Apuntes de cálculo diferencial en una y varias variables reales

Eduardo Liz Marzán

Septiembre de 2022

Índice general

Introducción

Se recogen aquí los apuntes de la asignatura "Cálculo I" adaptada a los grados iniciados en el curso 2010/2011. En lo que respecta a las titulaciones de *Ingeniería de la energía* e *Ingeniería de los recursos mineros y energéticos* en la Universidad de Vigo, el grado tiene tres asignaturas cuya docencia corresponde al área de Matemática Aplicada. En el primer cuatrimestre se imparten *Álgebra Lineal* y *Cálculo I* y en el segundo cuatrimestre *Cálculo II*.

La primera parte del cálculo (a la que corresponden estos apuntes) se dedica al estudio de la continuidad y diferenciabilidad de funciones de una y varias variables reales y sus aplicaciones.

Estas notas no pretenden sustituir ni a los apuntes de clase ni a los libros de la bibliografía (que incluyen mucho más material, en particular muchos más ejemplos y aplicaciones), sino servir de ayuda para que los alumnos tengan el material del curso organizado.

Agradezco la cuidadosa lectura y las observaciones de Guillermo García Lomba, que han ayudado a mejorar versiones anteriores.

Eduardo Liz Marzán
Vigo, septiembre de 2022.

Capítulo 1

Conceptos preliminares

1.1. Introducción

En la primera parte del curso se estudia el cálculo en una variable real, con lo que el conjunto protagonista es el de los números reales, que denotaremos por $\mathbb{R}$. En esta asignatura el conjunto $\mathbb{R}$ juega un papel "más dinámico"que en la de Álgebra ya que, aparte de las operaciones, nos interesa ver la recta real como un conjunto en el que se mueve una magnitud escalar (por ejemplo, el tiempo). En este capítulo preliminar recordaremos algunas cuestiones importantes que se utilizan en los temas siguientes, como la relación de orden total de los números reales y el concepto de función.

1.2. Intervalos y valor absoluto

Una característica fundamental en el conjunto $\mathbb{R}$ de los números reales es que existe una relación de orden total compatible con las operaciones. En particular, se cumplen las siguientes propiedades:

1. Si $x \leq y$ y $z \in \mathbb{R}$ entonces $x + z \leq y + z$.

2. Si $x \leq y$ y $\lambda > 0$ entonces $\lambda x \leq \lambda y$.

3. Si $x \leq y$ y $\lambda < 0$ entonces $\lambda x \geq \lambda y$. En particular, $x \leq y \Longrightarrow -x \geq -y$.

La relación de orden permite definir los intervalos, que son los subconjuntos de números reales que usaremos más habitualmente. Si a, b son dos números reales tales que $a < b$, se definen los siguientes intervalos:

- $(a, b) = \{x \in \mathbb{R} \,/\, a < x < b\}$ (Intervalo **abierto** de extremos a y b).

- $[a, b] = \{x \in \mathbb{R} \,/\, a \leq x \leq b\}$ (Intervalo **cerrado** de extremos a y b).

- $[a, b) = \{x \in \mathbb{R} \,/\, a \leq x < b\} \quad , \quad (a, b] = \{x \in \mathbb{R} \,/\, a < x \leq b\}$.

Estos dos últimos se llaman intervalos semiabiertos o semicerrados. Todos estos tipos de intervalos se llaman intervalos acotados. Los intervalos cerrados y acotados $[a, b]$ se llaman también **intervalos compactos**. Existen otros intervalos no acotados:

- **Abiertos:** $(a, \infty) = \{x \in \mathbb{R} \,/\, x > a\}$, $(-\infty, b) = \{x \in \mathbb{R} \,/\, x < b\}$.
- **Cerrados:** $[a, \infty) = \{x \in \mathbb{R} \,/\, x \geq a\}$, $(-\infty, b] = \{x \in \mathbb{R} \,/\, x \leq b\}$.

Obsérvese que $\mathbb{R} = (-\infty, \infty)$ también es un intervalo no acotado.

El valor absoluto

Se define el valor absoluto de un número real x como

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

Esto quiere decir que los números reales x y $-x$ tienen el mismo valor absoluto, que representa su distancia a cero: $d(x, 0) = d(-x, 0) = |x|$.

Se define la distancia entre dos números reales x e y como el valor absoluto de su diferencia: $d(x, y) = |x - y|$.

En $\mathbb{R}$ hay una clara relación entre la distancia y los intervalos. Por ejemplo, el intervalo $(1, 5)$ se puede caracterizar:

- en términos de orden, como los números mayores que 1 y menores que 5;
- en términos de distancia, como los números cuya distancia al centro 3 es menor que el radio 2.

Así:

$$(1, 5) = \{x \in \mathbb{R} \,/\, 1 < x < 5\} = \{x \in \mathbb{R} \,/\, d(x, 3) < 2\} = \{x \in \mathbb{R} \,/\, |x - 3| < 2\}.$$

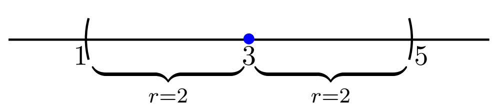


Figura 1.1: Intervalo abierto $(1, 5)$ de centro 3 y radio 2.

1.3. Funciones reales de una variable real

Las funciones reales de variable real son el objeto principal de los primeros temas del curso. En esta sección repasamos algunos conceptos relacionados.

Si D es un subconjunto de $\mathbb{R}$, una función real de una variable real (en adelante nos referiremos a ella únicamente como función) es una correspondencia $f : D \to \mathbb{R}$ que asigna a cada número $x \in D$ un número $f(x) \in \mathbb{R}$ que se llama imagen de x.

Informalmente, una función real de una variable real se puede entender como una "máquina"que transforma una variable de entrada $x \in \mathbb{R}$ en una variable de salida $f(x) \in \mathbb{R}$.

$$\xrightarrow{x} \boxed{\ f\ } \xrightarrow{f(x)}$$

Los ejemplos más usuales de funciones que utilizaremos a lo largo del curso son los polinomios, la función exponencial e^x, el logaritmo neperiano $\ln(x)$ y las funciones trigonométricas $\operatorname{sen}(x)$ y $\cos(x)$. Con estas funciones se pueden definir otras mediante operaciones de suma, producto, división y composición.

Elementos de una función

- El conjunto D donde está definida f se llama dominio de definición de f. Por simplicidad, entenderemos que el dominio $D(f)$ es un intervalo o una unión de intervalos. Por ejemplo, el dominio de definición de la función racional $f(x) = 1/x$ es

  $$D(f) = \{x \in \mathbb{R} \,/\, x \neq 0\} = (-\infty, 0) \cup (0, \infty).$$

  En particular, para las funciones que se obtienen como potencias de otras, el dominio de definición solo incluirá el conjunto de puntos donde la base es positiva. Por ejemplo, si $f(x) = x^x$ entonces su dominio de definición es $D(f) = \{x \in \mathbb{R} \,/\, x > 0\} = (0, \infty)$. Obsérvese que hay puntos fuera del dominio donde la función tiene sentido. Por ejemplo,

  $$f(-2) = (-2)^{-2} = 1/(-2)^2 = 1/4,$$

  pero sería imposible definir el dominio de definición como una unión de intervalos para los números reales negativos.

- Se llama imagen de f o rango de f al conjunto de valores que puede tomar $f(x)$, es decir, $f(D) = \{f(x) \,/\, x \in D\}$. Por ejemplo, el rango de la función $f(x) = \cos(x)$ es $[-1, 1]$.

- La gráfica de f es el subconjunto de $\mathbb{R}^2$ definido por $G(f) = \{(x, f(x)) \,/\, x \in D(f)\}$. Si proyectamos la gráfica sobre el eje horizontal obtenemos su dominio de definición y si la proyectamos sobre el eje vertical obtenemos su imagen. En la figura 1.2 se representa la gráfica de la función $f(x) = \sqrt{(1-x)/(1+x)}$, cuyo dominio es $D(f) = (-1, 1]$ y cuya imagen es $\operatorname{Im}(f) = [0, \infty)$ (en $x = -1$ tiene una asíntota vertical).

Composición de funciones

Consideremos dos funciones $f : D_1 \to \mathbb{R}$ y $g : D_2 \to \mathbb{R}$ Si $f(D_1) \subset D_2$, entonces se puede definir la composición $(g \circ f) : D_1 \to \mathbb{R}$ como $(g \circ f)(x) = g(f(x))$, $\forall\, x \in D_1$.

Por ejemplo, si $f(x) = x^2 + 1$ y $g(x) = \ln(x)$, se define $(g \circ f) : \mathbb{R} \to \mathbb{R}$ como

$$(g \circ f)(x) = g(f(x)) = \ln(x^2 + 1),\ \forall\, x \in \mathbb{R}.$$

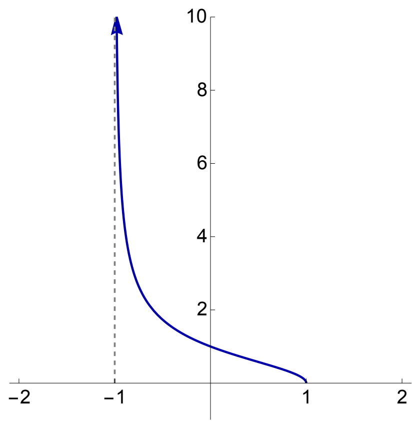


Figura 1.2: Gráfica de la función $f(x) = \sqrt{(1-x)/(1+x)}$.

Funciones inversas

Se dice que dos funciones $f : D_1 \to \mathbb{R}$ y $g : D_2 \to \mathbb{R}$ son **inversas** si $(g \circ f)(x) = x$, $\forall\, x \in D_1$ y $(f \circ g)(x) = x$, $\forall\, x \in D_2$. Si f y g son inversas, se denota $g = f^{-1}$.

Por ejemplo, las funciones $f(x) = e^x$ y $g(x) = \ln(x)$ son inversas ya que $\ln(e^x) = x$, $\forall\, x \in \mathbb{R}$ y $e^{\ln(x)} = x$, $\forall\, x \in (0, \infty)$.

Las funciones que tienen inversa son las funciones estrictamente crecientes o estrictamente decrecientes. Recordemos que una función $f : I \to \mathbb{R}$ definida en un intervalo real I es estrictamente **creciente** si para cualquier par de números reales distintos $x, y \in I$ tales que $x < y$ se cumple que $f(x) < f(y)$. Análogamente, f es estrictamente **decreciente** si $x < y \Longrightarrow f(x) > f(y)$.

Las gráficas de f y f^{-1} son simétricas respecto de la recta $y = x$. Es consecuencia de que el punto (x, y) está en la gráfica de f si y solo si el punto (y, x) está en la gráfica de f^{-1}:

$$(x, y) \in G(f) \Longleftrightarrow y = f(x) \Longleftrightarrow x = f^{-1}(y) \Longleftrightarrow (y, x) \in G(f^{-1}).$$

En la figura 1.3 se muestra la gráfica de $f(x) = \cos(x)$ (en azul) restringida al intervalo $[0, \pi]$, donde f es estrictamente decreciente. Su inversa $f^{-1}(x) = \arccos(x)$ se muestra en rojo; su dominio de definición es el intervalo $[-1, 1]$, que es la imagen de f. Obsérvese que los puntos $(0, 1)$ y $(\pi, -1)$ están en la gráfica de f, mientras que $(1, 0)$ y $(-1, \pi)$ están en la gráfica de f^{-1}.

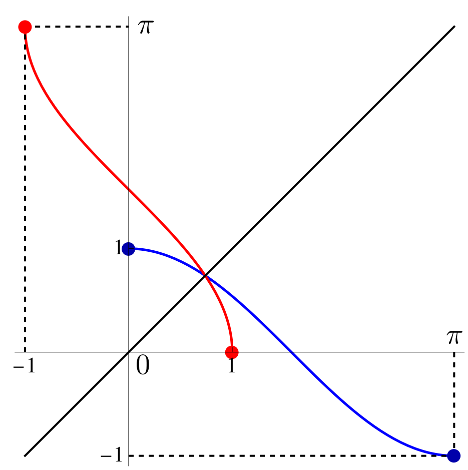


Figura 1.3: Gráficas de $f(x) = \cos(x)$ y su inversa $f^{-1}(x) = \arccos(x)$. La recta en color negro es $y = x$.

Capítulo 2

Límites y continuidad de funciones de una variable

2.1. Introducción

En este capítulo se introducen los conceptos de límite de una función en un punto y de límite en infinito. La noción de límite conduce a la importante definición de continuidad. Al final del tema se enuncian algunos resultados importantes que dependen de forma esencial de la continuidad.

2.2. Límite de una función en un punto

El concepto de límite permite, entre otras cosas, calcular el valor al que se aproxima una función cuando x se aproxima a un punto concreto x_0. Por ejemplo, la función $f(x) = \operatorname{sen}(x)/x$ no se puede evaluar en $x_0 = 0$ pero el cálculo de límites permite deducir que $f(x)$ se aproxima a 1 cuando x tiende a 0.

Para definir el concepto de límite en un punto x_0 es necesario que la función esté definida en las proximidades de x_0, es decir, en un intervalo $(x_0 - r, x_0)$, en $(x_0, x_0 + r)$, o en ambos, donde r es un número positivo.

Límites laterales

Si f una función y x_0 es un número real de tal forma que f está definida en un intervalo $(x_0 - r, x_0)$ para algún $r > 0$ entonces diremos que existe el **límite por la izquierda** de f en x_0 si ocurre una de las siguientes posibilidades:

1. El límite por la izquierda es un número real y_0, es decir, $f(x)$ se aproxima a y_0 cuando x se aproxima a x_0 por la izquierda. Formalmente:

$$\lim_{x \to x_0^-} f(x) = y_0 \iff [\forall\, \varepsilon > 0,\ \exists \delta > 0 \,/\, x \in (x_0 - \delta, x_0) \Longrightarrow |f(x) - y_0| < \varepsilon]\,.$$

2. El límite por la izquierda es infinito, es decir, f toma valores arbitrariamente grandes cuando x se aproxima a x_0 por la izquierda. Formalmente:

$$\lim_{x\to x_0^-} f(x) = \infty \iff [\forall\, M > 0,\, \exists\delta > 0 \,/\, x \in (x_0 - \delta, x_0) \Longrightarrow f(x) > M]\,.$$

3. El límite por la izquierda es menos infinito, es decir, f toma valores arbitrariamente pequeños (negativos con valor absoluto grande) cuando x se aproxima a x_0 por la izquierda. Formalmente:

$$\lim_{x\to x_0^-} f(x) = -\infty \iff [\forall\, M > 0,\, \exists\delta > 0 \,/\, x \in (x_0 - \delta, x_0) \Longrightarrow f(x) < -M]\,.$$

   Por ejemplo, $\lim\limits_{x\to 0^-} (1/x) = -\infty$.

Si no sucede ninguna de las situaciones anteriores, diremos que no existe el límite por la izquierda de f en x_0.

De modo completamente análogo (cambiando $(x_0-\delta, x_0)$ por $(x_0, x_0+\delta)$) se define el **límite por la derecha** de f en x_0 si f está definida en un intervalo de la forma (x_0, x_0+r) para algún $r > 0$. Se denota $\lim\limits_{x\to x_0^+} f(x)$. Por ejemplo, $\lim\limits_{x\to 0^+} (1/x) = \infty$.

Concepto de límite

Si $x_0 \in \mathbb{R}$ y f es una función definida en los intervalos $(x_0 - r, x_0)$ y $(x_0, x_0 + r)$ para algún $r > 0$, se dice que existe el **límite** de f en x_0 si existen los límites laterales, son finitos y coinciden, es decir, $\lim\limits_{x\to x_0^-} f(x) = \lim\limits_{x\to x_0^+} f(x)$.

Por ejemplo, consideremos la función $f : (-\infty, 0)\cup(0, \infty) \to \mathbb{R}$ definida por $f(x) = \operatorname{sen}(x)/x$. En este caso, existe $\lim\limits_{x\to 0} f(x) = 1$.

Nótese que para definir el límite en x_0 no es necesario que x_0 sea un punto del dominio de f.

Asíntotas verticales

Si alguno de los límites laterales de f en un punto x_0 es ∞ o $-\infty$ entonces se dice que la recta vertical $x = x_0$ es una **asíntota** vertical a la gráfica de f.

- Si $\lim\limits_{x\to x_0^-} f(x) = \pm\infty$, la recta $x = x_0$ es una asíntota vertical a la gráfica de f por la izquierda.

- Si $\lim\limits_{x\to x_0^+} f(x) = \pm\infty$, la recta $x = x_0$ es una asíntota vertical a la gráfica de f por la derecha.

Por ejemplo, como vimos en la figura 1.2, la recta $x = -1$ es una asíntota vertical por la derecha a la gráfica de la función $f(x) = \sqrt{(1-x)/(1+x)}$.

2.3. Continuidad

Consideremos una función $f : I \to \mathbb{R}$ definida en un intervalo abierto I. Se dice que f es **continua** en un punto $x_0 \in I$ si existe el límite de f en x_0 y además

$$\lim_{x \to x_0} f(x) = f(x_0).$$

Si f está definida en un intervalo $[a, b]$, diremos que f es continua en a si existe el límite por la derecha de f en a y coincide con $f(a)$. De modo análogo se define la continuidad en el extremo derecho b del intervalo. Se dice que una función es continua en un conjunto D si es continua en todos los puntos de D. Si f no es continua en x_0, diremos que f tiene una discontinuidad en x_0. Consideraremos dos tipos distintos de discontinuidades:

1. **Discontinuidad de salto:** si existen los límites laterales de f en x_0 pero alguno de ellos es infinito o ambos son finitos y no coinciden. En el primer caso se dice que el salto es infinito y en el segundo que el salto es finito.

2. **Discontinuidad esencial:** si no existe alguno de los límites laterales de f en x_0.

Para ganar generalidad, entenderemos que para que haya una discontinuidad en un punto x_0 no es preciso que x_0 esté en el dominio de f, basta que f esté definida a ambos lados de x_0. Por ejemplo, la función $f(x) = 1/x$ tiene una discontinuidad de salto infinito en 0. La función es continua en cada uno de los intervalos que componen su dominio de definición: $(-\infty, 0)$ y $(0, \infty)$.

Un ejemplo típico de discontinuidad esencial lo proporciona la función $f(x) = \text{sen}(1/x)$, definida en $(-\infty, 0) \cup (0, \infty)$. Existe una discontinuidad esencial en $x_0 = 0$ ya que no existe $\lim_{x \to 0} f(x)$: no existe ningún valor concreto al que se aproxime $f(x)$ cuando x se aproxima a cero. La gráfica se representa en la figura 2.1. En las proximidades de cero (en cualquier intervalo de la forma $(-\varepsilon, \varepsilon)$, con $\varepsilon > 0$ arbitrariamente pequeño), la función toma todos los valores entre -1 y 1. La razón es que los valores arbitrariamente grandes de $1/x$ para x cerca de 0 producen oscilaciones de frecuencias arbitrariamente grandes.

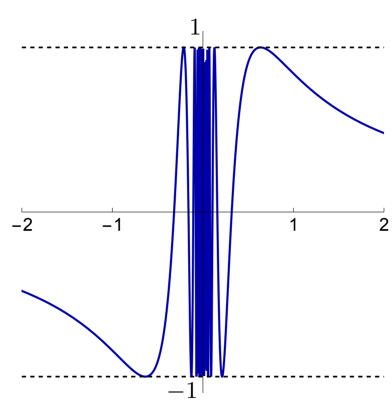


Figura 2.1: Gráfica de la función $f(x) = \text{sen}(1/x)$.

Algunos ejemplos de funciones continuas

Las funciones más comunes son continuas en sus dominios de definición. Por ejemplo:

- La función valor absoluto definida por
$$|x| = \begin{cases} -x & \text{si } x < 0 \\ x & \text{si } x \geq 0 \end{cases}$$
es continua en $\mathbb{R}$.

- Las funciones polinómicas $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$, donde $a_0, a_1, \ldots, a_n \in \mathbb{R}$, son continuas en $\mathbb{R}$.

