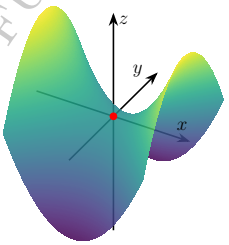
Diferenciación de funciones de varias variables

Prof. Renato Álvarez Nodarse

www: https://renato.ryn-fismat.es

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geq c(|\alpha_1| + \cdots + |\alpha_n|)$$



$$\frac{\partial (g \circ f)_i(a)}{\partial x_j} = \sum_{l=1}^{m} \frac{\partial g_i(f(a))}{\partial x_l} \frac{\partial f_l(a)}{\partial x_j}$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Versión 2.1 | 10/11/2026

Diferenciación de funciones de varias variables

Índice general

Prefacio II

1. Preliminares 1

1.1. Cálculo diferencial en $\mathbb{R}$ 1

1.2. Espacios vectoriales 13

1.3. $\mathbb{R}^n$ como espacio normado y métrico 16

1.4. Espacios normados de dimensión finita 27

1.5. Aplicaciones lineales 31

1.6. Espacios euclídeos 35

1.7. Problemas propuestos 38

2. Límite, continuidad y diferenciabilidad 41

2.1. Límite y continuidad de funciones de varias variables 41

2.2. Diferenciabilidad de funciones de varias variables 45

2.3. Otras propiedades de la diferenciación 57

2.4. Derivadas de orden superior 66

2.5. Diferenciales de orden superior 72

2.6. El Teorema de Taylor para funciones de varias variables . . . 76

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

3. El Teorema de la función implícita **83**

3.1. Planteamiento del problema 83

3.2. Caso de una única ecuación 85

3.3. Caso de un sistema de ecuaciones 91

3.4. Teorema de la función inversa 96

3.5. Cambio de variables 100

3.6. Dependencia funcional 102

4. Extremos de funciones de varias variables **107**

4.1. Extremos libres 107

4.2. Extremos condicionados 115

Bibliografía 125

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Prefacio

Estas notas contienen el contenido del curso *Diferenciación de funciones de varias variables* impartido por el autor en la Facultad de Matemáticas de la Universidad de Sevilla desde hace más de diez años. Dicho curso es una asignatura del segundo curso del Grado en matemáticas.

Las notas están divididas en cuatro capítulos. El primero es un resumen de algunos de los conceptos necesarios para entender las presentes notas. Incluye, tanto resultados del análisis matemático *elemental* en $\mathbb{R}$, como propiedades de $\mathbb{R}^n$ como espacio normado de dimensión finita.

El segundo capítulo constituye el núcleo teórico esencial del curso y en el se tratan los conceptos fundamentales: límites, continuidad y derivabilidad para funciones de varias variables. En el tercer y cuarto capítulos discutiremos dos de las aplicaciones más relevantes de las funciones de varias variables: el Teorema de la función implícita y el estudio de los extremos de dichas funciones, respectivamente.

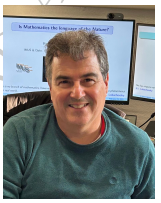
Quiero agradecer a todos los que de una forma u otra me han ayudado a que estas notas sean posibles. En primer lugar a mi familia, a los que les he robado un tiempo precioso y en segundo lugar a los alumnos a los que iba dirigidos. Sus muchas preguntas y comentarios a lo largo de los años me han permitido mejorar la forma de introducir los distintos conceptos y pruebas de los principales resultados.

También soy deudor de los autores de las monografías mencionadas en la bibliografía. Aunque estas notas no coinciden en su totalidad con ninguna de ellas (si ese fuese el caso no hubiese considerado necesario escribir este texto, hubiese bastado usar una de dichas monografías), de cada una de ellas he tomado prestado, entre otras cosas, demostraciones, ejemplos y problemas. Para ayudar al lector a la hora de consultar dichos textos en el

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

apartado dedicado a la bibliografía se menciona y/o recomiendan algunos para profundizar en los distintos capítulos de estas notas.

Espero que el estudiante, y el lector en general, le encuentre utilidad a estas notas y disfrute de su lectura como yo he disfrutado al escribirlas.



Renato Álvarez Nodarse

Sevilla, 1 de septiembre de 2026

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Todas las figuras que aparecen en el texto han sido hechas por el autor con el paquete de LaTeX TikZ o con en programa Maxima CAS.

Diferenciación de funciones de varias variables

Diferenciación de funciones de varias variables

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Capítulo 1

Preliminares

El principal objetivo de estas notas es aprender las técnicas de diferenciación para las funciones vectoriales de varias variables y sus aplicaciones. No obstante, en un intento de hacerlas *autocontenidas*, en este primer capítulo vamos a introducir algunos conceptos que nos serán de utilidad.

Así, en la sección 1.1, recordaremos algunos resultados del cálculo diferencial de funciones reales de una variable y, en el resto de las secciones 1.2-1.6, discutiremos algunas de las propiedades de los espacios vectoriales, los espacios métricos, normados y euclídeos, con especial hincapié en $\mathbb{R}^n$.

1.1. Cálculo diferencial en $\mathbb{R}$

Recordemos que el conjunto de los números reales, que denotaremos por $\mathbb{R}$, es el conjunto de todos los números racionales $\mathbb{Q}$ e irracionales $\mathbb{I}$, o sea $\mathbb{R} = \mathbb{Q} \cup \mathbb{I}$.

1.1.1. Sucesiones de números reales

En adelante denotaremos por $\mathbb{N} = \{1, 2, 3, \dots, n, n+1, \dots\}$ al conjunto de los números naturales que es el conjunto infinito más sencillo.

Definición 1.1.1 *La aplicación o función $f : \mathbb{N} \mapsto A$ que hace corresponder a cada número natural un elemento de cierto conjunto A se denomina sucesión de elementos de A.*

Por ejemplo, si $A = \mathbb{R}$, diremos que $f : \mathbb{N} \mapsto \mathbb{R}$ define una sucesión de números reales. Usualmente se usa la notación $x_n := f(n)$ para denotar a los distintos términos de la sucesión de forma que $(x_n)_n$ denotará a la propia sucesión.

Como ejemplo de sucesiones de números reales tenemos las siguientes: $x_n = n$ (la sucesión de los números naturales $1, 2, 3, \dots$), $x_n = 1/n$ (los inversos de los números naturales $(1, 1/2, 1/3, \dots)$), o $x_n = 1$ (todos los elementos son iguales a 1, $1, 1, 1 \dots$). Es importante tener en cuenta que una sucesión **siempre** tiene infinitos términos que son $x_1, x_2, x_3, \dots$, los cuales no necesariamente han de ser distintos. Por ejemplo la sucesión $x_n = 1/n$ tiene todos sus términos distintos, pero la sucesión definida por $x_n = 0$ si n es par y $1/n$ si n es impar tiene infinitos términos iguales (los correspondientes a n par que valen cero es este caso) e infinitos desiguales. Análogamente, la sucesión $x_n = 1$ tiene todos sus infinitos términos iguales.

Definición 1.1.2 *Un conjunto A es numerable si*

1. *A es vacío.*
2. *A tiene un número finito de elementos.*
3. *Si A tiene un número infinito de elementos, estos se pueden poner en correspondencia biunívoca con $\mathbb{N}$.*

Por ejemplo, $P = \{2, 4, 6, \dots\}$, el conjunto de todos los números pares es numerable. En efecto, cualquiera sea $n \in \mathbb{N}$ existe un único $p \in P$ tal que $p = 2n$ y viceversa, cualquiera sea $p \in P$, existe un único $n \in \mathbb{N}$ tal que $n = p/2$, es decir existe una correspondencia biunívoca entre ambos conjuntos, o lo que es lo mismo *hay tantos números pares como naturales*. En ejemplo menos intuitivo es el caso del conjunto de todos los números racionales $\mathbb{Q}$. Se puede probar que $\mathbb{Q}$ es numerable, sin embargo $\mathbb{R}$ no lo es.

En adelante llamaremos δ-vecindad o δ-entorno de un punto $x = a$, y lo denotaremos por $U_\delta(a)$, al intervalo abierto $(a - \delta, a + \delta)$.

Definición 1.1.3 *Sea S un conjunto infinito de números reales. Se denomina punto de acumulación o punto límite de un subconjunto $S \subset \mathbb{R}$, a un punto $x \in \mathbb{R}$ tal que en cualquier entorno $U_\delta(a)$ de x hay al menos un punto de S diferente de x.*[1]

[1] Esta definición es equivalente a decir que en cualquier entorno (vecindad) de x tiene que haber infinitos puntos de S.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Según esta definición x puede o no pertenecer a S. Por ejemplo, Si $S = (a, b)$, entonces x puede ser cualquier número real del intervalo $[a, b]$, en particular $x = a$ o b (que no pertenecen a (a, b)). Por otro lado, el conjunto $S = \{1/n, n \in \mathbb{N}\}$ tiene como único punto de acumulación el $x = 0$, mientras que el conjunto $S = \{n^2, n \in \mathbb{N}\}$ no tiene ningún punto de acumulación.

Definición 1.1.4 *Diremos que un conjunto $A \subset \mathbb{R}$ es acotado si existe un $\delta > 0$ tal que, para todo $x \in A$, $x \in U_\delta(0)$. En otras palabras si todo $x \in A$ es tal que $|x| < \delta$.*

El siguiente resultado es de gran relevancia en el análisis de funciones de una variable real.

Lema 1.1.5 *(Bolzano-Weierstrass) Cualquier subconjunto infinito, es decir, con un número infinito de elementos, acotado de $\mathbb{R}$ tiene por lo menos un punto de acumulación.*

Un concepto esencial del análisis es el concepto de límite.

Definición 1.1.6 *Se dice que una sucesión $(x_n)_n$ tiene límite $a \in \mathbb{R}$, y lo denotaremos por $\lim_{n\to\infty} x_n = a$, si para todo $\epsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$, se tiene que $|x_n - a| < \epsilon$. O sea,*

$$\lim_{n\to\infty} a_n = a \iff \forall \epsilon > 0, \exists N \in \mathbb{N},\ \forall n > N, \quad |x_n - a| < \epsilon.$$

Si una sucesión tiene límite diremos que es convergente, en caso contrario diremos que es divergente.

Las sucesiones convergentes tienen varias propiedades importantes entre ellas destacan las dos siguientes:

1. Si $(x_n)_n$ tiene límite este es único.
2. Toda sucesión convergente es acotada, es decir existe un $M > 0$ tal que, para todo $n \in \mathbb{N}$, $|x_n| \leq M$. Esto es lo mismo que decir que el conjunto $\{x_1, x_2, x_3, \dots\}$ es acotado.
3. Dada dos sucesiones $(a_n)_n$ y $(b_n)_n$ tales que $\lim_{n\to\infty} a_n = a$, y $\lim_{n\to\infty} b_n = b$. Entonces:

   a) $\lim_{n\to\infty} a_n + b_n = a + b$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

b) $\lim_{n\to\infty} a_n \cdot b_n = a \cdot b$. En particular, para todo $\alpha \in \mathbb{R}$, $\lim_{n\to\infty} \alpha\, a_n = \alpha\, a$.

c) Si $\forall n \in \mathbb{N}$, $b_n \neq 0$, $\quad b \neq 0$, entonces, $\lim_{n\to\infty} a_n/b_n = a/b$.

Definición 1.1.7 *Sea $\mathcal{N}$ un subconjunto de $\mathbb{N}$, $\mathcal{N} = \{n_1, n_2, ..., n_k, ...,\ k \in \mathbb{N}\}$ tal que $n_k < n_{k+1}$ para todo k. Sea $(x_n)_n$ una sucesión de números reales. Construyamos una nueva sucesión cuyos elementos sean los elementos de x_n correspondientes a los valores n_k de $\mathcal{N}$. La nueva sucesión así obtenida la denotaremos $(x_{n_k})_n$ y la llamaremos subsucesión de $(x_n)_n$*

Por ejemplo, sea $x_n = (-1)^n$. Si elegimos $\mathcal{N} = \{2, 4, ..., 2k, ...,\ k \in \mathbb{N}\}$ y $\mathcal{N} = \{1, 3, ..., 2k-1, ...,\ k \in \mathbb{N}\}$, obtendremos las subsucesiones de los elementos pares $a_{2k} = 1$ e impares $a_{2k-1} = -1$, respectivamente.

A partir de la definición 1.1.6 se puede probar que cualquier subsucesión $(x_{n_k})_n$ de una sucesión convergente $(x_n)_n$ es convergente y su límite es el mismo. O sea, si

$$\lim_{n\to\infty} x_n = a \quad \Rightarrow \quad \lim_{n_k\to\infty} x_{n_k} = a.$$

Una consecuencia del Lema 1.1.5 es el siguiente:

Teorema 1.1.8 *(de Bolzano-Weierstrass para las sucesiones) De toda sucesión acotada en $\mathbb{R}$ se puede extraer una subsucesión convergente.*

Dentro del conjunto de todas las sucesiones las sucesiones de Cauchy juegan un papel esencial.

Definición 1.1.9 *Una sucesión $(x_n)_n$ es de Cauchy si para todo $\epsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todos $n, m > N$, entonces $|a_n - a_m| < \epsilon$. O, equivalentemente, existe un $N \in \mathbb{N}$ tal que para todo $n > N \in \mathbb{N}$ y todo $p \in \mathbb{N}$, $|a_{n+p} - a_n| < \epsilon$.*

De hecho se tiene el siguiente teorema:

Teorema 1.1.10 *(Criterio de Cauchy para las sucesiones) Una sucesión $(x_n)_n$ es de Cauchy si y solo si es convergente.*

Una de las consecuencias del teorema anterior es el siguiente interesante resultado:

Teorema 1.1.11 *Para que una sucesión sea convergente es necesario y suficiente que cualquiera de sus subsucesiones sea convergente y tenga el mismo límite.*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

1.1.2. Funciones reales de una variable real

Recordemos a continuación algunas de los principales resultados para las funciones de una variable real.

Definición 1.1.12 *Diremos que una función $f : A \subset \mathbb{R} \mapsto \mathbb{R}$ tiene límite l cuando x tiende a a (punto de acumulación de A), y lo denotaremos por $\lim_{x\to a} f(x) = l$, si para cualquiera sea la sucesión $(x_n)_n \in A$ que converja a a se tiene que $f(x_n) \to l$. Si además $f(a) = l$, entonces se dice que f es continua en a. Si f es continua para todo x de A, se dice que f es continua en A.*

La definición de límite anterior es equivalente a decir que para todo $\epsilon > 0$ existe un $\delta > 0$ tal que para todo x que satisfaga $0 < |x - a| < \delta$, entonces $|f(x) - l| < \epsilon$. Gráficamente la definición anterior la podemos ver en la figura 1.1.1.

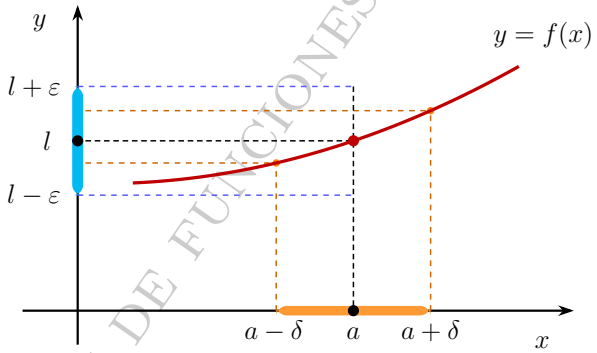


Figura 1.1.1: Significado geométrico de la definición de límite para una función $f : \mathbb{R} \mapsto \mathbb{R}$ tal que $\lim_{x\to a} f(x) = l$.

Nótese que de la definición 1.1.12 también se sigue que, si existe el límite, entonces es único así como las siguientes propiedades algebraicas de los límites:

Así, sean f, $g : A \mapsto \mathbb{R}$ dos funciones tales que $\lim_{x\to a} f(x) = l$ y $\lim_{x\to a} g(x) = m$. Entonces:

1. $\lim_{x\to a} f(x) + g(x) = l + m$,
2. $\lim_{x\to a} f(x) \cdot g(x) = l \cdot m$, y en particular $\lim_{x\to a} \alpha f(x) = \alpha\, l$, $\forall \alpha \in \mathbb{R}$,

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

3. Si $g(x) \neq 0$ en un entorno de a (excepto quizá en el propio punto a) y $m \neq 0$, entonces $\lim_{x\to a} f(x)/g(x) = l/m$.

Para las funciones continuas se tienen además las siguientes propiedades:

1. Si $f(x)$ y $g(x)$ son continuas en $x = a$, entonces $f(x) + g(x)$, $f(x) \cdot g(x)$ y $f(x)/g(x)$ son continuas en $x = a$, donde en el último caso se supone $g(x) \neq 0$ en un entorno de $x = a$.

2. Si $f(x)$ es continua en $x = a$, existe todo un entorno $U_\delta(a)$ de a, donde f es acotada. O sea, existe un $\delta > 0$ tal que si $|x - a| < \delta$, $|f(x)| < \infty$.

3. Sea $f : A \mapsto \mathbb{R}$ una función continua en $x = a$. Entonces existe un entorno de $x = a$ tal que si $f(a) > 0$ entonces $f(x)$ es mayor que cero en todo el entorno y si $f(a) < 0$ entonces $f(x)$ es menor que cero en dicho entorno. O sea, si f es continua en $x = a$, existe un $\delta > 0$, para todo $x \in (a - \delta, a + \delta)$, $f(x)$ tiene el mismo signo (positivo o negativo) que $f(a)$

Uno de los teoremas más importantes para funciones continuas es el Teorema de Bolzano, que mostramos gráficamente en la figura 1.1.2.

Teorema 1.1.13 *(Bolzano) Sea la función $f : [a, b] \subset \mathbb{R} \mapsto \mathbb{R}$ es continua en todo $[a, b]$, cerrado y acotado, y sea que los valores en los extremos $f(a)$, $f(b)$ tienen signos distintos ($f(a)f(b) < 0$). Entonces existe un punto c en el interior del intervalo $[a, b]$, $c \in (a, b)$, tal que $f(c) = 0$ –ver figura 1.1.2–.*

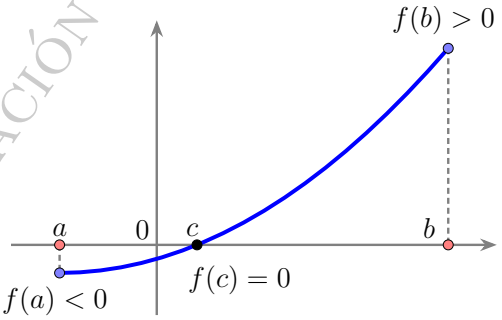


Figura 1.1.2: Esquema que ilustra el Teorema de Bolzano.

Este teorema se puede generalizar a un abierto (a, b) si los límites laterales $\lim_{y\to a+} f(x)$ y $\lim_{y\to b-} f(x)$ existen y tienen signos contrarios.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Definición 1.1.14 *Se dice que una función $f : A \subset \mathbb{R} \mapsto \mathbb{R}$ es derivable en $x = a$, punto de acumulación de A, si existe el límite*

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

Dicho límite, que denotaremos por $f'(a)$, se denomina derivada *de $f(x)$ en $x = a$. Si f es derivable para todo x de A, se dice que f es derivable en A.*

La derivada tiene un significado geométrico muy sencillo: el valor $f'(a)$ no es más que la pendiente de la recta tangente a la curva $y = f(x)$ en el punto $x = a$, tal y como se muestra en la figura 1.1.3. Así, la ecuación de la recta tangente $y = t(x)$ a la curva en $x = a$ se escribe como

$$y = f(a) + f'(a)(x - a). \tag{1.1.1}$$

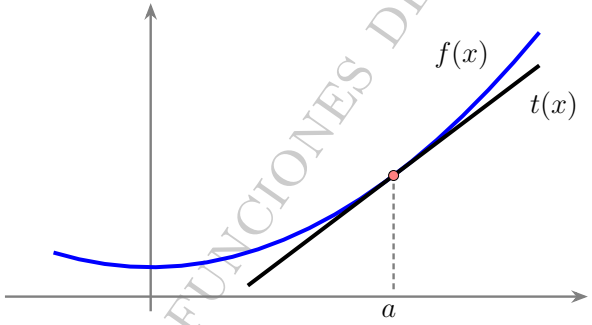


Figura 1.1.3: Significado geométrico de la derivada: $f'(a)$ es la pendiente de la recta $t(x)$ tangente (ver la ecuación (1.1.1)) a la curva definida por la función $f(x)$ en el punto $x = a$.

Es fácil comprobar usando la definición 1.1.14 que si f es derivable en $x = a$, entonces es continua en $x = a$, pero no a la inversa. Por ejemplo, la función $f : [-1, 1] \mapsto \mathbb{R}$, con $f(x) = |x|$ es continua en $x = 0$ pero claramente no es derivable en $x = 0$ (¿por qué?).

De la definición 1.1.14 se sigue que si f, $g : A \mapsto \mathbb{R}$ son dos funciones derivables en A, entonces las funciones $f(x) + g(x)$, $f(x) \cdot g(x)$ y $f(x)/g(x)$ son derivables en A y se tiene que

$$[f(x) + g(x)]' = f'(x) + g'(x), \quad [f(x) \cdot g(x)]' = f'(x)g(x) + f(x)g'(x),$$

$$\left[\frac{f(x)}{g(x)}\right]' = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}, \text{ siempre que } g(x) \neq 0.$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Las funciones derivables tienen un sinnúmero de interesantes propiedades algunas de las cuales enunciaremos a continuación.

Definición 1.1.15 *Se dice que una función $f : A \subset \mathbb{R} \mapsto \mathbb{R}$ es monótona creciente (decreciente) en un subconjunto $B \subset A$ si para todos $x_1, x_2 \in B$ tal que $x_1 < x_2$, $f(x_1) < f(x_2)$ $(f(x_1) > f(x_2))$.*

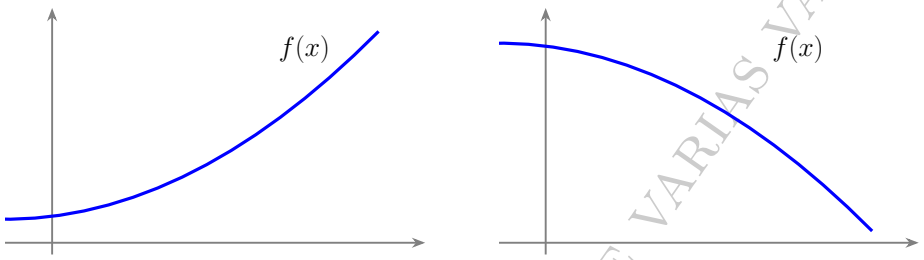


Figura 1.1.4: Ejemplo de función creciente (izquierda) y decreciente (derecha).

Proposición 1.1.16 *Sea $f(x) : [a, b] \mapsto \mathbb{R}$ una función continua en $[a, b]$ y derivable en (a, b). Entonces, si la función $f(x)$ es creciente en (a, b), $f'(x) \geq 0$ en todo $[a, b]$. Si por el contrario $f(x)$ es decreciente en (a, b), entonces $f'(x) \leq 0$ en todo (a, b).*

Definición 1.1.17 *Sea la función $f : A \subset \mathbb{R} \mapsto \mathbb{R}$ y sea a un punto interior de A. Diremos que f tiene un máximo local en el punto $x = a \in A$ si en todo un entorno $(a - \delta, a + \delta)$, $\delta > 0$, de $x = a$, $f(x) \leq f(a)$. Diremos que una función $f(x)$ tiene un mínimo local en el punto $x = a$ si en todo un entorno $(a-\delta, a+\delta)$, $\delta > 0$, de $x = a$, $f(x) \geq f(a)$. Si una función f tiene un máximo o un mínimo local en $x = a$, diremos que f tiene un extremo local en el punto $x = a$.*

En la figura 1.1.5 se muestra un máximo local en $x = c$ y un mínimo local en $x = d$.

Resulta que la derivada es una herramienta muy poderosa a la hora de estudiar los extremos locales de una función. EL siguiente resultado sa una condición necesaria de extremo:

Teorema 1.1.18 *(Lema de Fermat) Sea $f : A \subset \mathbb{R} \mapsto \mathbb{R}$ y sea a un punto interior de A. Si f es derivable en un entorno de $x = a$ y tiene un extremo local en $x = a$, entonces, $f'(a) = 0$.*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

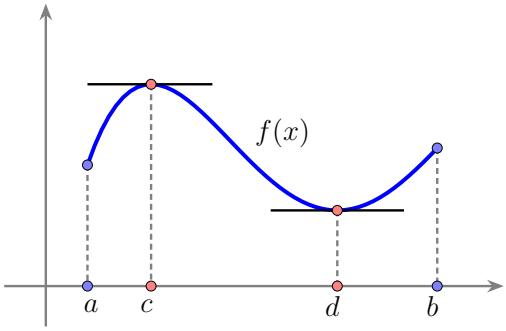


Figura 1.1.5: Interpretación geométrica del Lema de Fermat. Sea $f : [a, b] \mapsto \mathbb{R}$ una función diferenciable que un un máximo local en $x = c$ y un mínimo local en $x = d$. Como por el Lema de Fermat $f'(c) = f'(d) = 0$, entonces la recta tangente a f en ambos puntos es paralela al eje x.

Como se puede ver en la figura 1.1.5 la función f es creciente en (a, c) y (d, b), por tanto la Proposición 1.1.16 nos dice que en dichos intervalos $f'(x) \geq 0$, mientras que en (c, d) f es decreciente luego $f'(x) \leq 0$. Nótese además que en $x = c$, f tiene un máximo local mientras que en $x = d$ tiene un mínimo local.

Además se tiene el siguiente importante teorema:

Teorema 1.1.19 *(Teorema del valor medio de Lagrange) Sea la función $f(x) : [a, b] \mapsto \mathbb{R}$, continua en todo el intervalo cerrado y acotado $[a, b]$ y derivable en el intervalo abierto (a, b). Entonces, existe un c en el interior de del intervalo $[a, b]$, $c \in (a, b)$, tal que $f(b) - f(a) = f'(c)(b - a)$.*

Como consecuencia del Teorema del valor medio de Lagrange 1.1.19 se tienen, entre otras, las siguientes dos propiedades:

1. Sea $f(x) : [a, b] \mapsto \mathbb{R}$ una función continua en $[a, b]$ y derivable en (a, b). Si la función $f'(x) > 0$ $(f'(x) \geq 0)$ en (a, b), entonces f es creciente (no decreciente) en (a, b) y si $f'(x) < 0$ $(f'(x) \leq 0)$ en (a, b), entonces $f(x)$ es decreciente (no creciente) en todo (a, b).

2. Si $f(x)$ continua en $[a, b]$, derivable en (a, b) y $f'(x) = 0$ para todo x del intervalo (a, b), entonces $f(x) = const.$ Dado que si f es constante $f'(x) = 0$, entonces podemos afirmar que f es constante en $[a, b]$ si y solo si $f'(x) = 0$ para todo $x \in (a, b)$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Sea la función $f : A \subset \mathbb{R} \mapsto \mathbb{R}$. Si existe la derivada de $f(x)$ en todo un intervalo (que asumiremos abierto) $(a, b) \subset A$ podemos definir una nueva función $f'(x)$ que coincide con la primera derivada de f en dicho intervalo. Definiremos la segunda derivada, que denotaremos por $f''(x_0)$, de una función en un punto $x = x_0 \in (a, b)$ a la derivada de la función $f'(x)$ en $x = x_0$, o sea,

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

Si existe la segunda derivada para todo x comprendida en un intervalo (que asumiremos abierto) $B \subset (a, b)$ podemos definir la función segunda derivada de f en dicho intervalo, y a partir de ella definir la tercera derivada, y así sucesivamente. Así, si existen las derivadas hasta orden $n - 1$, $f^{(n-1)}(x)$, en un abierto $D \subset A$, podemos definir la n-ésima derivada de f en $x_0 \in D$, que denotaremos por $f^{(n)}(x)$, mediante la fórmula

$$f^{(n)}(x_0) = \lim_{h \to 0} \frac{f^{(n-1)}(x_0 + h) - f^{(n-1)}(x_0)}{h}.$$

La segunda derivada permite encontrar una condición muy útil para determinar si una función dos veces derivable tiene o no extremos locales en un punto $x = a$.

Teorema 1.1.20 *(Condición suficiente de extremo) Sea $f : A \mapsto \mathbb{R}$ una función dos veces derivable con segunda derivada continua en un entorno del punto $x = a$ tal que $f'(a) = 0$, entonces la función tendrá en $x = a$ un máximo local si $f''(a) < 0$ y un mínimo local si $f''(a) > 0$.*

Por ejemplo, si $f(x) = x^3 + 3x^2 - 2$, $f'(x) = 3x(x + 2) = 0$, luego el Lema de Fermat nos dice que en $x = 0$ y $x = -2$ son posibles extremos de f. Como $f''(0) > 0$ y $f''(-2) < 0$, entonces en $x = -2$ hay un máximo local y en $x = 0$ un mínimo local.

Las derivadas de orden superior nos permiten definir el polinomio de Taylor de una función.

Definición 1.1.21 *Dada una función $f(x)$ n-veces derivable en un entorno de $x = a$, llamaremos polinomio de Taylor de orden n de la función $f(x)$, y lo denotaremos por $P_n(x, a)$, al polinomio*

$$P_n(x, a) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x - a)^k. \tag{1.1.2}$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Nótese que el polinomio de Taylor de una función n-veces derivable es un polinomio de grado a lo más n. El siguiente teorema nos muestra el error que se comete al aproximar una función $n+1$ derivable usando su correspondiente polinomio de Taylor.

Teorema 1.1.22 *(Teorema de Taylor) Supongamos que la función $f(t)$ es n-veces derivable en el intervalo*[2] *$[a,x]$ y con las $n-$ésimas derivadas continuas en $[a,x]$ y supongamos que existe la derivada $n+1$ en (a,x) y sea $P_n(x,a)$ el polinomio de Taylor de la función $f(x)$ definido por (1.1.2). Entonces,*

$$f(x) = P_n(x,a) + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}, \quad c \in (a,x). \tag{1.1.3}$$

Teorema 1.1.23 *(Teorema local de Taylor) Supongamos que la función $f(x)$ es $n-$veces derivable en un entorno de $x=a$ y sea $P_n(x,a)$ el polinomio de Taylor de grado n de la función $f(x)$ definido en (1.1.2). Entonces,*

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

Antes de continuar conviene recordar algunas definiciones.

Definición 1.1.24 *Dos funciones $f : A \mapsto \mathbb{R}$ y $g : A \mapsto \mathbb{R}$ se denominan infinitésimos equivalentes en $x=a$ si $\lim_{x\to a} f(x) = 0$, $\lim_{x\to a} g(x) = 0$ y $\lim_{x\to a} \frac{f(x)}{g(x)} = 1$ y se escribe $f(x) \sim g(x)$ cuando x tiende a a.*

Por ejemplo, se tienen los siguientes infinitésimos equivalentes en $x=0$:

1. $\operatorname{sen} x \sim x$,
2. $1 - \cos x \sim \dfrac{x^2}{2}$,
3. $(1+x)^\alpha - 1 \sim \alpha\, x$,
4. $e^x - 1 \sim x, \quad b^x - 1 \sim x \ln b$,
5. $\ln(1+x) \sim x, \quad \log_b(1+x) \sim x \log_b e$,
6. $\tan x \sim x$,

[2] si $a > x$ se escogen los intervalos $[x,a]$ y (x,a).

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

7. $\operatorname{arc\,sen} x \sim x$,

8. $\arctan x \sim x$.

Definición 1.1.25 *(o pequeña) Dados dos funciones f y g infinitesimales en cierto $x = a$, diremos que $g(x)$ es un infinitésimo de orden mayor que $f(x)$ en $x = a$ si $\lim_{x\to a} \frac{g(x)}{f(x)} = 0$ y se escribe $g(x) = o(f(x))$ cuando x tiende a a.*

Por ejemplo, la función x^2 es un infinitésimo de mayor orden que x en $x = 0$, es decir $x^2 = o(x)$, y la función x^3 es un infinitésimo de mayor orden que x^2 en $x = 0$, o sea, $x^3 = o(x^2)$.

A partir de la definición 1.1.25 se puede comprobar que

1. Para todo $m \in \mathbb{R}$, $m \cdot o(x) = o(x)$.

2. La suma de un número finito infinitésimos equivalentes es un infinitésimo del mismo orden.

3. El producto de un número finito de infinitésimos es un infinitésimo de orden superior.

Usando la notación de las "o pequeñas" el Teorema local de Taylor 1.1.23 se puede escribir de la forma

$$\lim_{x\to a} \frac{f(x) - P_n(x,a)}{(x-a)^n} = 0 \quad \Longleftrightarrow \quad f(x) - P_n(x,a) = o[(x-a)^n].$$

Para aproximar las funciones elementales se utilizan principalmente los polinomios de McLaurin $P_n(x)$ que no son más que casos particulares de los polinomios de Taylor cuando $a = 0$, o sea, $P_n(x) \equiv P_n(x,0)$. Así, se tiene que

1. $\operatorname{sen} x = \sum_{k=1}^{n} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} + o(x^{2n})$,

2. $\cos x = \sum_{k=0}^{n} (-1)^{k} \frac{x^{2k}}{(2k)!} + o(x^{2k+1})$,

3. $(1+x)^{\alpha} = 1 + \sum_{k=1}^{n} \frac{(\alpha)_k}{k!} x^k + o(x^n)$,

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

4. $e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + o(x^n)$,

5. $\ln(1+x) = \sum_{k=1}^{n} (-1)^{k+1} \frac{x^k}{k} + o(x^n)$.

Los desarrollos anteriores son muy útiles a la hora de calcular límites de funciones. Así, por ejemplo,

$$\lim_{x\to 0} \frac{x - \operatorname{sen} x}{x^3} = \lim_{x\to 0} \frac{x - x + x^3/6 + o(x^3)}{x^3} = \frac{1}{6} + \lim_{x\to 0} \frac{o(x^3)}{x^3} = \frac{1}{6}.$$

Nótese que si solo usamos que $\operatorname{sen} x = x + o(x)$, entonces

$$\lim_{x\to 0} \frac{x - \operatorname{sen} x}{x^3} = \lim_{x\to 0} \frac{x - x + o(x)}{x^3} = \lim_{x\to 0} \frac{o(x)}{x^3},$$

que es indeterminado (¿por qué?)[3]. Otro ejemplo es

$$\lim_{x\to 0} \frac{e^x - 1 - x}{1 - \cos x} = \lim_{x\to 0} \frac{1 + x + x^2/2 + o(x^2) - 1 - x}{1 - (1 - x^2/2 + o(x^2))} = \lim_{x\to 0} \frac{1 + o(x^2)/x^2}{1 + o(x^2)/x^2} = 1.$$

Para más detalles sobre el cálculo diferencial de una variable, así como para ver las demostraciones de los teoremas aquí enunciados, se recomienda consultar, por ejemplo, el volumen I de [3] y [8].

1.2. Espacios vectoriales

Vamos a definir el espacio $\mathbb{R}^n$ como el espacio de las n-tuplas (vectores) $x = (x_1, \cdots, x_n)$ con $x_k \in \mathbb{R}$ para $k = 1, \ldots, n$. Para $n = 1$ tenemos el conjunto $\mathbb{R}$ de los números reales. Para $n = 2$ tenemos el conjunto de los vectores del plano (x, y), para $n = 3$ el conjunto de los vectores del espacio (x, y, z), etc.

Definiremos la suma de dos vectores $x = (x_1, \cdots, x_n)$ e $y = (y_1, \cdots, y_n)$ de $\mathbb{R}^n$ como el vector $z = x + y = (x_1 + y_1, \cdots, x_n + y_n)$. Definiremos el producto de un escalar (número real) λ por un vector x de $\mathbb{R}^n$ al vector $z = \lambda \cdot x = (\lambda x_1, \cdots, \lambda x_n)$.

Es fácil comprobar que $\mathbb{R}^n$ es un espacio vectorial, es decir se cumple la siguiente definición:

[3] Nótese que $x^2 = o(x)$, en cuyo caso $o(x)/x^3 = 1/x$ que no tiene límite en cero. Por otro lado también $x^4 = o(x)$, en cuyo caso $o(x)/x^3 = x$ que tiende a cero. Pero también $x^3 = o(x)$ que nos conduciría al valor $o(x)/x^3 = 1/3$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Definición 1.2.1 *Sea $\mathbb{V}$ un conjunto de elementos cualesquiera y $\mathbb{K}$ el cuerpo de los números reales $\mathbb{R}$ o complejos $\mathbb{C}$. Definiremos en $\mathbb{V}$ las operaciones* suma "+" *de dos elementos x, y de $\mathbb{V}$ y* multiplicación "$\cdot$" *de un elemento de $\mathbb{V}$ por un número (real o complejo) $\alpha \in \mathbb{K}$ por un elemento de $\mathbb{V}$. Diremos que $\mathbb{V}$ es un espacio vectorial sobre $\mathbb{K}$ (real o complejo si $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$, respectivamente), si se cumplen las siguientes propiedades (axiomas):*

1. *Para todos x e y,* vectores *de $\mathbb{V}$, el vector suma, $w = x + y$, también es un vector de $\mathbb{V}$ y para todos $x, y, z \in \mathbb{V}$ se cumple que:*
   a) $x + y = y + x$
   b) $(x + y) + z = x + (y + z)$
   c) *Existe un elemento "nulo" de $\mathbb{V}$, tal que $x + 0 = 0 + x = x$*
   d) *Cualquiera sea el vector x de $\mathbb{V}$, existe el elemento $(-x)$ "opuesto" a x, tal que $x + (-x) = (-x) + x = 0$.*
2. *Para todo x* vector *de $\mathbb{V}$, el vector que se obtiene al multiplicar por un escalar, $w = \alpha \cdot x$, también es un vector de $\mathbb{V}$ y para todos $x, y \in \mathbb{V}$, $\alpha, \beta \in \mathbb{K}$ se cumple que:*
   a) $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$
   b) $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$
   c) $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$
   d) $1 \cdot x = x$

Definición 1.2.2 *Sea $\mathbb{V}$ un espacio vectorial. Diremos que un subconjunto $H \subset \mathbb{V}$ de elementos de $\mathbb{V}$ es un subespacio vectorial de $\mathbb{V}$ si H es a su vez un espacio vectorial respecto a las mismas operaciones* suma "+" *y* multiplicación "$\cdot$" *que $\mathbb{V}$.*

Teorema 1.2.3 *Un subconjunto H de elementos de $\mathbb{V}$ es un subespacio vectorial de $\mathbb{V}$ si y sólo si se cumple que para todos x e y,* vectores *de H y α, $\beta \in \mathbb{K}$ el vector $w = \alpha x + \beta y$ también es un vector de H.*

Definamos ahora la envoltura lineal $\text{span}\,(v_1, v_2, ..., v_p)$ de los vectores $v_1, v_2, ..., v_p$ como el conjunto de todas las combinaciones lineales de dichos vectores:

$$\text{span}\,(v_1, v_2, ..., v_p) = \left\{ \sum_{k=1}^{p} \alpha_k v_k \;\middle|\; \alpha_k \in \mathbb{K},\ k = 1, 2, \ldots, p \right\}.$$

Usando el teorema anterior se deduce el siguiente

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Teorema 1.2.4 *Dado un conjunto de vectores* $\{v_1, v_2, ..., v_p\}$ *de un espacio vectorial* $\mathbb{V}$*, el conjunto* $\text{span}\,(v_1, v_2, ..., v_p)$ *es un subespacio vectorial de* $\mathbb{V}$*. Dicho subespacio vectorial comúnmente se denomina subespacio generado por los vectores* $v_1, v_2, ..., v_p$*.*

Un conjunto de vectores $v_1, v_2, ..., v_p$ de un espacio vectorial $\mathbb{V}$ se denomina linealmente independiente si la ecuación vectorial

$$x_1 v_1 + x_2 v_2 + \cdots + x_p v_p = 0,$$

tiene como única solución la trivial $x_1 = \cdots = x_p = 0$.

