begin{bmatrix} 4 \\ -8 \end{bmatrix}$$

$$(-3)\mathbf{v} = \begin{bmatrix} -6 \\ 15 \end{bmatrix}$$

$$4\mathbf{u} + (-3)\mathbf{v} = \begin{bmatrix} 4 \\ -8 \end{bmatrix} + \begin{bmatrix} -6 \\ 15 \end{bmatrix} = \begin{bmatrix} -2 \\ 7 \end{bmatrix}.$$

Observación 10.1. Obsérvese que

$$\begin{bmatrix} 4 \\ -8 \end{bmatrix} \neq (4, -8)$$

$$\begin{bmatrix} 4 \\ -8 \end{bmatrix} \neq \begin{bmatrix} 4 & -8 \end{bmatrix}$$

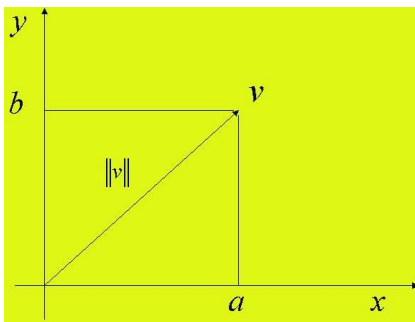
La notación $(4, -8)$ se emplea **por comodidad en la escritura**.

Descripciones geométricas de $\mathbb{R}^2$

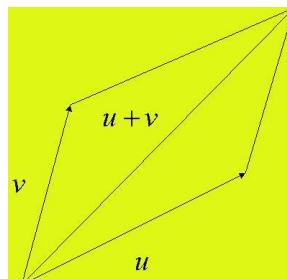
Identificaremos el punto geométrico (a, b) del plano coordenado con el vector

$$\begin{bmatrix} a \\ b \end{bmatrix}.$$

Visualización geométrica:



Regla del paralelogramo para la suma: Si u y v se representan como puntos del plano, entonces $u + v$ corresponde al cuarto vértice del paralelogramo cuyos otros vértices son u , 0 y v .

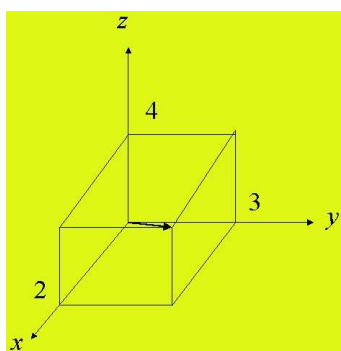


Vectores en $\mathbb{R}^3$

Los vectores en $\mathbb{R}^3$ son matrices columnas 3×1 :

$$\mathbf{a} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}.$$

Visualización geométrica:

**Vectores en $\mathbb{R}^n$**

Si n es un entero positivo (o número natural) $\mathbb{R}^n$ denota la colección de todas las listas (o n -uplas ordenadas) de n -números reales, que usualmente se escriben como matrices columna $n \times 1$:

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Propiedades algebraicas de $\mathbb{R}^n$: Para todo $\mathbf{u}, \mathbf{v}$ y $\mathbf{w}$ en $\mathbb{R}^n$ y cualesquiera escalares x y y :

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
3. $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$.
4. $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$, donde $-\mathbf{u}$ denota a $(-1)\mathbf{u}$.
5. $x(\mathbf{u} + \mathbf{v}) = x\mathbf{u} + x\mathbf{v}$.
6. $(x + y)\mathbf{u} = x\mathbf{u} + y\mathbf{u}$.
7. $(xy)\mathbf{u} = x(y\mathbf{u})$.
8. $1\mathbf{u} = \mathbf{u}$.

Observación 10.2. Por simplicidad

$$\mathbf{u} + (-1)\mathbf{v} \triangleq \mathbf{u} - \mathbf{v}.$$

Combinaciones lineales

Dados los vectores

$$\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p$$

en $\mathbb{R}^n$ y los escalares

$$x_1, x_2, \dots, x_p$$

al vector $\mathbf{b}$ definido como

$$\mathbf{b} \triangleq x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_p \mathbf{a}_p$$

se le llama **combinación lineal de $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p$** empleando los **pesos** $x_1, x_2, \dots, x_p$.

Ejemplo 10.2. Sean

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} \quad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}.$$

Determina cuando $\mathbf{b}$ puede estar generado (o bien puede escribirse) como una combinación lineal de $\mathbf{a}_1$ y $\mathbf{a}_2$. Esto es, determina qué pesos x_1 y x_2 existen de forma que

$$x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 = \mathbf{b}.$$

Sustituimos:

$$x_1 \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

Empleamos la multiplicación de vectores por escalares:

$$\begin{bmatrix} x_1 \\ -2x_1 \\ -5x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 5x_2 \\ 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

Empleamos la suma de vectores:

$$\begin{bmatrix} x_1 + 2x_2 \\ -2x_1 + 5x_2 \\ -5x_1 + 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

Empleamos que dos vectores son iguales si lo son componente a componente:

$$\begin{array}{rcl} x_1 & +2x_2 & = 7 \\ -2x_1 & +5x_2 & = 4 \\ -5x_1 & +6x_2 & = -3 \end{array}$$

Observación 10.3. Podemos afirmar entonces que los problemas

$$x_1 \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \tag{10.1}$$

y

$$\begin{array}{rcl} x_1 & +2x_2 & = 7 \\ -2x_1 & +5x_2 & = 4 \\ -5x_1 & +6x_2 & = -3 \end{array} \tag{10.2}$$

son equivalentes. En consecuencia si el sistema tiene solución podemos afirmar que el vector de términos independientes

$$\begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

es combinación lineal de los vectores

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} \quad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}.$$

¿Cómo podemos ligar ahora la interpretación obtenida con la codificación geométrica de un sistema lineal con la codificación algebraica?

10.1.2. Notación matricial

Los coeficientes del sistema lineal

$$\begin{array}{rrcr} x_1 & -2x_2 & +x_3 & = & 0 \\ & 2x_2 & -8x_3 & = & 8 \\ -4x_1 & +5x_2 & +9x_3 & = & -9 \end{array}$$

se pueden escribir como la **matriz de coeficientes**

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

y la **matriz aumentada del sistema**

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array} \right]$$

Una matriz $\mathbf{m} \times \mathbf{n}$ es una formación rectangular de números con m -filas y n -columnas. Diremos entonces que su **tamaño** es $m \times n$. El tamaño de una matriz nos dice cuantas filas y columnas tiene.

Empleando la notación vectorial y considerando que

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} 1 \\ -8 \\ 9 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ 8 \\ -9 \end{bmatrix}$$

podemos representar la matriz de coeficientes como

$$A = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} \triangleq [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3],$$

y la matriz ampliada del sistema

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array} \right] \triangleq [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3 \mid \mathbf{b}] = [A \mid \mathbf{b}]$$

Vamos ahora a resolver el sistema lineal empleando un mecanismo que usa la manipulación algebraica de la matriz ampliada del sistema:

Resolución de un sistema lineal mediante el algoritmo de Gauss–Jordan

Vamos a resolver el siguiente sistema lineal, al lado del mismo escribiremos las matrices ampliadas de los sistemas que vayamos obteniendo al realizar los cálculos pertinentes.

$$\begin{array}{rrcr} x_1 & -2x_2 & +x_3 & = & 0 \\ & 2x_2 & -8x_3 & = & 8 \\ -4x_1 & +5x_2 & +9x_3 & = & -9 \end{array}, \quad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array} \right]$$

- Queremos mantener a x_1 en la primera ecuación y eliminarla de las otras ecuaciones, para ello multiplicamos la primera ecuación (fila de la matriz ampliada) por 4 y se la sumamos a la tercera ecuación (fila de la matriz ampliada):

$$\begin{array}{rrcr} x_1 & -2x_2 & +x_3 & = & 0 \\ & 2x_2 & -8x_3 & = & 8 \\ & -3x_2 & +13x_3 & = & -9 \end{array}, \quad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{array} \right]$$

- Queremos que el coeficiente de x_2 en la segunda ecuación sea uno, para ello multiplicamos toda la misma por $1/2$:

$$\begin{array}{rrcr} x_1 & -2x_2 & +x_3 & = & 0 \\ & x_2 & -4x_3 & = & 4 \\ & -3x_2 & +13x_3 & = & -9 \end{array}, \quad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{array} \right]$$

- Queremos mantener x_2 en la segunda ecuación y eliminarla de las otras ecuaciones que hay por debajo de esta. Para ello multiplicamos por 3 la segunda ecuación y se la sumamos a la tercera:

$$\begin{array}{rrcr} x_1 & -2x_2 & +x_3 & = & 0 \\ & x_2 & -4x_3 & = & 4 \\ & & x_3 & = & 3 \end{array}, \quad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{array} \right]$$

- Queremos eliminar x_3 de las ecuaciones que están por encima de la tercera ecuación. Para ello en primer lugar multiplicamos por 4 la tercera ecuación y se la sumamos a la segunda:

$$\begin{array}{rrcr} x_1 & -2x_2 & +x_3 & = & 0 \\ & x_2 & & = & 16 \\ & & x_3 & = & 3 \end{array}, \quad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array} \right]$$

En segundo lugar multiplicamos por -1 la tercera ecuación y se la sumamos a la primera:

$$\begin{array}{rrcr} x_1 & -2x_2 & & = & -3 \\ & x_2 & & = & 16 \\ & & x_3 & = & 3 \end{array}, \quad \left[\begin{array}{ccc|c} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array} \right]$$

- Queremos eliminar x_2 de la ecuación que está por encima de la segunda ecuación, para ello multiplicamos por 2 la segunda ecuación y se la sumamos a la primera:

$$\left\{ \begin{array}{rrcr} x_1 & & & = & 29 \\ & x_2 & & = & 16 \\ & & x_3 & = & 3 \end{array} \right., \quad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array} \right]$$

Observación 10.4. Podemos ahora afirmar que el vector formado por los términos independientes de sistema

$$\mathbf{b} = \begin{bmatrix} 29 \\ 16 \\ 3 \end{bmatrix}$$

es combinación lineal de los vectores columna de la matriz de coeficientes A .

¿Qué tipos de operaciones por filas (ecuaciones) están permitidas?

1. (Reemplazo) Reemplazar una fila por la suma de sí misma y un múltiplo de una fila.
2. (Intercambio) Intercambiar dos filas.
3. (Escalamiento) Multiplicar todas las entradas de una fila por una constante diferente de cero.

Decimos que dos matrices son **equivalentes por filas** si existe una sucesión de operaciones que transforman una matriz en otra.

Podemos afirmar entonces que:

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{array} \right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array} \right]$$

Observación 10.5. Todas las operaciones por filas son invertibles.

Teorema 10.1. Si las matrices aumentadas de dos sistemas lineales son equivalentes por filas, entonces los dos sistemas tienen el mismo conjunto de soluciones.

¿Qué tipos de conjuntos de soluciones puede tener un sistema lineal?

10.2. La ecuación matricial $A\mathbf{x} = \mathbf{b}$

■ Definición

- Si A es una matriz $m \times n$, con columnas $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ y si $\mathbf{x}$ está en $\mathbf{R}^n$ entonces el producto de A por $\mathbf{x}$ que denotamos por

$$A\mathbf{x}$$

es la **combinación lineal de las columnas de A utilizando como pesos las correspondientes entradas de $\mathbf{x}$** , esto es,

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n.$$

■ Ejemplo:

1.

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7 \begin{bmatrix} -1 \\ 3 \end{bmatrix} \\ = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

2.

$$\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = 4 \begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7 \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} \\ = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$$

- Conocemos que el sistema

$$\begin{array}{rrcr} x_1 & +2x_2 & -x_3 & = & 4 \\ & -5x_2 & +3x_3 & = & 1 \end{array}$$

es equivalente a

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ -5 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

o bien en forma matricial

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

- Esta última ecuación tiene la forma

$$A\mathbf{x} = \mathbf{b}$$

y llamaremos a una ecuación así, **ecuación matricial**.

Teorema 10.2. Si A es una matriz $m \times n$, con columnas $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ y si $\mathbf{b}$ está en $\mathbf{R}^n$, la ecuación matricial

$$A\mathbf{x} = \mathbf{b}$$

tiene el mismo conjunto de soluciones que la ecuación vectorial

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = \mathbf{b}$$

la cual, a su vez tiene el mismo conjunto solución que el sistema lineal de ecuaciones cuya matriz aumentada es

$$\left[\begin{array}{cccc|c} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n & \mathbf{b} \end{array} \right]$$

10.2.1. Existencia de soluciones

- La ecuación $A\mathbf{x} = \mathbf{b}$ tiene solución si y solo si $\mathbf{b}$ es una combinación lineal de las columnas de A .

- **Aplicación:** Sean

$$A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & 6 \\ -3 & -2 & -7 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

¿Es consistente la ecuación $A\mathbf{x} = \mathbf{b}$ para todos los posibles valores de b_1, b_2 y b_3 ?

- *Solución:*

$$\begin{aligned} \left[\begin{array}{cccc|c} 1 & 3 & 4 & b_1 \\ -4 & 2 & 6 & b_2 \\ -3 & -2 & -7 & b_3 \end{array} \right] & \sim \left[\begin{array}{cccc|c} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 7 & 5 & b_3 + 3b_1 \end{array} \right] \\ & \sim \left[\begin{array}{cccc|c} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 0 & 0 & b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) \end{array} \right] \end{aligned}$$

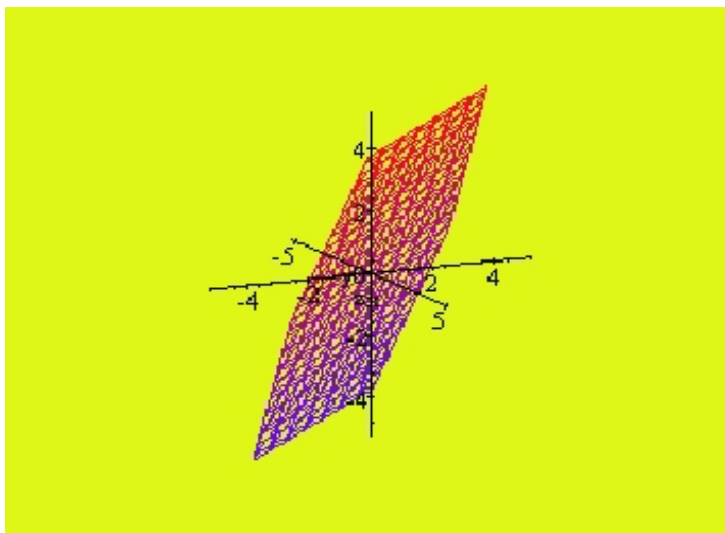
En consecuencia el sistema es consistente si y solo si

$$b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) = 0,$$

esto es,

$$b_1 - \frac{1}{2}b_2 + b_3 = 0$$

donde esto representa la ecuación de un plano que pasa por el origen.



Teorema 10.3. Sea A una matriz $m \times n$. Entonces las siguientes afirmaciones son lógicamente equivalentes. Es decir, para una matriz A en particular, todas estas afirmaciones son verdad o todas ellas son falsas.

1. Para cada $\mathbf{b}$ en $\mathbf{R}^m$ la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución.
2. Las columnas de A generan $\mathbf{R}^m$.
3. A tiene una posición pivote en cada fila.

10.2.2. Cálculo de $A\mathbf{x}$

- Si está definido el producto $A\mathbf{x}$, entonces la entrada i -ésima de $A\mathbf{x}$ es la suma de los productos de las entradas correspondientes de la fila i de A y del vector $\mathbf{x}$.

$$\begin{aligned}
 \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & 6 \\ -3 & -2 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= x_1 \begin{bmatrix} 1 \\ -4 \\ -3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 2 \\ -2 \end{bmatrix} + x_3 \begin{bmatrix} 4 \\ 6 \\ -7 \end{bmatrix} \\
 &= \begin{bmatrix} x_1 \\ -4x_1 \\ -3x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ 2x_2 \\ -2x_2 \end{bmatrix} + \begin{bmatrix} 4x_3 \\ 6x_3 \\ -7x_3 \end{bmatrix} \\
 &= \begin{bmatrix} x_1 + 3x_2 + 4x_3 \\ -4x_1 + 2x_2 + 6x_3 \\ -3x_1 - 2x_2 - 7x_3 \end{bmatrix}
 \end{aligned}$$

La primera fila de $A\mathbf{x}$ es

$$\begin{bmatrix} 1 & 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 + 3x_2 + 4x_3,$$

la segunda fila

$$\begin{bmatrix} -4 & 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -4x_1 + 2x_2 + 6x_3$$

y la tercera fila

$$\begin{bmatrix} -3 & -2 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -3x_1 - 2x_2 - 7x_3.$$

■ Definición

- Se llama **matriz identidad**, y se denota por I a la matriz $n \times n$ que tiene unos en la diagonal y cero en el resto de entradas.

10.2.3. Propiedades del producto matriz–vector $A\mathbf{x}$

Teorema 10.4. Si A es una matriz $m \times n$, $\mathbf{u}$ y $\mathbf{v}$ son vectores de $\mathbf{R}^n$ y c es un escalar, entonces

1. $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$.
2. $A(c\mathbf{u}) = c(A\mathbf{u})$.

■ Demostración

Consideremos $n = 3$.

$$\begin{aligned} A(\mathbf{u} + \mathbf{v}) &= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{bmatrix} \\ &= \mathbf{a}_1(u_1 + v_1) + \mathbf{a}_2(u_2 + v_2) + \mathbf{a}_3(u_3 + v_3) \\ &= (\mathbf{a}_1 u_1 + \mathbf{a}_2 u_2 + \mathbf{a}_3 u_3) + (\mathbf{a}_1 v_1 + \mathbf{a}_2 v_2 + \mathbf{a}_3 v_3) \\ &= A\mathbf{u} + A\mathbf{v}. \end{aligned}$$

y

$$\begin{aligned} A(c\mathbf{u}) &= \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{bmatrix} \begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix} \\ &= \mathbf{a}_1(cu_1) + \mathbf{a}_2(cu_2) + \mathbf{a}_3(cu_3) \\ &= c(\mathbf{a}_1 u_1 + \mathbf{a}_2 u_2 + \mathbf{a}_3 u_3) \\ &= c(A\mathbf{u}). \end{aligned}$$

QED

10.3. Conjuntos solución de sistemas lineales

10.3.1. Sistemas lineales homogéneos

Un sistema se dice **homogéneo** si es de la forma

$$A\mathbf{x} = \mathbf{0}$$

donde

1. A es una matriz $m \times n$,
2. $\mathbf{0}$ es el vector cero de $\mathbf{R}^m$.

■ $\mathbf{x} = \mathbf{0}$ siempre es solución del sistema homogéneo. A esta solución se le llama **solución trivial**.

Como consecuencia del Theorem 2 obtenemos:

- La ecuación homogénea $A\mathbf{x} = \mathbf{0}$ tiene una solución distinta de la trivial si y solo si la ecuación tiene por lo menos una variable libre.
- **Ejemplo:** Determine si el sistema homogéneo tiene una solución no trivial. Luego describa el conjunto solución

$$\begin{array}{rrcr} 3x_1 & +5x_2 & -4x_3 & = 0 \\ -3x_1 & -2x_2 & +4x_3 & = 0 \\ 6x_1 & +x_2 & -8x_3 & = 0 \end{array}$$

• *Solución:*

$$\begin{aligned} \begin{bmatrix} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & -8 & 0 \end{bmatrix} &\sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &\sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 0 & -4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 0 & -4/3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ (RREF)} \\ \begin{array}{rrcr} x_1 & -\frac{4}{3}x_3 & = & 0 \\ x_2 & & = & 0 \\ & 0 & = & 0 \end{array} &\left\{ \begin{array}{l} x_1 = \frac{4}{3}x_3 \\ x_2 = 0 \\ x_3 = \text{libre} \end{array} \right. \end{aligned}$$

La solución general

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$$

- **Ejemplo:** Describa todas las soluciones del sistema homogéneo

$$10x_1 - 3x_2 - 2x_3 = 0$$

• *Solución:*

$$\begin{cases} x_1 = \frac{3}{10}x_2 + \frac{1}{5}x_3 \\ x_2 = \text{libre} \\ x_3 = \text{libre} \end{cases}$$

La solución general

$$\begin{aligned} \mathbf{x} &= \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{3}{10}x_2 + \frac{1}{5}x_3 \\ x_2 \\ x_3 \end{bmatrix} \\ &= \begin{bmatrix} \frac{3}{10}x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{5}x_3 \\ 0 \\ x_3 \end{bmatrix} \\ &= x_2 \begin{bmatrix} \frac{3}{10} \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} \frac{1}{5} \\ 0 \\ 1 \end{bmatrix} \end{aligned}$$

Esto nos dice que

$$\mathbf{x} \in \text{Gen} \left\{ \begin{bmatrix} \frac{3}{10} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{5} \\ 0 \\ 1 \end{bmatrix} \right\}$$

10.3.2. Forma vectorial paramétrica

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} \frac{3}{10} \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} \frac{1}{5} \\ 0 \\ 1 \end{bmatrix} = x_2 \mathbf{u} + x_3 \mathbf{v}, x_2, x_3 \in \mathbf{R}.$$

Se le llama **forma vectorial paramétrica** de la solución.

10.3.3. Soluciones de sistemas no homogéneos

- **Ejemplo:** Describa todas las soluciones de $A\mathbf{x} = \mathbf{b}$, donde

$$A = \begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 7 \\ -1 \\ 4 \end{bmatrix}$$

- *Solución:*

$$\begin{bmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -4/3 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{cases} x_1 - \frac{4}{3}x_3 = -1 \\ x_2 = 2 \\ 0 = 0 \end{cases}$$

entonces la solución en forma paramétrica

$$\begin{aligned} \mathbf{x} &= \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 + \frac{4}{3}x_3 \\ 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} = \mathbf{p} + x_3 \mathbf{v} \end{aligned}$$

Se puede comprobar que

$$\mathbf{x} = x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$$

es la solución del sistema homogéneo

$$A\mathbf{x} = \mathbf{0}$$

y que

$$\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

cumple

$$\begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \\ 4 \end{bmatrix}$$

es decir es una **solución particular** de

$$A\mathbf{x} = \mathbf{b}.$$

Teorema 10.5. Suponga que la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para algún $\mathbf{b}$ dado y sea $\mathbf{p}$ una solución dada. Entonces el conjunto solución de $A\mathbf{x} = \mathbf{b}$ es el conjunto de vectores de la forma

$$\mathbf{w} = \mathbf{p} + \mathbf{v}_h$$

donde $\mathbf{v}_h$ es cualquier solución del sistema homogéneo

$$A\mathbf{x} = \mathbf{0}.$$

10.3.4. Un sistema homogéneo en economía

Carbón	Electricidad	Acero	Comprado por:
0.0	0.4	0.6	Carbón
0.6	0.1	0.2	Electricidad
0.4	0.5	0.2	Acero

p_C = precio de la producción anual de Carbón,

p_E = precio de la producción anual de Electricidad,

p_A = precio de la producción anual de Acero,

$$\text{Ingresos} = \text{Gastos}$$

$$p_C = +0.4p_E + 0.6p_A$$

$$p_E = 0.6p_C + 0.1p_E + 0.2p_A$$

$$p_A = 0.4p_C + 0.5p_E + 0.2p_A$$

$$0 = -p_C + 0.4p_E + 0.6p_A$$

$$0 = 0.6p_C - 0.9p_E + 0.2p_A$$

$$0 = 0.4p_C + 0.5p_E - 0.8p_A$$

Finalmente

$$\begin{bmatrix} -1 & 0.4 & 0.6 & 0 \\ 0.6 & -0.9 & 0.2 & 0 \\ 0.4 & 0.5 & -0.8 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -0.93939 & 0 \\ 0 & 1 & -0.84848 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución general es

$$\mathbf{p} = \begin{bmatrix} p_C \\ p_E \\ p_A \end{bmatrix} = p_A \begin{bmatrix} 0.93939 \\ 0.84848 \\ 1 \end{bmatrix}.$$

10.4. Independencia lineal

■ DEFINICIÓN

- Un conjunto de vectores indexado

$$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$$

en $\mathbf{R}^n$ es **linealmente independiente** si la ecuación vectorial

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_p\mathbf{v}_p = \mathbf{0}$$

tiene únicamente la solución trivial $\mathbf{x} = \mathbf{0}$. El conjunto

$$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$$

es **linealmente dependiente** si existen pesos

$$c_1, c_2, \dots, c_p$$

no todos cero tales que

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_p \mathbf{v}_p = \mathbf{0}.$$

A esta última ecuación se le llama **relación de dependencia lineal** entre $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$.

■ **Ejemplo:** Sean

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

Se pide (a) Determine si el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente. (b) Si no es posible encuentre una relación de dependencia lineal entre $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

• *Solución:*

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Hay variables libres (el sistema tiene solución distinta de la trivial).

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{array}{rcl} x_1 & -2x_3 & = 0 \\ x_2 & +x_3 & = 0 \\ & 0 & = 0 \end{array}$$

$$\begin{array}{rcl} x_1 & & = 2x_3 \\ x_2 & & = -x_3 \\ x_3 & & = \text{libre} \end{array}$$

Si elegimos

$$x_3 = 5$$

entonces

$$\begin{array}{rcl} x_1 & = & 10 \\ x_2 & = & -5 \end{array}$$

luego

$$10 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - 5 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + 5 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

10.4.1. Independencia lineal entre las columnas de una matriz

- Sea una matriz

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n].$$

La ecuación matricial

$$A\mathbf{x} = \mathbf{0}$$

se puede escribir

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$$

- Las columnas de una matriz son linealmente independientes si y solo si la ecuación

$$A\mathbf{x} = \mathbf{0}$$

tiene la solución trivial.

- **Ejemplo:** Determine si las columnas de

$$A = \begin{bmatrix} 0 & 1 & 4 \\ 1 & 2 & -1 \\ 5 & 8 & 0 \end{bmatrix}$$

son linealmente independientes.

- *Solución:*

$$\begin{bmatrix} 0 & 1 & 4 & 0 \\ 1 & 2 & -1 & 0 \\ 5 & 8 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 13 & 0 \end{bmatrix}$$

Como no hay variables libres la única solución es la trivial. En consecuencia las columnas son linealmente independientes.

10.4.2. Conjuntos de uno o dos vectores

- Un conjunto formado por un único vector

$$\{\mathbf{v}\}, \mathbf{v} \neq \mathbf{0}$$

es linealmente independiente

$$x\mathbf{v} = \mathbf{0} \implies x = 0.$$

- Un conjunto de vectores

$$\{\mathbf{v}_1, \mathbf{v}_2\}, \mathbf{v}_1, \mathbf{v}_2 \neq \mathbf{0}$$

el linealmente dependiente si y solo si un vector es múltiplo del otro.

- Demostración:

linealmente dependientes

$\Updownarrow$

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}, c_1 \neq 0, c_2 \neq 0$$

(nótese que si $c_1 = 0$ entonces

$$c_2\mathbf{v}_2 = \mathbf{0} \implies \mathbf{v}_2 = \mathbf{0} \text{ (una contradicción)}).$$

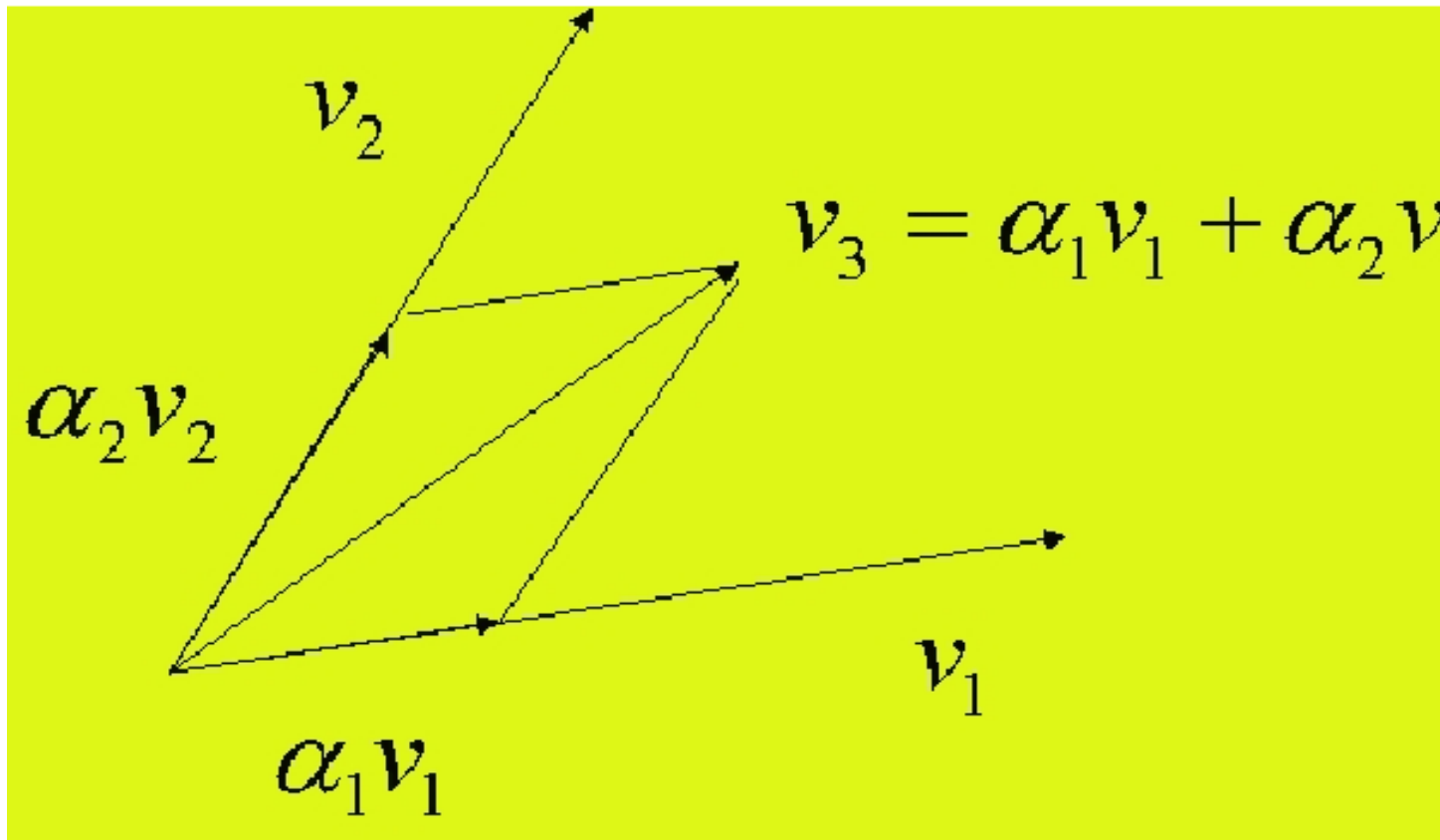
$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0}, c_1 \neq 0, c_2 \neq 0$$

$\Updownarrow$

$$c_1\mathbf{v}_1 = -c_2\mathbf{v}_2$$

$\Updownarrow$

$$\mathbf{v}_1 = -\frac{c_2}{c_1}\mathbf{v}_2.$$



10.4.3. Conjuntos de dos o más vectores

Teorema 10.6. Un conjunto indexado

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$$

de dos o más vectores es linealmente dependiente si y solo si al menos uno de los vectores en S es combinación lineal de los otros. De hecho, si S es linealmente dependiente y si $\mathbf{v}_1 \neq \mathbf{0}$, entonces algún $\mathbf{v}_j$ (con $j > 1$) es una combinación lineal de los vectores precedentes $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}$.

■ Gráficamente

Teorema 10.7. Si un conjunto tiene más vectores que entradas en cada vector, entonces el conjunto es linealmente dependiente. Esto es, cualquier conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$ en $\mathbf{R}^n$ es linealmente dependiente si

$$p > n.$$

Demostración:

$$n \begin{bmatrix} \blacksquare & * & * & * & * \\ * & \blacksquare & * & * & * \\ * & * & \blacksquare & * & * \end{bmatrix}^p$$

$p > n$ (número máximo de pivotes)

$\Updownarrow$

Existen variables libres

Teorema 10.8. Si un conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$ en $\mathbf{R}^n$ contiene el vector cero, entonces el conjunto es linealmente dependiente.

Demostración: Supongamos que

$$\mathbf{v}_1 = \mathbf{0}$$

entonces

$$c_1 \mathbf{v}_1 + 0 \mathbf{v}_2 + \dots + 0 \mathbf{v}_p = \mathbf{0}, \quad c_1 \neq 0.$$

10.5. Subespacios de $\mathbf{R}^n$

- Un subespacio de $\mathbf{R}^n$ es cualquier subconjunto H de vectores de $\mathbf{R}^n$ que cumple tres propiedades:

1. El vector cero está en H .
2. Si $\mathbf{u}$ y $\mathbf{v}$ están en H entonces $\mathbf{u} + \mathbf{v}$ está en H .
3. Para cada $\mathbf{u}$ en H y cualquier escalar c , el vector $c\mathbf{u}$ está en H .