- Las funciones racionales (cocientes de polinomios) $r(x) = p(x)/q(x)$ son continuas en su dominio de definición, es decir, en el conjunto de los números reales x tales que $q(x) \neq 0$. Si $p(x)$ y $q(x)$ no tienen factores comunes (fracción irreducible), la función $r(x)$ tiene discontinuidades de salto infinito en los ceros de $q(x)$, es decir, en los puntos donde $q(x) = 0$.

- La función exponencial e^x es continua en $\mathbb{R}$ y $\ln(x)$ es continua en $(0, \infty)$.

- Las funciones trigonométricas $\operatorname{sen}(x)$ y $\cos(x)$ son continuas en $\mathbb{R}$.

Los ejemplos anteriores, combinados con las propiedades que enunciamos a continuación, permiten probar la continuidad de muchas funciones.

Propiedades.

1. La composición de funciones continuas es una función continua. Por ejemplo, la función $f(x) = e^{|\cos(x)|}$ es continua en $\mathbb{R}$ ya que $f(x) = f_1(f_2(f_3(x)))$, donde $f_1(x) = e^x$, $f_2(x) = |x|$ y $f_3(x) = \cos(x)$ son continuas.

2. Si f y g son funciones continuas en x_0 entonces las funciones $(f+g)$ y $(f \cdot g)$ son continuas en x_0. La función f/g es continua en x_0 si $g(x_0) \neq 0$.

   Por ejemplo, la función
$$f(x) = \frac{\operatorname{sen}(x) + e^x}{x - 1}$$
   es continua para todo $x \neq 1$.

2.4. Límites en infinito

Consideremos una función f definida en el intervalo (a, ∞) para algún $a \in \mathbb{R}$. Diremos que existe el límite de f en ∞ si ocurre una de las siguientes posibilidades:

1. El límite en infinito es un número real y_0, es decir, $f(x)$ se aproxima a y_0 cuando x se hace suficientemente grande. Formalmente:
$$\lim_{x\to\infty} f(x) = y_0 \Longleftrightarrow [\forall\, \varepsilon > 0,\ \exists M > 0 \,/\, x > M \Longrightarrow |f(x) - y_0| < \varepsilon]\,.$$

2. El límite en infinito es infinito, es decir, $f(x)$ toma valores arbitrariamente grandes cuando x se hace suficientemente grande. Formalmente:

   $$\lim_{x\to\infty} f(x) = \infty \iff \left[\forall\, M > 0,\, \exists M' > 0 \,/\, x > M' \Longrightarrow f(x) > M\right].$$

3. El límite en infinito es menos infinito, es decir, $f(x)$ toma valores arbitrariamente pequeños (negativos con valor absoluto grande) cuando x se hace suficientemente grande. Formalmente:

   $$\lim_{x\to\infty} f(x) = -\infty \iff \left[\forall\, M > 0,\, \exists M' > 0 \,/\, x > M' \Longrightarrow f(x) < -M\right].$$

Si no sucede ninguna de las situaciones anteriores, diremos que no existe el límite de f en infinito.

De modo completamente análogo se define el límite de f en $-\infty$ si f está definida en $(-\infty, b)$ para algún $b \in \mathbb{R}$.

Si $y_0 \in \mathbb{R}$ es el límite de f en $\pm\infty$, entonces la recta horizontal $y = y_0$ es una **asíntota horizontal** de la gráfica de f.

Ejemplos:

1. $\lim\limits_{x\to\infty} e^x = \infty$, $\lim\limits_{x\to-\infty} e^x = 0$. Por tanto, la recta $y = 0$ es una asíntota horizontal de la gráfica de $f_1(x) = e^x$ en $-\infty$, pero f_1 no tiene una asíntota horizontal en ∞ porque crece indefinidamente.

2. $\lim\limits_{x\to\infty} (1+x)/x = \lim\limits_{x\to-\infty} (1+x)/x = 1$. En consecuencia, la recta $y = 1$ es una asíntota horizontal de la gráfica de $f_2(x) = (1+x)/x$ en ∞ y $-\infty$.

3. No existe $\lim\limits_{x\to\infty} \operatorname{sen}(x)$. La gráfica de la función $f_3(x) = \operatorname{sen}(x)$ es periódica y no tiene asíntotas.

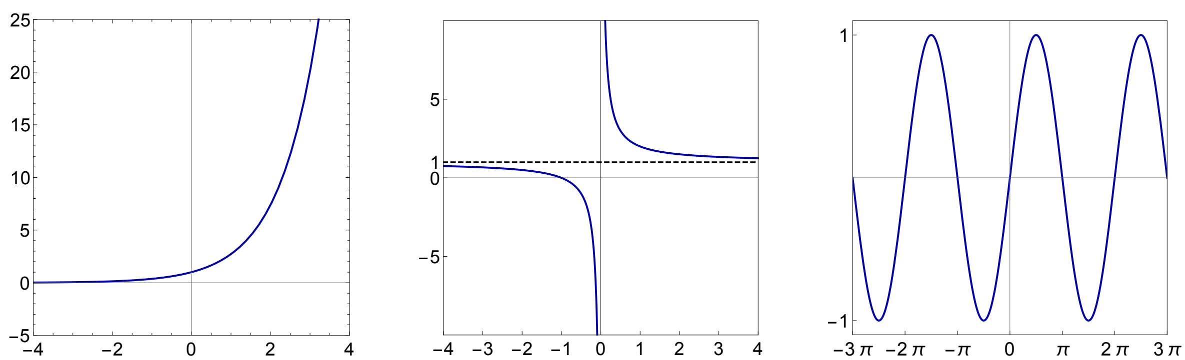


Figura 2.2: Gráficas de las funciones $f_1(x) = e^x$, $f_2(x) = (1+x)/x$ y $f_3(x) = \operatorname{sen}(x)$.

2.5. Cálculo de límites

Las siguientes propiedades son útiles para el cálculo de límites:

1. Si f y g son dos funciones y existen $\lim_{x \to x_0} f(x) \in \mathbb{R}$ y $\lim_{x \to x_0} g(x) \in \mathbb{R}$ (x_0 puede ser $\pm\infty$), entonces:

   a) $\lim_{x \to x_0} (f(x) + g(x)) = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x)$.

   b) $\lim_{x \to x_0} (f(x) \cdot g(x)) = \left(\lim_{x \to x_0} f(x)\right) \cdot \left(\lim_{x \to x_0} g(x)\right)$.

   c) $\lim_{x \to x_0} (f(x)/g(x)) = \left(\lim_{x \to x_0} f(x)\right) / \left(\lim_{x \to x_0} g(x)\right)$, si $\lim_{x \to x_0} g(x) \neq 0$.

   d) $\lim_{x \to x_0} \left(f(x)^{g(x)}\right) = \left(\lim_{x \to x_0} f(x)\right)^{\lim_{x \to x_0} g(x)}$, si $\lim_{x \to x_0} f(x) > 0$.

   Algunas de las propiedades anteriores se pueden extender al caso en que alguno de los límites es $\pm\infty$, teniendo en cuenta las relaciones formales $\infty + \infty = \infty$, $\infty \cdot \infty = \infty$, $\lambda \cdot \infty = \infty$ si $\lambda > 0$, $\lambda \cdot \infty = -\infty$ si $\lambda < 0$, $\infty^\infty = \infty$.

2. Si g es continua y existe $\lim_{x \to x_0} f(x)$ entonces $\lim_{x \to x_0} g(f(x)) = g\left(\lim_{x \to x_0} f(x)\right)$.

   Por ejemplo, $\lim_{x \to \infty} \ln\left(\frac{x+1}{x-1}\right) = \ln\left(\lim_{x \to \infty} \frac{x+1}{x-1}\right) = \ln(1) = 0$.

3. Si $\lim_{x \to x_0} f(x) = 0$ y g está acotada en un entorno de x_0 entonces $\lim_{x \to x_0} (f(x) \cdot g(x)) = 0$. Por ejemplo, $\lim_{x \to \infty} \operatorname{sen}(x)/x = 0$, ya que $\operatorname{sen}(x)$ está acotada y $\lim_{x \to \infty} 1/x = 0$.

En ocasiones el cálculo de límites conduce a indeterminaciones. Las más usuales son $(\infty - \infty)$, ∞/∞, $0/0$, $0 \cdot \infty$, 1^∞, ∞^0 y 0^0. En muchos casos las indeterminaciones se resuelven aplicando la regla de L'Hôpital, que veremos en el tema siguiente.

La función $f(x) = \operatorname{sen}(x)/x$ proporciona un ejemplo inusual de asíntota horizontal. Como $\lim_{x \to \infty} \operatorname{sen}(x)/x = 0$, la recta $y = 0$ es una asíntota horizontal a la gráfica de f en infinito. En este caso, la gráfica de f corta infinitas veces a la asíntota, algo que no habíamos observado en ejemplos anteriores, en los que la convergencia era monótona. En el caso de este ejemplo, la convergencia es oscilatoria (ver la figura 2.3).

2.6. Continuidad e intervalos compactos

La siguiente propiedad relaciona continuidad e intervalos:

Si I es un intervalo real y $f : I \to \mathbb{R}$ es una función continua entonces su conjunto imagen $f(I) = \{f(x) \,/\, x \in I\}$ también es un intervalo. Además, si $I = [a, b]$ es un intervalo compacto entonces $f([a, b]) = [c, d]$ también es un intervalo compacto.

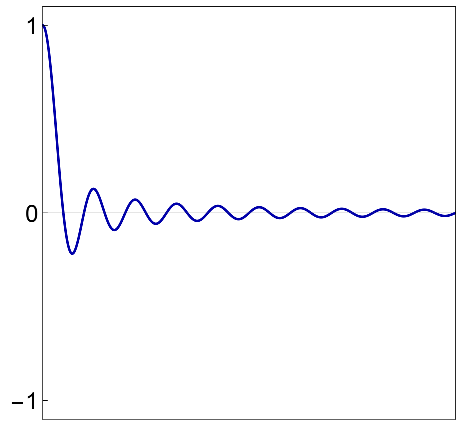


Figura 2.3: Gráfica de la función $f(x) = \text{sen}(x)/x$.

Como consecuencias de esta propiedad obtenemos dos resultados importantes:

Teorema de los valores extremos: Si $f : [a, b] \to \mathbb{R}$ es una función continua entonces f alcanza el máximo y el mínimo globales, es decir, existen $m, M \in \mathbb{R}$ tales que

$$m = \min_{x\in[a,b]} f(x) \ , \quad M = \max_{x\in[a,b]} f(x).$$

Por ejemplo, para la función $f : [0, 2] \to \mathbb{R}$ definida por $f(x) = x^4 - 2x^2$ es fácil probar que

$$\min_{x\in[0,2]} f(x) = f(1) = -1 \quad ; \quad \max_{x\in[0,2]} f(x) = f(2) = 8.$$

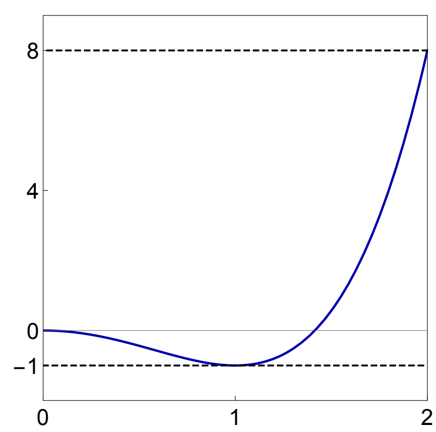


Figura 2.4: Gráfica de la función $f(x) = x^4 - 2x^2$ en el intervalo $[0, 2]$.

Este resultado no tiene por qué ser cierto si f no es continua o no está definida en un intervalo compacto. Por ejemplo, la función $f : (0, 2) \to \mathbb{R}$ definida por $f(x) = 1/x$ es continua pero no existe $\max_{x\in(0,2)} f(x)$ ya que $\lim_{x\to 0^+} f(x) = \infty$.

Teorema de Bolzano: Si $f : [a, b] \to \mathbb{R}$ es una función continua y $f(a)f(b) < 0$ entonces existe al menos un número $x \in (a, b)$ tal que $f(x) = 0$.

El teorema de Bolzano dice que si una función continua toma valores de distinto signo en dos puntos, entre ellos hay un punto donde la función vale cero. Es una herramienta útil para probar la existencia de ceros de una función (se dice que c es un cero o una raíz de f si $f(c) = 0$.)

Ejemplo: La función continua $f(x) = e^x - x^4$ tiene al menos un cero en el intervalo $(1, 2)$ ya que

$$f(1) = e - 1 > 0 \; ; \quad f(2) = e^2 - 16 < 0.$$

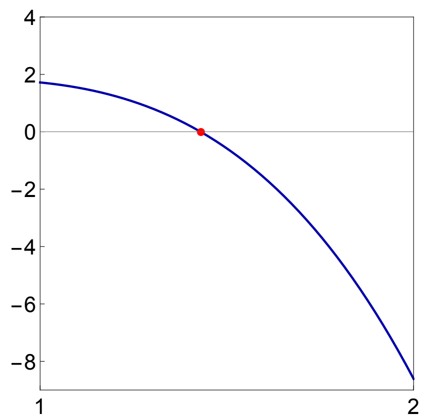


Figura 2.5: Gráfica de la función $f(x) = e^x - x^4$ en el intervalo $[1, 2]$. El cero de f se representa por el punto rojo en el que la gráfica corta a la recta $y = 0$ (eje horizontal).

Capítulo 3

Derivación de funciones de una variable

3.1. Introducción

En este capítulo se introduce el concepto de derivada de una función en un punto y algunas de sus propiedades. El cálculo diferencial es la herramienta más eficaz para obtener propiedades de una función como la determinación de sus extremos locales y sus intervalos de crecimiento, decrecimiento, concavidad y convexidad. También tiene importantes aplicaciones al cálculo de límites (y por tanto de asíntotas) y a la aproximación de funciones por polinomios.

3.2. Recta tangente y tasas de cambio

Una de las aplicaciones del cálculo de límites, que conduce al concepto de derivada, es la definición precisa de lo que se entiende por recta tangente a una curva en un punto. Esta definición se basa en que la pendiente de la recta tangente a una curva en el plano (x, y) representa la tasa de cambio instantáneo de y con respecto a x.

Si f es una función, la tasa promedio de cambio en un intervalo $[x_0, x_0 + h]$ de longitud h viene dada por el cociente de incrementos

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

Por tanto, tiene sentido definir la razón de cambio instantánea en un punto x_0 como

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Si este límite existe, se puede definir la recta tangente en x_0 como la recta que pasa por $(x_0, f(x_0))$ y tiene como pendiente dicho límite. Su interpretación es la siguiente: para cada h, el cociente de incrementos representa la pendiente de la recta que pasa por los puntos $(x_0, f(x_0))$ y $(x_0 + h, f(x_0 + h))$. La recta tangente se puede interpretar como el límite cuando h tiende a cero de estas rectas. Se ilustra en la figura 3.1 para el caso en que $h > 0$.

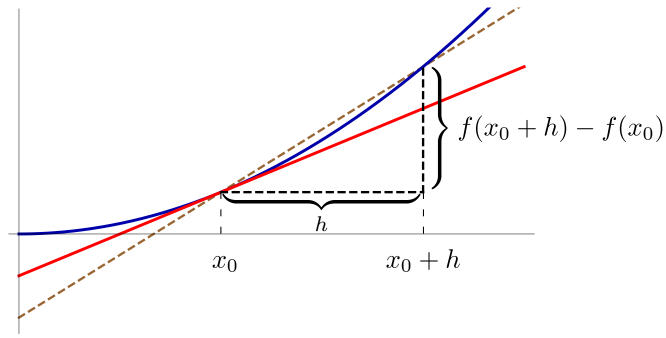


Figura 3.1: Gráfica de una función f (en azul) y su recta tangente en el punto $(x_0, f(x_0))$ (en rojo). La recta que une los puntos $(x_0, f(x_0))$ y $(x_0+h, f(x_0+h))$ se muestra en trazo discontinuo marrón.

La definición de tasa de cambio instantánea conduce al concepto de derivada de una función en un punto.

3.3. Concepto y propiedades de las derivadas

Consideremos una función continua $f : I \to \mathbb{R}$, donde I es un intervalo abierto. Se dice que f es derivable en un punto $x_0 \in I$, si existe el siguiente límite y es finito:

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}.$$

En caso de que exista, el límite se llama **derivada de f en x_0** y se denota por $f'(x_0)$. Si el límite no existe o es infinito, diremos que f no es derivable en x_0 o que no existe la derivada de f en x_0.

Una definición equivalente de $f'(x_0)$ se obtiene tomando $h = x - x_0$:

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

Ejemplo:
La función $f(x) = x^2$ es derivable en todo punto $x \in \mathbb{R}$ y además

$$f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \lim_{h\to 0} \frac{(x+h)^2-x^2}{h} = \lim_{h\to 0} \frac{h^2+2xh}{h} = \lim_{h\to 0}(h+2x) = 2x.$$

Ecuación de la recta tangente

Como hemos comentado antes, la derivada de f en x_0 representa la pendiente de la recta tangente a la gráfica de f en el punto $(x_0, f(x_0))$. Por tanto, la ecuación de dicha recta tangente es

$$y - f(x_0) = f'(x_0)(x - x_0).$$

Se puede interpretar la tangente como la recta que "mejor predice" el valor de $f(x_0 + h)$ o $f(x_0 - h)$ para $h > 0$ pequeño, conociendo el valor de $f(x_0)$.

Por ejemplo, la función $f(x) = 2 + x^2$ es derivable en $x_0 = 1$, $f(1) = 3$ y $f'(1) = 2$. Por tanto, la ecuación de la recta tangente a la gráfica de f en el punto $(1, 3)$ es $y - 3 = 2(x - 1)$.

Función derivada.

Si $f : I \to \mathbb{R}$ es una función definida en un intervalo real I, se dice que f es derivable en I si es derivable en todos los puntos de I. En este caso se puede definir la función derivada de f:

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

Por ejemplo, la función derivada de $f(x) = x^3$ es $f'(x) = 3x^2$, que está definida para todo $x \in \mathbb{R}$.

Derivadas laterales y derivadas de funciones definidas a trozos

También se pueden definir las derivadas laterales de f en x_0. Se llama derivada por la izquierda de f en x_0, y se denota $f'(x_0^-)$, al siguiente límite:

$$f'(x_0^-) = \lim_{h \to 0^-} \frac{f(x_0 + h) - f(x_0)}{h},$$

siempre que el límite exista y sea finito.

Análogamente, se define la derivada por la derecha $f'(x_0^+)$ tomando el límite por la derecha.

Ejemplo: La función $f : [0, \infty) \to \mathbb{R}$ definida por $f(x) = \sqrt{x}$ no es derivable (por la derecha) en 0 ya que

$$f'(0^+) = \lim_{h \to 0^+} \frac{f(0 + h) - f(0)}{h} = \lim_{h \to 0^+} \frac{\sqrt{h}}{h} = \lim_{h \to 0^+} \frac{1}{\sqrt{h}} = \infty.$$

En caso de que una función f esté definida en un intervalo compacto $[a, b]$, la derivada de f en a se entiende como la derivada por la derecha y la derivada de f en b se entiende como la derivada por la izquierda.

El siguiente resultado simplifica el estudio de la derivabilidad en funciones definidas a trozos:

Consideremos una función f continua en un intervalo $(x_0 - r, x_0 + r)$ y derivable en los intervalos $(x_0 - r, x_0)$ y $(x_0, x_0 + r)$ para algún $r > 0$. Supongamos que existen $\lim_{x \to x_0^-} f'(x)$ y $\lim_{x \to x_0^+} f'(x)$. Entonces f es derivable en x_0 si y solo si $\lim_{x \to x_0^-} f'(x) = \lim_{x \to x_0^+} f'(x)$.

Ejemplo: Consideremos la función $f : \mathbb{R} \to \mathbb{R}$ definida por

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

Claramente f es derivable para todo $x \neq 0$ y

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

Como $\lim\limits_{x \to 0^-} f'(x) = -1 \neq \lim\limits_{x \to 0^+} f'(x) = 1$, f no es derivable en 0.