Un conjunto de vectores $v_1, v_2, ..., v_p$ se denomina linealmente dependiente si existen los valores $x_1, x_2, \cdots, x_p$ no todos iguales a cero tales que se verifique la ecuación vectorial

$$x_1 v_1 + x_2 v_2 + \cdots + x_p v_p = 0.$$

Se dice que un conjunto infinito de vectores es linealmente independiente si cualquier subsistema finito del mismo es linealmente independiente. En caso contrario se dice que el sistema es dependiente.

Las siguientes propiedades se pueden verificar fácilmente:

1. Un conjunto $S = \{v_1, v_2, ..., v_p\}$ de dos o más vectores es linealmente dependiente si y sólo si al menos uno de los vectores del conjunto es combinación lineal de los demás.

2. Un conjunto $S = \{v_1, v_2, ..., v_p\}$ de dos o más vectores de $\mathbb{V}$ con alguno de los vectores $v_i = 0$ $(1 \leq i \leq p)$ es necesariamente un conjunto de vectores linealmente dependientes.

Los vectores linealmente independientes de un espacio vectorial juegan un papel fundamental en el estudio de los sistemas lineales gracias a la siguiente definición:

Definición 1.2.5 *Dado un subespacio vectorial* H *del espacio vectorial* $\mathbb{V}$ *diremos que el conjunto de vectores* $B = \{b_1, b_2, ..., b_p\}$ *de* $\mathbb{V}$ *es una base de* H *si*

i) B *es un conjunto de vectores linealmente independientes*

ii) $H = \text{span}\,(b_1, b_2, ..., b_p)$*, o sea,* B *genera a todo* H*.*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

En particular si H coincide con $\mathbb{V}$, entonces B es una base de todo el espacio vectorial $\mathbb{V}$.

Por ejemplo, si tomamos una matriz $n \times n$ invertible, entonces sus columnas $a_1, ..., a_n$ son linealmente independientes y además $\mathbb{R}^n = \text{span}\,(a_1, ..., a_n)$. Por tanto $B = a_1, ..., a_n$ es una base de $\mathbb{R}^n$. En particular, si $A = I_n$, la matriz identidad $n \times n$, las columnas $(e_k)_{k=1}^n$ de misma, o sea, los vectores

$$(1.2.1)\quad e_1 = (1, 0, 0, \ldots, 0, 0),\ e_2 = (0, 1, 0, \ldots, 0, 0),\ \ldots,\ e_n = (0, 0, 0, \ldots, 0, 1),$$

son una base de $\mathbb{R}^n$ la cual se conoce como base canónica de $\mathbb{R}^n$.

El siguiente teorema es de gran importancia en las aplicaciones.

Teorema 1.2.6 *Si un espacio vectorial $\mathbb{V}$ tiene una base de n vectores $B = \{b_1, b_2, ..., b_n\}$, entonces cualquier conjunto con más de n vectores de $\mathbb{V}$ es linealmente dependiente. Más aún, si un espacio vectorial $\mathbb{V}$ tiene una base de n vectores $B = \{b_1, b_2, ..., b_n\}$, entonces cualquier otra base de $\mathbb{V}$ tendrá que tener n vectores de $\mathbb{V}$.*

Por tanto el menor número de vectores linealmente independientes que generan un espacio vectorial es una propiedad intrínseca de dicho espacio. Dicho número se denomina *dimensión del espacio vectorial*.

Un espacio vectorial es de dimensión finita n si $\mathbb{V}$ está generado por una base de n elementos, es decir si $\mathbb{V} = \text{span}\,(b_1, ..., b_n)$, donde $B = \{b_1, ..., b_n\}$ es una base de $\mathbb{V}$ y lo escribiremos de la forma $\dim V = n$. En el caso que $\mathbb{V} = \{0\}$ sea el espacio vectorial nulo, $\dim\{0\} = 0$. Si $\mathbb{V}$ no puede ser generado por una base finita de vectores, entonces diremos que $\mathbb{V}$ es de dimensión infinita y lo denotaremos por $\dim V = \infty$.

Está claro que $\mathbb{R}^n$ es de dimensión finita y que $\dim \mathbb{R}^n = n$.

1.3. $\mathbb{R}^n$ como espacio normado y métrico

Definición 1.3.1 *Un espacio métrico es un par $(\mathbb{X}, \rho)$ donde $\mathbb{X}$ es un conjunto y $\rho := \rho(x, y)$ es una función real (univaluada) no negativa definida para todos $x, y, z \in \mathbb{X}$ tal que*

1. $\rho(x, y) = 0 \Longleftrightarrow x = y$,

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

2. $\rho(x,y) = \rho(y,x)$,

3. $\rho(x,z) \leq \rho(x,y) + \rho(y,z)$.

Si escogemos $\mathbb{X} = \mathbb{R}^n$, es decir el espacio de las n-tuplas $x = (x_1, x_2, \ldots, x_n)$ con la *métrica*

$$\rho(x,y) = \sqrt{\sum_{k=1}^{n} |x_k - y_k|^2}, \tag{1.3.1}$$

obtenemos un espacio métrico. Esta métrica se suele denominar métrica euclídea. De hecho, también son espacios métricos los espacios $\mathbb{R}^n$ con la p-métrica

$$\rho(x,y) = \left(\sum_{k=1}^{n} |x_k - y_k|^p\right)^{1/p}, \qquad p \geq 1, \tag{1.3.2}$$

y la métrica "infinita"

$$\rho(x,y) = \max_{k=1,\ldots,n} |x_k - y_k|, \tag{1.3.3}$$

respectivamente. Los espacios métricos anteriores se suelen denotar por $\mathbb{R}^n_p$ donde $p \in [1,\infty)$. En este curso nos restringiremos al espacio $\mathbb{R}^n_2$ que denotaremos simplemente por $\mathbb{R}^n$.

Para probar que (1.3.2) y (1.3.3) son dos métricas tenemos que comprobar los tres apartados de la definición 1.3.1. En ambos casos los dos primeros son evidentes, así como el tercero para la función (1.3.3). Para probar el tercer punto de la definición 1.3.1 en el caso de la función (1.3.2) necesitaremos probar el siguiente teorema:

Teorema 1.3.2 *(Desigualdad de Minkowski) Sea los números x_i e y_i, $i = 1, 2 \ldots, n$ no negativos. Entonces, para todo $p > 1$ se tiene*

$$\left(\sum_{i=1}^{n} (x_i + y_i)^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} y_i^p\right)^{\frac{1}{p}}, \tag{1.3.4}$$

donde la igualdad sólo tiene lugar si $x_i = c\, y_i$ para todo $i = 1, 2, \ldots, n$ (es decir si x_i y y_i son proporcionales).

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Demostración: Vamos a dar una prueba directa[4] de (1.3.4). El primer paso consiste en probar que para todos $u, v \geq 0$ y $1 \leq p < +\infty$ se tiene

$$(u+v)^p = \inf_{t\in(0,1)} \left\{ t^{1-p}u^p + (1-t)^{1-p}v^p; u, v \geq 0, 0 < t < 1 \right\}. \tag{1.3.5}$$

Para ello basta comprobar que el mínimo absoluto de la función $f : (0,1) \mapsto \mathbb{R}$, $f(t) = t^{1-p}u^p + (1-t)^{1-p}v^p$ con $u, v > 0$ y $1 < p < +\infty$ (en el caso de las igualdades es trivial) se alcanza en $t_* = u/(u+v)$ y $f(t_*) = (u+v)^p$. Sea $p > 1$. En primer lugar es fácil ver que, $f(t)$ es diferenciable y $f'(t) = (1-p)[= t^{-p}u^p + (-t)^{1-p}v^p]$. De lo anterior se sigue que $f'(t) = 0$ si y solo si $t = t^*$, es decir f tiene un único extremo, y dado que en los extremos del intervalo se tiene que (recordemos que $p > 1$), entonces $f(t) \to +\infty$ cuando $t \to 0+$ y $1+$, se sigue que este es un mínimo. Sustituyendo el valor t^* en f se obtiene el resultado.

Entonces, para todos x_i e y_i, usando (1.3.5) tenemos

$$\begin{aligned} |x_i + y_i|^p \leq (|x_i| + |y_i|)^p &= \inf_{t\in(0,1)} (t^{1-p}|x_i|^p + (1-t)^{1-p}|y_i|^p) \\ &\leq t^{1-p}|x_i|^p + (1-t)^{1-p}|y_i|^p, \end{aligned}$$

que, sumando en i, nos conduce a

$$\sum_{i=1}^n |x_i + y_i|^p \leq t^{1-p} \sum_{i=1}^n |x_i|^p + (1-t)^{1-p} \sum_{i=1}^n |y_i|^p.$$

Llamando $u^p = \sum_{i=1}^n |x_i|^p$, $v^p = \sum_{i=1}^n |y_i|^p$ y tomando el ínfimo en $t \in (0,1)$ en la desigualdad anterior y usando nuevamente (1.3.5) se sigue que

$$\sum_{i=1}^n |x_i + y_i|^p \leq \left(\left(\sum_{i=1}^n x_i^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n y_i^p \right)^{\frac{1}{p}} \right)^p,$$

de donde se obtiene la desigualdad (1.3.4). □

Como consecuencia de la desigualdad de Minkowski (1.3.4) se sigue fácilmente la desigualdad triangular (propiedad 3) de la definición 1.3.1, para ello basta elegir $x_i = x_i - z_i$ y $y_i = z_i - y_i$ en (1.3.4).

Definición 1.3.3 *Un espacio vectorial $\mathbb{X}$ se denomina espacio normado si $\forall x \in \mathbb{X}$ existe un número real denominado norma, y que denotaremos por $\|x\|$, que cumple con las condiciones*

[4] Esta prueba se debe a Heinz König (A simple proof of the Minkowski inequality. In: Walter, W. (eds) *General Inequalities 6. International Series of Numerical Mathematics*, Vol. 103. Birkhäuser, Basel, 1992, pág. 469).

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

1. *Para todo* $x \in \mathbb{X}$, $\|x\| \geq 0$ *y si* $\|x\| = 0$ *entonces* $x = 0$.

2. *Para todo* $x \in \mathbb{X}$ *y* $\lambda \in \mathbb{R}$, $\|\lambda x\| = |\lambda|\|x\|$,

3. *Para todos* $x, y \in \mathbb{X}$ *se tiene la desigualdad triangular*

$$\|x + y\| \leq \|x\| + \|y\|. \tag{1.3.6}$$

Es evidente que si en un espacio normado $\mathbb{X}$ definimos la función $\rho(x, y) = \|x-y\|$, esta satisface los axiomas de la definición 1.3.1, i.e., todo espacio normado es un espacio métrico. La función ρ anterior se denomina métrica inducida por la norma. En $\mathbb{R}^n$ tenemos las normas:

$$\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p}, \quad p \geq 1, \qquad \|x\|_\infty = \max_{k=1,\ldots,n} |x_k|, \tag{1.3.7}$$

denominadas habitualmente como norma p y norma infinito (o del máximo), respectivamente. El caso $p = 2$, corresponde a la norma euclídea que induce la métrica euclídea (1.3.1).

Probar los primeros dos puntos de la definición 1.3.3 para la normas p e infinito es inmediato, como lo es el punto tres para la norma infinito. La desigualdad triangular en el caso de la norma p se sigue directamente de la desigualdad Minkowski (1.3.4).

Un ejemplo de norma muy distinta a las anteriores es la definida en el siguiente ejercicio:

Ejercicio:[5] Sea $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$ y sea el polinomio $p_{n-1}(t, x) = x_1 + x_2 t + \cdots + x_n t^{n-1}$. Sea $w(t) \geq 0$ una función continua (se pueden elegir funciones mucho más generales) no idénticamente nula en un intervalo (a, b) tal que existan las integrales

$$\int_a^b t^k w(t) dt, \quad k = 0, 1, \ldots, n-1.$$

Prueba que entonces la función

$$\|x\| = \int_a^b |p_{n-1}(t, x)| w(t) dt$$

define una norma en $\mathbb{R}^n$.

[5] Este ejercicio está inspirado en un ejemplo encontrado por D. Antonio Medinilla Garófano.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Definición 1.3.4 *Por aplicación (operador) o función entenderemos una regla T que le hace corresponder a cada elemento del subconjunto $\mathcal{D}(T) \subset \mathbb{X}$ un único elemento del espacio métrico $\mathbb{Y}$. Así, $T : \mathbb{X} \mapsto \mathbb{Y}$, $y = Tx$ o $y = T(x)$, donde $x \in \mathcal{D}(T) \subset \mathbb{X}$ e $y \in \mathbb{Y}$. Al conjunto $\mathcal{D}(T) \subset \mathbb{X}$ se le denomina dominio de la aplicación.*

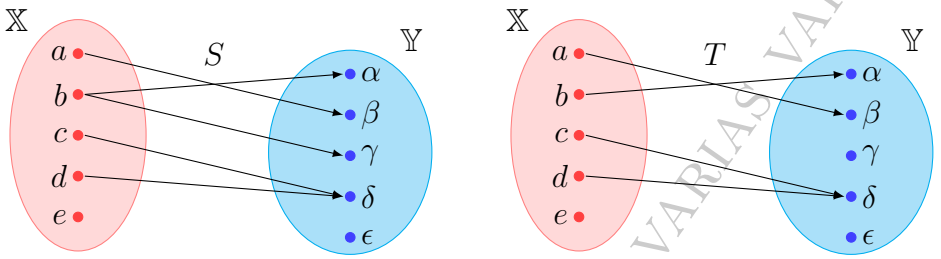


Figura 1.3.1: $T : \mathbb{X} \mapsto \mathbb{Y}$ (esquema de la derecha) es una aplicación pero S (a la izquierda) no lo es.

Definición 1.3.5 *Si a cada $x \in \mathcal{D}(T)$ le corresponde un valor $y = Tx \in \mathbb{Y}$ diremos que Tx es la imagen de x según T. Al conjunto de todas las imágenes Tx le denominaremos imagen de T y le denotaremos por $\mathcal{I}(T)$.*

Definición 1.3.6 (Composición de aplicaciones) *Sean $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathcal{I}(T)\mathbb{Y}$ y $U : \mathcal{D}(U) \subset \mathbb{Y} \mapsto \mathcal{I}(U)\mathbb{Z}$ dos aplicaciones tales que $\mathcal{I}(T) \subset \mathcal{D}(U)$. Entonces definiremos la aplicación $U \circ T : \mathbb{X} \mapsto \mathbb{Z}$ y la denominaremos aplicación compuesta de U y T a la aplicación que le hace corresponder a cada $x \in \mathcal{D}(T) \subset \mathbb{X}$ un elemento $z \in \mathbb{Z}$ tal que $z = U(Tx)$ $(z = U\,Tx)$.*

En general $UTx \neq TUx$. Más aún que exista $U \circ T$ no implica que exista $T \circ U$.

Definición 1.3.7 *Se dice que una aplicación $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$, $\mathbb{X}$ e $\mathbb{Y}$ espacios métricos, es continua en $x_0 \in \mathcal{D}(T)$ si para todo $\epsilon > 0$, existe un $\delta > 0$ tal que $\forall x \in \mathcal{D}(T)$ con $\rho(x, x_0) < \delta$ es tal que[6] $\sigma(Tx, Tx_0) < \epsilon$. Se dice que T es continua en todo $M \subset \mathcal{D}(T)$ si T es continua en todo $x \in M$.*

Nota 1.3.8 *La definición anterior es equivalente a decir que para toda sucesión $(x_n)_n \in \mathbb{X}$ con $x_n \overset{n\to\infty}{\longrightarrow} x_0 \in \mathbb{X}$, entonces $Tx_n \overset{n\to\infty}{\longrightarrow} Tx$.*

[6] Aquí ρ denota la métrica de $\mathbb{X}$ y σ la de $\mathbb{Y}$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

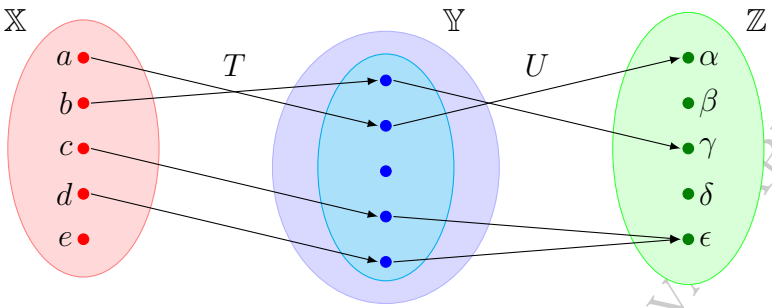


Figura 1.3.2: Composición de funciones $T \circ U$ de $T : \mathbb{X} \mapsto \mathcal{I}(T) \subset \mathbb{Y}$ y $U : \mathcal{D}(U) \subset \mathbb{Y} \mapsto \mathbb{Z}$, donde $\mathcal{I}(T) \subset \mathcal{D}(U)$.

En todo espació métrico (y por tanto en $\mathbb{R}^n$) podemos definir una *topología*. para ello necesitamos el concepto de bola abierta.

Definición 1.3.9 *Sea $\mathbb{X}$ un espacio métrico, $x_0 \in \mathbb{X}$ y $r > 0$. Definiremos la bola abierta $B(x_0, r)$ al conjunto $B(x_0, r) = \{x \in \mathbb{X};\ \rho(x_0, x) < r\}$.*

Definición 1.3.10 *Se dice que el conjunto $M \subset \mathbb{X}$ es abierto en $\mathbb{X}$ si todos sus puntos (elementos) se pueden encerrar en una bola abierta contenida completamente en $\mathbb{X}$. Un conjunto $M \subset \mathbb{X}$ es cerrado en $\mathbb{X}$ si es su complementario en $\mathbb{X}$, $\mathbb{X} \backslash M$ es abierto.*

Las bolas abiertas $B(x_0, \epsilon)$ se suelen denominar ϵ-vecindades (o entornos) de x_0. Es evidente que toda ϵ-vecindad de x_0 contiene al propio x_0.

Definición 1.3.11 *Un punto x_0 se denomina punto interior del conjunto $M \subset \mathbb{X}$ si existe un $\epsilon > 0$ tal que $B(x_0, \epsilon) \subset M$ (ver figura 1.3.4).*

De la definición anterior se deduce que el conjunto $M \subset \mathbb{X}$ es abierto si y sólo si todos sus puntos son interiores.

A partir de las definiciones anteriores se tiene la siguiente proposición cuya prueba omitiremos:

Proposición 1.3.12 *Sea Σ en conjunto de todos los subconjuntos abiertos de $\mathbb{X}$. Entonces*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

1. $\emptyset \in \Sigma$, $\mathbb{X} \in \Sigma$,

2. *la unión (finita o infinita) de subconjuntos abiertos de* $\mathbb{X}$ *es abierto: Si* U_k, $k = 1, 2, \ldots$ *son abiertos,* $\bigcup_k U_k \in \Sigma$

3. *La intersección de un número finito de abiertos es abierto: Si* U_k, $k = 1, 2, \ldots, n$ *son abiertos,* $\bigcap_{k=1}^{n} U_k \in \Sigma$.

Las tres propiedades anteriores son de extrema importancia. Tal es así que ellas definen un tipo de espacios muy generales: Los espacios topológicos. Así, el par, dados un conjunto $\mathbb{X}$ y una colección Σ de subconjuntos de $\mathbb{X}$, $(\mathbb{X}, \Sigma)$ se denomina espacio topológico si Σ cumple con los *axiomas* (propiedades) 1, 2 y 3 de la proposición anterior. Al conjunto Σ se le denomina *topología* de $\mathbb{X}$. Así pues, todo espacio métrico es un espacio topológico.

Está claro que una bola $B(x_0, r)$ es un conjunto abierto. Para probarlo basta probar que para cualquier $x \in B(x_0, r)$ existe una bola $B(x, \delta) \subset B(x_0, r)$. Para ello elegimos $\delta = r - \rho(x, x_0)$. Entonces para cualquiera $\xi \in B(x, \delta)$, $\rho(\xi, x) < \delta < r - \rho(x, x_0)$. Entonces $\rho(\xi, x_0) \leq \rho(\xi, x) + \rho(x, x_0) < \delta + (r - \delta) = r$, o sea, $\rho(\xi, x_0) < r$. Luego $B(x, \delta) \subset B(x_0, r)$. Como ejercicio probar que el conjunto $U(x_0, r) := \mathbb{X} \setminus S(x_0, r) = \{x \in \mathbb{X}, \rho(x, x_0) > r\}$ también es abierto y por tanto $S(x_0, r)$ es cerrado. Por otro lado, definamos el conjunto $\Lambda(x_0, r) := \{x \in \mathbb{X}, \rho(x, x_0) = r\}$. Está claro que $\mathbb{X} \setminus \Lambda(x_0, r) = U(x_0, r) \bigcup B(x_0, r)$ que es abierto, por lo que $\Lambda(x_0, r)$ es cerrado.

Antes de continuar vamos a mostrar como es la *geometría* de algunos espacios métricos. Como ejemplo tomaremos $\mathbb{X} = \mathbb{R}^2$ y escogeremos distintas métricas. En la figura 1.3.3 están representados las bolas $B(0, 1)$ para distintas métricas en $\mathbb{R}^2$.

Definición 1.3.13 *Sea* $M \subset \mathbb{X}$. *Diremos que* $x \in \mathbb{X}$ *es un punto de contacto (o adherente) de* M *si en cualquier bola* $B(x, \epsilon)$, $\epsilon > 0$ *hay al menos un elemento de* M. *Así mismo, diremos que* x *es un punto de acumulación (o punto límite) de* M *si en cualquier bola* $B(x, \epsilon)$, $\epsilon > 0$ *hay al menos un elemento de* M *distinto de* x, *o equivalentemente, en cada bola* $B(x, \epsilon)$, $\epsilon > 0$ *hay infinitos elementos de* M. *Un punto* x *se denomina aislado de* M *si existe una bola* $B(x, \epsilon)$, $\epsilon > 0$ *que no contiene ningún elemento* M *excepto el propio* x.

Definición 1.3.14 *Dado un subconjunto* $M \in \mathbb{X}$, *se denomina clausura de* M *al conjunto* $\overline{M}$ *formado por todos los puntos de contacto de* M.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

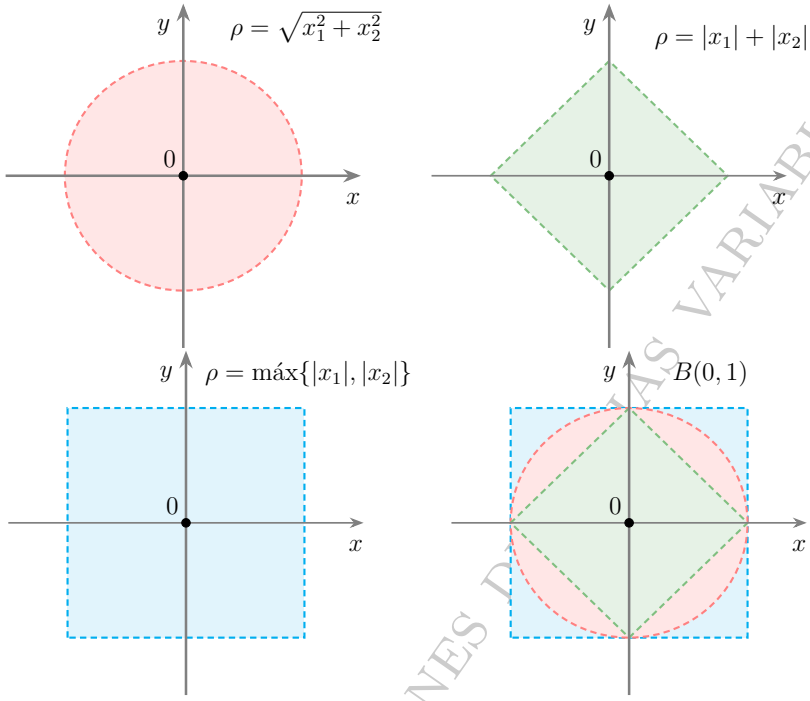


Figura 1.3.3: Entornos $B(0,1)$ según las métricas (1.3.2) para $p = 2$ (encima a la izquierda), $p = 1$ (encima a la derecha) y la métrica (1.3.3) (debajo a la izquierda). Debajo a la derecha se muestran los tres entornos en una única figura. Las distancias ρ corresponden a la distancia entre los puntos $x = (x_1, x_2)$ y $x_0 = (0,0)$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Nótese que si $x \in M$, entonces obviamente x es un punto de contacto de M. En particular, todo punto interior de M es un punto de contacto. Además, todo punto interior es obviamente un punto de acumulación. Por otro lado, x es un punto de acumulación de M si y sólo si es un punto de contacto de $M \setminus \{x\}$. De lo anterior se deduce que los puntos de contacto de M o bien son puntos límites, o bien son aislados. De hecho, si $x \in \mathbb{X}$ pero no es de acumulación, entonces es aislado.

De lo anterior se sigue que $\overline{M} = M \cup \{\text{conjunto de sus puntos límites}\}$, que usaremos en adelante como definición de clausura de M.

Por ejemplo, si $\mathbb{X} = \mathbb{Q}$, entonces $\overline{\mathbb{Q}} = \mathbb{R}$ pues todo $x \in \mathbb{R}$ es un punto límite de $\mathbb{Q}$ (¿por qué?).

Proposición 1.3.15 *Un conjunto $M \subset \mathbb{X}$ es cerrado en $\mathbb{X}$ si y sólo si $M = \overline{M}$.*

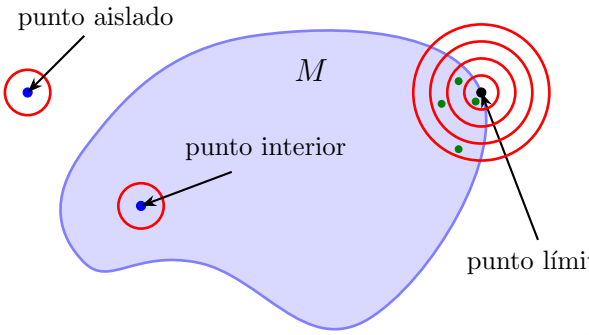


Figura 1.3.4: Esquema donde se muestra un punto interior, un punto límite y un punto aislado del conjunto M. Los círculos rojos representan los entornos (bolas) centrados en los diferentes.

Demostración: $\Rightarrow$ Sea M cerrado. Como $M \subset \overline{M}$ basta probar que $M \supset \overline{M}$. Como M es cerrado, entonces $\mathbb{X} \setminus M$ es abierto así que para todo $x \in \mathbb{X} \setminus M$ existe una bola $B(x,r)$ completamente contenida en $\mathbb{X} \setminus M$, i.e., $B(x,r)$ no contiene puntos de M, luego $x \notin \overline{M}$ (pues si $x \in \overline{M}$ en cualquier bola $B(x,r)$ tendría que haber al menos un x' de M distinto de x lo cual contradice lo anterior), es decir, $x \in \mathbb{X} \setminus \overline{M}$, así que $\mathbb{X} \setminus M \subset \mathbb{X} \setminus \overline{M}$, i.e. $\overline{M} \subset M$.

$\Leftarrow$ Asumamos que $M = \overline{M}$. Probaremos que $\mathbb{X} \setminus M$ es abierto. Sea $x \notin M$, entonces $x \notin \overline{M}$. Entonces $x \in \mathbb{X} \setminus M$. Como $x \notin \overline{M}$ entonces x no es un punto adherente de M así que debe existir al menos una bola $B(x,r)$ que no contiene a ningún elemento de M, es decir $x \in \mathbb{X} \setminus M$ es un punto interior del conjunto $\mathbb{X} \setminus M$. Como x es arbitrario, $\mathbb{X} \setminus M$ es abierto luego M es cerrado. □

Dado que $M \subset \overline{M}$, entonces, de la proposición anterior se deduce que $\overline{M}$ es el menor conjunto cerrado que contiene a M.

Definición 1.3.16 *Un subconjunto $M \subset \mathbb{X}$ es acotado si su diámetro $d(M) = \sup_{x,y \in M} \rho(x,y)$ es es finito.*

Definición 1.3.17 *Dada una sucesión $(x_n)_n$ de elementos de $\mathbb{X}$, diremos que $(x_n)_n$ es acotada si existe un subconjunto $M \subset \mathbb{X}$ acotado tal que $x_n \in M$ para todo $n \in \mathbb{N}$.*

Lo anterior es equivalente a que exista un $x \in \mathbb{X}$ y un número $K > 0$ tal que $\rho(x, x_n) < K$ para todo $n \in \mathbb{N}$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Definición 1.3.18 *Una sucesión $(x_n)_n$ de elementos de $\mathbb{X}$ es convergente, y lo denotaremos por $\lim_{n\to\infty} x_n = x$, si existe un $x \in \mathbb{X}$ tal que para todo $\epsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$, $\rho(x, x_n) < \epsilon$. En caso contrario diremos que $(x_n)_n$ es divergente.*

La siguiente propiedad, cuya demostración se deja como ejercicio, es de gran utilidad:

Teorema 1.3.19 *Sea M un subespacio no vacío de un espacio métrico $\mathbb{X}$, y sea $\overline{M}$ su clausura. Entonces*

1. *$x \in \overline{M}$ si y sólo si existe una sucesión $(x_n)_n$ de elementos de M (es decir, $x_n \in M$, para todo n) tal que $\lim_{n\to\infty} x_n = x$.*
2. *M es cerrado si y sólo si $\lim_{n\to\infty} x_n = x$ implica que $x \in M$.*

Una propiedad muy útil en los espacios normados (recuérdese que los espacios normados también son métricos con una métrica inducida por la norma) es la continuidad de la norma.

Lema 1.3.20 *Sea $\mathbb{X}$ un espacio normado. Entonces la norma $\|\cdot\| : \mathbb{X} \mapsto \mathbb{R}$ es una aplicación continua.*

Demostración: Usando que

$$\|x_n\| = \|x_n - x + x\| \leq \|x\| + \|x_n - x\| \text{ y } \|x\| = \|x - x_n + x_n\| \leq \|x_n\| + \|x_n - x\|,$$

se tiene que

$$|\|x_n\| - \|x\|| \leq \|x_n - x\|,$$

de donde se sigue que, si $x_n \overset{n\to\infty}{\longrightarrow} x$, entonces $\|x_n\| \overset{n\to\infty}{\longrightarrow} \|x\|$, es decir, la norma es una aplicación continua. □

Ejercicio: Prueba que también la métrica es una una aplicación continua.

Definición 1.3.21 *Una sucesión $(x_n)_n$ de elementos de $\mathbb{X}$ se denomina de Cauchy o fundamental si para todo $\epsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $n > N$ y todo $p \in \mathbb{N}$, $\rho(x_n, x_{n+p}) < \epsilon$.*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Una sucesión en $\mathbb{R}^n$ es de Cauchy si y sólo si lo son sus componentes, por tanto, en $\mathbb{R}^n$ toda sucesión es convergente si y sólo si es de Cauchy. Esta propiedad fundamental de $\mathbb{R}^n$ no es cierta para cualquier espacio métrico $\mathbb{X}$. Por ejemplo, sea el espacio métrico $(0, 1)$ con la métrica de $\mathbb{R}$. Está claro que $x_n = 1/n$ es de Cauchy, sin embargo, x_n no tiene límite (tendría que ser 0, pero este no pertenece al espacio). Así, tenemos la siguiente definición:

Definición 1.3.22 *Un espacio métrico $\mathbb{X}$ se denomina completo si y sólo si toda sucesión de Cauchy de elementos de $\mathbb{X}$ es convergente.*

Obviamente en los espacios normados podemos definir la convergencia de sucesiones, sucesiones de Cauchy, etc. Basta considerarlos como espacios métricos con la métrica ρ inducida por la norma: $\rho(x, y) = \|x - y\|$.

Definición 1.3.23 *Un espacio normado completo (en la métrica inducida por la norma) se denomina espacio de Banach.*

Definición 1.3.24 *Un subconjunto $M \subset \mathbb{X}$ es denso en $\mathbb{X}$ si su clausura $\overline{M} = \mathbb{X}$.*

De la definición anterior se infiere que si M es denso en $\mathbb{X}$ entonces cualquiera sea la bola $B(x, \epsilon)$ (por pequeño que sea $\epsilon > 0$) siempre contiene puntos de M. En otras palabras, cualquiera sea $x \in \mathbb{X}$, siempre tiene elementos de M tan cerca como se quiera. Por ejemplo $\mathbb{Q}$ es denso en $\mathbb{R}$.

Definición 1.3.25 *Un espacio métrico $\mathbb{X}$ es separable si contiene un subespacio numerable*[7] *$M \subset \mathbb{X}$ denso en $\mathbb{X}$.*

Así pues, $\mathbb{R}$ es separable pues $\mathbb{Q}$ es numerable y denso en $\mathbb{R}$. Como consecuencia $\mathbb{R}^n$ es también separable, cuya prueba dejamos como ejercicio al lector.

[7] Un conjunto M cualquiera se denomina *numerable* si se puede poner en correspondencia biunívoca con $\mathbb{N} = \{1, 2, 3, \dots\}$. Es decir, existe una correspondencia biunívoca entre los elementos de M y los números naturales. Por ejemplo, $\mathbb{Q}$ es numerable, pero $\mathbb{R}$ no lo es.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

1.4. Espacios normados de dimensión finita

Dado que $\mathbb{R}^n$ es un espacio normado de dimensión finita, veamos con algo más de detalle algunas de las propiedades de dichos espacios.

Comenzaremos probado un lema técnico que será de gran utilidad en adelante.