- **Ejemplo:** Consideremos dos vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ de $\mathbf{R}^n$. Sea

$$H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\} = \{\mathbf{y} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 : c_1, c_2 \in \mathbf{R}\}.$$

1. $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2$ está en H .
2. Si $\mathbf{u} = c_1^u \mathbf{v}_1 + c_2^u \mathbf{v}_2$ y $\mathbf{v} = c_1^v \mathbf{v}_1 + c_2^v \mathbf{v}_2$ están en H . Entonces,

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (c_1^u \mathbf{v}_1 + c_2^u \mathbf{v}_2) + (c_1^v \mathbf{v}_1 + c_2^v \mathbf{v}_2) \\ &= (c_1^u + c_1^v) \mathbf{v}_1 + (c_2^u + c_2^v) \mathbf{v}_2 \end{aligned}$$

está en H .

3. Si $\mathbf{u} = c_1^u \mathbf{v}_1 + c_2^u \mathbf{v}_2$ y c es un escalar

$$c\mathbf{u} = c(c_1^u \mathbf{v}_1 + c_2^u \mathbf{v}_2) = (cc_1^u) \mathbf{v}_1 + (cc_2^u) \mathbf{v}_2$$

está en H .

- El conjunto de todas las combinaciones lineales de vectores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ de $\mathbf{R}^n$ es un subespacio de $\mathbf{R}^n$. Esto es

$$H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\} \text{ es un subespacio de } \mathbf{R}^n.$$

10.5.1. Espacio de columnas y espacio nulo de una matriz

- **Definición:** El **espacio de columnas** de una matriz A $m \times n$, es el conjunto $\text{Col } A$ formado por todas las combinaciones lineales de las columnas de A . Esto es: si

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix}$$

entonces

$$\text{Col } A = \text{Gen}\{ \mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n \}.$$

- Diremos que el vector $\mathbf{b}$ de $\mathbf{R}^m$ está en $\text{Col } A$ si y solo si el sistema lineal

$$A\mathbf{x} = \mathbf{b}$$

tiene solución.

- **Definición:** El **espacio nulo** de una matriz A es el conjunto $\text{Nul } A$ formado por todas las soluciones de la ecuación homogénea

$$A\mathbf{x} = \mathbf{0}.$$

Teorema 10.9. El espacio nulo de una matriz A $m \times n$ es un subespacio de $\mathbf{R}^n$. De manera equivalente, el conjunto de todas las soluciones del sistema

$$A\mathbf{x} = \mathbf{0}.$$

de m -ecuaciones lineales homogéneas con n -incógnitas es un subespacio de $\mathbf{R}^n$.

- **Demostración:**

1. $\mathbf{0} \in \text{Nul } A$; $A\mathbf{0} = \mathbf{0}$.
2. $\mathbf{u}, \mathbf{v} \in \text{Nul } A$; $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$, entonces $\mathbf{u} + \mathbf{v} \in \text{Nul } A$.
3. $\mathbf{u} \in \text{Nul } A$ y $c \in \mathbf{R}$; $A(c\mathbf{u}) = c(A\mathbf{u}) = c\mathbf{0} = \mathbf{0}$, entonces $c\mathbf{u} \in \text{Nul } A$.

10.5.2. Bases para un subespacio

- **Definición:** Una **base** para un subespacio H de $\mathbf{R}^n$ es un conjunto linealmente independiente que genera cualquier vector de H .

- **Ejemplo:**

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

forman una base de $\mathbf{R}^n$ ya que

$$I_n = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \dots \quad \mathbf{e}_n]$$

e

$$I_n \mathbf{x} = \mathbf{x}.$$

- **Ejemplo:** Encuentre una base para el espacio nulo de la matriz

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & 4 & 5 & 8 & -4 \end{bmatrix}$$

- Solución:

$$\begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Entonces

$$\begin{array}{rrrrr} x_1 & -2x_2 & & -x_4 & +3x_5 & = & 0 \\ & & x_3 & +2x_4 & -2x_5 & = & 0 \\ & & & & 0 & = & 0 \end{array}$$

obtenemos la solución general

$$\begin{array}{l} x_1 = -2x_2 + x_4 - 3x_5 \\ x_2 = \text{libre} \\ x_3 = -2x_4 + 2x_5 \\ x_4 = \text{libre} \\ x_5 = \text{libre.} \end{array}$$

La solución en forma vectorial

$$\begin{aligned}
 \mathbf{x} &= \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} \\
 &= \begin{bmatrix} -2x_2 \\ x_2 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} x_4 \\ 0 \\ -2x_4 \\ x_4 \\ 0 \end{bmatrix} + \begin{bmatrix} -3x_5 \\ 0 \\ +2x_5 \\ 0 \\ x_5 \end{bmatrix} \\
 &= x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}
 \end{aligned}$$

Entonces

$$\text{Nul } A = \text{Gen} \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Basta comprobar que son linealmente independientes. Consideramos la matriz

$$\begin{bmatrix} -2 & 1 & -3 \\ 1 & 0 & 0 \\ 0 & -2 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

y se puede comprobar que tiene tres pivotes.

- Encontrar una base del espacio columna de una matriz necesita menos trabajo que encontrar una base del espacio nulo.
- **Ejemplo:** Encuentre una base para el espacio columna de la matriz

$$B = \begin{bmatrix} 1 & 0 & -3 & 5 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- Solución: Sea

$$B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \quad \mathbf{b}_4 \quad \mathbf{b}_5].$$

Entonces

$$\begin{aligned}
 -3\mathbf{b}_1 + 2\mathbf{b}_2 &= \mathbf{b}_3, \quad 5\mathbf{b}_1 + (-1)\mathbf{b}_2 = \mathbf{b}_4 \\
 (-3) \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} &= \begin{bmatrix} -3 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \quad 5 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + (-1) \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \\ 0 \\ 0 \end{bmatrix}
 \end{aligned}$$

Además

$$\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

son linealmente independientes. En consecuencia

$$\text{Col } A = \text{Gen } \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5\}.$$

Teorema 10.10. Las columnas pivote de una matriz A forman una base para el espacio columna de A .

10.5.3. La dimensión de un subespacio

- **Definición:** La **dimensión** de un subespacio H diferente de cero, denotada por $\dim H$, es el número de vectores que hay en cualquier base de H . La dimensión del subespacio cero $\{\mathbf{0}\}$ es por definición cero.

- **Ejemplo:**

$$\dim (\text{Col } A) = \dim \left(\text{Col } \begin{bmatrix} 1 & 0 & -3 & 5 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right) = 3.$$

- **Definición:** El **rango** de una matriz A , denotado por $\text{rango } A$, es la dimensión del espacio de columnas de A .

- **Ejemplo:**

$$\text{rango} \left(\begin{bmatrix} 1 & 0 & -3 & 5 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \right) = 3.$$

Teorema 10.11 (Teorema del rango). Si una matriz A tiene n columnas entonces

$$n = \text{rango } A + \dim \text{Nul } A$$

- Notese que

$$\text{columnas de } A = \text{variables principales} + \text{variables libres.}$$

Teorema 10.12 (Teorema de la base). Sea H un espacio p dimensional de $\mathbf{R}^n$. Cualquier conjunto linealmente independiente de exactamente p elementos en H , es automáticamente una base de H . De igual modo, cualquier conjunto de p elementos que genere H es automáticamente una base de H .

Capítulo 11

Operaciones con matrices. Inversa de una matriz

11.1. La ecuación matricial de un sistema lineal

Recordemos que podemos escribir el sistema lineal de ecuaciones:

$$\begin{array}{rrcr} 2x_1 & -x_2 & +5x_3 & = & 8 \\ x_1 & & -4x_3 & = & -7 \end{array}$$

como

$$\begin{bmatrix} 2 & -1 & 5 \\ 1 & 0 & -4 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ -4 \end{bmatrix} = \begin{bmatrix} 8 \\ -7 \end{bmatrix},$$

o bien

$$\begin{bmatrix} 2 & -1 & 5 \\ 1 & 0 & -4 \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \times x_1 + (-1) \times x_2 + 5 \times x_3 \\ 1 \times x_1 + 0 \times x_2 + (-4) \times x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ -7 \end{bmatrix},$$

Si la matriz de coeficientes es

$$A = \begin{bmatrix} 2 & -1 & 5 \\ 1 & 0 & -4 \end{bmatrix},$$

el vector de términos independientes se denota por

$$\mathbf{b} = \begin{bmatrix} 8 \\ -7 \end{bmatrix},$$

y finalmente el vector de incógnitas lo escribimos como

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

el sistema lineal se puede expresar en forma matricial como:

$$A\mathbf{x} = \mathbf{b}.$$

Observación 11.1. Esto nos permite definir una nueva operación algebraica, que es la multiplicación de una matriz de 2×3 por un vector 3×1 . En general, deberíamos poder mutiplicar una matriz $m \times n$ por un vector $n \times 1$.

Ejemplo 11.1.

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7 \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

Ejemplo 11.2.

$$\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = 4 \begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7 \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$$

Ejemplo 11.3.

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \times 4 + 2 \times 3 + (-1) \times 7 \\ 0 \times 4 + (-5) \times 3 + 3 \times 7 \end{bmatrix}$$

$$= \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

Ejemplo 11.4.

$$\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 2 \times 4 + (-3) \times 7 \\ 8 \times 4 + 0 \times 7 \\ (-5) \times 4 + 2 \times 7 \end{bmatrix}$$

$$= \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$$

11.1.1. La multiplicación de matrices

Consideremos una matriz $m \times n$

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n],$$

esto quiere decir que los vectores columna están en $\mathbb{R}^m$. Sabemos, dado cualquier vector $\mathbf{b} \in \mathbb{R}^n$ multiplicar

$$A\mathbf{b}$$

el resultado es un vector de $\mathbb{R}^m$. Tomemos pues una matriz $m \times p$ que podemos escribir como

$$B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_p]$$

y cuyos vectores columna están en $\mathbb{R}^n$. Sabemos multiplicar

$$A\mathbf{b}_1, \quad A\mathbf{b}_2, \quad \cdots, \quad A\mathbf{b}_p.$$

Empleando este argumento podemos definir el producto de dos matrices una de tamaño $m \times n$ por una de tamaño $n \times p$ de la forma siguiente:

$$AB \triangleq [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p]$$

el resultado es una matriz $m \times p$.

Ejemplo 11.5.

$$A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}_{2 \times 2}, \quad B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}_{2 \times 3}$$

$$AB = \left[\underbrace{\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix}}_{\text{columna1}} \quad \underbrace{\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix}}_{\text{columna2}} \quad \underbrace{\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \end{bmatrix}}_{\text{columna3}} \right]$$

Realizamos estas operaciones aparte:

$$\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 11 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 13 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 21 \\ -9 \end{bmatrix}.$$

Escribimos el resultado final

$$AB = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

Nótese que cada columna de la matriz producto AB es combinación lineal de las columnas de la matriz A . La matriz AB tiene el mismo número de filas que A y el mismo número de columnas que B .

Si denotamos por

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{i1} & \cdots & b_{ij} & \cdots & b_{ip} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{np} \end{bmatrix}$$

Regla fila-columna para calcular AB : Si $(AB)_{ij}$ denota la entrada de la fila i columna j de la matriz AB , entonces

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

$$= \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} \triangleq \mathbf{A}_{i,\bullet} \bullet \mathbf{B}_{\bullet,j},$$

donde $\mathbf{A}_{i,\bullet}$ representa el i -ésimo vector fila de la matriz A y $\mathbf{B}_{\bullet,j}$ representa el j -ésimo vector columna de la matriz B . Esto permite definir una nueva operación entre vectores, llamada **producto escalar**. Si

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

entonces

$$\mathbf{u} \bullet \mathbf{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

Ejemplo 11.6. Calcule $\mathbf{u} \bullet \mathbf{v}$ y $\mathbf{v} \bullet \mathbf{u}$ cuando

$$\mathbf{u} = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$$

$$\mathbf{u} \bullet \mathbf{v} = \begin{bmatrix} 2 & 5 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = 2 \times 3 + 5 \times 2 + (-1) \times (-3) = 19$$

$$\mathbf{v} \bullet \mathbf{u} = \begin{bmatrix} 3 & 2 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} = 2 \times 3 + 5 \times 2 + (-1) \times (-3) = 19$$

Es evidente que el producto escalar de dos vectores es conmutativo

$$\mathbf{u} \bullet \mathbf{v} = \mathbf{v} \bullet \mathbf{u}.$$

Además, empleando el producto escalar podemos escribir la multiplicación de matrices como

$$AB = \begin{bmatrix} \mathbf{A}_{1,\bullet} \bullet \mathbf{B}_{\bullet,1} & \mathbf{A}_{1,\bullet} \bullet \mathbf{B}_{\bullet,2} & \cdots & \mathbf{A}_{1,\bullet} \bullet \mathbf{B}_{\bullet,p} \\ \mathbf{A}_{2,\bullet} \bullet \mathbf{B}_{\bullet,1} & \mathbf{A}_{2,\bullet} \bullet \mathbf{B}_{\bullet,2} & \cdots & \mathbf{A}_{2,\bullet} \bullet \mathbf{B}_{\bullet,p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{m,\bullet} \bullet \mathbf{B}_{\bullet,1} & \mathbf{A}_{m,\bullet} \bullet \mathbf{B}_{\bullet,2} & \cdots & \mathbf{A}_{m,\bullet} \bullet \mathbf{B}_{\bullet,p} \end{bmatrix}.$$

11.1.2. La suma de matrices y la multiplicación de una matriz por un escalar

Recordemos estas dos propiedades de los vectores y empleemoslas para definir la suma de matrices. Sean

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \text{ matriz } m \times n$$

y

$$B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{bmatrix} \text{ matriz } m \times n$$

dos matrices $m \times n$ declaradas por columnas. Definimos entonces la matriz suma

$$A + B \triangleq \begin{bmatrix} (\mathbf{a}_1 + \mathbf{b}_1) & (\mathbf{a}_2 + \mathbf{b}_2) & \cdots & (\mathbf{a}_n + \mathbf{b}_n) \end{bmatrix} \text{ matriz } m \times n.$$

Sea x un escalar, definimos entonces

$$xA \triangleq \begin{bmatrix} x\mathbf{a}_1 & x\mathbf{a}_2 & \cdots & x\mathbf{a}_n \end{bmatrix} \text{ matriz } m \times n.$$

Ejemplo 11.7. Sean

$$A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} \quad C = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$$

Entonces

$$A + B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}.$$

Para el caso

$$A + C = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} + \begin{bmatrix} 2 & -3 & \text{falta una columna} \\ 0 & 1 & \end{bmatrix},$$

nótese, tal como se indica que falta una columna para que podamos realizar la operación. Finalmente,

$$\sqrt{2}B = \begin{bmatrix} \sqrt{2} & \sqrt{2} & \sqrt{2} \\ 3\sqrt{2} & 5\sqrt{2} & 7\sqrt{2} \end{bmatrix}$$

Definimos la matriz cero como

$$0 = \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & & 0 \end{bmatrix} = [\mathbf{0} \quad \mathbf{0} \quad \cdots \quad \mathbf{0}]$$

Teorema 11.1. Sean A, B y C matrices del mismo tamaño, y sean r y s escalares.

1. $A + B = B + A$.
2. $(A + B) + C = A + (B + C)$.
3. $A + 0 = A$.
4. $r(A + B) = rA + rB$.
5. $(r + s)A = rA + sA$.
6. $r(sA) = (rs)A$.

11.2. Propiedades de la multiplicación de matrices

El producto de matrices depende de la posición en la que se multiplican las mismas. Si A es una matriz 3×2 y B es una matriz 2×2 , entonces solo se pueden multiplicar como AB , ya que BA no tiene sentido. Para que esta situación no se de, necesitamos que sean compatibles

$$A_{(n \times p)} B_{(q \times r)}$$

y

$$B_{(q \times r)} A_{(n \times p)},$$

en consecuencia, $p = q$ y $r = n$, Entonces tendríamos que AB en una matriz $n \times n$ y BA una matriz $q \times q$. Si $n \neq q$ entonces AB tiene distinto tamaño que BA excepto que $p = q = r = n$.

Veamos que incluso en el caso de matrices $n \times n$ la propiedad conmutativa no se cumple.

Ejemplo 11.8. Sean

$$A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$$

Comprobemos que estas matrices no conmutan, esto es, $AB \neq BA$.

$$\begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix}$$

Definimos la matriz identidad $m \times m$ como

$$I_m = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Teorema 11.2. Sea A una matriz $m \times n$ y sean B y C dos matrices con los tamaños adecuados para que las sumas y productos estén definidos.

1. $A(BC) = (AB)C$.
2. $A(B + C) = AB + AC$.
3. $(B + C)A = BA + CA$.
4. $r(AB) = (rA)B = A(rB)$ para cualquier escalar r .
5. $I_m A = A = A I_n$.

11.3. La inversa de una matriz

Si A es una matriz $n \times n$ y existe una matriz C de forma que

$$AC = I_n, \quad CA = I_n,$$

diremos que C es la **matriz inversa** de A .

Si B fuera otra inversa para la matriz A , entonces cumpliría que

$$AB = I_n, \quad BA = I_n,$$

ahora bien

$$B = BI_n = B(AC) = (BA)C = I_n C = C.$$

Si A tiene inversa diremos que A es **invertible o no singular** y denotaremos a su inversa por A^{-1} , entonces

$$AA^{-1} = I_n, \quad A^{-1}A = I_n.$$

Si una matriz no tiene inversa se dice que es **singular**.

Ejemplo 11.9.

$$\begin{aligned} A &= \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}, \quad C = \begin{bmatrix} -7 & -5 \\ 3 & -2 \end{bmatrix}. \\ AC &= \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} \begin{bmatrix} -7 & -5 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ CA &= \begin{bmatrix} -7 & -5 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ C &= \begin{bmatrix} -7 & -5 \\ 3 & -2 \end{bmatrix} = A^{-1}. \end{aligned}$$

Teorema 11.3. Sea

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Si $ad - bc \neq 0$, entonces A es invertible y

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

Si $ad - bc = 0$ entonces es no invertible.

A la cantidad $ad - bc$ se le llama **determinante de la matriz A** , esto es,

$$\det A = ad - bc.$$

Ejemplo 11.10.

$$A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$$

$$\det A = 18 - 20 = -2 \neq 0$$

$$A^{-1} = -\frac{1}{2} \begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ \frac{5}{2} & -\frac{3}{2} \end{bmatrix}.$$

Teorema 11.4. Si A es una matriz invertible $n \times n$, entonces para cada $\mathbf{b}$ en $\mathbf{R}^n$, la ecuación

$$A\mathbf{x} = \mathbf{b}$$

tiene solución única

$$\mathbf{x} = A^{-1}\mathbf{b}.$$

Demostración.

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \implies A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b} \\ \implies (A^{-1}A)\mathbf{x} &= A^{-1}\mathbf{b} \implies I_n\mathbf{x} = \mathbf{x} = A^{-1}\mathbf{b}. \end{aligned}$$

■

Ejemplo 11.11.

$$\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$$

$$\begin{aligned} 3x_1 + 4x_2 &= 3 \\ 5x_1 + 6x_2 &= 7 \end{aligned}$$

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 7 \end{bmatrix} \\ &= \begin{bmatrix} -3 & 2 \\ \frac{5}{2} & -\frac{3}{2} \end{bmatrix} \begin{bmatrix} 3 \\ 7 \end{bmatrix} \\ &= \begin{bmatrix} 5 \\ -3 \end{bmatrix}. \end{aligned}$$

11.4. El cálculo de la inversa de una matriz

Ejemplo 11.12. Encuentra la matriz inversa de

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}.$$

Para calcular A^{-1} necesito construir una matriz

$$\begin{bmatrix} x & p & u \\ y & q & v \\ z & r & w \end{bmatrix}$$

de forma que

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} x & p & u \\ y & q & v \\ z & r & w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Si nos fijamos para resolver este problema necesitamos resolver los siguientes sistemas lineales:

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

Para ello tenemos que emplear el algoritmo de reducción por filas a las siguientes matrices ampliadas:

$$\left[\begin{array}{ccc|c} 0 & 1 & 2 & 1 \\ 1 & 0 & 3 & 0 \\ 4 & -3 & 8 & 0 \end{array} \right].$$

.

$$\left[\begin{array}{ccc|c} 0 & 1 & 2 & 0 \\ 1 & 0 & 3 & 1 \\ 4 & -3 & 8 & 0 \end{array} \right].$$

.

$$\left[\begin{array}{ccc|c} 0 & 1 & 2 & 0 \\ 1 & 0 & 3 & 0 \\ 4 & -3 & 8 & 1 \end{array} \right].$$

Como las operaciones afectan por igual a la matriz de coeficientes podemos escribir los anteriores tres sistemas en uno de la forma

$$\left[\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array} \right].$$

y efectuar operaciones por filas a esta matriz (triplemente ampliada), lo que nos permitirá resolver los tres sistemas en uno solo.

$$\begin{aligned}
 [A \mid I_n] &= \left[\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array} \right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array} \right] \\
 &\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{array} \right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{array} \right] \\
 &\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{array} \right] \\
 &\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{array} \right] \\
 &\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{array} \right] = [I_n \mid A^{-1}].
 \end{aligned}$$

En consecuencia

$$A^{-1} = \begin{bmatrix} -\frac{9}{2} & 7 & -\frac{3}{2} \\ -2 & 4 & -1 \\ \frac{3}{2} & -2 & \frac{1}{2} \end{bmatrix}.$$

Teorema 11.5. Sea A una matriz cuadrada $n \times n$. Entonces los enunciados que siguen son todos equivalentes.

1. A es una matriz invertible.
2. A es equivalente por filas a la matriz identidad.
3. A tiene n posiciones pivote.
4. La ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial.
5. La ecuación lineal $A\mathbf{x} = \mathbf{b}$ tiene una solución única para cada $\mathbf{b}$ en $\mathbf{R}^n$.
6. Las columnas de A generan $\mathbf{R}^n$.
7. Existe una matriz C de forma que $CA = I$.
8. Existe una matriz C de forma que $AC = I$.

Capítulo 12

Problemas de mínimos cuadrados

12.1. La solución por mínimos cuadrados de un sistema lineal

El objetivo de esta lectura es intentar resolver el problema

$$A\mathbf{x} = \mathbf{b}$$

cuando $\mathbf{b}$ no es combinación lineal de las columnas de la matriz

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}.$$

El problema se reduce a plantear el sistema en forma vectorial

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b},$$

que al no tener solución, nos lleva a intentar resolver el problema

$$\min_{x_1, x_2, \dots, x_n} \text{dist}(x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n, \mathbf{b}),$$

esto es, encontrar los valores $x_1, x_2, \dots, x_n$ que hacen mínima la distancia entre el vector $\mathbf{b}$ y el conjunto generado por las columnas de la matriz A .

Se define la **longitud** o **norma** de un vector

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

como el escalar no negativo definido por

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \bullet \mathbf{v}} = \sqrt{v_1^2 + \cdots + v_n^2}$$

y donde

$$\|\mathbf{v}\|^2 = \mathbf{v} \bullet \mathbf{v} = v_1^2 + \cdots + v_n^2$$

A un vector $\mathbf{v}$ cuya longitud es igual a 1 lo llamaremos **vector unitario**.

Si dividimos a un vector $\mathbf{v}$ distinto de cero por su longitud (esto es, lo multiplicamos por

$$\frac{1}{\|\mathbf{v}\|}$$

obtenemos un vector unitario $\mathbf{u}$ que es igual a

$$\left(\frac{1}{\|\mathbf{v}\|} \right) \mathbf{v}.$$

A $\mathbf{u}$ se le llama la **normalización de $\mathbf{v}$** .

Ejemplo 12.1. Sea

$$\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix}.$$

Encuentre un vector unitario en la misma dirección que $\mathbf{v}$.

$$\|\mathbf{v}\| = \sqrt{1^2 + (-2)^2 + 2^2 + 0^2} = \sqrt{9} = 3,$$

entonces

$$\mathbf{u} = \left(\frac{1}{\|\mathbf{v}\|} \right) \mathbf{v} = \frac{1}{3} \begin{bmatrix} 1 \\ -2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \\ 0 \end{bmatrix}$$

es un vector unitario en la dirección de $\mathbf{v}$.

Para $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$ la **distancia entre $\mathbf{u}$ y $\mathbf{v}$** , escrita como

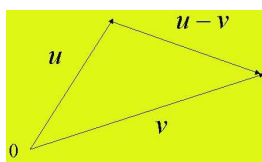
$$\text{dist}(\mathbf{u}, \mathbf{v})$$

es la longitud del vector

$$\mathbf{u} - \mathbf{v}.$$

Esto es

$$\text{dist}(\mathbf{u}, \mathbf{v}) \triangleq \|\mathbf{u} - \mathbf{v}\|$$



Ejemplo 12.2. Calcule la distancia entre los vectores

$$\mathbf{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

$$\mathbf{u} - \mathbf{v} = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

entonces

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{4^2 + 1^2} = \sqrt{17}.$$

Nótese que

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})}$$

Teorema 12.1. Sean $\mathbf{u}, \mathbf{v}$ y $\mathbf{w}$ tres vectores de $\mathbb{R}^n$ y sea c un escalar. Entonces

1. $\mathbf{u} \bullet \mathbf{v} = \mathbf{v} \bullet \mathbf{u}$.
2. $(\mathbf{u} + \mathbf{v}) \bullet \mathbf{w} = \mathbf{u} \bullet \mathbf{w} + \mathbf{v} \bullet \mathbf{w}$.
3. $(c\mathbf{u}) \bullet \mathbf{v} = \mathbf{u} \bullet (c\mathbf{v}) = c(\mathbf{u} \bullet \mathbf{v})$.
4. $\mathbf{u} \bullet \mathbf{u} \geq 0$ y $\mathbf{u} \bullet \mathbf{u} = 0$ si y solo si $\mathbf{u} = \mathbf{0}$.

Combinando las propiedades (2) y (3) se obtiene

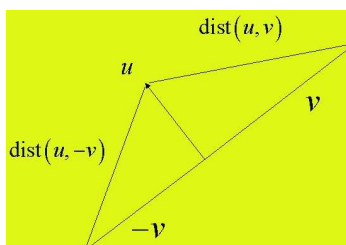
$$(c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p) \bullet \mathbf{w} = c_1(\mathbf{u}_1 \bullet \mathbf{w}) + \cdots + c_p(\mathbf{u}_p \bullet \mathbf{w}).$$

12.2. La propiedad de la ortogonalidad

Deseamos establecer una relación de forma que dado un punto identificado por un vector $\mathbf{u}$ encontrar la condición que caracterize cuál es la distancia mínima de ese punto a una recta con dirección $\mathbf{v}$.

La condición se puede expresar como que el punto identificado con el vector $\mathbf{u}$ representa la distancia óptima a la recta con dirección $\mathbf{v}$ si

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \text{dist}(\mathbf{u}, -\mathbf{v})$$



$$\begin{aligned}
 \text{dist}(\mathbf{u}, \mathbf{v})^2 &= (\mathbf{u} - \mathbf{v}) \bullet (\mathbf{u} - \mathbf{v}) = \mathbf{u} \bullet (\mathbf{u} - \mathbf{v}) - \mathbf{v} \bullet (\mathbf{u} - \mathbf{v}) \\
 &= \{(\mathbf{u} \bullet \mathbf{u}) - (\mathbf{u} \bullet \mathbf{v})\} - \{(\mathbf{v} \bullet \mathbf{u}) - (\mathbf{v} \bullet \mathbf{v})\} \\
 &= \left\{ \|\mathbf{u}\|^2 - (\mathbf{u} \bullet \mathbf{v}) \right\} - \left\{ (\mathbf{v} \bullet \mathbf{u}) - \|\mathbf{v}\|^2 \right\} \\
 &= \|\mathbf{u}\|^2 - 2(\mathbf{u} \bullet \mathbf{v}) + \|\mathbf{v}\|^2
 \end{aligned}$$

$$\begin{aligned}
 \text{dist}(\mathbf{u}, -\mathbf{v})^2 &= (\mathbf{u} - (-\mathbf{v})) \bullet (\mathbf{u} - (-\mathbf{v})) \\
 &= (\mathbf{u} + \mathbf{v}) \bullet (\mathbf{u} + \mathbf{v}) \\
 &= \mathbf{u} \bullet (\mathbf{u} + \mathbf{v}) + \mathbf{v} \bullet (\mathbf{u} + \mathbf{v}) \\
 &= \{(\mathbf{u} \bullet \mathbf{u}) + (\mathbf{u} \bullet \mathbf{v})\} + \{(\mathbf{v} \bullet \mathbf{u}) + (\mathbf{v} \bullet \mathbf{v})\} \\
 &= \left\{ \|\mathbf{u}\|^2 + (\mathbf{u} \bullet \mathbf{v}) \right\} + \left\{ (\mathbf{v} \bullet \mathbf{u}) + \|\mathbf{v}\|^2 \right\} \\
 &= \|\mathbf{u}\|^2 + 2(\mathbf{u} \bullet \mathbf{v}) + \|\mathbf{v}\|^2
 \end{aligned}$$

En consecuencia

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \text{dist}(\mathbf{u}, -\mathbf{v})$$

si y solo si

$$\|\mathbf{u}\|^2 - 2(\mathbf{u} \bullet \mathbf{v}) + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2(\mathbf{u} \bullet \mathbf{v}) + \|\mathbf{v}\|^2.$$

Es decir, si y solo si

$$-2(\mathbf{u} \bullet \mathbf{v}) = 2(\mathbf{u} \bullet \mathbf{v})$$

o bien, si y solo si

$$4(\mathbf{u} \bullet \mathbf{v}) = 0.$$

o bien

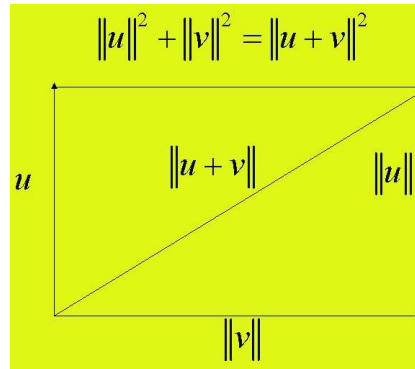
$$\mathbf{u} \bullet \mathbf{v} = 0.$$

Dos vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$ son **ortogonales** (perpendiculares, entre si) si

$$\mathbf{u} \bullet \mathbf{v} = 0.$$

Teorema 12.2 (Teorema de Pitágoras). Dos vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$ son ortogonales si y solo si

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$



12.3. El método de los mínimos cuadrados

Para obtener el vector $\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_n \end{bmatrix}$ que resuelve el problema

$$\begin{aligned} \min_{x_1, x_2, \dots, x_n} \text{dist}(x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n, \mathbf{b}) &= \\ \min_{x_1, x_2, \dots, x_n} \|x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n - \mathbf{b}\|. \end{aligned}$$

Tenemos que encontrar un vector $\hat{\mathbf{b}}$ que sea combinación lineal de las columnas de la matriz de coeficientes A , esto es,

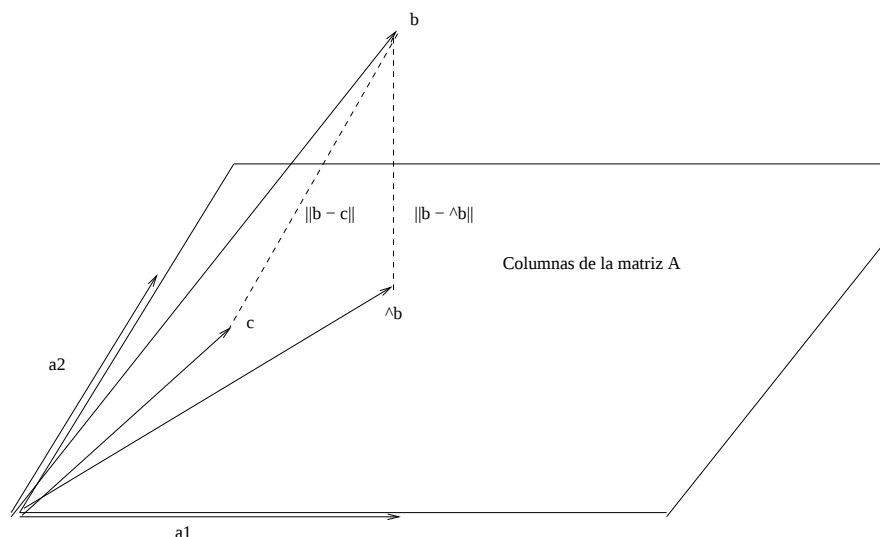
$$\hat{\mathbf{b}} = \hat{x}_1 \mathbf{a}_1 + \hat{x}_2 \mathbf{a}_2 + \dots + \hat{x}_n \mathbf{a}_n$$

de forma que se cumpla que la distancia entre el vector $\hat{\mathbf{b}}$ y el vector $\mathbf{b}$ sea la menor posible entre cualquier otro de los vectores que son combinación lineal de las columnas de la matriz A , esto es,

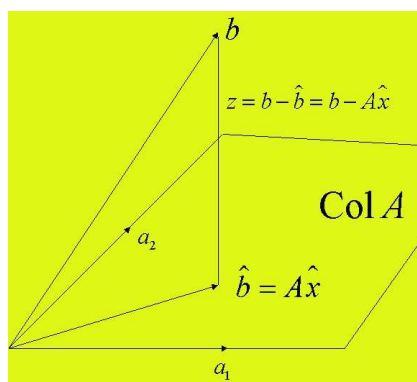
$$\|\mathbf{b} - \hat{\mathbf{b}}\| \leq \|\mathbf{b} - \mathbf{c}\|$$

para cualquier otro vector $\mathbf{c}$ que sea combinación lineal de las columnas de la matriz A , esto es

$$\mathbf{c} = x_1 \mathbf{a}_1 + x_2 \mathbf{a}_2 + \dots + x_n \mathbf{a}_n.$$



La cuestión radica en como podemos calcular este vector $\hat{\mathbf{b}}$? Si nos fijamos en la gráfica el vector $\mathbf{z} = \mathbf{b} - \hat{\mathbf{b}}$ tiene que ser ortogonal al conjunto formado por todas las combinaciones lineales de las columnas de la matriz A .