Derivadas sucesivas y regularidad de una función

Si f es derivable en I y la función f' es también derivable en I entonces se define la derivada segunda (o derivada de orden 2) de f como $f''(x) = (f')'(x)$, $\forall\, x \in I$. Por ejemplo, la derivada segunda de $f(x) = x^3$ es $f''(x) = 6x$, que también está definida para todo $x \in \mathbb{R}$. Este proceso se puede seguir inductivamente para obtener todas las derivadas sucesivas de f mientras sea posible.

Cuantas más veces se pueda derivar una función, se dice que tiene mayor regularidad. En cierto modo, la regularidad hace que la gráfica de f sea suave. Por ejemplo, si una función f es dos veces derivable entonces su gráfica y la de su derivada son "suaves". La regularidad de una función se suele medir con el concepto de clase: se dice que una función es de clase $\mathcal{C}^n$ en I si f es n veces derivable en I y su derivada de orden n es continua. En particular, las funciones de clase $\mathcal{C}^1$ son las funciones derivables con derivada continua.

Si f tiene derivadas de todos los órdenes entonces se dice que f es de clase $\mathcal{C}^\infty$ (se lee "f es de clase clase c infinito"). Por ejemplo, la función $f(x) = e^x$ es de clase $\mathcal{C}^\infty$ en $\mathbb{R}$ ya que existen todas las derivadas sucesivas de f. De hecho, $f^{n)}(x) = e^x$, $\forall\, n \in \mathbb{N}$.

Observación. La continuidad de una función es una condición necesaria para que sea derivable. Sin embargo, no toda función continua es derivable. Por ejemplo, la función $f(x) = |x|$ es continua en $\mathbb{R}$ pero no es derivable en $x = 0$.

Propiedades de las derivadas

Si $f : I \to \mathbb{R}$ y $g : I \to \mathbb{R}$ son dos funciones derivables en un punto $x_0 \in I$, entonces:

1. $(f + g)$ es derivable en x_0 y $(f + g)'(x_0) = f'(x_0) + g'(x_0)$.
2. (λf) es derivable en x_0 y $(\lambda f)'(x_0) = \lambda f'(x_0)$, $\forall\, \lambda \in \mathbb{R}$.
3. $(f \cdot g)$ es derivable en x_0 y $(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$.
4. Si $g(x_0) \neq 0$, (f/g) es derivable en x_0 y

$$(f/g)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}.$$

Regla de la cadena

La regla de la cadena es la fórmula para la derivada de la composición de dos funciones.

Regla de la cadena: Consideremos dos funciones $f : D_1 \to \mathbb{R}$ y $g : D_2 \to \mathbb{R}$ tales que $f(D_1) \subset D_2$. Si f es derivable en x_0 y g es derivable en $f(x_0)$ entonces $(g \circ f)$ es derivable en x_0 y además

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

Ejemplo. La función $f(x) = \ln(x^2 + 1)$ se puede escribir como $f(x) = g(h(x))$, con $g(x) = \ln(x)$ y $h(x) = x^2 + 1$. Por tanto, f es derivable en $\mathbb{R}$ y

$$f'(x) = g'(h(x))h'(x) = \frac{h'(x)}{h(x)} = \frac{2x}{x^2+1}\,,\ \forall\, x \in \mathbb{R}.$$

3.4. Derivación de funciones implícitas

En algunas ocasiones la variable y no está expresada como una función explícita de x (es decir, $y = f(x)$), sino de forma implícita mediante una expresión $F(x, y) = 0$. En general no es posible despejar y en función de x, pero muchas veces se puede calcular la derivada de y respecto de x utilizando derivación implícita.

Ejemplo: Consideremos la curva definida implícitamente por la ecuación $e^{x+y} + xy^2 = 1$. Vamos a calcular su recta tangente en el punto $(0, 0)$.

Escribiendo $y = y(x)$ y aplicando la regla de la cadena, se tiene

$$e^{x+y(x)} + x(y(x))^2 = 1 \Longrightarrow (1 + y'(x))e^{x+y(x)} + (y(x))^2 + 2xy(x)y'(x) = 0.$$

Para $x = y = 0$, se obtiene $y'(0) = -1$, y por tanto la ecuación de la recta tangente en $(0, 0)$ es $y = -x$.

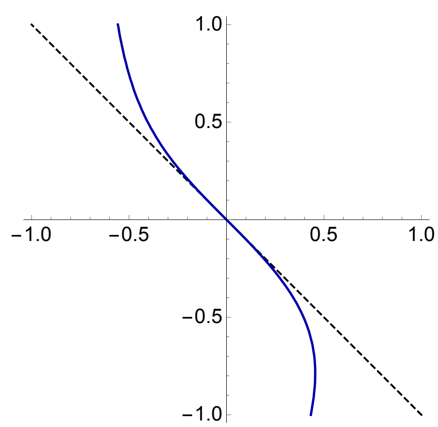


Figura 3.2: Gráficas de la curva $e^{x+y} + xy^2 = 1$ (trazo continuo en azul) y su recta tangente $y = -x$ (trazo discontinuo) en el punto $(0, 0)$.

3.5. La regla de L'Hôpital.

En esta sección se introduce la regla de L'Hôpital, que es el método más efectivo para el cálculo de límites en el caso de indeterminaciones.

Regla de L'Hôpital: Consideremos un punto $x_0 \in \mathbb{R}$ y dos funciones f y g definidas y derivables en los intervalos $(x_0 - r, x_0)$ y $(x_0, x_0 + r)$ para algún $r > 0$, de tal manera que g no se anula en esos intervalos. Supongamos que se cumplen las siguientes condiciones:

(a) $\lim\limits_{x \to x_0} f(x) = \lim\limits_{x \to x_0} g(x) = 0$ o $\lim\limits_{x \to x_0} f(x) = \lim\limits_{x \to x_0} g(x) = \infty$.

(b) Existe $\lim\limits_{x \to x_0} (f'(x)/g'(x)) = l$ (l puede ser finito o infinito).

Entonces existe $\lim\limits_{x \to x_0} (f(x)/g(x))$ y además

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}.$$

El resultado del teorema sigue siendo cierto si $x_0 = \pm\infty$ y las funciones f y g son derivables en intervalos de la forma (a, ∞) o $(-\infty, b)$. También se aplica para el cálculo de límites laterales en caso de que f/g solo esté definida a la izquierda o a la derecha de x_0.

Ejemplos:

1. Como $\lim\limits_{x \to 0} \operatorname{sen}(x) = \lim\limits_{x \to 0} x = 0$, se puede aplicar la regla de L'Hôpital para calcular

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

2. Es importante que el límite del cociente de las derivadas exista. Por ejemplo,

$$\lim_{x \to \infty} \frac{x + \operatorname{sen}(x)}{1 + x} = \lim_{x \to \infty} \frac{x}{1 + x} + \lim_{x \to \infty} \frac{\operatorname{sen}(x)}{1 + x} = 1 + 0 = 1,$$

pero no se puede aplicar la regla de L'Hôpital porque el límite del cociente de las derivadas es $\lim\limits_{x \to \infty} (1 + \cos(x))$, que no existe porque el coseno es una función periódica.

3. Un ejemplo interesante de que la regla de L'Hôpital no siempre es la manera adecuada de resolver una indeterminación es el siguiente:

$$\lim_{x \to \infty} \frac{\sqrt{1 + x^2}}{x} = 1,$$

pero la aplicación de la regla de L'Hôpital no lo resuelve. De hecho, aplicando la regla dos veces se obtiene:

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

Órdenes de crecimiento

Hay muchas funciones que tienen límite infinito en infinito. Sin embargo, también es importante la magnitud del crecimiento. Por ejemplo, es sabido que el crecimiento exponencial es más rápido que el crecimiento logarítmico.

Si f y g son dos funciones tales que

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) = \infty,$$

diremos que f y g tienen el mismo orden de crecimiento cuando cuando x tiende a infinito si se cumple que

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = c > 0.$$

Diremos que el orden de crecimiento de f en infinito es mayor que el de g cuando

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

Los siguientes órdenes de crecimiento en infinito están ordenados de mayor a menor:

1. Crecimiento exponencial: f crece exponencialmente cuando x tiende a infinito si existen constantes $a > 0$ y $c > 0$ tales que $\lim_{x\to\infty} (f(x)/e^{ax}) = c$.

2. Crecimiento superlineal: f crece superlinealmente cuando x tiende a infinito si existen constantes $a > 1$ y $c > 0$ tales que $\lim_{x\to\infty} (f(x)/x^a) = c$.

3. Crecimiento lineal: f crece linealmente cuando x tiende a infinito si existe una constante $c > 0$ tal que $\lim_{x\to\infty} (f(x)/x) = c$.

4. Crecimiento sublineal: f crece sublinealmente cuando x tiende a infinito si existen constantes $a \in (0,1)$ y $c > 0$ tales que $\lim_{x\to\infty} (f(x)/x^a) = c$.

5. Crecimiento logarítmico: f tiene crecimiento logarítmico cuando x tiende a infinito si existe una constante $c > 0$ tal que $\lim_{x\to\infty} (f(x)/\ln(x)) = c$.

Por ejemplo, veamos que el crecimiento sublineal es más rápido que el logarítmico usando la regla de L'Hôpital. En efecto,

$$\lim_{x\to\infty} \frac{x^a}{\ln x} = \lim_{x\to\infty} \frac{ax^{a-1}}{x^{-1}} = \lim_{x\to\infty} ax^a = \infty, \ \forall\, a > 0.$$

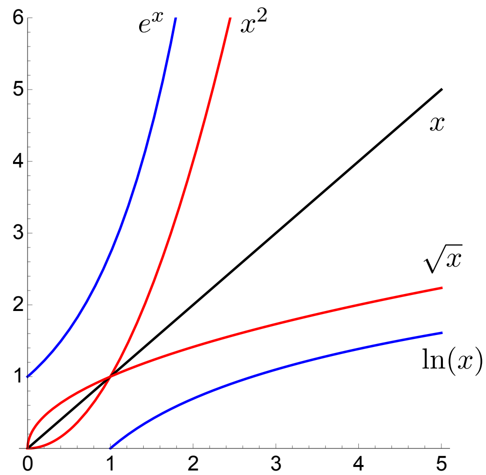


Figura 3.3: Gráficas de funciones con distintos órdenes de crecimiento.

3.6. Extremos locales y globales de una función

Las siguientes aplicaciones que veremos del cálculo diferencial se dirigen en primer lugar al estudio cualitativo de las funciones, con especial atención al crecimiento, decrecimiento, concavidad y convexidad. Pero los resultados que se incluyen en esta sección tienen otras aplicaciones.

Empezamos recordando los conceptos de máximo y mínimo local de una función.

Si I es un intervalo real y $f : I \to \mathbb{R}$ es una función, se dice que f alcanza un **máximo local** en un punto $x_0 \in I$ si existe un número $r > 0$ tal que $f(x) \leq f(x_0)$ para todos los puntos $x \in I$ tales que $x \in (x_0 - r, x_0 + r)$.

Análogamente, f alcanza un **mínimo local** en un punto $x_0 \in I$ si existe un número $r > 0$ tal que $f(x) \geq f(x_0)$ para todos los puntos $x \in I$ tales que $x \in (x_0 - r, x_0 + r)$.

Observaciones:

1. En cualquiera de los dos casos anteriores se dice que f tiene un **extremo local** en x_0.
2. El extremo es **estricto** si las desigualdades que relacionan $f(x)$ y $f(x_0)$ son estrictas.
3. El extremo es **global** si las desigualdades $f(x) \leq f(x_0)$ o $f(x) \geq f(x_0)$ se cumplen para todo $x \in I$.

En virtud del teorema de los valores extremos, si $f : [a, b] \to \mathbb{R}$ es una función continua entonces siempre se alcanzan el máximo y el mínimo globales, es decir, existen $x_1, x_2 \in [a, b]$ tales que $m = f(x_1) \leq f(x) \leq f(x_2) = M$, $\forall\, x \in [a, b]$.

El siguiente resultado facilita la búsqueda de los extremos locales de una función.

Teorema del extremo local: Si $f : I \to \mathbb{R}$ es una función derivable en un intervalo abierto I y f tiene un extremo local en un punto $x_0 \in I$ entonces $f'(x_0) = 0$.

Observación: Este teorema solo proporciona una condición necesaria para la existencia de extremos. El hecho de que $f'(x_0) = 0$ no garantiza que f tenga un extremo en x_0. Por ejemplo, la función $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = x^3$ no tiene ningún extremo local y sin embargo $f'(0) = 0$.

Extremos globales de una función en un intervalo compacto

Para calcular los extremos globales de una función continua f en un intervalo compacto $[a, b]$, seguiremos dos pasos:

1. Identificamos los posibles candidatos a extremos globales, que son:
   - Los puntos a y b de la frontera del intervalo.
   - Los puntos donde se anula la derivada.
   - Los puntos donde la función no es derivable.
2. Evaluamos f en cada uno de los posibles candidatos y elegimos el mínimo y el máximo.

Ejemplo: Calcular los extremos globales de la función $f(x) = |x^2 - 4|$ en el intervalo $[-1, 3]$. Los posibles extremos son:

- los puntos frontera -1 y 3;
- el punto 0, donde f es derivable y $f'(0) = 0$;
- el punto 2, donde f no es derivable.

Como $f(-1) = 3$, $f(0) = 4$, $f(2) = 0$ y $f(3) = 5$, el mínimo global es $f(2) = 0$ y el máximo global es $f(3) = 5$ (ver la figura 3.4).

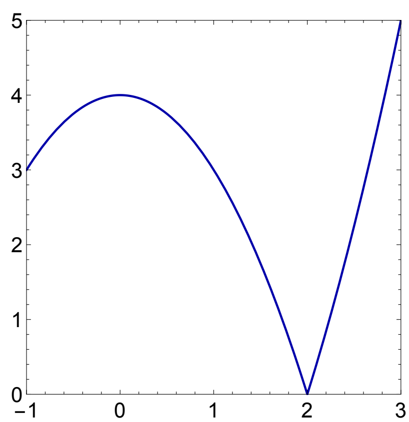


Figura 3.4: Gráfica de la función $f(x) = |x^2 - 4|$ en el intervalo $[-1, 3]$.

Una de las consecuencias del teorema del extremo local es el teorema de Rolle.

Teorema de Rolle: Si $f : [a, b] \to \mathbb{R}$ es una función derivable y $f(a) = f(b)$ entonces existe al menos un punto $x_0 \in [a, b]$ tal que $f'(x_0) = 0$.

La prueba es una simple consecuencia del teorema de los extremos globales y el teorema del extremo local. Si f es constante en $[a, b]$ entonces $f'(x) = 0$, $\forall\, x \in (a, b)$. En otro caso, el mínimo y el máximo globales de f en $[a, b]$ son distintos. Como $f(a) = f(b)$, necesariamente uno de ellos se alcanza en un punto $x_0 \in (a, b)$. El teorema del extremo local garantiza que $f'(x_0) = 0$.

Como consecuencia del teorema de Rolle se puede limitar el número de ceros de una función derivable f si conocemos el número de ceros de f'.

Si $f : I \to \mathbb{R}$ es una función derivable definida en un intervalo real I entonces entre dos ceros consecutivos de f existe al menos un cero de f'.

Como consecuencia de este resultado, si f' tiene n ceros en I entonces f no puede tener más de $(n+1)$ ceros en I. En particular, si f' no tiene ceros en I entonces f no puede tener más de uno.

Ejemplo. Veamos que la función $f(x) = 2e^x - x - 3$ tiene exactamente un cero en el intervalo $(0, 1)$. En primer lugar, como $f(0) = -1 < 0$ y $f(1) = 2e - 4 > 0$, el teorema de Bolzano permite afirmar que f tiene al menos un cero en el intervalo $(0, 1)$. Por otra parte,

$$f'(x) = 2e^x - 1 = 0 \iff e^x = 1/2 \iff x = \ln(1/2).$$

Como $\ln(1/2) < 0$, f' no tiene raíces en $(0, 1)$ y por tanto f tiene exactamente un cero en el intervalo $(0, 1)$.

3.7. Crecimiento, decrecimiento, convexidad y concavidad

Una de las consecuencias principales del teorema de Rolle es el teorema del valor medio.

Teorema del valor medio: Si $f : [a, b] \to \mathbb{R}$ es una función derivable entonces existe al menos un punto $x_0 \in (a, b)$ tal que $f(b) - f(a) = f'(x_0)(b - a)$.

Una de las interpretaciones del teorema del valor medio es que el error al aproximar el valor de $f(b)$ por el de $f(a)$ depende de la distancia entre a y b y de la tasa de cambio de f. Otra forma de leerlo es que la tasa media de cambio de f en el intervalo $[a, b]$ coincide con la tasa instantánea de cambio en algún punto del intervalo.

A continuación se obtienen algunas consecuencias del teorema del valor medio.

Intervalos de crecimiento y decrecimiento

Consideremos una función f derivable en un intervalo (a, b). Entonces se cumplen las siguientes propiedades:

(a) Si $f'(x) > 0$ para todo $x \in (a, b)$ entonces f es estrictamente creciente en (a, b).

(b) Si $f'(x) < 0$ para todo $x \in (a, b)$ entonces f es estrictamente decreciente en (a, b).

Como ejemplo del uso del teorema del valor medio probamos el apartado (a): consideramos dos puntos $x, y \in (a, b)$ tales que $x < y$. Tenemos que demostrar que $f(x) < f(y)$.

El teorema del valor medio garantiza la existencia de un punto $x_0 \in (x, y)$ tal que

$$f(y) - f(x) = f'(x_0)(y - x).$$

Entonces:

$$f'(x_0)(y - x) > 0 \Longrightarrow f(y) - f(x) > 0 \Longrightarrow f(x) < f(y).$$

Determinación de máximos y mínimos locales

Consideremos una función f definida en un intervalo abierto I. Supongamos que f es dos veces derivable en un entorno de un punto $x_0 \in I$ y $f'(x_0) = 0$.

(a) Si $f''(x_0) > 0$ entonces f alcanza un mínimo local en x_0.

(b) Si $f''(x_0) < 0$ entonces f alcanza un máximo local en x_0.

Como antes, probamos el primer apartado. Como $f''(x_0) > 0$, la función f'' toma valores positivos en un intervalo $(x_0 - r, x_0 + r)$ para algún $r > 0$. Por el resultado anterior, f' es estrictamente creciente en $(x_0 - r, x_0 + r)$. En consecuencia, $f'(x) > f'(x_0) = 0$ si $x > x_0$ y $f'(x) < f'(x_0) = 0$ si $x < x_0$. Por tanto, f es estrictamente decreciente en $(x_0 - r, x_0)$ y estrictamente creciente en $(x_0, x_0 + r)$. Esto quiere decir que f alcanza en x_0 un mínimo local.

Los resultados anteriores se pueden aplicar para determinar los intervalos de concavidad y convexidad de la gráfica de una función.

Recordamos estos conceptos. Si $f : I \to \mathbb{R}$ es una función derivable en un intervalo I entonces se puede definir la recta tangente a la gráfica de f en cada punto $(x_0, f(x_0))$, $x_0 \in I$. Se dice que la función es **convexa** en I si la gráfica de f queda por encima de cualquier recta tangente a dicha gráfica en los puntos del intervalo I. Si la gráfica de f queda por debajo de cualquier recta tangente a dicha gráfica en los puntos del intervalo I se dirá que f es **cóncava** en I. Si f es convexa a un lado de x_0 y cóncava a otro se dirá que f tiene en x_0 un **punto de inflexión**.

Concavidad y convexidad: Consideremos una función f dos veces derivable en un intervalo abierto I.

(a) Si $f''(x) > 0$, $\forall x \in I$ entonces f es convexa en I.

(b) Si $f''(x) < 0$, $\forall x \in I$ entonces f es cóncava en I.

(c) Si f tiene en x_0 un punto de inflexión entonces necesariamente $f''(x_0) = 0$.

Obsérvese que las funciones lineales $f(x) = ax + b$ tienen derivada segunda cero. Se pueden interpretar las funciones convexas como aquellas en las que la recta tangente predice un crecimiento menor que el real (como x^2 o e^x) y las cóncavas como aquellas en las que la recta tangente predice un crecimiento mayor (como $\sqrt{x}$ o $\ln(x)$). Se ilustran en la figura 3.5.