Lema 1.4.1 *Sean n vectores cualesquiera $x_1, \ldots, x_n$ linealmente independientes de un espacio normado $\mathbb{X}$. Entonces, existe un número real $c > 0$ tal que cuales quiera sean los escalares $\alpha_1, \ldots, \alpha_n$,*

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geq c(|\alpha_1| + \cdots + |\alpha_n|). \tag{1.4.1}$$

Demostración: Sea $s = |\alpha_1| + \cdots + |\alpha_n|$. Si $s = 0$ el lema es trivial así que asumiremos $s > 0$. Dividiendo por s (1.4.1) se sigue que (1.4.1) es equivalente a probar que si $x_1, \ldots, x_n$ son linealmente independientes, entonces existe un número real $c > 0$ tal que cuales quiera sean los números $\beta_1, \ldots, \beta_n$, con $\sum_{k=1}^n |\beta_k| = 1$

$$\|\beta_1 x_1 + \cdots + \beta_n x_n\| \geq c.$$

Supongamos que la desigualdad anterior es falsa. Entonces ha de existir (¿por qué?) una sucesión $(y_m)_m \subset \mathbb{X}$ tal que

$$y_m = \beta_1^{(m)} x_1 + \cdots + \beta_n^{(m)} x_n, \quad \sum_{k=1}^n |\beta_k^{(m)}| = 1, \text{ y } \|y_m\| \overset{m\to\infty}{\longrightarrow} 0.$$

De la condición $\sum_{k=1}^n |\beta_k^{(m)}| = 1$ se sigue que las n sucesiones numéricas $(\beta_k^{(m)})_m$, $k = 1, \ldots, n$, son acotadas. Sea la sucesión $(\beta_1^{(m)})_m$ acotada, entonces por el Teorema de Bolzano-Weierstrass 1.1.8 de ella se puede extraer una subsucesión convergente $\beta_1^{(m_j)} \overset{j\to\infty}{\longrightarrow} \beta_1$. Escojamos de cada una de las sucesiones restantes $(\beta_k^{(m)})_m$, $k = 2, \ldots, n$, las subsucesiones definidas por los índices m_j de antes. Entonces la sucesión $(\beta_2^{(m_j)})_j$ es acotada y por Bolzano-Weierstrass de ella se puede extraer una subsucesión convergente $\beta_2^{(j_l)} \overset{l\to\infty}{\longrightarrow} \beta_2$. Además, si escogemos los índices j_l definidos por esta sucesión, la subsucesión $(\beta_1^{(j_l)})_j \overset{l\to\infty}{\longrightarrow} \beta_1$ (¿por qué?). Continuando este proceso n veces tenemos que existe una sucesión de índices $(l_i)_i$ tales que $\beta_k^{(l_i)} \overset{i\to\infty}{\longrightarrow} \beta_k$ para todos los $k = 1, 2, \ldots, n$. Dicha sucesión de índices define una subsucesión $(y_{l_i})_i$ de $(y_m)_m$ tal que

$$y_{l_i} = \sum_{k=1}^n \beta_k^{(l_i)} x_k, \qquad \beta_k^{(l_i)} \overset{i\to\infty}{\longrightarrow} \beta_k.$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Luego

$$\sum_{k=1}^{n} |\beta_k| = 1, \quad \text{y} \quad \lim_{i\to\infty} y_{l_i} = \sum_{k=1}^{n} \beta_k x_k := y.$$

De lo anterior se sigue que no todos los β_k pueden ser ceros al mismo tiempo. Como los vectores $x_1, \ldots, x_n$ son linealmente independientes entonces $y \neq 0$ (¿por qué?). Ahora bien, como la norma es una aplicación continua ($\lim_n \|x_n\| = \|\lim_n x_n\|$), entonces se tiene

$$\lim_{i\to\infty} y_{l_i} = y \quad \Longrightarrow \quad \lim_{i\to\infty} \|y_{l_i}\| = \|y\|,$$

pero como $\|y_m\| \overset{m\to\infty}{\longrightarrow} 0$, entonces $\lim_{i\to\infty} \|y_{l_i}\| = 0$, luego $\|y\| = 0$ de donde se sigue que $y = 0$ lo cual es una contradicción. □

Como corolario del Lema 1.4.1 anterior tenemos el siguiente teorema de completitud:

Teorema 1.4.2 *Todo subespacio M de dimensión finita de un espacio normado es completo. En particular, todo espacio normado de dimensión finita es completo.*

Demostración: Sea $(y_m)_m$ una sucesión de Cauchy arbitraria de M. Sea $\dim M = n$. Entonces para todo $y_m \in M$, existen los números (únicos) $\alpha_k^{(m)}$, $k = 1, 2, \ldots, n$, tales que $y_m = \sum_{k=1}^{n} \alpha_k^{(m)} e_k$. Como $(y_m)_m$ es de Cauchy, entonces, por un lado

$$\forall \epsilon > 0, \ \exists N \in \mathbb{N}, \ | \ \forall m > N, \ \forall p \in \mathbb{N}, \quad \|y_m - y_{m+p}\| < \epsilon,$$

y por el otro, existe un $c > 0$ (por el lema anterior) tal que

$$\epsilon > \|\sum_{k=1}^{n} (\alpha_k^{(m)} - \alpha_k^{(m+p)}) e_k\| \geq c \sum_{k=1}^{n} |\alpha_k^{(m)} - \alpha_k^{(m+p)}| \quad \Longrightarrow$$

$$\sum_{k=1}^{n} |\alpha_k^{(m)} - \alpha_k^{(m+p)}| < \epsilon/c,$$

i.e., las sucesiones $(\alpha_k^{(m)})_m$ son de Cauchy, y por tanto convergen (los $\alpha_k^{(m)}$ son números reales o complejos), a ciertos α_k, $k = 1, 2, \ldots, n$. Sea $y = \sum_{k=1}^{n} \alpha_j e_j \in M$. Entonces

$$\|y_m - y\| = \left\| \sum_{k=1}^{n} (\alpha_k^{(m)} - \alpha_k) e_k \right\| \leq \sum_{k=1}^{n} |\alpha_k^{(m)} - \alpha_k| \|e_k\| \overset{m\to\infty}{\longrightarrow} 0 \quad \Longrightarrow$$

$$\lim_{m\to\infty} y_m = y \in M,$$

i.e., M es completo. □

Como consecuencia del teorema anterior se tiene que $\mathbb{R}^n$ es un espacio de Banach.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Definición 1.4.3 *Una norma $\|\cdot\|$ en un espacio vectorial $\mathbb{X}$ es equivalente a otra norma $\|\cdot\|'$ si existen dos números reales a, b positivos ($a > 0$, $b > 0$) tales que para todo $x \in \mathbb{X}$*

$$a\|x\|' \leq \|x\| \leq b\|x\|'.$$

De lo anterior se sigue que si dos normas son equivalentes entonces toda sucesión de Cauchy en $(\mathbb{X}, \|\cdot\|)$ también lo es en $(\mathbb{X}, \|\cdot\|')$, y viceversa. Usando el Lema 1.4.1 se puede probar el siguiente teorema:

Teorema 1.4.4 *Sea $\mathbb{X}$ un espacio vectorial de dimensión finita. Entonces cualquier norma $\|\cdot\|$ en $\mathbb{X}$ es equivalente a cualquier otra norma en $\mathbb{X}$.*

Demostración: Sea $\dim \mathbb{X} = n$ y sea $e_1, \ldots, e_n$ una base de $\mathbb{X}$. Entonces para todo $x \in \mathbb{X}$ existen (y son únicos) los números $\alpha_1, \ldots, \alpha_n$ tales que $x = \sum_{j=1}^{n} \alpha_j e_j$. Entonces, por el Lema 1.4.1

$$\|x\| \geq c(|\alpha_1| + \cdots + |\alpha_n|) \quad \Longrightarrow |\alpha_1| + \cdots + |\alpha_n| \leq \frac{\|x\|}{c}.$$

Pero

$$\|x\|' = \left\| \sum_{j=1}^{n} \alpha_j e_j \right\|' \leq \sum_{j=1}^{n} |\alpha_j| \|e_j\|' \leq \max_j \|e_j\|' \sum_{j=1}^{n} |\alpha_j| \leq \frac{k}{c}\|x\|,$$

donde hemos denotado por k al $\max_j \|e_j\|$. Intercambiando las normas obtenemos $\|x\| \leq \frac{k'}{c'}\|x\|'$, de donde se sigue el resultado. □

De lo anterior se sigue que en $\mathbb{R}^n$ todas las normas son equivalentes. Aunque, en general vamos a usar la norma $\|\cdot\|_2$ conocida como norma 2 o norma euclídea, en teorema anterior nos permite usar en cada momento la norma más conveniente. Como ejemplos de aplicación de lo anterior tenemos el siguiente resultado:

Proposición 1.4.5 *Una sucesión $(x_k)_k$ de elementos de $\mathbb{R}^n$ es convergente si y solo si lo son sus componentes.*

Demostración: Como sabemos, una sucesión $(x_k)_k$ de elementos de $\mathbb{R}^n$ es convergente a $x \in \mathbb{R}^n$ si para todo $\epsilon > 0$ existe un $N \in \mathbb{N}$ tal que para todo $k > N$, $\|x - x_k\| < \epsilon$, donde $x_k = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$. Si usamos la norma del máximo $\|x\|_\infty = \max_{k=1,\ldots,n} |x_k|$, entonces

$$\|x - x_n\| < \epsilon \quad \Longleftrightarrow \quad \max_{m=1,\ldots,n} |x_m^{(k)} - x_m| < \epsilon,$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

si y sólo si sus componentes $x_m^{(k)}$ convergen a las componentes del límite x_m, $m = 1, 2, \ldots, n$. □

Ejercicio: Adapta la prueba para la norma euclídea (1.3.7) ($p = 2$).

Para $\mathbb{R}^n$ se tiene el siguiente resultado:

Teorema 1.4.6 (Bolzano-Weierstrass) *Todo conjunto infinito y acotado de $\mathbb{R}^n$ tiene al menos un punto de acumulación.*

Demostración: En primer lugar, si $M \subset \mathbb{R}^n$ es un conjunto acotado, entonces existe un $K > 0$ tal que todos los vectores de $x \in M$ son tales que $\|x\| < K$, luego todas las componentes x_k, $k = 1, 2, \ldots, n$ de x están acotadas. El teorema se deduce entonces de aplicar el Teorema de Bolzano-Weierstrass en $\mathbb{R}$ 1.1.8a cada una de las componente de los vectores x. □

Definición 1.4.7 *Un espacio métrico $\mathbb{X}$ se denomina compacto si cualquier sucesión $(x_n)_n$ de elementos de $\mathbb{X}$ tiene una subsucesión convergente.*

Entenderemos que $M \subset \mathbb{X}$ es compacto si M es compacto como subconjunto de $\mathbb{X}$, i.e., cualquier $(x_n)_n$ de elementos de M tiene una subsucesión convergente en M.

Lema 1.4.8 *Si $M \subset \mathbb{X}$ es compacto, entonces M es cerrado y acotado.*

Demostración: Sea M compacto y sea $x \in \overline{M}$ cualquiera. Como $x \in \overline{M}$ entonces existe una sucesión $(x_n)_n$ en M tal que $x_n \overset{n\to\infty}{\longrightarrow} x \in \mathbb{X}$. Como M es compacto, entonces $x \in M$, luego $M = \overline{M}$ por lo que M es cerrado.

Supongamos que M es no acotado. Entonces existe al menos una sucesión $(x_n)_n$ de elementos de M tal que, fijado un $b \in M$ arbitrario, se tiene que $\rho(x_n, b) > n$ (¿por qué?). Dicha sucesión obviamente no puede tener ninguna subsucesión convergente (pues caso que la tuviera ésta fuese acotada) y por tanto M no puede ser compacto. □

El recíproco, en general, es falso. No obstante, en el caso de dimensión finita se tiene el siguiente teorema:

Teorema 1.4.9 *En un espacio normado $\mathbb{X}$ de dimensión finita (y por tanto en $\mathbb{R}^n$), todo subconjunto es compacto si y sólo si es cerrado y acotado.*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Demostración: $\Rightarrow$ Si M es compacto, entonces por el lema anterior es cerrado y acotado.

$\Leftarrow$ Sea M cerrado y acotado. Probemos que es compacto. Sea $n = \dim \mathbb{X}$ y $e_1, \dots, e_n$ una base. Sea la sucesión $(x_m)_m$ de M, entonces $x_m = \sum_{k=1}^{n} \alpha_k^{(m)} e_k \in \mathbb{X}$. Como M es acotado, x_m también lo es, i.e., existe $K > 0$ tal que, para todo $m \in \mathbb{N}$, $\|x_m\| \leq K$, luego, por el Lema 1.4.1 tenemos que

$$K \geq \|x_m\| = \left\| \sum_{k=1}^{n} \alpha_k^{(m)} e_k \right\| \geq c \sum_{k=1}^{n} |\alpha_k^{(m)}|.$$

Luego las sucesiones $(\alpha_k^{(m)})_m$ son acotadas para cada k, y por tanto existe una subsucesión convergente (Teorema de Bolzano-Weierstrass en $\mathbb{R}$ 1.1.8) a α_k, para todo $k = 1, 2, \dots, n$ –debemos hacer una construcción similar a la de la prueba del Lema 1.4.1–. Así, existe una subsucesión $(x_{n_k})_k$ que converge a $x = \sum_{k=1}^{n} \alpha_k e_k$. Probemos que dicho $x \in M$. Está claro que $x \in \overline{M}$ (ver el apartado 1 del Teorema 1.3.19). Como M es cerrado, $M = \overline{M}$, luego $x \in M$. Luego toda sucesión $(x_n)_n$ de M tiene una subsucesión convergente en M, i.e., M es compacto. □

1.5. Aplicaciones lineales

En adelante asumiremos que los espacios $\mathbb{X}$ e $\mathbb{Y}$ son espacios vectoriales reales y T es el operador $A : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$. $\mathcal{D}(T)$ denotará el dominio de la aplicación T e $\mathcal{I}(T)$ la imagen de T.

Definición 1.5.1 *Diremos que una aplicación (operador) $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ es lineal si, para todos $\alpha, \beta \in \mathbb{K}$ y para todos $x, y \in \mathcal{D}(T)$*

$$T(\alpha z + \beta y) = \alpha T(x) + \beta T(y).$$

Ejemplos de operadores lineales son:

1. El operador identidad $I : \mathbb{X} \mapsto \mathbb{X}$, tal que $y = Ix = x$ para todo $x \in \mathbb{X}$.
2. El operador nulo $\Theta : \mathbb{X} \mapsto \mathbb{Y}$, tal que $y = \Theta x = 0$ para todo $x \in \mathbb{X}$.
3. El operador derivada D definido por $D : \mathbb{P} \mapsto \mathbb{P}$, tal que $y(t) = Dp(t) = p'(t)$, donde $\mathbb{P}$ es el espacio de los polinomios reales $p(t)$ de cualquier grado.
4. El operador $T : \mathbb{R}^n \mapsto \mathbb{R}^m$, tal que $y = Tx = A \cdot x$, donde A es una matriz $n \times m$, x e y son los correspondientes vectores de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente, y $\cdot$ denota la multiplicación usual de matrices.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Representación matricial de los operadores lineales $T : \mathbb{X} \mapsto \mathbb{Y}$

Veamos el caso de un operador lineal $T : \mathbb{R}^n \mapsto \mathbb{R}^m$. El caso general cuando $\mathbb{X}$ e $\mathbb{Y}$ son de dimensiones n y m, respectivamente, es equivalente.

Sea $(e_k)_k$ una base de $\mathbb{R}^n$ y $(e'_i)_i$ una de $\mathbb{R}^m$. Entonces para todo $x \in \mathbb{R}^n$

$$x = \sum_{k=1}^{n} x_k e_k, \quad y = \sum_{i=1}^{m} y_i e'_i \quad \Longrightarrow \quad y = Tx = \sum_{k=1}^{n} x_k T e_k.$$

Como $Te_k \in \mathbb{R}^m$,

$$Te_k = \sum_{i=1}^{m} a_{i,k} e'_i \quad \Longrightarrow \quad y = \sum_{i=1}^{m} \left(\sum_{k=1}^{n} a_{i,k} x_k \right) e'_i = \sum_{i=1}^{n} y_i e'_i \quad \Longrightarrow$$

$$y_i = \sum_{k=1}^{n} a_{i,k} x_k, \qquad i = 1, 2, \dots n.$$

Es decir, si consideramos los vectores $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$ con coordenadas x_i, y_j, $i = 1, \dots, n$, $j = 1, \dots, m$, respectivamente, entonces el operador T se puede representar usando una matriz $(a_{i,j})_{i,j=1}^{n}$, i.e., tenemos la aplicación $T : \mathbb{R}^n \mapsto \mathbb{R}^m$, $y = Ax$, donde A es la matriz $m \times n$

$$A = (Te_1\, Te_2 \cdots Te_n) = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n} \end{pmatrix}, \tag{1.5.1}$$

obtenida a partir de la transformaciones de los elementos de la base $(e_k)_k$ de $\mathbb{R}^n$.

Definición 1.5.2 *Sean $\mathbb{X}$ e $\mathbb{Y}$ dos espacios normados y sea el operador $T : \mathcal{D}(T) \mapsto \mathbb{Y}$ lineal. T es acotado si existe $c \geq 0$ tal que*[8]

$$\|Tx\| \leq c\|x\|, \qquad \forall x \in \mathcal{D}(T). \tag{1.5.2}$$

De lo anterior se sigue que si T es acotado, entonces para todo $x \neq 0$,

$$\frac{\|Tx\|}{\|x\|} \leq c, \qquad \forall x \in \mathcal{D}(T),\ x \neq 0. \tag{1.5.3}$$

[8] Se sobrentiende que $\|x\|$ es la norma en $\mathbb{X}$ y $\|Tx\|$ es en $\mathbb{Y}$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

El menor valor de c para el cual (1.5.2) se cumple lo denotaremos por $\|T\|$ y se denomina *norma del operador* lineal T. Tomando supremos en $x \neq 0$ en (1.5.3) e ínfimos en c tenemos

$$\sup_{x\in\mathbb{X}\setminus\{0\}} \frac{\|Tx\|}{\|x\|} \leq \|T\|.$$

Por otro lado, para todo $y \neq 0$

$$\frac{\|Ty\|}{\|y\|} \leq \sup_{x\in\mathbb{X}\setminus\{0\}} \frac{\|Tx\|}{\|x\|} := c',$$

luego $\|Ty\| \leq c'\|y\|$ por lo tanto

$$\|T\| = \text{ínf}\{c : \|Ty\| \leq c\|y\|, \quad \forall y \in \mathbb{X}\} \leq c' = \sup_{x\in\mathbb{X}\setminus\{0\}} \frac{\|Tx\|}{\|x\|},$$

de donde se sigue que

$$\|T\| = \sup_{x\in\mathbb{X}\setminus\{0\}} \frac{\|Tx\|}{\|x\|}. \tag{1.5.4}$$

Si $T = 0$ obviamente $\|T\| = 0$. Además de (1.5.2), tomando ínfimos en c se tiene que para todo $y \in \mathbb{X}$

$$\frac{\|Ty\|}{\|y\|} \leq \|T\| \iff \|Ty\| \leq \|T\|\|y\|.$$

Teorema 1.5.3 *Toda aplicación lineal $T : \mathbb{X} \mapsto \mathbb{Y}$ de un espacio normado de dimensión finita $\mathbb{X}$ en otro espacio normado cualquiera $\mathbb{Y}$ es acotada.*

Demostración: Sea $\dim \mathbb{X} = n$ y sea $(e_1, \ldots, e_n)$ una base de $\mathbb{X}$. Para todo $x \in \mathbb{X}$, $x = \sum_{k=1}^n x_k e_k$. Entonces

$$\|Tx\| = \left\| T \sum_{k=1}^n x_k e_k \right\| \leq \sum_{k=1}^n |x_k| \|Te_k\| \leq \max_k \|Te_k\| \sum_{k=1}^n |x_k|.$$

Por otro lado, usando el Lema 1.4.1 tenemos que existe un $c > 0$ tal que

$$\|x\| = \left\| \sum_{k=1}^n x_k e_k \right\| \geq c \sum_{k=1}^n |x_k|.$$

Combinando ambas tenemos

$$\|Tx\| \leq \max_k \|Te_k\| \frac{\|x\|}{c} \quad \Longrightarrow \quad \|Tx\| \leq \gamma\|x\|,$$

con $\gamma = \max_k \|Te_k\|/c$. □

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Teorema 1.5.4 *Sea $T : \mathcal{D}(T) \subset \mathbb{X} \mapsto \mathbb{Y}$ una aplicación lineal de un espacio normado $\mathbb{X}$ a otro espacio normado $\mathbb{Y}$. Entonces*

1. *T es continuo si y sólo si T es acotado.*
2. *Si T es continuo en algún $x_0 \in \mathcal{D}(T)$, T es continuo en $\mathcal{D}(T)$.*

Demostración: Asumiremos que T no es el operador nulo.

1. $\Rightarrow$ Sea T acotado y sea $x_0 \in \mathcal{D}(T)$ cualquiera. Como T es lineal y acotado, entonces

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leq \|T\|\|x - x_0\|.$$

Entonces, para todo $\epsilon > 0$, existe un $\delta = \epsilon/\|T\| > 0$ tal que, para todo x con $\|x - x_0\| < \delta$, $\|Tx - Tx_0\| < \epsilon$, i.e., T es continuo en $\mathcal{D}(T)$.

$\Leftarrow$ Sea T lineal y continuo en $x_0 \in \mathcal{D}(T)$ cualquiera. Entonces para todo $\epsilon > 0$, existe un $\delta > 0$ tal que, para todo x con $\|x - x_0\| < \delta$, $\|Tx - Tx_0\| < \epsilon$. Sea $y \neq 0$ en $\mathcal{D}(T)$ cualquiera. Escojamos x tal que

$$x = x_0 + \frac{\delta}{2\|y\|}y \quad \Longrightarrow \quad x - x_0 = \frac{\delta}{2\|y\|}y \quad \Longrightarrow \quad \|x - x_0\| < \delta \quad \Longrightarrow$$

$$\|Tx - Tx_0\| < \epsilon.$$

Además, para dichos x tenemos, usando la linealidad de T, que

$$\|Tx - Tx_0\| = \|T(x - x_0)\| = \left\|T\frac{\delta}{2\|y\|}y\right\| = \frac{\delta}{2\|y\|}\|Ty\| \leq \epsilon \quad \Longrightarrow$$

$$\|Ty\| \leq \frac{2\epsilon}{\delta}\|y\|,$$

luego T es acotado.

2. Nótese que en la prueba anterior se probó que si T era continuo en un punto $x_0 \in \mathcal{D}(T)$, entonces era acotado en $\mathcal{D}(T)$. Pero entonces por 1, al ser acotado en $\mathcal{D}(T)$, es continuo en $\mathcal{D}(T)$. □

Nota 1.5.5 *Sea $\mathcal{L}(\mathbb{X}, \mathbb{Y})$ el espacio de todas las aplicaciones lineales de $\mathbb{X}$ en $\mathbb{Y}$. Definamos en dicho espacio la suma de dos aplicaciones $A+B$ y la multiplicación por un escalar λ de la forma habitual:*

$$\forall x \in \mathbb{X}, \quad (A + B)x = Ax + By, \quad (\lambda A)x = \lambda(Ax).$$

Con esta definición es fácil ver que $\mathcal{L}(\mathbb{X}, \mathbb{Y})$ es un espacio lineal (el elemento nulo de $\mathcal{L}(\mathbb{X}, \mathbb{Y})$ es el operador nulo).

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Si en dicho espacio definimos la cantidad $\|T\| = \sup_{x\in\mathbb{X},\|x\|=1} \|Tx\|$*, es sencillo probar (lo que dejamos como ejercicio al lector) que* $\|T\|$ *satisface los axiomas de la definición, i.e.,* $\|T\|$ *es una norma. Es decir, el subespacio* $\mathcal{B}(\mathbb{X},\mathbb{Y}) \subset \mathcal{L}(\mathbb{X},\mathbb{Y})$ *de todas las aplicaciones lineales acotadas (y, por tanto, según el Teorema 1.5.4, continuas) es un espacio normado.*

1.6. Espacios euclídeos

Para terminar esta introducción recordemos la definición de espacios euclídeos.

Definición 1.6.1 *Se dice que un espacio vectorial* $\mathbb{E}$ *es un espacio euclídeo si dados dos elementos cualesquiera* $x, y \in \mathbb{E}$ *existe un número denominado* producto escalar*, que denotaremos por* $\langle x, y\rangle$*, tal que*[9]

1. *Para todo* $x, y \in \mathbb{E}$, $\langle x, y\rangle = \overline{\langle y, x\rangle}$.
2. *Para todo* $x, y, z \in \mathbb{E}$, $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$.
3. *Para todo* $x, y \in \mathbb{E}$ *y* $\lambda \in \mathbb{C}$, $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$
4. *Para todo* $x \in \mathbb{E}$, $x \neq 0$, $\langle x, x\rangle > 0$ *y si* $\langle x, x\rangle = 0$*, entonces* $x = 0$.

De lo anterior se sigue que:

1. Para todos $x, y, z \in \mathbb{E}$, $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$.
2. Para todos $x, y \in \mathbb{E}$ y $\lambda \in \mathbb{C}$, $\langle x, \lambda y\rangle = \overline{\lambda}\langle x, y\rangle$.
3. Para todo $x \in \mathbb{E}$, $\langle x, 0\rangle = \langle 0, x\rangle = 0$.
4. Si $\langle x, z\rangle = \langle y, z\rangle$ para todos los $z \in \mathbb{E}$, entonces $x = y$.

El ejemplo más sencillo de espacio euclídeo es el espacio $\mathbb{R}^n$ con el producto escalar estándar: dados $x = (x_1, \ldots, x_n)$, e $y = (y_1, \ldots, y_n)$

$$\langle x, y\rangle = \sum_{k=1}^{n} x_k y_k.$$

[9] Si $\mathbb{E}$ es complejo denotaremos por $\overline{z}$ al complejo conjugado de z.

DIFERENCIACION DE FUNCIONES DE VARIAS VARIABLES

Teorema 1.6.2 (desigualdad de Cauchy-Schwarz) *Sea $\mathbb{E}$ un espacio euclídeo. Entonces para todos $f, g \in \mathbb{E}$,*

$$|\langle f, g\rangle|^2 \leq \langle f, f\rangle\langle g, g\rangle. \tag{1.6.1}$$

Demostración: Para demostrarla basta usar que para todo $\lambda \in \mathbb{C}$ y $f, g \in \mathbb{E}$, $\langle \lambda f + g, \lambda f + g\rangle \geq 0$, o equivalentemente,

$$|\lambda|^2\langle f, f\rangle + \lambda\langle f, g\rangle + \overline{\lambda}\langle g, f\rangle + \langle g, g\rangle = |\lambda|^2\langle f, f\rangle + 2\Re(\lambda\langle f, g\rangle) + \langle g, g\rangle \geq 0.$$

Escogiendo $\lambda = \alpha\overline{\langle f, g\rangle}/|\langle f, g\rangle|$, $\alpha \in \mathbb{R}$ tenemos

$$0 \leq |\lambda|^2\langle f, f\rangle + 2\Re(\overline{\lambda}\langle f, g\rangle) + \langle g, g\rangle = |\alpha|^2\langle f, f\rangle + 2|\alpha||\langle f, g\rangle| + \langle g, g\rangle,$$

De la desigualdad $|\alpha|^2\langle f, f\rangle + 2|\alpha||\langle f, g\rangle| + \langle g, g\rangle \geq 0$, se sigue que el discriminante de la ecuación cuadrática en α, $|\alpha|^2\langle f, f\rangle + 2|\alpha||\langle f, g\rangle| + \langle g, g\rangle = 0$, ha de ser negativo o nulo (¿por qué?). Luego $|\langle f, g\rangle|^2 - \langle f, f\rangle\langle g, g\rangle \leq 0$, de donde se sigue (1.6.1). □

Teorema 1.6.3 *Todo espacio euclídeo $\mathbb{E}$ es normado si en él definimos la norma mediante la fórmula $\|f\| = \sqrt{\langle f, f\rangle}$. Además, $|\langle f, g\rangle| \leq \|f\| \cdot \|g\|$.*

Demostración: Los dos primeros axiomas de la definición de norma 1.3.3 son inmediatos. Para probar el tercero nótese que

$$\begin{aligned}\|f + g\|^2 &= \langle f + g, g + g\rangle = \langle f, f\rangle + 2\Re(\langle f, g\rangle) + \langle g, g\rangle \\ &\leq \langle f, f\rangle + 2|\langle f, g\rangle| + \langle g, g\rangle \leq \langle f, f\rangle + 2\sqrt{\langle f, f\rangle\langle g, g\rangle} + \langle g, g\rangle \\ &= (\sqrt{\langle f, f\rangle} + \sqrt{\langle g, g\rangle})^2,\end{aligned}$$

donde hemos usado la desigualdad de Cauchy-Schwarz (1.6.1). Tomando ahora raíces cuadradas se sigue el resultado. □

De lo anterior se sigue que todo espacio euclídeo $\mathbb{E}$ es un espacio métrico con la métrica inducida por el producto escalar mediante la fórmula

$$\rho(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y\rangle}.$$

Así, en $\mathbb{R}^n$ tenemos que la norma inducida por el producto escalar es

$$\|x\|_2 = \sqrt{\sum_{k=1}^{n} |x_k|^2},$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Definición 1.6.4 *Un espacio euclídeo* $\mathbb{E}$ *completo*[10] *se denomina espacio de Hilbert.*

Luego $\mathbb{R}^n$ es un espacio de Hilbert.

Definición 1.6.5 *Sea el sistema de vectores* $(\phi_n)_n$ *(finito o infinito) de un espacio euclídeo* $\mathbb{E}$*. Diremos que* $(\phi_n)_{n=1}^{\infty}$ *es un sistema ortogonal dos a dos si*

$$\langle \phi_n, \phi_m \rangle = \delta_{n,m} \|\phi_n\|^2. \tag{1.6.2}$$

Si además $\|\phi_n\| = 1$ *para todo* $n \in \mathbb{N}$*, se dice que el sistema es ortonormal.*

Por ejemplo, el sistema de los vectores canónicos de $\mathbb{R}^n$, $(e_k)_{k=1}^n$, definido en (1.2.1) es un sistema ortonormal.

El siguiente teorema es sencillo de probar (por reducción al absurdo) y se deja como ejercicio.

Teorema 1.6.6 *Si los vectores* $x_1, \ldots, x_n$ *de un espacio euclídeo son ortogonales, entonces son linealmente independientes.*

Teorema 1.6.7 (Gram-Schmidt) *En un espacio de Hilbert* $\mathbb{H}$ *de cualquier conjunto de vectores linealmente independiente se puede construir un conjunto de vectores ortonormales (ortogonales).*

Demostración: Para probar el teorema tomamos un sistema de vectores linealmente independiente $(\phi_n)_n$ de $\mathbb{H}$ cualquiera y definimos un nuevo sistema de vectores $(\psi_n)_{n=1}^{\infty}$ de la siguiente forma:

1. Tomamos $\psi_1 = \phi_1/\|\phi_1\|$.
2. A continuación escogemos $\tilde{\psi}_2$ de la forma

$$\tilde{\psi}_2 = \phi_2 + \alpha_{2,1}\psi_1,$$

donde $\alpha_{2,1}$ es tal que $\tilde{\psi}_2$ sea ortogonal al vector ϕ_1, i.e. $\langle \psi_1, \psi_2 \rangle = 0$, de donde se deduce que $\alpha_{2,1} = -\langle \psi_1, \phi_2 \rangle$. Entonces $\psi_2 = \tilde{\psi}_2/\|\tilde{\psi}_2\|$ es ortonormal a ϕ_1.

[10] Es decir, un espacio $\mathbb{E}$ donde cualquier sucesión de Cauchy converge a un vector de $\mathbb{E}$ (en la norma inducida por el producto escalar).

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

3. Paso n. Escogemos $\tilde{\psi}_k$, $k \geq 3$, de la forma

$$\tilde{\psi}_n = \phi_n + \sum_{k=1}^{n-1} \alpha_{n,k} \psi_k,$$

donde los coeficientes $\alpha_{n,k}$, $n \in \mathbb{N}$, $k = 1, \ldots, n-1$ son tales que $\tilde{\psi}_n$ sea ortogonal a todos los vectores ϕ_k, $k = 1, 2, \ldots, n-1$, anteriores. Usando la ortogonalidad es fácil comprobar que $\alpha_{n,k} = -\langle \psi_k, \phi_n \rangle$, $k = 1, 2, \ldots, n-1$. Finalmente definimos $\psi_n = \tilde{\psi}_n / \|\tilde{\psi}_n\|$ que es ortonormal a todos los vectores anteriores ϕ_k, $k = 1, 2, \ldots, n-1$. Y así sucesivamente. □

Si quieres saber más sobre la teoría de espacios métricos, normados y de Hilbert, puedes consultar, por ejemplo, [4, 7].

1.7. Problemas propuestos

Problema 1.7.1 Prueba la desigualdad de Young: Dados dos números reales $a, b > 0$ y $p > 1$, entonces,

$$a^{\frac{1}{p}} b^{\frac{1}{q}} \leq \frac{1}{p} a + \frac{1}{q} b, \qquad \frac{1}{p} + \frac{1}{q} = 1. \tag{1.7.1}$$

Además la igualdad sólo tiene lugar si $a = b$. **Ayuda**: Encuentra los extremos de la función $f : [0, \infty) \mapsto \mathbb{R}$, $f(x) = x^\alpha - \alpha x + \alpha - 1$, $\alpha \in (0, 1)$ y prueba que $f(x) \leq 0$ para todo $x \geq 0$. Escogiendo $x = a/b$, $a, b > 0$ y $\alpha = 1/p$, $p > 1$ se deduce el resultado.

Problema 1.7.2 Prueba la desigualdad de Hölder: Sean los números x_i e y_i, $i = 1, 2 \ldots, n$ no negativos. Entonces, para todo $p > 1$ y $q > 1$ tales que $\frac{1}{p} + \frac{1}{q} = 1$, se tiene

$$\sum_{i=1}^{n} x_i \, y_i \leq \left(\sum_{i=1}^{n} x_i^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} y_i^q \right)^{\frac{1}{q}}, \tag{1.7.2}$$

donde la igualdad sólo tiene lugar si $x_i^p = c \, y_i^q$ para todo $i = 1, 2, \ldots, n$ (es decir si x_i^p y y_i^q son proporcionales). Hölder. **Ayuda**: Usa la desigualdad de Young.

Problema 1.7.3 Usando la desigualdad de Hölder (1.7.2) prueba la desigualdad de Minkowski (1.3.4).

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Problema 1.7.4 Prueba que si dos elementos x e y de un espacio euclídeo son ortogonales, entonces

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

La igualdad anterior es una generalización del Teorema de Pitágoras.

Problema 1.7.5 Sea $\mathbb{E}$ un espacio euclídeo y $\|\cdot\|$ la norma inducida por el producto escalar. Prueba que para todos $x, y \in \mathbb{E}$,

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2). \tag{1.7.3}$$

Esta igualdad se suele denominar *ley del paralelogramo*.

Problema 1.7.6 *Se sabe que todo espacio euclídeo es un espacio normado con la norma inducida por el producto escalar. ¿Y al contrario? Encuentra un ejemplo de espacio normado que no es euclídeo. No obstante un espacio normado $\mathbb{X}$ es euclídeo si y sólo si para todos $x, y \in \mathbb{E}$,*

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

Es decir, la ley del paralelogramo caracteriza los espacios euclídeos. Prueba dicha afirmación para el caso de espacios normados reales. ***Ayuda****: Prueba que si se cumple la ley del paralelogramo entonces la función $f(x,z) := \langle x, z\rangle = \frac{1}{4}(\|x+z\|^2 - \|x-z\|^2)$, $z \in \mathbb{X}$, f cumple con las las cuatro propiedades de la Definición 1.6.1.*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Diferenciación de funciones de varias variables

Capítulo 2

Límite, continuidad y diferenciabilidad

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

2.1. Límite y continuidad de funciones de varias variables

Una función vectorial de n-variables es una aplicación $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$. Está claro que como

$$f(x_1, \cdots, x_n) = (f_1(x_1, \cdots, x_n), f_2(x_1, \cdots, x_n), \ldots, f_m(x_1, \cdots, x_n))$$

para estudiar las propiedades de f podemos restringirnos al estudio de cada una de las componentes de f. Es decir, basta con estudiar las funciones $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$. En adelante, a no ser que se diga lo contrario, asumiremos que $A \in \mathbb{R}^n$ es un abierto.

Definición 2.1.1 *Diremos que $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ tiene límite l cuando x tiende a a, punto límite de A, y lo denotaremos por $\lim_{x\to a} f(x) = l$, si*

$$\forall \epsilon > 0 \quad \exists \delta > 0; \quad 0 < \|x - a\| < \delta \quad \Longrightarrow \quad \|f(x) - l\| < \epsilon.$$

Si además $l = f(a)$ diremos que f es continua en a.

En la figura 2.1.1 mostramos esquemáticamente el significado geométrico de la definición anterior para el caso de una función $f : \mathbb{R}^2 \mapsto \mathbb{R}$ tal que $\lim_{(x,y)\to(a,b)} f(x) = l$. En ella vemos representada la función como una superficie

de $\mathbb{R}^3$. En el eje z se muestra el valor del límite l que es el centro del intervalo $(l-\epsilon, l+\epsilon)$, mientras que en el plano xy se muestra la bola de radio δ centrada en el punto $A=(a,b)$, $B(A,\delta)$. Sobre la propia superficie se muestra el punto (a,b,l). Así, el conjunto $f(B(A,\delta)\setminus\{(a,b)\})\subset(l-\epsilon, l+\epsilon)$.

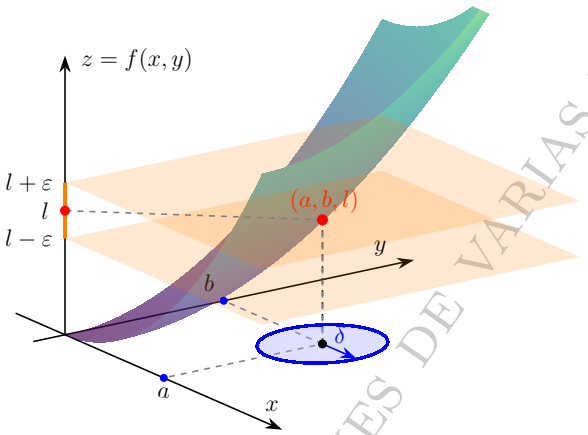


Figura 2.1.1: Significado geométrico de la definición de límite para una función $f:\mathbb{R}^2\mapsto\mathbb{R}$ tal que $\lim_{(x,y)\to(a,b)} f(x)=l$.

De la definición anterior se sigue que

1. Si una función $f:A\subset\mathbb{R}^n\mapsto\mathbb{R}^m$ tiene límite en $x=a$, entonces este es único.

2. Una función $f:A\subset\mathbb{R}^n\mapsto\mathbb{R}^m$ tiene límite si y sólo si tienen límite cada una de sus componentes.

3. Una función $f:A\subset\mathbb{R}^n\mapsto\mathbb{R}^m$ es continua si y sólo si son continuas sus componentes.

4. Si $f:A\subset\mathbb{R}^n\mapsto\mathbb{R}^m$ tiene límite (y en particular si es continua) en $x=a$ entonces existe un entorno $B(a,\delta)$ de a donde f es acotada.

Demostración:

1. Supongamos que existen dos límites distintos $l_1\neq l_2$. Entonces, por un lado, $d=\|l_1-l_2\|>0$, y por otro, como $f(x)\to l_1$ y $f(x)\to l_2$, entonces para todo

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

$\varepsilon > 0$ existe un $\delta > 0$ tal que se cumple, al mismo tiempo (¿por qué?), que $\|f(x) - l_1\| < \frac{\varepsilon}{2}$ y $\|f(x) - l_2\| < \frac{\varepsilon}{2}$, por tanto,

$$\|l_1 - l_2\| \leq \|f(x) - l_1\| + \|f(x) - l_2\| < \varepsilon \quad \Rightarrow \quad l_1 = l_2,$$

lo cual es una contradicción.

2. Razonando como en la prueba de la Proposición 1.4.5 se sigue el resultado.

3. Es consecuencia de 2.

4. Una forma sencilla de probarlo es la siguiente:

$$\|f\| = \|f - l + l\| \leq \|f - l\| + \|l\|, \quad \|l\| = \|l - f + f\| \leq \|f - l\| + \|f\| \Longrightarrow$$
$$|\|f\| - \|l\|| \leq \|f - l\|.$$

Si f tiene límite l entonces en todo un entorno $\|f - l\| < \epsilon$, luego en dicho entorno

$$|\|f\| - \|l\|| \leq \|f - l\| < \epsilon \quad \Longrightarrow \quad \|l\| - \epsilon \leq \|f\| \leq \|l\| + \epsilon.$$

Es conveniente hacer notar que si $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ es continua en $x = a$ y $f(a) \neq 0$, entonces existe todo un entorno de a donde f mantiene su signo. La prueba es similar al caso de una sola variable y la omitiremos.

Teorema 2.1.2 (de Weierstrass para funciones continuas) *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ una función continua y sea $C \subset A \subset \mathbb{R}^n$ un compacto (cerrado y acotado). Entonces f es acotada en C y alcanza su máximo y su mínimo absolutos en C.*

Demostración: Sea $C \subset A$ un intervalo cerrado y acotado, y por tanto compacto, (ver Teorema 1.4.9) y sea f continua en C pero no acotada. Entonces, para todo $n \in \mathbb{N}$ existe un $x_n \in C$ tal que $|f(x_n)| > n$, es decir existe una sucesión tal que $x_n \in C$ y $f(x_n)$ es no acotada. Pero $\{x_n\}$ está definida en un compacto C luego, por definición, se puede extraer de la misma una subsucesión $\{x_{n_k}\}$ convergente, i.e., $x_{n_k} \overset{k\to\infty}{\longrightarrow} c \in C$. Pero al ser f continua en C, $f(x_{n_k}) \overset{k\to\infty}{\longrightarrow} f(c)$, luego $\{f(x_{n_k}\}$ es acotada, lo cual es una contradicción pues $|f(x_{n_k})| > n_k$.

Probemos ahora que f alcanza su máximo y mínimo en C. Como f es acotada superiormente existe el supremo de f que llamaremos S. Demostremos que S es alcanzable, es decir que existe un $c \in C$ tal que $f(c) = S$. Supongamos que dicho c no existe. Entonces la función $S - f(x)$ es continua y mayor que cero en todo $[a, b]$, por tanto la función

$$g(x) = \frac{1}{S - f(x)} > 0,$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

es continua y acotada superiormente. Luego existe un $M > 0$ tal que para todo $x \in C$, $g(x) \leq M$, es decir, que para todo $x \in C$, $f(x) \leq S - \frac{1}{M}$, lo cual contradice que S sea el supremo de f. La prueba para el mínimo es análoga y se deja como ejercicio. □

Antes de discutir algunos ejemplos debemos recalcar que, en general, vamos a usar la norma euclídea. Así que, para las funciones escalares de dos variables, i.e., $f : \mathbb{R}^2 \mapsto \mathbb{R}$, la norma en $\mathbb{R}^2$ será $\|(x, y)\| = \sqrt{x^2 + y^2}$ y en $\mathbb{R}$, $\|x\| = |x|$.