Como sabemos que $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ entonces conocemos que el vector

$$\mathbf{z} = \mathbf{b} - \hat{\mathbf{b}} = \mathbf{b} - A\hat{\mathbf{x}}$$

tiene que ser ortogonal a todas las combinaciones lineales de las columnas de la matriz A . Para encontrar esta condición necesitamos introducir el concepto siguiente.

Dada una matriz A $m \times n$, la **transpuesta de A** es la matriz $n \times m$ que denotaremos por A^T se obtiene al intercambiar las filas de A por sus columnas.

Ejemplo 12.3.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

$$B = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix} \quad B^T = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}$$

Proposición 12.1. Un vector $\mathbf{z}$ es ortogonal a todas las columnas de la matriz A si y solo si $A^T \mathbf{z} = \mathbf{0}$.

Demostración. En primer lugar fijémonos que

$$\mathbf{a} \bullet \mathbf{z} = a_1 z_1 + \cdots + a_n z_n = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = \mathbf{a}^T \mathbf{z}.$$

Entonces si

$$A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$$

se tiene que

$$A^T = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}.$$

Luego

$$A^T \mathbf{z} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix} \mathbf{z} = \begin{bmatrix} \mathbf{a}_1^T \mathbf{z} \\ \vdots \\ \mathbf{a}_n^T \mathbf{z} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 \bullet \mathbf{z} \\ \vdots \\ \mathbf{a}_n \bullet \mathbf{z} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

de donde deducimos el resultado. ■

Volviendo al problema original, como el vector $\mathbf{z} = \mathbf{b} - \hat{\mathbf{b}}$ tiene que ser ortogonal al conjunto formado por todas las combinaciones lineales de las columnas de la matriz A , se ha de cumplir que

$$A^T (\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}.$$

Multiplicando se obtiene que

$$A^T \mathbf{b} - A^T A \hat{\mathbf{x}} = \mathbf{0},$$

despejando $A^T \mathbf{b}$ se tiene que $\hat{\mathbf{x}}$ tiene que ser solución del sistema lineal

$$A^T \mathbf{b} = \mathbf{0} + A^T A \hat{\mathbf{x}} = A^T A \hat{\mathbf{x}}.$$

Al sistema lineal de ecuaciones

$$A^T A \hat{\mathbf{x}} = A^T \mathbf{b},$$

se les llama **ecuaciones normales para $\hat{\mathbf{x}}$** .

Teorema 12.3. El conjunto de soluciones por mínimos cuadrados de

$$A\mathbf{x} = \mathbf{b}$$

coincide con el conjunto no vacío de soluciones de las ecuaciones normales

$$A^T A \hat{\mathbf{x}} = A^T \mathbf{b}.$$

Ejemplo 12.4. Consideremos el sistema inconsistente

$$\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Calcula la solución de mínimos cuadrados.

Consideremos las ecuaciones normales

$$\begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

esto es,

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

entonces

$$\begin{bmatrix} 17 & 1 & 19 \\ 1 & 5 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

obtenemos como solución

$$\begin{aligned} \hat{x} &= 1 \\ \hat{y} &= 2 \end{aligned}.$$

Ejemplo 12.5. Encuentre una solución por mínimos cuadrados para el sistema inconsistente

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$

Calculamos las ecuaciones normales

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

y

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

esto es

$$\begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}.$$

Como

$$\begin{bmatrix} 6 & 2 & 2 & 2 & 4 \\ 2 & 2 & 0 & 0 & -4 \\ 2 & 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 2 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

obtenemos que la forma general de la solución es

$$\begin{cases} \hat{x}_1 &= 3 - \hat{x}_4 \\ \hat{x}_2 &= -5 + \hat{x}_4 \\ \hat{x}_3 &= -2 + \hat{x}_4 \\ \hat{x}_4 &= \text{libre} \end{cases}$$

Capítulo 13

Diagonalización. Valores y vectores propios

13.1. Valores y vectores propios

- **Definición:** Un **vector propio** de una matriz A $n \times n$ es un vector $\mathbf{x}$ distinto de cero tal que

$$A\mathbf{x} = \lambda\mathbf{x}$$

para algún escalar λ . Un escalar λ se llama **valor propio** de A si existe una solución no trivial $\mathbf{x}$ de

$$A\mathbf{x} = \lambda\mathbf{x};$$

un tal $\mathbf{x}$ se llama **vector propio correspondiente a λ** .

- **Ejemplo:** Sean

$$A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

¿Son $\mathbf{u}$ y $\mathbf{v}$ vectores propios de la matriz A ?

- Solución:

$$A\mathbf{u} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} \left(= \lambda\mathbf{u} = \begin{bmatrix} 6\lambda \\ -5\lambda \end{bmatrix} \right)$$

entonces el posible valor propio λ debe cumplir

$$\begin{bmatrix} -24 \\ 20 \end{bmatrix} = \begin{bmatrix} 6\lambda \\ -5\lambda \end{bmatrix}$$

$$-24 = 6\lambda \implies \lambda = -4$$

$$20 = -5\lambda \implies \lambda = -4.$$

En consecuencia

$$\begin{bmatrix} 6 \\ -5 \end{bmatrix} \text{ es un vector propio con valor propio asociado } \lambda = -4.$$

En el segundo caso

$$A\mathbf{v} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \left(= \lambda\mathbf{v} = \lambda \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right)$$

entonces el posible valor propio λ debe cumplir

$$\begin{aligned} -9 &= 3\lambda \implies \lambda = -3 \\ 11 &= -2\lambda \implies \lambda = -\frac{11}{2} \end{aligned}$$

como nos da dos valores distintos no puede existir un tal valor λ . Finalmente,

$$\begin{bmatrix} 3 \\ -2 \end{bmatrix} \text{ no es un vector propio.}$$

- **Ejemplo:** Demuestre que 7 es un valor propio de A en el anterior ejemplo, y encuentre los vectores propios correspondientes.

- Solución: Hay que encontrar los $\mathbf{x} = \begin{bmatrix} x & y \end{bmatrix}^T$ de forma que

$$A\mathbf{x} = 7\mathbf{x}.$$

Nótese que la expresión anterior se puede escribir como

$$A\mathbf{x} = 7I_2\mathbf{x},$$

o bien

$$A\mathbf{x} - 7I_2\mathbf{x} = \mathbf{0}$$

simplificandose esta expresión a un sistema homogéneo

$$(A - 7I_2)\mathbf{x} = \mathbf{0}.$$

Esto es

$$\left(\begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - 7 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

o bien

$$\begin{bmatrix} 1-7 & 6 \\ 5 & 2-7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

La matriz ampliada del sistema homogéneo

$$\begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

es decir

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

luego

$$\begin{cases} x = y \\ y = \text{libre} \end{cases}$$

Los vectores propios son solución de un sistema homogéneo, es decir

$$\mathbf{x} \text{ está en } \text{Nul} \left(\begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix} \right)$$

y en consecuencia

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

- Podemos afirmar entonces que λ es un valor propio de A si y solo si la ecuación homogénea

$$(A - \lambda I_n) \mathbf{x} = \mathbf{0}$$

tiene infinitas soluciones, esto es, si y solo si

$$\mathbf{x} \text{ está en } \text{Nul}(A - \lambda I_n)$$

y

$$\text{Nul}(A - \lambda I_n) \neq \{\mathbf{0}\}.$$

- Al subespacio

$$\text{Nul}(A - \lambda I_n)$$

se le llama **espacio propio** de A correspondiente a λ .

- Ejemplo:** Sea

$$A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$$

Un valor propio de A es 2. Encuentre una base para el espacio propio correspondiente.

- Solución:

$$\begin{aligned} A - \lambda I_3 &= \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix} \end{aligned}$$

Calculamos entonces

$$\text{Nul}(A - \lambda I_3) = \text{Nul}\left(\begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}\right)$$

para ello

$$\begin{bmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -\frac{1}{2} & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

entonces el sistema lineal equivalente es

$$\begin{cases} x - \frac{1}{2}y + 3z = 0 \\ 0 = 0 \\ 0 = 0 \end{cases}$$

luego

$$\begin{cases} x = \frac{1}{2}y - 3z \\ y = \text{libre} \\ z = \text{libre} \end{cases}$$

es decir los vectores del espacio nulo se escriben como

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{2}y - 3z \\ y \\ z \end{bmatrix} = y \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}.$$

Finalmente

$$\text{Nul}(A - \lambda I_3) = \text{Gen} \left\{ \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Teorema 13.1. Los valores propios de una matriz triangular son las entradas de su diagonal principal.

■ **Demostración:**

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$$

entonces los valores propios λ hacen que el sistema homogéneo

$$\begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

tenga ∞ soluciones. Para eso la matriz $A - \lambda I$ tiene que tener menos de 3 columnas pivotes. Esto ocurre si

$$\begin{aligned} a_{11} - \lambda &= 0, \\ a_{22} - \lambda &= 0, \\ a_{33} - \lambda &= 0, \end{aligned}$$

cualquiera o todas estas igualdades se cumplen.

■ **Ejemplo:** Calcula los valores propios de

$$A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix} \text{ y } B = \begin{bmatrix} 3 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{bmatrix}$$

- Solución: Los valores propios de A son

3, 0 y 2.

Los valores propios de B son

3, 1 y 4.

Teorema 13.2. Sea A una matriz $n \times n$. Entonces A es invertible si y solo si 0 no es un valor propio de A .

■ **Demostración:** Existe una solución no trivial $\mathbf{x}$ de

$$A\mathbf{x} = 0\mathbf{x} = \mathbf{0}.$$

Teorema 13.3. Si

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$$

son vectores propios que corresponden a distintos valores propios

$$\lambda_1, \lambda_2, \dots, \lambda_r$$

de una matriz A $n \times n$, entonces el conjunto

$$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$$

es linealmente independiente.

■ **Demostración:**

13.2. La ecuación característica

- ¿Cómo podemos calcular los valores propios de una matriz A que no sea una matriz triangular?
- **Ejemplo:** Encuentre los valores propios de

$$A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}.$$

- Solución: Tenemos que encontrar todos los escalares λ de forma que

$$(A - \lambda I_2) \mathbf{x} = \mathbf{0}$$

tenga una solución no trivial. Por el Teorema de caracterización de matrices invertibles, esto es equivalente a demostrar que

$$(A - \lambda I_2)$$

no es invertible. El Teorema 4 de la Sección 2 del Tema 2, sabemos que la matriz $(A - \lambda I_2)$ no es invertible si y solo si su determinante vale 0, esto es,

$$\det(A - \lambda I_2) = 0.$$

Luego

$$\det \begin{bmatrix} 2 - \lambda & 3 \\ 3 & -6 - \lambda \end{bmatrix} = 0$$

es decir,

$$(2 - \lambda)(-6 - \lambda) - 9 = 0;$$

$$\lambda^2 + 4\lambda - 21 = 0.$$

Resolviendo obtenemos los valores propios:

$$\lambda = \frac{-4 \pm \sqrt{16 - 4 \times 1 \times (-21)}}{2} = \begin{cases} 3 \\ -7 \end{cases}.$$

13.2.1. Determinantes

- Sea

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

una matriz 2×2 se define el determinante como

$$\det A = a_{11}a_{22} - a_{21}a_{12}.$$

- Sea

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

se define entonces

$$\begin{aligned} \det A &= a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} \\ &\quad + a_{13} \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \end{aligned}$$

donde

$$\det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$$

se obtiene quitando la primera fila y la primera columna de la matriz A , el

$$\det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}$$

se obtiene quitando la primera fila y la segunda columna a la matriz A y, finalmente,

$$\det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

se obtiene quitando la primera fila y la tercera columna a la matriz A .

- Luego en general si se tiene una matriz A de tamaño $n \times n$ se define la matriz A_{ij} como aquella matriz que resta de quitar la fila i y la columna j de la matriz A .
- Se define entonces el determinante de la matriz A a partir del desarrollo por la fila i de la matriz A como

$$\det A \triangleq \sum_i (-1)^{i+1} a_{i1} \det A_{i1} + (-1)^{i+2} a_{i2} \det A_{i2} + \dots + (-1)^{i+n} a_{in} \det A_{in}.$$

- Se define entonces el determinante de la matriz A a partir del desarrollo por la columna j de la matriz A como

$$\det A \triangleq \sum_j (-1)^{1+j} a_{1j} \det A_{1j} + (-1)^{2+j} a_{2j} \det A_{2j} + \dots + (-1)^{n+j} a_{nj} \det A_{nj}.$$

Teorema 13.4. Sea A una matriz de tamaño $n \times n$, entonces

$$\det A = \det A = \det A$$

para $i = 1, 2, \dots, n$ y $j = 1, 2, \dots, n$.

- **Ejemplo:** Calcular

$$\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

- **Solución:** Desarrollamos por la tercera columna de la matriz A

$$\begin{aligned} & \det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \\ &= (-1)^3 2 \det \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} + (-1)^4 5 \det \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix} + (-1)^5 8 \det \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix} \\ &= -2 \det \begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} + 5 \det \begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix} - 8 \det \begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix} \\ &= -2 \times (-6) + 5 \times (-12) - 8 \times (-6) = 0. \end{aligned}$$

Propiedades de los determinantes

Teorema 13.5. Sea A una matriz cuadrada $n \times m$. Entonces

1. Si se suma un múltiplo de una fila de A a otra fila para producir una matriz B entonces

$$\det A = \det B.$$

2. Si se intercambian dos filas de A para producir una matriz B entonces

$$\det A = -\det B.$$

3. Si a una fila de A se multiplica por k para producir una matriz B entonces

$$\det B = k \det A.$$

■ **Ejemplo enunciado 1:**

$$\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \det \begin{bmatrix} 1 & 2 & 3 \\ 5 & 7 & 9 \\ 7 & 8 & 9 \end{bmatrix} = \det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 2 & -1 & 0 \end{bmatrix}.$$

■ **Ejemplo enunciado 2:**

$$\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = -\det \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{bmatrix}.$$

■ **Ejemplo enunciado 3:**

$$\sqrt{2} \det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \det \begin{bmatrix} 1 & 2 & 3 \\ 4\sqrt{2} & 5\sqrt{2} & 6\sqrt{2} \\ 7 & 8 & 9 \end{bmatrix}$$

■ **Ejemplo:** Calcular

$$\begin{aligned} & \det \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix} \\ &= 2 \det \begin{bmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix} = 2 \det \begin{bmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & -12 & 10 & 10 \\ 0 & 0 & -3 & 2 \end{bmatrix} \\ &= 2 \det \begin{bmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{bmatrix} = 4 \det \begin{bmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & -3 & 2 \end{bmatrix} \\ &= 4 \det \begin{bmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = 4 \times (-9) = -36. \end{aligned}$$

Teorema 13.6. Sea A una matriz $n \times n$, y sea U la forma escalonada obtenida a partir de A efectuando exclusivamente las operaciones por filas de reemplazo (reemplazar una fila por la suma de sí misma y un múltiplo de una fila) e intercambio (intercambiar dos filas) esto es,

$$A \sim U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

y sea

$r =$ número de intercambios de fila realizados.

Entonces

$$\det A = (-1)^r u_{11} u_{22} \cdots u_{nn}.$$

Además

$$\det U = u_{11} u_{22} \cdots u_{nn}.$$

■ **Ejemplo:**

$$\det \begin{bmatrix} 1 & 123 & 345 & 456 \\ 0 & 2 & \sqrt{234} & 321 \\ 0 & 0 & 3 & 232 \\ 0 & 0 & 0 & 4 \end{bmatrix} = 1 \times 2 \times 3 \times 4 = 24.$$

■ **Ejemplo:** Calcule el $\det A$ para

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

• Solución:

$$\begin{aligned} \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix} &\sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & -6 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \end{aligned}$$

entonces

$$\det \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix} = (-1)^1 (1 \times (-2) \times (-1)) = -2.$$

Otra forma alternativa (desarrollar por la última fila):

$$\begin{aligned} \det \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix} \\ = (-1)^4 0 \det \begin{bmatrix} 5 & 0 \\ 4 & -1 \end{bmatrix} + (-1)^5 (-2) \det \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \\ + (-1)^6 0 \det \begin{bmatrix} 1 & 5 \\ 2 & 4 \end{bmatrix} = 2 \det \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} = -2. \end{aligned}$$

Teorema 13.7. Sean A y B matrices $n \times n$:

1. A es invertible si y solo si $\det A \neq 0$.

$$2. \det(AB) = \det A \det B.$$

$$3. \det A^T = \det A.$$

■ **Ejemplo enunciado 2:**

$$\det \left(\begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} \right) = \det \begin{bmatrix} 25 & 20 \\ 14 & 13 \end{bmatrix} = 45.$$

$$\left(\det \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix} \right) \left(\det \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} \right) = 9 \times 5 = 45.$$

■ **Ejemplo enunciado 3:**

$$\det \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix} = 9 = \det \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}^T = \det \begin{bmatrix} 6 & 3 \\ 1 & 2 \end{bmatrix}$$

13.2.2. La ecuación característica

- Empleando el Teorema 4-(1) podemos utilizar el determinante para decidir cuando la matriz

$$(A - \lambda I_n)$$

no es invertible.

- La ecuación escalar

$$\det(A - \lambda I_n) = 0$$

es la **ecuación característica** de A .

- Un escalar λ es un valor propio de la matriz A $n \times n$, si y solo si satisface la ecuación característica

$$\det(A - \lambda I_n) = 0.$$

- **Ejemplo:** Encuentre la ecuación característica de

$$A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

- Solución:

$$A - \lambda I_4 = \begin{bmatrix} 5 - \lambda & -2 & 6 & -1 \\ 0 & 3 - \lambda & -8 & 0 \\ 0 & 0 & 5 - \lambda & 4 \\ 0 & 0 & 0 & 1 - \lambda \end{bmatrix}$$

entonces

$$\det(A - \lambda I_4) = (5 - \lambda)(3 - \lambda)(5 - \lambda)(1 - \lambda).$$

La ecuación característica es

$$(5 - \lambda)^2(3 - \lambda)(1 - \lambda) = 0$$

los valores propios son

$$\lambda_1 = 5, \lambda_2 = 3, \lambda_3 = 1.$$

Además, si desarrollamos la anterior expresión obtenemos

$$(5 - \lambda)^2(3 - \lambda)(1 - \lambda) = 75 - 130\lambda + 68\lambda^2 - 14\lambda^3 + \lambda^4 = 0$$

- Se puede demostrar que si A es una matriz $n \times n$, entonces

$$\det(A - \lambda I_n)$$

es un polinomio en λ de grado n que se le llama **polinomio característico** de A .

- El valor propio $\lambda_1 = 5$ del ejemplo anterior se dice que tiene multiplicidad 2 porque $(5 - \lambda)$ aparece dos veces como factor del polinomio característico.
- La **multiplicidad algebraica** de un valor propio λ es su multiplicidad como raíz de la ecuación característica.

13.2.3. Semejanza

- Sean A y B dos matrices $n \times n$, entonces A es **semejante** a B si existe una matriz invertible P de forma que

$$A = PBP^{-1}.$$

Teorema 13.8. Si dos matrices A y B $n \times n$ son semejantes, entonces tienen el mismo polinomio característico y por lo tanto los mismos valores propios (con las mismas multiplicidades).

- **Demostración:** Si

$$A = PBP^{-1}$$

entonces

$$A - \lambda I_n = PBP^{-1} - P(\lambda I_n)P^{-1} = P(B - \lambda I_n)P^{-1}.$$

Entonces empleando el Teorema 4-(2) obtenemos

$$\begin{aligned} \det(A - \lambda I_n) &= \det(P(B - \lambda I_n)P^{-1}) \\ &= \det P \det(B - \lambda I_n) \det P^{-1} \\ &= \det(B - \lambda I_n) \end{aligned}$$

ya que

$$1 = \det I_n = \det(PP^{-1}) = \det P \det P^{-1}$$

luego

$$\det P^{-1} = \frac{1}{\det P}.$$

13.2.4. Aplicación un Modelo dinámico de consumo

- Consideremos el problema

$$\begin{bmatrix} X_A(t+1) \\ X_B(t+1) \end{bmatrix} = \begin{bmatrix} 0.9821 & 0.0029 \\ 0.0179 & 0.9971 \end{bmatrix} \begin{bmatrix} X_A(t) \\ X_B(t) \end{bmatrix},$$

$$\begin{bmatrix} X_A(0) \\ X_B(0) \end{bmatrix} = \begin{bmatrix} 100000 \\ 300000 \end{bmatrix}$$

$\mathbf{x}_{t+1} = A\mathbf{x}_t$, $\mathbf{x}_0$ es la condición inicial.

$$\mathbf{x}_{t+1} = \begin{bmatrix} X_A(t+1) \\ X_B(t+1) \end{bmatrix}, A = \begin{bmatrix} 0.9821 & 0.0029 \\ 0.0179 & 0.9971 \end{bmatrix},$$

$$\mathbf{x}_t = \begin{bmatrix} X_A(t) \\ X_B(t) \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} X_A(0) \\ X_B(0) \end{bmatrix} = \begin{bmatrix} 100000 \\ 300000 \end{bmatrix}$$

para $t = 0, 1, \dots$. Nos interesa estudiar el comportamiento a largo plazo, esto es,

$$\lim_{t \rightarrow \infty} \begin{bmatrix} X_A(t) \\ X_B(t) \end{bmatrix}.$$

■ Nótese que si

$\mathbf{v}_1$ es un vector propio de A con valor propio λ_1 ,

$\mathbf{v}_2$ es un vector propio de A con valor propio λ_2 ,

y podemos escribir la condición inicial como

$$\mathbf{x}_0 = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2$$

entonces

$$\begin{aligned} \mathbf{x}_1 &= A\mathbf{x}_0 = A(x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2) = x_1 (A\mathbf{v}_1) + x_2 (A\mathbf{v}_2) \\ &= x_1 \lambda_1 \mathbf{v}_1 + x_2 \lambda_2 \mathbf{v}_2 \end{aligned}$$

En un segundo paso

$$\begin{aligned} \mathbf{x}_2 &= A\mathbf{x}_1 = A(x_1 \lambda_1 \mathbf{v}_1 + x_2 \lambda_2 \mathbf{v}_2) = x_1 \lambda_1 (A\mathbf{v}_1) + x_2 \lambda_2 (A\mathbf{v}_2) \\ &= x_1 \lambda_1^2 \mathbf{v}_1 + x_2 \lambda_2^2 \mathbf{v}_2 \end{aligned}$$

en general obtenemos la fórmula de recurrencia

$$\mathbf{x}_t = x_1 \lambda_1^t \mathbf{v}_1 + x_2 \lambda_2^t \mathbf{v}_2.$$

Finalmente

$$\lim_{t \rightarrow \infty} \mathbf{x}_t = \lim_{t \rightarrow \infty} (x_1 \lambda_1^t \mathbf{v}_1 + x_2 \lambda_2^t \mathbf{v}_2).$$

1. Calculamos polinomio característico

$$\det \begin{bmatrix} 0.9821 - \lambda & 0.0029 \\ 0.0179 & 0.9971 - \lambda \end{bmatrix} = 0.9792 - 1.9792\lambda + \lambda^2.$$

2. Calculamos las raíces del polinomio característico

$$0.9792 - 1.9792\lambda + \lambda^2 = 0$$

donde se obtiene que

$$\{\lambda_1 = 0.9792\}, \{\lambda_2 = 1.0\}$$

3. Calculamos los vectores propios para $\lambda_1 = 0.9792$ que son las soluciones del sistema homogéneo

$$\begin{bmatrix} 0.9821 - 0.9792 & 0.0029 \\ 0.0179 & 0.9971 - 0.9792 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

obtenemos que

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

luego

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

4. Calculamos los vectores propios para $\lambda_2 = 1.0$ que son las soluciones del sistema homogéneo

$$\begin{bmatrix} 0.9821 - 1.0 & 0.0029 \\ 0.0179 & 0.9971 - 1.0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

obtenemos

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} -0.16201 \\ 1 \end{bmatrix}$$

luego

$$\mathbf{v}_2 = \begin{bmatrix} -0.16201 \\ 1 \end{bmatrix}.$$

5. La condición inicial se escribe como combinación lineal de los vectores propios:

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 100000 \\ 300000 \end{bmatrix} = x_1 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} -0.16201 \\ 1 \end{bmatrix}$$

es decir

$$\begin{bmatrix} -1 & -0.16201 & 100000 \\ 1 & 1 & 300000 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1.7733 \times 10^5 \\ 0 & 1 & 4.7733 \times 10^5 \end{bmatrix}$$

lo que nos dice

$$x_1 = -1.7733 \times 10^5, x_2 = 4.7733 \times 10^5.$$

6. En consecuencia

$$\begin{aligned} \mathbf{x}_t &= \begin{bmatrix} X_A(t) \\ X_B(t) \end{bmatrix} = \\ &= (-1.7733 \times 10^5) (0.9792)^t \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &\quad + (4.7733 \times 10^5) (1.0)^t \begin{bmatrix} -0.16201 \\ 1 \end{bmatrix} \\ &= (-1.7733 \times 10^5) (0.9792)^t \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \\ &\quad + (4.7733 \times 10^5) \begin{bmatrix} -0.16201 \\ 1 \end{bmatrix} \end{aligned}$$

como

$$\lim_{t \rightarrow \infty} (0.9792)^t = 0$$

obtenemos que

$$\lim_{t \rightarrow \infty} \mathbf{x}_t = (4.7733 \times 10^5) \begin{bmatrix} -0.16201 \\ 1 \end{bmatrix} = \begin{bmatrix} -77332. \\ 4.7733 \times 10^5 \end{bmatrix}.$$

13.3. Diagonalización

- Dada una matriz A $n \times n$ podemos construir una matriz diagonal D que sea semejante a la matriz A : esto es, ¿podemos construir una matriz invertible P de forma que

$$A = PDP^{-1}.$$

- Recordemos que una matriz diagonal es de la forma

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix}.$$

- Una de las ventajas que ofrece una matriz diagonal es la siguiente: sea

$$D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}.$$

Entonces

$$D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 0 \\ 0 & 9 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix},$$

$$D^3 = DD^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 3^3 \end{bmatrix}.$$

En general

$$D^n = \overbrace{D \cdot D \cdots D}^{n\text{-veces}} = \begin{bmatrix} 5^n & 0 \\ 0 & 3^n \end{bmatrix} \text{ para } n \geq 1.$$

- **Ejemplo:** Sea

$$A = \begin{bmatrix} 7 & 2 \\ 1 & 4 \end{bmatrix}.$$

Encuentre una fórmula para A^k , dado que $A = PDP^{-1}$ donde

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}, D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}.$$

- Solución: En primer lugar calculamos

$$P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

a partir de

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ -1 & -2 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & -1 & -1 \end{bmatrix}.$$

Entonces, como

$$\begin{aligned} A^2 &= (PDP^{-1})^2 = (PDP^{-1})(PDP^{-1}) \\ &= PD(P^{-1}P)DP^{-1} \\ &= P(DI_n D)P^{-1} \\ &= PD^2P^{-1} \\ &= \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \end{aligned}$$

Si empleamos el mismo argumento una vez más:

$$\begin{aligned}
 A^3 &= AA^2 = (PDP^{-1})(PD^2P^{-1}) \\
 &= PD(P^{-1}P)D^2P^{-1} \\
 &= P(DI_n D^2)P^{-1} \\
 &= PD^3P^{-1}.
 \end{aligned}$$

En general, obtenemos la fórmula de recurrencia

$$A^k = PD^k P^{-1}.$$

Que nos da

$$\begin{aligned}
 \begin{bmatrix} 7 & 2 \\ 1 & 4 \end{bmatrix}^k &= \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 2 \times 5^k & 5^k \\ -3^k & -3^k \end{bmatrix} \\
 &= \begin{bmatrix} 2 \times 5^k - 3^k & 5^k - 3^k \\ -2 \times 5^k + 2 \times 3^k & -5^k + 2 \times 3^k \end{bmatrix}.
 \end{aligned}$$

- **Definición:** Una matriz cuadrada $n \times n$, A es **diagonalizable** si A es semejante a una matriz diagonal, esto es, si existe una matriz invertible P y una matriz D diagonal de forma que

$$A = PDP^{-1}.$$

Teorema 13.9 (Teorema de diagonalización). Una matriz A $n \times n$ es diagonalizable si y solo si A tiene n -vectores propios linealmente independientes. De hecho,

$$A = PDP^{-1}$$

si y solo si las columnas de la matriz P son n -vectores propios linealmente independientes. En este caso, las entradas diagonales de D son valores propios de A que corresponden, respectivamente, a los vectores propios de P .