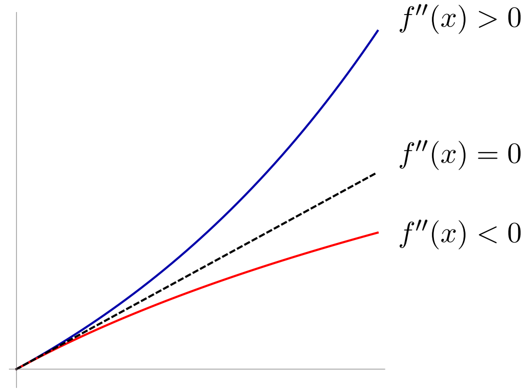


Figura 3.5: Gráficas de una función lineal (trazo discontinuo), una convexa (en azul) y una cóncava (en rojo).

3.8. Polinomios de Taylor

Consideremos una función f definida en un intervalo real I. El teorema de Taylor establece la forma de aproximar la función f por un polinomio en un entorno de un punto $x_0 \in I$. El caso más sencillo consiste en aproximar por la recta tangente a la gráfica de f en $(x_0, f(x_0))$, es decir, por el polinomio de grado uno $p_1(x) = f(x_0) + f'(x_0)(x - x_0)$. Obsérvese que $p_1(x)$ es el único polinomio de grado 1 que cumple las relaciones $p_1(x_0) = f(x_0)$, $p_1'(x_0) = f'(x_0)$ (f y p_1 coinciden en x_0 en posición y velocidad).

La recta tangente es una "primera aproximación"de la función f en un entorno del punto x_0. Cabe esperar que podamos mejorar esta aproximación si imponemos condiciones adicionales al polinomio (y por tanto incrementamos su grado).

Para una función n veces derivable en I se define el **polinomio de Taylor** de grado n de f centrado en x_0 como el único polinomio $p_n(x)$ de grado menor o igual que n que cumple las $n+1$ ecuaciones $p_n(x_0) = f(x_0)$, $p_n'(x_0) = f'(x_0)$, $\ldots$, $p_n^{n)}(x_0) = f^{n)}(x_0)$. Su expresión es la siguiente:

$$\boxed{p_n(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{n)}(x_0)}{n!}(x - x_0)^n.}$$

En forma abreviada:

$$p_n(x) = f(x_0) + \sum_{k=1}^{n} \frac{f^{k)}(x_0)}{k!}(x - x_0)^k.$$

Por ejemplo, si $f(x) = e^x$ y $x_0 = 0$, el polinomio de Taylor de grado 3 de f centrado en cero es:

$$p_3(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}\,x^2 + \frac{f'''(0)}{3!}\,x^3 = 1 + x + \frac{x^2}{2} + \frac{x^3}{6},$$

ya que $f(0) = f'(0) = f''(0) = f'''(0) = e^0 = 1$.

Usando este polinomio, podemos aproximar e^x en puntos próximos a cero. Por ejemplo, $\sqrt{e} = e^{1/2} \approx p_3(1/2) = 1+1/2+1/8+1/48 = 1.64583$. La calculadora proporciona $\sqrt{e} = 1.64872$, con lo que las dos primeras cifras decimales son correctas. En la figura 3.6 se muestran en el intervalo $[0, 1.5]$ las aproximaciones de Taylor de grados 1, 2 y 3 de e^x centradas en 0.

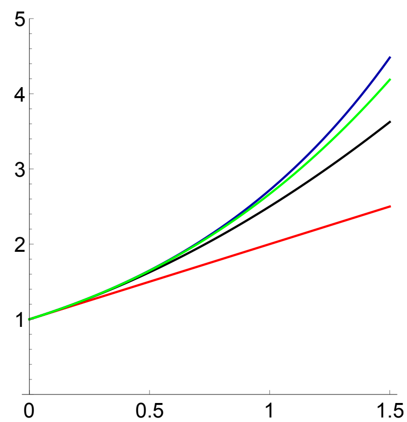


Figura 3.6: Gráfica de la función $f(x) = e^x$ (en azul) en el intervalo $[0, 1.5]$, junto con sus aproximaciones mediante los polinomios de Taylor centrados en cero y de grados 1 (en rojo), 2 (en negro) y 3 (en verde).

El teorema de Taylor permite estimar el error cometido cuando aproximamos una función por el polinomio de Taylor.

Teorema de Taylor: Si $f : [a, b] \to \mathbb{R}$ es una función derivable $n + 1$ veces y $x_0 \in [a, b]$ entonces, para cada $x \in [a, b]$, existe un número c entre x_0 y x tal que

$$f(x) = p_n(x) + r_n(x), \text{ con } \quad r_n(x) = \frac{f^{n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}.$$

El término $r_n(x)$ se llama **resto del polinomio de Taylor** de grado n de f centrado en x_0 y proporciona el error cometido en la aproximación, ya que $|f(x) - p_n(x)| = |r_n(x)|$, $\forall\, x \in I$.

Ejemplo: El polinomio de Taylor de grado tres de $f(x) = \operatorname{sen}(x)$ centrado en $x_0 = 0$ es $p_3(x) = x - x^3/6$. El error cometido al aproximar $\operatorname{sen}(1/2)$ por $p_3(1/2)$ es

$$|r_3(1/2)| = \left| \frac{f^{iv)}(c)}{4!} \left(\frac{1}{2}\right)^4 \right|, \; c \in (0, 1/2).$$

Dado que $|f^{iv)}(c)| = |\operatorname{sen}(c)| \leq 1$, $\forall\, c \in (0, 1/2)$, se tiene que:

$$|\operatorname{sen}(1/2) - p_3(1/2)| = |r_3(1/2)| \leq \frac{1}{4!\, 2^4} = 0.00260417.$$

De hecho $p_3(1/2) = 1/2 - 1/48 = 0.479167$ y la calculadora proporciona $\operatorname{sen}(1/2) = 0.479426$.

Finalizamos el tema con dos aplicaciones del teorema de Taylor para obtener condiciones precisas para la existencia de extremos locales y puntos de inflexión de una función f suficientemente regular (con suficientes derivadas sucesivas).

Criterio para la existencia de extremos: Consideremos una función $f : I \to \mathbb{R}$ de clase $\mathcal{C}^n$ en un intervalo abierto I. Si $x_0 \in I$ cumple que $f'(x_0) = 0$ y n es el orden de la primera derivada de f que no se anula en x_0, es decir $f^{k)}(x_0) = 0$ si $1 \leq k < n$ y $f^{n)}(x_0) \neq 0$, entonces:

(a) Si n es par y $f^{n)}(x_0) > 0$, f alcanza un mínimo local en x_0.

(b) Si n es par y $f^{n)}(x_0) < 0$, f alcanza un máximo local en x_0.

(c) Si n es impar, f tiene en x_0 un punto de inflexión.

Criterio para la existencia de puntos de inflexión: Consideremos una función $f : I \to \mathbb{R}$ de clase n en un intervalo abierto I. Si $x_0 \in I$ cumple que $f''(x_0) = 0$ y n es el orden de la primera derivada de f mayor que dos que no se anula en x_0, es decir $f^{k)}(x_0) = 0$ si $2 \leq k < n$ y $f^{n)}(x_0) \neq 0$, entonces f tiene un punto de inflexión en x_0 si y solo si n es impar.

Ejemplo: Consideremos la función $f(x) = x^4 - x^3$.

$$f'(x) = 4x^3 - 3x^2 = 0 \Longleftrightarrow x = 0 \text{ ó } x = 3/4$$
$$f''(x) = 12x^2 - 6x = 0 \Longleftrightarrow x = 0 \text{ ó } x = 1/2.$$

Para $x = 0$ se tiene que

$$f'(0) = f''(0) = 0,\ f'''(0) = -6 \neq 0.$$

Por tanto, f tiene en 0 un punto de inflexión.
Para $x = 3/4$ se tiene que

$$f'(3/4) = 0,\ f''(3/4) = 9/4 > 0.$$

Por tanto, f tiene en 3/4 un mínimo local.
Para $x = 1/2$ se tiene que

$$f''(1/2) = 0,\ f'''(1/2) = 6 \neq 0.$$

Por tanto, f tiene en 1/2 un punto de inflexión.

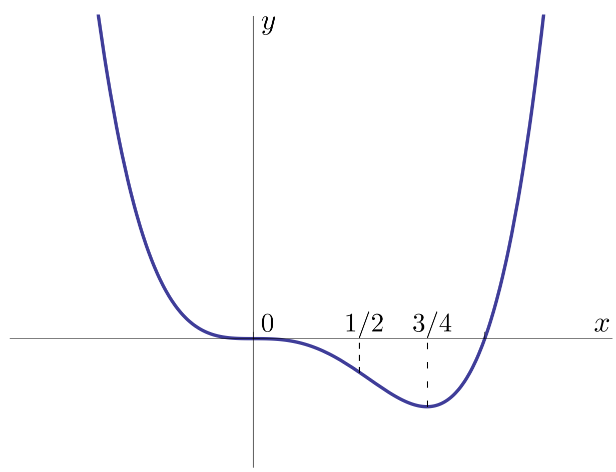


Figura 3.7: Gráfica de $f(x) = x^4 - x^3$.

Capítulo 4

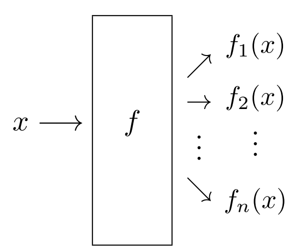
Introducción a las funciones vectoriales

4.1. Funciones vectoriales de una variable. Curvas y vectores tangente

Consideremos un intervalo real I. Una función $f : I \to \mathbb{R}^n$ se llama función vectorial de una variable. Para cada $x \in I$, $f(x)$ es un vector de $\mathbb{R}^n$ y por tanto se puede escribir en la forma

$$f(x) = (f_1(x), f_2(x), \ldots, f_n(x)),$$

donde $f_i : I \to \mathbb{R}$ es una función escalar para cada $i = 1, 2, \ldots, n$. Las funciones f_i se llaman funciones componentes de f. Informalmente, una función vectorial de una variable real se puede entender como una "máquina"que transforma una variable de entrada $x \in \mathbb{R}$ en n variables de salida $f_1(x)$, $f_2(x)$, $\ldots$, $f_n(x)$.



Por ejemplo, la función $f : [0, 2\pi] \to \mathbb{R}^2$ dada por $f(x) = (\cos(x), \operatorname{sen}(x))$ es una función vectorial con valores en $\mathbb{R}^2$. Sus componentes son $f_1(x) = \cos(x)$ y $f_2(x) = \operatorname{sen}(x)$.

Las funciones componentes permiten extender fácilmente los conceptos de límite, continuidad y derivada. Si $x_0 \in \mathbb{R}$ y f está definida en los intervalos $(x_0 - r, x_0)$ y $(x_0, x_0 + r)$ para algún $r > 0$, diremos que existe el límite de f cuando x tiende a x_0 si existe el límite de cada una de las componentes. En este caso, se define

$$\lim_{x \to x_0} f(x) = \left(\lim_{x \to x_0} f_1(x), \lim_{x \to x_0} f_2(x), \ldots, \lim_{x \to x_0} f_n(x) \right).$$

Una función f es continua en $x_0 \in I$ si todas sus componentes son continuas en x_0.

Se dice que $f : I \to \mathbb{R}^n$ es derivable en un punto x_0 si lo es cada una de sus componentes; en ese caso, la derivada de f en x_0 es

$$f'(x_0) = \left(f_1'(x_0), f_2'(x_0), \ldots, f_n'(x_0)\right).$$

Por ejemplo, la derivada de $f(x) = (\cos(x), \operatorname{sen}(x))$ es $f'(x) = (-\operatorname{sen}(x), \cos(x))$.

Curvas y vector tangente

Si $f : I \to \mathbb{R}^n$ es una función continua definida en un intervalo real I, con componentes $f(x) = (f_1(x), f_2(x), \ldots, f_n(x))$, entonces el conjunto

$$C = \{(f_1(x), f_2(x), \ldots, f_n(x)) \,/\, x \in I\}$$

se llama curva en $\mathbb{R}^n$ y se dice que la función f es una parametrización de la curva.

Ejemplo: La función $f : [0, 2\pi] \to \mathbb{R}^2$ definida por $f(x) = (\operatorname{sen}(x), \operatorname{sen}(2x))$ describe la lemniscata de Bernoulli, es decir, la lemniscata está definida como:

$$\mathcal{L} = \{(\operatorname{sen}(x), \operatorname{sen}(2x) \,/\, x \in [0, 2\pi]\}.$$

Su representación gráfica se puede ver en la figura 4.1.

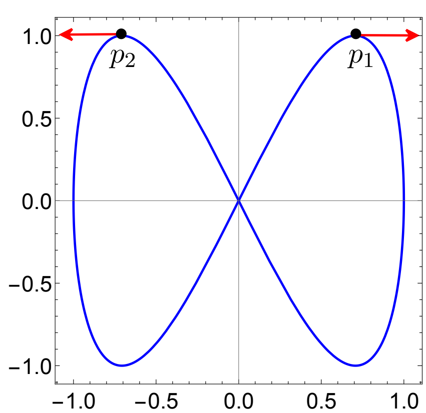


Figura 4.1: Lemniscata de Bernoulli junto con sus vectores tangente en los puntos p_1 y p_2.

Si $f : I \to \mathbb{R}^n$ es derivable en x_0 y $f'(x_0) \neq (0, 0, \ldots, 0)$ entonces $f'(x_0)$ es un vector tangente a la curva C en el punto $f(x_0)$. El vector tangente indica en qué sentido se recorre la curva (su orientación). Se suele representar con origen en el punto $f(x_0)$.

Por ejemplo, la lemniscata pasa por los puntos

$$p_1 = f(\pi/4) = (\mathrm{sen}(\pi/4), \mathrm{sen}(\pi/2)) = (1/\sqrt{2}, 1);$$
$$p_2 = f(5\pi/4) = (\mathrm{sen}(5\pi/4), \mathrm{sen}(5\pi/2)) = (-1/\sqrt{2}, 1).$$

Como $f'(x) = (\cos(x), 2\cos(2x))$, el vector tangente en p_1 es $f'(\pi/4) = (1/\sqrt{2}, 0)$, que apunta hacia la derecha, y el vector tangente en p_2 es $f'(5\pi/4) = (-1/\sqrt{2}, 0)$, que apunta hacia la izquierda. Se representan en rojo en la figura 4.1.

Cuando $n = 2$ o $n = 3$, una curva se puede pensar como la trayectoria que sigue una partícula en el plano o en el espacio. Si denotamos por $r(t) = (x(t), y(t), z(t))$ el vector de posición en un instante t, entonces el vector tangente $v(t) = r'(t) = (x'(t), y'(t), z'(t))$ es el vector velocidad y $a(t) = v'(t) = r''(t) = (x''(t), y''(t), z''(t))$ es el vector aceleración.

4.2. Campos escalares y vectoriales. Curvas de nivel

Si D es un subconjunto de $\mathbb{R}^n$, con $n > 1$, una función $f : D \to \mathbb{R}^p$ se llama función de varias variables reales. El conjunto D se llama dominio de definición de f.

En el caso particular $p = 1$, una función $f : D \subset \mathbb{R}^n \to \mathbb{R}$ se suele llamar campo escalar. Por ejemplo, la función que asigna a cada punto (x, y, z) de un recinto tridimensional D su temperatura es un campo escalar $T : D \subset \mathbb{R}^3 \to \mathbb{R}$ llamado campo de temperaturas.

Informalmente, un campo escalar en $\mathbb{R}^n$ se puede entender como una "máquina" que transforma una variable de entrada $P = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ en una variable de salida $f(P) \in \mathbb{R}$.

$$\begin{array}{ll} x_1 & \nearrow \\ x_2 & \rightarrow \\ \vdots & \vdots \\ x_n & \searrow \end{array} \boxed{\quad f \quad} \longrightarrow f(x_1, x_2, \dots, x_n)$$

Si $p > 1$, una función $f : D \subset \mathbb{R}^n \to \mathbb{R}^p$ se llama función vectorial de varias variables. Especialmente en el caso $p = n$, la función vectorial $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ suele llamarse campo de vectores.

Una función $f : D \subset \mathbb{R}^n \to \mathbb{R}^p$ se puede expresar en función de sus componentes:

$$f(x) = f(x_1, x_2, \dots, x_n) = (f_1(x_1, x_2, \dots, x_n), f_2(x_1, x_2, \dots, x_n), \dots, f_p(x_1, x_2, \dots, x_n)).$$

Para cada $k = 1, 2, \dots, p$, la componente $f_k : \mathbb{R}^n \to \mathbb{R}$ es un campo escalar.

Por ejemplo, la función $f : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $f(x, y) = (x + y, e^{xy}, \mathrm{sen}(x - y))$ tiene tres componentes $f_1(x, y) = x + y$, $f_2(x, y) = e^{xy}$, $f_3(x, y) = \mathrm{sen}(x - y)$.

La gráfica de una función $f : D \subset \mathbb{R}^n \to \mathbb{R}^p$ se define como el conjunto

$$G(f) = \{(x, f(x)) \,/\, x \in D\}.$$

Nótese que los puntos de la gráfica son de la forma

$$(x, f(x)) = (x_1, x_2, \dots, x_n, f_1(x_1, x_2, \dots, x_n), f_2(x_1, x_2, \dots, x_n), \dots, f_p(x_1, x_2, \dots, x_n)) \in \mathbb{R}^{n+p},$$

y por tanto la gráfica de f es un subconjunto de $\mathbb{R}^{n+p}$.

Por ejemplo, la gráfica de la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x, y) = x^2 + y^2$ es el paraboloide definido por la ecuación $z = x^2 + y^2$ en $\mathbb{R}^3$ (representado en la figura 4.2).

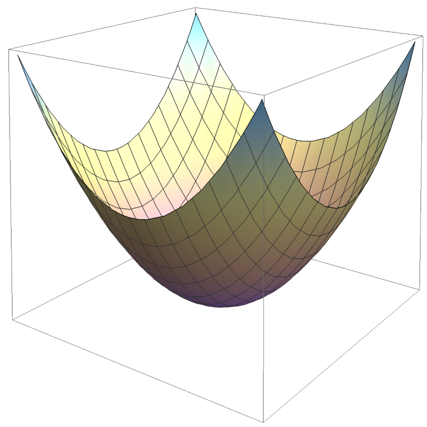


Figura 4.2: Gráfica del campo escalar $f(x, y) = x^2 + y^2$.

En general, las gráficas de funciones de $\mathbb{R}^2$ en $\mathbb{R}$ son superficies en $\mathbb{R}^3$ difíciles de representar sin ayuda de software adecuado. Por ejemplo, la gráfica de la función $f(x, y) = \operatorname{sen}(x) + \cos(y)$ tiene la forma de huevera representada en la figura 4.3.

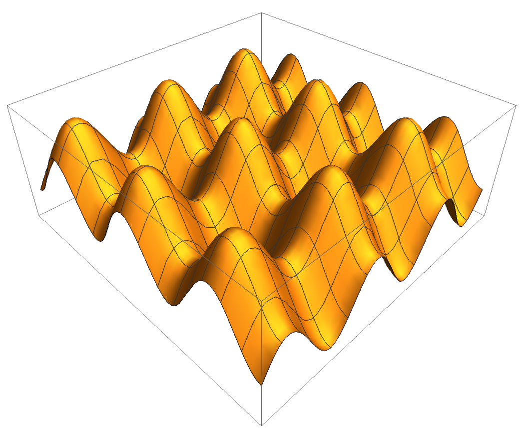


Figura 4.3: Gráfica del campo escalar $f(x, y) = \operatorname{sen}(x) + \cos(y)$.

Para dimensiones mayores, la gráfica de una función de varias variables $f : \mathbb{R}^n \to \mathbb{R}^p$ estaría en un espacio de dimensión mayor que 3 y por tanto no se podría representar.

Curvas de nivel

Si $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ es un campo escalar definido en un dominio D del plano, se definen sus curvas de nivel como los conjuntos de puntos sobre los que f toma el mismo valor, es decir, si $K \in \mathbb{R}$, la curva de nivel K es

$$C_K = \{(x, y) \in \mathbb{R}^2 \,/\, f(x, y) = K\}.$$

Ejemplos típicos de curvas de nivel son los mapas topográficos, donde cada curva de nivel se corresponde con los puntos que tienen la misma altura, y los mapas de presión, donde cada curva de nivel (isobara) representa los puntos donde la presión atmosférica es la misma.