Además, tenemos que tener en cuenta que el problema de calcular límites en $\mathbb{R}^n$ es *mucho* más complicado que el caso de $\mathbb{R}$. La razón principal es que el $\mathbb{R}$ cuando $x \to a$ sólo hay dos formas de aproximarse a a: por la izquierda o por la derecha, mientras que en $\mathbb{R}^n$ existen una infinidad de maneras (trayectorias) de hacerlo. En cualquier caso, lo que está claro es que si $\lim_{x\to a} f(x) = l$, entonces, independientemente de la forma que nos acerquemos a a, $f(x)$ tiene que acercarse a l. Si resulta que dada una función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, cuando nos acercamos a a siguiendo distintas *trayectorias* obtenemos resultados distintos, entonces f no tiene límite en a.

Veamos dos ejemplos que muestran esto último.

Ejemplo 2.1.3 Calcular el límite cuando $(x, y) \to (0, 0)$ de la función $f(x, y)$ definida por

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}, \quad \text{si } (x, y) \neq (0, 0).$$

Si elegimos $y = mx$, $m \neq 0$, con $x \to 0$, está claro que $f(x, mx) = (1 - m^2)/(1 + m^2)$, luego el límite cuando $x \to 0$ (y por tanto $y \to 0$) depende de la dirección que tomemos, de donde se deduce que no existe el límite de $f(x, y)$ cuando $(x, y) \to (0, 0)$ (pues este ha de ser único e independiente de la trayectoria). □

Ejemplo 2.1.4 Calcular el límite cuando $(x, y) \to (0, 0)$ de la función $f(x, y)$ definida por

$$f(x, y) = \frac{x^2 y}{x^4 + y^2}, \quad \text{si } (x, y) \neq (0, 0).$$

Podemos comprobar que $\lim_{x\to 0} f(x, mx) = 0$ para todo m, sin embargo si elegimos la trayectoria $y = x^2$, entonces $\lim_{x\to 0} f(x, x^2) = 1/2 \neq 0$, luego no existe el límite de $f(x, y)$ cuando $(x, y) \to (0, 0)$. □

Veamos ahora un ejemplo cuando existe el límite.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Ejemplo 2.1.5 Sea la función

$$f(x,y)=\frac{|x|^{3/2}y}{x^2+y^2},\qquad \text{si } (x,y)\neq(0,0).$$

En primer lugar es sencillo comprobar que $\lim_{x\to 0} f(x,mx)=0$ para todo m, y que $\lim_{x\to 0} f(x,x^2)=0$. Incluso es cero tomando $y=|x|^\alpha$, $\alpha>0$. Eso nos induce a pensar que probablemente tenga límite y que este es cero. Para probar que existe el límite tenemos que comprobar que, para todo $\epsilon>0$ existe un $\delta>0$ tal que si $\sqrt{x^2+y^2}<\delta$, entonces $|f(x,y)|<\epsilon$, o equivalentemente, si $\sqrt{x^2+y^2}\to 0$, $|f(x,y)|\to 0$.

Para ello vamos a intentar encontrar acotaciones de $|f|$.

Teniendo en cuenta que $2|xy|\leq x^2+y^2$ cualesquiera sean $x,y\in\mathbb{R}$, tenemos

$$0\leq\left|\frac{x^{3/2}y}{x^2+y^2}\right|=|x|^{1/2}\left|\frac{|xy|}{x^2+y^2}\right|\leq\frac{1}{2}|x|^{1/2}. \tag{2.1.1}$$

Ahora bien, como $|x|\leq\sqrt{x^2+y^2}<\delta$, entonces, $|f(x,y)|\leq\delta/2<\delta$, así que dado $\epsilon>0$ cualquiera, si elegimos $\delta=\epsilon$ se tiene que $|f(x,y)|<\epsilon$, es decir, de la acotación anterior (2.1.1) se sigue que cuando $(x,y)\to(0,0)$ (en la norma euclídea), $|f(x,y)|\to 0$ lo cual también implica que $f(x,y)$ tiende a cero si $(x,y)\to(0,0)$. Luego $\lim\limits_{(x,y)\to(0,0)} f(x,y)=0$.

Nótese que si definimos $f(0,0)=0$, entonces la función es continua en $(0,0)$ y, por tanto, es continua en todo $\mathbb{R}^2$. □

2.2. Diferenciabilidad de funciones de varias variables

Recordemos que[1] dada una función de una variable $f:A\subset\mathbb{R}\mapsto\mathbb{R}$, la derivada $f'(a)$ en $x=a$ se define como el límite, si existe, del siguiente cociente:

$$\lim_{h\to 0}\frac{\Delta f(a)}{\Delta x}=\lim_{h\to 0}\frac{f(a+h)-f(a)}{h},$$

donde el numerador representa la diferencia de los valores de f en los puntos $a+h$ y a y el denominador la diferencia entre los valores de $x=a+h$ y $x=a$.

[1] Ver Definición 1.1.14.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

El problema de la derivación en varias variables es algo más sutil. Hay muchos candidatos para definir la derivada de una función de varias variables. El candidato más sencillo es considerar el caso cuando solo cambia una de las variables. Así, tenemos la siguiente definición:

Definición 2.2.1 *Sea A un subconjunto abierto de $\mathbb{R}^n$, $a \in A$ y f una función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$. La derivada parcial i-ésima ($1 \leq i \leq n$) de f en a se define como el límite, si existe,*

$$\lim_{x_i \to a_i} \frac{f(a_1, \cdots, a_{i-1}, x_i, a_{i+1}, \cdots, a_n) - f(a_1, \cdots, a_i, \cdots, a_n)}{x_i - a_i} = \lim_{h \to 0} \frac{f(a_1, \cdots, a_{i-1}, a_i + h, a_{i+1}, \cdots, a_n) - f(a_1, \cdots, a_i, \cdots, a_n)}{h}.$$

A dicha derivada la denotaremos por $D_i f(a)$ o $\dfrac{\partial f(a)}{\partial x_i}$.

De la definición anterior se deduce que para calcular las derivadas parciales se procede de forma análoga al caso del cálculo de la derivada de una función de una variable. De hecho todas las reglas de la derivación de una variable se traspasan al caso de varias variables, simplemente al calcular $\frac{\partial f(a)}{\partial x_i}$ se consideran todas las variables x_k, $k = 1, \ldots, n$, $k \neq i$ constantes, de forma que a todos los efectos $f(x_1, \ldots, x_n)$ es una función de una única variable x_i. Y así con cada una de las derivadas parciales. Nótese que las operaciones en la definición anterior están bien definidas: el numerador es una diferencia de vectores de $\mathbb{R}^m$ y en denominador es un escalar.

Veamos un ejemplo. Sea $f : \mathbb{R}^2 \mapsto \mathbb{R}$, $f(x, y) = \exp(x^2 + y^2)$ entonces

$$\frac{\partial f(x,y)}{\partial x} = 2x \exp(x^2 + y^2) \quad \text{y} \quad \frac{\partial f(x,y)}{\partial y} = 2y \exp(x^2 + y^2).$$

Otra opción, más general, es definir las derivadas en la dirección definida por un vector u. Para ello tomamos un punto a del dominio de nuestra función y trazamos la recta r que pasa por dicho punto y cuyo vector director coincida con el vector u.

En este caso podemos calcular sin problema la diferencia de de los valores de f en los puntos $a + hu$ y a: $\Delta f(a) = f(a + hu) - f(a)$, que es un vector de $\mathbb{R}^n$, pero ¿qué elegimos como diferencia entre los valores de $x = a + hu$ y $x = a$? Esta claro que no puede ser hu pues no tiene sentido dividir $\Delta f(a)$ por un vector, así que el candidato más natural sería dividir $\Delta f(a)$ por $h\|u\|$. Gráficamente

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

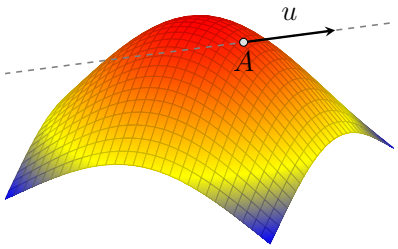


Figura 2.2.1: Cálculo de la derivada direccional de $z = f(x)$ en el punto A según la dirección definida por el vector u.

representamos el proceso del cálculo en la figura 2.2.1, donde el punto A es el punto de la superficie correspondiente a $x = a$.

Así tenemos la siguiente definición:

Definición 2.2.2 *Para cada vector normalizado $u \in \mathbb{R}^n$, $\|u\| = 1$, denominaremos derivada direccional de f en a según la dirección u, y lo denotamos por $D_u f(a)$, al límite, si existe,*

$$\lim_{h\to 0} \frac{f(a+hu) - f(a)}{h}. \tag{2.2.1}$$

Nótese que si denotamos por e_i, $i = 1, \ldots, n$ a los vectores de la base canónica de $\mathbb{R}^n$ entonces

$$D_{e_i} f(a) = \frac{\partial f(a)}{\partial x_i}.$$

Es importante destacar que la existencia de derivadas direccionales ni siquiera garantiza la continuidad de la función. Por ejemplo, para la función

$$f(x,y) = \begin{cases} \dfrac{x^2 y}{x^4+y^2}, & \text{si } (x,y) \neq (0,0), \\ 0, & \text{si } (x,y) = (0,0), \end{cases}$$

tenemos

$$D_u f(0,0) = \lim_{h\to 0} \frac{f(hu_x, hu_y) - f(0,0)}{h} = \lim_{h\to 0} \frac{h^3 u_x^2 u_y}{h(h^4 u_x^4 + u_y^2 h^2)} = \begin{cases} 0, & u_y = 0, \\ \dfrac{u_x^2}{u_y} & u_y \neq 0. \end{cases}$$

Es decir, existen todas sus derivadas direccionales en $(0,0)$ pero, como ya vimos en el Ejemplo 2.1.4, f ni siquiera es continua en dicho punto (no existe el límite).

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Lo anterior implica que, a diferencia del caso de una variable, la existencia de las derivadas direccionales de una función no garantiza siquiera la continuidad de la misma, así que es natural preguntarse si hay un concepto de derivada en varias variables que implique la continuidad. Así, por similitud con el caso de una variable, vamos a definir la diferenciabilidad de una función f de la siguiente forma:

Definición 2.2.3 *Sea A un subconjunto abierto de $\mathbb{R}^n$, y $a \in A$. Una función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ es diferenciable en a si existe una aplicación lineal de $D : \mathbb{R}^n \mapsto \mathbb{R}^m$, a la que denotaremos por $Df(a)$, tal que*

$$\lim_{x \to a} \frac{f(x) - f(a) - Df(a)(x-a)}{\|x-a\|} = 0,$$

o, equivalentemente,

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - Df(a)(h)}{\|h\|} = 0.$$

Lo anterior suele escribirse como

$$f(a+h) - f(a) - Df(a)(h) = o(\|h\|), \tag{2.2.2}$$

donde hemos usado el símbolo[2] "*o pequeña*" que significa que

$$\lim_{h \to 0} \frac{o(\|h\|)}{\|h\|} = 0,$$

es decir, que cada una de las componentes del vector $o(\|h\|) \in \mathbb{R}^m$ tiende a cero más rápido que $\|h\|$ cuando $h \to 0$.

Es importante recalcar que la diferencial o derivada *total* de una función de f en a, que denotamos por $Df(a)$, es un operador lineal. Nótese además que de la definición anterior se deducen la siguientes propiedades:

1. $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ es diferenciable en a si y sólo si lo son sus funciones componentes.
2. Si f es diferenciable en a, entonces f continua en a.
3. Si f es diferenciable en a, entonces existen todas las derivadas direccionales de f en a y $D_u f(a) = Df(a)(u)$. Lo anterior implica que, dada una función f diferenciable en $x = a$, la aplicación $u \mapsto D_u f(a)$ que a cada vector $u \neq 0$ le hace corresponder la derivada direccional $D_u f(a)$ es una aplicación lineal.

[2] Ver la Definición 1.1.25 en la página 12.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

4. Si f es diferenciable en a, entonces existen todas las derivadas parciales de f en a y

$$\frac{\partial f(a)}{\partial x_i} = Df(a)(e_i), \tag{2.2.3}$$

donde e_i es el i-ésimo vector de la base canónica de $\mathbb{R}^n$.

5. Si f y g son diferenciables en a, entonces también lo es la suma $f+g$ y λf, $\lambda \in \mathbb{R}$, y se verifica que $D(f+g)(a) = Df(a)+Dg(a)$, $D(\lambda f)(a) = \lambda Df(a)$.

Demostración:

1. Basta aplicar la formula (2.2.2) componente a componente.

2. Para probar esta implicación basta recordar que toda aplicación lineal sobre un espacio normado de dimensión finita es acotada. De (2.2.2), se tiene que, cuando $\|h\| \to 0$

$$\|f(a+h) - f(a)\| = \|Df(a)(h) + o(\|h\|)\| \leq \|Df(a)\|\|h\| + \|o(\|h\|)\| \to 0,$$

pues, por el Teorema 1.5.3, $Df(a)$ es una aplicación lineal acotada.

3. Es inmediato de la definición (2.2.1) y (2.2.2) pues

$$\frac{f(a+hu) - f(a)}{h} = \frac{Df(a)(hu) + o(\|hu\|)}{h} = Df(a)(u) + \frac{o(h)}{h} \to Df(a)(u),$$

pues $o(\|hu\|) = o(|h|\|u\|) = o(|h|) = o(h)$.

La prueba de que la aplicación $u \mapsto D_u f(a)$ es lineal se deja como ejercicio al lector.

4. Se sigue de 3. eligiendo $u = e_i$, $i = 1, \ldots, n$, los vectores de la base canónica de $\mathbb{R}^n$.

5. Es inmediata de la definición (2.2.2) y se deja como ejercicio.

Nótese que de (2.2.3) se deduce que si no existe alguna de las derivadas parciales de f, entonces la función no puede ser diferenciable. Es decir, la existencia de derivadas parciales es una condición necesaria para la diferenciabilidad. Además, de la diferenciabilidad de f se deduce la existencia de las derivadas parciales y direccionales, pero no a la inversa.

Para mostrar lo anterior definamos la función

$$f(x,y) = \begin{cases} 0, & \text{si } xy = 0, \\ 1, & \text{si } xy \neq 0. \end{cases}$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Claramente esta función es discontinua en el origen, luego no puede ser diferenciable en $(0,0)$ y sin embargo $\frac{\partial f(0,0)}{\partial x} = \frac{\partial f(0,0)}{\partial y} = 0$.

Por otro lado, la función $f(x,y) = \dfrac{x^3 y}{x^4+y^2}$ si $(x,y) \neq (0,0)$, $f(0,0) = 0$, no es diferenciable en $(0,0)$ y sin embargo todas sus derivadas direccionales en $(0,0)$ son cero. En efecto, dado un vector unitario $u = (u_x, u_y)$ cualquiera

$$D_u f(0,0) = \lim_{h\to 0} \frac{f(hu_x, hu_y) - f(0,0)}{h} = \lim_{h\to 0} \frac{h^4 u_x^3 u_y}{h(h^4 u_x^4 + u_y^2 h^2)} = 0.$$

En particular, para las derivadas parciales tenemos $\dfrac{\partial f(0,0)}{\partial x} = \dfrac{\partial f(0,0)}{\partial y} = 0$.

Nótese que de lo anterior se deduce que si f fuese diferenciable en $(0,0)$, entonces $Df(0,0) = (0\,0)$, pues las derivadas parciales de f en $(0,0)$ son cero. Así, tomando $h = (x,y)$ tenemos

$$\frac{f(x,y) - f(0,0) - Df(0,0)(h)}{\sqrt{x^2+y^2}} = \frac{x^3 y}{(x^4+y^2)\sqrt{x^2+y^2}}.$$

Si elegimos la trayectoria $y = x^2$ obtenemos que, cuando $x \to 0$,

$$\frac{x^5}{2x^4\sqrt{x^2+x^4}} = \frac{x}{2|x|\sqrt{1+x^2}} \not\to 0,$$

luego f no puede ser diferenciable en $(0,0)$.

Un caso similar ocurre con la función $f(x,y) = x^3/y$ si $(x,y) \neq (0,0)$, $f(0,0) = 0$, que tiene derivadas direccionales en $(0,0)$ pero ni siquiera es continua en dicho punto. La comprobación de este hecho se deja como ejercicio al lector.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Representación matricial del diferencial.

Si elegimos en $\mathbb{R}^n$ y $\mathbb{R}^m$ las base canónicas e_k, $k = 1,\ldots,n$, y e'_i $j = 1,\ldots,m$, respectivamente, entonces de (2.2.3) y (1.5.1) se deduce que la matriz asociada a

la aplicación lineal $Df(a)$ tiene la forma:

$$Df(a) = \begin{pmatrix} \frac{\partial f_1(a)}{\partial x_1} & \frac{\partial f_1(a)}{\partial x_2} & \cdots & \frac{\partial f_1(a)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m(a)}{\partial x_1} & \frac{\partial f_m(a)}{\partial x_2} & \cdots & \frac{\partial f_m(a)}{\partial x_n} \end{pmatrix} = \begin{pmatrix} D_1 f_1(a) & D_2 f_1(a) & \ldots & D_n f_1(a) \\ \vdots & \vdots & \ddots & \vdots \\ D_1 f_m(a) & D_2 f_m(a) & \ldots & D_n f_m(a) \end{pmatrix}. \tag{2.2.4}$$

A la matriz $Df(a)$ se le denomina matriz jacobiana de f en a y se suele denotar por también por $J_f(a)$. Al determinante de la matriz $J_f(a)$ se le denomina jacobiano de f en a.

Nótese que, dado que las derivadas parciales son únicas, lo anterior implica que si f es diferenciable en $x = a$, entonces el diferencial $Df(a)$ es único. Se deja como ejercicio al lector probar la unicidad de la diferencial directamente a partir de la propia definición (2.2.2).

El vector gradiente y la interpretación geométrica de la diferencial.

Supongamos que $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ es diferenciable en a. Entonces existen todas sus derivadas parciales. Se define al vector $\nabla f(a)$ por

$$\nabla f(a) = \left(\frac{\partial f(a)}{\partial x_1}, \ldots, \frac{\partial f(a)}{\partial x_n} \right)$$

y se le denomina gradiente de f en $x = a$. Nótese que la derivada direccional $D_u f(a)$ según el vector normalizado u se puede expresar de la siguiente forma:

$$D_u f(a) = \langle \nabla f(a), u \rangle.$$

De la expresión anterior se deduce que la derivada direccional es máxima en la dirección del gradiente y si $\nabla f(a)$ es ortogonal a u, entonces $D_u f(a) = 0$. En otras palabras, la dirección donde f varía (crece o decrece) más rápidamente es en la dirección del gradiente ($\nabla f(a)$ y $-\nabla f(a)$, respectivamente). Cuando el gradiente es cero ello indica que en esa dirección la función no cambia pues, en ese caso, $D_u f(a) = 0$ lo que implica que $f(a + hu) \approx f(a)$. Este hecho es importante pues es una forma de encontrar las curvas de nivel, es decir el conjunto de puntos del dominio de f tales que $f(x) = k$, con k una constante dada, pues el vector gradiente de f es siempre perpendicular a dicha curva (se asume que la curva es *suave*, o sea no tiene *picos*).

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

En la figura 2.2.2 mostramos las curvas de nivel de una función de dos variables $f(x, y)$. Cada plano representa un valor de f. Sobre el plano xy se representan las curvas correspondientes obtenidas al intersecarse la superficie definida por f y los planos correspondientes a los valores de f.

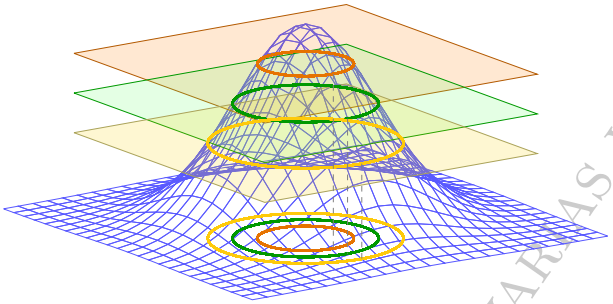


Figura 2.2.2: Curvas de nivel de una función $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}$. En la figura se representa una función y tres planos paralelos al plano xy a distintas *alturas*. La intersección de dichos planos con f, o sea, las curvas de nivel, se muestran proyectadas bajo la propia función.

Veamos ahora la interpretación geométrica de la diferencial $Df(a)$. Por simplicidad usaremos una función $f : \mathbb{R}^2 \mapsto \mathbb{R}$. En Si f es diferenciable en (a, b) entonces, como ya hemos visto

$$f(x, y) = f(a, b) + \frac{\partial f(a, b)}{\partial x}(x - a) + \frac{\partial f(a, b)}{\partial y}(y - b) + o(\|h\|),$$

donde $\|h\| = \sqrt{(x - a)^2 + (y - b)^2}$. Sea $c = f(a, b)$ y $z = f(x, y)$. Con esta notación la expresión anterior se puede reescribir como

$$\frac{\partial f(a, b)}{\partial x}(x - a) + \frac{\partial f(a, b)}{\partial y}(y - b) - (z - c) = o(\|h\|).$$

Sea ahora σ la superficie definida por los puntos $(x, y, f(x, y))$. Si f es diferenciable en (a, b) entonces, despreciando la $o(\|h\|)$ en la expresión anterior, esta se convierte en la ecuación

$$\frac{\partial f(a, b)}{\partial x}(x - a) + \frac{\partial f(a, b)}{\partial y}(y - b) - (z - c) = 0,$$

que define un plano π que contiene al punto (a, b, c) y cuyo vector normal en dicho punto es $\vec{n} = \left(\frac{\partial f(a,b)}{\partial x}, \frac{\partial f(a,b)}{\partial y}, -1\right)$.

Demostremos que, efectivamente, dicho plano π es tangente a σ en $\xi = (a, b, c)$. Para ello basta probar que cualquier punto $s = (x, y, f(x, y))$ de la superficie σ es tal que el vector $\vec{v} = s - \xi$ forma un ángulo de 90 grados con el vector normal $\vec{n}$ anterior cuando s se acerca a ξ.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

En efecto, sea $s = (x, y, z) \neq \xi = (a, b, c)$ un punto de la superficie σ y sea $\vec{v} = s - \xi$. Entonces, el coseno del ángulo entre los vectores $\vec{n}$ y $\vec{v}$ es

$$\cos\theta = \frac{\langle \vec{v}, \vec{n} \rangle}{\|\vec{n}\|\|\vec{v}\|}.$$

Por un lado, un cálculo directo da

$$\langle \vec{v}, \vec{n} \rangle = \frac{\partial f(a,b)}{\partial x}(x-a) + \frac{\partial f(a,b)}{\partial y}(y-b) - (f(x,y) - f(a,b)) = o(\|h\|).$$

Por otro lado, como

$$\|\vec{n}\| = \sqrt{1 + \left(\frac{\partial f(a,b)}{\partial x}\right)^2 + \left(\frac{\partial f(a,b)}{\partial y}\right)^2} \geq 1,$$

y

$$\|\vec{v}\| = \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2} \geq \|h\|,$$

luego $|\cos\theta| \leq o(\|h\|)/\|h\| \to 0$ si $\|h\| \to 0$ y, por tanto, $\theta \to \pi/2$.

Lo anterior lo vemos gráficamente en la figura 2.2.3.

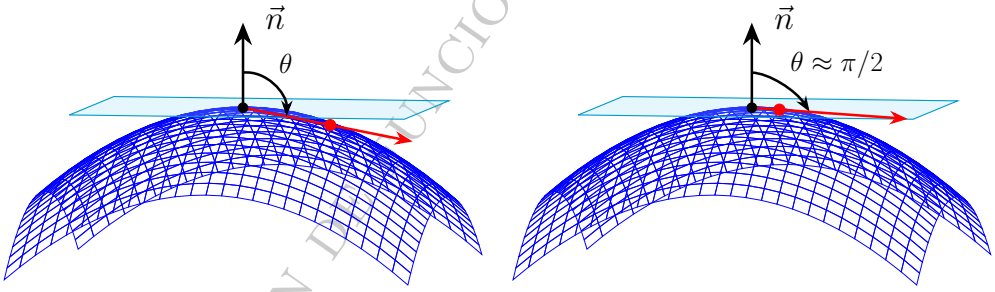


Figura 2.2.3: Plano tangente a la superficie definida por $z = f(x, y)$. El vector negro representa al vector normal $\vec{n}$ al plano (y a la superficie) en el punto $\xi = (a, b, c)$. Si el punto $s = (x, y, f(x, y))$ se acerca a (a, b, c), el ángulo entre $\vec{n}$ y $\vec{v} = (x - a, y - b, f(x, y) - f(a, b))$ se aproxima a $\pi/2$.

En otras palabras, el plano π es tangente a la superficie σ en el punto (a, b, c) como queríamos probar.

Sea ahora el caso de una función $f : \mathbb{R}^n \mapsto \mathbb{R}$. Supongamos que f es diferenciable en $a = (a_1, \ldots, a_n)$ y sea $c = f(a)$. Entonces, el vector normal

$$\vec{n} = \left(\frac{\partial f(a)}{\partial x_1}, \frac{\partial f(a)}{\partial x_2}, \ldots, \frac{\partial f(a)}{\partial x_n}, -1\right),$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

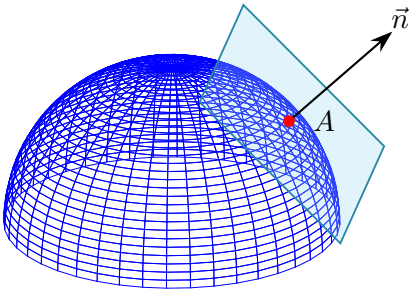


Figura 2.2.4: Plano tangente a una superficie definida por una función diferenciable. El vector representa al vector normal al plano (y a la superficie) en el punto A.

y denotando por $\vec{v} = (x_1 - a_1, \ldots, x_n - a_n, z - c)$, la ecuación del plano tangente a la superficie $z = f(x)$ vendrá dada por $\langle \vec{n}, \vec{v} \rangle = 0$, i.e.,

$$\frac{\partial f(a)}{\partial x_1}(x_1 - a_1) + \frac{\partial f(a)}{\partial x_2}(x_2 - a_2) + \cdots + \frac{\partial f(a)}{\partial x_n}(x_n - a_n) - (z - c) = 0.$$

En este caso el plano se denomina *hiperplano* tangente a la superficie $n + 1$ dimensional definida por f.

Para terminar este apartado, mostraremos en la figura 2.2.4 el plano tangente a la superficie

$$z = f(x, y) = \sqrt{1 - x^2 - y^2}$$

en el punto $A = (\sqrt{2}/2, 1/2, 1/2)$, dado por la ecuación

$$(x - \sqrt{2}/2)\sqrt{2} + (y - 1/2) + (z - 1/2) = 0,$$

siendo $v = (\sqrt{2}, 1, 1)$ el vector normal a la superficie en dicho punto.

2.2.1. Problemas propuestos

Problema 2.2.1 Calcula, si existen, los siguientes límites:

$$\lim_{(x,y)\to(0,2)} (1 + xy)^{\frac{2}{x^2+xy}}, \qquad \lim_{(x,y)\to(\infty,\infty)} \frac{2x + y}{2x^2 + 2xy + y^2},$$

$$\lim_{(x,y)\to(0,0)} \frac{x}{y} \operatorname{sen}(x^2 + y^2), \qquad \lim_{(x,y)\to(0,0)} \frac{(1 - e^x)(y - \operatorname{sen} y)}{x^2 + y^4},$$

$$\lim_{(x,y)\to(0,0)} \frac{xy}{|x| + |y|}, \qquad \lim_{(x,y)\to(0,0)} \frac{xy}{|x + y|}, \qquad \lim_{(x,y)\to(0,0)} \frac{x^a(y - \operatorname{sen} y)}{x^4 + y^4},$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

para $a = 1$ y $a = 2$.

Problema 2.2.2 Sea $f : A \subset \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = (x^2 - y^2)/(x^2 + y^2)$, $A = \{(x, y) \in \mathbb{R}^2 : |y| < x^2\}$, demostrar que existe $\lim_{(x,y)\to(0,0)} f(x, y)$ cuando $(x, y) \in A$. ¿Cuál es su valor? Compara el resultado con el del Ejemplo 2.1.3.

Problema 2.2.3 Prueba que existe el siguiente límite y calcúlalo:

$$\lim_{(x,y)\to(0,0)} f(x, y) = \frac{y(x - \arctan(x))}{x^4 + y^2}.$$

Problema 2.2.4 Estudiar la continuidad de las siguientes funciones (definidas como 0 donde no tenga sentido la expresión correspondiente):

$$f(x, y) = \frac{xy}{|x| + |y|}, \qquad f(x, y) = \frac{xy^2}{x^2 + y^4}, \qquad f(x, y) = \frac{x^3 y}{x^4 + 2y^2},$$

$$f(x, y) = \frac{y + \operatorname{sen}(x + y)}{2x + y}, \qquad f(x, y) = (x + y) \operatorname{sen} \frac{1}{x} \operatorname{sen} \frac{1}{y},$$

$$f(x, y) = \begin{cases} \dfrac{\operatorname{sen} x - \operatorname{sen} y}{\tan x - \tan y}, & \text{si } x \neq y, \\ \cos^3 x, & \text{en otro caso,} \end{cases} \qquad (x, y) \in \left(\frac{-\pi}{2}, \frac{\pi}{2}\right) \times \left(\frac{-\pi}{2}, \frac{\pi}{2}\right).$$

Problema 2.2.5 Sea la función

$$f(x, y) = \begin{cases} \dfrac{x^5}{(y - x^2)^2 + x^8}, & \text{si } (x, y) \neq (0, 0), \\ 0, & \text{si } (x, y) = (0, 0). \end{cases}$$

Calcula todas las derivadas direccionales de f en $(0, 0)$. La aplicación $u \to D_u f(0, 0)$, ¿es lineal? ¿Es f continua en $(0, 0)$? ¿y diferenciable en $(0, 0)$?

Problema 2.2.6 Sea la función

$$f(x, y) = \begin{cases} x^2 + 2y + \dfrac{2x^3 y}{x^4 + 3y^2}, & \text{si } (x, y) \neq (0, 0), \\ 0, & \text{si } (x, y) = (0, 0). \end{cases}$$

Estudia su continuidad y diferenciabilidad en $(0, 0)$. La aplicación $u \to D_u f(0, 0)$ ¿es lineal?

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Problema 2.2.7 Calcula la matriz de Jacobi (diferencial) si existe de la función $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}^2$

$$f(x,y) = \left(\arctan\left(\frac{x+y}{1-xy}\right), \quad \tan\left(\frac{x^2}{y}\right)\right).$$

Encuentra el mayor conjunto A donde está definida f. Calcula el valor de la diferencial en los puntos $(0,1)$ y $(\sqrt{\pi},1)$.

Problema 2.2.8 Encontrar la derivada de la función $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}^2$, $f(x,y) = \text{sen}\left(\frac{x^2+y^2}{1+x^2+y^2}\right)$ en el punto $A = (-1,1)$ según la dirección $\vec{u} = (-4,3)$.

Problema 2.2.9 Encontrar la derivada de la función $f : \{(x,y,z) \,|\, x > 0, y > 0, z > 0\} \mapsto \mathbb{R}^2$, $f(x,y,z) = (x/y)^z$ en el punto $(1,1,1)$ según la dirección $\vec{u} = (2,1,-1)$.

Problema 2.2.10 Encontrar la derivada de la función $f : \{(x,y,z) \,|\, z > 0\} \mapsto \mathbb{R}^2$, $f(x,y,z) = (x^2+y^2)^z$ en el punto $(1,0,1)$ según la dirección $\vec{u} = (1,-1,1)$.

Problema 2.2.11 Encuentra el plano tangente de las siguientes funciones:

1. $f(x,y) = 1 - x^2 - y^2$ en el punto $(-1,1)$.
2. $f(x,y) = \dfrac{x^2y^2}{x^2+y^2}$ en el punto $(-2,1)$.
3. $f(x,y) = \arctan(x^2+y^2)$ en el punto $(0,1)$.
4. $f(x,y) = (x^2+y^2)\sin(x^2+y^2)$ en los puntos $(0,0)$ y $(0,1)$.
5. $f(x,y) = \exp(-x^2-y^2)$ en los puntos $(0,0)$ y $(1/2,-1/2)$.

Problema 2.2.12 Sea la función

$$f(x,y) = \begin{cases} xy\,\text{sen}\left(\dfrac{1}{2y}\right), & \text{si } y \neq 0, \\ 0, & \text{si } y = 0. \end{cases}$$

Decide si es continua y diferenciable en $(0,0)$. Haz el mismo estudio para la función

$$f(x,y) = \begin{cases} \dfrac{x\,\text{sen}(y) - y\,\text{sen}(x)}{2x^2+y^2}, & \text{si } (x,y) \neq (0,0), \\ 0, & \text{si } (x,y) = (0,0). \end{cases}$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Problema 2.2.13 Sea $f(x,y) = (x^2+y^2)\operatorname{sen}\frac{1}{\sqrt{x^2+y^2}}$, $f(0,0)=0$. Probar que f es diferenciable en en $(0,0)$, pero sus derivadas parciales no son continuas en dicho punto.

Problema 2.2.14 Sea la función $f:\mathbb{R}^2\to\mathbb{R}$, $\alpha\geq 0$, definida por

$$f(x,y)=\begin{cases}\dfrac{|x|^\alpha \arctan y}{\sqrt{x^2+y^2}}, & \text{si } (x,y)\neq(0,0),\\ 0, & \text{si } (x,y)=(0,0).\end{cases}$$

1. Demuestra que para $\alpha>0$ la función f es continua en $(0,0)$.
2. Para $\alpha>0$, escribe los valores de $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial y}$ en $(0,0)$.
3. Demuestra que f es diferenciable en $(0,0)$ cuando $\alpha>1$. ¿Es f diferenciable en $(0,0)$, cuando $\alpha=1$? Razona la respuesta.

Problema 2.2.15 Sea $f:\mathbb{R}^2\setminus\{x=0\}\mapsto\mathbb{R}$ definida por

$$f(x,y)=\frac{|x|^\alpha\cos(x)\arctan(xy)}{\sqrt{3x^2+y^2}},\quad \alpha\in\mathbb{R}.$$

1. Demuestra que para $\alpha>-1$ existe el límite $\lim\limits_{(x,y)\to(0,0)} f(x,y)$.
2. ¿Qué valor ha de tomar f en $(0,0)$, $\alpha>-1$, para que sea continua en $(0,0)$? Justifica la respuesta
3. Sea $\alpha>-1$. Calcula, si es posible, las derivadas parciales de f en $(0,0)$.
4. ¿Para que valores de α f es diferenciable en $(0,0)$? Justifica la respuesta y calcula la derivada (diferencial) de f en $(0,0)$ si existe. ¿Es diferenciable para $\alpha=0$? Justifica la respuesta.

2.3. Otras propiedades de la diferenciación

Como ya vimos antes, si $f,g:A\subset\mathbb{R}^n\mapsto\mathbb{R}^m$, entonces cualquier combinación lineal de las mismas es diferenciable y

$$D(\alpha f+\beta g)(a)=\alpha Df(a)+\beta Dg(a),\quad \forall\alpha,\beta\in\mathbb{R}.$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

En el caso particular cuando $f, g : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ son diferenciables en a, entonces se tiene también que tanto el producto como el cociente también son diferenciables en a. Más concretamente,

$$D(fg)(a) = g(a)Df(a) + f(a)Dg(a),$$

$$D(f/g)(a) = \frac{g(a)Df(a) - f(a)Dg(a)}{(g(a))^2}, \qquad g(a) \neq 0$$

Dejamos al lector interesado la prueba de ambas expresiones.

Teorema 2.3.1 (Regla de la cadena) *Sean $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ y $g : B \subset \mathbb{R}^m \to \mathbb{R}^k$, A, B abiertos tales que $f(A) \subset B$. Supongamos que f es diferenciable en a y g es diferenciable en $f(a)$. Entonces la función compuesta $g \circ f : A \subset \mathbb{R}^n \to \mathbb{R}^k$ es diferenciable en a y $D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$. Lo anterior se puede escribir en coordenadas de la siguiente forma:*

$$D_j(g \circ f)_i(a) = \sum_{k=1}^{m} D_k g_i(f(a)) D_j f_k(a), \qquad \frac{\partial (g \circ f)_i(a)}{\partial x_j} = \sum_{l=1}^{m} \frac{\partial g_i(f(a))}{\partial x_l} \frac{\partial f_l(a)}{\partial x_j},$$

donde $i = 1, \ldots, k$, $j = 1, \cdots, n$. Matricialmente lo anterior se escribe como: $D(g \circ f)(a) = Dg(f(a)) \cdot Df(a)$ o $J_{g \circ f}(a) = J_g(f(a)) \cdot J_f(a)$.

Demostración: Vamos a probar que

$$(g \circ f)(a+h) - (g \circ f)(a) = g(f(a+h)) - g(f(a)) = [Dg(f(a)) \cdot Df(a)](h) + o(\|h\|),$$

de donde se sigue el teorema. Usaremos que f es diferenciable (y por tanto continua), que g es diferenciable y, por tanto, su diferencial Dg es una aplicación lineal. Calculamos

$$\begin{aligned} &g(f(a+h)) - g(f(a)) = Dg(f(a))(f(a+h) - f(a)) + o(\|f(a+h) - f(a)\|) \\ &= Dg(f(a))[Df(a)(h) + o(\|h\|)] + o(\|f(a+h) - f(a)\|) \\ &= [Dg(f(a)) \cdot Df(a)](h) + Dg(f(a))o(\|h\|) + o(\|f(a+h) - f(a)\|). \end{aligned}$$

Ahora bien, como $Dg(f(a))$ es acotada, entonces

$$\|Dg(f(a))v\| \leq \|Dg(f(a))\| \|v\|,$$

de donde se sigue que $Dg(f(a))o(\|h\|) = o(\|h\|)$. Por otro lado $f(a+h) - f(a) = Df(a)(h) + o(\|h\|)$, por tanto

$$\|f(a+h) - f(a)\| \leq \|Df(a)\| \|h\| + o(\|h\|) = \left(\|Df(a)\| + \frac{o(\|h\|)}{\|h\|} \right) \|h\|,$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

pero la expresión entre paréntesis de la última igualdad está acotada pues $Df(a)$ es acotada y $o(\|h\|)/\|h\| \to 0$ si $h \to 0$. De lo anterior se deduce que existe un $M > 0$ tal que $\|f(a+h) - f(a)\| \leq M\|h\|$, o equivalentemente, que $\|f(a+h) - f(a)\|$ es del mismo orden que $\|h\|$. Luego $o(\|f(a+h) - f(a)\|) = o(\|h\|)$, de donde se sigue el resultado. □

Una forma alternativa de probar que $o(\|f(a+h) - f(a)\|) = o(\|h\|)$ es la siguiente. Dividimos el entorno $B(a,h)$ en dos subconjuntos U_1 y U_2 tales que $f(a+h) - f(a) = 0$ y $f(a+h) - f(a) \neq 0$, respectivamente. En U_1 obviamente $o(\|f(a+h) - f(a)\|) = 0$. En U_2 hacemos lo siguiente:

$$\begin{aligned}\frac{|o(\|f(a+h) - f(a)\|)|}{\|h\|} &= \frac{|o(\|f(a+h) - f(a)\|)|}{\|f(a+h) - f(a)\|} \frac{\|f(a+h) - f(a)\|}{\|h\|} \\ &\leq \left(\|Df(a)\| + \frac{|o(\|h\|)|}{\|h\|}\right) \frac{|o(\|f(a+h) - f(a)\|)|}{\|f(a+h) - f(a)\|} \to 0\end{aligned}$$

si $h \to 0$. Luego, en $B(a,h)$, $o(\|f(a+h) - f(a)\|) = o(\|h\|)$.