- A una base como la que describe en Teorema de diagonalización se le llama **base de vectores propios**.
- **Demostración:** Sean

$$\begin{cases} \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \text{ vectores propios de } A \\ \lambda_1, \lambda_2, \dots, \lambda_n \text{ valores propios asociados} \end{cases}$$

Sea

$$P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix},$$

entonces

$$\begin{aligned}
 AP &= A \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} \\
 &= \begin{bmatrix} A\mathbf{v}_1 & A\mathbf{v}_2 & \cdots & A\mathbf{v}_n \end{bmatrix} \\
 &= \begin{bmatrix} \lambda_1\mathbf{v}_1 & \lambda_2\mathbf{v}_2 & \cdots & \lambda_n\mathbf{v}_n \end{bmatrix} \\
 &= \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \\ 0 & \cdots & \lambda_n \end{bmatrix} \\
 &= PD \implies AP = PD \iff APP^{-1} = PDP^{-1} \\
 &\iff A = PDP^{-1}.
 \end{aligned}$$

13.3.1. Diagonalización de matrices

- **Ejemplo:** Diagonalice, si es posible, la matriz siguiente:

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

- Solución:

1. Encontrar los valores propios de A .

$$\det(A - \lambda I) = \det \begin{bmatrix} 1 - \lambda & 3 & 3 \\ -3 & -5 - \lambda & -3 \\ 3 & 3 & 1 - \lambda \end{bmatrix}$$

$$\begin{aligned}
 &\begin{bmatrix} 1 - \lambda & 3 & 3 \\ -3 & -5 - \lambda & -3 \\ 3 & 3 & 1 - \lambda \end{bmatrix} \xrightarrow{r=1} \begin{bmatrix} 3 & 3 & 1 - \lambda \\ -3 & -5 - \lambda & -3 \\ 1 - \lambda & 3 & 3 \end{bmatrix} \\
 &\sim \begin{bmatrix} 3 & 3 & 1 - \lambda \\ 0 & -2 - \lambda & -2 - \lambda \\ 1 - \lambda & 3 & 3 \end{bmatrix} \text{ multiplicamos la 1ª fila por } \frac{1}{3} \\
 &\sim \begin{bmatrix} 1 & 1 & \frac{1}{3}(1 - \lambda) \\ 0 & -2 - \lambda & -2 - \lambda \\ 1 - \lambda & 3 & 3 \end{bmatrix} \\
 &\sim \begin{bmatrix} 1 & 1 & \frac{1}{3}(1 - \lambda) \\ 0 & -(2 + \lambda) & -(2 + \lambda) \\ 0 & 2 + \lambda & -\frac{1}{3}(1 - \lambda)^2 + 3 \end{bmatrix} \text{ multiplicamos la 1ª fila por 3} \\
 &\sim \begin{bmatrix} 3 & 3 & (1 - \lambda) \\ 0 & -(2 + \lambda) & -(2 + \lambda) \\ 0 & 0 & -\frac{1}{3}(1 - \lambda)^2 + 3 - (2 + \lambda) \end{bmatrix} = U \text{ con } r = 1.
 \end{aligned}$$

Entonces el polinomio característico

$$\begin{aligned}
 & \det \begin{bmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{bmatrix} \\
 &= (-1)^1 \times \left(3 \times (- (2+\lambda)) \times \left(-\frac{1}{3} (1-\lambda)^2 + 3 - (2+\lambda) \right) \right) \\
 &= 3(2+\lambda) \left(-\frac{1}{3} (1-\lambda)^2 + (1-\lambda) \right) \\
 &= 3(2+\lambda)(1-\lambda) \left(-\frac{1}{3} (1-\lambda) + 1 \right) \\
 &= 3(2+\lambda)(1-\lambda) \left(\frac{1}{3}\lambda + \frac{2}{3} \right) \\
 &= 3(2+\lambda)(1-\lambda) \frac{1}{3}(\lambda+2) \\
 &= (2+\lambda)^2(1-\lambda).
 \end{aligned}$$

En consecuencia los valores propios son

$$\begin{aligned}
 \lambda_1 &= -2 \text{ con multiplicidad algebraica } 2, \\
 \lambda_1 &= 1.
 \end{aligned}$$

2. Encontrar tres vectores propios de A linealmente independientes:

a) Para $\lambda_1 = -2$ hay que resolver

$$\begin{bmatrix} 1-(-2) & 3 & 3 \\ -3 & -5-(-2) & -3 \\ 3 & 3 & 1-(-2) \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

como

$$\begin{aligned}
 \begin{bmatrix} 1-(-2) & 3 & 3 \\ -3 & -5-(-2) & -3 \\ 3 & 3 & 1-(-2) \end{bmatrix} &= \begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} \\
 &\sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
 \end{aligned}$$

que nos da como solución

$$\begin{aligned}
 x + y + z &= 0, \\
 0 &= 0, \\
 0 &= 0,
 \end{aligned}$$

esto es

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = y \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Los dos vectores propios son

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

b) Para $\lambda_1 = 1$ hay que resolver

$$\begin{bmatrix} 1-1 & 3 & 3 \\ -3 & -5-1 & -3 \\ 3 & 3 & 1-1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

como

$$\begin{bmatrix} 1-1 & 3 & 3 \\ -3 & -5-1 & -3 \\ 3 & 3 & 1-1 \end{bmatrix} = \begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix} \\ \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

obtenemos

$$\begin{aligned} x - z &= 0, \\ y + z &= 0, \\ 0 &= 0 \end{aligned}$$

esto es

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = z \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

El vector propio es

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

3. Construir P con los vectores obtenidos con el paso anterior:

$$P = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

entonces se calcula

$$P^{-1} = \begin{bmatrix} -1 & -2 & -1 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

obteniendo finalmente que

$$\begin{bmatrix} 1 & -1 & 1 \\ -1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & -2 & -1 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \\ = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

- **Ejemplo de una matriz no diagonalizable:** Sea

$$A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

El polinomio característico es

$$\det(A - \lambda I) = (1 - \lambda)(\lambda + 2)^2.$$

Luego los valores propios son los mismos, sin embargo para $\lambda_1 = -2$ obtenemos como vector propio

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

y para $\lambda_2 = 1$ obtenemos

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

En este caso

$$P = \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & 0 \end{bmatrix}$$

no es una matriz cuadrada y en consecuencia los vectores propios no forman una base para $\mathbf{R}^3$. Podemos afirmar entonces que la matriz A no es diagonalizable.

- Como consecuencia del Teorema 2 obtenemos:

Teorema 13.10. Una matriz $n \times n$ con n -valores propios distintos es diagonalizable.

13.3.2. Matrices cuyos valores propios no son distintos

Teorema 13.11. Sea A una matriz $n \times n$ cuyos valores propios distintos son

$$\lambda_1, \lambda_2, \dots, \lambda_p.$$

1. Para $1 \leq k \leq p$, la dimensión del espacio propio para λ_k es menor o igual que la multiplicidad algebraica del valor propio λ_k .
2. La matriz A es diagonalizable si y solo si la suma de las dimensiones de los distintos espacios propios es igual a n y esto sucede si y solo si la dimensión del espacio propio para cada λ_k es igual a la multiplicidad algebraica de λ_k .
3. Si A es diagonalizable y $\mathcal{B}_k$ es una base para el espacio correspondiente a λ_k para cada k , entonces la colección total de vectores de los conjuntos

$$\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_p$$

forma una base de vectores propios de $\mathbf{R}^n$.

- **Ejemplo:** De ser posible diagonalice la matriz

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

- **Solución:** Como A es una matriz triangular los valores propios son

$$\lambda_1 = 5 \text{ multiplicidad algebraica } 2$$

y

$$\lambda_2 = -3 \text{ multiplicidad algebraica } 2$$

Entonces obtenemos

$$\mathcal{B}_{-3} = \text{Gen} \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\} \leftrightarrow \dim \mathcal{B}_{-3} = 2,$$

$$\mathcal{B}_5 = \text{Gen} \left\{ \begin{bmatrix} \frac{8}{3} \\ \frac{3}{4} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{16}{3} \\ -\frac{4}{3} \\ 0 \\ 1 \end{bmatrix} \right\} \leftrightarrow \dim \mathcal{B}_5 = 2.$$

El Teorema 11 (2) nos dice que como la dimensión de cada subespacio coincide con la multiplicidad algebraica del valor propio entonces la matriz A es diagonalizable.

Capítulo 14

Funciones de dos o más variables

14.1. Funciones de dos o más variables

Una función de dos variables x e y con dominio $D \subset \mathbb{R}^2$ es una regla que asigna un número específico $z = f(x, y)$ a cada vector

$$\begin{bmatrix} x \\ y \end{bmatrix} \text{ en } D.$$

Ejemplo 14.1. Consideremos $z = 2x + x^2y^3$, entonces $f(x, y) = 2x + x^2y^3$. Podemos entonces calcular

$$\begin{aligned} f(0, 0) &= 2 \cdot 0 + 0^2 \cdot 0^3 = 0, \\ f(1, 0) &= 2 \cdot 1 + 1^2 \cdot 0^3 = 2, \\ f(0, 1) &= 2 \cdot 0 + 0^2 \cdot 1^3 = 0, \\ f(1, 1) &= 2 \cdot 1 + 1^2 \cdot 1^3 = 3, \\ f(1, 2) &= 2 \cdot 1 + 1^2 \cdot 2^3 = 10. \end{aligned}$$

Ejemplo 14.2. Un estudio de la demanda de leche realizado por R. Frisch y T. Haavelmo, encontró la relación

$$x = A \frac{r^{2.08}}{p^{1.5}},$$

donde x es el consumo de leche, p su precio relativo y r la renta por familia y siendo A es una constante positiva. En este caso podemos escribir que

$$x = f(r, p) = A \frac{r^{2.08}}{p^{1.5}}.$$

Ejemplo 14.3. Una función que aparece en muchos modelos económicos es la función de Cobb–Douglas, esta se expresa como

$$F(x, y) = Ax^a y^b \text{ siendo } A, a \text{ y } b \text{ constantes.}$$

Nótese que la podemos expresar como

$$F(x, y) = Ae^{a \ln x} e^{b \ln y} = e^{a \ln x + b \ln y},$$

luego solo está definida para valores $x > 0$ e $y > 0$.

Ejercicio 14.1. ¿Podemos afirmar que la función que aparece en el Ejemplo 2 es una función de Cobb–Douglas?

Ejercicio 14.2. Demuestra que el logaritmo neperiano de una función Cobb–Douglas es lineal en la variables x e y , esto es, el $\ln F(x, y)$ es de la forma $Cx + Dy + F$ con C, D y F constantes.

Una función f de n -variables $x_1, x_2, \dots, x_n$ con dominio $D \subset \mathbb{R}^n$ es una regla que asigna un número específico $z = f(x_1, x_2, \dots, x_n)$ a cada vector

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ en } D.$$

Ejemplo 14.4. T. W. Schultz calculó que la demanda de azúcar x en E.E.U.U durante el periodo 1929–1935 se podía describir aproximadamente empleando la expresión

$$x = 108.83 - 6.0294p + 0.164w - 0.4217t,$$

siendo p el precio del azúcar, w un índice de producción y t el año, tomando como referencia $t = 0$ en 1929. En este caso

$$x = f(p, w, t) = 108.83 - 6.0294p + 0.164w - 0.4217t.$$

Ejemplo 14.5. R. Stone halló la fórmula siguiente para la demanda de cerveza x en Inglaterra:

$$x = 1.058x_1^{0.136}x_2^{-0.727}x_3^{0.914}x_4^{0.816},$$

donde x_1 es la renta del individuo, x_2 el precio de la cerveza, x_3 un índice general de precios y otros bienes y x_4 la fortaleza de la cerveza. En este caso

$$x = f(x_1, x_2, x_3, x_4) = 1.058x_1^{0.136}x_2^{-0.727}x_3^{0.914}x_4^{0.816}$$

y el dominio de definición es $x_1 > 0, x_2 > 0, x_3 > 0$ y $x_4 > 0$.

Una función de n -variables se dice lineal si

$$f(x_1, x_2, \dots, x_n) = a_1x_1 + a_2x_2 + \dots + a_nx_n + b,$$

siendo $a_1, a_2, \dots, a_n$ y b constantes.

Ejemplo 14.6.

$$f(x_1, x_2, x_3) = 3x_1 + 2x_2 - x_3 + 2.$$

Dominios

Al igual que en el caso de las funciones de una variable, supondremos, a menos que se diga lo contrario, que el dominio de una función definida por una fórmula es el mayor dominio en el cuál la fórmula tiene sentido y da un valor único.

Ejemplo 14.7. Consideremos

$$f(x, y) = \sqrt{x-1} + \sqrt{y}.$$

En este caso el dominio D , estará formado por los vectores de $\mathbb{R}^2$ que cumplen que $x \geq 1$ e $y \geq 0$.

Ejemplo 14.8. Consideremos

$$g(x, y) = \frac{1}{(x^2 + y^2 - 4)^{1/2}} + \sqrt{9 - (x^2 + y^2)}.$$

Se tiene que cumplir para que esté correctamente definida que $9 - (x^2 + y^2) \geq 0$ y que $x^2 + y^2 - 4 \neq 0$. Es decir $x^2 + y^2 \leq 9$ y $x^2 + y^2 \neq 4$.

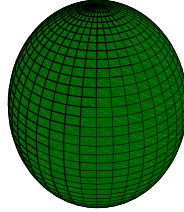


Figura 14.1: La esfera $x^2 + y^2 + z^2 = 4$.

14.2. Representación geométrica de las funciones de varias variables

Una ecuación en dos variables del tipo $f(x, y) = c$, siendo c una constante se puede representar en el plano mediante una curva.

Ejemplo 14.9. Si $f(x, y) = x^2 + y^2$, entonces tomando $c = 4$ se tiene que

$$f(x, y) = x^2 + y^2 = c = 4.$$

La ecuación $x^2 + y^2 = 4$ describe los puntos situados sobre una circunferencia de radio 2.

Una ecuación en tres variables del tipo $f(x, y, z) = c$, siendo c una constante se puede representar en el espacio mediante una superficie-

Ejemplo 14.10. Si $f(x, y, z) = x^2 + y^2 + z^2$, entonces tomando $c = 4$ se tiene que

$$f(x, y, z) = x^2 + y^2 + z^2 = c = 4.$$

La ecuación $x^2 + y^2 + z^2 = 4$ describe los puntos situados sobre una esfera de radio 2.

La gráfica de una función de dos variables

Sea $z = f(x, y)$ una función de dos variables definida en un dominio D del plano $\mathbb{R}^2$. La gráfica de la función f se puede describir como la superficie generada por la función $F(x, y, z) = z - f(x, y)$ tomando

$$F(x, y, z) = z - f(x, y) = c = 0.$$

Ejemplo 14.11. Consideremos $z = f(x, y) = x^2 + y^2$ y definamos

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

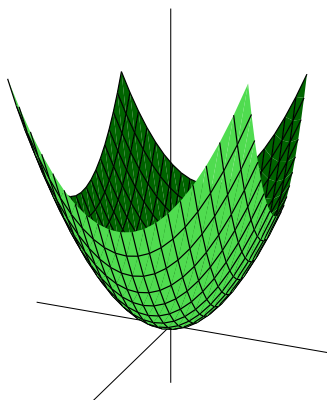


Figura 14.2: La superficie $z = x^2 + y^2$.

si dibujamos

$$F(x, y) = z - (x^2 + y^2) = 0 \Leftrightarrow z = x^2 + y^2,$$

obtenemos la gráfica de la Figura 2.

Este modo de representar gráficamente una función de dos variables nos ayuda a visualizar el comportamiento en términos generales. El inconveniente es que se requiere o bien una cierta habilidad artística o bien el empleo de un programa de ordenador. La Figura 2 se generó empleando el software Euler (licencia GPL) que se puede descargar gratuitamente en:

<http://mathsrv.ku-eichstaett.de/MGF/homes/grothmann/euler/>

simplemente ejecutando la instrucción:

```
>f3d("x^2+y^2")
```

Curvas de nivel para $z = f(x, y)$

Los cartógrafos dibujan características topográficas de la superficie terrestre, como colinas y valles en un mapa plano. Para ello dibujan *curvas de nivel*, que son contornos que unen puntos del mapa que representan posiciones en el terreno con la misma altitud sobre el nivel del mar. Podemos emplear la misma idea para representar geoméricamente una función $z = f(x, y)$ arbitraria. Para ello se corta la gráfica de una función por planos horizontales $z = \text{constante}$, esto es, paralelos al plano xy . Luego, se proyectan esas intersecciones perpendicularmente sobre el plano xy . Si el plano por el que hemos cortado es $z = c$, la proyección sobre el plano xy se le llama *curva de nivel* de f de cota c . La curva de nivel será, pues, la curva plana de ecuación

$$f(x, y) = c.$$

Ejemplo 14.12. Consideremos la función de dos variables

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

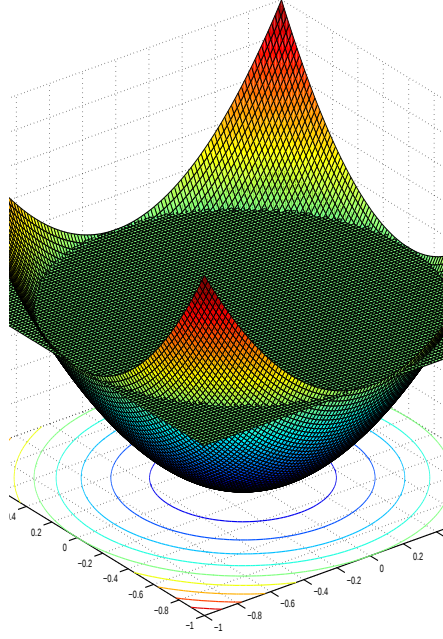


Figura 14.3: La superficie $z = x^2 + y^2$ intersección el plano $z = 1$, y las curvas de nivel asociadas.

definida para $-1 \leq x \leq 1$ y $-1 \leq y \leq 1$. Si consideramos $z = 1$, en la Figura 3 representamos el corte por el plano $z = 1$, y la proyección correspondiente que nos da la curva de nivel $x^2 + y^2 = 1$ en el plano xy . Como podemos ver en la misma figura, las curvas de nivel son del tipo $x^2 + y^2 = c$, que representan circunferencias centradas en el origen con radio $\sqrt{c}$.

Ejemplo 14.13. Consideremos la función de dos variables

$$z = x^3 + y^2 + x + y - 1.$$

definida para $-1 \leq x \leq 1$ y $-1 \leq y \leq 1$. En la Figura 4 representamos la intersección con el plano $z = 1$ y las curvas de nivel del tipo $x^3 + y^2 + x + y - 1 = c$.

Funciones de n -variables en $\mathbb{R}^n$

Solo podemos obtener algún tipo de representación geométrica para funciones de $n = 1, 2, 3$ -variables, o bien empleando su gráfica o bien mediante curvas o superficies de nivel.

Si $z = f(x_1, x_2, \dots, x_n)$, llamamos gráfica de f al lugar geométrico de $\mathbb{R}^{n+1}$ formado por los puntos de la forma

$$(x_1, x_2, \dots, x_n, f(x_1, x_2, \dots, x_n)).$$

tales que $(x_1, x_2, \dots, x_n)$ están en el dominio D . A esta gráfica se le llama por similitud *superficie*, aunque su nombre técnico es el de *hipersuperficie*.

Si c es un número real cualquiera, lugar geométrico de $\mathbb{R}^n$ formado por los puntos que cumplen que

$$f(x_1, x_2, \dots, x_n) = c$$

se le llama *superficie de nivel* de f .

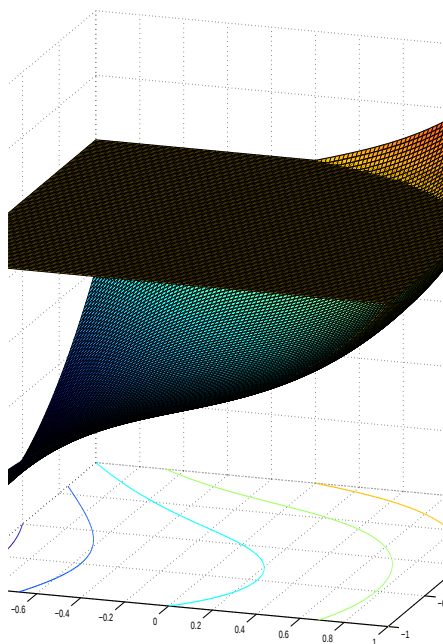


Figura 14.4: La superficie $z = x^3 + y^2 + x + y - 1$ intersección con $z = 1$ y las curvas de nivel asociadas.

14.3. Continuidad

Se puede generalizar a funciones de varias variables el concepto de continuidad para funciones de una variable, aunque la noción de límite es claramente mucho más compleja. Veamos como ilustración la definición en el caso de funciones de dos variables.

Sea $f(x, y)$ una función y (x_0, y_0) un punto dado, de forma que f está definida en el conjunto de puntos (x, y) de forma que

$$\text{dist} \left(\begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \right) = \sqrt{(x - x_0)^2 + (y - y_0)^2} < r.$$

Diremos que

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = L$$

si para cada $\varepsilon > 0$ existe $\delta > 0$ de forma que si

$$\text{dist} \left(\begin{bmatrix} x \\ y \end{bmatrix}, \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \right) = \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta$$

entonces

$$|f(x, y) - L| < \varepsilon.$$

En este caso se han sustituido los intervalos abiertos de la recta real $\mathbb{R}$ del tipo

$$|x - x_0| < \delta \Leftrightarrow x \text{ está en } (x_0 - \delta, x_0 + \delta)$$

por conjuntos de $\mathbb{R}^2$ del tipo

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta \Leftrightarrow (x - x_0)^2 + (y - y_0)^2 < \delta^2.$$

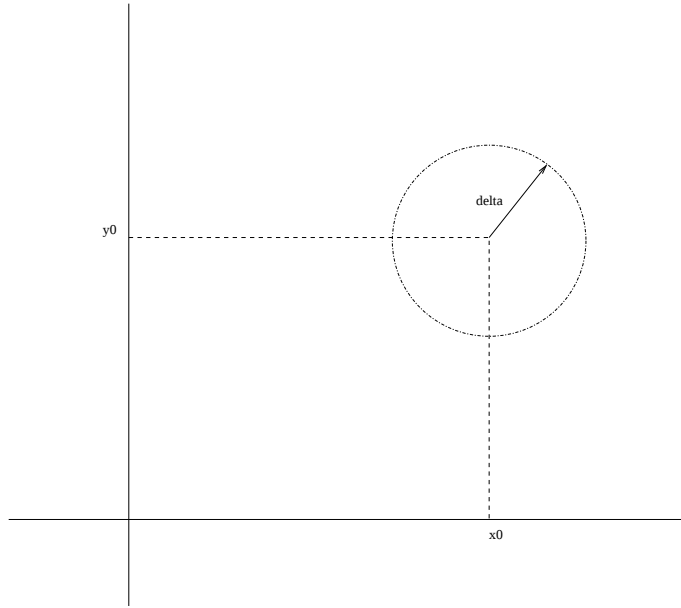


Figura 14.5: El disco centrado en (x_0, y_0) de radio δ .

Este conjunto es la parte interior de un disco centrado en (x_0, y_0) y radio δ (ver la Figura 5).

La mayor parte de funciones con las que trabajaremos son derivables y como consecuencia de ello continuas, por eso al igual que en el caso de una variable emplearemos la siguiente regla útil: *Toda función de n -variables que se pueda construir por operaciones de adición, sustracción, multiplicación, división y composición de funciones es continua allí donde esté definida.*

Ejemplo 14.14. La función

$$f(x, y, z) = x^2y + 8x^2y^5z - xy + 8z,$$

es continua para cualquier punto de $\mathbb{R}^3$.

Ejemplo 14.15. La función

$$g(x, y) = \frac{xy - 3}{x^2 + y^2 - 4}$$

es continua en todos los puntos de $\mathbb{R}^2$ excepto aquellos que cumplen que

$$x^2 + y^2 = 4.$$

14.4. Derivadas parciales

Derivadas parciales en dos variables

Consideremos la función

$$z = x^3 + 2y^2 \tag{14.1}$$

si suponemos que y es constante, podemos calcular

$$\frac{dz}{dx} = 3x^2.$$

Si en la expresión (14.1) suponemos ahora que x es constante, podemos calcular

$$\frac{dz}{dy} = 4y.$$

Cuando se trabaja con funciones de más de una variable se escribe que

$$\frac{\partial z}{\partial x} = 3x^2 \text{ y que } \frac{\partial z}{\partial y} = 4y.$$

Sea $z = f(x, y)$. La *derivada parcial* de z o f con respecto a x , y que se denota por

$$\frac{\partial z}{\partial x}(x, y) \text{ ó } \frac{\partial f}{\partial x}(x, y)$$

o bien para simplificar como

$$\frac{\partial z}{\partial x} \text{ ó } \frac{\partial f}{\partial x}$$

es la derivada de $f(x, y)$ con respecto a x cuando y se mantiene constante. La derivada parcial de z o f con respecto a y , y que se denota por

$$\frac{\partial z}{\partial y}(x, y) \text{ ó } \frac{\partial f}{\partial y}(x, y),$$

o bien para simplificar como

$$\frac{\partial z}{\partial y} \text{ ó } \frac{\partial f}{\partial y},$$

es la derivada de $f(x, y)$ con respecto a y cuando x se mantiene constante.

Ejemplo 14.16. Calcular las derivadas parciales de

$$f(x, y) = x^3y + x^2y^2 + x + y^2.$$

Tenemos

$$\frac{\partial f}{\partial x} = 3x^2y + 2xy^2 + 1,$$

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

Ejemplo 14.17. Calcular las derivadas parciales de

$$f(x, y) = \frac{xy}{x^2 + y^2}.$$

Tenemos

$$\frac{\partial f}{\partial x} = \frac{y(x^2 + y^2) - (2x)xy}{(x^2 + y^2)^2} = \frac{y^3 - x^2y}{(x^2 + y^2)^2},$$

$$\frac{\partial f}{\partial y} = \frac{x^3 - y^2x}{(x^2 + y^2)^2}.$$

Ejemplo 14.18. Calcular el valor de $\frac{\partial f}{\partial x}(1, 2)$ y $\frac{\partial f}{\partial y}(1, 2)$ para la función

$$f(x, y) = \frac{xy}{x^2 + y^2}.$$

Empleando los resultados obtenidos el ejemplo anterior tenemos

$$\frac{\partial f}{\partial x}(1, 2) = \frac{2^3 - 1^2 \cdot 2}{(1^2 + 2^2)^2} = \frac{6}{25},$$

y

$$\frac{\partial f}{\partial y}(1, 2) = \frac{1^3 - 2^2 \cdot 1}{(1^2 + 2^2)^2} = \frac{-3}{25}.$$

A menudo se emplean también las siguientes notaciones alternativas para las derivadas parciales:

$$\frac{\partial f}{\partial x}(x, y) = f_x(x, y) = z_x(x, y) = D_x f(x, y),$$

y

$$\frac{\partial f}{\partial y}(x, y) = f_y(x, y) = z_y(x, y) = D_y f(x, y).$$

La manera formal de definir las derivadas parciales es: Sea $z = f(x, y)$ definida en un disco sin incluir el perímetro del mismo (es lo que se conoce con un *disco abierto*), entonces si existe

$$\lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} = \frac{\partial f}{\partial x}(x, y).$$

Del mismo modo si existe

$$\lim_{\Delta y \rightarrow 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} = \frac{\partial f}{\partial y}(x, y).$$

Derivadas parciales en dos variables de orden superior

Recordemos que para la función del Ejemplo 16 las derivadas parciales

$$\frac{\partial f}{\partial x}(x, y) = 3x^2y + 2xy^2 + 1,$$

y

$$\frac{\partial f}{\partial y}(x, y) = x^3 + 2x^2y + 2y.$$

son funciones que podemos volver a derivar como sigue:

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x}(x, y) \right) = \frac{\partial}{\partial x} (3x^2y + 2xy^2 + 1) = 6xy + 2y^2,$$

$$\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x}(x, y) \right) = \frac{\partial}{\partial y} (3x^2y + 2xy^2 + 1) = 3x^2 + 4xy,$$

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y}(x, y) \right) = \frac{\partial}{\partial x} (x^3 + 2x^2y + 2y) = 3x^2 + 4xy,$$

$$\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y}(x, y) \right) = \frac{\partial}{\partial y} (x^3 + 2x^2y + 2y) = 2x^2 + 2.$$

La notación que se emplea es la siguiente:

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial x^2} \quad \text{y} \quad \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial y \partial x},$$

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial x \partial y} \quad \text{y} \quad \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial y^2}.$$

Podemos además calcular para la función del Ejemplo 16:

$$\frac{\partial^2 f}{\partial x^2}(1, 2) = 6 \cdot 1 \cdot 2 + 2 \cdot 2^2 = 20,$$

$$\frac{\partial^2 f}{\partial y \partial x}(1, 2) = 3 \cdot 1^2 + 4 \cdot 1 \cdot 2 = 11,$$

$$\frac{\partial^2 f}{\partial x \partial y}(1, 2) = 3 \cdot 1^2 + 4 \cdot 1 \cdot 2 = 11,$$

$$\frac{\partial^2 f}{\partial y^2}(1, 2) = 2 \cdot 1^2 + 2 = 4.$$

14.5. La diferencial

Conocemos que si una función de una variable $y = f(x)$ es derivable entonces su diferencial es

$$dy = f'(x)dx.$$

En el caso de una función de dos variables $z = f(x, y)$ podemos extender de forma natural esta definición. Supongamos que existen las derivadas parciales $f_x(x, y)$ y $f_y(x, y)$ y son funciones continuas. Entonces se define la diferencial

$$dz = f_x(x, y)dx + f_y(x, y)dy.$$

Ejemplo 14.19. Calcula la diferencial de la función $f(x, y) = 2x \sin y - 3x^2y^2$.

Conocemos que

$$f_x(x, y) = 2 \sin y - 6xy^2,$$

$$f_y(x, y) = 2x \cos y - 6x^2y.$$

Entonces

$$dz = (2 \sin y - 6xy^2) dx + (2x \cos y - 6x^2y) dy.$$

Nótese que se puede escribir empleando notación matricial y el producto escalar

$$\begin{aligned} dz &= \begin{bmatrix} 2 \sin y - 6xy^2 & 2x \cos y - 6x^2y \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} \\ &= \begin{bmatrix} f_x(x, y) & f_y(x, y) \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} \\ &= \begin{bmatrix} f_x(x, y) \\ f_y(x, y) \end{bmatrix} \bullet \begin{bmatrix} dx \\ dy \end{bmatrix} \end{aligned}$$

Al vector formado por las derivadas parciales se le llama *vector gradiente* y se le denota por

$$\nabla f(x, y) \triangleq \begin{bmatrix} f_x(x, y) \\ f_y(x, y) \end{bmatrix}.$$

A la matriz

$$Df(x, y) \triangleq \begin{bmatrix} f_x(x, y) & f_y(x, y) \end{bmatrix} = \nabla f(x, y)^T,$$

se le llama *derivada de la función f* .

En general para una función de n -variables se define el vector gradiente como

$$\nabla f(x_1, x_2, \dots, x_n) \triangleq \begin{bmatrix} f_{x_1}(x_1, x_2, \dots, x_n) \\ f_{x_2}(x_1, x_2, \dots, x_n) \\ \vdots \\ f_{x_n}(x_1, x_2, \dots, x_n) \end{bmatrix},$$

y la derivada como

$$Df(x_1, x_2, \dots, x_n) = \nabla f(x_1, x_2, \dots, x_n)^T.$$

Sea $f(x, y)$ una función definida en un dominio D . Diremos que f es diferenciable en (x_0, y_0) si

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \varepsilon_1\Delta x + \varepsilon_2\Delta y$$

donde $\varepsilon_1, \varepsilon_2 \rightarrow 0$ cuando $(\Delta x, \Delta y) \rightarrow (0, 0)$. La función se dice diferenciable en D si es diferenciable en todo punto de D .

Teorema 14.1. Si f es una función para la cual las derivadas parciales existen y son continuas en un dominio abierto D , entonces f es diferenciable en D .

Teorema 14.2. Si f es diferenciable en un punto del dominio, entonces es continua en dicho punto.

14.6. La Regla de la cadena

El caso de una única variable independiente

Consideremos $w = f(x, y) = x^2 y - y^2$. Supongamos que

$$x = x(t) = \sin t, \quad y = y(t) = e^t,$$

y deseamos calcular

$$\frac{dw}{dt}$$

en $t = 0$.

Está claro que podemos sustituir

$$w = w(t) = f(x(t), y(t)) = x(t)^2 y(t) - y(t)^2 = \sin^2 t \, e^t - e^{2t}$$

y derivar

$$\frac{dw}{dt} = 2 \sin t \cos t \, e^t + \sin^2 t \, e^t - 2e^{2t}.$$

Sustituyendo $t = 0$ obtenemos

$$\left. \frac{dw}{dt} \right|_{t=0} = -2.$$

Ahora bien, en el ejemplo anterior conocemos

$$f_x(x, y) = 2xy \quad f_y(x, y) = x^2 - 2y$$

y

$$\frac{dx}{dt} = \cos t, \quad \frac{dy}{dt} = e^t$$

¿Podemos relacionar $\frac{dw}{dt}$ con estas derivadas que podemos calcular previamente? La respuesta viene determinada por el siguiente argumento:

$$\frac{dw}{dt} = \lim_{\Delta t \rightarrow 0} \frac{f(x(t + \Delta t), y(t + \Delta t)) - f(x(t), y(t))}{\Delta t}.$$

Ahora bien, empleando la definición de diferencial, tomando

$$\Delta x = x(t + \Delta t) - x(t) \quad \text{y} \quad \Delta y = y(t + \Delta t) - y(t)$$

se tiene que

$$f(x(t + \Delta t), y(t + \Delta t)) - f(x(t), y(t)) =$$

$$f_x(x(t), y(t))(x(t + \Delta t) - x(t)) + f_y(x(t), y(t))(y(t + \Delta t) - y(t))$$

$$+ \varepsilon_1(x(t + \Delta t) - x(t)) + \varepsilon_2(y(t + \Delta t) - y(t))$$

donde

$$\lim_{\Delta t \rightarrow 0} \varepsilon_i = 0,$$

para $i = 1, 2$. Nótese que cuando $\Delta t \rightarrow 0$ se tiene que $(\Delta x, \Delta y) \rightarrow (0, 0)$. Dividiendo por Δt , la expresión anterior obtenemos

$$\begin{aligned} & \frac{f(x(t + \Delta t), y(t + \Delta t)) - f(x(t), y(t))}{\Delta t} = \\ & f_x(x(t), y(t)) \frac{x(t + \Delta t) - x(t)}{\Delta t} + f_y(x(t), y(t)) \frac{y(t + \Delta t) - y(t)}{\Delta t} \\ & + \varepsilon_1 \frac{(x(t + \Delta t) - x(t))}{\Delta t} + \varepsilon_2 \frac{(y(t + \Delta t) - y(t))}{\Delta t} \end{aligned}$$

Si tomamos el límite cuando $\Delta t \rightarrow 0$ obtenemos la regla de la cadena:

$$\frac{dw}{dt} = f_x(x(t), y(t)) \frac{dx}{dt} + f_y(x(t), y(t)) \frac{dy}{dt}.$$

Esta se expresa usualmente como

$$\frac{dw}{dt} = \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt}.$$

Ejemplo 14.20. Consideremos $w = f(x, y) = x^2y - y^2$. Supongamos que

$$x = x(t) = \sin t, \quad y = y(t) = e^t.$$

Calcular entonces

$$\frac{dw}{dt}$$

en $t = 0$.