Por ejemplo, las curvas de nivel de la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x, y) = x^2 + y^2$ solo tienen sentido para $K = r^2 > 0$ y son circunferencias de centro $(0, 0)$ y radio r:

$$C_K = \{(x, y) \in \mathbb{R}^2 \,/\, x^2 + y^2 = K = r^2\}.$$

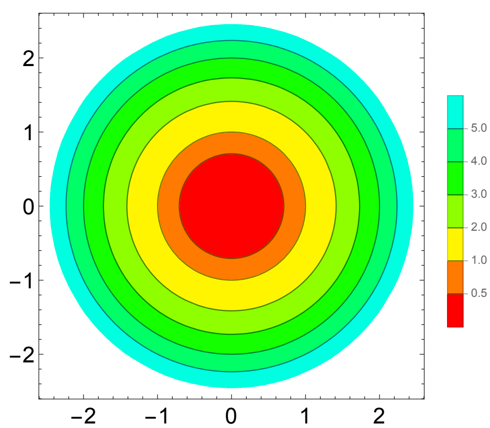


Figura 4.4: Curvas de nivel del campo escalar $f(x, y) = x^2 + y^2$.

4.3. Nociones básicas de topología en $\mathbb{R}^n$

En esta sección generalizaremos a $\mathbb{R}^n$ algunos conceptos que en $\mathbb{R}$ proporcionan de forma natural los intervalos y el valor absoluto. El papel de un intervalo real $(x_0 - r, x_0 + r)$ centrado en $x_0 \in \mathbb{R}$ y radio $r > 0$ lo juega en $\mathbb{R}^n$ la bola abierta de centro $P \in \mathbb{R}^n$ y radio $r > 0$, que se define como el conjunto de puntos de $\mathbb{R}^n$ cuya distancia a P es menor que r, es decir,

$$B(P, r) = \{x \in \mathbb{R}^n \,/\, \|x - P\| < r\},$$

donde $\|(x_1, x_2, \ldots, x_n)\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$.

Por ejemplo, en $\mathbb{R}^2$ la bola abierta de centro $(1, 1)$ y radio r está formada por los puntos que quedan dentro de la circunferencia de centro $(1, 1)$ y radio r, es decir:

$$B((1, 1), r) = \left\{(x, y) \in \mathbb{R}^2 \,/\, (x - 1)^2 + (y - 1)^2 < r^2\right\}.$$

Las bolas abiertas permiten definir el interior y la frontera de un subconjunto A de $\mathbb{R}^n$:

- Se dice que $P \in A$ es un punto interior de A si existe alguna bola abierta centrada en P que queda dentro de A. El conjunto de los puntos interiores de A se llama interior de A.
- Diremos que un conjunto A es abierto si todos sus puntos son interiores.
- Diremos que un punto P está en la frontera de A si cualquier bola abierta centrada en P contiene puntos de A y puntos que no están en A.
- Diremos que un conjunto A es cerrado si contiene a todos los puntos de su frontera.
- Si A es cerrado, distinguiremos los puntos interiores de A de los puntos de la frontera de A. El conjunto de puntos de la frontera de A lo denotaremos por $\mathrm{Fr}(A)$.
- Diremos que A está acotado si existe una constante positiva K tal que $\|x\| \leq K$, $\forall\, x \in A$.
- Diremos que A es compacto si está acotado y es cerrado. Los conjuntos compactos de $\mathbb{R}^n$ juegan el papel de los intervalos compactos $[a, b]$ de $\mathbb{R}$.

Por ejemplo, los conjuntos

$$D_1 = \{(x, y) \in \mathbb{R}^2 \,/\, x^2 + y^2 < 1\} \quad , \quad D_2 = \{(x, y) \in \mathbb{R}^2 \,/\, x \geq 0\}$$

no son compactos: D_1 es un círculo, que está acotado pero no es cerrado, mientras que D_2 es un semiplano, que es cerrado pero no está acotado.

En general, los conjuntos cerrados que manejaremos están definidos por desigualdades no estrictas y los conjuntos abiertos están definidos por desigualdades estrictas (obsérvese la analogía con los intervalos abiertos y cerrados).

En la figura 4.5 se representa el conjunto compacto

$$A = \{(x, y) \in \mathbb{R}^2 \,/\, 9 \leq x^2+y^2 \leq 25\,,\ x+y \leq 1\}.$$

El interior de A está coloreado en verde. La frontera de A (en trazo azul) está formada por 2 arcos de circunferencia y 2 segmentos de recta.

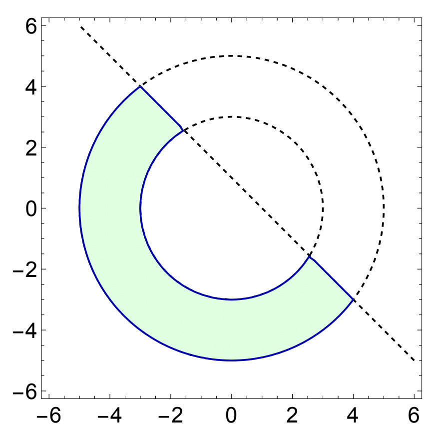


Figura 4.5: Representación del conjunto A.

Capítulo 5

Continuidad y cálculo diferencial de funciones de varias variables

5.1. Límites y continuidad de funciones de varias variables

Consideremos una función vectorial $f : D \subset \mathbb{R}^n \to \mathbb{R}^p$ definida en un subconjunto abierto D de $\mathbb{R}^n$. Si P es un punto de D o de su frontera, diremos que $L \in \mathbb{R}^p$ es el límite de f cuando x tiende a P si $f(x)$ se aproxima a L cuando x se aproxima a P. Esta definición se puede escribir en términos de normas y es muy similar a la de límite de funciones escalares de una variable:

$$\lim_{x \to P} f(x) = L \Longleftrightarrow \left[\forall \varepsilon > 0,\, \exists \delta > 0 \,/\, \|x - P\| < \delta, x \neq P \Longrightarrow \|f(x) - L\| < \varepsilon\right].$$

Al igual que pasaba con las funciones vectoriales de una variable, si $f : D \subset \mathbb{R}^n \to \mathbb{R}^p$ y $P \in \mathbb{R}^n$ entonces existe el límite de f en P si y solo si existen los límites en P de cada una de las componentes $f_1(x), f_2(x), \ldots, f_p(x)$. Por tanto, podemos restringir el estudio a campos escalares.

Continuidad

El concepto de continuidad para funciones de varias variables es similar al de funciones de una variable. Si $f : D \subset \mathbb{R}^n \to \mathbb{R}^p$ es una función definida en un conjunto D y $P \in D$, se dice que f es continua en P si existe el límite de f en P y además

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

Diremos que f es continua en D si es continua en todos los puntos de D.

Es claro que $f(x) = (f_1(x), f_2(x), \ldots, f_p(x))$ es continua en P si y solo si todas sus componentes son continuas en P. Por ello podemos centrarnos de nuevo en campos escalares.

Las propiedades relacionadas con la continuidad para funciones de varias variables son similares a las de las funciones de una variable:

1. La composición de funciones continuas es una función continua.

2. Si $f : \mathbb{R}^n \to \mathbb{R}$ y $g : \mathbb{R}^n \to \mathbb{R}$ son funciones continuas en P entonces las funciones $(f+g)$ y $(f \cdot g)$ son continuas en P. La función f/g es continua en P si $g(P) \neq 0$.

Por ejemplo, la función $f(x,y) = \dfrac{e^{x+y}}{xy-1}$ es continua para todos los puntos de $\mathbb{R}^2$ que no están sobre la hipérbola $xy = 1$.

Los puntos en los que la función no es continua dan lugar a discontinuidades. Un ejemplo típico de discontinuidad es aquel en el que el límite es infinito. Por ejemplo, el campo escalar $f(x,y) = -1/(x^2+y^2)$ tiene una discontinuidad en $(0,0)$ ya que

$$\lim_{(x,y)\to(0,0)} f(x,y) = -\infty.$$

Se muestra su aspecto en la figura 5.1. En términos físicos, se relaciona este tipo de discontinuidades con singularidades y agujeros negros.

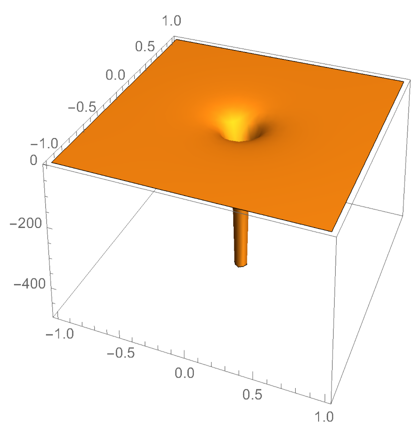


Figura 5.1: Gráfica de $f(x,y) = -1/(x^2+y^2)$ junto con una imagen típica de singularidad usada en astrofísica.

Límites direccionales

Una de las diferencias fundamentales es que el límite de una función de una variable en un punto $x_0 \in \mathbb{R}$ existe si y solo si existen los límites laterales y coinciden. Sin embargo, las posibles formas de aproximarse a un punto en $\mathbb{R}^n$ son infinitas.

Esta observación resulta útil como criterio para probar que el límite en P no existe y por tanto hay una discontinuidad: basta encontrar dos direcciones de tal forma que los límites de $f(x)$ cuando x tiende a P a lo largo de esas direcciones sean distintos. Estos límites se llaman límites direccionales.

Ejemplos: El límite

$$\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2}$$

no existe porque cuando nos aproximamos a $(0,0)$ a lo largo de los dos ejes de coordenadas la función se aproxima a valores diferentes:

$$\lim_{\substack{(x,y)\to(0,0)\\ y=0}} \frac{x^2-y^2}{x^2+y^2} = \lim_{x\to 0} \frac{x^2}{x^2} = \lim_{x\to 0} 1 = 1.$$

$$\lim_{\substack{(x,y)\to(0,0)\\ x=0}} \frac{x^2-y^2}{x^2+y^2} = \lim_{y\to 0} \frac{-y^2}{y^2} = \lim_{y\to 0}(-1) = -1.$$

En el siguiente ejemplo estos dos límites coinciden y consideramos las rectas $y=\lambda x$ que pasan por $(0,0)$. El límite

$$\lim_{(x,y)\to(0,0)} \frac{2xy}{x^2+y^2}$$

no existe porque cuando nos aproximamos a $(0,0)$ a lo largo de las rectas $y=\lambda x$ con distintos valores de λ la función se aproxima a valores diferentes:

$$\lim_{\substack{(x,y)\to(0,0)\\ y=\lambda x}} \frac{2xy}{x^2+y^2} = \lim_{x\to 0} \frac{2\lambda x^2}{(1+\lambda^2)x^2} = \frac{2\lambda}{1+\lambda^2}.$$

Para $\lambda=1$, el límite vale 1 mientras que para $\lambda=-1$ el límite vale -1. Se puede visualizar esta discontinuidad en la figura 5.2.

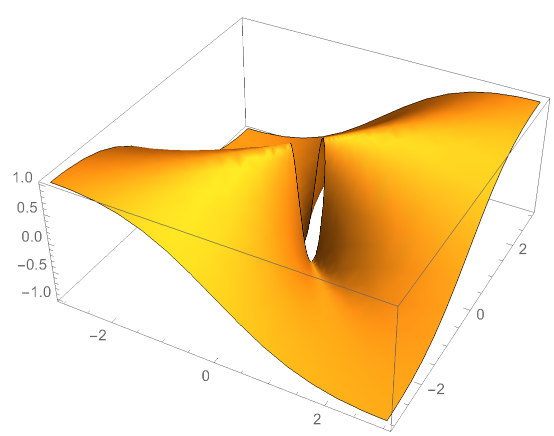


Figura 5.2: Gráfica de $f(x,y)=2xy/(x^2+y^2)$ ilustrando dos límites direccionales diferentes en $(0,0)$. El límite es 1 a través de la recta $y=x$ y es -1 a través de la recta $y=-x$.

El ejemplo de la figura 5.2 ilustra una discontinuidad en $(0,0)$ de naturaleza diferente a la de la figura 5.1. En ese caso, los límites direccionales son finitos y la función se mantiene acotada (de hecho $f(x,y) = 2xy/(x^2+y^2)$ solo toma valores entre -1 y 1).

5.2. Derivadas parciales y plano tangente.

Para definir el concepto de diferenciabilidad de funciones de varias variables, comenzaremos con campos escalares en $\mathbb{R}^2$.

Si $f : \mathbb{R}^2 \to \mathbb{R}$ es un campo escalar, se llama **derivada parcial** de $f(x,y)$ respecto a la variable x a la función que resulta de considerar la variable y constante y aplicar la derivación usual respecto de x. La representaremos por $\partial f/\partial x$. La derivada parcial respecto de y se define de modo análogo.

Por ejemplo, si $f(x,y) = x^3y + xy^2 - \ln(x+2y)$ entonces:

$$\frac{\partial f}{\partial x} = 3x^2y + y^2 - \frac{1}{x+2y} \quad ; \quad \frac{\partial f}{\partial y} = x^3 + 2xy - \frac{2}{x+2y}\,.$$

Formalmente, la derivada parcial de f respecto de x en un punto (x,y) se define como

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

La derivada parcial de f respecto de x en el punto (x_0,y_0) representa la tasa de cambio instantánea de f en la dirección paralela al eje x, cuando y_0 se mantiene constante. Por ejemplo, si la función temperatura $T(t,x)$ depende del tiempo t y del desplazamiento en la dirección de x, entonces la derivada parcial de T respecto de t en el punto (t_0,x_0) representa la tasa de cambio de la temperatura con el tiempo manteniendo constante la posición.

La interpretación geométrica es la siguiente: denotemos por $z = f(x,y)$ la gráfica de f; entonces la derivada parcial respecto de x en un punto (x_0,y_0) representa la pendiente de la recta tangente a la curva $z = f(x,y_0)$ en el punto $(x_0,y_0,f(x_0,y_0))$, es decir, la pendiente de la tangente a la curva intersección de la gráfica de f con el plano $y = y_0$.

De modo análogo, la parcial respecto de y en el punto (x_0,y_0) representa la pendiente de la tangente a la curva intersección de la gráfica de f con el plano $x = x_0$ (ver la figura 5.3).

Si existen las derivadas parciales respecto de x y de y en un punto (x_0,y_0) entonces el vector fila

$$\nabla f(x_0,y_0) = \left(\frac{\partial f}{\partial x}(x_0,y_0), \frac{\partial f}{\partial y}(x_0,y_0)\right) \in \mathbb{R}^2$$

se llama **vector gradiente** de f en (x_0,y_0).

Por ejemplo, si $f(x,y) = x^3y + xy^2 - \ln(x+2y)$ entonces

$$\nabla f(x,y) = \left(3x^2y + y^2 - \frac{1}{x+2y}\,,\; x^3 + 2xy - \frac{2}{x+2y}\right) \Longrightarrow \nabla f(-1,1) = (3,-5).$$

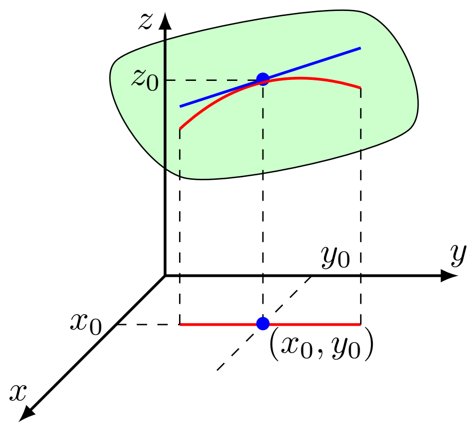


Figura 5.3: Interpretación geométrica de la derivada parcial de f con respecto a y en el punto (x_0, y_0). La región en verde representa la gráfica de f. La recta azul es la tangente a la curva intersección de la gráfica de f con el plano $x = x_0$ (curva en trazo rojo) y la pendiente de esa recta es la derivada parcial.

Plano tangente

La interpretación geométrica de las derivadas parciales motiva la idea de definir el plano tangente a la superficie $z = f(x, y)$ definida por la gráfica de un campo escalar $f : \mathbb{R}^2 \to \mathbb{R}$ en un punto (x_0, y_0, z_0) como el plano que pasa por (x_0, y_0, z_0) y contiene a las rectas tangentes definidas por $(\partial f/\partial x)(x_0, y_0)$ y $(\partial f/\partial y)(x_0, y_0)$ respectivamente.

Recordemos que la ecuación de la recta tangente a la gráfica de una una función $f : \mathbb{R} \to \mathbb{R}$ en un punto $(x_0, f(x_0))$ es $y - f(x_0) = f'(x_0)(x - x_0)$. La ecuación del plano tangente es similar:

$$\boxed{z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)\,(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)\,(y - y_0) = \nabla f(x_0, y_0) \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix},}$$

donde $z_0 = f(x_0, y_0)$.

Nótese que si hacemos $y = y_0$ entonces la intersección del plano tangente con el plano $y = y_0$ es precisamente la recta tangente a la intersección de la gráfica de f con dicho plano.

Ejemplo: Calculamos la ecuación del plano tangente a la gráfica de $f(x, y) = x^2 + y^2 + y$ en el punto $(1, 0, 1)$. Como $\nabla f(x, y) = (2x, 2y + 1)$, en particular $\nabla f(1, 0) = (2, 1)$ y la ecuación del plano tangente es

$$z - 1 = (2, 1) \begin{pmatrix} x - 1 \\ y - 0 \end{pmatrix} = 2(x - 1) + y \Longleftrightarrow 2x + y - z = 1.$$

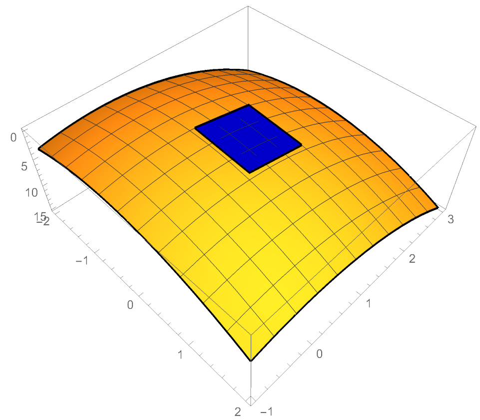


Figura 5.4: Ilustración de la gráfica de $f(x, y) = x^2 + y^2 + y$ y su plano tangente en el punto $(1, 0, 1)$.

5.3. Diferenciabilidad

El plano tangente permite extender el concepto de función diferenciable en un punto a un campo escalar en $\mathbb{R}^2$. Informalmente, un campo escalar continuo $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ definido en un conjunto abierto D es diferenciable en un punto $(x_0, y_0) \in D$ si el plano tangente en el punto $(x_0, y_0, f(x_0, y_0))$ es una buena aproximación de la gráfica de f. Formalmente, diremos que f es diferenciable si existen las derivadas parciales de f en (x_0, y_0) y además:

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x, y) - f(x_0, y_0) - \nabla f(x_0, y_0) \cdot (x - x_0, y - y_0)}{\|(x, y) - (x_0, y_0)\|} = 0.$$

Si ocurre esto, se define la **derivada** de f en (x_0, y_0) como el vector gradiente:

$$Df(x_0, y_0) = \nabla f(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right).$$

Una condición suficiente de diferenciabilidad más fácil de comprobar es la siguiente:

Condición suficiente de diferenciabilidad: Consideremos un subconjunto abierto D de $\mathbb{R}^2$ y un campo escalar continuo $f : D \to \mathbb{R}$. Si existen las derivadas parciales de f y son continuas en todo punto de D, entonces f es diferenciable en todos los puntos de D.