Ejemplo 2.3.2 Sean las función $g : \mathbb{R}^2 \mapsto \mathbb{R}^2$, $g(x,y) = (x^2 + y^2, 2x + y)$, y $f : \mathbb{R}^2 \mapsto \mathbb{R}^3$, $f(u,v) = (u^2, u + v, v^2)$. Estudiemos la función $f \circ g$.

Esta claro que existe la función compuesta $h(x,y) = (f \circ g)(x,y) : \mathbb{R}^2 \mapsto \mathbb{R}^3$, y $h(x,y) = ((x^2+y^2)^2, x^2 + y^2 + 2x + y, (2x+y)^2)$. Es *obvio* que h es diferenciable en todo $\mathbb{R}^2$ y en particular en $(1,0)$. Su derivada (matriz de Jacobi) viene dada por (2.2.4)

$$Dh(1,0) = \left.\begin{pmatrix} 4x(x^2+y^2) & 4y(x^2+y^2) \\ 2x+2 & 2y+1 \\ 4(2x+y) & 2(2x+y) \end{pmatrix}\right|_{(x,y)=(1,0)} = \begin{pmatrix} 4 & 0 \\ 4 & 1 \\ 8 & 4 \end{pmatrix}.$$

Por otro lado tanto f como g son diferenciables (g en $(0,1)$ y f en $(1,2)$) y

$$Dg(1,0) = \left.\begin{pmatrix} 2x & 2y \\ 2 & 1 \end{pmatrix}\right|_{(x,y)=(1,0)} = \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix},$$

$$Df(1,2) = \left.\begin{pmatrix} 2u & 0 \\ 1 & 1 \\ 0 & 2v \end{pmatrix}\right|_{(u,v)=(1,2)} = \begin{pmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 4 \end{pmatrix}.$$

Un cálculo directo muestra que $Dh(1,0) = D(f \circ g)(1,0) = Df(1,2) \cdot Dg(1,0)$.□

Como ya hemos visto la existencia de derivadas parciales en un punto no implica la diferenciabilidad de f en dicho punto. No obstante imponiendo ciertas condiciones extra se puede probar la diferenciabilidad de f. De hecho se tiene el siguiente teorema:

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Teorema 2.3.3 (Condición suficiente de diferenciabilidad I) *Sea la función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, con A abierto y sea $a \in A$. Supongamos que existen las derivadas parciales de cada una de las componentes de f en un entorno de a con respecto a cada una de las variables y son continuas en a, entonces f es diferenciable en a.*

Demostración: Para probarlo es suficiente considerar el caso de una función escalar (¿por qué?).

$$\begin{aligned}
f(a+h) - f(a) &= f(a_1+h_1, a_2+h_2, \cdots, a_n+h_n) - f(a_1, a_2, \cdots, a_{n-1}, a_n) \\
&= [f(a_1+h_1, a_2+h_2, \cdots, a_{n-1}+h_{n-1}, a_n+h_n) - f(a_1, a_2+h_2, \cdots, a_n+h_n)] \\
&\quad + [f(a_1, a_2+h_2, \cdots, a_n+h_n) - f(a_1, a_2, a_3+h_3, \cdots, a_n+h_n)] \\
&\quad + \qquad \cdots \\
&\quad + [f(a_1, a_2, \cdots, a_{n-1}+h_{n-1}, a_n+h_n) - f(a_1, a_2, \cdots, a_{n-1}, a_n+h_n)] \\
&\quad + [f(a_1, a_2, \cdots, a_{n-1}, a_n+h_n) - f(a_1, a_2, \cdots, a_{n-1}, a_n)].
\end{aligned}$$

A cada uno de los sumandos entre corchetes se le puede aplicar el Teorema del valor medio de Lagrange 1.1.19 (sólo cambia una única variable y f como función de cada variable es continua y diferenciable en todo un intervalo). Así pues

$$f(a+h) - f(a) = \sum_{k=1}^{n} \frac{\partial f(y_k)}{\partial a_k} h_k,$$

donde $y_1 = (\xi_1, a_2+h_2, a_3+h_3, \ldots, a_n+h_n)$, $y_2 = (a_1, \xi_2, a_3+h_3, \ldots, a_n+h_n)$, $\ldots$, $y_n = (a_1, a_2, a_3, \ldots, \xi_n)$, con $\xi_k \in (a_k, a_k+h_k)$, $k = 1, 2, \ldots, n$. Calculamos ahora

$$\begin{aligned}
f(a+h) - f(a) - Df(a)(h) &= f(a+h) - f(a) - \sum_{k=1}^{n} \frac{\partial f(a)}{\partial x_k} h_k \\
&= \sum_{k=1}^{n} \left[\frac{\partial f(y_k)}{\partial x_k} - \frac{\partial f(a)}{\partial x_k} \right] h_k.
\end{aligned}$$

Teniendo en cuenta que $|h_k| \leq \|h\|$ para todo $k = 1, \ldots, n$, tenemos entonces que

$$\frac{|f(a+h) - f(a) - Df(a)(h)|}{\|h\|} \leq \sum_{k=1}^{n} \left| \frac{\partial f(y_k)}{\partial x_k} - \frac{\partial f(a)}{\partial x_k} \right| \to 0 \text{ cuando } h \to 0,$$

pues $y_k \to a$, $k = 1, \ldots, n$ cuando $h \to 0$ y las derivadas parciales son continuas por hipótesis. □

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Las condiciones del teorema 2.3.3 son suficientes pero no necesarias. En efecto si escogemos la función

$$f(x,y) = (x^2+y^2)\,\operatorname{sen}\left(\frac{1}{x^2+y^2}\right), \quad f(0,0)=0. \tag{2.3.1}$$

es fácil ver que es continua en $(0,0)$ pues la función $\operatorname{sen}(z)$ es acotada y x^2+y^2 tiende a cero cuando $(x,y)\to(0,0)$.

Por otro lado,

$$\frac{\partial f}{\partial x}(0,0) = \lim_{x\to 0}\frac{f(x,0)-f(0,0)}{x} = \lim_{x\to 0}\frac{x^2\sin(1/x^2)}{x} = 0,$$

y análogamente (basta cambiar x por y) $\frac{\partial f}{\partial y}(0,0)=0$, luego, si f fuese diferenciable en $(0,0)$, su diferencial sería la matriz $Df(0,0)=[0,0]$. Así, usando la definición de diferenciabilidad 2.2.3 tendríamos que probar que

$$\lim_{(x,y)\to(0,0)}\frac{f(x,y)-f(0,0)-Df(0,0)\begin{pmatrix}x\\y\end{pmatrix}}{\sqrt{x^2+y^2}} = 0.$$

Pero

$$\frac{f(x,y)-f(0,0)-Df(0,0)\begin{pmatrix}x\\y\end{pmatrix}}{\sqrt{x^2+y^2}} = \sqrt{x^2+y^2}\,\operatorname{sen}\left(\frac{1}{x^2+y^2}\right)\to 0,$$

luego f es diferenciable el $(0,0)$.

Calculemos la derivada parcial $\frac{\partial f}{\partial x}$

$$\frac{\partial f}{\partial x}(x,y) = 2x\sin\left(\frac{1}{x^2+y^2}\right) - \frac{2x}{x^2+y^2}\cos\left(\frac{1}{x^2+y^2}\right).$$

Está claro que el primer sumando de $\frac{\partial f}{\partial x}$ tiende a cero cuando $(x,y)\to(0,0)$, sin embargo el segundo no tiene límite pues la función $\cos(1/z)$ no tiene límite si $z\to 0$, de hecho oscila entre -1 y 1, y la función $2x/(x^2+y^2)$ ni siquiera es acotada en un entorno de $(0,0)$ (tómese la trayectoria $y=x$ y se obtiene que dicha función se comporta como $1/x$ que es no acotada). Es decir, la derivada parcial $\frac{\partial f}{\partial x}$ no es continua en $(0,0)$. De forma análoga se prueba que $\frac{\partial f}{\partial y}$ tampoco es continua en $(0,0)$, lo que dejamos como ejercicio al lector. Así, las derivadas parciales de la función 2.3.1 están bien definidas en todo $\mathbb{R}^2$ pero no son continuas en $(0,0)$, sin embargo, como ya vimos, f es diferenciable en $(0,0)$ siendo su derivada el operador con matriz jacobiana $(0\ 0)$.

El Teorema 2.3.3 se puede generalizar como sigue:

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Teorema 2.3.4 (Condición suficiente de diferenciabilidad II) *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, con A abierto y sea $a \in A$. Si existe la derivada parcial de f en a con respecto a una de las variables y las restantes $n-1$ derivadas parciales existen en un entorno de a y son continuas en a, entonces f es diferenciable en a. En el caso de funciones con valores en $\mathbb{R}^m$, el teorema se aplica asumiendo las hipótesis para cada una de sus componentes.*

Las condiciones del Teorema 2.3.4 son suficientes. La misma función (2.3.1) de antes nos vale para probar que no son necesarias.

Demostración: Nuevamente vamos a probarlo para el caso escalar. Sin pérdida de generalidad asumiremos que las derivadas parciales de las primeras $n-1$ variables son continuas y que la derivada respecto a la última variable existe. Repitiendo el proceso de antes, y usando la misma notación que en la prueba del Teorema 2.3.3, tenemos

$$\begin{aligned} f(a+h) - f(a) &= f(a_1+h_1, a_2+h_2, \cdots, a_n+h_n) - f(a_1, a_2, \cdots, a_n) \\ &= \sum_{k=1}^{n-1} \frac{\partial f(y_k)}{\partial x_k} h_k + [f(a_1, a_2, \cdots, a_{n-1}, a_n+h_n) - f(a_1, a_2, \cdots, a_{n-1}, a_n)]. \end{aligned}$$

Claramente la expresión dentro del corchete se puede considerar (de hecho lo es) una función de la última variable $f(a_1, a_2, \cdots, x_n)$. Ahora bien, como existe la derivada parcial respecto a x_n entonces, $f(a_1, a_2, \cdots, a_n)$ es derivable respecto a x_n y por tanto diferenciable como función de la variable x_n, luego se tiene, $h_n \neq 0$,

$$f(a_1, a_2, \cdots, a_{n-1}, a_n+h_n) - f(a_1, a_2, \cdots, a_{n-1}, a_n) = \frac{\partial f(a)}{\partial x_n} h_n + o(h_n).$$

Juntando esto con lo anterior obtenemos

$$f(a+h) - f(a) - Df(a)(h) = \sum_{k=1}^{n-1} \left[\frac{\partial f(y_k)}{\partial x_k} - \frac{\partial f(a)}{\partial x_k} \right] h_k + o(h_n).$$

de donde, usando nuevamente que $|h_k| \leq \|h\|$, se sigue que

$$\frac{|f(a+h) - f(a) - Df(a)(h)|}{\|h\|} \leq \sum_{k=1}^{n-1} \left| \frac{\partial f(y_k)}{\partial x_k} - \frac{\partial f(a)}{\partial x_k} \right| + \frac{o(h_n)}{\|h\|} \to 0$$

si $h \to 0$. □

A continuación enunciaremos el análogo del Teorema del valor medio 1.1.19 para el caso de varias variables.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Definición 2.3.5 *Diremos que un abierto $A \in \mathbb{R}^n$ es convexo si, dados dos puntos a y b cualesquiera de A, el segmento $s = \{(1-t)a + tb : t \in [0,1]\}$ que los une también pertenece a A.*



Figura 2.3.1: Conjunto convexo (izquierda) y no convexo (derecha).

Teorema 2.3.6 (del valor medio) *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, diferenciable en A abierto y convexo. Sean $a, b \in A$ y sea $s = \{(1-t)a + tb : t \in [0,1]\}$ el segmento que los une. Entonces, para cada vector $v \in \mathbb{R}^m$ existe un punto z en el interior del segmento s tal que*

$$\langle v, f(b) - f(a)\rangle = \langle v, Df(z)(b-a)\rangle,$$

donde $\langle \cdot, \cdot \rangle$ denota el producto escalar en $\mathbb{R}^m$.

Demostración: Primero hemos de notar que como A es abierto y convexo, y $a, b \in s \subset A$, entonces el conjunto $U = \{x = a + (b-a)t, t \in [0,1]\} \subset A$. Sea $v \in \mathbb{R}^m$ cualquier vector distinto de cero. Definamos la función $F : [0,1] \mapsto \mathbb{R}$ (por f_k denotamos las componentes de f) definida por

$$F(t) = \langle v, f(a + (b-a)t)\rangle = \sum_{k=1}^{m} v_k f_k(a + (b-a)t).$$

Como f es diferenciable en A, entonces la función compuesta $f(a + (b-a)t)$ también lo es (¿por qué?), luego F es continua en $[0,1]$ y diferenciable en $(0,1)$ por lo que podemos aplicar el Teorema del valor medio de Lagrange de una variable 1.1.19. Así, $F(1) - F(0) = \langle v, f(b) - f(a)\rangle = F'(\xi)$ con $\xi \in (0,1)$. Pero, usando la regla de la cadena (ver Teorema 2.3.1) y las propiedades de la diferenciación de la suma y multiplicación

$$\frac{d}{dt} f_k(a + (b-a)t) = \sum_{j=1}^{n} \frac{\partial}{\partial x_j} f_k(a + (b-a)t) \cdot (b_j - a_j) = Df_k[a + (b-a)t](b-a).$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Así que

$$F'(\xi) = \sum_{k=1}^{m} v_k \frac{d}{dt} f_k(a + (b-a)\xi) = \sum_{k=1}^{m} v_k Df_k[a + (b-a)\xi](b-a)$$
$$= \langle v, Df(z)(b-a)\rangle, \quad z = a + (b-a)\xi.$$

Como $\xi \in (0,1)$ entonces z pertenece al interior de s. □

Nótese que para la prueba del Teorema del valor medio 2.3.6 basta con que el segmento $s = \{(1-t)a + tb : t \in [0,1]\}$ pertenezca a A.

Una consecuencia inmediata del teorema anterior (la prueba se deja como ejercicio) es la siguiente propiedad:

Si $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ (o sea, f es una función escalar) y f es diferenciable en A, abierto y convexo, entonces existe un punto z en el interior del segmento s que une a con b tal que

$$f(b) - f(a) = Df(z)(b-a) = \langle \nabla f(z), b-a\rangle.$$

Aplicando lo anterior a cada componente de una función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, diferenciable en A abierto y convexo, se tiene que existen los puntos $z_1, \ldots, z_n$ en el interior del segmento s que une a con b tales que

$$f_k(b) - f_k(a) = Df_k(z_k)(b-a) = \langle \nabla f(z_k), b-a\rangle, \quad k = 1, 2, \ldots, m.$$

Nótese los z_k anteriores no tienen por qué coincidir unos con otros.

Como corolarios del Teorema del valor medio 2.3.6 tenemos:

Corolario 2.3.7 *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, diferenciable en A abierto y convexo. Entonces, existe un $M > 0$ tal que para todos a y b de A*

$$\|f(b) - f(a)\| \leq M\|b-a\|.$$

Demostración: Asumiremos que $f(b) \neq f(a)$. Escogiendo $v = f(b) - f(a)$ y obtenemos

$$\|f(b) - f(a)\|^2 = \langle f(b) - f(a), Df(z)(b-a)\rangle \leq \|f(b) - f(a)\|\|Df(z)(b-a)\|.$$

Usando que $Df(z)$ es acotada (ver Teorema 1.5.3), es decir, que existe $M > 0$ tal que $\|Df(z)\| \leq M$, y que $\|f(b) - f(a)\| > 0$, obtenemos

$$\|f(b) - f(a)\| \leq \|Df(z)(b-a)\| \leq \|Df(z)\|\|b-a\| \leq M\|b-a\|,$$

de donde se sigue el resultado deseado. □

Nótese que para que todos los corolarios anteriores sean ciertos basta con que el segmento $s = \{(1-t)a + tb : t \in [0,1]\}$ pertenezca a A. No así el siguiente, donde la convexidad de A juega un papel determinante.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Corolario 2.3.8 *Sea A un conjunto abierto y convexo y sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ una función diferenciable en A tal que $Df(x) = 0$, para todo $x \in A$, entonces f es constante en A.*

Demostración: Del Teorema 2.3.6 se sigue que para cualesquiera a y b en A y $v \in \mathbb{R}^m$,

$$\langle v, f(b) - f(a)\rangle = \langle v, Df(z)(b-a)\rangle = \langle v, 0\rangle = 0.$$

Escogiendo $v = f(b) - f(a)$ se sigue que $\|f(b) - f(a)\|^2 = 0$ luego $f(a) = f(b)$ para cualesquiera sean a y b de A. □

2.3.1. Problemas propuestos

Problema 2.3.1 Sea la función

$$f(x,y) = \begin{cases} f(x,y) = x\sqrt{x^2+y^2}\ \text{sen}\dfrac{1}{x^2+y^2}, & \text{si } (x,y) \neq (0,0), \\ 0, & \text{si } (x,y) = (0,0), \end{cases}$$

Estudia la continuidad, existencia de derivadas parciales y diferenciabilidad de f en $\mathbb{R}^2$.

Problema 2.3.2 Dada la función con $a > 0$

$$f(x,y) = \begin{cases} \dfrac{|xy|^a}{x^2 \pm xy + y^2}, & \text{si } (x,y) \neq (0,0), \\ 0, & \text{si } (x,y) = (0,0), \end{cases}$$

estudia su continuidad y diferenciabilidad en $\mathbb{R}^2$. Repite el estudio para la función

$$f(x,y) = \begin{cases} \dfrac{x^4+y^4}{x^2 \pm xy + y^2}, & \text{si } (x,y) \neq (0,0), \\ 0, & \text{si } (x,y) = (0,0). \end{cases}$$

Problema 2.3.3 Sea

$$f(x,y) = \frac{1-\cos xy}{y^2}, \quad \text{si } y \neq 0.$$

¿Como definir f en $y = 0$ para que sea continua en $\mathbb{R}^2$? Calcula las derivadas parciales de f y prueba que son continuas en todo $\mathbb{R}^2$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Problema 2.3.4 Sea $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \begin{cases} \dfrac{|x|^\alpha}{2x^2+y^2}, & \text{si } (x,y) \neq (0,0), \\ 0, & \text{si } (x,y) = (0,0). \end{cases} \qquad \alpha \geq 0.$$

Estudia la continuidad y diferenciabilidad de f en $\mathbb{R}^2$ en función de α. Hacer el mismo estudio para la función $f(x,y) = \dfrac{2|x|^\alpha y}{x^2+y^2}$ si $(x,y) \neq (0,0)$, definiéndola, si es posible, en $(0,0)$ de forma que sea continua en $\mathbb{R}^2$.

2.4. Derivadas de orden superior

Veamos ahora las derivadas de orden superior. Comenzaremos con las derivadas parciales. Supongamos que $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, A es un abierto de $\mathbb{R}^n$ tiene derivadas parciales $D_i f = \dfrac{\partial f(x)}{\partial x_i}$ en A, $i = 1, \ldots, n$. Supongamos que dichas derivadas parciales $D_i f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ admiten a su vez derivadas parciales $D_j(\cdot)$ en A. Dichas derivadas parciales se denominan derivadas parciales de segundo orden y se denotan por

$$D_j(D_i f)(x) = D_{j,i} f(x) = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}, \quad i,j = 1,2,\ldots,n.$$

Si las funciones $D_{j,i} f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ admiten derivadas parciales entonces podemos definir las derivadas parciales de orden 3

$$D_k(D_j(D_i f))(x) = D_{k,j,i} f(x) = \frac{\partial^3 f(x)}{\partial x_k \partial x_j \partial x_i}, \quad i,j,k = 1,2,\ldots,n.$$

Y así, sucesivamente.

Una pregunta natural es cuando las derivadas cruzadas son iguales, i.e.,

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

Veamos un ejemplo.

Ejemplo 2.4.1 Sea $f : \mathbb{R}^2 \mapsto \mathbb{R}$ definida por

$$f(x,y) = \begin{cases} x^2 \arctan \dfrac{y}{x} - y^2 \arctan \dfrac{x}{y}, & \text{si } xy \neq 0, \\ 0, & \text{si } xy = 0. \end{cases}$$

Probar que

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Nótese que como la función arctan es acotada, entonces cuando $(x,y) \to (0,0)$ $f(x,y) \to 0$, luego f es continua en $(0,0)$ y, por tanto, en todo $\mathbb{R}^2$.

Se puede comprobar además que en todo $\mathbb{R}^2 \setminus \{(0,0)\}$ las derivadas cruzadas de esta función son iguales a

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

Veamos que ocurre en el punto $(0,0)$. Ante todo notemos que, para $xy \neq 0$,

$$\frac{\partial f(x,y)}{\partial x} = -y + 2x \arctan \frac{y}{x}, \qquad \frac{\partial f(x,y)}{\partial y} = x - 2y \arctan \frac{x}{y}.$$

Además, usando la definición, obtenemos

$$\frac{\partial f(x,0)}{\partial y} = \lim_{h\to 0} \frac{f(x,h)-f(x,0)}{h} = \lim_{h\to 0} \frac{x^2 \arctan h/x}{h} - \underbrace{\lim_{h\to 0} h \arctan \frac{x}{h}}_{\to 0} = x.$$

El cálculo de $\frac{\partial f(0,y)}{\partial x}$ es análogo y se deja como ejercicio. De esta forma tenemos que

$$\frac{\partial f(x,0)}{\partial y} = x, \qquad \frac{\partial f(0,y)}{\partial x} = -y.$$

Usando lo anterior calculamos las derivadas cruzadas en $(0,0)$. Así, tenemos

$$\frac{\partial^2 f(0,0)}{\partial x \partial y} = \frac{\partial}{\partial x}\left[\frac{\partial f(0,0)}{\partial y}\right] = \lim_{h\to 0} \frac{1}{h}\left[\frac{\partial f(h,0)}{\partial y} - \frac{\partial f(0,0)}{\partial y}\right] = \lim_{h\to 0} \frac{-h}{h} = -1,$$
$$\frac{\partial^2 f(0,0)}{\partial y \partial x} = \frac{\partial}{\partial y}\left[\frac{\partial f(0,0)}{\partial x}\right] = \lim_{h\to 0} \frac{1}{h}\left[\frac{\partial f(0,h)}{\partial x} - \frac{\partial f(0,0)}{\partial x}\right] = \lim_{h\to 0} \frac{h}{h} = 1.$$

Luego, en el punto $(0,0)$ se obtiene

$$\frac{\partial^2 f(0,0)}{\partial x \partial y} \neq \frac{\partial^2 f(0,0)}{\partial y \partial x}.$$

Es decir, no siempre las derivadas cruzadas de una función son iguales. □

No obstante, bajo ciertas condiciones *razonables* se tiene el siguiente resultado:

Teorema 2.4.2 (Schwarz) *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ con A abierto, y sea $x_0 \in A$. Si en A existen las derivadas parciales $\frac{\partial f(x)}{\partial x_i}$, $\frac{\partial f(x)}{\partial x_j}$ y $\frac{\partial^2 f(x)}{\partial x_j \partial x_i}$ y la derivada $\frac{\partial^2 f(x)}{\partial x_j \partial x_i}$ es continua en x_0, entonces en $x_0 \in A$ existe la derivada $\frac{\partial^2 f(x_0)}{\partial x_i \partial x_j}$ y*

$$\frac{\partial^2 f(x_0)}{\partial x_j \partial x_i} = \frac{\partial^2 f(x_0)}{\partial x_i \partial x_j}.$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Demostración: Sin pérdida de generalidad vamos a probarlo para una función escalar de dos variables (¿por qué?).

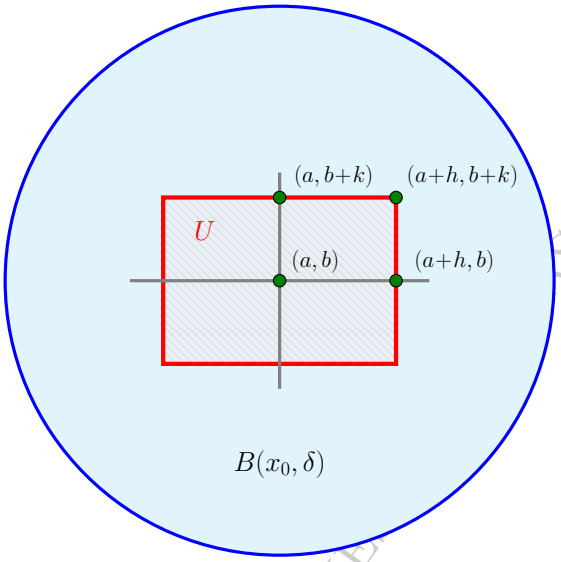


Figura 2.4.1: El entorno U y la bola $B(x_0, \delta) \supset U$.

Sean h y k suficientemente pequeños de forma que el rectángulo $U = \{(x, y),\ x \in [a-h, a+h], y \in [b-k, b+k]\}$ esté contenido en un δ-entorno del punto $x_0 = (a, b) \in A$ (ver figura 2.4.1). En la prueba jugará un papel esencial la cantidad

$$\Delta f(a,b) = [f(a+h, b+k) - f(a+h, b)] - [f(a, b+k) - f(a, b)].$$

Sea la función $u(x) = f(x, b+k) - f(x, b)$, definida en U. Entonces

$$\Delta f(a,b) = u(a+h) - u(a).$$

Por hipótesis existe la derivada parcial de f respecto a x en todo A, luego f, como función de x, es continua y diferenciable en A (basta que lo sea en un entorno del punto x_0) y, por tanto, u también lo es así que podemos aplicar el Teorema del valor medio de Lagrange para funciones de una variable 1.1.19 a $u(x)$,

$$\Delta f(a,b) = u(a+h) - u(a) = u'(\xi)h = \left(\frac{\partial f(\xi, b+k)}{\partial x} - \frac{\partial f(\xi, b)}{\partial x}\right) h,$$

donde $\xi \in (a, a+h)$. Por hipótesis existe la derivada $\frac{\partial}{\partial y}\left[\frac{\partial f(\xi,y)}{\partial x}\right]$, luego a la función $\frac{\partial f(\xi,y)}{\partial x}$ podemos aplicarle el Teorema del valor medio de Lagrange 1.1.19

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

en la variable y que nos conduce a

$$\Delta f(a,b) = hk\frac{\partial}{\partial y}\left(\frac{\partial f(\xi,\eta)}{\partial x}\right) = hk\frac{\partial^2 f(\xi,\eta)}{\partial y \partial x}, \quad \xi \in (a, a+h),\ \eta \in (b, b+k).$$

Como $\frac{\partial^2 f(x,y)}{\partial y \partial x}$ es continua en (a,b) entonces para cualquier $\epsilon > 0$ existe una bola $B(x_0,\delta)$ tal que para todos $(x,y) \in B(x_0,\delta)$

$$\left|\frac{\partial^2 f(x,y)}{\partial y \partial x} - \frac{\partial^2 f(a,b)}{\partial y \partial x}\right| < \frac{\epsilon}{2}.$$

Por tanto, si escogemos $U \subset B(x_0,\delta)$ (el rectángulo definido anteriormente) tenemos que $(\xi,\eta) \in U$, y la desigualdad anterior nos conduce a

$$\left|\frac{\partial^2 f(\xi,\eta)}{\partial y \partial x} - \frac{\partial^2 f(a,b)}{\partial y \partial x}\right| = \left|\frac{\Delta f(a,b)}{hk} \frac{\partial^2 f(a,b)}{\partial y \partial x}\right| < \frac{\epsilon}{2}. \tag{2.4.1}$$

Como existe $\dfrac{\partial f}{\partial y}$, entonces

$$\begin{aligned}\lim_{k\to 0}\frac{\Delta f(a,b)}{hk} =& \lim_{k\to 0}\frac{1}{h}\left[\frac{f(a+h,b+k)-f(a+h,b)}{k} - \frac{f(a,b+k)-f(a,b)}{k}\right]\\ =& \frac{1}{h}\left[\frac{\partial f(a+h,b)}{\partial y} - \frac{\partial f(a,b)}{\partial y}\right].\end{aligned}$$

Luego, tomando el límite $k \to 0$ en (2.4.1), tenemos

$$\left|\frac{1}{h}\left[\frac{\partial f(a+h,b)}{\partial y} - \frac{\partial f(a,b)}{\partial y}\right] - \frac{\partial^2 f(a,b)}{\partial y \partial x}\right| \leq \frac{\epsilon}{2} < \epsilon.$$

Lo anterior indica que existe el el límite $h \to 0$ de

$$\lim_{h\to 0}\frac{1}{h}\left[\frac{\partial f(a+h,b)}{\partial y} - \frac{\partial f(a,b)}{\partial y}\right] = \frac{\partial^2 f(a,b)}{\partial y \partial x}$$

pero, por definición, si existe el límite del miembro izquierdo este es igual a $\frac{\partial^2 f(a,b)}{\partial x \partial y}$, lo que demuestra el teorema. □

El teorema anterior solamente da condiciones suficientes como veremos más adelante.

Un corolario extremadamente útil del Teorema de Schwarz es el siguiente resultado:

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Corolario 2.4.3 (Bonnet) *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ con A abierto, y sea $a \in A$ tal que existen las derivadas parciales $\frac{\partial^2 f(x)}{\partial x_j \partial x_i}$ y $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$ en un entorno de $x_0 \in A$ y ambas son continuas en a. Entonces*

$$\frac{\partial^2 f(x_0)}{\partial x_j \partial x_i} = \frac{\partial^2 f(x_0)}{\partial x_i \partial x_j}.$$

Finalmente, probemos otra condición suficiente distinta a la anterior.

Teorema 2.4.4 (Heffter-Young) *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, A abierto y sea $a \in A$. Supongamos que existen las derivadas parciales $\frac{\partial f(x)}{\partial x_i}$, y $\frac{\partial f(x)}{\partial x_j}$ en un entorno de x_0 y son diferenciables en x_0. Entonces*

$$\frac{\partial^2 f(x_0)}{\partial x_j \partial x_i} = \frac{\partial^2 f(x_0)}{\partial x_i \partial x_j}.$$

Demostración: Nuevamente, y sin pérdida de generalidad, vamos a probarlo para una función escalar de dos variables y sea $x_0 = (a, b)$. Definamos la cantidad

$$\Delta f(a,b) = [f(a+h, b+h) - f(a+h, b)] - [f(a, b+h) - f(a,b)],$$

y sea la función $G(x) = f(x, b+h) - f(x,b)$. Está claro que $\Delta f(a,b) = G(a+h) - G(a)$ y que G es una función de una variable que por condición es continua y diferenciable en un entorno de (a,b). luego existe un $\xi \in (a, a+h)$ tal que

$$\Delta f(a,b) = hG'(\xi) = h\left[\frac{\partial f(\xi, b+h)}{\partial x} - \frac{\partial f(\xi, b)}{\partial x}\right].$$

Como $\frac{\partial f}{\partial x}$ es diferenciable por hipótesis entonces

$$\begin{aligned}\frac{\partial f(\xi, b+h)}{\partial x} =& \frac{\partial f(a,b)}{\partial x} + \frac{\partial}{\partial x}\left(\frac{\partial f(a,b)}{\partial x}\right)(\xi - a) + \frac{\partial}{\partial y}\left(\frac{\partial f(a,b)}{\partial x}\right)h \\ &+ o\left(\sqrt{(\xi-a)^2 + h^2}\right), \\ \frac{\partial f(\xi, b)}{\partial x} =& \frac{\partial f(a,b)}{\partial x} + \frac{\partial}{\partial x}\left(\frac{\partial f(a,b)}{\partial x}\right)(\xi - a) + o(\xi - a),\end{aligned}$$

que, teniendo en cuenta que $o\left(\sqrt{(\xi-a)^2 + h^2}\right) = o(h)$ y que $o(\xi - a) = o(h)$ (pues $|\xi - a| < |h|$, y por tanto podemos tomar $\xi - a = th$ con $|t| < 1$) nos conduce a

$$\Delta f(a,b) = \frac{\partial}{\partial y}\left(\frac{\partial f(a,b)}{\partial x}\right)h^2 + ho(h),$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

de donde se sigue que

$$\lim_{h\to 0}\frac{\Delta f(a,b)}{h^2}=\frac{\partial}{\partial y}\left(\frac{\partial f(a,b)}{\partial x}\right)=\frac{\partial^2 f(a,b)}{\partial y\partial x}.$$

Sea ahora la función $H(y)=f(a+h,y)-f(a,y)$. Como $\Delta f(a,b)=H(b+h)-H(b)$ tenemos, de forma análoga que existe un $\xi\in(b,b+h)$ tal que

$$\Delta f(a,b)=hH'(\xi)=h\left[\frac{\partial f(a+h,\xi)}{\partial y}-\frac{\partial f(a,\xi)}{\partial y}\right].$$

Si usamos la diferenciabilidad de $\frac{\partial f}{\partial y}$ obtenemos, de forma completamente análoga

$$\Delta f(a,b)=\frac{\partial}{\partial x}\left(\frac{\partial f(a,b)}{\partial y}\right)h^2+ho(h)\implies\lim_{h\to 0}\frac{\Delta f(a,b)}{h^2}=\frac{\partial}{\partial x}\left(\frac{\partial f(a,b)}{\partial y}\right),$$

de donde, por la unicidad del límite, se sigue el resultado. □

Tanto el Teorema de Schwarz como el de Heffter-Young dan condiciones suficientes. Para comprobarlo definamos la función

$$f(x,y)=\begin{cases}x^2y\cos\dfrac{1}{x}, & \text{si } x\neq 0,\\ 0, & \text{si } x=0.\end{cases}$$

que cumple con las condiciones de ambos teoremas en $\mathbb{R}^2\setminus\{(0,y),\forall y\in\mathbb{R}\}$. Veamos que ocurre en $(0,0)$. Es fácil comprobar que

$$\frac{\partial f(0,y)}{\partial x}=\frac{\partial f(0,y)}{\partial y}=0,\quad\frac{\partial^2 f(0,y)}{\partial y\partial x}=\frac{\partial}{\partial y}\left(\frac{\partial f(0,y)}{\partial x}\right)=\frac{\partial}{\partial y}0=0,$$

$$\frac{\partial^2 f(0,y)}{\partial x\partial y}=\lim_{x\to 0}\frac{1}{x}\left[\frac{\partial f(x,y)}{\partial y}-\frac{\partial f(0,y)}{\partial y}\right]=\lim_{x\to 0}x\sin\frac{1}{x}=0.$$

Luego

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

Ahora bien, un sencillo cálculo nos muestra que

$$\frac{\partial^2 f(x,y)}{\partial y\partial x}=2x\cos\frac{1}{x}+\sin\frac{1}{x},$$

que no es continua en $(0,y)$, y en particular en $(0,0)$, luego las condiciones del Teorema de Schwarz no son necesarias. Por otro lado, si consideramos la función

$$\frac{\partial f}{\partial x}=2xy\cos\frac{1}{x}+y\sin\frac{1}{x},$$

esta no es diferenciable en $(0,0)$, luego tampoco las condiciones del Teorema de Heffter-Young son necesarias.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

2.5. Diferenciales de orden superior

Supongamos que $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en A. Entonces, para cada $x \in A$ podemos definir la derivada (diferencial) de f en A. Denotemos por $\mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$ al espacio de todas las aplicaciones lineales acotadas de $\mathbb{R}^n$ en $\mathbb{R}^m$. Como para todo $x \in A$, $Df(x)$ es una aplicación lineal y acotada (ver Teorema 1.5.3), i.e., $Df(x) \in \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$, entonces podemos definir la aplicación $Df : A \subset \mathbb{R}^n \mapsto \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$. Recordemos que, como ya mencionamos en el apartado 1.4, el espacio $\mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$ es un espacio normado[3] (en la norma de las aplicaciones lineales (1.5.4)).

Antes definir las derivadas (diferenciales) de orden superior conviene generalizar el concepto de diferenciabilidad a cualquier espacio normado.

Definición 2.5.1 *Diremos que una aplicación $g : E \subset \mathbb{X} \to \mathbb{Y}$, siendo $\mathbb{X}$ e $\mathbb{Y}$ dos espacios normados, es diferenciable en $a \in \mathbb{X}$ si existe una aplicación lineal acotada $L(a) : \mathbb{X} \to \mathbb{Y}$ tal que*

$$g(a+h) - g(a) - L(a)h = o(\|h\|).$$

Nótese que si $\mathbb{X} = \mathbb{R}^n$ y $\mathbb{Y} = \mathbb{R}^m$, recuperamos la definición de diferenciabilidad de la definición 2.2.3.

Supongamos que Df está definido en todo un entorno $A \subset \mathbb{R}^n$. Como $Df \in \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$ y este espacio es normado entonces podemos definir la diferencial de Df. Así tenemos la siguiente:

Definición 2.5.2 *Diremos que f es dos veces diferenciable en un punto $a \in A \in \mathbb{R}^n$ si $Df(x)$ es diferenciable en a y denotaremos a la derivada segunda de f en a por $D^2 f(a)$.*

Nótese que de lo anterior se sigue que $D^2 f(a)$ es una aplicación lineal de $\mathbb{R}^n$ en $\mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$, que además es acotada (la prueba de la acotación la omitiremos), o sea,

$$D^2 f(a) \in \mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)),$$

donde por $\mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))$ denotamos el espacio de las aplicaciones lineales acotadas que van de $\mathbb{R}^n$ en $\mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$.