Si empleamos ahora la regla de la cadena tenemos que

$$\begin{aligned} \frac{dw}{dt} &= \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} \\ &= (2xy) \cos t + (x^2 - 2y)e^t. \end{aligned}$$

Si ahora empleamos que $x = \sin t$ e $y = e^t$ obtenemos

$$\begin{aligned} &= 2 \sin t e^t \cos t + (\sin^2 t - 2e^t)e^t \\ &= 2 \sin t \cos t e^t + \sin^2 t e^t - 2e^{2t}. \end{aligned}$$

que es el mismo resultado que habíamos obtenido anteriormente.

En el caso de una función $f(x_1, x_2, \dots, x_n)$ con

$$x_1 = x_1(t), x_2 = x_2(t), \dots, x_n = x_n(t)$$

se tiene

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1} \frac{dx_1}{dt} + \frac{\partial f}{\partial x_2} \frac{dx_2}{dt} + \dots + \frac{\partial f}{\partial x_n} \frac{dx_n}{dt}.$$

El caso de más de una variable independiente

Consideremos la función $w = f(x, y) = 2xy$ de forma que

$$x = x(s, t) = s^2 + t^2 \text{ e } y = y(s, t) = s/t.$$

Entonces

$$\begin{aligned} w = f(s, t) = f(x(s, t), y(s, t)) &= 2x(s, t)y(s, t) = 2(s^2 + t^2)(s/t) \\ &= 2\left(\frac{s^3}{t} + st\right). \end{aligned}$$

Podemos calcular directamente entonces

$$\begin{aligned} \frac{\partial w}{\partial s} &= \frac{6s^2 + 2t^2}{t}, \\ \frac{\partial w}{\partial t} &= \frac{2st^2 - 2s^3}{t^2}. \end{aligned}$$

Ahora bien a partir de la expresión

$$w = f(x(s, t), y(s, t))$$

si consideramos t constante y derivamos con respecto a s empleando la regla de la cadena con una única variable independiente obtenemos

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial s}.$$

Empleando el mismo argumento con t , dejando s constante obtenemos

$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial w}{\partial y} \frac{\partial y}{\partial t}.$$

Si volvemos a nuestro ejemplo inicial se tiene

$$\frac{\partial w}{\partial x} = 2y, \quad \frac{\partial w}{\partial y} = 2x$$

y

$$\begin{aligned} \frac{\partial x}{\partial s} &= 2s, \quad \frac{\partial y}{\partial s} = \frac{1}{t}, \\ \frac{\partial x}{\partial t} &= 2t, \quad \frac{\partial y}{\partial t} = -\frac{s}{t^2}. \end{aligned}$$

Sustituyendo se obtiene finalmente

$$\begin{aligned} \frac{\partial w}{\partial s} &= (2y)(2s) + (2x)\frac{1}{t} \text{ sustituyendo } x = s^2 + t^2 \text{ e } y = s/t \\ &= \left(\frac{2s}{t}\right)(2s) + 2(s^2 + t^2)\frac{1}{t} \\ &= \frac{6s^2 + 2t^2}{t}. \end{aligned}$$

y

$$\begin{aligned} \frac{\partial w}{\partial t} &= (2y)(2t) + 2x\left(-\frac{s}{t^2}\right) \text{ sustituyendo } x = s^2 + t^2 \text{ e } y = s/t \\ &= \left(\frac{2s}{t}\right)(2t) - 2(s^2 + t^2)\frac{s}{t^2} \\ &= \frac{2st^2 - 2s^3}{t^2}. \end{aligned}$$

14.7. Derivación implícita

Consideremos la ecuación

$$w = f(x, y) = 0$$

entonces la diferencial cumple

$$dw = 0 = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy.$$

Luego,

$$\frac{\partial f}{\partial y} dy = -\frac{\partial f}{\partial x} dx,$$

es decir

$$\frac{dy}{dx} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}$$

siempre que

$$\frac{\partial f}{\partial y} \neq 0.$$

Esto viene a decirnos que siempre que se cumpla $f(x, y) = 0$ y cuando $\frac{\partial f}{\partial y} \neq 0$ podemos despejar $y = y(x)$ como función de x localmente, y la función es derivable siendo capaces de calcular la derivada implícitamente.

Ejemplo 14.21. Consideremos

$$w = y^3 + y^2 - 5y - x^2 + 4 = 0$$

Entonces

$$\begin{aligned} dw &= \frac{\partial}{\partial x}(y^3 + y^2 - 5y - x^2 + 4)dx + \frac{\partial}{\partial y}(y^3 + y^2 - 5y - x^2 + 4)dy \\ &= -2xdx + (3y^2 + 2y)dy \\ &= 0. \end{aligned}$$

De aquí se obtiene

$$(3y^2 + 2y)dy = 2xdx \Rightarrow \frac{dy}{dx} = \frac{2x}{3y^2 + 2y} \text{ siempre que } 3y^2 + 2y \neq 0.$$

¿Qué ocurre en el caso de funciones de más de dos variables?

Ejemplo 14.22. Consideremos

$$w = 3x^2z - x^2y^2 + 2z^3 + 3yz - 5 = 0$$

Si calculamos la diferencial

$$\begin{aligned} dw &= \frac{\partial w}{\partial x} dx + \frac{\partial w}{\partial y} dy + \frac{\partial w}{\partial z} dz \\ &= (6xz - 2xy^2)dx + (2x^2y + 3z)dy + (3x^2 + 6z^2 + 3y)dz \\ &= 0. \end{aligned}$$

De aquí deducimos que

$$(6xz - 2xy^2)dx + (2x^2y + 3z)dy + (3x^2 + 6z^2 + 3y)dz = 0.$$

Si despejamos dz obtenemos la expresión

$$dz = -\frac{6xz - 2xy^2}{3x^2 + 6z^2 + 3y}dx - \frac{2x^2y + 3z}{3x^2 + 6z^2 + 3y}dy,$$

esto nos dice que si consideramos a $z = z(x, y)$ entonces

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy.$$

en consecuencia siempre que $3x^2 + 6z^2 + 3y \neq 0$ podemos escribir

$$\frac{\partial z}{\partial x} = -\frac{6xz - 2xy^2}{3x^2 + 6z^2 + 3y} \text{ y } \frac{\partial z}{\partial y} = -\frac{2x^2y + 3z}{3x^2 + 6z^2 + 3y}.$$

Estas derivadas son la derivadas parciales implícitas de z con respecto a x y a y , respectivamente.

Nótese que si se hubiese despejado dx entonces

$$dx = -\frac{2x^2y + 3z}{6xz - 2xy^2}dy - \frac{3x^2 + 6z^2 + 3y}{6xz - 2xy^2}dz.$$

Si consideramos que $x = x(y, z)$ entonces

$$dx = \frac{\partial x}{\partial y}dy + \frac{\partial x}{\partial z}dz.$$

Luego siempre que $6xz - 2xy^2 \neq 0$ podemos calcular las derivadas parciales implícitas

$$\frac{\partial x}{\partial y} = -\frac{2x^2y + 3z}{6xz - 2xy^2} \text{ y } \frac{\partial x}{\partial z} = -\frac{3x^2 + 6z^2 + 3y}{6xz - 2xy^2}.$$

14.8. Derivadas direccionales y gradientes

Sea $w = f(x, y)$ una función de dos variables. Entonces si consideramos una recta descrita en forma paramétrica

$$\gamma(t) = \mathbf{x}_0 + t\mathbf{u} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + t \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} x_0 + tu_1 \\ y_0 + tu_2 \end{bmatrix}$$

donde (x_0, y_0) y (u_1, u_2) son conocidos. Podemos evaluar la función f a lo largo de esta recta, simplemente calculando

$$w(t) = f(\gamma(t)) = f(x_0 + tu_1, y_0 + tu_2).$$

Nótese que esta función cumple que $w(0) = f(x_0, y_0)$ y, empleando la regla de la cadena

$$w'(t) = f_x(x_0, y_0)u_1 + f_y(x_0, y_0)u_2 = \nabla f(x_0, y_0) \bullet \mathbf{u}.$$

Ejemplo 14.23. Consideremos $w = f(x, y) = 2xy$, y sea

$$\gamma(t) = (1 + 2t, 3 + 5t) = (1, 3) + t(2, 5).$$

En este caso se trata de una recta que pasa por el punto $(1, 2)$ y lleva la dirección del vector $(2, 5)$. Si calculamos

$$w(t) = f(\gamma(t)) = f((1 + 2t, 3 + 5t)) = 2(1 + 2t)(3 + 5t) = 20t^2 + 22t + 6$$

entonces

$$w'(t) = 40t + 22.$$

Esto nos da la derivada de f en la dirección de la recta $\gamma(t)$. Si igualamos a cero, obtenemos que $t = -22/40$ la derivada en esa dirección se anula, como $w''(t) = 40 > 0$, se tiene que en esa dirección la función f tiene un mínimo.

Se define entonces la *derivada direccional de f en la dirección de $\mathbf{u}$* y se denota por

$$D_{\mathbf{u}}f(x, y) = \nabla f(x, y) \bullet \mathbf{u}.$$

Teorema 14.3. El gradiente $\nabla f(x, y)$ siempre indica la dirección de máximo crecimiento de la función de la función f en el punto (x, y) .

Demostración. Consideremos todas las direcciones $\mathbf{u}$ de forma que $\|\mathbf{u}\| = 1$. Si evaluamos la derivada direccional para calcular la dirección de máximo crecimiento para la función f obtenemos que

$$\begin{aligned} D_{\mathbf{u}}f(x, y) &= \nabla f(x, y) \bullet \mathbf{u} = \|\nabla f(x, y)\| \|\mathbf{u}\| \cos \angle(\nabla f(x, y), \mathbf{u}) \\ &= \|\nabla f(x, y)\| \cos \angle(\nabla f(x, y), \mathbf{u}). \end{aligned}$$

Ahora bien la derivada direccional tomará el valor máximo en un punto (x, y) cuando

$$\cos \angle(\nabla f(x, y), \mathbf{u}) = 1,$$

es decir cuando $\nabla f(x, y)$ y $\mathbf{u}$ apunten en la misma dirección. ■

Sea $f(\mathbf{x})$ una función definida en un dominio D . Dado un punto $\mathbf{x}^*$ en D diremos que es un mínimo (respectivamente, un máximo) si

$$f(\mathbf{x}^*) \leq f(\mathbf{x}) \text{ (respectivamente, } f(\mathbf{x}^*) \geq f(\mathbf{x}) \text{)}.$$

para todo $\mathbf{x}$ en D

Corolario 14.1. Si $\mathbf{x}^*$ es un máximo o un mínimo para f , entonces

$$\nabla f(\mathbf{x}^*) = \mathbf{0}.$$

Teorema 14.4. El gradiente es perpendicular a las curvas de nivel.

Demostración. Consideremos la curva de nivel de una función f caracterizada por la expresión

$$f(x, y) = c$$

siendo c una constante fijada. Supongamos que podemos describir localmente la curva empleando una parametrización

$$\gamma(t) = (x(t), y(t))$$

para t tomando valores en un cierto intervalo I , de forma que

$$f(x(t), y(t)) = c.$$

Derivando mediante el empleo de la regla de la cadena obtenemos

$$f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) = 0,$$

esto es,

$$\begin{bmatrix} f_x(x(t), y(t)) \\ f_y(x(t), y(t)) \end{bmatrix} \bullet \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = 0.$$

Como el vector

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix}$$

es la tangente a la curva nivel, obtenemos que el vector gradiente

$$\begin{bmatrix} f_x(x(t), y(t)) \\ f_y(x(t), y(t)) \end{bmatrix}$$

es perpendicular a la misma (ver la Figura 1).

■

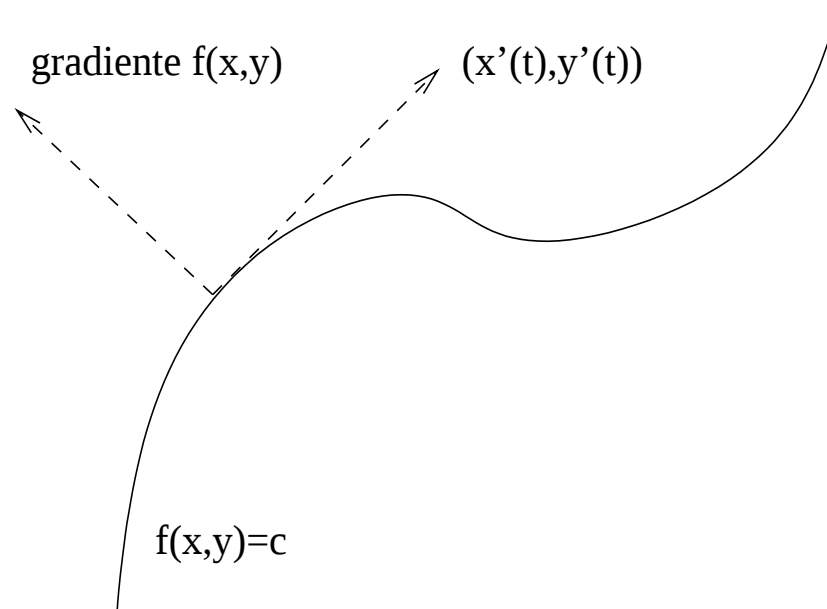


Figura 14.6: El gradiente es perpendicular a las curvas de nivel.

Capítulo 15

Optimización

15.1. Introducción

Supongamos que deseamos resolver el siguiente programa

$$\max_{(x,y)} z,$$

siendo $z = f(x, y) = -2x^2 - y^2 + 4x + 4y - 2$. Conocemos que los máximos y mínimos son puntos en los que el gradiente se tiene que anular, esto es, (x, y) será un candidato a máximo o mínimo si cumple que:

$$\nabla f(x, y) = \begin{bmatrix} -4x + 4 \\ -2y + 4 \end{bmatrix} = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Igualando componente a componente tenemos que resolver el sistema lineal

$$\begin{aligned} -4x + 4 &= 0 \\ -2y + 4 &= 0 \end{aligned}$$

cuya solución es $x = 1$ e $y = 2$. En consecuencia, el único punto candidato a máximo o mínimo es el punto $(1, 2)$. ¿Cómo podemos determinar si es máximo o mínimo? ¿y si es local o global?

Una primera observación importante es que el conjunto posibles soluciones del anterior problema es todo el dominio D de la función f que es $\mathbb{R}^2$. Deberíamos pues haber planteado el problema del modo siguiente:

$$\max_{(x,y) \in \mathbb{R}^2} z.$$

Se dice entonces que el conjunto *factible* de soluciones del problema de optimización es $\mathbb{R}^2$.

Veamos un resultado útil que extiende a varias variables, el resultado que nos dice que una función continua definida en un intervalo cerrado tiene máximo y mínimo absoluto. En primer lugar extenderemos la definición de intervalo cerrado para conjuntos de $\mathbb{R}^n$.

Dado un conjunto D llamaremos *frontera* del conjunto D al conjunto de puntos que separan a D de los puntos que no están dentro del conjunto (ver Figura 1). Se llaman *puntos interiores* a los puntos que no están situados en la frontera. Un conjunto D en $\mathbb{R}^n$ se dice *abierto* si no tiene puntos frontera. Si contiene a sus puntos frontera se dice que es un conjunto *cerrado*.

Ejemplo 15.1. Un intervalo abierto (a, b) sus puntos frontera están formados por el conjunto de puntos $\{a, b\}$, como no contiene los puntos frontera es un conjunto abierto.

Ejemplo 15.2. Un intervalo cerrado $[a, b]$ sus puntos frontera están formados por el conjunto de puntos $\{a, b\}$, como contiene los puntos frontera es un conjunto cerrado.

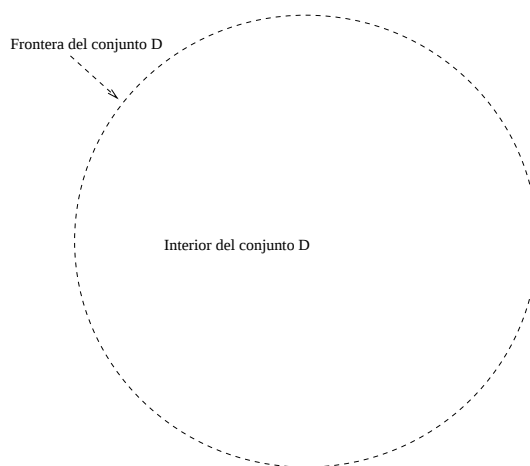
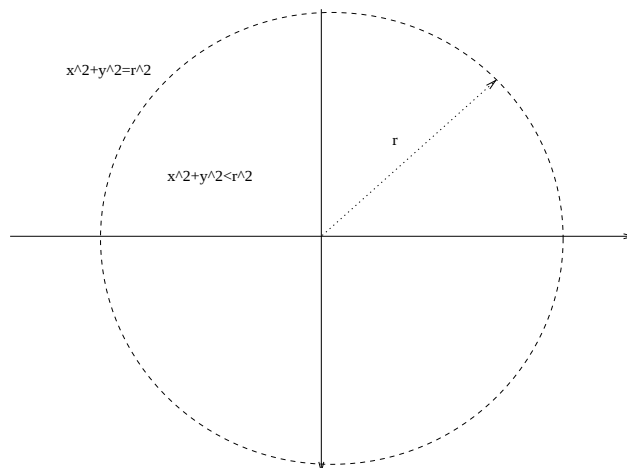


Figura 15.1: Los puntos frontera y los puntos interiores.

Figura 15.2: El conjunto abierto $x^2 + y^2 < r^2$.

Ejemplo 15.3. Un intervalo $(a, b]$ sus puntos frontera están formados por el conjunto de puntos $\{a, b\}$, como contiene a un punto frontera y a otro no, no es ni abierto ni cerrado.

Ejemplo 15.4. En $\mathbb{R}^2$ el prototipo de conjunto abierto es un disco centrado en el origen de radio r que no contiene a los puntos del perímetro que cumplen que $x^2 + y^2 = r^2$. Los puntos interiores cumplen $x^2 + y^2 < r^2$ (ver la Figura 2)

Diremos que un conjunto D de $\mathbb{R}^n$ es *acotado* si existe un radio $r > 0$ de forma que D está contenido en un conjunto del tipo $x_1^2 + x_2^2 + \dots + x_n^2 < r^2$.

Teorema 15.1 (Teorema de los valores extremos). Sea $f(\mathbf{x})$ una función continua definida en un dominio D de $\mathbb{R}^n$ cerrado y acotado. Entonces f tiene máximo y mínimo absoluto en D , esto es, existe $\mathbf{x}_m$ y $\mathbf{x}_M$ en D de forma que

$$f(\mathbf{x}_m) \leq f(\mathbf{x}) \leq f(\mathbf{x}_M) \text{ para cualquier } \mathbf{x} \text{ en } D.$$

Ejemplo 15.5. Resolver el problema

$$\min_{(x,y) \in D} z,$$

siendo $z = f(x, y) = x^2 + y^2 + y - 1$ y $D = \{(x, y) : x^2 + y^2 \leq 1\}$.

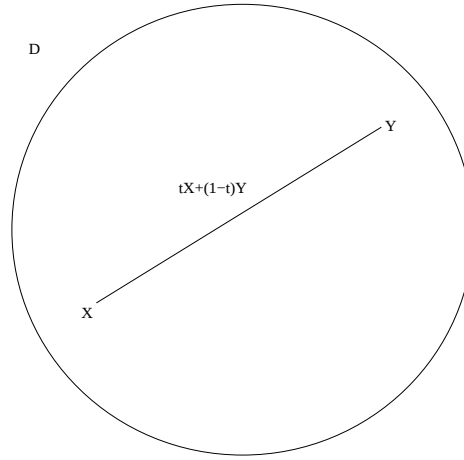


Figura 15.3: La propiedad de la convexidad de un conjunto.

Como el conjunto D es cerrado y acotado y la función z es continua en D , el Teorema de los valores extremos nos asegura que existe máximo y mínimo absoluto en D . Para poder calcularlos, en primer lugar encontraremos los puntos candidatos a máximos o mínimos en el interior de D , que son puntos en los que el gradiente se tiene que anular, y en una segunda fase buscaremos los máximos y mínimos de f sobre los puntos frontera.

Calculemos los puntos en D donde el gradiente de f se anula:

$$\begin{aligned} 2x &= 0, \\ 2y + 1 &= 0. \end{aligned}$$

Luego $x = 0$ e $y = -1/2$ como se cumple que $x^2 + y^2 = 0^2 + (-1/2)^2 = 1/4 < 1$, tenemos que el punto $(0, 1/2)$ está en el interior del conjunto D .

Veamos el comportamiento de f en la frontera de D formado por los puntos que cumplen $x^2 + y^2 = 1$, entonces $f(x, y) = (x^2 + y^2) + y - 1 = 1 + y - 1 = y$ cuando D toma valores en la frontera. como y solo toma valores en el intervalo cerrado $[-1, 1]$, la función alcanza un mínimo para $y = -1$ y un máximo para $y = 1$, para los que x solo puede tomar el valor 0 (recuerda que $x^2 + y^2 = 1$).

Los tres puntos a testar son entonces $(0, -1/2)$, $(0, 1)$, $(0, -1)$. Sustituyendo en la función obtenemos:

$$\begin{aligned} f(0, -1/2) &= -5/4 \\ f(0, 1) &= 1 \\ f(0, -1) &= -1 \end{aligned}$$

Luego se alcanza el mínimo absoluto en $(0, -1/2)$ y el máximo absoluto en $(0, 1)$.

15.2. Funciones cóncavas y convexas

Un conjunto D es *convexo* si para cualquier par de puntos del mismo el segmento que une a ambos puntos está contenido dentro del mismo. Esto quiere decir que dados $\mathbf{x}$ e $\mathbf{y}$ en D se tiene que $t\mathbf{x} + (1-t)\mathbf{y}$ está en D para cualquier t del intervalo $[0, 1]$ (ver la Figura 3).

Una función f definida sobre un conjunto D se dice *convexa*, si D es convexo y el segmento que une dos puntos cualesquiera de la gráfica, esto es,

$$tf(\mathbf{x}) + (1-t)f(\mathbf{y})$$

está siempre por encima de la gráfica de la función, esto es,

$$tf(\mathbf{x}) + (1-t)f(\mathbf{y}) \geq f(t\mathbf{x} + (1-t)\mathbf{y})$$

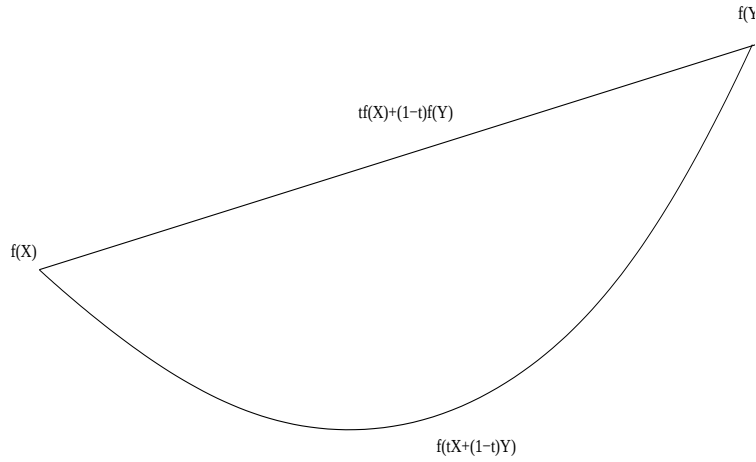


Figura 15.4: La propiedad de la convexidad de una función.

(ver la Figura 4).

Si la desigualdad se cumple al contrario, esto es,

$$tf(\mathbf{x}) + (1-t)f(\mathbf{y}) \leq f(t\mathbf{x} + (1-t)\mathbf{y})$$

se dice que la función es *cóncava*.

De forma evidente se cumple que si f es cóncava, entonces $-f$ es convexa y viceversa.

Ejemplo 15.6. La función lineal

$$f(x_1, x_2, \dots, x_n) = f(\mathbf{x}) = \mathbf{a} \bullet \mathbf{x} + b = a_1x_1 + a_2x_2 + \dots + a_nx_n + b$$

donde $a_1, a_2, \dots, a_n$ y b son constantes es cóncava y convexa a la vez. Nótese que

$$\begin{aligned} f(t\mathbf{x} + (1-t)\mathbf{y}) &= (\mathbf{a} \bullet (t\mathbf{x} + (1-t)\mathbf{y})) + b \\ &= (\mathbf{a} \bullet t\mathbf{x} + \mathbf{a} \bullet (1-t)\mathbf{y}) + (1-t)b + tb \\ &= t(\mathbf{a} \bullet \mathbf{x} + b) + (1-t)(\mathbf{a} \bullet \mathbf{y} + b) \\ &= tf(\mathbf{x}) + (1-t)f(\mathbf{y}). \end{aligned}$$

El resultado siguiente no permitirá determinar condiciones necesarias y suficientes para que un punto sea máximo o mínimo global cuando el conjunto factible es convexo.

Teorema 15.2. Supongamos que f tiene derivadas parciales continuas en un conjunto convexo D de $\mathbb{R}^n$ y sea $\mathbf{x}^*$ un punto interior de D . Entonces

1. Si f es convexa, $\mathbf{x}^*$ es un mínimo global de f en D si y solo si $\nabla f(\mathbf{x}^*) = \mathbf{0}$.
2. Si f es cóncava, $\mathbf{x}^*$ es un máximo global de f en D si y solo si $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Demostración. Veamos 1. la demostración de 2. se obtiene tomando $-f$ que sabemos que es convexa. Si $\mathbf{x}^*$ es un mínimo global sabemos que se tiene que cumplir que $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Veamos ahora que si $\nabla f(\mathbf{x}^*) = \mathbf{0}$, entonces $\mathbf{x}^*$ es un mínimo global. Para ello emplearemos el siguiente lema

Lema 15.1. f es convexa en D si y solo si para cualquier $\mathbf{x}$ e $\mathbf{y}$ en D se cumple

$$f(\mathbf{x}) - f(\mathbf{y}) \geq \nabla f(\mathbf{y}) \bullet (\mathbf{x} - \mathbf{y}).$$

De forma clara al ser f convexa obtenemos del lema tomando $\mathbf{y} = \mathbf{x}^*$ que

$$f(\mathbf{x}) - f(\mathbf{x}^*) \geq \nabla f(\mathbf{x}^*) \bullet (\mathbf{x} - \mathbf{x}^*) = \mathbf{0} \bullet (\mathbf{x} - \mathbf{x}^*) = 0,$$

para cualquier otro $\mathbf{x}$ en D , en consecuencia

$$f(\mathbf{x}) \geq f(\mathbf{x}^*)$$

para cualquier otro $\mathbf{x}$ en D y $\mathbf{x}^*$ es un mínimo global en D . ■

Ejemplo 15.7. La función $z = f(x, y) = x^2 + y^2 + y - 1$ definida sobre el dominio $D = \{(x, y) : x^2 + y^2 \leq 1\}$ convexo. ¿Es cóncava o convexa?

El problema se reduce ahora a determinar bajo que condiciones una función dada en un dominio convexo es cóncava o convexa.

Supongamos que f es convexa. Si definimos para $\mathbf{x}$ en D la función

$$w(t) = f(\mathbf{x} + t\mathbf{h}),$$

donde $\|\mathbf{h}\| < \delta$, entonces como la función $g(t) = \mathbf{x} + t\mathbf{h}$ es lineal es a su vez cóncava y convexa, además es fácil comprobar que entonces w es a su vez una función convexa de una variable. Luego se tiene que cumplir que $w''(t) > 0$, en particular,

$$w''(0) > 0$$

para cualquier $\mathbf{x}$ en D . Conocemos que

$$w'(t) = \nabla f(\mathbf{x} + t\mathbf{h}) \bullet \mathbf{h}.$$

La cuestión es que vale $w''(t)$. Veamos que ocurre con dos variables:

$$w'(t) = f_x(\mathbf{x} + t\mathbf{h})h_1 + f_y(\mathbf{x} + t\mathbf{h})h_2,$$

tenemos que derivar entonces

$$\frac{d}{dt} f_x(\mathbf{x} + t\mathbf{h}) \text{ y } \frac{d}{dt} f_y(\mathbf{x} + t\mathbf{h}).$$

Si volvemos a aplicar la regla de la cadena obtenemos

$$\frac{d}{dt} f_x(\mathbf{x} + t\mathbf{h}) = f_{xx}(\mathbf{x} + t\mathbf{h})h_1 + f_{xy}(\mathbf{x} + t\mathbf{h})h_2,$$

y

$$\frac{d}{dt} f_y(\mathbf{x} + t\mathbf{h}) = f_{yx}(\mathbf{x} + t\mathbf{h})h_1 + f_{yy}(\mathbf{x} + t\mathbf{h})h_2.$$

Entonces

$$\begin{aligned} w''(t) &= \frac{d}{dt} f_x(\mathbf{x} + t\mathbf{h})h_1 + \frac{d}{dt} f_y(\mathbf{x} + t\mathbf{h})h_2 \\ &= (f_{xx}(\mathbf{x} + t\mathbf{h})h_1 + f_{xy}(\mathbf{x} + t\mathbf{h})h_2)h_1 \\ &\quad + (f_{yx}(\mathbf{x} + t\mathbf{h})h_1 + f_{yy}(\mathbf{x} + t\mathbf{h})h_2)h_2 \\ &= f_{xx}(\mathbf{x} + t\mathbf{h})h_1^2 + f_{yy}(\mathbf{x} + t\mathbf{h})h_2^2 \\ &\quad + (f_{xy}(\mathbf{x} + t\mathbf{h}) + f_{yx}(\mathbf{x} + t\mathbf{h}))h_1h_2 \\ &= \begin{bmatrix} h_1 & h_2 \end{bmatrix} \begin{bmatrix} f_{xx}(\mathbf{x} + t\mathbf{h}) & f_{xy}(\mathbf{x} + t\mathbf{h}) \\ f_{yx}(\mathbf{x} + t\mathbf{h}) & f_{yy}(\mathbf{x} + t\mathbf{h}) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}. \end{aligned}$$

Luego se ha de cumplir que

$$w''(0) = \begin{bmatrix} h_1 & h_2 \end{bmatrix} \begin{bmatrix} f_{xx}(\mathbf{x}) & f_{xy}(\mathbf{x}) \\ f_{yx}(\mathbf{x}) & f_{yy}(\mathbf{x}) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} > 0.$$

para todo $\mathbf{x}$ en D y para cualquier $\|\mathbf{h}\| < \delta$.

Observación 15.1 (Importante). Fijémonos que si la matriz

$$\begin{bmatrix} f_{xx}(\mathbf{x}) & f_{xy}(\mathbf{x}) \\ f_{yx}(\mathbf{x}) & f_{yy}(\mathbf{x}) \end{bmatrix}$$

es simétrica, entonces la función

$$q(h_1, h_2) = \begin{bmatrix} h_1 & h_2 \end{bmatrix} \begin{bmatrix} f_{xx}(\mathbf{x}) & f_{xy}(\mathbf{x}) \\ f_{yx}(\mathbf{x}) & f_{yy}(\mathbf{x}) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$$

es una forma cuadrática. Conocemos, en este caso, que si los valores propios de la matriz de derivadas segundas son positivos entonces se cumple que $q(h_1, h_2) > 0$ y en consecuencia se cumplirá la condición de convexidad para la función f .

Ejemplo 15.8. Para la función $z = f(x, y) = x^2 + y^2 + y - 1$ se tiene

$$\begin{aligned} f_x &= 2x & f_y &= 2y + 1 \\ f_{xx} &= 2 & f_{xy} &= 0 \\ f_{yx} &= 0 & f_{yy} &= 2. \end{aligned}$$

Entonces la matriz

$$\begin{bmatrix} f_{xx}(\mathbf{x}) & f_{xy}(\mathbf{x}) \\ f_{yx}(\mathbf{x}) & f_{yy}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

es simétrica y diagonal. Los dos valores propios son $\lambda_1 = \lambda_2 = 2$ positivos. Luego la función f es convexa sobre cualquier dominio que sea convexo, en particular cualquier dominio convexo que contenga al mínimo de esta función este será mínimo global sobre el dominio considerado.

El siguiente resultado nos proporciona las condiciones bajo las cuales la matriz de derivadas parciales segundas va a ser una matriz simétrica.

Teorema 15.3 (Teorema de Young). Supongamos que dos derivadas parciales de orden m de la función f se obtenido con el mismo número de derivaciones con respecto a cada una de las variables y son continuas en un dominio abierto D . Entonces las dos derivadas parciales son iguales en todo punto de D .

En particular este resultado afirma que si $f_{xy}(x, y)$ y $f_{yx}(x, y)$ son continuas en D , entonces

$$f_{xy}(x, y) = f_{yx}(x, y).$$

Esto nos dice que bajo estas hipótesis la matriz de derivadas segundas

$$\begin{bmatrix} f_{xx}(\mathbf{x}) & f_{xy}(\mathbf{x}) \\ f_{yx}(\mathbf{x}) & f_{yy}(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} f_{xx}(\mathbf{x}) & f_{xy}(\mathbf{x}) \\ f_{xy}(\mathbf{x}) & f_{yy}(\mathbf{x}) \end{bmatrix}$$

es simétrica.