Los conceptos de vector gradiente y diferenciabilidad se generalizan a un campo escalar definido en un subconjunto abierto D de $\mathbb{R}^n$: si existen las derivadas parciales de f y son continuas en todo punto de D, entonces f es diferenciable en cada punto $P \in D$ y se define la derivada de f en P como el vector gradiente

$$Df(P) = \nabla f(P) = \left(\frac{\partial f}{\partial x_1}(P), \frac{\partial f}{\partial x_2}(P), \ldots, \frac{\partial f}{\partial x_n}(P) \right).$$

Ejemplo: Si $f : \mathbb{R}^3 \to \mathbb{R}$ está definida por $f(x, y, z) = xyz - \operatorname{sen}(xy) + 3z$ entonces el gradiente de f en (x, y, z) es

$$\nabla f(x, y, z) = (yz - y\cos(xy), xz - x\cos(xy), xy + 3)\,.$$

Como todas las derivadas parciales son continuas, f es diferenciable en $\mathbb{R}^3$. En particular, la derivada de f el punto $P = (0, 1, 2)$ es $\nabla f\,(0, 1, 2) = (1, 0, 3)\,.$

Por el momento hemos definido el concepto de derivada para campos escalares. Si queremos extender este concepto a funciones vectoriales $f : \mathbb{R}^n \to \mathbb{R}^p$, usaremos de nuevo las componentes $f_1, f_2, \ldots, f_p$.

Recordemos que si $f : \mathbb{R} \to \mathbb{R}^p$ entonces la derivada se definía como el vector formado por las derivadas de las componentes. Escribiremos esta derivada como vector columna:

$$Df(x_0) = f'(x_0) = \begin{pmatrix} f_1'(x_0) \\ f_2'(x_0) \\ \vdots \\ f_p'(x_0) \end{pmatrix}.$$

Si $f : \mathbb{R}^n \to \mathbb{R}^p$ y $P \in \mathbb{R}^n$ entonces diremos que f es diferenciable en P si cada una de las componentes de f es diferenciable en P. Se define la derivada de f en P como la matriz cuyas filas son los vectores gradiente de cada una de las componentes:

$$Df(P) = \begin{pmatrix} \nabla f_1(P) \\ \nabla f_2(P) \\ \vdots \\ \nabla f_p(P) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(P) & \frac{\partial f_1}{\partial x_2}(P) & \cdots & \frac{\partial f_1}{\partial x_n}(P) \\ \frac{\partial f_2}{\partial x_1}(P) & \frac{\partial f_2}{\partial x_2}(P) & \cdots & \frac{\partial f_2}{\partial x_n}(P) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_p}{\partial x_1}(P) & \frac{\partial f_p}{\partial x_2}(P) & \cdots & \frac{\partial f_p}{\partial x_n}(P) \end{pmatrix}.$$

Obsérvese que si $f : \mathbb{R}^n \to \mathbb{R}^p$ es diferenciable entonces la derivada en cada punto es una matriz con p filas y n columnas. Esta matriz también se llama **matriz jacobiana** o matriz de las derivadas parciales. Cada elemento $a_{ij} = \partial f_i / \partial x_j(P)$ representa la tasa de variación instantánea de la componente f_i respecto de la variable x_j en el punto P.

Ejemplo: Consideremos la función $f : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $f(x, y, z) = (xyz, \operatorname{sen}(xy) + z)$. Vamos a calcular la derivada de f en el punto $(0, 1, 2)$. Para ello necesitamos las derivadas parciales de las componentes $f_1(x, y, z) = xyz,\ f_2(x, y, z) = \operatorname{sen}(xy) + z$.

$$\frac{\partial f_1}{\partial x} = yz \quad ; \quad \frac{\partial f_1}{\partial y} = xz \quad ; \quad \frac{\partial f_1}{\partial z} = xy\,;$$

$$\frac{\partial f_2}{\partial x} = y\cos(xy) \quad ; \quad \frac{\partial f_2}{\partial y} = x\cos(xy) \quad ; \quad \frac{\partial f_2}{\partial z} = 1\,.$$

Entonces:

$$Df(x, y, z) = \begin{pmatrix} yz & xz & xy \\ y\cos(xy) & x\cos(xy) & 1 \end{pmatrix} \Longrightarrow Df(0, 1, 2) = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Linealidad de la derivada

Si $f : \mathbb{R}^n \to \mathbb{R}^p$ y $g : \mathbb{R}^n \to \mathbb{R}^p$ son funciones diferenciables en $P \in \mathbb{R}^n$ y $\lambda, \mu \in \mathbb{R}$ entonces $(\lambda f + \mu g)$ es diferenciable en P y $D(\lambda f + \mu g)(P) = \lambda Df(P) + \mu Dg(P)$.

5.4. Regla de la cadena

Como en el caso de funciones escalares de una variable, una herramienta fundamental del cálculo de la derivada es la regla de la cadena.

Si $f : \mathbb{R}^n \to \mathbb{R}^p$ y $g : \mathbb{R}^p \to \mathbb{R}^q$ son dos funciones, f es diferenciable en P y g es diferenciable en $f(P)$, entonces $(g \circ f)$ es diferenciable en P y además

$$D(g \circ f)(P) = Dg(f(P))\, Df(P).$$

Nótese que el producto de matrices está bien definido porque $Dg(f(P))$ tiene p columnas y $Df(P)$ tiene p filas.

Ejemplo: Consideremos las funciones $f : \mathbb{R}^3 \to \mathbb{R}^2$ y $g : \mathbb{R}^2 \to \mathbb{R}^2$ definidas por

$$f(x, y, z) = (xyz, \operatorname{sen}(xy) + z)\,, \quad g(u, v) = (uv^2, v\, e^u).$$

Vamos a calcular la derivada de $g \circ f$ en el punto $(0, 1, 2)$. Por la regla de la cadena, teniendo en cuenta que $f(0, 1, 2) = (0, 2)$, se tiene:

$$D(g \circ f)(0, 1, 2) = Dg(0, 2)Df(0, 1, 2).$$

Ya hemos calculado

$$Df(0, 1, 2) = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

A continuación calculamos $Dg(0, 2)$: como $g_1(u, v) = uv^2$ y $g_2(u, v) = v\, e^u$, se tiene:

$$Dg(u, v) = \begin{pmatrix} \partial g_1/\partial u & \partial g_1/\partial v \\ \partial g_2/\partial u & \partial g_2/\partial v \end{pmatrix} = \begin{pmatrix} v^2 & 2uv \\ v\, e^u & e^u \end{pmatrix} \Longrightarrow Dg(0, 2) = \begin{pmatrix} 4 & 0 \\ 2 & 1 \end{pmatrix}$$

Finalmente,

$$D(g \circ f)(0, 1, 2) = Dg(0, 2)Df(0, 1, 2) = \begin{pmatrix} 4 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 8 & 0 & 0 \\ 5 & 0 & 1 \end{pmatrix}.$$

Regla de la cadena: una variable independiente

Consideremos un campo escalar f definido en un subconjunto abierto D de $\mathbb{R}^n$ y una función vectorial de una variable $g : I \subset \mathbb{R} \to \mathbb{R}^n$, con componentes $g(t) = (x_1(t), x_2(t), \ldots, x_n(t))$. Entonces la acción del campo f sobre la función vectorial g es la función escalar de una variable

$$h(t) = f(g(t)) = f(x_1(t), x_2(t), \ldots, x_n(t)).$$

La derivada de h se puede calcular usando la regla de la cadena:

$$h'(t) = \nabla f(g(t)) \cdot g'(t) = \frac{\partial f}{\partial x_1}\, x_1'(t) + \frac{\partial f}{\partial x_2}\, x_2'(t) + \cdots + \frac{\partial f}{\partial x_n}\, x_n'(t).$$

Las variables $x_1, x_2, \ldots, x_n$ son variables dependientes y t es la variable independiente.

Ejemplo: Un cilindro circular recto varía de tal manera que su radio r crece a razón de 2 cm/hora y su altura decrece a razón de 4 cm/hora. Calcular la relación entre el radio y la altura en el momento en el que el volumen empieza a decrecer. (Inicialmente la altura es mayor que el radio).

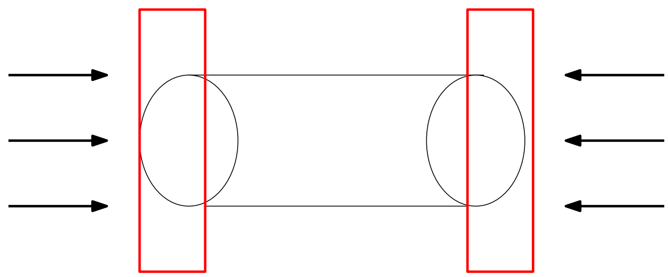


Figura 5.5: Ilustración de la deformación ejercida en el cilindro, reduciendo la altura y aumentando el radio.

El volumen es $V(r,h) = \pi r^2 h$. El radio r y la altura h son variables dependientes y t es la variable independiente (el diagrama de dependencias se muestra en la figura 5.6).

Las tasas de variación son $r'(t) = 2$, $h'(t) = -4$. Entonces:

$$V'(t) = \frac{\partial V}{\partial r}\, r'(t) + \frac{\partial V}{\partial h}\, h'(t) = 2\pi r h r'(t) + \pi r^2 h'(t) = 4\pi r h - 4\pi r^2.$$

El volumen empieza a decrecer cuando su tasa de variación pasa de positiva a negativa. Como

$$V'(t) = 4\pi r h - 4\pi r^2 > 0 \Longleftrightarrow h > r,$$

el volumen empieza a decrecer cuando la longitud del radio iguala a la de la altura.

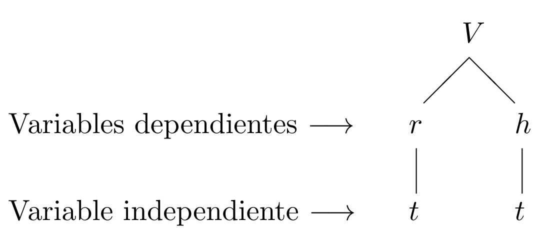


Figura 5.6: Diagrama de dependencias para el volumen.

Regla de la cadena: varias variables independientes

Supongamos que $z = f(x, y)$, donde x e y dependen de u y v, es decir, $z = f(x(u, v), y(u, v))$.

Si f es diferenciable y existen las derivadas parciales de $x = x(u, v)$ e $y = y(u, v)$ respecto de u y v entonces se pueden calcular las derivadas parciales de z respecto de u y v del siguiente modo:

$$\frac{\partial z}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} \quad ; \quad \frac{\partial z}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}.$$

En este caso, el diagrama de dependencias sería el mostrado en la figura 5.7.

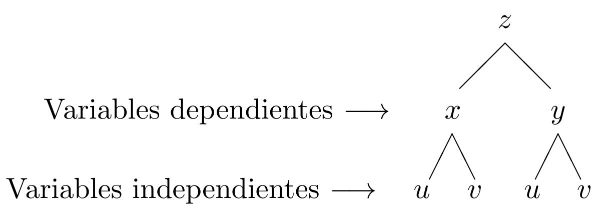


Figura 5.7: Diagrama de dependencias para dos variables dependientes y dos independientes.

Estas fórmulas se extienden de manera natural al caso de un número arbitrario de variables dependientes y variables independientes: si $z = f(y_1, y_2, \ldots, y_n)$ e $y_i = y_i(x_1, x_2, \ldots, x_p)$ para cada $i = 1, 2, \ldots, n$, entonces, para cada $j = 1, 2, \ldots, p$,

$$\frac{\partial z}{\partial x_j} = \frac{\partial f}{\partial y_1}\frac{\partial y_1}{\partial x_j} + \frac{\partial f}{\partial y_2}\frac{\partial y_2}{\partial x_j} + \cdots + \frac{\partial f}{\partial y_n}\frac{\partial y_n}{\partial x_j}.$$

5.5. Derivadas direccionales y dirección de máximo crecimiento

Consideremos un campo escalar $f : D \subset \mathbb{R}^n \to \mathbb{R}$ definido en un subconjunto abierto D de $\mathbb{R}^n$, $P \in D$ y $\vec{u} \in \mathbb{R}^n$ un vector unitario, es decir $\|\vec{u}\| = 1$. Se considera la recta $y(t) = P + t\vec{u}$ que pasa por P y tiene la dirección de $\vec{u}$. Si $g(t) = f(P + t\vec{u})$ es la acción del campo f sobre la recta, aplicando la regla de la cadena, se obtiene:

$$g'(t) = \nabla f(P + t\vec{u}) \cdot \vec{u}$$

Se define la **derivada direccional** de f en P según el vector $\vec{u}$ como

$$D_{\vec{u}} f(P)) = g'(0) = \nabla f(P) \cdot \vec{u}.$$

La derivada direccional mide la tasa de cambio de $f(x)$ en el punto $(P, f(P))$ cuando x pasa por P siguiendo la dirección de $\vec{u}$. Nótese que en los casos particulares de que $\vec{u} = (1, 0)$ o $\vec{u} = (0, 1)$, las derivadas direccionales son respectivamente la derivada parcial de f respecto de x y la derivada parcial de f respecto de y.

Ejemplo: Calcular la derivada direccional de $f(x,y) = x^2y + y^3$ en el punto $(2,1)$ en la dirección de $\vec{v} = (1,1)$. En primer lugar, $\nabla f(x,y) = (2xy, x^2 + 3y^2) \Longrightarrow \nabla f(2,1) = (4,7)$. Por otra parte, el vector unitario en la dirección de $\vec{v}$ es

$$\vec{u} = \frac{\vec{v}}{\|\vec{v}\|} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right).$$

Por tanto, la derivada direccional es

$$D_{\vec{u}}f(2,1) = (4,7)\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} = \frac{11}{\sqrt{2}}.$$

Dirección de máximo crecimiento

Consideremos un campo escalar diferenciable $f : \mathbb{R}^n \to \mathbb{R}$. Si $\nabla f(P) \neq 0$ entonces el máximo valor de la derivada direccional se alcanza para el vector unitario en la dirección del gradiente

$$\vec{u} = \frac{\nabla f(P)}{\|\nabla f(P)\|}.$$

Es decir, el gradiente de f apunta en la dirección de $\mathbb{R}^n$ en la que f crece más rápidamente.

Esta propiedad es una consecuencia directa de la fórmula del producto escalar. La derivada direccional en la dirección de un vector unitario u es

$$\nabla f(P) \cdot u = \|\nabla f(P)\|\,\|u\|\,\cos(\phi) = \|\nabla f(P)\|\,\cos(\phi),$$

donde ϕ es el ángulo entre $\nabla f(P)$ y u. Claramente, la derivada direccional alcanza su máximo valor cuando $\cos(\phi) = 1$, es decir, cuando u apunta en la dirección y sentido del gradiente.

Obsérvese que el mismo razonamiento indica que la dirección de máximo decrecimiento es la opuesta del gradiente, es decir, $-\nabla f(P)$. Así, el valor máximo de la derivada direccional de f en el punto P es $\|\nabla f(P)\|$ y el mínimo es $-\|\nabla f(P)\|$.

Ejemplo: Calcular la dirección de máximo crecimiento del campo escalar $f(x,y) = x^2 + y^2$ en el punto $(1,1)$.

$$\nabla f(x,y) = (2x, 2y) \Longrightarrow \nabla f(1,1) = (2,2).$$

La dirección de máximo crecimiento es la del vector unitario $\vec{u} = (1/\sqrt{2}, 1/\sqrt{2})$.

En el siguiente ejemplo vemos una aplicación de este resultado a la propagación del flujo de calor.

Ejemplo: Consideremos una placa cuadrada $[0,5] \times [0,5]$ y supongamos que la temperatura en un punto (x,y) es $T(x,y) = x^2 + y^2/4$. Calcular en qué sentido irá el flujo de calor $C(x,y)$ en el punto $(2,4)$.

Teniendo en cuenta que el calor fluye de puntos de mayor temperatura a puntos de menor temperatura, se tiene que

$$C(x,y) = -k\,\nabla T(x,y),$$

donde k es una constante llamada conductividad térmica del medio. En este caso:

$$\nabla T(x,y) = (2x, y/2) \Longrightarrow C(2,4) = -k\,\nabla T(2,4) = -k(4,2).$$

Por tanto, el calor fluye en el sentido del vector unitario $\vec{u} = \left(\dfrac{-2}{\sqrt{5}}, \dfrac{-1}{\sqrt{5}}\right)$.

5.6. Gradiente y curvas de nivel

Consideremos un campo escalar en el plano $f : D \subset \mathbb{R}^2 \to \mathbb{R}$. El siguiente resultado establece que el gradiente de f en un punto P apunta en la dirección ortogonal a la curva de nivel de f que pasa por P.

Si f es diferenciable en (x_0, y_0) entonces el gradiente de f en (x_0, y_0) es ortogonal a la curva de nivel que pasa por (x_0, y_0).

El razonamiento es sencillo. Si denotamos por $r(t) = (x(t), y(t))$ a la curva de nivel que pasa por (x_0, y_0), de la definición de curva de nivel se obtiene que $f(r(t)) = K$ para una constante K. Aplicando la regla de la cadena se tiene:

$$\nabla f(r(t)) \cdot r'(t) = 0\,.$$

Como $r'(t)$ es el vector tangente a la curva, se deduce que el gradiente de f es ortogonal a la curva de nivel en cada punto de la curva.

En el ejemplo anterior, la curva de nivel correspondiente al punto $(2,4)$ viene dada por la elipse $x^2 + y^2/4 = 8$. En la figura 5.8 se representa la curva de nivel junto con el vector en la dirección opuesta al gradiente, que marca la dirección del flujo de calor.

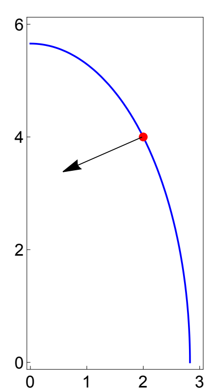


Figura 5.8: Curva de nivel de la temperatura $T(x,y) = x^2 + y^2/4$ pasando por el punto $(2,4)$ y la dirección del flujo de calor, marcada por el vector $-\nabla T(2,4)$.

Ecuación normal de una recta tangente

Usando que el gradiente es un vector normal a la curva de nivel, la ecuación de la recta tangente a la curva de nivel en el punto (x_0, y_0) se puede expresar en la forma

$$\nabla f(x_0, y_0) \cdot (x - x_0, y - y_0) = 0.$$

Este resultado se puede aplicar al cálculo de la recta tangente a una curva dada en forma implícita.
Ejemplo: Calcular la recta tangente a la circunferencia $x^2 + y^2 = 2$ en el punto $(1, 1)$.

La circunferencia es una curva de nivel del campo escalar en $\mathbb{R}^2$ definido por $f(x, y) = x^2+y^2$. Como $\nabla f(x, y) = (2x, 2y)$, se tiene que $\nabla f(1, 1) = (2, 2)$ y por tanto la ecuación de la recta tangente es

$$(2,2)\begin{pmatrix} x-1 \\ y-1 \end{pmatrix} = 0 \Longleftrightarrow 2(x-1) + 2(y-1) = 0 \Longleftrightarrow y = 2 - x.$$

5.7. Derivación implícita

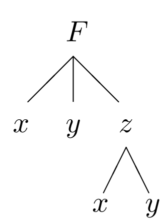
Ya hemos descrito en la sección 3.4 del capítulo 3 cómo derivar de forma implícita una expresión $F(x, y) = 0$, donde $y = y(x)$. Una manera alternativa de resolver el problema es aplicar la regla de la cadena a la expresión $F(x, y(x)) = 0$, lo que proporciona:

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\, y'(x) = 0 \Longrightarrow y'(x) = -\frac{\partial F/\partial x}{\partial F/\partial y}.$$

Este procedimiento se puede generalizar a funciones de más variables dependientes e independientes utilizando la forma general de la regla de la cadena.

Describimos el proceso para una expresión con tres variables x, y, z en las que una de ellas se despeja en función de las otras dos (por ejemplo, $z = z(x, y)$). En este caso tenemos una expresión $F(x, y, z) = 0$ que define implícitamente una función $z : \mathbb{R}^2 \to \mathbb{R}$ dada por $z = z(x, y)$. Es posible calcular las derivadas parciales de z respecto de x e y sin necesidad de despejar z (lo cual es a veces imposible).