[3] Ver nota 1.5.5 de la página 34.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Si existe D^2f en todo un abierto $B \subset A$, entonces podemos extender el procedimiento anterior y definir la derivada tercera $D^3f(a)$ que será una aplicación lineal de $\mathbb{R}^n$ en $\mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))$, que además es acotada luego

$$D^3f(a) \in \mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))),$$

y así sucesivamente.

Conviene hacer notar que las derivadas sucesivas de una función $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ se pueden interpretar como aplicaciones multilineales[4] de $\mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ en $\mathbb{R}^m$. De hecho se puede probar que el espacio $\mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))$ es isométrico al espacio $\mathcal{B}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^m)$, siendo este último el espacio de las aplicaciones bilineales acotadas de $\mathbb{R}^n \times \mathbb{R}^n$ en $\mathbb{R}^m$, el espacio $\mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)))$ es isométrico al espacio de las aplicaciones trilineales acotadas $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ en $\mathbb{R}^m$, y así sucesivamente.

Veamos esto con más detalle para el caso especial (y de gran importancia en las aplicaciones) de la segunda derivada. Asumiremos que f tiene derivadas parciales hasta orden dos en A y que estas son continuas en A, y que $a \in A$.

Como ya hemos mencionado $D^2f(a)$ es una aplicación lineal de $\mathbb{R}^n$ en $\mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$, i.e., $D^2f(a) \in \mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))$ y dicho espacio $\mathcal{B}(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m))$ es isométrico al espacio de las aplicaciones bilineales $\mathcal{B}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^m)$.

Así pues, dados $x, y \in \mathbb{R}^n$ tenemos

$$D^2f(a)(x) \in \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m) \text{ y } \big(D^2f(a)(x)\big)(y) \in \mathbb{R}^m.$$

Es decir, $D^2f(a)$ es la aplicación bilineal $D^2f(a) : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}^m$, definida por

$$D^2f(a)(x, y) = \big(D^2f(a)(x)\big)(y).$$

De manera análoga se pueden construir las formas multilineales (k-multilineales) correspondientes a las derivadas (diferenciales) de orden k. El lector interesado puede consultar el capítulo 10 de [9].

La siguiente definición será de gran utilidad en lo que sigue del curso:

Definición 2.5.3 *Sea una función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, A abierto. Diremos que $f \in C^{(k)}(A)$ si f es tal que existen todas sus derivadas parciales hasta orden k y estas son continuas en A. Es decir por $C^{(k)}(A)$ denotaremos al espacio de todas las funciones $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$ que tienen todas sus derivadas parciales hasta orden k continuas en A.*

[4] Una aplicación $B(x_1, x_2, \ldots, x_n)$ es multilineal si es lineal en cada una de sus variables.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

2.5.1. El caso de funciones escalares

En el caso particular $f : A \subset \mathbb{R}^n \to \mathbb{R}$, la derivada segunda en un punto se puede representar mediante una matriz cuadrada $n \times n$, que se denomina ***matriz hessiana***. En efecto, si f es dos veces diferenciable en a, hemos visto que $D^2 f(a)$ puede ser interpretada como una aplicación bilineal $B(x, y)$ de $\mathbb{R}^n \times \mathbb{R}^n$ en $\mathbb{R}$. Ahora bien, las aplicaciones bilineales $B(x, y)$ de $\mathbb{R}^n \times \mathbb{R}^n$ en $\mathbb{R}$ se identifican con las matrices cuadradas $n \times n$ (fijada una base de $\mathbb{R}^n$, que en nuestro caso será la base canónica e_k, $k = 1, \dots, n$) mediante la expresión $B(x, y) = x^T B y$, donde x^T denota al vector transpuesto de x siendo B una matriz $\mathbb{R}^{n\times n}$. Puesto que $D^2 f(a)$ se obtiene derivando la función Df es fácil comprobar (lo veremos más adelante) que la matriz $H_f(a)$ asociada a $D^2 f(a)$ tiene por entradas las derivadas parciales segundas de f, es decir,

$$H_f(a) = \begin{pmatrix} D_{11}f(a) & \cdots & D_{n1}f(a) \\ \vdots & \ddots & \vdots \\ D_{1n}f(a) & \cdots & D_{nn}f(a) \end{pmatrix} = \begin{pmatrix} \dfrac{\partial^2 f(a)}{\partial^2 x_1} & \cdots & \dfrac{\partial^2 f(a)}{\partial x_n \partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(a)}{\partial x_1 \partial x_n} & \cdots & \dfrac{\partial^2 f(a)}{\partial^2 x_n} \end{pmatrix}. \tag{2.5.1}$$

Si $f \in C^{(2)}(A)$ entonces las derivadas cruzadas no dependen del orden de derivación y por tanto la matriz anterior es una matriz simétrica. En general, la forma bilineal correspondiente a la segunda diferencial tiene la forma

$$\begin{aligned} D^2 f(a)(w, z) &= \begin{pmatrix} w_1 & \dots & w_n \end{pmatrix} \begin{pmatrix} \dfrac{\partial^2 f(a)}{\partial^2 x_1} & \cdots & \dfrac{\partial^2 f(a)}{\partial x_n \partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(a)}{\partial x_1 \partial x_n} & \cdots & \dfrac{\partial^2 f(a)}{\partial^2 x_n} \end{pmatrix} \begin{pmatrix} z_1 \\ \\ \vdots \\ \\ z_n \end{pmatrix} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i \partial x_j} w_i z_j. \end{aligned} \tag{2.5.2}$$

De especial interés es el caso cuando la segunda diferencial actúa sobre un único vector $h = (h_1\, h_2\, \dots,\, h_n)^T$, i.e., $D^2 f(a)(h, h)$ lo que, por simplicidad, denotaremos por $D^2 f(a)(h)$. En este caso tenemos

$$D^2 f(a)(h) := D^2 f(a)(h, h) = \big(D^2 f(a)(h)\big)(h) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i \partial x_j} h_i h_j,$$

de donde además se deduce que $D^2 f(a)(h)$ es un polinomio de grado dos en h_1, h_2.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Existe una forma muy útil para calcular los valores de $D^k f(a)(h) := D^k f(a)(h, \dots, h)$ que mostraremos a continuación. Para ello, nótese que

$$Df(a)(h) = \sum_{i=1}^{n} \frac{\partial f(a)}{\partial x_i} h_i = \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \right) f(x) \Big|_{x=a},$$

donde

$$L := h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n},$$

es una aplicación lineal que a cada función $f \in C^{(1)}(A)$ le hace corresponder la función $h_1 \frac{\partial f(x)}{\partial x_1} + \cdots + h_n \frac{\partial f(x)}{\partial x_n}$. Se puede comprobar mediante un cálculo directo que $D^2 f(a)(h) = L(Lf(x))|_{x=a}$, es decir,

$$D^2 f(a)(h) = \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \right)^2 f(x) \Big|_{x=a}.$$

En adelante usaremos la siguiente notación

$$D^k f(a)(h) := D^k f(a)(h, h, \dots, h)$$

para las diferenciales (derivadas) de orden k de las funciones $f \in C^{(k)}(A)$.

Por inducción es posible probar que si $f \in C^{(k)}(A)$ entonces la derivada k-ésima de f en $a \in A$ aplicada a un vector $h \in \mathbb{R}^n$ se expresa por

$$\begin{aligned} D^k f(a)(h) &= \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{\partial^k f(a)}{\partial x_{i_1} \cdots \partial x_{i_k}} h_{i_1} \cdots h_{i_k} \\ &= \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \right)^k f(a), \end{aligned} \tag{2.5.3}$$

donde se define la potencia $k \in \mathbb{N}$ del operador L de la siguiente forma recurrente $L^k f(x) = L(L^{k-1} f(x))$, i.e.,

$$\begin{aligned} &\left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \right)^k f(x) \\ &\quad = \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \right) \left[\left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \right)^{k-1} f(x) \right]. \end{aligned}$$

Nótese que $D^k f(a)(h)$ es, en general, un polinomio de grado k en las variables $h_1, \dots h_n$.

La utilidad de lo anterior quedará más claro al final de este apartado.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

2.5.2. El caso de funciones escalares de dos variables

Un cálculo directo permite comprobar que la expresión (2.5.3) en el caso de funciones de dos variables $f:\mathbb{R}^2\mapsto\mathbb{R}$, $f(x,y)\in C^{(k)}(A)$ se reduce a la siguiente expresión:

$$\begin{aligned} D^k f(x_0,y_0)(h) &= \sum_{\substack{i,j\geq 0\\ i+j=k}} \frac{\partial^k f(x_0,y_0)}{\partial^i x\,\partial^j y}(x-x_0)^i(y-y_0)^j,\\ &= \sum_{i=0}^{k}\binom{k}{i}\frac{\partial^k f(x_0,y_0)}{\partial^i x\,\partial^{k-i} y}(x-x_0)^i(y-y_0)^{k-i}. \end{aligned}$$

donde $h=(x-x_0,y-y_0)$.

Si tomamos $k=1$ y 2 en la formula anterior recuperamos las expresiones para los diferenciales de orden 1 y 2, cuyas expresiones son, respectivamente

$$\begin{aligned} Df(x_0,y_0)(h) =& \frac{\partial f(x_0,y_0)}{\partial x}(x-x_0)+\frac{\partial f(x_0,y_0)}{\partial y}(y-y_0),\\ D^2f(x_0,y_0)(h) =& \frac{\partial^2 f(x_0,y_0)}{\partial x^2}(x-x_0)^2+2\frac{\partial^2 f(x_0,y_0)}{\partial x\partial y}(x-x_0)(y-y_0)\\ &+\frac{\partial^2 f(x_0,y_0)}{\partial y^2}(y-y_0)^2, \end{aligned}$$

donde, como antes, $h=(x-x_0,y-y_0)$.

Antes de continuar vamos a introducir una notación muy útil a la hora de trabajar las derivadas parciales de orden superior. Así, dada $f:\mathbb{R}^2\mapsto\mathbb{R}$ denotaremos por f_x a la función $f_x(x,y)=D_1f(x,y)=\frac{\partial f}{\partial x}(x,y)$, por f_y a $f_y(x,y)=D_2f(x,y)=\frac{\partial f}{\partial y}(x,y)$ y, en general,

$$f_{xx}(x,y)=D_{11}f(x,y)=\frac{\partial^2 f}{\partial x^2}(x,y),\ \ f_{yy}(x,y)=D_{22}f(x,y)=\frac{\partial^2 f}{\partial y^2}(x,y),$$

$$f_{xy}(x,y)=D_{12}f(x,y)=\frac{\partial^2 f}{\partial x\partial y}(x,y),\ \ f_{yx}(x,y)=D_{21}f(x,y)=\frac{\partial^2 f}{\partial y\partial x}(x,y),$$

$$f_{xxy}(x,y)=D_{112}f(x,y)=\frac{\partial^3 f}{\partial x\partial x\partial y}(x,y)=\frac{\partial^3 f}{\partial x^2\partial y}(x,y),\quad \text{etc.}$$

2.6. El Teorema de Taylor para funciones de varias variables

Comenzaremos con un ejemplo.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Ejemplo 2.6.1 Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, $f \in C^k(A)$. Sea la función $l : \mathbb{R} \mapsto \mathbb{R}^n$, $l(t) = x + th$, $t \in \mathbb{R}$, $h, x \in \mathbb{R}^n$. Definamos la función compuesta $\phi : \mathbb{R} \mapsto \mathbb{R}$, $\phi(t) = (f \circ l)(t) = f(l(t)) = f(x + th)$. Probar que $\phi \in C^k([0, 1])$.

Por hipótesis todas las derivadas de f hasta orden k son continuas, luego diferenciables, y está claro que l también es k-veces diferenciable. Usando lo anterior, y la regla de la cadena (ver Teorema 2.3.1) probemos que $\phi \in C^k([0, 1])$. Para ello calculemos las sucesivas derivadas de ϕ.

Así, para la primera derivada tenemos

$$\phi'(t) = Df(x + ht) \cdot h = \sum_{i=1}^{n} \frac{\partial f(x + ht)}{\partial x_i} h_i,$$

que es una función continua (composición y suma de funciones continuas). Para la segunda derivada obtenemos

$$\begin{aligned}\phi''(t) = \frac{d}{dt}\left[\sum_{i=1}^{n} \frac{\partial f(x + ht)}{\partial x_i} h_i\right] &= \sum_{i=1}^{n}\left[\sum_{j=1}^{n} \frac{\partial}{\partial x_j}\left(\frac{\partial f(x + ht)}{\partial x_i}\right) h_j\right] h_i \\ &= \sum_{i,j=1}^{n} \frac{\partial^2 f(x + ht)}{\partial x_j \partial x_i} h_j h_i,\end{aligned}$$

que por las mismas razones de antes (composición y suma de funciones continuas) es continua.

Por inducción, se tiene que

$$\begin{aligned}\phi^{(k)}(t) &= \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{\partial^k f(x + ht)}{\partial x_{i_1} \cdots \partial x_{i_k}} h_{i_1} \cdots h_{i_k} \\ &= \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n}\right)^k f(x + ht),\end{aligned} \tag{2.6.1}$$

de donde se sigue que $\phi \in C^k([0, 1])$. □

En general, el cálculo de la derivada k-ésima para una función arbitraria es muy engorroso, por lo que usar un paquete de cálculo simbólico (por ejemplo MAXIMA CAS) puede se de gran ayuda.

Definición 2.6.2 *Dado un $a \in A$ y $h \in \mathbb{R}^n$ definiremos al intervalo (cerrado) $[a, a+h]$ como el conjunto (intervalo) $[a_1, a_1+h_1] \times [a_2, a_2+h_2] \times \cdots \times [a_n, a_n+h_n]$.*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Teorema 2.6.3 (de Taylor con resto de Lagrange) *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, $f \in C^k(A)$. Sea $a \in A$ y asumamos que el intervalo $[a, a+h] \subset A$ para cierto $h \neq 0$. Entonces*

$$f(a+h) = P_{k-1}(x,a) + r_k(a,h),$$

donde

$$P_{k-1}(x,a) := f(a) + \sum_{l=1}^{k-1} \frac{1}{l!} D^l f(a)(h) \text{ y } r_k(a,h) = \frac{1}{k!} D^k f(a+\xi h)(h), \quad \xi \in (0,1),$$

denotan al polinomio de Taylor de orden $k-1$ y al resto, respectivamente.

Nótese que el polinomio de Taylor de orden $k-1$ es un polinomio de grado a lo más $k-1$ en $h_1, \ldots, h_n$.

Demostración: Definamos la función $\phi(t) = f(a+ht)$ estudiada en el ejemplo anterior. Como hemos visto es continuamente diferenciable hasta orden k en $[0,1]$. Entonces podemos usar el Teorema de Taylor para funciones de una variable 1.1.22. Sea $\tau \in (0,1]$ cualquiera, entonces existe un $\xi \in (0,\tau)$ tal que

$$\phi(\tau) = \phi(0) + \phi'(0)\tau + \cdots + \frac{1}{(k-1)!}\phi^{(k-1)}(0)\tau^{k-1} + r_k(\tau),$$

donde tomamos el resto r_k, por ejemplo, en forma de Lagrange

$$r_k(\tau) = \frac{1}{k!}\phi^{(k)}(\xi)\tau^k.$$

Escogiendo $\tau = 1$ y usando (2.5.3) y (2.6.1) con $x = a$ se deduce el resultado. □

Corolario 2.6.4 (Teorema local de Taylor) *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, $f \in C^k(A)$ y $[a, a+h] \subset A$ para cierto $h \neq 0$, entonces*

$$f(a+h) = f(a) + \sum_{l=1}^{k} \frac{1}{l!} D^l f(a)(h) + o(\|h\|^k). \tag{2.6.2}$$

Demostración: Calculamos el límite

$$\lim_{h\to 0} \frac{D^k f(a+\xi h)(h) - D^k f(a)(h)}{\|h\|^k} =$$

$$\lim_{h\to 0} \sum_{i_1,\cdots,i_k=1}^{n} \left[\frac{\partial^k f(a+\xi h)}{\partial x_{i_1}\cdots\partial x_{i_k}} - \frac{\partial^k f(a)}{\partial x_{i_1}\cdots\partial x_{i_k}}\right] \frac{h_{i_1}}{\|h\|}\cdots\frac{h_{i_k}}{\|h\|} = 0,$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

pues las k-ésimas derivadas parciales son continuas y las cantidades $h_j/\|h\|$, $j = 1, \ldots, k$ están acotadas. De lo anterior se sigue que $D^k f(a+\xi h)(h) = D^k f(a)(h) + o(\|h\|^k)$, de donde se deduce el resultado. □

Nótese que del corolario anterior se sigue la unicidad del polinomio de Taylor.

Tanto el Teorema de Taylor 2.6.3 como su corolario se pueden extender a funciones vectoriales $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^m$, si las consideramos componente a componente.

Nótese además que en el caso especial cuando $k = 2$ la expresión (2.6.2) se transforma en

$$f(x) = f(a) + Df(a)(x-a) + \frac{1}{2}D^2 f(a)(x-a) + o(\|x-a\|^2). \tag{2.6.3}$$

Lo anterior nos permite *redefinir* el concepto de diferenciabilidad para funciones $f : \mathbb{R}^n \mapsto \mathbb{R}$ de una forma más sencilla no solo de entender si no también de usar. Por simplicidad lo discutiremos para el caso de una función dos veces diferenciable siendo fácilmente generalizable al caso general.

Así pues, asumamos que f tiene derivadas parciales de orden dos y estas son continuas entonces, (2.6.3) nos conduce a

$$f(a+h) - f(a) - Df(a)(h) - \frac{1}{2}D^2 f(a)(h) = o(\|h\|^2), \tag{2.6.4}$$

donde $D^2 f(a)(h)$ es la forma bilineal (ver (2.5.3))

$$D^2 f(a)(h) = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \frac{\partial^2 f(a)}{\partial x_{i_1}\partial x_{i_2}} h_{i_1} h_{i_2} = h^T H_f(a) h. \tag{2.6.5}$$

Por tanto podemos decir que una función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, es dos veces diferenciable si existen una aplicación lineal $Df(a)$ y una bilineal $D^2 f(a)$ tales que (2.6.4) sea cierta. Este razonamiento es fácilmente generalizable para cualquier $k \geq 3$.

Lo anterior nos permite restringirnos, por simplicidad, al caso cuando las funciones $f \in C^k(A)$. Así pues diremos que $f : A \in \mathbb{R}^n \mapsto \mathbb{R}$, es k veces diferenciable en a si f es $C^k(A)$ siendo A un abierto tal que $a \in A$, de forma que, por el Teorema de Taylor tenemos asegurado que f es k veces diferenciable en A en el sentido antes explicado.

Como veremos más adelante, el Teorema de Taylor juega un papel fundamental en el cálculo de los extremos de las funciones de varias variables.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Ejemplo 2.6.5 Encuentra el polinomio de Taylor de orden 2 alrededor del punto $(0,0)$ de la función $f(x,y) = e^{ax+ay}$.

Ante todo notemos que f es una función $C^\infty(\mathbb{R}^2)$ así que podemos definir las diferenciales de cualquier orden de f. Está claro que $f(0,0) = 1$, $f_x(0,0) = f_y(0,0) = a$, $f_{xx}(0,0) = f_{xy}(0,0) = f_{yx}(0,0) = f_{yy}(0,0) = a^2$. Luego,

$$Df(0,0) = (a,a), \qquad H_f(0,0) = \begin{pmatrix} a^2 & a^2 \\ a^2 & a^2 \end{pmatrix},$$

de donde, usando (2.6.3), tenemos

$$e^{ax+ay} = 1 + a(x+y) + \frac{1}{2}a^2(x+y)^2 + o(x^2+y^2).$$

□

Para esta función se puede encontrar fácilmente el polinomio de cualquier orden. La idea es usar la unicidad del polinomio de Taylor. En nuestro ejemplo consideremos la función de una variable $g(z) = e^z$. Para ella se tiene

$$e^z = 1 + z + \frac{z^2}{2!} + \cdots + \frac{z^n}{n!} + o(z^n) = P_n(z) + o(z^n).$$

Entonces, sustituyendo $z = a(x+y)$ obtenemos el polinomio de orden n para nuestra función de dos variables e^{ax+ay}

$$P_n(x,y) = 1 + a(x+y) + \frac{a^2(x+y)^2}{2!} + \cdots + \frac{a^n(x+y)^n}{n!},$$

de donde se sigue que

$$e^{ax+ay} = 1 + a(x+y) + \frac{a^2(x+y)^2}{2!} + \cdots + \frac{a^n(x+y)^n}{n!} + o(\|h\|^n), \tag{2.6.6}$$

con $\|h\| = \sqrt{x^2+y^2}$.

Como ejercicio puede el lector encontrar los diferenciales de orden 3 y 4 y obtener los correspondientes términos en el desarrollo (2.6.3) y compararlos con los obtenidos en (2.6.6).

2.6.1. Problemas propuestos

Problema 2.6.1 Sea la función

$$f(x,y) = \begin{cases} \arctan\left(\dfrac{x^4+y^4}{x^2+y^2}\right), & \text{si } (x,y) \neq (0,0); \\ 0, & \text{si } (x,y) = (0,0). \end{cases}$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Decide si es continua y diferenciable en $\mathbb{R}^2$. Decide si $f \in C^1(\mathbb{R}^2)$. Calcula $D_{12}f(0,0)$ y $D_{21}f(0,0)$. Decide si $f \in C^2(\mathbb{R}^2)$.

Problema 2.6.2 Sea la función definida por

$$f(x,y) = \begin{cases} xy \operatorname{sen} \dfrac{1}{2}\left(\dfrac{x-y}{x+y}\right), & \text{si } x+y \neq 0; \\ 0, & \text{si } x+y = 0. \end{cases}$$

Estudia la continuidad y diferenciabilidad de esta función. Calcula $D_{12}f(0,0)$ y $D_{21}f(0,0)$.

Problema 2.6.3 Sea la función

$$f(x,y) = \begin{cases} \dfrac{\operatorname{sen}(xy) - xy}{x^2 y}, & \text{si } x \neq 0, y \neq 0, \\ 0, & \text{en otro caso.} \end{cases}$$

Estudia la continuidad y diferenciabilidad de f en $\mathbb{R}^2$. Calcular $D_{12}f(0,0)$ y $D_{21}f(0,0)$. ¿Es la función dos veces diferenciable en $\mathbb{R}^2$? Si es dos veces diferenciable en un entorno de $(0,0)$ calcula su polinomio de Taylor de orden 2 en $(0,0)$.

Problema 2.6.4 Sea $f : \mathbb{R}^2 \to \mathbb{R}$, definida por

$$f(x,y) = \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2}, & (x,y) \neq (0,0), \\ 0, & (x,y) = (0,0). \end{cases}$$

Probar que $f \in C^1(\mathbb{R}^2)$. Calcular $D_{12}f(0,0)$ y $D_{21}f(0,0)$. ¿La función f es dos veces diferenciable en $\mathbb{R}^2$?

Problema 2.6.5 Sea la función $g : \mathbb{R}^2 \mapsto \mathbb{R}$, $g(x,y,z) = x^2 + ye^{x\,z}$.

1. Decide si g es diferenciable en $\mathbb{R}^2$. Justifica la respuesta.
2. Calcula las derivadas parciales $\frac{\partial g}{\partial x}$, $\frac{\partial g}{\partial y}$ y $\frac{\partial g}{\partial z}$.
3. ¿Cuánto vale la derivada g en el punto $(1,2,0)$ según la dirección del vector $(3/4, \sqrt{3}/2, 1/2)$.
4. ¿En que dirección es máxima la variación de g en dicho punto $(1,2,0)$? ¿Y mínima?

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

5. Escribe la expresión para el plano tangente a la superficie definida por $w = g(x, y, z)$ en el punto $(1, 2, 0)$.

6. Escribe el polinomio de Taylor de orden 2 en el punto $(1, 2, 0)$.

Problema 2.6.6 Sea $g : \mathbb{R}^2 \mapsto \mathbb{R}$, $g(x, y) = x^2 \cos(y) + y^2 \cos(x)$.

1. Decide si g es diferenciable en $\mathbb{R}^2$. Justifica la respuesta.

2. Calcula las derivadas parciales $\frac{\partial g}{\partial x}$ y $\frac{\partial g}{\partial y}$.

3. ¿Cuánto vale la derivada de g en el punto $(0, \pi/4)$ según la dirección del vector $(1, -1)$.

4. ¿En que dirección es máxima la variación de g en dicho punto $(0, \pi/4)$? ¿Y mínima?

5. Escribe la expresión para el plano tangente a la superficie definida por $z = g(x, y)$ en el punto $(0, \pi/4)$?

6. Escribe el polinomio de Taylor de orden 2 en el punto $(0, \pi/4)$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Capítulo 3

El Teorema de la función implícita

Este capítulo lo dedicaremos a uno de los problemas más importante del análisis: el Teorema de la función implícita. Comenzaremos con un ejemplo my sencillo pero a la vez revelador.

3.1. Planteamiento del problema

Sea una función escalar de dos variables $F : \mathbb{R}^2 \mapsto \mathbb{R}$. Comenzaremos estudiando el problema de cuándo la ecuación $F(x, y) = 0$ permite definir una función $y = f(x)$ tal que $F(x, y) = 0$ si y sólo si $y = f(x)$. ¿Qué propiedades tiene f? ¿Es continua, diferenciable, etc?

Por ejemplo, $F(x, y) = x^2 + y^2 - 1 = 0$ define una circunferencia en $\mathbb{R}^2$. Ahora bien, si queremos despejar la y tenemos $y = \pm\sqrt{1 - x^2}$. ¿Cuál de las dos ramas tomamos? Supongamos que elegimos $y = f(x) = \sqrt{1 - x^2}$. Esta función es continua en $[-1, 1]$ pero no es diferenciable en los extremos ± 1. Formalmente podríamos haber elegido también la función $f(x) = \sqrt{1 - x^2}$ si $x \in \mathbb{Q}$ y $f(x) = -\sqrt{1 - x^2}$ si $x \in \mathbb{I}$ que no es continua en ningún punto de $[-1, 1]$ y que sin embargo satisface la ecuación $F(x, f(x)) = 0$.

Una opción para resolver el problema consiste en aproximar $F(x, y)$ por el plano (recta) tangente a un determinado punto (x_0, y_0) que cumple con que $F(x_0, y_0) = 0$. Si F es diferenciable en (x_0, y_0) entonces en un entorno de (x_0, y_0)

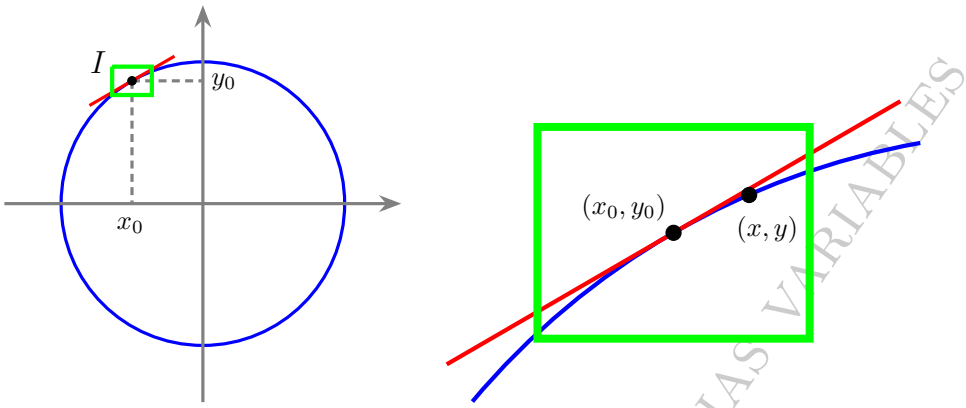


Figura 3.1.1: Entorno I (en verde) de (x_0, y_0) (ampliado en la figura de la derecha) donde podemos construir la función implícita $f(x)$ tal que $F(x, f(x)) = 0$ para $F(x, y) = x^2 + y^2 - 1$ (en azul). En rojo se representa el plano (recta) tangente a $F(x, y)$ en (x_0, y_0).

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

tenemos

$$F(x,y) = F(x_0,y_0) + \frac{\partial F(x_0,y_0)}{\partial x}(x - x_0) + \frac{\partial F(x_0,y_0)}{\partial y}(y - y_0) + o(\|h\|),$$

donde $\|h\| = \sqrt{(x-x_0)^2 + (y-y_0)^2}$. Como $F(x_0, y_0) = 0$, y queremos que $F(x, y) = 0$, entonces

$$\frac{\partial F(x_0,y_0)}{\partial x}(x - x_0) + \frac{\partial F(x_0,y_0)}{\partial y}(y - y_0) \approx 0.$$

De lo anterior deducimos un valor aproximado para y en función de la x

$$y - y_0 \approx -\left[\frac{\partial F(x_0,y_0)}{\partial y}\right]^{-1} \frac{\partial F(x_0,y_0)}{\partial x}(x - x_0).$$

Como $y = f(x)$ y $y_0 = f(x_0)$ tenemos además que

$$\frac{\Delta f(x)}{\Delta x} \approx -\left[\frac{\partial F(x_0,y_0)}{\partial y}\right]^{-1} \frac{\partial F(x_0,y_0)}{\partial x}$$

que al tomar límites cuando $\Delta x \to 0$ nos genera una expresión para calcular la derivada $f'(x_0)$. Nótese que de lo anterior se deduce además que para que podamos despejar la y necesitamos que $\frac{\partial F(x_0,y_0)}{\partial y} \neq 0$.

Si aplicamos lo anterior al ejemplo $F(x, y) = x^2 + y^2 - 1 = 0$ tenemos que, en general, podemos definir una función f en cualquier entorno de (x_0, y_0), $x_0 \in (-1, 1)$ que escojamos siempre que $\frac{\partial F(x_0,y_0)}{\partial y} = 2y_0 \neq 0$, o sea, siempre que $y_0 \neq 0$ (véase la figura 3.1.1).

3.2. Caso de una única ecuación

Pasemos a enunciar el teorema que resuelve el problema de una función implícita definida por una única ecuación.

Teorema 3.2.1 (de la función implícita) *Sea $F : A \subset \mathbb{R}^n \times \mathbb{R} \mapsto \mathbb{R}$ definida en un entorno del punto*[1] *$(x_0, y_0) \in A$, A abierto de $\mathbb{R}^n \times \mathbb{R}$. Supongamos que:*

1. $F(x, y) := F(x_1, x_2, \ldots, x_n, y) \in C^{(p)}(A)$, $p \geq 1$,
2. $F(x_0, y_0) := F(x_{01}, x_{02}, \ldots, x_{0n}, y_0) = 0$,
3. $F'_y(x_0, y_0) = \dfrac{\partial F(x_{01}, x_{02}, \ldots, x_{0n}, y_0)}{\partial y} \neq 0$.

Entonces existe un abierto $I = I_x \times I_y = (x_0 - h, x_0 + h) \times (y_0 - k, y_0 + k)$[2] *alrededor del punto (x_0, y_0), $I \subset A$, y una función $f : I_x \subset \mathbb{R}^n \mapsto I_y \subset \mathbb{R}$ tal que:*

1. *$F(x, y) = 0$ en I si y sólo si $f(x) = y$,*
2. *$f(x) \in C^{(p)}(I_x)$.*
3. *Para todo $x \in I_x$, las derivadas parciales $\frac{\partial f(x)}{\partial x_i}$, $i = 1, 2, \ldots, n$, de $f(x)$ se calculan por la fórmula*

$$\frac{\partial f(x)}{\partial x_i} := \frac{\partial f(x_1, \ldots x_n)}{\partial x_i} = -[F'_y(x, f(x))]^{-1} \cdot [F'_{x_i}(x, f(x))], \tag{3.2.1}$$

donde por F'_{x_i} denotamos la derivada parcial $\frac{\partial F}{\partial x_i}$.

Demostración: Vamos a probar el teorema en cuatro pasos:

1. Asumamos, sin pérdida de generalidad, que $F'_y(x_0, y_0) > 0$, entonces, como $F'_y(x, y)$ es continua en $(x_0, y_0) \in A$, se tiene que existe todo un entorno $I = I_x \times I_y$ de (x_0, y_0) donde $F'_y(x, y) > 0$ (ver figura).

[1] En este apartado usaremos la siguiente notación $(x, y) \in \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1} \Rightarrow (x, y) = (x_1, x_2, \ldots, x_n, y)$.

[2] Análogamente al caso de los intervalos definidos justo antes del Teorema 2.6.3, definiremos el abierto $(x_0 - h, x_0 + h)$ como $(x_{01} - h_2, x_{01} + h_1) \times (x_{02} + -h_2, x_{02} + h_2) \times \cdots \times (x_{0n} - h_n, x_{0n} + h_n)$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

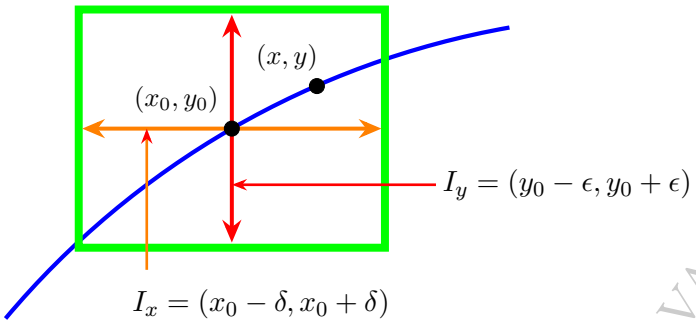


Figura 3.2.1: El entorno $I = I_x \times I_y$ de (x_0, y_0).

Luego, $F(x, y)$ como función de y es estrictamente creciente. Como $F(x_0, y_0) = 0$ entonces tenemos que

$$0 > F(x_0, y_0 - k) < F(x_0, y_0) < F(x_0, y_0 + k) > 0.$$

Como $F(x, y)$ es continua en A entonces los signos de F se mantienen en todo un entorno de cada punto, i.e., existe un entorno I_x (que por simplicidad asumimos igual al de antes) tal que

$$0 > F(x, y_0 - k) < F(x_0, y_0) < F(x, y_0 + k) > 0, \qquad \forall x \in I_x.$$

Pero para cada $x \in I_x$ la función $h(y) := F(x, y)$ cambia de signo en los extremos del intervalo I_y, luego por el Teorema de Bolzano 1.1.13, para cada $x \in I_x$ existe un único $y \in I_y$ tal que $F(x, y) = 0$ (la unicidad es consecuencia de la monotonía de $h(y)$). Definiendo la función $f : I_x \mapsto I_y \subset \mathbb{R}$ de forma que $y = f(x)$ obtenemos una función que cumple con que $F(x, f(x)) = 0$.

2. Que $f(x)$ así definida es continua es inmediato de la construcción anterior pues de esta se sigue que (ver figura 1) para todo $\epsilon > 0$ (suficientemente pequeño de forma que $y_0 \pm \epsilon \in I_y$) siempre podemos encontrar un $\delta > 0$ tal que $\|x - x_0\| < \delta$ y $x \in I_x$ de forma que $F(x, y)$ cambie de signo en los extremos de $[y_0 - \epsilon, y_0 + \epsilon]$, luego $f(x) \in (y_0 - \epsilon, y_0 + \epsilon)$, i..e., para todo $\epsilon > 0$, existe $\delta > 0$, tal que $|f(x) - f(x_0)| < \epsilon$, que es justo lo que se quería probar.

3. Probemos que $f \in C^{(1)}(I_x)$. Hemos probado que para todo $x \in I_x$ existe una función continua $y = f(x)$ tal que $F(x, f(x)) = 0$ en I_x. Sean los puntos x y $x + \Delta x = x + he_i$, e_i i-ésimo vector de la base canónica de $\mathbb{R}^n$ con h tal que $x + he_i \in I_x$, y sea $\Delta y = f(x + \Delta x) - f(x)$, i.e., $f(x + \Delta x) = y + \Delta y$ (ver 3).

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

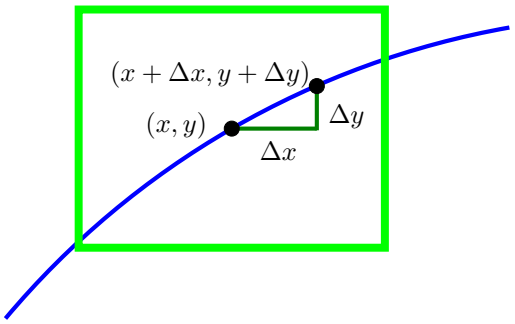


Figura 3.2.2: Construcción para la prueba de que $f \in C^{(1)}(I_x)$.

Nótese que como f es continua en I_x, entonces $\Delta y = f(x+\Delta x) - f(x) \to 0$ si $\Delta x \to 0$.

Como $F(x,y) = F(x+\Delta x, y+\Delta y) = F(x+\Delta x, f(x+\Delta x)) = 0$ entonces, aplicando el Teorema del valor medio 2.3.6 para funciones escalares sabemos que existe un $\theta \in (0,1)$ tal que

$$\begin{aligned} 0 =& F(x+\Delta x, y+\Delta y) - F(x,y) = DF(x+\theta h e_i, y+\theta\Delta y)\begin{pmatrix}\Delta x\\ \Delta y\end{pmatrix} \\ =& \frac{\partial F(x+\theta h e_i, y+\theta\Delta y)}{\partial x_i} h + \frac{\partial F(x+\theta h e_i, y+\theta\Delta y)}{\partial y}\Delta y. \end{aligned}$$

Como $F(x,y) \in C^{(1)}(I)$ entonces tomando límites $\Delta x \to 0$ obtenemos

$$0 = \frac{\partial F(x,y)}{\partial x_i} + \frac{\partial F(x,y)}{\partial y} \lim_{h\to 0} \frac{\Delta y}{h}$$

de donde, teniendo en cuenta que $\lim_{h\to 0} \Delta y/h = \frac{\partial f(x)}{\partial x_i}$, y que, como hemos visto, en todo un entorno de (x_0, y_0), $\frac{\partial F(x,y)}{\partial y} \neq 0$, se sigue el resultado (3.2.1).

4. Para probar que $f(x) \in C^{(p)}(I_x)$ usaremos inducción. Para $p = 1$ ya lo hemos probado. Supongamos que $F \in C^{(2)}(A)$. Entonces, si derivamos

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

(3.2.1) respecto a x_j tenemos

$$\frac{\partial^2 f(x_1,\dots x_n)}{\partial x_j \partial x_i} = \frac{\overbrace{[F''_{x_i x_j}(x,f(x)) + F''_{x_i y}(x,f(x))f_{x_j}(x)]}^{\frac{\partial F_{x_i}(x,f(x))}{\partial x_j}} F'_y(x,f(x))}{[F'_y(x,f(x))]^2}$$

$$- \frac{F_{x_i}(x,f(x))\overbrace{[F''_{yx_j}(x,f(x)) + F''_{yy}(x,f(x))f_{x_j}(x)]}^{\frac{\partial F_y(x,f(x))}{\partial x_j}}}{[F'_y(x,f(x))]^2}. \tag{3.2.2}$$

Como vemos, en parte derecha de la expresión (3.2.2) aparecen las segundas derivadas parciales de F que son, por hipótesis, continuas. Como f es continua entonces la composición de las segundas derivadas parciales de F y f también es continua. También aparecen las derivadas primeras de f, que como hemos probado, son continuas, luego las segundas derivadas parciales de F son continuas (¿por qué?) y por tanto $f \in C^{(2)}(I_x)$. Y así, sucesivamente.