Podemos ahora establecer las condiciones necesarias y suficientes para determinar si una función en un dominio convexo es cóncava o convexa.

Dada una función $f(x_1, x_2, \dots, x_n)$ se llama *matriz hessiana* a la matriz formada por todas las derivadas parciales segundas

$$Hf(x_1, x_2, \dots, x_n) = Hf(\mathbf{x}) = \begin{bmatrix} f_{x_1 x_1}(\mathbf{x}) & f_{x_1 x_2}(\mathbf{x}) & \cdots & f_{x_1 x_n}(\mathbf{x}) \\ f_{x_2 x_1}(\mathbf{x}) & f_{x_2 x_2}(\mathbf{x}) & \cdots & f_{x_2 x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_n x_1}(\mathbf{x}) & f_{x_n x_2}(\mathbf{x}) & \cdots & f_{x_n x_n}(\mathbf{x}) \end{bmatrix}.$$

Teorema 15.4. Supongamos que f tiene derivadas parciales segundas continuas y está definida en un dominio abierto y convexo D de $\mathbb{R}^n$ entonces:

1. f es convexa si y solo si todos los valores propios de la matriz $Hf(\mathbf{x})$ son positivos para cualquier $\mathbf{x}$ en D .
2. f es cóncava si y solo si todos los valores propios de la matriz $Hf(\mathbf{x})$ son negativos para cualquier $\mathbf{x}$ en D .

15.3. Las condiciones de segundo orden

Vamos a establecer ahora las condiciones para determinar si un punto donde el gradiente se anula es un máximo o mínimo local.

Supongamos que $\mathbf{x}^*$ es un punto del dominio D donde

$$\nabla f(\mathbf{x}^*) = \mathbf{0}.$$

Consideremos la función de una variable

$$w(t) = f(\mathbf{x}^* + t\mathbf{h}),$$

donde $\|\mathbf{h}\| < \delta$. Entonces empleando el Teorema de Taylor tenemos que

$$w(t) = w(0) + w'(0)t + w''(0)t^2 + \varepsilon$$

donde ε mide el error que se comete al aproximar la función. Sustituyendo se tiene que

$$\begin{aligned} f(\mathbf{x}^* + t\mathbf{h}) &= f(\mathbf{x}^*) + (\nabla f(\mathbf{x}^*) \bullet \mathbf{h})t + (\mathbf{h}^T Hf(\mathbf{x}^*)\mathbf{h})t^2 + \varepsilon \\ &= f(\mathbf{x}^*) + (\mathbf{h}^T Hf(\mathbf{x}^*)\mathbf{h})t^2 + \varepsilon, \text{ recordemos que } \nabla f(\mathbf{x}^*) = \mathbf{0}, \end{aligned}$$

entonces

$$f(\mathbf{x}^* + t\mathbf{h}) - f(\mathbf{x}^*) = (\mathbf{h}^T Hf(\mathbf{x}^*)\mathbf{h})t^2 + \varepsilon$$

si consideramos que el efecto del error es pequeño, si todos los valores propios de la matriz hessiana $Hf(\mathbf{x}^*)$ en el punto donde el gradiente se anula son positivos, empleando coordenadas canónicas podemos afirmar que

$$\mathbf{h}^T Hf(\mathbf{x}^*)\mathbf{h} > 0$$

para todo $\|\mathbf{h}\| < \delta$ distinto del vector cero. En consecuencia, se cumple que

$$f(\mathbf{x}^* + t\mathbf{h}) - f(\mathbf{x}^*) > 0 \Leftrightarrow f(\mathbf{x}^* + t\mathbf{h}) > f(\mathbf{x}^*).$$

Esto nos dice que $\mathbf{x}^*$ es un mínimo relativo, ya que la desigualdad solo se cumple en puntos cercanos a $\mathbf{x}^*$.

Podemos enunciar entonces el siguiente teorema.

Teorema 15.5. Sea f una función con derivadas parciales segundas continuas en un dominio D . Supongamos que $\mathbf{x}^*$ es un punto interior de D de forma que $\nabla f(\mathbf{x}^*) = \mathbf{0}$. Entonces

1. Si todos los valores propios de la matriz hessiana $Hf(\mathbf{x}^*)$ son positivos entonces la función f tiene en $\mathbf{x}^*$ un mínimo relativo.
2. Si todos los valores propios de la matriz hessiana $Hf(\mathbf{x}^*)$ son negativos entonces la función f tiene en $\mathbf{x}^*$ un máximo relativo.

Ejemplo 15.9. Hallar los máximos, mínimos relativos de la función

$$f(x, y) = \log(x^2 + y^2 + 1).$$

Calculamos los puntos donde el gradiente se anula, para ello igualamos las derivadas parciales a cero y resolvemos:

$$f_x(x, y) = \frac{2x}{(x^2 + y^2 + 1)} = 0$$

$$f_y(x, y) = \frac{2y}{(x^2 + y^2 + 1)} = 0$$

entonces el único punto es

$$\mathbf{x}_0 = (0, 0).$$

Calculamos las derivadas segundas

$$f_{xx}(x, y) = -2 \frac{x^2 - y^2 - 1}{(x^2 + y^2 + 1)^2}, \quad f_{xx}(0, 0) = 2,$$

$$f_{xy}(x, y) = -4 \frac{xy}{(x^2 + y^2 + 1)^2}, \quad f_{xy}(0, 0) = 0,$$

$$f_{yy}(x, y) = 2 \frac{x^2 - y^2 + 1}{(x^2 + y^2 + 1)^2}, \quad f_{yy}(0, 0) = 2.$$

La matriz hessiana es entonces

$$Hf(0, 0) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

como la matriz es diagonal, los valores propios coinciden con los elementos de la diagonal, entonces el punto $(0, 0)$ es un mínimo local ya que todos son positivos.

Ejemplo 15.10. Analizar el comportamiento de $z = x^5y + xy^5 + xy$ en los puntos donde el gradiente se anula.

En primer lugar

$$z_x = 5x^4y + y^5 + y = 0;$$

$$z_y = x^5 + 5xy^4 + x = 0;$$

esto es

$$z_x = y(5x^4 + y^4 + 1) = 0;$$

$$z_y = x(x^4 + 5y^4 + 1) = 0;$$

como

$$5x^4 + y^4 + 1 > 0; \quad x^4 + 5y^4 + 1 > 0$$

entonces el único punto donde las derivadas parciales se anulan es

$$\mathbf{x}_0 = (0, 0).$$

Calculemos las derivadas segundas:

$$z_{xx}(x, y) = 20x^3y; z_{xx}(0, 0) = 0,$$

$$z_{xy}(x, y) = 5x^4 + 5y^4 + 1; z_{xy}(0, 0) = 1,$$

$$z_{yy}(x, y) = 20xy^3; z_{yy}(0, 0) = 0.$$

La matriz hessiana es entonces

$$Hf(0, 0) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

Como es diagonal y uno de los valores de la diagonal es cero, nos encontramos frente a un punto sobre el cual no podemos decidir.

Ejemplo 15.11. Calcular los extremos relativos y clasificarlos para la función

$$z = x_1 + 2x_3 + x_2x_3 - x_1^2 - x_2^2 - x_3^2.$$

En primer lugar calculamos los valores donde la derivada se anula:

$$\begin{aligned} z_{x_1} &= 1 - 2x_1 = 0, \\ z_{x_2} &= x_3 - 2x_2 = 0, \\ z_{x_3} &= 2 + x_2 - 2x_3 = 0. \end{aligned}$$

De la primera ecuación tenemos que $x_1 = 1/2$. En la segunda ecuación se tiene que $x_3 = 2x_2$, sustituyendo en la tercera se tiene que

$$2 + x_2 - 4x_2 = 0 \Rightarrow x_2 = 2/3 \Rightarrow x_3 = 4/3.$$

Luego el único punto donde se anula el gradiente es $\mathbf{x}^* = (1/2, 2/3, 4/3)$.

Calculamos ahora las derivadas parciales segundas

$$\begin{aligned} z_{x_1x_1}(x_1, x_2, x_3) &= -2, \quad z_{x_1x_1}(1/2, 2/3, 4/3) = -2, \\ z_{x_1x_2}(x_1, x_2, x_3) &= 0, \quad z_{x_1x_2}(1/2, 2/3, 4/3) = 0, \\ z_{x_1x_3}(x_1, x_2, x_3) &= 0, \quad z_{x_1x_3}(1/2, 2/3, 4/3) = 0, \\ z_{x_2x_2}(x_1, x_2, x_3) &= -2, \quad z_{x_2x_2}(1/2, 2/3, 4/3) = -2, \\ z_{x_2x_3}(x_1, x_2, x_3) &= 1, \quad z_{x_2x_3}(1/2, 2/3, 4/3) = 1, \\ z_{x_3x_3}(x_1, x_2, x_3) &= -2, \quad z_{x_3x_3}(1/2, 2/3, 4/3) = -2. \end{aligned}$$

La matriz hessiana es entonces

$$Hf(1/2, 2/3, 4/3) = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix}.$$

Para calcular los valores propios tenemos que calcular en primer lugar el polinomio característico

$$\begin{aligned} \det(Hf(1/2, 2/3, 4/3) - \lambda I) &= \det \begin{bmatrix} -2 - \lambda & 0 & 0 \\ 0 & -2 - \lambda & 1 \\ 0 & 1 & -2 - \lambda \end{bmatrix} \\ &\quad \text{desarrollamos por la primera columna} \\ &= (-2 - \lambda) \det \begin{bmatrix} -2 - \lambda & 1 \\ 1 & -2 - \lambda \end{bmatrix} \\ &= (-2 - \lambda)((-2 - \lambda)^2 - 1) \\ &= (-2 - \lambda)(\lambda^2 + 4\lambda + 3). \end{aligned}$$

A partir de esta expresión ya podemos conocer que $\lambda = -2$ es una raíz del polinomio característico. Para obtener las otras dos necesitamos resolver

$$\lambda^2 + 4\lambda + 3 = 0 \Rightarrow \lambda = \frac{-4 \pm \sqrt{16 - 12}}{2} = \begin{cases} -3 \\ -1 \end{cases}$$

Como los tres valores propios son negativos, la función z alcanza en el punto $(1/2, 2/3, 4/3)$ un máximo relativo.

15.4. Extremos restringidos y multiplicadores de Lagrange

- Recordemos como solucionábamos problemas del tipo

$$\begin{array}{ll} \text{máx} & x^2 + y^2 - 2xy \\ \text{s.a} & x^2 + y^2 - 1 = 0. \end{array}$$

- A la función a optimizar

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

se le llama **función objetivo**. A la función

$$g(x, y) = x^2 + y^2 - 1$$

se le llama **función restricción**.

- Este tipo de problemas se escriben de forma más general como

$$\begin{array}{ll} \text{máx} & f(\mathbf{x}) \\ \text{s.a} & g_1(\mathbf{x}) = 0 \\ & g_2(\mathbf{x}) = 0 \\ & \vdots \\ & g_m(\mathbf{x}) = 0 \end{array}$$

donde

$$f, g_1, g_2, \dots, g_m : \mathbb{R}^n \longrightarrow \mathbb{R}.$$

- Otra forma de expresarlo es

$$\text{máx } f(\mathbf{x})$$

donde

$$f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$$

siendo

$$\Omega = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) = 0, g_2(\mathbf{x}) = 0, \dots, g_m(\mathbf{x}) = 0\}$$

el conjunto de soluciones de m -ecuaciones con n -incógnitas.

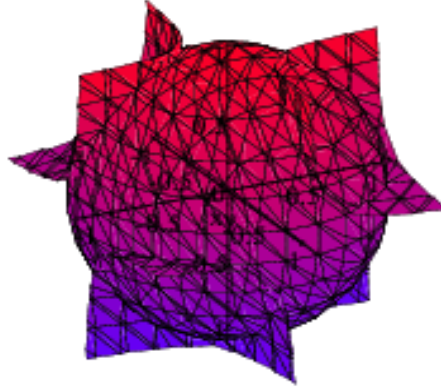
- Nótese que los $\mathbf{x}$ que satisfacen

$$g_1(\mathbf{x}) = 0$$

están situados en una superficie de nivel para la función g_1 con constante 0. Entonces

$$\Omega = \{\mathbf{x} : g_1(\mathbf{x}) = 0\} \cap \{\mathbf{x} : g_2(\mathbf{x}) = 0\} \cap \dots \cap \{\mathbf{x} : g_m(\mathbf{x}) = 0\}$$

es la intersección de m -superficies de nivel.



■ **Ejemplo:** Sean

$$g_1(x, y, z) = x^2 + y^2 + z^2 - 1 = 0,$$

$$g_2(x, y, z) = xyz - 0.05 = 0.$$

gráficamente

Teorema 15.6 (Método de los multiplicadores de Lagrange). Sean $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ y $g : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ funciones de clase $\mathcal{C}^1$ dadas. Sean $\mathbf{x}_0 \in U$ de forma que $g_1(\mathbf{x}_0) = 0$. Supongamos que

$$\nabla g(\mathbf{x}_0) \neq \mathbf{0}.$$

Si $\mathbf{x}_0$ es una solución del programa

$$\begin{array}{ll} \text{máx} & f(\mathbf{x}) \\ \text{s.a} & g(\mathbf{x}) = 0 \end{array}$$

entonces existe un número real λ de forma que

$$\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0),$$

esto es,

$$\nabla f(\mathbf{x}_0) \in \text{Gen} \{ \nabla g(\mathbf{x}_0) \}.$$

■ **Demostración:** Consideremos trayectorias

$$\mathbf{c} : [-\varepsilon, \varepsilon] \rightarrow \{ \mathbf{x} : g(\mathbf{x}) = 0 \}$$

de forma que $\mathbf{c}(0) = \mathbf{x}_0$. Como $g[\mathbf{c}(t)] = 0$ para todo $t \in [-\varepsilon, \varepsilon]$, se tiene que

$$\left. \frac{d}{dt} g[\mathbf{c}(t)] \right|_{t=0} = \nabla g(\mathbf{x}_0) \cdot \mathbf{c}'(0) = 0.$$

En consecuencia $\mathbf{c}'(0)$ es ortogonal a $\nabla g(\mathbf{x}_0)$. Si f tiene un máximo en $\{ \mathbf{x} : g(\mathbf{x}) = 0 \}$ situado en $\mathbf{x}_0$, entonces $f[\mathbf{c}(t)]$ tiene un máximo en $t = 0$. En consecuencia

$$\left. \frac{d}{dt} f[\mathbf{c}(t)] \right|_{t=0} = \nabla f(\mathbf{x}_0) \cdot \mathbf{c}'(0) = 0,$$

y por lo tanto $\mathbf{c}'(0)$ es ortogonal a $\nabla f(\mathbf{x}_0)$. Como $\nabla f(\mathbf{x}_0)$ es ortogonal al vector tangente de cualquier curva en $\{\mathbf{x} : g(\mathbf{x}) = 0\}$ que pase por $\mathbf{x}_0$, entonces tiene que ser ortogonal al plano tangente a $\{\mathbf{x} : g(\mathbf{x}) = 0\}$ en $\mathbf{x}_0$. Del mismo modo como $\nabla g(\mathbf{x}_0)$ es ortogonal al vector tangente de cualquier curva en $\{\mathbf{x} : g(\mathbf{x}) = 0\}$ que pase por $\mathbf{x}_0$, entonces tiene que ser ortogonal al plano tangente a $\{\mathbf{x} : g(\mathbf{x}) = 0\}$ en $\mathbf{x}_0$. Como dos vectores ortogonales a un plano tienen que ser linealmente dependientes, obtenemos que $\nabla f(\mathbf{x}_0)$ y $\nabla g(\mathbf{x}_0)$ tienen que ser linealmente dependientes. Entonces debe de existir un número real λ de forma que $\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0)$ ■

- Si ahora consideramos el problema más general

$$\begin{array}{ll} \text{máx} & f(\mathbf{x}) \\ \text{s.a} & g_1(\mathbf{x}) = 0 \\ & g_2(\mathbf{x}) = 0 \\ & \vdots \\ & g_m(\mathbf{x}) = 0 \end{array}$$

podemos escribirlo del modo siguiente: Sea

$$\mathbf{g} : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

definida por

$$\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_m(\mathbf{x}))$$

entonces el programa puede escribirse como

$$\begin{array}{ll} \text{máx} & f(\mathbf{x}) \\ \text{s.a} & \mathbf{g}(\mathbf{x}) = \mathbf{0} \end{array}$$

En este caso

$$\{\mathbf{x} : \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$$

también define un conjunto de nivel. Dado un punto $\mathbf{x}_0 \in U$ de forma que

$$\mathbf{g}(\mathbf{x}_0) = \mathbf{0}$$

¿Cuál es el espacio tangente al conjunto de nivel en este punto?

- En el caso anterior un vector $\mathbf{v}$ pertenecía al plano tangente $\{\mathbf{x} : g(\mathbf{x}) = 0\}$ en $\mathbf{x}_0$ si

$$\nabla g(\mathbf{x}_0) \cdot \mathbf{v} = 0$$

o bien

$$Dg(\mathbf{x}_0) \mathbf{v} = 0$$

esto es

$$\mathbf{v} \in \text{Nul } Dg(\mathbf{x}_0).$$

- Se define entonces el espacio tangente a $\{\mathbf{x} : \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$ en $\mathbf{x}_0$ como el conjunto de soluciones del sistema homogéneo

$$D\mathbf{g}(\mathbf{x}_0) \mathbf{x} = \mathbf{0}.$$

Esto es, $\mathbf{v}$ pertenece al espacio tangente a $\{\mathbf{x} : \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$ en $\mathbf{x}_0$ si y solo si

$$D\mathbf{g}(\mathbf{x}_0) \mathbf{v} = \mathbf{0}.$$

Recordemos que

$$\text{Nul } D\mathbf{g}(\mathbf{x}_0) = (\text{fila } D\mathbf{g}(\mathbf{x}_0))^\perp.$$

- **Ejemplo:** Consideremos la superficie de nivel definida a partir de las soluciones del sistema

$$g_1(x, y, z) = x^2 + y^2 + z^2 - 1 = 0,$$

$$g_2(x, y, z) = \sqrt{3}x + \sqrt{3}y + \sqrt{3}z - 3 = 0.$$

y el punto

$$\mathbf{x}_0 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right)$$

que satisface ambas ecuaciones. Entonces el espacio tangente se calcula definiendo

$$\mathbf{g}(x, y, z) = (x^2 + y^2 + z^2 - 1, \sqrt{3}x + \sqrt{3}y + \sqrt{3}z - 3).$$

Calculamos en primer lugar

$$D\mathbf{g}(\mathbf{x}) = \begin{bmatrix} 2x & 2y & 2z \\ \sqrt{3} & \sqrt{3} & \sqrt{3} \end{bmatrix}$$

y sustituimos

$$D\mathbf{g}\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) = \begin{bmatrix} \frac{2}{\sqrt{3}} & \frac{2}{\sqrt{3}} & \frac{2}{\sqrt{3}} \\ \sqrt{3} & \sqrt{3} & \sqrt{3} \end{bmatrix}.$$

Entonces los vectores del espacio tangente a la superficie de nivel son soluciones del sistema homogéneo

$$D\mathbf{g}\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)\mathbf{x} = \begin{bmatrix} \frac{2}{\sqrt{3}} & \frac{2}{\sqrt{3}} & \frac{2}{\sqrt{3}} \\ \sqrt{3} & \sqrt{3} & \sqrt{3} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

esto es,

$$\begin{bmatrix} \frac{2}{\sqrt{3}} & \frac{2}{\sqrt{3}} & \frac{2}{\sqrt{3}} \\ \sqrt{3} & \sqrt{3} & \sqrt{3} \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

en consecuencia el sistema en forma escalonada reducida es

$$\begin{aligned} x + y + z &= 0, \\ 0 &= 0. \end{aligned}$$

Entonces

$$x = -y - z$$

y los vectores $\mathbf{x}$ del espacio tangente se escriben como

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -y - z \\ y \\ z \end{bmatrix} = y \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

- Empleando el mismo argumento que en la demostración del Teorema 9 obtenemos que $\nabla f(\mathbf{x}_0)$ tiene que ser ortogonal al espacio tangente al conjunto de nivel en $\mathbf{x}_0$ siempre y cuando $\mathbf{x}_0$ sea una solución del programa. Por otro lado, cualquier vector ortogonal al espacio tangente tiene que estar generado por el espacio fila de $D\mathbf{g}(\mathbf{x}_0)$, que esta generado a su vez por

$$\{\nabla g_1(\mathbf{x}_0), \nabla g_2(\mathbf{x}_0), \dots, \nabla g_m(\mathbf{x}_0)\}.$$

En consecuencia obtenemos el siguiente resultado:

Teorema 15.7 (Multiplicadores de Lagrange). Sean

$$f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$$

y

$$g_1, \dots, g_m : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$$

funciones de clase $\mathcal{C}^1$ dadas. Sean $\mathbf{x}_0 \in U$ de forma que

$$g_1(\mathbf{x}_0) = 0, g_2(\mathbf{x}_0) = 0, \dots, g_m(\mathbf{x}_0) = 0.$$

Supongamos que

$$\nabla g_1(\mathbf{x}_0), \nabla g_2(\mathbf{x}_0), \dots, \nabla g_m(\mathbf{x}_0)$$

son linealmente independientes (luego $m \leq n$). Si $\mathbf{x}_0$ es una solución del programa

$$\begin{array}{ll} \text{máx} & f(\mathbf{x}) \\ \text{s.a} & \mathbf{g}(\mathbf{x}) = 0 \end{array}$$

entonces existen m -números reales $\lambda_1, \lambda_2, \dots, \lambda_m$ de forma que

$$\nabla f(\mathbf{x}_0) = \lambda_1 \nabla g_1(\mathbf{x}_0) + \lambda_2 \nabla g_2(\mathbf{x}_0) + \dots + \lambda_m \nabla g_m(\mathbf{x}_0),$$

esto es,

$$\nabla f(\mathbf{x}_0) \in \text{Gen} \{ \nabla g_1(\mathbf{x}_0), \nabla g_2(\mathbf{x}_0), \dots, \nabla g_m(\mathbf{x}_0) \}.$$

Corolario 15.1. Las posibles soluciones del programa

$$\begin{array}{ll} \text{máx} & f(\mathbf{x}) \\ \text{s.a} & \mathbf{g}(\mathbf{x}) = 0 \end{array}$$

tienen que ser soluciones del sistema de ecuaciones

$$\begin{aligned} f_{x_1}(\mathbf{x}) &= \lambda_1 (g_1)_{x_1}(\mathbf{x}) + \lambda_2 (g_2)_{x_1}(\mathbf{x}) + \dots + \lambda_m (g_m)_{x_1}(\mathbf{x}) \\ f_{x_2}(\mathbf{x}) &= \lambda_1 (g_1)_{x_2}(\mathbf{x}) + \lambda_2 (g_2)_{x_2}(\mathbf{x}) + \dots + \lambda_m (g_m)_{x_2}(\mathbf{x}) \\ &\vdots \\ f_{x_n}(\mathbf{x}) &= \lambda_1 (g_1)_{x_n}(\mathbf{x}) + \lambda_2 (g_2)_{x_n}(\mathbf{x}) + \dots + \lambda_m (g_m)_{x_n}(\mathbf{x}) \\ g_1(\mathbf{x}) &= 0 \\ g_2(\mathbf{x}) &= 0 \\ &\vdots \\ g_m(\mathbf{x}) &= 0 \end{aligned}.$$

donde las incógnitas son

$$(\mathbf{x}, \lambda) = (x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m).$$

■ **Ejemplo:** Hallar las posibles soluciones del programa

$$\begin{array}{ll} \text{máx} & x + y + z \\ \text{s.a.} & x^2 + y^2 = 2 \\ & x + z = 1 \end{array}$$

• **Solución:** En este caso

$$f(x, y, z) = x + y + z$$

$$g_1(x, y, z) = x^2 + y^2 - 2 = 0,$$

$$g_2(x, y, z) = x + z - 1 = 0$$

Calculamos

$$\nabla f(x, y, z) = (1, 1, 1),$$

$$\nabla g_1(x, y, z) = (2x, 2y, 0),$$

$$\nabla g_2(x, y, z) = (1, 0, 1).$$

Planteamos ahora

$$\nabla f(\mathbf{x}) = \lambda_1 \nabla g_1(\mathbf{x}) + \lambda_2 \nabla g_2(\mathbf{x})$$

esto es

$$(1, 1, 1) = \lambda_1 (2x, 2y, 0) + \lambda_2 (1, 0, 1) = (2x\lambda_1 + \lambda_2, 2y\lambda_1, \lambda_2)$$

igualando componente a componente obtenemos

$$\begin{aligned} 1 &= 2x\lambda_1 + \lambda_2 \\ 1 &= 2y\lambda_1 \\ 1 &= \lambda_2 \end{aligned}$$

y añadimos las restricciones

$$\begin{aligned} x^2 + y^2 &= 2 \\ x + z &= 1 \end{aligned}.$$

Hay pues que resolver

$$\begin{aligned} 1 &= 2x\lambda_1 + \lambda_2 \\ 1 &= 2y\lambda_1 \\ 1 &= \lambda_2 \\ 2 &= x^2 + y^2 \\ 1 &= x + z \end{aligned}$$

Entonces

$$\lambda_2 = 1 \implies (\text{sustituimos en la 1ª ec.}) 1 = 2x\lambda_1 + 1 \implies 2x\lambda_1 = 0,$$

(si $\lambda_1 = 0 \implies$ (sustituimos en la 2ª ec.) $1 = 2y\lambda_1 = 0$ una contradicción) luego $\lambda_1 \neq 0$, esto implica

$$x = 0 \implies (\text{sustituimos en la 5ª ec.}) z = 1$$

y en la 4ª ecuación obtenemos

$$y = \pm\sqrt{2},$$

y en la 2ª ecuación obtenemos

$$\lambda_1 = \pm \frac{1}{2\sqrt{2}}.$$

Los posibles óptimos son

$$\mathbf{x}_0^1 = (0, \sqrt{2}, 1) \text{ con } \lambda_1 = \frac{1}{2\sqrt{2}}, \lambda_2 = 1$$

y

$$\mathbf{x}_0^2 = (0, -\sqrt{2}, 1) \text{ con } \lambda_1 = -\frac{1}{2\sqrt{2}}, \lambda_2 = 1$$

■ **Cuestión:** ¿Tiene solución el programa de optimización anterior?

15.4.1. Condiciones suficientes de segundo orden

- Supongamos que $\mathbf{x}_0$ es un máximo, entonces para cualquier trayectoria

$$\mathbf{c} : [-\varepsilon, \varepsilon] \longrightarrow \{\mathbf{x} : g(\mathbf{x}) = 0\}$$

de forma que $\mathbf{c}(0) = \mathbf{x}_0$ y donde $\mathbf{g}[\mathbf{c}(t)] = 0$ para todo $t \in [-\varepsilon, \varepsilon]$, se tiene que

$$\left. \frac{d^2}{dt^2} f[\mathbf{c}(t)] \right|_{t=0} \leq 0.$$

Empleando la regla de la cadena se tiene que

$$\left. \frac{d^2}{dt^2} f[\mathbf{c}(t)] \right|_{t=0} = \nabla f(\mathbf{x}_0) \cdot \mathbf{c}''(0) + \mathbf{c}'(0)^T Hf(\mathbf{x}_0) \mathbf{c}'(0).$$

Además si derivamos dos veces la expresión

$$\lambda_1 g_1[\mathbf{c}(t)] + \lambda_2 g_2[\mathbf{c}(t)] + \cdots + \lambda_m g_m[\mathbf{c}(t)] = 0$$

obtenemos

$$\sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}_0) \cdot \mathbf{c}''(0) + \sum_{i=1}^m \lambda_i \mathbf{c}'(0)^T Hg_i(\mathbf{x}_0) \mathbf{c}'(0) = 0.$$

Restando ambas derivadas segundas obtenemos

$$\begin{aligned} & \left. \frac{d^2}{dt^2} f[\mathbf{c}(t)] \right|_{t=0} \\ &= \nabla f(\mathbf{x}_0) \cdot \mathbf{c}''(0) + \mathbf{c}'(0)^T Hf(\mathbf{x}_0) \mathbf{c}'(0) \\ & \quad - \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}_0) \cdot \mathbf{c}''(0) - \sum_{i=1}^m \lambda_i \mathbf{c}'(0)^T Hg_i(\mathbf{x}_0) \mathbf{c}'(0) \\ &= \left(\nabla f(\mathbf{x}_0) - \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}_0) \right) \cdot \mathbf{c}''(0) \\ & \quad + \mathbf{c}'(0)^T \left(Hf(\mathbf{x}_0) - \sum_{i=1}^m \lambda_i Hg_i(\mathbf{x}_0) \right) \mathbf{c}'(0) \\ &= \mathbf{c}'(0)^T \left(Hf(\mathbf{x}_0) - \sum_{i=1}^m \lambda_i Hg_i(\mathbf{x}_0) \right) \mathbf{c}'(0) \leq 0. \end{aligned}$$

ya que al ser $\mathbf{x}_0$ un máximo tiene que cumplir que

$$\nabla f(\mathbf{x}_0) - \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}_0) = \mathbf{0}.$$

Teorema 15.8 (Condición de segundo orden). Si $\mathbf{x}_0$ es una solución del programa

$$\begin{array}{ll} \text{máx} & f(\mathbf{x}) \\ \text{s.a} & \mathbf{g}(\mathbf{x}) = 0 \end{array}$$

entonces

$$\mathbf{v}^T \left(Hf(\mathbf{x}_0) - \sum_{i=1}^m \lambda_i Hg_i(\mathbf{x}_0) \right) \mathbf{v} \leq 0$$

para cualquier vector $\mathbf{v}$ del espacio tangente a $\{\mathbf{x} : \mathbf{g}(\mathbf{x}) = 0\}$, esto es para cualquier $\mathbf{v}$ que cumple

$$D\mathbf{g}(\mathbf{x}_0) \mathbf{v} = \mathbf{0}.$$

■ La condición

$$\mathbf{v}^T \left(Hf(\mathbf{x}_0) - \sum_{i=1}^m \lambda_i Hg_i(\mathbf{x}_0) \right) \mathbf{v} \leq 0$$

nos dice que

$$\left(Hf(\mathbf{x}_0) - \sum_{i=1}^m \lambda_i Hg_i(\mathbf{x}_0) \right)$$

es la matriz de una forma cuadrática semidefinida negativa sobre el espacio tangente a $\{\mathbf{x} : \mathbf{g}(\mathbf{x}) = 0\}$.

Teorema 15.9 (Condición suficiente). Supongamos que $\mathbf{x}_0$ es solución del sistema de ecuaciones

$$\begin{aligned} f_{x_1}(\mathbf{x}) &= \lambda_1 (g_1)_{x_1}(\mathbf{x}) + \lambda_2 (g_2)_{x_1}(\mathbf{x}) + \cdots + \lambda_m (g_m)_{x_1}(\mathbf{x}) \\ f_{x_2}(\mathbf{x}) &= \lambda_1 (g_1)_{x_2}(\mathbf{x}) + \lambda_2 (g_2)_{x_2}(\mathbf{x}) + \cdots + \lambda_m (g_m)_{x_2}(\mathbf{x}) \\ &\vdots \\ f_{x_n}(\mathbf{x}) &= \lambda_1 (g_1)_{x_n}(\mathbf{x}) + \lambda_2 (g_2)_{x_n}(\mathbf{x}) + \cdots + \lambda_m (g_m)_{x_n}(\mathbf{x}) \quad . \\ g_1(\mathbf{x}) &= 0 \\ g_2(\mathbf{x}) &= 0 \\ &\vdots \\ g_m(\mathbf{x}) &= 0 \end{aligned}$$

para unos números reales $\lambda_1, \lambda_2, \dots, \lambda_m$ adecuados. Además

$$\mathbf{v}^T \left(Hf(\mathbf{x}_0) - \sum_{i=1}^m \lambda_i Hg_i(\mathbf{x}_0) \right) \mathbf{v} < 0$$

para cualquier vector $\mathbf{v}$ del espacio tangente a $\{\mathbf{x} : \mathbf{g}(\mathbf{x}) = 0\}$, esto es para cualquier $\mathbf{v}$ que cumple

$$D\mathbf{g}(\mathbf{x}_0) \mathbf{v} = \mathbf{0},$$

entonces $\mathbf{x}_0$ es solución del programa

$$\begin{array}{ll} \text{máx} & f(\mathbf{x}) \\ \text{s.a} & \mathbf{g}(\mathbf{x}) = \mathbf{0} \end{array} .$$

■ ¿Cómo podemos obtener la matriz

$$L(\mathbf{x}_0, \lambda) = \left(Hf(\mathbf{x}_0) - \sum_{i=1}^m \lambda_i Hg_i(\mathbf{x}_0) \right)$$

proyectada sobre el espacio tangente a $\{\mathbf{x} : \mathbf{g}(\mathbf{x}) = 0\}$?