En este caso, el diagrama de dependencias sería el siguiente:



Aplicando la regla de la cadena a la expresión $F(x, y, z(x, y)) = 0$, tenemos:

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} = 0 \Longrightarrow \frac{\partial z}{\partial x} = -\frac{\partial F/\partial x}{\partial F/\partial z}$$

$$\frac{\partial F}{\partial y} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial y} = 0 \Longrightarrow \frac{\partial z}{\partial y} = -\frac{\partial F/\partial y}{\partial F/\partial z}\,.$$

Ejemplo: La temperatura T de un fluido depende del instante t y de la densidad d según la relación

$$T^2 e^{-dt} + d\ln(t+T) = d(1+t).$$

Calcular la densidad d y la tasa de cambio de temperatura respecto al tiempo cuando $t = 0$ y $T = 1$.

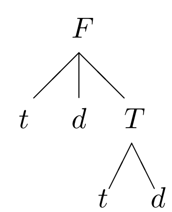
Para $t = 0$ y $T = 1$, se obtiene de la igualdad del enunciado que $d = 1$.

La tasa de cambio de temperatura respecto al tiempo es $\partial T/\partial t$. Para calcularla, usaremos derivación implícita. Para ello, consideramos la función auxiliar

$$F(t, d, T) = T^2 e^{-dt} + d\ln(t+T) - d(1+t),$$

de modo que la relación se escribe como $F(t, d, T) = 0$.

El diagrama de dependencias es



Derivando respecto a t:

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial T}\frac{\partial T}{\partial t} = 0 \Longrightarrow \frac{\partial T}{\partial t} = -\frac{\partial F/\partial t}{\partial F/\partial T}\,.$$

Las derivadas parciales son:

$$\frac{\partial F}{\partial t} = -dT^2 e^{-dt} + \frac{d}{t+T} - d \quad ; \quad \frac{\partial F}{\partial T} = 2Te^{-dt} + \frac{d}{t+T}.$$

Evaluamos en $t = 0$, $T = 1$, $d = 1$ y obtenemos

$$\frac{\partial T}{\partial t} = -\frac{\partial F/\partial t}{\partial F/\partial T} = -\frac{(-1)}{3} = \frac{1}{3}.$$

5.8. Derivadas parciales de orden superior

Si $f : \mathbb{R}^n \to \mathbb{R}$ es un campo escalar y las derivadas parciales $\partial f / \partial x_i$ son funciones derivables entonces se pueden definir las derivadas parciales segundas (o de segundo orden). Se define la derivada parcial segunda de f respecto de x_i y de x_j como

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right).$$

Ejemplo: Si $f(x, y) = xy + x^2y^2 + ye^x$ entonces

$$\frac{\partial f}{\partial x} = y + 2xy^2 + ye^x \quad ; \quad \frac{\partial f}{\partial y} = x + 2x^2y + e^x\,.$$

Por tanto, las derivadas parciales segundas son:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = 2y^2 + ye^x \quad ; \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = 1 + 4xy + e^x\,;$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = 2x^2 \quad ; \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = 1 + 4xy + e^x\,.$$

Un campo $f : \mathbb{R}^n \to \mathbb{R}$ es de clase $\mathcal{C}^1$ si existen las derivadas parciales primeras y son continuas. Si existen las derivadas parciales segundas y son continuas, se dice que f es de clase $\mathcal{C}^2$. De modo análogo se definen las derivadas de orden superior a 2 y las funciones de clase $\mathcal{C}^k$ para $k > 2$.

Por ejemplo, para la función del ejemplo anterior,

$$\frac{\partial^3 f}{\partial y \partial x^2} = \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x^2}\right) = \frac{\partial}{\partial y}\left(2y^2 + ye^x\right) = 4y + e^x\,;\ \frac{\partial^3 f}{\partial x^3} = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x^2}\right) = \frac{\partial}{\partial x}\left(2y^2 + ye^x\right) = ye^x\,.$$

Para funciones de clase $\mathcal{C}^2$, las derivadas cruzadas no dependen del orden de derivación:

Teorema de las derivadas cruzadas: Si $f : \mathbb{R}^n \to \mathbb{R}$ es un campo escalar de clase $\mathcal{C}^2$ entonces, para cada par de variables independientes x_i, x_j, se cumple que

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}\,.$$

Este teorema se extiende a derivadas de orden superior. Si una función es de clase $\mathcal{C}^k$ su derivada parcial $k-$ésima con respecto de k variables es independiente del orden de derivación.

Matriz hessiana

Si $f : \mathbb{R}^n \to \mathbb{R}$ es un campo escalar de clase $\mathcal{C}^2$, se llama **matriz hessiana** de f a la matriz cuadrada $Hf(x)$ de tamaño $n \times n$ que tiene en el lugar (i, j) la derivada parcial segunda $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$, es decir,

$$Hf(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}.$$

En cada punto $P \in \mathbb{R}^n$, $Hf(P) \in \mathcal{M}_{n\times n}(\mathbb{R})$ y, en virtud del teorema de las derivadas cruzadas, es una matriz simétrica.

Por ejemplo, para la función $f(x,y) = xy + x^2y^2 + ye^x$ del ejemplo anterior, la matriz hessiana es

$$Hf(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2y^2 + ye^x & 1 + 4xy + e^x \\ 1 + 4xy + e^x & 2x^2 \end{pmatrix}.$$

La matriz hessiana en el punto $(0,1)$ es

$$Hf(0,1) = \begin{pmatrix} 3 & 2 \\ 2 & 0 \end{pmatrix}.$$

Aproximación de grado 2

Al igual que para una función escalar f de una variable las derivadas segundas permiten obtener condiciones suficientes para asegurar que un punto crítico es un máximo o un mínimo local, la matriz hessiana (que contiene las derivadas parciales segundas) juega ese papel para campos escalares en $\mathbb{R}^n$. La razón es que si la mejor aproximación de grado 2 para una función $f : \mathbb{R} \to \mathbb{R}$ cerca de un punto x_0 es el polinomio de Taylor de grado 2 centrado en x_0:

$$p_2(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2,$$

para funciones $f : D \subset \mathbb{R}^n \to \mathbb{R}$ la mejor aproximación de grado 2 de f en un entorno de un punto $P \in D$ viene dada por

$$f(x) \approx q(x) = f(P) + \nabla f(P) \cdot (x - P) + \frac{1}{2}\,(x - P)^t\,(Hf(P))\,(x - P).$$

Por ejemplo, consideremos el campo escalar $f(x,y) = -(x^2+3)y^2 - e^x - 2y$. La mejor aproximación de grado 1 en un entorno de $(0,0)$ es el plano tangente

$$z = q_1(x,y) = f(0,0) + \nabla f(0,0) \begin{pmatrix} x \\ y \end{pmatrix} = -1 - x - 2y.$$

La mejor aproximación de grado 2 está dada por

$$z = q_2(x,y) = f(0,0) + \nabla f(0,0) \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2}(x,y) Hf(0,0) \begin{pmatrix} x \\ y \end{pmatrix} = -1 - x - 2y - \frac{x^2}{2} - 3y^2.$$

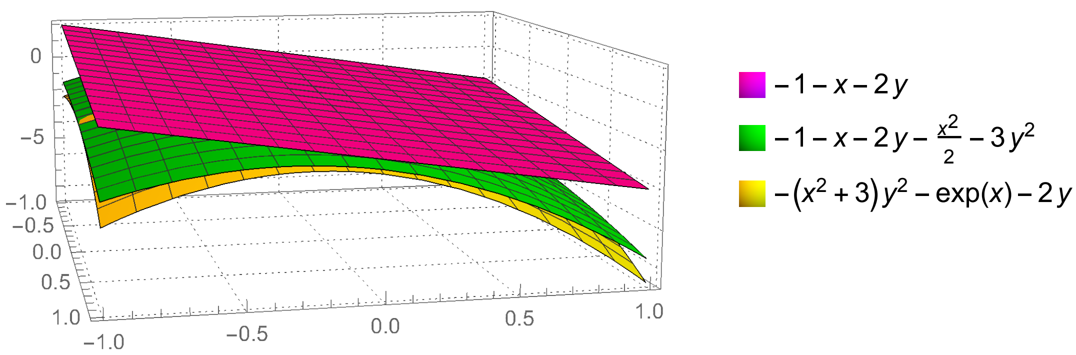


Figura 5.9: Ilustración de la gráfica de $f(x,y) = -(x^2+3)y^2 - e^x - 2y$ (en amarillo) junto con su plano tangente en el punto $(0,0,-1)$ (en rosa) y su aproximación de grado 2 (en verde).

5.9. Extremos locales de un campo escalar

Recordemos que si $f : I \to \mathbb{R}$ es una función diferenciable definida en un intervalo real I y un punto x_0 del interior de I es un extremo local, necesariamente $f'(x_0) = 0$. Además, si $f''(x_0) \neq 0$ entonces el signo de la derivada segunda permite saber si f alcanza en x_0 un máximo local o un mínimo local.

Estos argumentos se pueden generalizar a campos escalares en $\mathbb{R}^n$. En primer lugar, introducimos la definición de máximo y mínimo local.

Si $f : D \subset \mathbb{R}^n \to \mathbb{R}$ es un campo escalar definido en un subconjunto D de $\mathbb{R}^n$, se dice que f alcanza un **máximo local** en un punto $P \in D$ si existe un entorno B_0 de P tal que $f(x) \leq f(P)$, $\forall\, x \in B_0$.

La definición de **mínimo local** se obtiene cambiando el sentido de la desigualdad en la definición anterior. Diremos que P es un punto de **extremo local** para f si f alcanza en P un máximo local o un mínimo local.

Observación: Un entorno del punto P se puede definir como la intersección de una bola abierta centrada en P con el dominio de definición D.

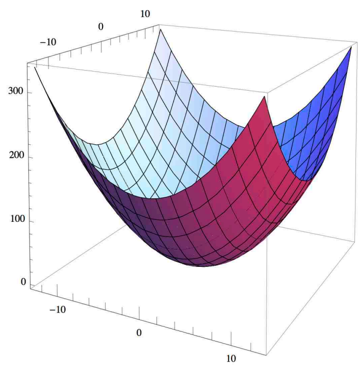
Condición necesaria de extremo local: Si $f : D \subset \mathbb{R}^n \to \mathbb{R}$ es un campo escalar de clase $\mathcal{C}^1$ definido en un subconjunto D de $\mathbb{R}^n$ y f alcanza en un punto interior $P \in D$ un extremo local entonces $\nabla f(P) = (0, 0, \ldots, 0)$.

Esta condición es muy intuitiva: si el gradiente no se anulase, alguna de las derivadas parciales sería positiva o negativa en P. Entonces la gráfica de f sería creciente o decreciente cuando x se acerca a P en la dirección de esa variable. En consecuencia, en P no podría haber ni un máximo ni un mínimo local.

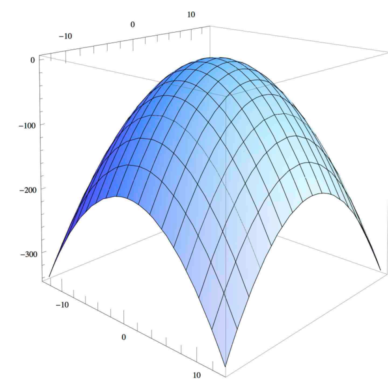
Nótese que si $f : \mathbb{R}^2 \to \mathbb{R}$ es un campo escalar en el plano, la condición necesaria es equivalente a decir que el plano tangente a la gráfica de f en el punto $(P, f(P))$ es paralelo al plano horizontal $z = 0$.

Por tanto, los extremos locales de f se buscarán entre los puntos que tienen gradiente cero o aquellos donde f no es diferenciable. Estos puntos se llaman **puntos críticos** de f.

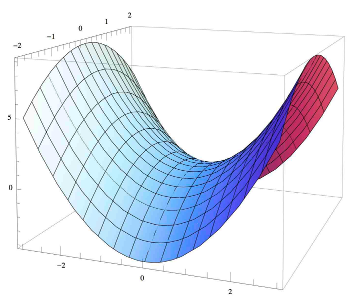
Si $\nabla f(P) = (0, 0, \ldots, 0)$ pero f no alcanza un extremo en P, entonces se dice que f tiene en P un **punto de silla.** En estos puntos, la gráfica de f crece en algunas direcciones cerca de P y decrece en otras. En la figura 5.10 se presentan cuatro ejemplos en los que f tiene un punto crítico en $(0, 0)$: dos extremos locales y dos puntos de silla.



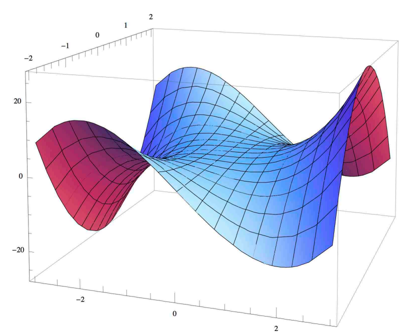
Mínimo: $f(x, y) = x^2 + y^2$



Máximo: $f(x, y) = -x^2 - y^2$



Silla clásica: $f(x, y) = x^2 - y^2$



Silla de mono: $f(x, y) = x^3 - 3xy^2$

Figura 5.10: En las cuatro gráficas $(0, 0)$ es un punto crítico de f. En las superiores f alcanza en $(0, 0)$ un extremo local y en las inferiores sendos puntos de silla.

<u>Observación:</u> El teorema solo vale para puntos interiores. Los posibles extremos locales en la frontera hay que determinarlos de otro modo.

<u>Ejemplo:</u> Calcular los puntos críticos de la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = 4x^3 - 12xy + y^2$.

$$\nabla f(x, y) = (12x^2 - 12y, -12x + 2y) = (0, 0) \Longleftrightarrow \left\{ \begin{array}{c} y = x^2 \\ 6x = y \end{array} \right\}$$

Por tanto,

$$x^2 = y = 6x \Longrightarrow x^2 - 6x = 0 \Longrightarrow x(x-6) = 0.$$

Las únicas raíces reales son $x = 0$, $x = 6$. Como $y = x^2$, se obtienen los puntos críticos $(0,0)$ y $(6,36)$.

Usando la expresión de la aproximación cuadrática

$$f(x) \approx q(x) = f(P) + \nabla f(P) \cdot (x - P) + \frac{1}{2}(x-P)^t \ (Hf(P)) \ (x - P)$$

se puede deducir que, cuando P es un punto crítico (es decir, $\nabla f(P) = (0, 0, \ldots, 0)$), lo que determina el signo de $f(x) - f(P)$ para puntos x próximos a P es el signo de la forma cuadrática $\omega(x) = x^t H x$, donde $H = Hf(P)$ es una matriz simétrica $n \times n$.

Recordemos la clasificación de matrices simétricas $H \in \mathcal{M}_{n\times n}(\mathbb{R})$:

- H es *definida positiva* si $x^t H x > 0\,,\ \forall\, x \neq 0$,
- H es *definida negativa* si $x^t H x < 0\,,\ \forall\, x \neq 0$,
- H es *indefinida* si existen dos vectores $x, y \in \mathbb{R}^n$ tales que $x^t H x > 0\,,\ y^t H y < 0$.

Se obtiene así el siguiente criterio para clasificar los puntos críticos de un campo escalar:

Criterio de la derivada segunda para extremos locales: Si f es un campo escalar de clase $\mathcal{C}^2$ definido en un subconjunto abierto D de $\mathbb{R}^n$ y $P \in D$ es un punto crítico de f, es decir, $\nabla f(P) = (0, 0, \ldots, 0)$, entonces:

1. Si la matriz $H = Hf(P)$ es definida positiva entonces f alcanza en P un mínimo local.
2. Si la matriz $H = Hf(P)$ es definida negativa entonces f alcanza en P un máximo local.
3. Si el determinante de H no es cero y la matriz $H = Hf(P)$ es indefinida entonces f alcanza en P un punto de silla.

Recordemos que si H es simétrica y $|H| \neq 0$ entonces se puede determinar si H es definida positiva, definida negativa o indefinida analizando el signo de los menores principales de la matriz.

Clasificación de formas cuadráticas no degeneradas: Supongamos que H es simétrica y $|H| \neq 0$.

(a) Si todos los menores principales de H son positivos entonces H es definida positiva.

(b) Si los menores principales de orden impar son negativos y los de orden par son positivos entonces H es definida negativa.

(c) En cualquier otro caso, H es indefinida.

Ejemplo: Clasificar los puntos críticos del campo escalar $f : \mathbb{R}^2 \to \mathbb{R}$ definido por

$$f(x,y) = x^2 + y^4 - y^2 + 2xy.$$

Calculamos los puntos críticos:

$$\nabla f(x,y) = (0,0) \Longleftrightarrow (2x+2y, 4y^3 - 2y + 2x) = (0,0) \Longleftrightarrow \left\{ \begin{array}{l} y = -x \\ 4y^3 - 4y = 0 \end{array} \right\}.$$

Como $4y^3 - 4y = 4y(y^2-1)$, las soluciones de $y^3 - 4y = 0$ son 0, 1 y -1. Como $x = -y$, se obtienen los puntos críticos $P_1 = (0,0)$, $P_2 = (1,-1)$ y $P_3 = (-1,1)$. La matriz hessiana de f es

$$Hf(x,y) = \begin{pmatrix} 2 & 2 \\ 2 & 12y^2 - 2 \end{pmatrix}.$$

Sustituyendo en los puntos críticos:

$$Hf(0,0) = \begin{pmatrix} 2 & 2 \\ 2 & -2 \end{pmatrix} \quad ; \quad Hf(1,-1) = Hf(-1,1) = \begin{pmatrix} 2 & 2 \\ 2 & 10 \end{pmatrix}.$$

$Hf(0,0)$ es indefinida y por tanto $P_1 = (0,0)$ es un punto de silla. $Hf(1,-1) = Hf(-1,1)$ es definida positiva y por tanto f alcanza en P_2 y P_3 sendos mínimos locales.

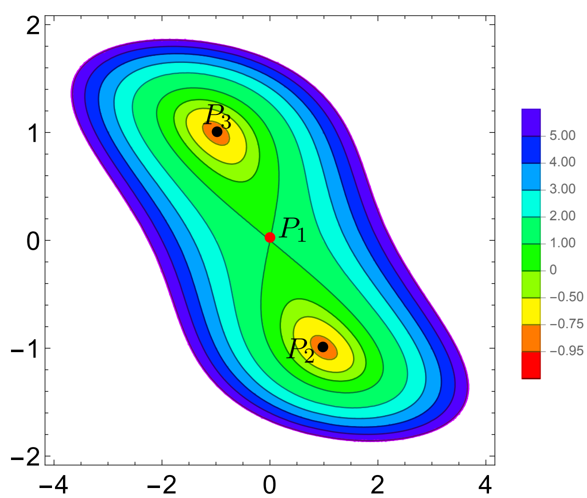


Figura 5.11: Mapa de curvas de nivel del campo escalar $f(x,y) = x^2 + y^4 - y^2 + 2xy$, con dos mínimos locales y un punto de silla.

Es interesante ver el mapa de curvas de nivel de f para visualizar los puntos críticos. En la figura 5.11, los puntos P_2 y P_3 se sitúan en los círculos naranja, donde f toma sus valores más bajos, lo que se corresponde con su condición de mínimos locales. Sin embargo, para el punto de silla $P_1 = (0,0)$, la función crece si nos movemos desde P_1 en la dirección horizontal (hacia las zonas azules) y decrece si nos movemos en la dirección vertical (hacia las zonas amarillas y naranjas).

Observación: Los puntos críticos de f para los que $|Hf(P)| \neq 0$ se llaman **puntos críticos no degenerados.** Los puntos críticos degenerados se suelen examinar directamente. Por ejemplo, la silla de mono de la figura 5.10 es punto crítico degenerado. Para deducir que es un punto de silla basta observar que f restringida a la recta $y = 0$ es la función escalar $f(x, 0) = x^3$, que es una función creciente, por lo que no puede haber ni un máximo ni un mínimo local en $(0, 0)$.

Extremos globales

Si $f : D \subset \mathbb{R}^n \to \mathbb{R}$ es un campo escalar definido en un conjunto D de $\mathbb{R}^n$ y $P \in D$, se dice que f alcanza en P su **máximo global** en D si $f(P) \geq f(x)$, $\forall\, x \in D$. De modo análogo se define el mínimo global (cambiando de sentido la desigualdad). El siguiente resultado es similar al de existencia de extremos globales en intervalos compactos de $\mathbb{R}$:

Existencia de extremos globales: Si $f : D \subset \mathbb{R}^n \to \mathbb{R}$ es un campo escalar definido en un conjunto D de $\mathbb{R}^n$, D es compacto y f es continua entonces f alcanza en D su máximo global y su mínimo global.