De hecho, es relativamente sencillo comprobar que la expresión de las k-ésimas derivadas parciales de f solo contienen las ℓ-ésimas derivadas parciales de F, $\ell = 1,\dots,k$ (que son continuas por hipótesis) compuestas con $f(x)$ (que es continua) y las m-ésimas derivadas parciales de f, $m = 1,\dots,k-1$, que lo son por la hipótesis de inducción, luego $f \in C^{(2)}(I_x)$. □

Ejemplo 3.2.2 Sea la ecuación $z^3 + 2(x+y)^2 z + e^{z-1} - 4 = 0$.

1. Prueba que la ecuación anterior define una función $z = f(x,y)$ en el entorno U del punto $(0,-1,1)$ y que dicha función es una función $C^{(\infty)}(U)$ en dicho entorno U.
2. Calcula las derivadas parciales $z_x = \frac{\partial f}{\partial x}$ y $z_y = \frac{\partial f}{\partial y}$ en dicho punto.
3. Escribe el polinomio de Taylor de orden 2 de f en $(0,-1,1)$.

Sea la función $F : \mathbb{R}^3 \mapsto \mathbb{R}$, $F(x,y,z) = z^3 + 2(x+y)^2 z + e^{z-1} - 4$. En primer lugar, está claro que en el punto $(0,-1,1)$ se verifica la ecuación $F(0,-1,1) = 0$. Además la función F es $C^{(p)}(\mathbb{R}^3)$ para todo $p \in \mathbb{N}$ y $F'_z(0,-1,1) = 6 \neq 0$, es decir, se cumplen todas las condiciones del Teorema 3.2.1 por lo que tenemos que existe en todo un entorno de $(0,-1,1)$ una función $z = f(x,y)$, $f \in C^{(p)}(\mathbb{R}^2)$ para todo $p \in \mathbb{N}$ tal que $F(x,y,f(x,y)) = 0$ en dicho entorno de $(0,-1,1)$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Para calcular las derivadas usamos la fórmula (3.2.1). Teniendo en cuenta que $F'_x(0,-1,1) = F'_y(0,-1,1) = 4(x+y)z = -4$ tenemos

$$z_x(0,-1) = -\frac{F'_x(0,-1,1)}{F'_z(0,-1,1)} = \frac{2}{3}, \quad z_x(0,-1) = -\frac{F'_y(0,-1,1)}{F'_z(0,-1,1)} = \frac{2}{3}.$$

Como f es $C^{(\infty)}(\mathbb{R}^2)$ en un entorno de $(0,-1)$ entonces es diferenciable tantas veces como se quiera. Así que podemos encontrar su polinomio de Taylor. Derivando dos veces respecto a x la ecuación $F(x,y) = 0$ y considerando z como función de x, y tenemos:

$$2z_{xx}\,(y+x)^2 + 8z_x\,(y+x) + 3z^2 z_{xx} + e^{z-1}z_{xx} + (6z + e^{z-1})z_x^2 + 4z = 0 \quad \Longrightarrow$$

$z_{xx} = -\frac{8}{27}$. Derivando respecto a y dos veces obtenemos

$$2z_{yy}\,(y+x)^2 + 8z_y\,(y+x) + 3z^2 z_{yy} + e^{z-1}z_{yy} + (6z + e^{z-1})z_y^2 + 4z = 0 \quad \Longrightarrow$$

$z_{yy} = -\frac{8}{27}$. Finalmente, derivando respecto a x y luego a y tenemos

$$2z_{xy}(y+x)^2 + 4(z_y + z_x)\,(y+x) + 6zz_x z_y + e^{z-1}z_x z_y + 3z^2 z_{xy} + e^{z-1}z_{xy} + 4z = 0$$

de donde se obtiene $z_{xy} = -\frac{8}{27}$.

Entonces, el Teorema de Taylor 2.6.3 y las expresiones (2.6.4) y (2.6.5), nos conducen a la expresión

$$\begin{aligned} z(x,y) =& z(0,-1) + Dz(0,-1)(x, y+1) + D^2 z(0,-1)(x, y+1) \\ =& 1 + \begin{pmatrix} \frac{2}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} x \\ y+1 \end{pmatrix} + \begin{pmatrix} x & y+1 \end{pmatrix} \begin{pmatrix} -\frac{8}{27} & -\frac{8}{27} \\ -\frac{8}{27} & -\frac{8}{27} \end{pmatrix} \begin{pmatrix} x \\ y+1 \end{pmatrix} + o\left(\|h\|\right), \end{aligned}$$

donde $\|h\| = \sqrt{x^2 + (y-1)^2}$. □

Ejemplo 3.2.3 Sean $x, y > 0$ y sea la ecuación $\log y - \log x + xy = 0$, $x, y > 0$. ¿Define dicha ecuación una función $y = f(x)$?

Definamos la función $F(x,y) = \log y - \log x + xy$ en $D = \{(x,y),\, x, y > 0\}$. Está claro que F es continua y diferenciable en todo su dominio (de hecho es $C^{(k)}(D)$ para todo $k \in \mathbb{N}$) y que $F_y(x,y) = x + 1/y > 0$, para todos $x, y > 0$. Luego se cumplen las dos últimas condiciones del Teorema de la función implícita 3.2.1. Nos falta comprobar la primera, es decir, que para cualquiera sea $x_0 > 0$, existe un $y_0 > 0$ tal que $F(x_0, y_0) = 0$. Sea la función $h(y) := F(x_0, y)$, que es continua para todo $y > 0$. Si y es lo suficientemente pequeña, entonces $h(y) < 0$ ($\lim_{y\to 0+} h(y) = -\infty$), mientras que si y es muy grande, $h(y) > 0$ ($\lim_{y\to\infty} h(y) =$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

$+\infty$). Entonces si aplicamos el Teorema de Bolzano 1.1.13 y teniendo en cuenta que $h'(y) = x + 1/y > 0$, sabemos que existe un único y_0 tal que $h(y_0) = 0$.

Luego, el Teorema de la función implícita nos asegura que existe en el entorno $U(x_0)$ de cada $x_0 > 0$ una función $y = f(x)$, $f \in C^{(k)}(U(x_0))$ que satisface la ecuación $\log y - \log x + xy = 0$. Dicha función está representada en la figura 3.2.

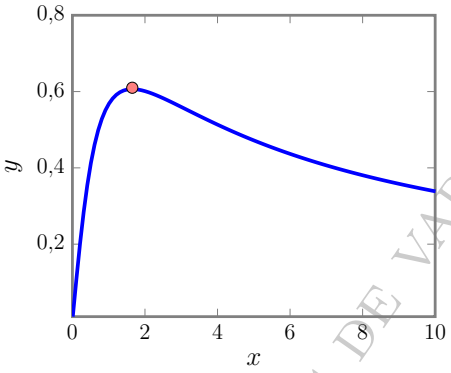


Figura 3.2.3: La función $f(x)$ (en azul) definida por la ecuación $\log y - \log x + xy = 0$, $x, y > 0$. El punto rojo corresponde al valor máximo $f(\sqrt{e}) = 1/\sqrt{e}$.

Probemos que dicha función, de la que no conocemos su forma explícita, tiene un máximo local. Como f es diferenciable (de hecho es k-veces diferenciable, $k \in \mathbb{N}$) su derivada ha de anularse:

$$f'(x) = -\frac{F'_x(x,y)}{F'_y(x,y)} = -\frac{y(xy-1)}{x(xy+1)} = 0 \quad \Longleftrightarrow \quad xy = 1,$$

pues $y = 0$ no pertenece al dominio de f. Para calcular los valores de x donde se anula usamos la ecuación $F(x_0, 1/x_0) = 1 - 2\log x = 0$, de donde se sigue que $x_0 = \sqrt{e}$ e $y_0 = f(x_0) = 1/\sqrt{e}$.

Veamos ahora que es máximo local. Para ello calculamos la segunda derivada de f en x_0, lo cual haremos como sigue. Primero derivamos la ecuación $F(x, y) = 0$ respecto a x asumiendo que $y = y(x)$

$$y + xy' + \frac{y'}{y} - \frac{1}{x} \quad \Longleftrightarrow \quad y^2 + xyy' + y' - \frac{y}{x} = 0.$$

Derivando respecto a x la ecuación anterior y teniendo en cuenta que $y' = 0$, obtenemos

$$0 = 3yy' + x(y')^2 + xyy'' + y'' - \frac{xy' - y}{x^2} = (1 + xy)y'' + \frac{y}{x^2}$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

de donde se sigue que

$$y'' = -\frac{y}{x^2(1+xy)} < 0, \quad y''(\sqrt{e}) = -\frac{1}{2e\sqrt{e}} < 0.$$

Dado que existe un solo punto x_0 donde se anula la derivada de f, este es el único extremo de f, que además es un máximo local. De la ecuación $\log f(x) - \log x + xf(x) = 0$ teniendo en cuenta que f es continua y diferenciable podemos deducir además que dicho máximo es el máximo absoluto de f. En efecto, como $f'(x_0) = 0$, y en x_0 f tiene un máximo local, entonces en $(0, x_0)$, f es estrictamente creciente y en (x_0, ∞) es estrictamente decreciente. De hecho se tiene que $f(x) \to 0$, cuando $x \to 0$ y $x \to +\infty$, lo cual dejamos como ejercicio para el lector interesado. □

3.3. Caso de un sistema de ecuaciones

Sea ahora el sistema de ecuaciones

$$\left\{\begin{array}{l} F_1(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0, \\ F_2(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0, \\ \qquad\qquad \vdots \\ F_m(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0, \end{array}\right. \tag{3.3.1}$$

donde $F_k : A \subset \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$, $k = 1, 2, \ldots, m$, o, equivalentemente, la ecuación $F(x, y) = 0$, donde la función $F : A \subset \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^m$ tiene por componentes las funciones F_k definidas en el sistema (3.3.1).

El problema consiste en saber si existe una $f : I_x \subset \mathbb{R}^n \mapsto I_y \subset \mathbb{R}^m$ tal que $F(x, f(x)) = 0$. Es decir, si existen m funciones $f_k : \mathbb{R}^n \mapsto \mathbb{R}$, $y_k = f_k(x) := f_k(x_1, \cdots, x_n)$ tales que para todos $k = 1, \ldots, m$, se cumple

$$F_k(x_1, \ldots, x_n, f_1(x), \ldots, f_m(x)) = 0,$$

en todo un entorno de $(x_0, y_0) \in A$.

3.3.1. Un caso particular representativo

Antes de enunciar el teorema vamos a mostrar un ejemplo de como resolver el problema anterior en el caso $n = m = 2$. Sea el sistema de ecuaciones

$$F_1(x_1, x_2, y_1, y_2) = 0, \quad F_2(x_1, x_2, y_1, y_2) = 0.$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Si derivamos ambas ecuaciones respecto a x_1 y x_2 y asumimos que y_1 e y_2 dependen de x_1 y x_2 obtenemos, usando la regla de la cadena (ver Teorema 2.3.1),

$$\frac{\partial F_1}{\partial x_i}+\frac{\partial F_1}{\partial y_1}\frac{\partial y_1}{\partial x_i}+\frac{\partial F_1}{\partial y_2}\frac{\partial y_2}{\partial x_i}=0,\quad \frac{\partial F_2}{\partial x_i}+\frac{\partial F_2}{\partial y_1}\frac{\partial y_1}{\partial x_i}+\frac{\partial F_2}{\partial y_2}\frac{\partial y_2}{\partial x_i}=0,\quad i=1,2,$$

que no es más que un sistema de cuatro ecuaciones con cuatro incógnitas que en forma matricial toma la forma

$$\begin{pmatrix}\frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2}\\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2}\end{pmatrix}\begin{pmatrix}\frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2}\\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2}\end{pmatrix}=-\begin{pmatrix}\frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2}\\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2}\end{pmatrix}.$$

Si la matriz del sistema anterior $F'_y:=\begin{pmatrix}\frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2}\\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2}\end{pmatrix}$ es invertible (i.e., si $\det F'_y\neq 0$) entonces el sistema es compatible y su solución se expresa mediante la fórmula

$$\begin{pmatrix}\frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2}\\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2}\end{pmatrix}=-\begin{pmatrix}\frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2}\\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2}\end{pmatrix}^{-1}\begin{pmatrix}\frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2}\\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2}\end{pmatrix}.$$

3.3.2. El Teorema de la función implícita para un sistema de ecuaciones.

Sea $x_0:=(x_{01},x_{02},\ldots,x_{0n})$ e $y_0:=(y_{01},y_{02},\ldots,y_{0m})$ y denotemos por I_x el intervalo (x_0-h,x_0+h) y por I_y el intervalo (y_0-k,y_0+k).

Definamos las matrices (aplicaciones lineales)

$$f':\mathbb{R}^n\mapsto\mathbb{R}^m,\quad f'(x)=\begin{pmatrix}\frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \cdots & \frac{\partial f_1(x)}{\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial f_m(x)}{\partial x_1} & \frac{\partial f_m(x)}{\partial x_2} & \cdots & \frac{\partial f_m(x)}{\partial x_n}\end{pmatrix},\tag{3.3.2}$$

$$F'_x:\mathbb{R}^n\mapsto\mathbb{R}^m,\quad F'_x(x,y)=\begin{pmatrix}\frac{\partial F_1(x,y)}{\partial x_1} & \frac{\partial F_1(x,y)}{\partial x_2} & \cdots & \frac{\partial F_1(x,y)}{\partial x_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial F_m(x,y)}{\partial x_1} & \frac{\partial F_m(x,y)}{\partial x_2} & \cdots & \frac{\partial F_m(x,y)}{\partial x_n}\end{pmatrix},\tag{3.3.3}$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

$$(3.3.4)\quad F'_y : \mathbb{R}^m \mapsto \mathbb{R}^m,\ \ F'_y(x,y) = \begin{pmatrix} \dfrac{\partial F_1(x,y)}{\partial y_1} & \dfrac{\partial F_1(x,y)}{\partial y_2} & \cdots & \dfrac{\partial F_1(x,y)}{\partial y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial F_m(x,y)}{\partial y_1} & \dfrac{\partial F_m(x,y)}{\partial y_2} & \cdots & \dfrac{\partial F_m(x,y)}{\partial y_m} \end{pmatrix}.$$

Además $F'_y(x,y)$ es una matriz cuadrada que será invertible si y sólo si $\det F'_y(x,y) \neq 0$.

Usando la notación anterior tenemos el siguiente teorema:

Teorema 3.3.1 (de sistemas de funciones implícitas) *Sea $F : A \subset \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^m$ definida en un entorno del punto*[3] *$(x_0, y_0) \in A$, A abierto de $\mathbb{R}^n \times \mathbb{R}^m$. Supongamos que:*

1. *$F(x,y) \in C^{(p)}(A)$, $p \geq 1$,*
2. *$F(x_0, y_0) = 0$,*
3. *$\det F'_y(x_0, y_0) \neq 0$ o sea,$F'_y(x,y)$ es una matriz invertible.*

Entonces existe un intervalo $I = I_x \times I_y = (x_0 - h, x_0 + h) \times (y_0 - k, y_0 + k)$ alrededor del punto (x_0, y_0), $I \subset A$, y una función $f : I_x \subset \mathbb{R}^n \mapsto I_y \subset \mathbb{R}^m$ tal que:

1. *$F(x,y) = 0$ en I si y sólo si $f(x) = y$,*
2. *$f(x) \in C^{(p)}(I_x)$.*
3. *Para todo $x \in I_x$, las derivadas parciales de $f(x)$ se calculan por la fórmula*

$$(3.3.5)\qquad f'(x) = -[F'_y(x, f(x))]^{-1} \cdot [F'_x(x, f(x))], \quad i = 1, 2, \ldots, n.$$

Demostración: La prueba es por inducción y la omitiremos. El lector interesado puede consultarla, por ejemplo, en [2, 8].

Ejemplo 3.3.2 Sea el sistema

$$(3.3.6)\qquad \begin{cases} (x-1)^2 + y^2 - z = 0, \\ x^2 + y^2 + z^2 = 1. \end{cases}$$

Decidir si este sistema se puede resolver de forma que existan las funciones $y = y(x)$ y $z = z(x)$ y calcular los valores de $y'(x)$ y $z'(x)$.

[3] Aquí entenderemos que $(x,y) \in \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$, i.e., $(x,y) = (x_1, x_2, \ldots, x_n, y_1 \cdots, y_m)$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Definamos la función $F : \mathbb{R} \times \mathbb{R}^2 \mapsto \mathbb{R}^2$

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

Está claro que $F \in C^{(\infty)}(\mathbb{R}^3)$. Su matriz $F'_y : \mathbb{R}^2 \mapsto \mathbb{R}^2$, tiene la forma (3.3.4)

$$F'_y(x, y, z) = \begin{pmatrix} 2y & -1 \\ 2y & 2z \end{pmatrix}, \qquad \det F'_y(x, y, z) = 2y(2z + 1).$$

$\det F'_y = 0$ en los puntos (a, b, c) tales que $b = 0$ o $c = -1/2$. Teniendo en cuenta la primera ecuación $c = (a-1)^2 + y^2$ el punto $c = -1/2$ queda descartado. Si $b = 0$ entonces $c = (a-1)^2$ y $a^2 + c^2 = 1$. De lo anterior se sigue que $a^2 + (a-1)^4 = 1$ que sólo tiene dos raíces reales: $a = 0$ y $a = 1$. Usando entonces la expresión $c = (a-1)^2$ tenemos que el Teorema de la función implícita no es aplicable en los puntos $(0, 0, 1)$ y $(1, 0, 0)$. Si asumimos que existe algún punto (a, b, c) distintos de los anteriores donde el sistema tenga solución entonces podemos aplicar el Teorema de la función implícita que nos asegura que existen las funciones $y = y(x)$ y $z = z(x)$ definidas por el sistema y que además podemos calcular sus derivadas por la fórmula (3.3.5)

$$f'(x) = \begin{pmatrix} y'(x) \\ z'(x) \end{pmatrix} = -\begin{pmatrix} 2y & -1 \\ 2y & 2z \end{pmatrix}^{-1} \begin{pmatrix} 2(x-1) \\ 2x \end{pmatrix} = -\frac{1}{y(2z+1)} \begin{pmatrix} 2zx - 2z + x \\ 2y \end{pmatrix}.$$

Este ejemplo lo podemos resolver por el método descrito en el apartado 3.3.1. Si derivamos la ambas ecuaciones en (3.3.6) por x obtenemos el sistema

$$\text{(3.3.7)} \qquad \begin{cases} 2(x-1) + 2yy_x - z_x = 0, \\ 2x + 2yy_x + 2zz_x = 0. \end{cases}$$

Restando la ambas obtenemos

$$(2z+1)z_x + 2 = 0 \quad \Rightarrow \quad z_x = -\frac{2}{2z+1}, \quad z \neq -\frac{1}{2}.$$

Multiplicando la primera por $2z$ y sumando

$$y(2z+1)y_x + 2z(x-1) - x = 0 \;\Rightarrow\; y_x = -\frac{2z(x-1) - x}{y(2z+1)}, \quad y \neq 0,\ z \neq -\frac{1}{2},$$

que coincide con el resultado anterior. □

3.3.3. Problemas propuestos

Problema 3.3.1 Prueba que la ecuación $x^2/a^2 + y^2/b^2 + z^2/c^2 = r^2$ define una función $z = f(x, y)$ diferenciable en el entorno de cualquier punto $A = (x_0, y_0, z_0)$ con $z_0 \neq 0$. Encuentra la ecuación de plano tangente a la superficie en dicho punto A. ¿Encuentra la ecuación del plano en los puntos con $z_0 = 0$?

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Problema 3.3.2 Decide para qué valores de la constante a, la ecuación $x^2+y^3+xy+x^3+ay=0$ define y como función implícita de x en un entorno del punto $(0,0)$. Hallar $y'(0)$. ¿Define la ecuación a x como función implícita de y en un entorno del mismo punto? Decide si $y(x)$ tiene un extremo en $x=0$. Encuentra el polinomio de Taylor de orden tres de y en un entorno de $x=0$.

Problema 3.3.3 Estudiar para qué valores de la constante a, la ecuación $x^2+y^3+xy+x^3+ay=0$ define y como función implícita de x en un entorno del punto (0,0). Hallar $y'(0)$. ¿Define la ecuación a x como función implícita de y en un entorno del mismo punto? Prueba que $y(x)$ tiene un extremo en $x=0$ y decide si es máximo o mínimo local.

Problema 3.3.4 Probar que existen funciones f y g de clase $C^{(\infty)}$ definidas en un entorno de $(1,1)$, tales que $f(1,1)=-1$, $g(1,1)=0$ y verificando las ecuaciones: $f(x,y)^3+xg(x,y)^2+y=0$, $g(x,y)^3+yg(x,y)+f(x,y)^2=x$. Encuentra las derivadas de f y g en dicho punto.

Problema 3.3.5 Sea la ecuación $F(x,y,z)=x^2y^2z^2+\exp(x+y+2z)+5y^3-4y-2=0$.

1. Prueba que la ecuación anterior define una función $z=f(x,y)$ en el entorno U del punto $A=(-1,1,0)$ y que dicha función es una función $C^{(\infty)}(U)$ en dicho entorno U.
2. Calcula las derivadas parciales z_x y z_y en dicho punto.
3. Calcula la derivada direccional de f en el punto $(-1,1)$ según la dirección $(1,1)$. ¿En que dirección dicha derivada direccional es mínima? ¿y máxima?
4. Encuentra la ecuación del plano tangente a la superficie $z=f(x,y)$ en el punto A.
5. Escribe el polinomio de Taylor de orden 2 de $z=f(x,y)$ alrededor del punto A.

Problema 3.3.6 Sea la ecuación $x^2z-yz^2+x\cos(xz^2)-1=0$.

1. Para que valores de $a,b\in\mathbb{R}$ la ecuación anterior define una función $z=f(x,y)$ en el entorno U del punto $(a,b,0)$. ¿Es para alguno de dichos puntos la función $z=f(x,y)$ diferenciable? ¿Cuántas veces? Justifica la respuesta.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

2. Calcula, si es posible, las derivadas z_x y z_y en los puntos $(a, b, 0)$.

3. En que dirección es máxima la variación de $z = f(x, y)$ en $(1, 1, 0)$.

Problema 3.3.7 Sea la ecuación $F(x, y, z) = x^2z^2 + \exp(x+y+z) - 3y^3 - ay - 2 = 0$, $a \in \mathbb{R}$.

1. ¿Para qué valores de a la ecuación anterior define una función $z = f(x, y)$ en el entorno U del punto $A = (-1, 1, 0)$ y cuántas veces podemos derivar dicha función $z = f(x, y)$ en dicho entorno? Justifica la respuesta.

2. Calcula las derivadas parciales z_x y z_y en un entorno de A y en el propio punto A.

3. Calcula la derivada direccional de $z = f(x, y)$ en $a = (-1, 1)$ según la dirección $u = (1, 1)$.

4. Escribe, si es posible, la ecuación del plano tangente a la superficie definida por la ecuación $F(x, y, z) = 0$ en el punto A.

5. Encuentra el polinomio de Taylor de orden 2 de $z = f(x, y)$ alrededor del punto A.

3.4. Teorema de la función inversa

Veamos un caso particular de especial importancia. Sea $f : A \subset \mathbb{R} \mapsto \mathbb{R}$ una función de una variable, sea $y \in \mathbb{R}$ y supongamos que queremos resolver la ecuación $f(x) = y$. Para ello la reescribiremos de la forma $F(x, y) = f(x) - y = 0$. Lo que queremos es saber si esta ecuación es resoluble respecto a x, i.e., si existe una función $x = g(y)$ de forma tal que $F(g(y), y) = 0$ para todo y de cierto intervalo dado. Es obvio que si en cierto intervalo I_y existe la solución definiendo I_x el conjunto de las x tales que $x = g(y)$ tendremos dos funciones $f(x)$ y $g(y)$ que son mutuamente inversas. Es decir, encontrando las condiciones que nos permiten resolver la ecuación $F(x, y) = 0$ respecto a x, sabremos en que condiciones $f(x)$ es invertible. Pero eso es justo lo que nos afirma el Teorema de la función implícita. Por ejemplo, basta que F sea $C^{(p)}(A)$, con A cierto entorno abierto de cierto (x_0, y_0) que satisface la ecuación $f(x_0) = y_0$ y que $F'_x(x_0, y_0) = f'(x_0) \neq 0$ para asegurar que f tiene en un cierto entorno de x_0 inversa, que va a ser además $C^{(p)}$ y su derivada se expresará por

$$g'(y_0) = -\frac{F'_x(x_0, y_0)}{F'_y(x_0, y_0)} = \frac{1}{f'(x_0)}.$$

DERIVACIÓN DE FUNCIONES DE VARIAS VARIABLES

Enunciemos a continuación el resultado general:

Teorema 3.4.1 (de la función inversa) *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ definida en un entorno del punto $x_0 \in A$ tal que*

1. *$f(x) \in C^{(p)}(A)$, $p \geq 1$,*
2. *$f(x_0) = y_0$, en x_0,*
3. *$f'(x_0)$ es una aplicación invertible.*

Entonces existe un entorno abierto $U(x_0) \subset A$ de $x_0 \in A$ y otro $V(y_0) \subset f(A)$ de $y_0 \in f(A)$ tal que f es invertible en $U(x_0)$, i.e., existe su inversa $f^{-1} : V(y_0) \mapsto U(x_0)$, $f^{(-1)} \in C^{(p)}(V(y_0))$, además, para todo $x \in U(x_0)$ e $y = f(x) \in V(y_0)$ se tiene que $(f^{-1}(y))' := Df^{-1}(y) = [f'(x)]^{-1} := [Df(x)]^{-1}$.

Veamos dos ejemplos.

Ejemplo 3.4.2 Sea la función $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ definida por $y = f(x) = Ax$, donde A es una matriz real $n \times n$. Es obvio que f es $C^{(p)}(\mathbb{R}^n)$ para todo $p \in \mathbb{N}$. Podemos además tomar cualquier $x \in \mathbb{R}^n$ y definir $y = Ax$. Obviamente la derivada (total) de f es la matriz A. Entonces si A es invertible (o equivalentemente, si el Jacobiano de f, que es $\det A$ es diferente de cero), entonces f es invertible. Además $Df^{-1} = [Df]^{-1}$, i.e., $[Df(x)]^{-1} = A^{-1}$. □

Ejemplo 3.4.3 Consideremos la función

$$f(x, y) = \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -3x + y^3 \\ -3y + x^3 \end{pmatrix}. \tag{3.4.1}$$

Calculemos la matriz jacobiana (2.2.4) de f así como su jacobiano

$$J_f(x, y) = \begin{pmatrix} -3 & 3y^2 \\ 3x^2 & -3 \end{pmatrix}, \quad \det J_f(x, y) = 9(1 - x^2y^2).$$

De lo anterior se sigue que podemos aplicar el Teorema de la función inversa 3.4.1 para todo (x_0, y_0) tal que $x_0^2 y_0^2 \neq 1$. Luego, f es localmente invertible para todo (x_0, y_0) tal que $x_0^2 y_0^2 \neq 1$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Una pregunta natural es qué ocurre en los casos donde el teorema no es aplicable. Por ejemplo tomemos $x = y = 1$. En este caso $u = -2$ y $v = -2$, así que tenemos el sistema

$$\begin{pmatrix} -3x + y^3 \\ -3y + x^3 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \end{pmatrix}.$$

Podemos comprobar que aparte de los valores $x = 1$ e $y = 1$ anteriores, también es solución $x = -2$ e $y = -2$, es decir, la función f, en general no es inyectiva.

Uno podría pensar que entonces cuando $x_0^2 y_0^2 \neq 1$ la función es inyectiva ya que es localmente invertible. Veamos que, en general, esta afirmación también es falsa. Para ello tomemos $u = 5$ y $v = -5$. Se puede comprobar que hay al menos dos soluciones $x = -2$, $y = -1$ y $x = 1$, $y = 2$. Es decir, el teorema es un resultado local.

Encontremos ahora el polinomio de Taylor de orden dos de la función inversa $f^{-1}(u, v) = (x(u, v), y(u, v))$ alrededor del punto $x = 1$, $y = 2$, $u = 5$ y $v = -5$, donde sabemos que existe la inversa local. Para ello reescribiremos (3.4.1) como

$$u + 3x - y^3 = 0, \quad v + 3y - x^3 = 0,$$

y asumiremos que x e y son funciones de u y v.

Derivando cada ecuación respecto a u y a v para obtenemos el sistema lineal

$$\begin{cases} -3y^2 y_u + 3x_u + 1 &= 0, \\ 3y_u - 3x^2 x_u &= 0, \\ 3x_v - 3y^2 y_v &= 0, \\ 3y_v - 3x^2 x_v + 1 &= 0. \end{cases}$$

Sustituyendo los valores de $x = 1$, $y = 2$, en dicho sistema y resolviendo obtenemos a los valores de las primeras derivadas parciales de x e y

$$x_u = \tfrac{1}{9}, \quad x_v = \tfrac{4}{9}, \quad y_u = \tfrac{1}{9}, \quad y_v = \tfrac{1}{9}.$$

A continuación derivamos las dos ecuaciones dos veces respecto a u, dos respecto a v y una vez respecto a u y a v obteniendo el sistema

$$\begin{cases} -3y^2 y_{uu} - 6y y_u^2 + 3x_{uu} &= 0, \\ 3y_{uu} - 3x^2 x_{uu} - 6x x_u^2 &= 0, \\ -3y^2 y_{vv} - 6y y_v^2 + 3x_{vv} &= 0, \\ 3y_{vv} - 3x^2 x_{vv} - 6x x_v^2 &= 0, \\ -6y y_u y_v - 3y^2 y_{uv} + 3x_{uv} &= 0, \\ 3y_{uv} - 6x x_u x_v - 3x^2 x_{uv} &= 0. \end{cases}$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

En dicho sistema sustituimos los valores de $x = 1$, $y = 2$, así como los valores de las las derivadas x_u, x_v, y_x e y_v obtenidos antes cuya solución es

$$x_{uu} = -\tfrac{4}{81},\ x_{vv} = -\tfrac{44}{81},\ x_{uv} = -\tfrac{4}{27},\ y_{uu} = -\tfrac{2}{81},\ y_{vv} = -\tfrac{4}{27},\ y_{uv} = -\tfrac{4}{81}.$$

Así, el polinomio de Taylor de orden dos de $f^{-1}(u, v)$ alrededor de $u = 5$, $v = -5$ para la primera componente de f^{-1} es

$$x(u,v) = 5 + \frac{1}{9}(u-5) + \frac{4}{9}(v+5) + \frac{1}{2}(u-5\,v+5)\begin{pmatrix} -\frac{4}{81} & -\frac{4}{27} \\ -\frac{4}{27} & -\frac{44}{81} \end{pmatrix}\begin{pmatrix} u-5 \\ v+5 \end{pmatrix},$$

mientras que para la segunda es

$$y(u,v) = -5 + \frac{1}{9}(u-5) + \frac{1}{9}(v+5) + \frac{1}{2}(u-5\,v+5)\begin{pmatrix} -\frac{2}{81} & -\frac{4}{81} \\ -\frac{4}{81} & -\frac{4}{27} \end{pmatrix}\begin{pmatrix} u-5 \\ v+5 \end{pmatrix}.$$

□

3.4.1. Problemas propuestos

Problema 3.4.1 Sea $f : \mathbb{R}^2 \longmapsto \mathbb{R}^2$, definida por $f(u, v) = (e^u + e^v, e^u - e^v)$.

a) Determinar los puntos de $\mathbb{R}^2$ en los que f es localmente invertible.

b) ¿Es f globalmente invertible? En caso afirmativo caracterizar la imagen de f y calcular la función inversa f^{-1}.

Problema 3.4.2 Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $f(x, y) = (e^{x+y}, x^2)$.

a) ¿En qué puntos es aplicable el teorema de la función inversa. Calcula la derivada de la función inversa local.

b) Estudia la inversión global de f en regiones de $\mathbb{R}^2$.

Problema 3.4.3 Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $f(x, y) = (xe^y, xe^{-y})$.

1. ¿En qué puntos de $\mathbb{R}^2$ se puede garantizar la existencia de inversa local?

2. ¿Es f inyectiva? En caso negativo, encuentra el mayor subconjunto de $\mathbb{R}^2$ donde lo es, caracteriza su imagen y da una expresión de f^{-1}.

Problema 3.4.4 Sea $f(x, y) = (e^x \cos y, e^x \operatorname{sen} y)$.

a) Probar que f es localmente invertible, pero no globalmente, aunque su jacobiano es distinto de 0 siempre.

b) Comprobar que si $U = \{(x, y) \in \mathbb{R}^2 : -\pi \leq y < \pi\}$ entonces $f_{|U}$ es invertible, siendo su inversa continua. ¿Es dicha región U la más grande posible donde para la que existe la inversa de f?

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

3.5. Cambio de variables

Supongamos que tenemos una expresión del tipo

$$\Phi(x, y, z, z_x, z_y, z_{xx}, z_{xy}, z_{yy}, \dots) = 0$$

donde x e y son variables independientes y z es una función $z : \mathbb{R}^2 \mapsto \mathbb{R}$, $z = z(x, y)$ y queremos escribirlas en las nuevas variables u, v y $w = w(u, v)$ asumiendo que las variables nuevas y viejas se relacionan mediante el sistema

$$g_i(x, y, z, u, v, w) = 0, \qquad i = 1, 2, 3,$$

que denominaremos expresiones del cambio de variables, donde las funciones g_i, $i = 1, 2, 3$ se asumen diferenciables tantas veces como haga falta.

Hay dos opciones de especial interés y es cuando el cambio de variables es de la forma (variables viejas en función de las nuevas)

$$x = f_1(u, v, w), \quad y = f_2(u, v, w), \quad z = f_3(u, v, w), \tag{3.5.1}$$

o (variables nuevas en función de las viejas)

$$u = f_1(x, y, z), \quad v = f_2(x, y, z), \quad w = f_3(x, y, z). \tag{3.5.2}$$

Aquí nos centraremos en el primer caso que suele ser el más usado en la práctica. Diferenciando (3.5.1) tenemos

$$\begin{aligned} dx =& D_u f_1 du + D_v f_1 dv + D_w f_1 dw, \\ dy =& D_u f_2 du + D_v f_2 dv + D_w f_2 dw, \\ dz =& D_u f_3 du + D_v f_3 dv + D_w f_3 dw, \end{aligned} \tag{3.5.3}$$

donde D_u, D_v y D_w son las correspondientes derivadas parciales respecto a las variables u, v y w, respectivamente. Si usamos que $dw = D_u w du + D_v w dv$ tenemos

$$\begin{aligned} dx =& \mathcal{D}_u f_1 du + \mathcal{D}_v f_1 dv, \\ dy =& \mathcal{D}_u f_2 du + \mathcal{D}_v f_2 dv, \\ dz =& \mathcal{D}_u f_3 du + \mathcal{D}_v f_3 dv \end{aligned} \tag{3.5.4}$$

donde

$$\mathcal{D}_u = \frac{\partial}{\partial u} + \frac{\partial w}{\partial u}\frac{\partial}{\partial w} = \frac{\partial}{\partial u} + w_u \frac{\partial}{\partial w}, \quad \mathcal{D}_v = \frac{\partial}{\partial v} + \frac{\partial w}{\partial v}\frac{\partial}{\partial w} = \frac{\partial}{\partial v} + w_v \frac{\partial}{\partial w}.$$

Si el determinante

$$\Delta = \begin{vmatrix} \mathcal{D}_u f_1 & \mathcal{D}_v f_1 \\ \mathcal{D}_u f_2 & \mathcal{D}_v f_2 \end{vmatrix} \neq 0$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

entonces las dos primeras ecuaciones de (3.5.4) se pueden resolver expresándose las diferenciales du y dv en función de las dx y dy

$$\begin{aligned} du =& \frac{1}{\Delta}\Big(\mathcal{D}_v f_2 dx - \mathcal{D}_v f_1 dy\Big), \\ dv =& \frac{1}{\Delta}\Big(- \mathcal{D}_u f_2 dx + \mathcal{D}_u f_1 dy\Big), \end{aligned} \tag{3.5.5}$$

que sustituimos en la tercera expresión de (3.5.4) obteniendo

$$dz = \frac{1}{\Delta}\Big(\mathcal{D}_u f_3 \mathcal{D}_v f_2 - \mathcal{D}_v f_3 \mathcal{D}_u f_2\Big)dx + \frac{1}{\Delta}\Big(- \mathcal{D}_u f_3 \mathcal{D}_v f_1 + \mathcal{D}_v f_3 \mathcal{D}_u f_1\Big)dy,$$

de donde deducimos

$$\begin{aligned} z_x =& \frac{1}{\Delta}\Big(\mathcal{D}_u f_3 \mathcal{D}_v f_2 - \mathcal{D}_v f_3 \mathcal{D}_u f_2\Big) = F_1(u, v, w, w_u, w_v), \\ z_y =& \frac{1}{\Delta}\Big(- \mathcal{D}_u f_3 \mathcal{D}_v f_1 + \mathcal{D}_v f_3 \mathcal{D}_u f_1\Big) = F_2(u, v, w, w_u, w_v). \end{aligned} \tag{3.5.6}$$

Si queremos obtener las expresiones de las segundas derivadas procedemos como sigue: Calculamos

$$d(z_x) = z_{xx}dx + z_{xy}dy = D_u F_1 du + D_v F_1 dv + D_w F_1 dw + D_{w_u} F_1 dw_u + D_{w_v} F_1 dw_v.$$

A continuación sustituimos en la parte derecha los valores de las diferenciales nuevas

$$dw = w_u du + w_v dv, \quad dw_u = w_{uu} du + w_{vu} dv, \quad dw_v = w_{uv} du + w_{vv} dv$$

y en la expresión resultante sustituimos los valores de las diferenciales du y dv obtenidos en (3.5.5). Esto nos da una expresión de $d(z_x)$ en función de las diferenciales antiguas. Igualando las expresiones delante de las diferenciales dx y dy obtenemos los valores z_{xx} y z_{xy} respectivamente. Para obtener z_{yy} se procede de forma análoga pero partiendo de la segunda ecuación de (3.5.6).

Veamos un ejemplo.

Ejemplo 3.5.1 Sea la expresión $\Phi(x, y, z, z_x, z_y, z_{xx}, z_{xy}, z_{yy}, \dots) = z_{xx} + z_{xy} + z_x - z = 0$, con $z = z(x, y)$. Hagamos en cambio $x = u+v$, $y = u-v$ y $z = we^{v-u}$. Encontrar la expresión de Φ en las nuevas variables.