(a) Obtenemos la solución en forma vectorial del sistema homogéneo

$$D\mathbf{g}(\mathbf{x}_0) \mathbf{x} = \mathbf{0},$$

de forma que obtenemos una base $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$ del espacio nulo de la matriz $D\mathbf{g}(\mathbf{x}_0)$.

(b) Proyectamos todas las columnas de la matriz

$$L(\mathbf{x}_0, \lambda) = \left(Hf(\mathbf{x}_0) - \sum_{i=1}^m \lambda_i Hg_i(\mathbf{x}_0) \right)$$

sobre el espacio generado por

$$\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$$

esto es construimos

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_k]$$

y calculamos la matriz simétrica

$$\hat{L}(\mathbf{x}_0, \lambda) = (A^T A)^{-1} A^T L(\mathbf{x}_0, \lambda) A.$$

que aparece como solución de las ecuaciones normales del problema de mínimos cuadrados

$$A^T A \hat{\mathbf{x}} = A^T L(\mathbf{x}_0, \lambda) A.$$

(c) Calculamos los valores propios de $\hat{L}(\mathbf{x}_0, \lambda)$, si son todos negativos entonces $\mathbf{x}_0$ es la solución del programa.

■ **Ejemplo:** Del programa

$$\begin{array}{ll} \text{máx} & x + y + z \\ \text{s.a.} & x^2 + y^2 = 2 \\ & x + z = 1 \end{array}$$

conociamos que los posibles óptimos son

$$\mathbf{x}_0^1 = (0, \sqrt{2}, 1) \text{ con } \lambda_1 = \frac{1}{2\sqrt{2}}, \lambda_2 = 1$$

y

$$\mathbf{x}_0^2 = (0, -\sqrt{2}, 1) \text{ con } \lambda_1 = -\frac{1}{2\sqrt{2}}, \lambda_2 = 1.$$

Además,

$$f(x, y, z) = x + y + z$$

$$g_1(x, y, z) = x^2 + y^2 - 2 = 0,$$

$$g_2(x, y, z) = x + z - 1 = 0$$

y

$$\nabla f(x, y, z) = (1, 1, 1),$$

$$\nabla g_1(x, y, z) = (2x, 2y, 0),$$

$$\nabla g_2(x, y, z) = (1, 0, 1).$$

Entonces

$$\begin{aligned} Hf(x, y, z) &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ Hg_1(x, y, z) &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ Hg_2(x, y, z) &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

En esta situación calculamos

$$\begin{aligned} L(x, y, z, \lambda_1, \lambda_2) &= Hf(x, y, z) - \lambda_1 Hg_1(x, y, z) - \lambda_2 Hg_2(x, y, z) \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \lambda_1 \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \lambda_2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -2\lambda_1 & 0 & 0 \\ 0 & -2\lambda_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

y sustituimos

$$\begin{aligned} L\left(0, \sqrt{2}, 1, \frac{1}{2\sqrt{2}}, 1\right) &= \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ L\left(0, -\sqrt{2}, 1, -\frac{1}{2\sqrt{2}}, 1\right) &= \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Ahora calculamos el espacio tangente a las restricciones a partir de

$$\mathbf{g}(x, y, z) = (x^2 + y^2 - 2, x + z - 1)$$

Calculamos

$$D\mathbf{g}(x, y, z) = \begin{bmatrix} 2x & 2y & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

y sustituimos

$$\begin{aligned} D\mathbf{g}(0, \sqrt{2}, 1) &= \begin{bmatrix} 0 & 2\sqrt{2} & 0 \\ 1 & 0 & 1 \end{bmatrix}, \\ D\mathbf{g}(0, -\sqrt{2}, 1) &= \begin{bmatrix} 0 & -2\sqrt{2} & 0 \\ 1 & 0 & 1 \end{bmatrix} \end{aligned}$$

en el primer caso el espacio tangente

$$D\mathbf{g}(0, \sqrt{2}, 1) = \begin{bmatrix} 0 & 2\sqrt{2} & 0 \\ 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

en consecuencia el sistema asociado es

$$\begin{array}{rcl} x & +z & = 0 \\ y & & = 0 \end{array}$$

luego

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

y la base del espacio tangente está únicamente formada por la base del espacio columna de la matriz

$$A = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

En el segundo caso

$$D\mathbf{g}(0, -\sqrt{2}, 1) = \begin{bmatrix} 0 & -2\sqrt{2} & 0 \\ 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

en consecuencia el sistema asociado es

$$\begin{array}{rcl} x & +z & = 0 \\ y & & = 0 \end{array}$$

luego

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

y la base del espacio tangente está únicamente formada por la base del espacio columna de la matriz

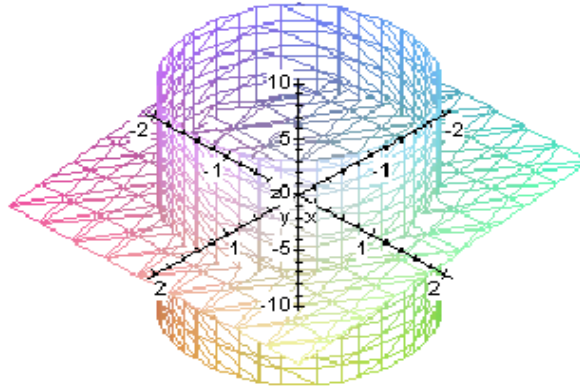
$$A = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Vamos ahora a calcular $\hat{L}(\mathbf{x}_0, \lambda)$, para ello necesitamos

$$\begin{aligned} (A^T A)^{-1} A^T &= \left(\begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \\ &= (2)^{-1} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \end{aligned}$$

Entonces,

$$\begin{aligned} \hat{L}\left(0, \sqrt{2}, 1, \frac{1}{2\sqrt{2}}, 1\right) &= \frac{1}{2} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \\ &= -\frac{1}{4}\sqrt{2} \end{aligned}$$



y el valor propio es

$$-\frac{1}{4}\sqrt{2}$$

negativo, luego se trata de un máximo. La solución del problema es

$$\mathbf{x}_0 = (0, \sqrt{2}, 1), \quad f(0, \sqrt{2}, 1) = 1 + \sqrt{2}.$$

En el segundo caso

$$\begin{aligned} \hat{L}\left(0, -\sqrt{2}, 1, -\frac{1}{2\sqrt{2}}, 1\right) &= \frac{1}{2} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \\ &= \frac{1}{4}\sqrt{2} \end{aligned}$$

el valor propio es

$$\frac{1}{4}\sqrt{2},$$

en este caso se trata de un mínimo. Ahora, bien si analizamos las restricciones

$$\begin{aligned} x^2 + y^2 &= 2 \\ x + z &= 1 \end{aligned}$$

obtenemos que

la intersección es un conjunto cerrado y acotado, en consecuencia

$$(0, -\sqrt{2}, 1), (0, \sqrt{2}, 1)$$

son el máximo absoluto y el mínimo absoluto respectivamente

$$f(0, \sqrt{2}, 1) = 1 + \sqrt{2}, \quad f(0, -\sqrt{2}, 1) = 1 - \sqrt{2}.$$

15.5. Condiciones de Kuhn–Tucker

- Consideremos ahora el problema

$$\begin{array}{ll} \text{máx} & f(\mathbf{x}) \\ \text{s.a.} & g(\mathbf{x}) = 0, \\ & h(\mathbf{x}) \leq 0. \end{array}$$

- Podemos entonces plantear

$$\nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x}) + \mu \nabla h(\mathbf{x}),$$

$$g(\mathbf{x}) = 0,$$

$$\mu h(\mathbf{x}) = 0$$

en donde la tercera ecuación nos dice que o bien se cumple

$$\nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x}),$$

$$g(\mathbf{x}) = 0,$$

$$\mu = 0;$$

o bien se cumple

$$\nabla f(\mathbf{x}) = \lambda \nabla g(\mathbf{x}) + \mu \nabla h(\mathbf{x}),$$

$$g(\mathbf{x}) = 0,$$

$$h(\mathbf{x}) = 0.$$

Teorema 15.10 (Condiciones de Kuhn–Tucker). Sea $\mathbf{x}_0$ un punto solución del problema

$$\begin{array}{ll} \text{máx} & f(\mathbf{x}) \\ \text{s.a.} & g(\mathbf{x}) = 0, \\ & h(\mathbf{x}) \leq 0. \end{array}$$

Supongamos que $\{\nabla g(\mathbf{x}_0), \nabla h(\mathbf{x}_0)\}$ es un conjunto linealmente independiente. Entonces existe un número real λ y un número real $\mu \geq 0$ de forma que

$$\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0) + \mu \nabla h(\mathbf{x}_0),$$

$$g(\mathbf{x}_0) = 0,$$

$$\mu h(\mathbf{x}_0) = 0.$$

Capítulo 16

Integración de funciones de dos o más variables

16.1. Introducción: Principio de Cavalieri

Conocemos que

$$\int_a^b 1 \times dx = (b - a)$$

si consideramos la función

$$f(x, y) = \begin{cases} 1 & \text{si } a \leq x \leq b, c \leq y \leq d \\ 0 & \text{si resto de puntos del plano.} \end{cases}$$

es decir la función toma el valor uno cuando

$$x \in [a, b] \text{ e } y \in [c, d]$$

esto se escribe como

$$(x, y) \in [a, b] \times [c, d]$$

donde

$$R = [a, b] \times [c, d]$$

es un rectángulo de base $[a, b]$ y de altura $[c, d]$. Entonces la gráfica de la función es si tomamos $R = [0, 1] \times [0, 1]$,

entonces el volumen contenido debajo de la gráfica es igual a

$$1 \times [(b - a) \times (d - c)] = \text{Altura} \times \text{Área de la base.}$$

Definición (Integral doble): Se define la integral doble de una función

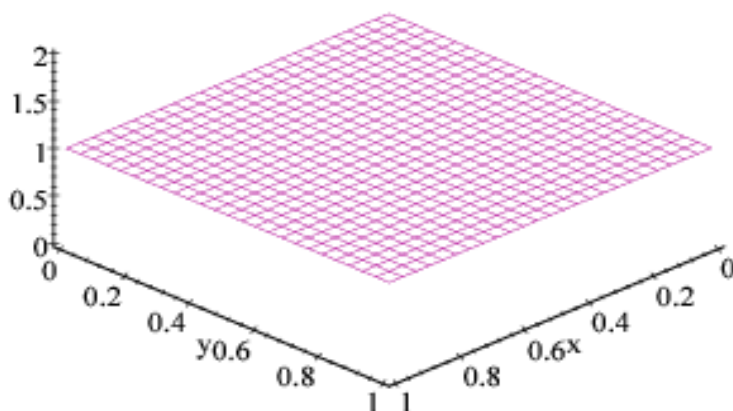
$$f : R = [a, b] \times [c, d] \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$$

y que toma valores no negativos

$$f(x, y) \geq 0,$$

como el volumen de la región

$$\Omega = \{(x, y, z) : (x, y) \in R, 0 \leq z \leq f(x, y)\} \subset \mathbb{R}^3,$$



y se denota este volumen por

$$\int \int_R f(x, y) dA = \int \int_R f(x, y) dx dy$$

donde

$$dA = dx dy$$

es la llamada “diferencial de área”. En este caso

$$dx dy \neq dy dx.$$

Ejemplo 16.1. Entonces

$$1 \times [(b - a) \times (d - c)] = \int \int_R 1 \times dA = \int \int_R 1 \times dx dy$$

Ejemplo 16.2. Si

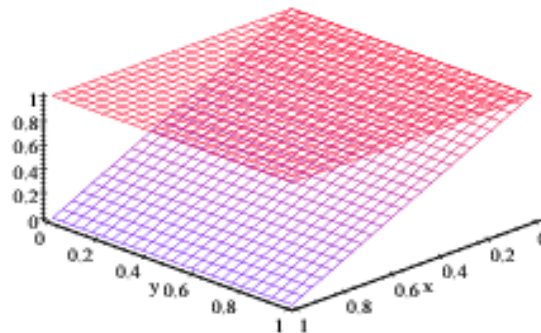
$$f : R = [a, b] \times [c, d] \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$$

se define como

$$f(x, y) = k > 0$$

entonces

$$\begin{aligned} \int \int_R f(x, y) dA &= \int \int_R f(x, y) dx dy \\ &= \int \int_R k dx dy \\ &= k \times \int \int_R dx dy \\ &= k \times [(b - a) \times (d - c)]. \end{aligned}$$



Ejemplo 16.3. Consideremos

$$f : R = [0, 1] \times [0, 1] \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$$

se define como

$$f(x, y) = 1 - x$$

entonces hay que calcular ese volumen, que es igual a la mitad del área de la base por la altura máxima que alcanza la figura, que en este caso es uno, esto es,

$$\int \int_R (1 - x) dA = \int \int_R (1 - x) dx dy = \frac{1}{2}$$

Ejemplo 16.4. Cómo sería el volumen definido por

$$\int \int_{[-1,1] \times [0,1]} (x^2 + y^2) dA?$$

En este caso

Para calcular este tipo de integrales la idea que da el llamado “Principio de Cavalieri” es la siguiente tomemos “cortemos” la figura anterior por el plano $y = 0.5$, esto es,

entonces la sección transversal tiene un área que depende de $y = 0.5$, que podemos calcular como

$$\int_{-1}^1 (x^2 + (0.5)^2) dx = 1.16666667$$

en general y para cada y podríamos calcular

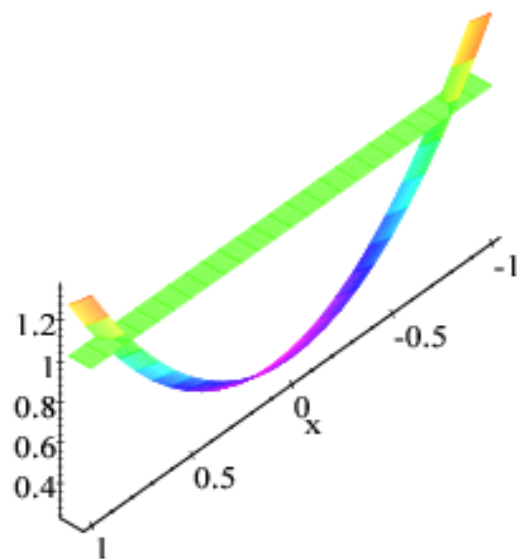
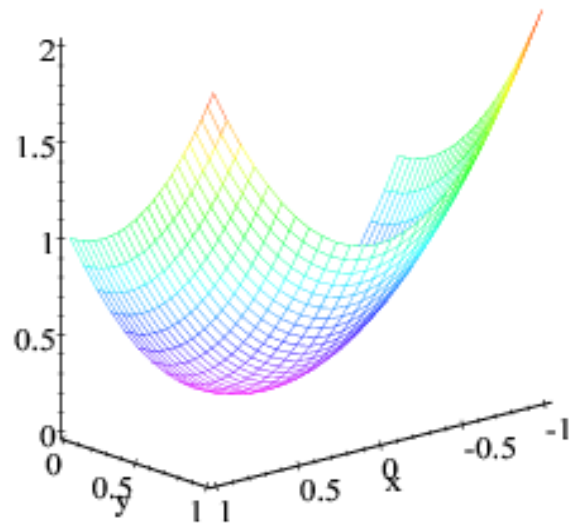
$$\int_{-1}^1 (x^2 + y^2) dx = \left[\frac{x^3}{3} + xy^2 \right]_{-1}^1 = \frac{2}{3} + 2y^2$$

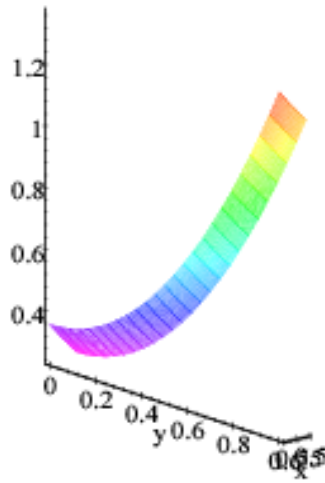
y ahora “sumamos” todas estas áreas para los valores de y en el intervalo $[0, 1]$, esto es,

$$\int_0^1 \left(\frac{2}{3} + 2y^2 \right) dy = \left[\frac{2}{3}y + \frac{2}{3}y^3 \right]_0^1 = \frac{4}{3}.$$

También se habrían podido emplear planos del tipo $x = cte$, por ejemplo $x = 0.5$, entonces habríamos calculado en primer lugar

$$\int_0^1 (x^2 + y^2) dy = \left[x^2y + \frac{y^3}{3} \right]_0^1 = x^2 + \frac{1}{3}$$





y en segundo lugar

$$\int_{-1}^1 \left(x^2 + \frac{1}{3} \right) dx = \left[\frac{x^3}{3} + \frac{1}{3}x \right]_{-1}^1 = \frac{4}{3}$$

Empleando, pues, el “Principio de Cavalieri” calculamos en primer lugar el área de la sección transversal

$$A(x) = \int_c^d f(x, y) dy$$

y en segundo lugar “sumamos” las áreas de todas las secciones transversales

$$\int_a^b A(x) dx = \int_a^b \left[\int_c^d f(x, y) dy \right] dx$$

que por abuso de escritura se suele escribir como

$$\int_a^b A(x) dx = \int_a^b \int_c^d f(x, y) dy dx.$$

Del mismo modo se puede calcular en primer lugar

$$A(y) = \int_a^b f(x, y) dx$$

y en segundo lugar

$$\int_c^d A(y) dy = \int_c^d \left[\int_a^b f(x, y) dx \right] dy$$

que por abuso de escritura se suele escribir como

$$\int_c^d A(y) dy = \int_c^d \int_a^b f(x, y) dx dy.$$

Se define

$$\iint_R f(x, y) dA = \int_a^b \left[\int_c^d f(x, y) dy \right] dx$$

o bien

$$\int \int_R f(x, y) dA = \int_c^d \left[\int_a^b f(x, y) dx \right] dy.$$

Ejemplo 16.5. Calcular

$$\int \int_{[0, \pi/2] \times [0, \pi/2]} (\sin x \cos y) dA$$

En este caso

$$\int \int_{[0, \pi/2] \times [0, \pi/2]} (\sin x \cos y) dA = \int_0^{\pi/2} \left[\int_0^{\pi/2} (\sin x \cos y) dy \right] dx.$$

Calculamos en primer lugar

$$\begin{aligned} \int_0^{\pi/2} (\sin x \cos y) dy &= \sin x \int_0^{\pi/2} \cos y dy \\ &= \sin x [\sin y]_0^{\pi/2} \\ &= \sin x. \end{aligned}$$

En segundo lugar

$$\int_0^{\pi/2} \sin x dx = [-\cos x]_0^{\pi/2} = 1.$$

Finalmente

$$\int \int_{[0, \pi/2] \times [0, \pi/2]} (\sin x \cos y) dA = 1.$$

16.2. La integral doble sobre un rectángulo

Al igual que en el caso de funciones de una variable podemos definir las funciones indicador de dos variables

$$I_{[a_1, b_1] \times [c_1, d_1]} : R = [a, b] \times [c, d] \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$$

como

$$I_{[a_1, b_1] \times [c_1, d_1]}(x, y) = \begin{cases} 1 & \text{si } [a_1, b_1] \times [c_1, d_1] \\ 0 & \text{si resto de puntos} \end{cases}$$

donde

$$a \leq a_1 \leq b_1 \leq b; c \leq c_1 \leq d_1 \leq d.$$

Entonces

$$\begin{aligned} \int \int_{[a, b] \times [c, d]} I_{[a_1, b_1] \times [c_1, d_1]}(x, y) dA &= \int \int_{[a_1, b_1] \times [c_1, d_1]} dA \\ &= (b_1 - a_1) \times (d_1 - c_1). \end{aligned}$$

Podemos entonces tomar una “cuadrícula” del rectángulo R como

$$R = ([a_1, b_1] \times [c_1, d_1]) \cup ([a_2, b_2] \times [c_2, d_2]) \cup \cdots \cup [a_n, b_n] \times [c_n, d_n]$$

y definir funciones escalonadas como

$$f(x, y) = \sum_{i=1}^n c_i I_{[a_i, b_i] \times [c_i, d_i]}(x, y)$$

donde $c_1, c_2, \dots, c_n$ son números reales no negativos. Se define entonces

$$\begin{aligned} \int \int_R f(x, y) dA &= \int \int_R \left(\sum_{i=1}^n c_i I_{[a_i, b_i] \times [c_i, d_i]}(x, y) \right) dA \\ &= \sum_{i=1}^n c_i \left(\int \int_R I_{[a_i, b_i] \times [c_i, d_i]}(x, y) dA \right) \\ &= \sum_{i=1}^n c_i [(b_i - a_i) \times (d_i - c_i)]. \end{aligned}$$

En general para funciones $f(x, y)$ continuas, se define mediante un paso al límite

$$\begin{aligned} &\int \int_R f(x, y) dA \\ &= \lim_{\Delta x_i \rightarrow 0, \Delta y_i \rightarrow 0} \sum f\left(\frac{x_i + x_{i+1}}{2}, \frac{y_i + y_{i+1}}{2}\right) [\Delta x_i \times \Delta y_i]. \end{aligned}$$

Teorema 16.1 (Integrabilidad de funciones acotadas). Sea $f : R \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ una función acotada y no negativa, esto es, existe un número $M < \infty$ de forma que

$$f(x, y) \leq M$$

para todo $(x, y) \in R$. Si el conjunto de puntos donde f tiene discontinuidades está formado por una unión finita de gráficas de funciones continuas, entonces f es integrable sobre R , esto es, podemos calcular

$$\int \int_R f(x, y) dA$$

mediante el paso al límite anterior.

16.2.1. Propiedades de la integral doble sobre un rectángulo

1. Linealidad

$$\begin{aligned} &\int \int_R (f(x, y) + g(x, y)) dA \\ &= \int \int_R f(x, y) dA + \int \int_R g(x, y) dA. \end{aligned}$$

2. Homogeneidad

$$\int \int_R (\lambda f(x, y)) dA = \lambda \left(\int \int_R f(x, y) dA \right).$$

3. Monotonía: Si $f(x, y) \geq g(x, y)$, entonces

$$\int \int_R f(x, y) dA \geq \int \int_R g(x, y) dA.$$

4. Aditividad: Si $R = R_1 \cup R_2$ donde R_1 y R_2 son dos rectángulos de forma que solo coinciden en la frontera, entonces

$$\int \int_{R_1 \cup R_2} f(x, y) dA = \int \int_{R_1} f(x, y) dA + \int \int_{R_2} f(x, y) dA.$$

16.2.2. El Teorema de Fubini

Bajo que circunstancias podemos asegurar que

$$\int_a^b \left[\int_c^d f(x, y) dy \right] dx = \int_c^d \left[\int_a^b f(x, y) dx \right] dy?$$

Teorema 16.2 (Teorema de Fubini). Sea f una función continua con dominio rectangular $R = [a, b] \times [c, d]$. Entonces

$$\begin{aligned} & \int \int_R f(x, y) dA \\ &= \int_a^b \left[\int_c^d f(x, y) dy \right] dx = \int_c^d \left[\int_a^b f(x, y) dx \right] dy. \end{aligned}$$

Teorema 16.3. Sea $f : R \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ una función acotada y no negativa, esto es, existe un número $M < \infty$ de forma que

$$f(x, y) \leq M$$

para todo $(x, y) \in R$. Supongamos que el conjunto de puntos donde f tiene discontinuidades está formado por una unión finita de gráficas de funciones continuas. Si la integral $\int_c^d f(x, y) dy$ existe para cada $x \in [a, b]$, entonces

$$\int \int_R f(x, y) dA = \int_a^b \left[\int_c^d f(x, y) dy \right] dx.$$

Del mismo modo si existe la integral $\int_a^b f(x, y) dx$ para cada $y \in [c, d]$, entonces

$$\int \int_R f(x, y) dA = \int_c^d \left[\int_a^b f(x, y) dx \right] dy.$$

Si todas estas condiciones se cumplen simultáneamente, entonces

$$\begin{aligned} & \int \int_R f(x, y) dA \\ &= \int_a^b \left[\int_c^d f(x, y) dy \right] dx = \int_c^d \left[\int_a^b f(x, y) dx \right] dy. \end{aligned}$$

Escribiremos abusando de la notación

$$\int_a^b \left[\int_c^d f(x, y) dy \right] dx = \int_a^b \int_c^d f(x, y) dy dx$$

y

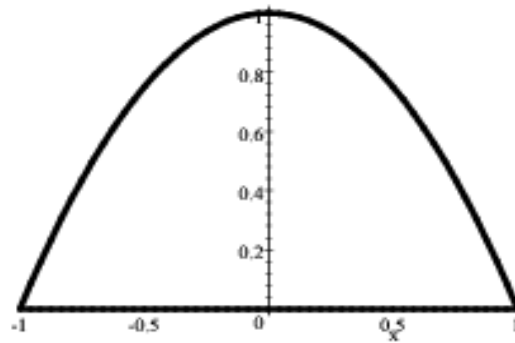
$$\int_c^d \left[\int_a^b f(x, y) dx \right] dy = \int_c^d \int_a^b f(x, y) dx dy.$$

Ejemplo 16.6. Calcular

$$\int \int_{[0,1] \times [0,1]} (x^2 + y) dA.$$

Empleando el Teorema de Fubini

$$\int \int_{[0,1] \times [0,1]} (x^2 + y) dA = \int_0^1 \int_0^1 (x^2 + y) dx dy.$$



Calcularemos en primer lugar

$$\int_0^1 (x^2 + y) dx = \left[\frac{x^3}{3} + xy \right]_0^1 = \frac{1}{3} + y.$$

Ahora

$$\int_0^1 \left(\frac{1}{3} + y \right) dy = \left[\frac{1}{3}y + \frac{y^2}{2} \right]_0^1 = \frac{5}{6}.$$

16.3. La integral doble sobre regiones más generales

Consideremos dos funciones

$$\phi_1, \phi_2 : [a, b] \longrightarrow \mathbb{R}$$

de forma que

$$\phi_1(x) \leq \phi_2(x)$$

para todo $x \in [a, b]$. Por ejemplo:

entonces estas funciones definen una región del plano

$$D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x)\}.$$

A este tipo de región del plano la llamaremos **región de tipo I**. Por ejemplo

$$D = \{(x, y) \in \mathbb{R}^2 : -1 \leq x \leq 1, 0 \leq y \leq -x^2 + 1\}.$$

Consideremos dos funciones

$$\psi_1, \psi_2 : [c, d] \longrightarrow \mathbb{R}$$

de forma que

$$\psi_1(y) \leq \psi_2(y)$$

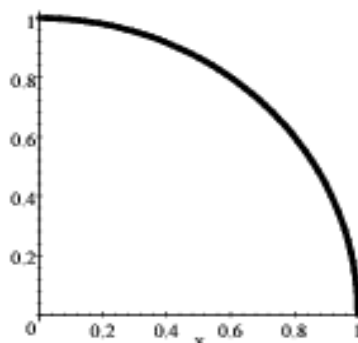
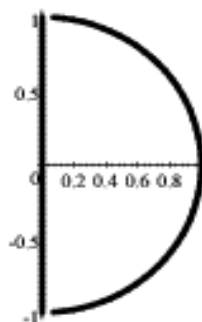
para todo $y \in [c, d]$. Por ejemplo:

entonces estas funciones definen una región del plano

$$D = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\}.$$

A este tipo de región del plano la llamaremos **región de tipo II**. Por ejemplo

$$D = \{(x, y) \in \mathbb{R}^2 : -1 \leq y \leq 1, 0 \leq x \leq \sqrt{1 - y^2}\}.$$



Finalmente llamaremos **región de tipo III** aquellas regiones que son a la vez de tipo I y de tipo II. Por ejemplo:

$$D = \left\{ (x, y) \in \mathbb{R}^2 : -1 \leq y \leq 1, 0 \leq x \leq \sqrt{1 - y^2} \right\}$$

también se puede escribir como

$$D = \left\{ (x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, -\sqrt{1 - x^2} \leq y \leq \sqrt{1 - x^2} \right\}$$

en este último caso

Teorema 16.4 (Integrales iteradas para regiones de tipo I). Sea $f : D \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ una función continua siendo D una región de tipo I, esto es,

$$D = \left\{ (x, y) \in \mathbb{R}^2 : a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x) \right\}.$$

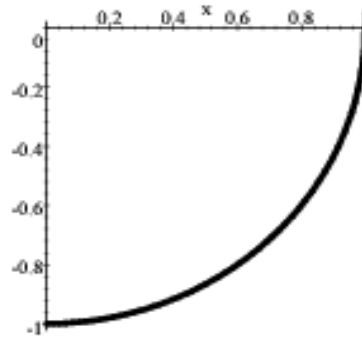
Entonces

$$\iint_D f(x, y) dA = \int_a^b \left[\int_{\phi_1(x)}^{\phi_2(x)} f(x, y) dy \right] dx.$$

Ejemplo 16.7. Hallar

$$\iint_T (x^3 y + \cos y) dA$$

siendo T el triángulo que consta de los puntos (x, y) tales que $0 \leq x \leq \pi/2$ y $0 \leq y \leq x$.



Como se trata de una región de tipo I, con

$$\phi_1(x) = 0 \leq y \leq x = \phi_2(x)$$

el anterior Teorema nos dice que

$$\iint_T (x^3 y + \cos y) dA = \int_0^{\pi/2} \left[\int_0^x (x^3 y + \cos y) dy \right] dx.$$

Calculamos en primer lugar

$$\begin{aligned} \int_0^x (x^3 y + \cos y) dy &= \left[x^3 \frac{y^2}{2} + \sin y \right]_0^x \\ &= \left(x^3 \frac{x^2}{2} + \sin x \right) - \left(x^3 \frac{0^2}{2} + \sin 0 \right) \\ &= \frac{x^5}{2} + \sin x. \end{aligned}$$

En segundo lugar

$$\int_0^{\pi/2} \left(\frac{x^5}{2} + \sin x \right) dx = \left[\frac{x^6}{12} - \cos x \right]_0^{\pi/2} = \frac{1}{768} \pi^6 + 1.$$

Luego

$$\iint_T (x^3 y + \cos y) dA = \frac{1}{768} \pi^6 + 1.$$

Teorema 16.5 (Integrales iteradas para regiones de tipo II). Sea $f : D \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ una función continua siendo D una región de tipo II, esto es,

$$D = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\}.$$

Entonces

$$\iint_D f(x, y) dA = \int_c^d \left[\int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx \right] dy.$$

En el caso de que $f(x, y) = 1$ para todo $(x, y) \in D$, entonces

$$\iint_D dA = A(D),$$

es el área de la región D . Si D es de tipo I se tiene que

$$\int \int_D dA = \int_a^b \left[\int_{\phi_1(x)}^{\phi_2(x)} dy \right] dx = \int_a^b (\phi_2(x) - \phi_1(x)) dx.$$

En el caso de que la región D fuese de tipo II se tiene que

$$\int \int_D dA = \int_c^d \left[\int_{\psi_1(y)}^{\psi_2(y)} dx \right] dy = \int_c^d (\psi_2(y) - \psi_1(y)) dy.$$

16.3.1. Cambio en el orden de integración

Supongamos que D es una región de tipo III. Entonces

$$D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x)\}.$$

y

$$D = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\}$$

Por lo tanto se tienen las fórmulas

$$\int \int_D f(x, y) dA = \int_a^b \left[\int_{\phi_1(x)}^{\phi_2(x)} f(x, y) dy \right] dx$$

y

$$\int \int_D f(x, y) dA = \int_c^d \left[\int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx \right] dy.$$

En consecuencia:

$$\begin{aligned} \int \int_D f(x, y) dA &= \int_a^b \left[\int_{\phi_1(x)}^{\phi_2(x)} f(x, y) dy \right] dx \\ &= \int_c^d \left[\int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx \right] dy. \end{aligned}$$

Ejemplo 16.8. Evaluar

$$\int_0^a \int_0^{(a^2-x^2)^{1/2}} (a^2-y^2) dy dx$$

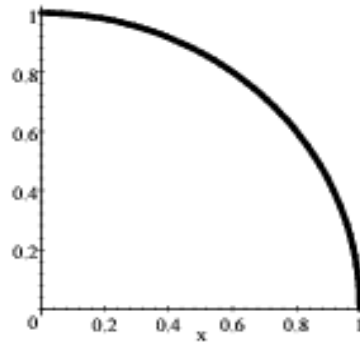
cambiando el orden de integración.