Para localizar los extremos globales de f se siguen estos pasos:

1. Localizar los puntos críticos de f en el interior de D.
2. Hallar los posibles extremos de f considerada como función definida en la frontera de D.
3. Calcular el valor de f en todos los puntos críticos y seleccionar el mayor y el menor.

5.10. Extremos condicionados. Multiplicadores de Lagrange

Para determinar los posibles extremos en la frontera usaremos el método de los multiplicadores de Lagrange, que introducimos a continuación para el cálculo de extremos condicionados. En estos problemas se considera una función objetivo (de la que se quieren calcular los puntos donde se alcanzan los máximos y mínimos locales) y una o varias restricciones que deben cumplir esos puntos.

Multiplicadores de Lagrange con una restricción

Comenzamos explicando el método de los multiplicadores de Lagrange cuando hay una única restricción. Para motivarlo, consideremos el siguiente ejemplo:

Ejemplo. Calcular los puntos de la elipse de ecuación $x^2 + 2y^2 - x = 6$ más próximos y más alejados del punto $(0, 0)$.

La función que hay que minimizar es la distancia de un punto (x, y) de la recta a $(0, 0)$, es decir, $d(x, y) = \sqrt{x^2 + y^2}$. Como la distancia es siempre positiva, esto es equivalente a minimizar el cuadrado de la distancia, por lo que consideramos la función objetivo $f(x, y) = x^2 + y^2$.

Trazamos las circunferencias $x^2 + y^2 = r^2$ de centro $(0, 0)$ y radio r, que son las curvas de nivel de f. Como se observa en la figura 5.12, los puntos más cercanos y más alejados de $(0, 0)$ corresponden a puntos de tangencia de la elipse con alguna de las curvas de nivel de f.

Para determinar analíticamente estos puntos, observemos que en los puntos buscados la elipse y la circunferencia comparten recta tangente. Como la recta tangente a la curva de nivel de un campo escalar es ortogonal al gradiente del campo en ese punto, los vectores gradiente de los campos $f(x, y) = x^2 + y^2$ y $g(x, y) = x^2 + 2y^2 - x$ en los puntos de tangencia deben ir en la misma dirección (es claro que la elipse es la curva de nivel 6 del campo g).

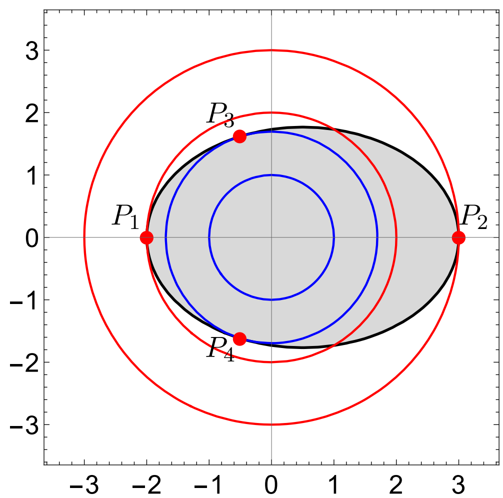


Figura 5.12: Localización de los puntos de tangencia de la elipse $x^2 + 2y^2 - x = 6$ con las curvas de nivel de la función objetivo $f(x, y) = x^2 + y^2$.

Esto lleva a lo que llamaremos la ecuación de multiplicadores:

$$\nabla f(x, y) = \lambda \nabla g(x, y),\ \lambda \in \mathbb{R}.$$

Es decir, para que (x, y) sea un punto de tangencia, debe existir $\lambda \in \mathbb{R}$ tal que

$$(2x, 2y) = \lambda(2x - 1, 4y).$$

Como (x, y) también debe ser un punto de la elipse, se llega al siguiente sistema de 3 ecuaciones con 3 incógnitas:

$$\begin{cases} 2x = \lambda\,(2x - 1) \\ 2y = 4\lambda y \\ x^2 + 2y^2 - x = 6. \end{cases}$$

Las únicas soluciones son $P_1 = (-2, 0)$ (para $\lambda = 4/5$), $P_2 = (3, 0)$ (para $\lambda = 6/7$), $P_3 = \left(-1/2, \sqrt{21/8}\right)$ y $P_4 = \left(-1/2, -\sqrt{21/8}\right)$ (los dos últimos para $\lambda = 1/2$).

Como $d(P_1, (0, 0)) = 2$, $d(P_2, (0, 0)) = 3$ y $d(P_3, (0, 0)) = d(P_4, (0, 0)) \approx 1.695$, los puntos más próximos son P_3 y P_4 y el más alejado es P_2.

Este ejemplo se generaliza a campos escalares en $\mathbb{R}^n$ para definir el método de los multiplicadores de Lagrange.

Teorema de multiplicadores de Lagrange con una restricción: Supongamos que los campos escalares f y g definidos en un subconjunto abierto D de $\mathbb{R}^n$ son de clase $\mathcal{C}^1$ y

$$\nabla g(x_1, x_2, \ldots, x_n) \neq (0, 0, \ldots, 0) \text{ si } g(x_1, x_2, \ldots, x_n) = 0.$$

Si $P \in D$ es un punto de máximo o mínimo local de f sujeto a la restricción $g(x_1, \ldots, x_n) = 0$, entonces existe un $\lambda \in \mathbb{R}$ tal que

$$\nabla f(P) = \lambda \nabla g(P).$$

Observación: Para $\lambda = 0$ la ecuación de multiplicadores $\nabla f(P) = \lambda \nabla g(P)$ se convierte en $\nabla f(P) = 0$, de modo que el multiplicador cero corresponde al caso en que P es un punto crítico de f (sin restricciones).

Ejemplo: Calcular los extremos globales del campo escalar $f(x, y) = x^2 + 3x + y^2 + 1$ en la región

$$D = \left\{(x, y) \in \mathbb{R}^2 \,/\, (x+1)^2 + y^2 \leq 1\right\}.$$

Es claro que D es compacto porque es cerrado y acotado.

En primer lugar, calculamos los puntos críticos de f en el interior de D.

$$\nabla f(x, y) = (2x + 3, 2y) = (0, 0) \Longleftrightarrow x = \frac{-3}{2}\,,\ y = 0.$$

El único punto crítico en el interior de D es $P_1 = (-3/2, 0)$.

La frontera de D está formada por los puntos (x, y) que cumplen la restricción

$$(x+1)^2 + y^2 = 1.$$

Para calcular los posibles extremos en la frontera planteamos un problema de extremos condicionados con función objetivo f y restricción $g(x, y) = (x+1)^2 + y^2 - 1 = 0$.

La ecuación de los multiplicadores es:

$$\nabla f(x, y) = \lambda \nabla g(x, y) \Longleftrightarrow (2x + 3, 2y) = \lambda(2x + 2, 2y) \Longleftrightarrow \left\{\begin{array}{l} 2x + 3 = \lambda\,(2x + 2) \\ 2y = \lambda\, 2y. \end{array}\right.$$

De la segunda ecuación se obtiene que $y = 0$ o $\lambda = 1$, así que distinguimos esos dos casos:

Caso 1: $\lambda = 1$. En ese caso la primera ecuación queda $2x + 3 = 2x + 2$, que obviamente no tiene soluciones.

Caso 2: $y = 0$. La restricción proporciona $(x+1)^2 = 1$, y por tanto $x + 1 = \pm 1$, lo que da las soluciones $x = 0$ y $x = 2$. Por tanto, las soluciones de la ecuación de multiplicadores son $P_2 = (0, 0)$ y $P_3 = (2, 0)$.

Finalmente, como $f(-3/2, 0) = -5/4$, $f(-2, 0) = -1$ y $f(0, 0) = 1$, se deduce que el máximo global se alcanza en $(0, 0)$ y el mínimo global se alcanza en $(-3/2, 0)$.

En la figura 5.13 se representan los extremos globales de f en D junto con el mapa de curvas de nivel de f.

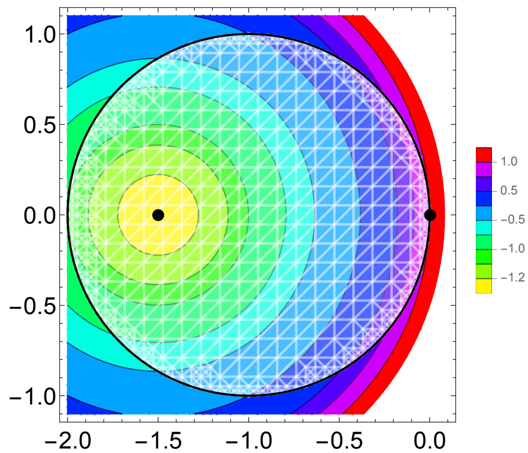


Figura 5.13: Mapa de curvas de nivel de f junto con el dominio D (círculo blanco delimitado por la circunferencia negra). El mínimo global se alcanza en el interior de D (punto negro en el centro del círculo amarillo) y el máximo global global se alcanza en la frontera (cuando el borde de D alcanza la curva de nivel 1, en color rojo).

Multiplicadores de Lagrange con varias restricciones

El método de los multiplicadores se extiende al cálculo de extremos de una función sujeta a varias restricciones. El resultado general es el siguiente:

Teorema de multiplicadores de Lagrange con varias restricciones: Supongamos que los campos escalares f y $g_1, g_2, \ldots, g_k$, definidos en un dominio D de $\mathbb{R}^n$, son de clase $\mathcal{C}^1$. Si $P \in D$ es un punto de máximo o mínimo local de f sujeto a las restricciones

$$\begin{aligned} g_1(x_1, \ldots, x_n) &= 0 \\ g_2(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ g_k(x_1, \ldots, x_n) &= 0, \end{aligned}$$

entonces existen $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ tales que

$$\nabla f(P) = \lambda_1 \nabla g_1(P) + \lambda_2 \nabla g_2(P) + \cdots + \lambda_k \nabla g_k(P).$$

Ejemplo: Calcular las alturas máxima y mínima que alcanza la curva intersección del plano $x + y + z = 0$ con la esfera $x^2 + y^2 + z^2 = 1$.

Tenemos que encontrar los extremos de la función $f(x, y, z) = z$ restringidos a la intersección del plano con la esfera, es decir, sujetos a las restricciones

$$g_1(x, y, z) = x + y + z = 0 \quad ; \quad g_2(x, y, z) = x^2 + y^2 + z^2 - 1 = 0\,.$$

Según el resultado anterior, en los puntos de extremo deben existir $\lambda, \mu \in \mathbb{R}$ tales que:

$$\nabla f(x,y,z) = \lambda \nabla g_1(x,y,z) + \mu \nabla g_2(x,y,z) \Longleftrightarrow (0,0,1) = \lambda(1,1,1) + \mu(2x,2y,2z) \Longleftrightarrow$$

$$\Longleftrightarrow \left\{ \begin{array}{l} \lambda + 2\mu x = 0 \\ \lambda + 2\mu y = 0 \\ \lambda + 2\mu z = 1 \end{array} \right\} \Longleftrightarrow 2\mu x = 2\mu y = 1 - 2\mu z\,.$$

Es claro que $\mu \neq 0$ y por tanto $x = y$. Como $x+y+z = 0$, se obtiene que $z = -x-y = -2y$. Sustituyendo en la ecuación $x^2 + y^2 + z^2 = 1$, tenemos :

$$1 = x^2 + y^2 + z^2 = y^2 + y^2 + (-2y)^2 = 6y^2 \Longrightarrow y = \pm\frac{1}{\sqrt{6}}\,.$$

Si $y = 1/\sqrt{6}$, entonces $x = 1/\sqrt{6}$, $z = -2/\sqrt{6}$. Si $y = -1/\sqrt{6}$, entonces $x = -1/\sqrt{6}$, $z = 2/\sqrt{6}$. Por tanto los puntos de extremo son

$$P_1 = \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}\right) \quad ; \quad P_2 = \left(\frac{-1}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right)\,.$$

En consecuencia, la altura máxima es $z = 2/\sqrt{6}$ y la mínima es $z = -2/\sqrt{6}$.

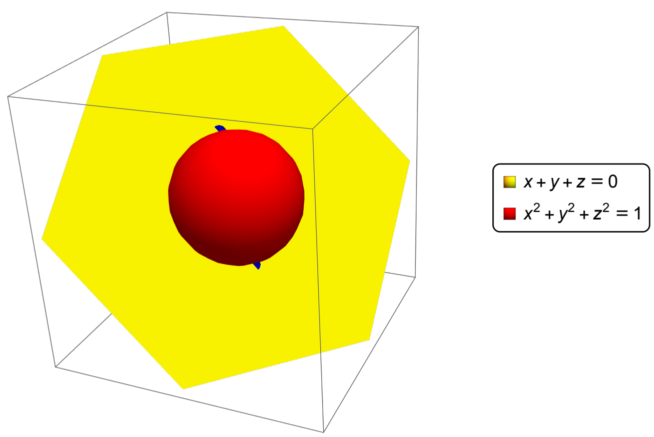


Figura 5.14: Ilustración de la intersección del plano $x+y+z = 0$ con la esfera $x^2 + y^2 + z^2 = 1$. En azul se destacan los puntos donde alcanza las alturas máxima y mínima.

Condición suficiente para la existencia de extremos condicionados

Hay que tener en cuenta que el método de los multiplicadores de Lagrange detecta puntos de tangencia, que pueden corresponder a mínimos, máximos o puntos de silla. Si la función objetivo f es continua y el conjunto de puntos que cumplen las restricciones es compacto, entonces se alcanzan el mínimo y el máximo globales de modo que, para saber a qué puntos corresponden de entre los que cumplen la ecuación de multiplicadores, basta evaluar f en cada uno de ellos y escoger los valores extremos. De esta forma se ha procedido en los ejemplos anteriores. En otros casos se puede usar el siguiente criterio:

Condición suficiente de extremos condicionados: Supongamos que los campos escalares f y $g_1, g_2, \ldots, g_k$ definidos en un dominio D de $\mathbb{R}^n$ son de clase $\mathcal{C}^2$ y que que para un punto $P \in D$ existen $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{R}$ tales que

$$\nabla f(P) = \lambda_1 \nabla g_1(P) + \lambda_2 \nabla g_2(P) + \cdots + \lambda_k \nabla g_k(P).$$

Se considera la función definida por $F(x) = f(x) - (\lambda_1 g_1(x) + \lambda_2 g_2(x) + \cdots + \lambda_k g_k(x))$. Denotemos por $U = \operatorname{Ker}(Dg(P)) = \{x \in \mathbb{R}^n \,/\, Dg(P)\,x = 0\}$, donde $g : \mathbb{R}^n \to \mathbb{R}^k$ es la función definida por $g(x) = (g_1(x), g_2(x), \ldots, g_k(x))$. Si $H = HF(P)$ es la matriz hessiana de F en P y $\omega : U \to \mathbb{R}$ es la forma cuadrática definida en U por $\omega(x) = x^t H x$, entonces:

- Si ω es definida positiva entonces P es un mínimo local de f restringido a la condición $g(x) = (0, 0, \ldots, 0)$.
- Si ω es definida negativa entonces P es un máximo local de f restringido a la condición $g(x) = (0, 0, \ldots, 0)$.
- Si ω es indefinida y no degenerada entonces P es un punto de silla.

Nuestro último ejemplo es una adaptación de un problema sugerido por Roberto Agromayor Otero, antiguo alumno de la asignatura.

Ejemplo: Se desea multiplicar por 8 la presión de un gas en 3 etapas con un coeficiente de compresión $r_i > 0$ en cada una de ellas, de modo que $r_1 r_2 r_3 = 8$.

Sabiendo que la potencia que consumen los compresores viene dada por

$$P = \frac{1}{3}\left(r_1^3 + r_2^3 + r_3^3\right),$$

calcular los coeficientes de compresión r_1, r_2, r_3 que minimizan la potencia consumida.

Por simplificar, denotamos $x = r_1$, $y = r_2$, $z = r_3$. La función objetivo es el consumo de potencia

$$P(x, y, z) = \frac{1}{3}\left(x^3 + y^3 + z^3\right).$$

Claramente P es un campo escalar de clase $\mathcal{C}^2$ definido en el conjunto abierto

$$D = \{(x, y, z) \in \mathbb{R}^3 \,/\, x > 0,\ y > 0,\ z > 0\}.$$

Usamos el método de multiplicadores de Lagrange para calcular el mínimo de P sujeto a la restricción

$$g(x, y, z) = xyz - 8 = 0.$$

Planteamos la ecuación de multiplicadores:

$$\nabla P(x,y,z) = \lambda \nabla g(x,y,z) \Longleftrightarrow (x^2,y^2,z^2) = \lambda(yz,xz,xy) \Longleftrightarrow \begin{cases} x^2 = \lambda yz \\ y^2 = \lambda xz \\ z^2 = \lambda xy. \end{cases}$$

Multiplicando la primera ecuación por x, la segunda por y y la tercera por z, se obtiene

$$x^3 = y^3 = z^3 = \lambda xyz.$$

De aquí se deduce que $x = y = z$. Juntando esta condición con la restricción $xyz = 8$, se obtiene

$$x^3 = y^3 = z^3 = 8 \Longrightarrow x = y = z = 2.$$

Veamos que los coeficientes de compresión $r_1 = r_2 = r_3 = 2$ minimizan la potencia consumida. Para ello consideramos el punto $(2,2,2)$ y aplicamos la condición suficiente para extremos condicionados.

En primer lugar, de la ecuación $x^2 = \lambda yz$ se obtiene $4 = 4\lambda^2$ y por tanto $\lambda = 1$. Definimos la función

$$F(x,y,z) = P(x,y,z) - \lambda g(x,y,z) = \frac{1}{3}\left(x^3+y^3+z^3\right) - xyz + 8.$$

Teniendo en cuenta que $\nabla F(x,y,z) = (x^2-yz, y^2-xz, z^2-xy)$, calculamos la matriz hessiana:

$$HF(x,y,z) = \begin{pmatrix} 2x & -z & -y \\ -z & 2y & -x \\ -y & -x & 2z \end{pmatrix} \Longrightarrow HF(2,2,2) = \begin{pmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{pmatrix}.$$

Los menores principales de $HF(2,2,2)$ son $\Delta_1 = 4$, $\Delta_2 = 12$, $\Delta_3 = 0$, que no permiten clasificar el punto crítico. Restringimos la forma cuadrática al subespacio

$$U = \text{Ker}\ (\nabla g(2,2,2)) = \text{Ker}\,(4,4,4) = \left\{(x,y,z)\in\mathbb{R}^3 \,/\, 4x+4y+4z=0\right\} =$$
$$= \left\{(x,y,z)\in\mathbb{R}^3 \,/\, z=-x-y\right\} = \left\{(x,y,-x-y)\,/\,x,y\in\mathbb{R}\right\}.$$

Para los vectores de U la forma cuadrática se expresa como:

$$\omega(x,y,-x-y) = (x,y,-x-y)\begin{pmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ -x-y \end{pmatrix} =$$
$$= 4x^2+4y^2+4(-x-y)^2-4xy-4x(-x-y)-4y(-x-y) =$$
$$= 12x^2+12y^2+12xy = (x,y)\begin{pmatrix} 12 & 6 \\ 6 & 12 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Como los menores principales de la matriz resultante son $\Delta_1 = 12 > 0$, $\Delta_2 = 108 > 0$, ω es definida positiva y en el punto $(2,2,2)$ se alcanza el mínimo de la potencia $P(x,y,z)$ condicionado a la restricción $xyz = 8$.

Referencias

- Gerald L. Bradley y Karl J. Smith, "Cálculo de una y varias variables" (dos volúmenes), Prentice Hall, 1998.
- Serge Lang, "Cálculo", Addison-Wesley Iberoamericana, 1990.
- Jerrold E. Marsden y Anthony J. Tromba, "Cálculo vectorial", 5a. ed., Pearson, 2004.
- Jon Rogawski, "Cálculo: Varias variables", Reverte, 2012.
- James Stewart, "Cálculo. Conceptos y contextos", 4a. ed., Thompson, 2010.