Conocemos que

$$\begin{aligned} &\int_0^a \int_0^{(a^2-x^2)^{1/2}} (a^2-y^2) dy dx \\ &= \int_0^a \left[\int_0^{(a^2-x^2)^{1/2}} (a^2-y^2)^{1/2} dy \right] dx \end{aligned}$$

entonces la región de integración D que es de tipo I se escribe como

$$D = \{(x, y) : 0 \leq x \leq a, 0 \leq y \leq \sqrt{a^2-x^2}\}.$$



En este caso

$$y = \sqrt{a^2 - x^2},$$

o bien

$$y^2 = a^2 - x^2 \implies x^2 = a^2 - y^2,$$

que representa el cuadrante positivo de una circunferencia de radio a centrada en el origen. Gráficamente entonces

$$0 \leq y \leq 1$$

e

$$0 \leq x \leq \sqrt{a^2 - y^2}.$$

Luego D como región de tipo II se escribe como

$$D = \left\{ (x, y) : 0 \leq y \leq a, 0 \leq x \leq \sqrt{a^2 - y^2} \right\}$$

entonces

$$\begin{aligned} & \int_0^a \int_0^{(a^2 - y^2)^{1/2}} (a^2 - y^2)^{1/2} dy dx \\ &= \int \int_D (a^2 - y^2)^{1/2} dA \\ &= \int_0^a \left[\int_0^{\sqrt{a^2 - y^2}} (a^2 - y^2)^{1/2} dx \right] dy, \end{aligned}$$

en este caso

$$\begin{aligned} \int_0^{\sqrt{a^2 - y^2}} (a^2 - y^2) dx &= (a^2 - y^2)^{1/2} \int_0^{\sqrt{a^2 - y^2}} dx \\ &= (a^2 - y^2)^{1/2} [x]_0^{(a^2 - y^2)^{1/2}} \\ &= (a^2 - y^2)^{1/2} (a^2 - y^2)^{1/2} \\ &= (a^2 - y^2). \end{aligned}$$

Finalmente,

$$\int_0^a (a^2 - y^2) dy = \left[a^2 y - \frac{y^3}{3} \right]_0^a = \frac{2a^3}{3}.$$

16.4. El Teorema del cambio de variables

El problema que nos planteamos es el siguiente. Usualmente calculamos un volumen empleando la “diferencial de área”

$$dA = dx dy.$$

Supongamos que cambiamos las variables (x, y) empleando

$$x = x(u, v)$$

$$y = y(u, v)$$

por las variables (u, v) . ¿Qué relación existe entre

$$dA = dx dy$$

y

$$d\hat{A} = du dv?$$

Nótese que podemos definir

$$\mathbf{T} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

como

$$\mathbf{T}(u, v) = (x(u, v), y(u, v)) = (x, y).$$

Ejemplo 16.9. Consideremos el cambio a coordenadas polares

$$x = r \cos \theta = x(r, \theta)$$

$$y = r \sin \theta = y(r, \theta),$$

esto define una transformación

$$\mathbf{T}(r, \theta) = (r \cos \theta, r \sin \theta) = (x, y)$$

Entonces si $(r, \theta) \in [0, 1] \times [0, 2\pi]$, tenemos que

$$x^2 + y^2 = r^2 \leq 1$$

es decir (x, y) pertenecen al disco de radio 1 centrado en el origen

Entonces podemos considerar que

$$\mathbf{T} : R \longrightarrow D$$

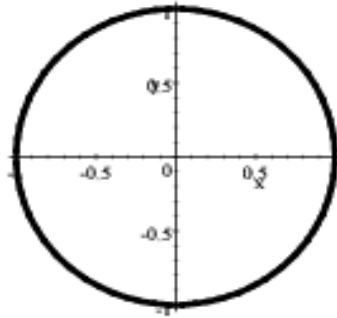
donde

$$R = [0, 1] \times [0, 2\pi], D = \{(x, y) : x^2 + y^2 \leq 1\}.$$

Además como

$$x = r \cos \theta = x(r, \theta)$$

$$y = r \sin \theta = y(r, \theta),$$



empleando la regla de la cadena obtenemos

$$\begin{aligned} dx &= \frac{\partial x}{\partial r} dr + \frac{\partial x}{\partial \theta} d\theta = \cos \theta dr - r \sin \theta d\theta, \\ dy &= \frac{\partial y}{\partial r} dr + \frac{\partial y}{\partial \theta} d\theta = \sin \theta dr + r \cos \theta d\theta \end{aligned}$$

en forma matricial

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} \begin{bmatrix} dr \\ d\theta \end{bmatrix}$$

o bien

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix} \begin{bmatrix} dr \\ d\theta \end{bmatrix}.$$

Nótese que

$$\det \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix} = r \cos^2 \theta + r \sin^2 \theta = r (\cos^2 \theta + \sin^2 \theta) = r,$$

y que

$$\begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix}^{-1} = \frac{1}{r} \begin{bmatrix} r \cos \theta & r \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

siempre que $r > 0$, luego

$$\frac{1}{r} \begin{bmatrix} r \cos \theta & r \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} dr \\ d\theta \end{bmatrix}.$$

Ejemplo 16.10. Consideremos

$$\mathbf{T}(x, y) = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ -2x + y \end{bmatrix} = (x + 2y, -2x + y)$$

esto es,

$$u = x + 2y = u(x, y),$$

$$v = -2x + y = v(x, y).$$

Además, como

$$\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}$$

y la matriz es invertible

$$\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix}$$

entonces

$$\begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix},$$

esto es,

$$\begin{aligned} x &= \frac{1}{5}u - \frac{2}{5}v = x(u, v), \\ y &= \frac{2}{5}u + \frac{1}{5}v = y(u, v). \end{aligned}$$

Diferenciando obtenemos en el primer caso

$$\begin{aligned} du &= \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy = 1dx + 2dy, \\ dv &= \frac{\partial v}{\partial x}dx + \frac{\partial v}{\partial y}dy = -2dx + 1dy \end{aligned}$$

y en el segundo

$$\begin{aligned} dx &= \frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv = \frac{1}{5}du - \frac{2}{5}dv, \\ dy &= \frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv = \frac{2}{5}du + \frac{1}{5}dv. \end{aligned}$$

En forma matricial

$$\begin{bmatrix} du \\ dv \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix}$$

y

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix}$$

Supongamos que

$$x = x(u, v),$$

$$y = y(u, v).$$

que define una transformación

$$\mathbf{T} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

como

$$\mathbf{T}(u, v) = (x(u, v), y(u, v)) = (x, y).$$

Entonces llamamos **determinante jacobiano** de la transformación a

$$\frac{\partial(x, y)}{\partial(u, v)} = \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}.$$

Además si

$$dA = dx dy$$

y

$$d\hat{A} = du dv$$

se tiene que

$$dA = dxdy = \left| \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} \right| dudv = \left| \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} \right| d\hat{A}.$$

Ejemplo 16.11. Sea

$$x = r \cos \theta = x(r, \theta)$$

$$y = r \sin \theta = y(r, \theta),$$

entonces

$$dxdy = \left| \det \begin{bmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{bmatrix} \right| drd\theta = |r| drd\theta.$$

Ejemplo 16.12. Sea

$$u = x + 2y = u(x, y),$$

$$v = -2x + y = v(x, y).$$

entonces

$$dudv = \left| \det \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \right| dxdy = |5| dxdy = 5dxdy.$$

Ejemplo 16.13. Sea

$$x = \frac{1}{5}u - \frac{2}{5}v = x(u, v),$$

$$y = \frac{2}{5}u + \frac{1}{5}v = y(u, v).$$

entonces

$$dxdy = \left| \det \begin{bmatrix} \frac{1}{5} & -\frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{bmatrix} \right| dudv = \left| \frac{1}{5} \right| dudv = \frac{1}{5}dudv.$$

Teorema 16.6 (Teorema del cambio de variables). Sean D y D^* dos regiones elementales del plano y sea $\mathbf{T} : D^* \longrightarrow D$ de clase $\mathcal{C}^1$, de forma que $\mathbf{T}(D^*) = D$ y que

$$\mathbf{T}(u, v) = (x(u, v), y(u, v)) = (x, y)$$

cumple que

$$\frac{\partial(x, y)}{\partial(u, v)} = \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} \neq 0$$

para todo $(u, v) \in D^*$. Entonces

$$\begin{aligned} & \int \int_D f(x, y) dxdy \\ &= \int \int_{D^*} f(x(u, v), y(u, v)) \left| \det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} \right| dudv. \end{aligned}$$

El problema consiste en encontrar los extremos de integración, esto es, si

$$\mathbf{T}(r, \theta) = (x(r, \theta), y(r, \theta)) = (r \cos \theta, r \sin \theta) = (x, y)$$

y

$$D^* = [0, 1] \times [0, 2\pi]$$

entonces

$$D = ?$$

Nótese que

$$\begin{aligned} & \int \int_{D^*} f(x(r, \theta), y(r, \theta)) |r| dr d\theta \\ &= \int_0^1 \left[\int_0^{2\pi} f(x(r, \theta), y(r, \theta)) |r| d\theta \right] dr \end{aligned}$$

se puede calcular facilmente, sin embargo para

$$\int \int_D f(x, y) dx dy = ?$$

O bien el problema inverso: ¿qué transformación tenemos que emplear sobre D de forma que D^* , sea una región elemental de fácil integración.

Ejemplo 16.14 (Cálculo del área de una circunferencia de radio unidad). Consideremos la circunferencia de radio unidad

$$D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$$

entonces el área se calcula como

$$A(D) = \int \int_D dx dy.$$

Si empleamos el Teorema del cambio de variable asumiendo

$$x = r \cos \theta,$$

$$y = r \sin \theta$$

y

$$f(x, y) = 1,$$

siendo el radio

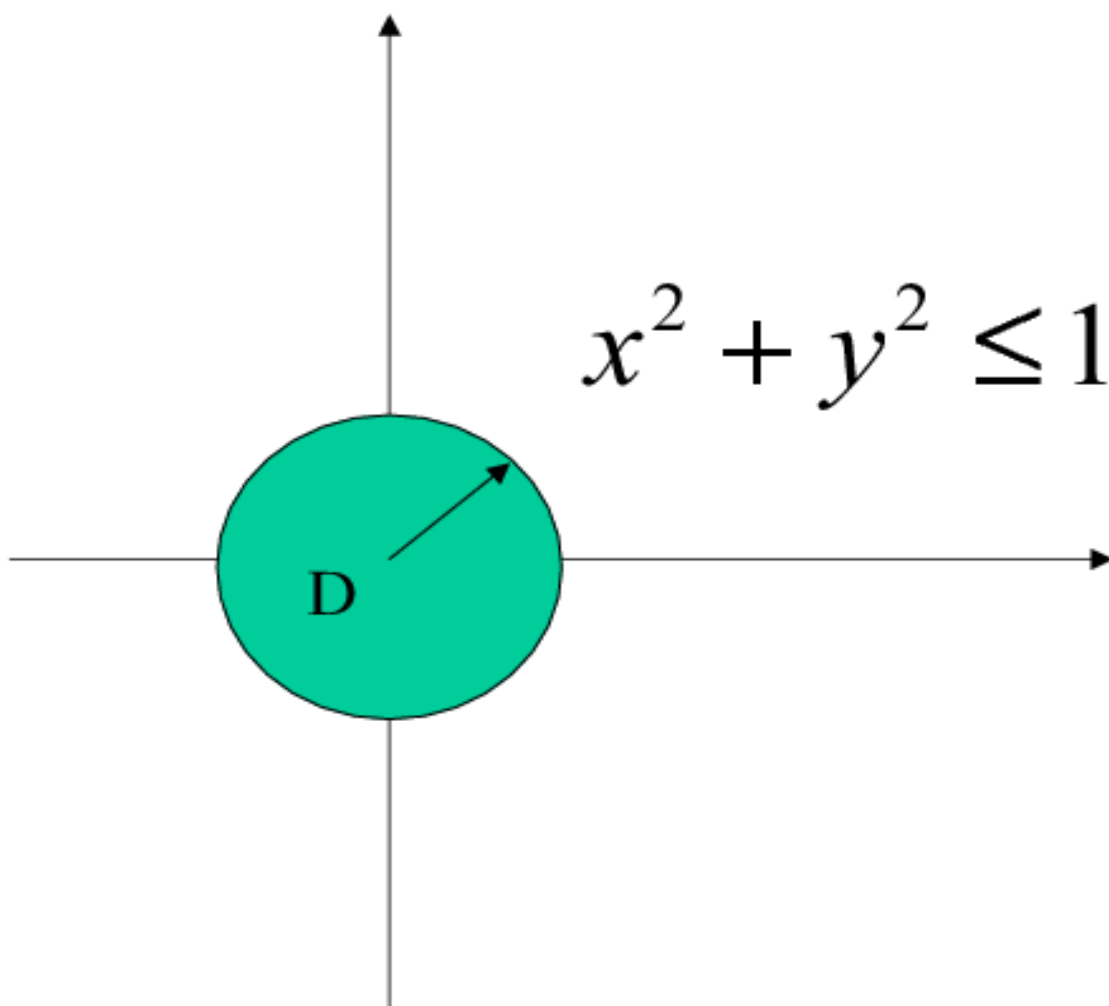
$$0 \leq r \leq 1$$

y el ángulo

$$0 \leq \theta \leq 2\pi$$

obtenemos que

$$A(D) = \int \int_D dx dy = \int \int_{[0,1] \times [0,2\pi]} |r| dr d\theta$$



en este caso

$$\begin{aligned}
 \int \int_{[0,1] \times [0,2\pi]} |r| dr d\theta &= \int_0^1 \left[\int_0^{2\pi} r d\theta \right] dr \\
 &= \int_0^1 [r\theta]_0^{2\pi} dr \\
 &= \int_0^1 [2\pi r - 0] dr \\
 &= 2\pi \left[\frac{r^2}{2} \right]_0^1 \\
 &= \pi.
 \end{aligned}$$

16.4.1. Geometría de las funciones de $\mathbb{R}^2$ a $\mathbb{R}^2$

Consideremos una matriz A 2×2 invertible y sea A^{-1} su matriz inversa. Supongamos que tenemos una transformación

$$\mathbf{T}(x, y) = (u, v)$$

definida a partir de

$$\begin{bmatrix} u \\ v \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix}$$

además podemos definir

$$\mathbf{T}^{-1}(u, v) = (x, y)$$

a partir de

$$\begin{bmatrix} x \\ y \end{bmatrix} = A^{-1} \begin{bmatrix} u \\ v \end{bmatrix}.$$

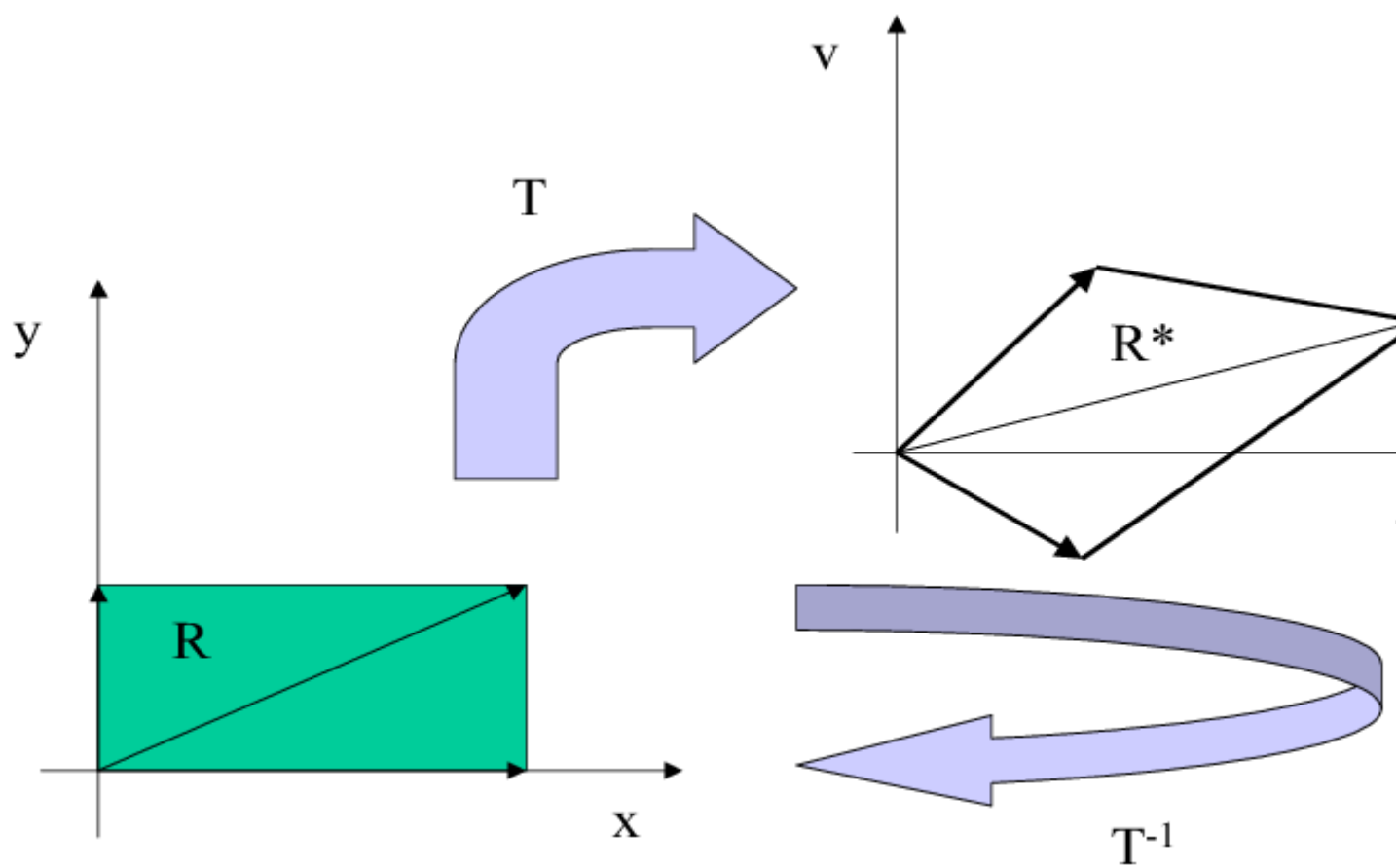
Si tenemos un rectángulo $R = [a, b] \times [c, d]$ en coordenadas (x, y) ¿qué tipo de figura geométrica será $\mathbf{T}(R) = R^*$ en coordenadas (u, v) ? Del mismo modo si conocemos que tipo de figura geométrica es R^* en coordenadas (u, v) entonces conocemos que $\mathbf{T}^{-1}(R^*) = R$ es un rectángulo en coordenadas (x, y) .

Sea $R = [-1, 2] \times [-3, 4]$ un rectángulo que contiene al origen. Sea

$$A = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

una matriz invertible. Entonces D puede describirse empleando cuatro vectores

$$\begin{bmatrix} -1 \\ -3 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$



que determinan los cuatro vértices del rectángulo R . Entonces

$$\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -3 \end{bmatrix} = \begin{bmatrix} -7 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} -4 \\ -7 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$$

Se puede comprobar que estos nuevos cuatro puntos

$$\begin{bmatrix} -7 \\ -1 \end{bmatrix}, \begin{bmatrix} 7 \\ 6 \end{bmatrix}, \begin{bmatrix} -4 \\ -7 \end{bmatrix}, \begin{bmatrix} 10 \\ 0 \end{bmatrix}$$

definen un paralelogramo que llamaremos R^* . Entonces la transformación $\mathbf{T} : R \longrightarrow R^*$ definida por

$$\mathbf{T} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ -2x + y \end{bmatrix}$$

transforma la región R en la región R^* punto a punto, esto es,

$$\mathbf{T}(R) = R^*.$$

Teorema 16.7. Sea A una matriz 2×2 invertible y sea $\mathbf{T}$ la transformación definida por

$$\mathbf{T}(\mathbf{x}) = A\mathbf{x}.$$

Entonces $\mathbf{T}$ transforma paralelogramos en paralelogramos y vértices en vértices. Además, si $T(R^*)$ es un paralelogramo, entonces R^* tiene que ser un paralelogramo.

Ejemplo 16.15. Sea P el paralelogramo acotado por las rectas $y = 2x, y = 2x - 2, y = x$ y $y = x + 1$. Evaluar entonces la integral doble

$$\int \int_P xy dx dy$$

haciendo el cambio de variable

$$x = u - v,$$

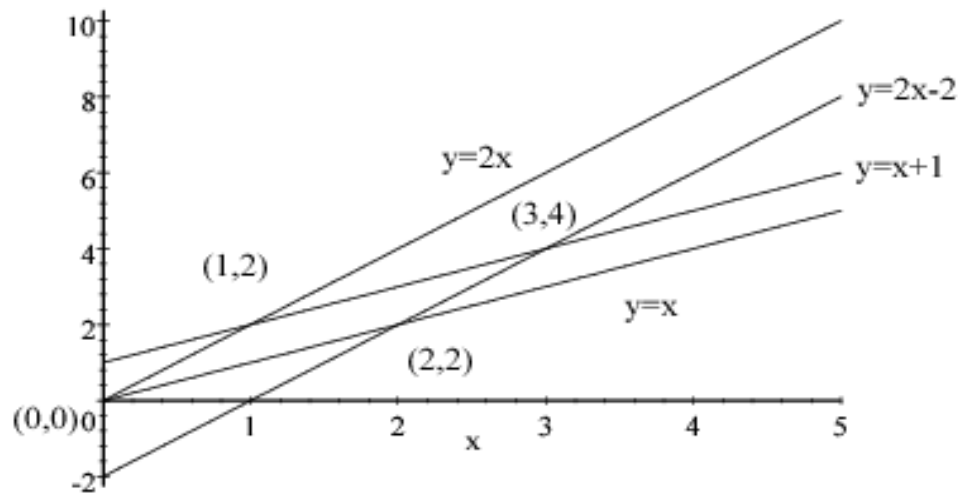
$$y = 2u - v,$$

esto es,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

Como la imagen de un paralelogramo es un paralelogramo y conocemos los vértices en coordenadas (x, y) , vamos a calcular los vértices en coordenadas (u, v) . Como

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$



entonces

$$\begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix},$$

como

$$\begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 1 \\ -2 & 1 \end{bmatrix}$$

tenemos

$$\begin{bmatrix} -1 & 1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}$$

Entonces como los vértices son $(1,2)$, $(2,2)$, $(3,4)$ y $(0,0)$ obtenemos

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

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

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

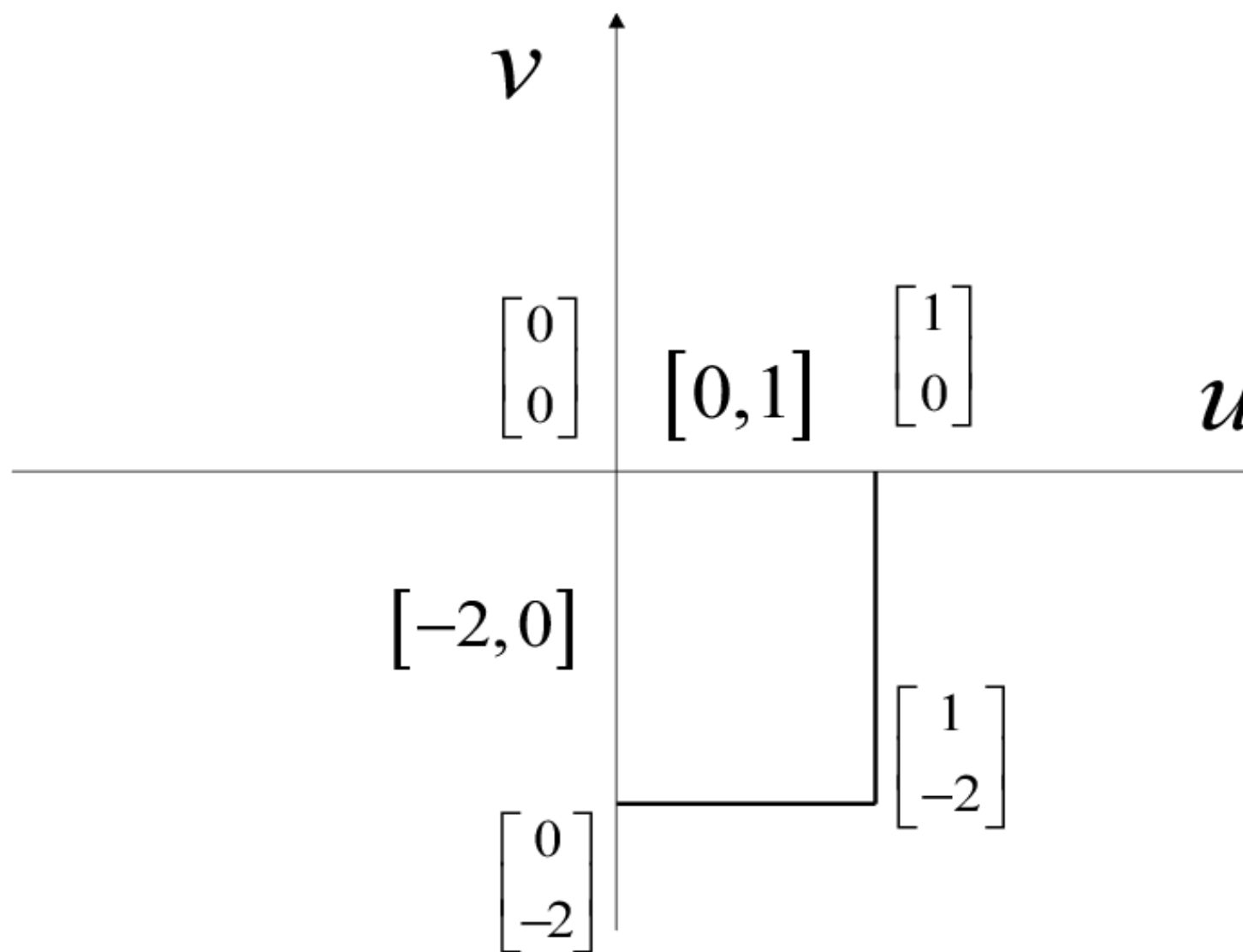
y

$$\begin{bmatrix} -1 & 1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Luego los vértices en coordenadas (u,v) son

$$(1,0), (0,-2), (1,-2), (0,0)$$

gráficamente



este rectángulo lo podemos escribir como

$$[0, 1] \times [-2, 0]$$

Además como

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

entonces

$$x = x(u, v) = u - v,$$

$$y = y(u, v) = 2u - v$$

de donde

$$dx = 1du - 1dv,$$

$$dy = 2du - 1dv$$

es decir

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix}$$

luego

$$dxdy = \left| \det \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} \right| dudv$$

la región en coordenadas (x, y) es el paralelogramo P y en coordenadas (u, v) el rectángulo $[1, 0] \times [-2, 0]$, entonces

$$\begin{aligned} & \int \int_P xy dxdy \\ &= \int \int_{[1,0] \times [-2,0]} x(u, v) y(u, v) \left| \det \begin{bmatrix} 1 & -1 \\ 2 & -1 \end{bmatrix} \right| dudv \\ &= \int \int_{[1,0] \times [-2,0]} (u - v) (2u - v) |1| dudv \\ &= \int_0^1 \left[\int_{-2}^0 (u - v) (2u - v) dv \right] du \\ &= 7. \end{aligned}$$

Ejemplo 16.16. Calcular

$$\int \int_D \log(x^2 + y^2) dxdy$$

donde D es la región del primer cuadrante que está entre los arcos de círculo

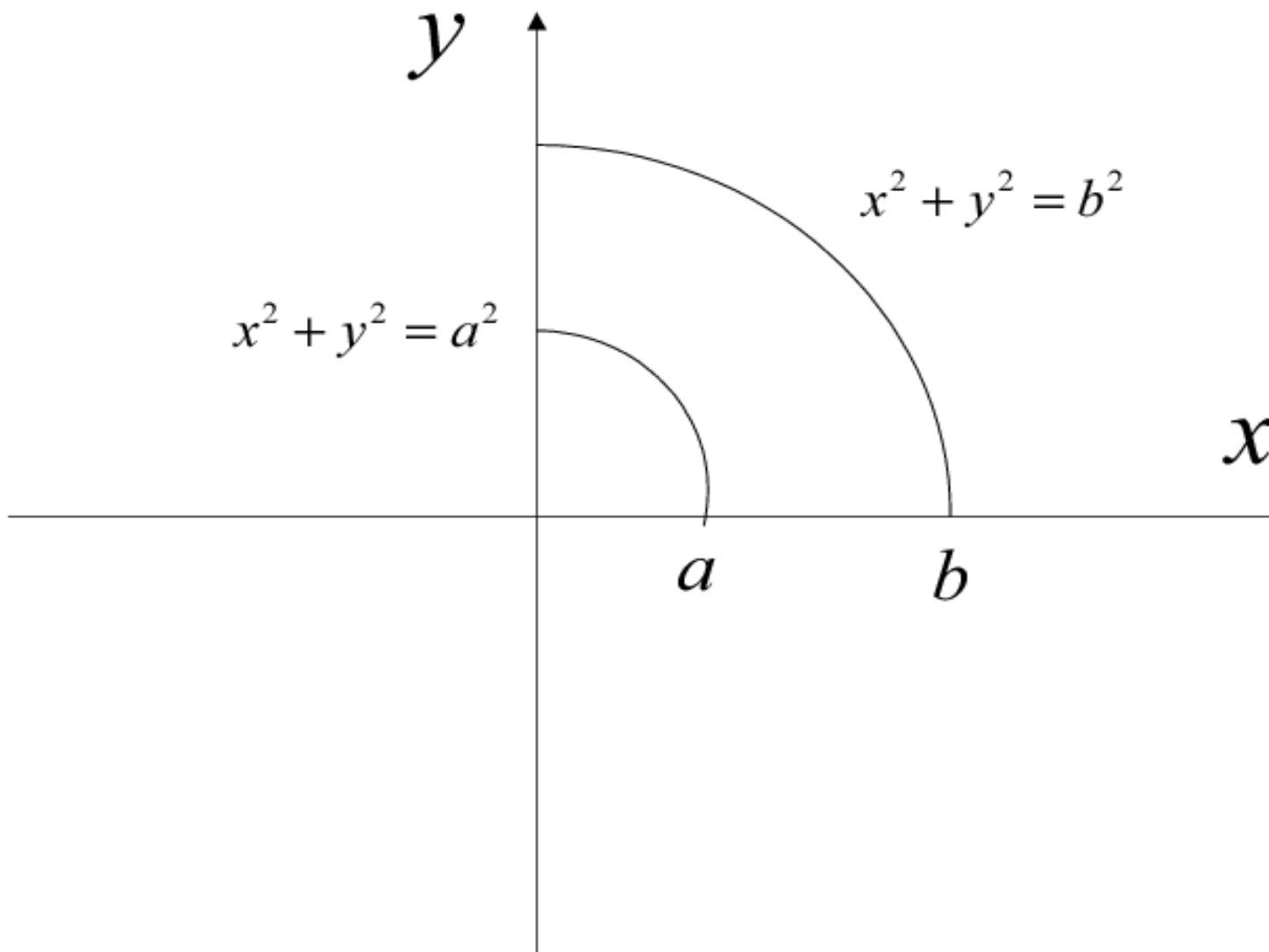
$$x^2 + y^2 = a^2$$

y

$$x^2 + y^2 = b^2$$

donde

$$0 < a < b.$$



$$D = \{(x, y) : a^2 \leq x^2 + y^2 \leq b^2\}$$

Si empleamos el cambio a coordenadas polares

$$x = x(r, \theta) = r \cos \theta,$$

$$y = y(r, \theta) = r \sin \theta$$

entonces como

$$x^2 + y^2 = r^2 \cos^2 \theta + r^2 \sin^2 \theta = r^2$$

se tiene que

$$a^2 \leq r^2 \leq b,$$

entonces

$$a \leq r \leq b.$$

Además,

$$0 \leq \theta \leq \frac{\pi}{2}.$$

Entonces

$$(r, \theta) \in [a, b] \times \left[0, \frac{\pi}{2}\right].$$

Empleando el Teorema del cambio de variable se tiene que

$$\begin{aligned} & \int \int_D \log(x^2 + y^2) \, dx dy \\ &= \int \int_{[a, b] \times [0, \frac{\pi}{2}]} \log(x(r, \theta)^2 + y(r, \theta)^2) |r| \, dr d\theta \\ &= \int_a^b \left[\int_0^{\pi/2} \log(r^2 \cos^2 \theta + r^2 \sin^2 \theta) r d\theta \right] dr \\ &= \int_a^b \left[\int_0^{\pi/2} \log(r^2) r d\theta \right] dr \\ &= \int_a^b \left[\int_0^{\pi/2} (2r \log r) d\theta \right] dr \\ &= \pi \int_a^b r \log r \, dr \\ &= \pi \left[\left(\frac{b^2}{2} \log b - \frac{b^2}{4} \right) - \left(\frac{a^2}{2} \log a - \frac{a^2}{4} \right) \right]. \end{aligned}$$