Realizar todos los cálculos requiere bastante trabajo así que es recomendable usar un programa de cálculo simbólico como, por ejemplo, MAXIMA CAS. Por completitud resumiremos los cálculos aquí.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Diferenciando las ecuaciones del cambio de variables tenemos

$$\begin{aligned} dx =& du + dv, \quad dy = du - dv, \\ dz =& \left(e^{v-u} w_v + e^{v-u}\, w\right)\, dv + \left(e^{v-u}\, w_u - e^{v-u}\, w\right)\, du. \end{aligned} \tag{3.5.7}$$

De las dos primeras deducimos

$$du = \frac{dy + dx}{2}, \qquad dv = \frac{dx - dy}{2}.$$

Sustituyendo lo anterior en la tercera y usando $dz = z_x dx + z_y dy$ obtenemos

$$z_x = \frac{e^{v-u}\, (w_u + w_v)}{2}, \quad z_y = -\frac{e^{v-u}\, (w_v - w_u + 2w)}{2}.$$

Para obtener z_{xx} tenemos que calcular la diferencial de dz_x tal y como se explicó anteriormente. Tras las correspondientes simplificaciones en resultado es

$$z_{xx} = \frac{e^{v-u}\, (w_{uu} + w_{vv} + 2w_{uv})}{4}, \quad z_{xy} = -\frac{e^{v-u}\, (w_{vv} - w_{uu} + 2\,(w_v + w_u))}{4}.$$

Sustituyendo los valores obtenidos para las distintas derivadas en la expresión de Φ obtenemos la nueva ecuación $w_{uu} + w_{uv} = 2w$. □

3.5.1. Problemas propuestos

Encontrar las expresiones en las nuevas variables de las siguientes expresiones:

1. $(x\, z_x)^2 + (y\, z_y)^2 = z^2\, z_x\, z_y$ con $x = u \exp(w)$, $y = v \exp(w)$, y $z = w \exp(w)$;
2. $z_{xx} - z_{yy} = 0$ con $x = u + v$, $y = u - v$, y $z = w$;
3. $z_{xx} + z_{yy} = 0$ con $x = r \cos(\phi)$, $y = r \sin(\phi)$, $z = w$;
4. $z_{xx} - z_y = 0$ con $x = -u/v$, $y = -1/v$, $z = \sqrt{-v} e^{u^2/(4v)} w$.

3.6. Dependencia funcional

Consideremos ahora el problema de cuando un conjunto de funciones diferenciables es funcionalmente dependiente o no. Para ello necesitamos las siguientes dos definiciones.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Definición 3.6.1 *Sean las funciones $\phi_k : \mathbb{R}^n \mapsto \mathbb{R}$, $k = 1, 2, \ldots, m$, $m \leq n$ diferenciables en cierto abierto $A \subset \mathbb{R}^n$ y sea $u_k = \phi_k(x_1, \ldots, x_n)$. Diremos que una de las funciones anteriores u_k (para cierto $k = 1, 2, \ldots, m$) depende funcionalmente de las demás en A si para todos los $x \in A$ existe una función $\Phi : \mathbb{R}^{m-1} \mapsto \mathbb{R}$, diferenciable en el correspondiente dominio[4] $B \subset \mathbb{R}^{m-1}$, tal que $u_k = \Phi(u_1, \ldots, u_{k-1}, u_{k+1}, \ldots, u_m)$.*

Definición 3.6.2 *Dado el sistema de funciones diferenciables*

$$\left\{\begin{array}{l} u_1 = \phi_1(x_1, \ldots, x_n), \\ u_2 = \phi_2(x_1, \ldots, x_n), \\ \qquad \vdots \\ u_m = \phi_m(x_1, \ldots, x_n), \end{array}\right. \qquad m \leq n, \tag{3.6.1}$$

diremos que es funcionalmente dependiente en el abierto $A \subset \mathbb{R}^n$ si al menos una de las funciones u_k depende funcionalmente de las demás, en caso contrario diremos que el sistema es funcionalmente independiente.

Ejemplo 3.6.3 El sistema

$$\left\{\begin{array}{l} u_1 = x_1^2 + x_2^2 + x_3^2, \\ u_2 = x_1 + x_2 + x_3, \\ u_3 = 2x_1x_2 + 2x_1x_3 + 2x_2x_3, \end{array}\right.$$

es funcionalmente dependiente pues $u_1 = \Phi(u_2, u_3) = u_2^2 - u_3$. □

Ejemplo 3.6.4 El sistema lineal

$$\left\{\begin{array}{l} u_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ u_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ \qquad \vdots \\ u_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n, \end{array}\right. \qquad m \leq n, \tag{3.6.2}$$

es funcionalmente independiente si el rango de la matriz del sistema es igual a m y dependiente si es menor que m. □

[4] B es un abierto contenido en el conjunto definido por los vectores $(\phi_1(A), \ldots, \phi_{k-1}(A), \phi_{k+1}(A), \ldots, \phi_m(A))$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Sea el sistema de ecuaciones lineales

$$\left\{\begin{array}{l} 0 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \\ 0 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \\ \qquad \vdots \\ 0 = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n, \end{array}\right. \qquad m \leq n, \tag{3.6.3}$$

o, en forma matricial, $Ax = 0$, donde

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}.$$

Si el rango de la matriz A del sistema es $r \leq m$ entonces existe al menos un menor de orden r, A_r, con determinante distinto de cero, que por simplicidad tomaremos el que aparece en el ángulo superior de la izquierda

$$A_r = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1r} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2r} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & a_{r3} & \cdots & a_{rr} \end{pmatrix}, \quad \det A_r \neq 0.$$

Entonces, las r primeras filas de la matriz A son linealmente independientes y las restantes $m - r$ filas se pueden expresar como combinaciones lineales de dichas r filas, i.e., las r primera ecuaciones del sistema (3.6.3) son independientes y las $m - r$ restantes son combinaciones lineales de dichas r ecuaciones. De lo anterior se deduce que las ecuaciones lineales (3.6.2) son linealmente independientes (y por tanto funcionalmente) si y solo si el rango de A es m. Este resultado es un teorema conocido del álgebra lineal. □

Teorema 3.6.5 *Supongamos que tenemos m funciones de n variables $m \leq n$, $\phi_k : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, $k = 1, 2, \ldots, m$, definidas en (3.6.1) que son diferenciables en cierto abierto $A \subset \mathbb{R}^n$ y sea $a \in A$. Entonces, si el jacobiano (o sea el determinante de la correspondiente matriz jacobiana) formado por las funciones $u_1, \ldots, u_m$ respecto a alguna combinación de m variables (digamos $x_{n_1}, \ldots, x_{n_m}$) es distinto de cero en a, las funciones $u_1, \ldots, u_m$ son funcionalmente independientes en todo un entorno $U(a)$ de a.*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Demostración: Por simplicidad asumamos que[5]

$$\det J = \det \begin{pmatrix} D_1u_1(a) & D_2u_1(a) & \dots & D_mu_1(a) \\ \vdots & \vdots & \ddots & \vdots \\ D_1u_m(a) & D_2u_m(a) & \dots & D_mu_m(a) \end{pmatrix} \neq 0. \tag{3.6.4}$$

Supongamos que en estas condiciones las funciones $u_1, \dots, u_m$ son funcionalmente dependientes. Entonces, una de ellas, que supondremos sin pérdida de generalidad u_m, dependerá de las demás. Es decir, existe una función Φ diferenciable en el entorno $U(a)$ tal que $u_m = \Phi(u_1, u_2, \dots, u_{m-1})$. Luego

$$D_k u_m = \frac{\partial u_m}{\partial x_k} = \frac{\partial \Phi}{\partial u_1}\frac{\partial u_1}{\partial x_k} + \dots + \frac{\partial \Phi}{\partial u_{m-1}}\frac{\partial u_{m-1}}{\partial x_k}, \quad k = 1, 2, \dots m.$$

Pero entonces la última fila (la m-ésima) del jacobiano (3.6.4) es combinación lineal de las filas anteriores $1, 2, \dots, m-1$, lo cual, por las propiedades de los determinantes implica que $\det J = 0$, lo cual es una contradicción. □

Ejemplo 3.6.6 Sea el sistema de funciones $u = x + y$ y $v = x - y$. Como

$$\det \begin{pmatrix} D_x u & D_y u \\ D_x v & D_y v \end{pmatrix} = \det \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -2 \neq 0,$$

estas funciones son funcionalmente independientes. □

Como corolario inmediato del teorema anterior la siguiente condición necesaria de dependencia funcional:

Corolario 3.6.7 *Si el sistema de m funciones diferenciables de n variables $m \leq n$ en un entorno $U(a) \subset A$, A abierto, de cierto punto $a \in A$ $\phi_k : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, $k = 1, 2, \dots, m$, definido en (3.6.1) es un sistema funcionalmente dependiente, entonces su jacobiano (3.6.4) respecto a cualquier combinación de $m \leq n$ variables es nulo.*

La condición anterior se puede parafrasear diciendo que si el sistema (3.6.1) es funcionalmente dependiente entonces el rango de la correspondiente matriz jacobiana del sistema es menor que m.

[5]Siempre podemos renombrar las variables para que dicho determinante sea distinto de cero.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Ejemplo 3.6.8 Si tomamos las funciones del ejemplo 1, estas son dependientes. En efecto, un cálculo directo nos da:

$$\det \begin{pmatrix} D_1u_1(a) & D_2u_1(a) & D_3u_1(a) \\ D_1u_2(a) & D_2u_2(a) & D_3u_2(a) \\ D_1u_3(a) & D_2u_3(a) & D_3u_3(a) \end{pmatrix}$$
$$= \det \begin{pmatrix} 2x_1 & 2x_2 & 2x_3 \\ 1 & 1 & 1 \\ 2(x_2+x_3) & 2(x_1+x_3) & 2(x_1+x_2) \end{pmatrix} = 0.$$

de donde se sigue el resultado. □

Para terminar este apartado enunciaremos un teorema más general cuya prueba omitiremos. El lector interesado puede consultar, por ejemplo, [5, §42.2, pág. 92].

Teorema 3.6.9 (del rango) *Sea el sistema de funciones* (3.6.1). *Asumiremos que las funciones están definidas y son diferenciables en el entorno de cierto* $a \in A \subset \mathbb{R}^n$. *Por sencillez asumiremos que todas las derivadas parciales* $\frac{\partial u_k}{\partial x_l}$, $k = 1, \ldots, m$, $l = 1, 2, \ldots, n$ *son continuas en* a. *Sea* J *la matriz jacobiana del sistema*

$$J = \begin{pmatrix} D_1u_1(a) & D_2u_1(a) & \ldots & D_nu_1(a) \\ \vdots & \vdots & \ddots & \vdots \\ D_1u_m(a) & D_2u_m(a) & \ldots & D_nu_m(a) \end{pmatrix}.$$

Supongamos que en la matriz J *anterior cierto menor de orden* $r \leq m \leq n$ *es distinto de cero en* a[6] *y que cualquier menor de orden* $r+1$ *es nulo en el entorno de* a *(en otras palabras, que el rango de la matriz* J *es* r*). Entonces las* r *funciones correspondientes al menor distinto de cero son funcionalmente independientes en el entorno de* a *y cada una de las demás funciones depende, en dicho entorno, de las* r *funciones independientes.*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

[6] De la continuidad de las derivadas parciales se deduce que dicho menor será distinto de cero en todo un entorno de a.

Capítulo 4

Extremos de funciones de varias variables

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Vamos a estudiar ahora el problema de encontrar los máximos y mínimos de las funciones de varias variables. Cuando queremos estudiar los extremos de una función de varias variables hay dos posibilidades: o bien las variables son independientes, entonces se dice que el extremo es un *extremo libre*, o las variables no son independientes, en cuyo caso se dice que el extremo es un *extremo condicionado*.

Por ejemplo, sea la función $f : \mathbb{R}^2 \mapsto \mathbb{R}$, $f(x, y) = x^2 + y^2$. El extremo de f suponiendo que las variables están definidas sobre la elipse $x^2 + 3y^2 = 1$ sería un ejemplo de un *extremo condicionado*. Si no hay restricciones sobre las variables, es decir si (x, y) es cualquier punto de $\mathbb{R}^2$, entonces tendríamos un ejemplo de un *extremo libre*.

Vamos a estudiar ambos casos por separados. Comenzaremos por los extremos libres.

4.1. Extremos libres

En general cuando f tiene un máximo y mínimo en cierto punto $x = a$ diremos que f tiene un extremo en a.

Definición 4.1.1 *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ definida en cierto subconjunto $A \subset \mathbb{R}^n$ que puede ser abierto o cerrado.*

1. *Si $f(x) \leq f(a)$ (respectivamente $f(x) \geq f(a)$), para todo $x \in A$, $x \neq a$, decimos que f alcanza en el punto a el máximo (respectivamente mínimo) absoluto en A.*

2. *Si existe una bola $B(a,\delta) \subset A$ tal que para todo $x \in B$, $x \neq a$, $f(x) \leq f(a)$ (respectivamente $f(x) \geq f(a)$) decimos que f alcanza en a un máximo (respectivamente mínimo) relativo.*

En el caso de que las desigualdades sean estrictas diremos que los extremos son estrictos.

De la definición anterior se deduce que todo extremo absoluto es un extremo relativo si este se encuentra en el interior de A. No obstante es conveniente tener en cuenta que, en general, los extremos absolutos no tienen porque ser extremos relativos (por ejemplo si el extremo absoluto se alcanza en $x = a$ con a en la frontera del dominio A no tiene por qué existir ninguna bola $B(a,\delta) \subset A$) ni los extremos relativos tienen por que ser absolutos (el extremo absoluto puede alcanzarse en la frontera de A). Para convencerse de ello basta recurrir a ejemplos sencillos de funciones de una variable (que se dejan como ejercicio al lector).

Teorema 4.1.2 (Condición necesaria de extremo relativo) *Sea la función $f : A \subset \mathbb{R}^n \to \mathbb{R}$, A abierto, $a \in A$. Supongamos que f tiene en a un extremo relativo. Entonces, si existen las derivadas parciales $\frac{\partial f}{\partial x_k}$, $k = 1, \ldots, n$ éstas son iguales a cero en a, i.e., $\frac{\partial f}{\partial x_k}(a) = 0$, $k = 1, \ldots, n$. En particular si f es diferenciable en a, entonces $Df(a) = 0$.*

Demostración: Sea $a = (a_1, \cdots, a_n) \in A$, A abierto, un extremo de la función $f : A \subset \mathbb{R}^n \to \mathbb{R}$ diferenciable en a. Definamos la función $\phi_k(x_k) := f(a_1, \ldots, a_{k-1}, x_k, a_{k+1}, \ldots, a_n)$. Es obvio que ϕ_k es diferenciable en su única variable y tiene un extremo en $x_k = a_k$, luego el Lema de Fermat para funciones escalares de una variable 1.1.18 nos asegura que $\phi_k'(a_k) = 0$, de donde se sigue el resultado. □

Definición 4.1.3 *Sea una función $f : A \subset \mathbb{R}^n \to \mathbb{R}$, A abierto que admite todas sus derivadas parciales en A. Sea $a \in A$ tal que $\frac{\partial f}{\partial x_k}(a) = 0$, $k = 1, \ldots, n$. Un punto a que cumple lo anterior se denomina punto crítico de f. Nótese que si f es diferenciable en un punto crítico a, entonces su diferencial $Df(a) = 0$.*

Ejemplo 4.1.4 Sea la función $f : \mathbb{R}^2 \mapsto \mathbb{R}$, $f(x,y) = -x^2 - y^2$. Es *obvio* que tiene un máximo en $(0,0)$. Un cálculo directo muestra que $\frac{\partial f}{\partial x}(0,0) = 0$, $\frac{\partial f}{\partial y}(0,0) = 0$. □

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Ejemplo 4.1.5 Sea ahora la función $f : \mathbb{R}^2 \mapsto \mathbb{R}$, $f(x, y) = x^2 + y^2$. Está claro que esta función tiene un mínimo en $(0, 0)$ y que también para ella $\frac{\partial f}{\partial x}(0, 0) = 0$, $\frac{\partial f}{\partial y}(0, 0) = 0$. □

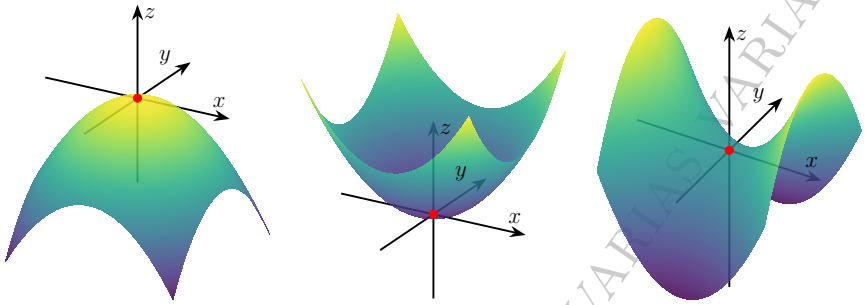


Figura 4.1.1: De izquierda a derecha se representan funciones $f : \mathbb{R}^2 \mapsto \mathbb{R}$ con: un máximo local (Ejemplo 4.1.4), un mínimo local (Ejemplo 4.1.5) y un punto silla (Ejemplo 4.1.6), respectivamente.

Ejemplo 4.1.6 Sea la función $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}$, $f(x, y) = x^2 - y^2$. Es obvio que $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$. Sin embargo, en cualquier entorno de $(0, 0)$ que escojamos f toma valores tanto positivos como negativos. En este caso el punto $(0, 0)$ se denomina *punto silla* de f. □

¿Cómo saber si un punto crítico es un extremo local o un punto silla? Para ello tenemos un teorema similar al del caso de una variable. Antes de enunciarlo conviene recordar que la segunda diferencial de una función de varias variables $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, $f \in C^{(2)}(A)$ es la forma bilineal simétrica (2.6.5) que escribiremos, convenientemente, de la forma

$$d^2 f(a) := D^2 f(a)(x) = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \frac{\partial^2 f(a)}{\partial x_{i_1} \partial x_{i_2}} x_{i_1} x_{i_2} = x^T H_f(a) x, \tag{4.1.1}$$

donde $H_f(a)$ es la matriz hessiana (2.5.1) y $x = (x_1, x_2, \ldots, x_n)^T$.

Teorema 4.1.7 (Condición suficiente de extremo relativo) *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ dos veces diferenciable (i.e., $f \in C^{(2)}(A)$) en $a \in A$, A abierto, y sea $x = a$ un punto crítico de f, i.e., $Df(a) = 0$. Entonces*

1. *Si la segunda diferencial $D^2 f(a)(x)$ es definida positiva en a, entonces f tiene un mínimo relativo en a.*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

2. *Si la segunda diferencial $D^2f(a)(x)$ es definida negativa, entonces f tiene un máximo relativo en a.*

3. *Si la segunda diferencial $D^2f(a)(x)$ es indefinida, i.e., si existen $x, y \in \mathbb{R}^n$ tales que $D^2f(a)(x) < 0 < D^2f(a)(y)$, entonces f tiene un punto de silla en a.*

Demostración: Usamos la fórmula local de Taylor (2.6.2) con $x = a+h$ teniendo en cuenta que $Df(a) = 0$ (se asume que $h \neq 0$)

$$f(x) - f(a) = \frac{1}{2}D^2f(a)(h) + o(\|h\|^2) = \frac{1}{2}\|h\|^2\left(\sum_{i,j=1}^{n}\frac{\partial^2 f(a)}{\partial x_i \partial x_j}\frac{h_i}{\|h\|}\frac{h_j}{\|h\|} + o(1)\right),$$

donde $o(1)$ es tal que $\lim_{x\to a} o(1) = \lim_{h\to 0} o(1) = 0$. De la expresión anterior se sigue que el signo de $f(x) - f(a)$ coincide con el signo de la expresión dentro de los paréntesis. Denotemos por $Q(h', a)$ la forma cuadrática

$$Q(h', a) = \sum_{i,j=1}^{n}\frac{\partial^2 f(a)}{\partial x_i \partial x_j}\frac{h_i}{\|h\|}\frac{h_j}{\|h\|}, \qquad h' = \frac{h}{\|h\|}, \quad \|h'\| = 1.$$

Nótese que Q está definida sobre el vector $h' = h/\|h\|$ con norma uno, i.e., Q es una función continua (polinomio) sobre un cerrado y acotado (la esfera $S = \{x \in \mathbb{R}^n | \|x-a\| = 1\}$), luego, por el Teorema de Weierstrass 2.1.2, Q alcanza su máximo y mínimo absolutos, que denotaremos por M y m respectivamente. Luego para todo $h' \in S$ tenemos $m \leq Q(h', a) \leq M$. Supongamos ahora que Q es definida positiva, entonces $Q(h', a) > 0$ para todo h', y por tanto $m > 0$. Ahora bien como $\lim_{h\to 0} o(1) = 0$, eso indica que para todo $\epsilon > 0$ existe un $\delta > 0$ tal que si $\|x - a\| = \|h\| < \delta$ entonces $|o(1)| < \epsilon$. Si tomamos $\epsilon = m/2$ entonces existe un entorno de $x = a$ tal que $Q(h', a) + o(1) > 0$, y por tanto en dicho entorno de a $f(x) - f(a) > 0$, i.e., f tiene un mínimo local. Si por el contrario Q es definida negativa entonces $M < 0$. Razonando igual que antes y tomando $\epsilon < |M|/2$, tenemos que $f(x) - f(a) < 0$, i.e., f tiene un máximo. Lo anterior prueba 1 y 2.

Probemos ahora 3. Sean a_m y a_M los puntos donde Q alcanza su mínimo y máximo absolutos sobre S (nótese que $\|a_m\| = \|a_M\| = 1$), respectivamente. Como Q es indefinida entonces $m < 0 < M$ (¿por qué?). Por tanto tenemos $0 > m \leq Q(h', a) \leq M > 0$. Si elegimos ahora $h = x - a = ta_m$ con t suficientemente pequeño tendremos (recordemos que $\|a_m\| = 1$)

$$\begin{aligned} f(x) - f(a) =& \frac{1}{2}D^2f(a)(h) + o(\|h\|^2) = \frac{1}{2}D^2f(a)(a_m t) + o(\|a_m t\|^2) \\ =& \frac{t^2}{2}(m + o(1)), \end{aligned}$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

y para ciertos x en un entorno de a el signo de $f(x) - f(a)$ es igual al signo de m, o sea, negativo. Repitiendo el razonamiento para $h = x - a = ta_M$ con t suficientemente pequeño obtendremos que para ciertos x en un entorno de a $f(x) - f(a) > 0$, luego en a no puede haber ningún extremo de f. □

Nota 4.1.8 *Conviene hacer notar que las condiciones del Teorema 4.1.7 son suficientes. Por ejemplo, para la función definida en todo $\mathbb{R}^2$, $f(x,y) = x^4 + y^4$, tanto la diferencial $Df(0,0)$ como la segunda diferencial $D^2f(0,0)$ son nulas pero claramente la función tiene un mínimo relativo (de hecho absoluto) en $(0,0)$. Análogamente, si la forma bilineal (cuadrática) $D^2f(a)(h)$ es semidefinida positiva (o negativa), es decir si para $h \neq 0$, $D^2f(a)(h) \geq 0$ (≤ 0), el teorema no asegura la existencia de extremos.*

Una pregunta natural es cuándo la forma bilineal $D^2f(a)(x)$ es definida positiva, negativa o indefinida. Ello nos los da el conocido Criterio de Sylvester del álgebra lineal.

Criterio 4.1.9 (de Sylvester) *Sea $B(x,y)$ una aplicación bilineal simétrica y sea $B = [b_{i,j}]_{i,j=1,n}$ su matriz. Las siguientes condiciones son equivalentes:*

1. *B es definida positiva.*
2. *Todos los autovalores de B son positivos.*
3. *Los menores principales Δ_k de B son positivos, i.e. $\Delta_k > 0$, $k = 1, 2, \ldots, n$ donde*

$$\Delta_k := \det \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,k} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,k} \\ & & & \\ \vdots & \ddots & \vdots & \\ b_{k,1} & b_{k,2} & \cdots & b_{k,k} \end{pmatrix}, \quad \forall k = 1, 2, \ldots n. \tag{4.1.2}$$

Análogamente se tiene para las formas bilineales definidas negativas las siguientes condiciones equivalentes:

1. *B es definida negativa.*
2. *Todos los autovalores de B son negativos.*
3. *Los menores principales Δ_k de B son tales que $(-1)^k\Delta_k > 0$, $k = 1, 2, \ldots, n$.*

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

El criterio anterior junto al Teorema 4.1.7 nos conduce al siguiente:

Corolario 4.1.10 (Condición suficiente de extremo) *Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ dos veces diferenciable en $a \in A$, A abierto, y sea $x = a$ un punto crítico de f, i.e., $Df(a) = 0$ y sean los determinantes*

$$\Delta_k(a) := \det \begin{pmatrix} \dfrac{\partial^2 f(a)}{\partial^2 x_1} & \cdots & \dfrac{\partial^2 f(a)}{\partial x_k \partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(a)}{\partial x_1 \partial x_k} & \cdots & \dfrac{\partial^2 f(a)}{\partial^2 x_k} \end{pmatrix}. \tag{4.1.3}$$

1. *Si todos los menores principales $\Delta_k > 0$, $k = 1, 2, \ldots, n$, entonces f tiene un mínimo relativo en a.*
2. *Si todos los menores principales son tales que $(-1)^k \Delta_k > 0$, $k = 1, 2, \ldots, n$, entonces f tiene un máximo relativo en a.*

En el caso especial de dos variables se puede ir más allá:

Corolario 4.1.11 *Sea $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ dos veces diferenciable en $a \in A$, A abierto, $Df(a) = 0$, y sea $\Delta_2(a)$ el determinante*

$$\Delta_2(a) = \det \begin{pmatrix} \dfrac{\partial^2 f(a)}{\partial^2 x_1} & \dfrac{\partial^2 f(a)}{\partial x_2 \partial x_1} \\ \dfrac{\partial^2 f(a)}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f(a)}{\partial^2 x_2} \end{pmatrix}.$$

1. *Si $\frac{\partial^2 f(a)}{\partial^2 x_1} > 0$ y $\Delta_2(a) > 0$, f tiene un mínimo relativo en a.*
2. *Si $\frac{\partial^2 f(a)}{\partial^2 x_1} < 0$ y $\Delta_2(a) > 0$, f tiene un máximo relativo en a.*
3. *Si $\Delta_2(a) < 0$, f tiene un punto de silla en a.*
4. *Si $\Delta_2(a) = 0$, nada puede decirse.*

Ejemplo 4.1.12 Consideremos la función $f : \mathbb{R}^2 \mapsto \mathbb{R}$, $f(x, y) = x^4 + y^4 - (x + y)^2 + 2$ y estudiemos si tiene extremos en $\mathbb{R}^2$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Calculamos sus derivadas parciales y las igualamos cero

$$\frac{\partial f(x,y)}{\partial x} = 4\,x^3 - 2\,(y+x) = 0, \quad \frac{\partial f(x,y)}{\partial y} = 4\,y^3 - 2\,(y+x) = 0.$$

Las soluciones reales son tres: $A(-1,-1)$, $B(1,1)$ y $C(0,0)$. Para determinar si son extremos calculamos la matriz Hessiana:

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

y la evaluamos en los distintos puntos. En el caso de los puntos A y B obtenemos

$$H_f(A) = H_f(B) = \begin{pmatrix} 10 & -2 \\ -2 & 10 \end{pmatrix},$$

que es definida positiva por el Corolario 4.1.11 y por tanto en ambos casos tenemos un mínimo local. Veamos que ocurren en el tercer punto $C(0,0)$. En este caso

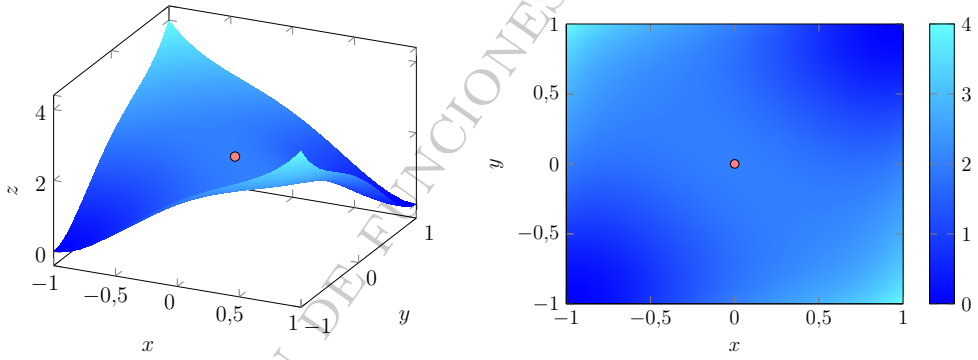


Figura 4.1.2: Representación de la función $f(x,y) = x^4 + y^4 - (x+y)^2 + 2$ en un entorno del $(0,0)$ (izquierda) y su dibujo de contorno (derecha).

tenemos

$$H_f(C) = \begin{pmatrix} -2 & -2 \\ -2 & -2 \end{pmatrix},$$

luego $\det H_f(C) = 0$ y según el Corolario 4.1.11 no podemos decir nada. Ahora bien, en este caso podemos ver que $f(x,x)$ tiene un máximo en $(0,0)$ mientras que $f(x,-x)$ tiene un mínimo, o sea, tenemos un punto silla como se puede ver el la figura 4.1.2. □

Ejemplo 4.1.13 Consideremos ahora la función $f : \mathbb{R}^2 \mapsto \mathbb{R}$, $f(x,y) = x^4 + y^4 + (x+y)^2 + 2$ y estudiemos si tiene extremos en $\mathbb{R}^2$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Calculamos sus derivadas parciales y las igualamos a cero

$$\frac{\partial f(x,y)}{\partial x} = 4\,x^3 + 2\,(y+x) = 0, \quad \frac{\partial f(x,y)}{\partial y} = 4\,y^3 + 2\,(y+x) = 0.$$

En este caso solo hay una solución real: $A(0,0)$. Para dicha solución la matriz

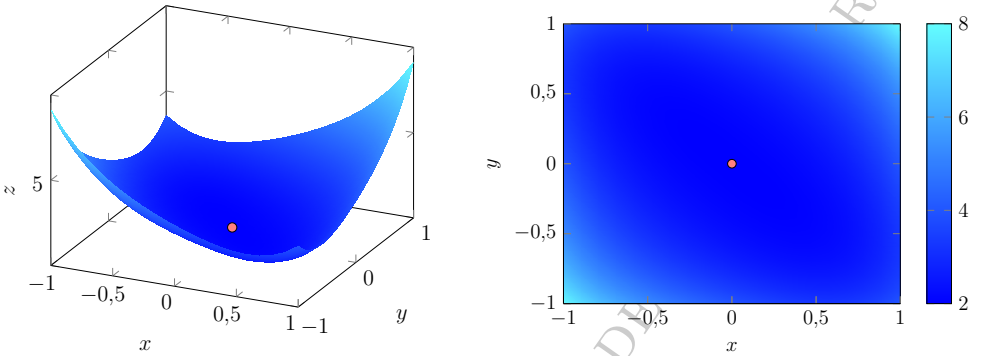


Figura 4.1.3: Representación de la función $f(x,y) = x^4 + y^4 + (x+y)^2 + 2$ en un entorno del $(0,0)$ (izquierda) y su dibujo de contorno (derecha).

Hessiana

$$H_f(x,y) = \begin{pmatrix} 12\,x^2 + 2 & 2 \\ 2 & 12\,y^2 + 2 \end{pmatrix},$$

toma el valor

$$H_f(C) = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix},$$

luego $\det H_f(C) = 0$ y según el Corolario 4.1.11 no podemos decir nada. Ahora bien, en este caso es sencillo ver que $f(x,y)$ tiene un mínimo en $(0,0)$ como se puede ver el la figura 4.1.3 pues $f(x,y) \geq 2$ para todos $(x,y) \in \mathbb{R}^2$. □

4.1.1. Problemas propuestos

En todos los casos decide si hay extremos absolutos y/o relativos así como puntos silla para las siguientes funciones $f : \mathbb{R}^2 \mapsto \mathbb{R}$

1. $f(x,y) = x^4 + y^4 + 6x^2y^2 + 8x^3$.
2. $f(x,y) = x^4 + y^4 \pm (x+y)^2 + 2$.
3. $f(x,y,z) = x^4 + y^4 + z^4 - 2(x^2+z^2) + y^2$.

DIFERENCIACIÓN DE FUNCIONES DE VARIABLES

4. $f(x,y) = xy\log(x^2+y^2)$ si $(x,y) \neq (0,0)$ y 0 si $(x,y) = (0,0)$.
5. $f(x,y) = (x \pm y)\exp(-x^2-y^2)$.
6. $f(x,y) = xy\exp(-x^2-y^2)$ y $f(x,y) = xy\exp(x^2+y^2)$.
7. $f(x,y) = x^3 - 9xy + y^3 + 27$.

4.2. Extremos condicionados

Pasemos ahora a considerar un problema muy relacionado con el anterior. Imaginemos que queremos encontrar los extremos de una función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ donde las variables no son todas independientes sino que han de satisfacer una serie de condiciones de *ligadura* $\Phi_k(x_1,\ldots,x_n) = 0$, $k = 1,\ldots,m$, $m < n$. Este problema es un problema natural cuando queremos encontrar los extremos de f sobre una determinada curva o trayectoria, o si queremos encontrar distancias máximas o mínimas entre superficies geométricas o entre un punto y una superficie geométrica, etc. y se conoce como el problema de *extremos condicionados*.

Veamos un ejemplo muy sencillo para aclarar ideas: Queremos encontrar el máximo y/o mínimo absolutos de cierta función $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}$, si sus variables satisfacen la ecuación $\Phi(x,y) = 0$. Vamos a suponer que tanto f como Φ son funciones lo suficientemente *buenas* (por ejemplo $C^{(2)}(A)$). Una forma de resolver el problema es como sigue: resolvemos la ecuación $\Phi(x,y) = 0$ respecto a una variable, digamos $y = g(x)$, y sustituimos la función resultante en nuestra f. Así obtenemos una función de una variable $F(x) = f(x,g(x))$ a la que podemos calcularle los extremos al ser x una variable *libre*. Nótese que, por el Teorema de la función implícita bastaría que $\Phi'_y(x,y) \neq 0$ en A para tener garantizado que exista la función $y = g(x)$. Esta idea aunque resuelve el problema, al menos formalmente, no es muy buena pues no siempre es posible encontrar explícitamente la función g aún sabiendo que ello es posible. Por otro lado hay una clara asimetría entre las variables siendo una dependiente de la otra. Mostremos, con este mismo ejemplo, una forma más elegante de proceder.

En la figura 4.2.1 representamos en rojo las curvas de nivel de la función f, i.e., las curvas que define la ecuación $f(x,y) = c$. Así mismo en negro representamos la curva que define la expresión $\Phi(x,y) = 0$. Imaginemos que recorremos la curva $\Phi(x,y) = 0$ en contra de las manecillas del reloj tal y como muestra la flecha de la figura 4.2.1. A medida que avanzamos en la curva Φ, esta va cortando las curvas de nivel de f. Supongamos que sabemos que f tiene un extremo a lo largo de la curva Φ. Entonces a lo largo de recorrido los valores de c irán aumentando o disminuyendo hasta que alcancemos el punto $P = (x_0,y_0)$ donde

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

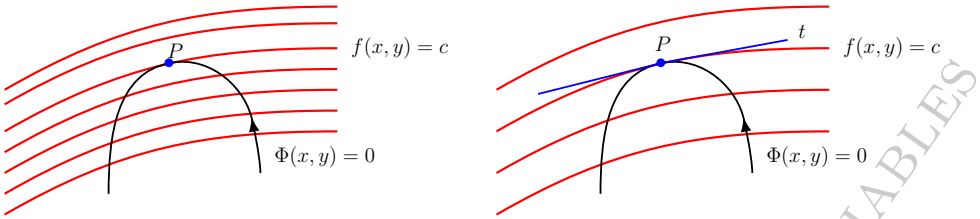


Figura 4.2.1: Curvas de nivel de $f(x,y)$ (en rojo) y curva $\Phi(x,y)=0$ (negro). En la figura de la derecha además está representada la recta tangente t (azul) en el punto $P=(x_0,y_0)$.

cambiará la tendencia (si c aumentaba, ahora disminuirá, o viceversa). Está claro que el punto P donde ocurre el cambio de la monotonía de c es un extremo de f. Denotemos dicho c por c_P. Si suponemos además que tanto la curva $f(x,y)=c_P$ como $\Phi(x,y)=0$ son *suaves* (f y Φ son funciones $C^{(1)}(U_P)$ en un entorno U_P de P) entonces ambas tienen la misma recta tangente en P. La pendiente de dicha tangente se calcula, en general, por la fórmula $-f'_x(x_0,y_0)/f'_y(x_0,y_0)$ o bien $-\Phi'_x(x_0,y_0)/\Phi'_y(x_0,y_0)$, donde vamos a asumir por simplicidad que todas las derivadas son distintas de cero. Lo anterior nos conduce a

$$\frac{f'_x(x_0,y_0)}{f'_y(x_0,y_0)}=\frac{\Phi'_x(x_0,y_0)}{\Phi'_y(x_0,y_0)}\quad\Leftrightarrow\frac{f'_x(x_0,y_0)}{\Phi'_x(x_0,y_0)}=\frac{f'_y(x_0,y_0)}{\Phi'_y(x_0,y_0)}=-\lambda.$$

Es decir, que si en P hay un extremo de f cuando nos restringimos a la curva $\Phi(x,y)=0$, entonces han de cumplirse las siguientes condiciones:

$$\begin{cases} f'_x(x_0,y_0)+\lambda\Phi'_x(x_0,y_0)=0,\\ f'_y(x_0,y_0)+\lambda\Phi'_y(x_0,y_0)=0,\\ \Phi(x,y)=0,\end{cases}$$

donde λ es cierta constante. O sea, P ha de ser un punto crítico de la función F de tres variables $L(x,y,\lambda)=f(x,y)+\lambda\Phi(x,y)$.

La función L anterior se suele denominar función de Lagrange y la forma de encontrar el extremo según el sistema anterior es conocido como el método de los coeficientes indeterminados de Lagrange. Nótese que lo anterior sólo nos da condiciones necesarias. Si queremos una condición suficiente tenemos que calcular el segundo diferencial de f en el punto crítico y luego usar la identidad $d\Phi(x,y)=\Phi'_x(x_0,y_0)dx+\Phi'_y(x_0,y_0)dy=0$ que relaciona los diferenciales de las dos variables. Sustituyendo esta última relación en la expresión (4.1.1) de $d^2f(x_0,y_0)$ obtendremos una forma cuadrática (en este caso de una única variable) cuyo signo determinará el tipo de extremo.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

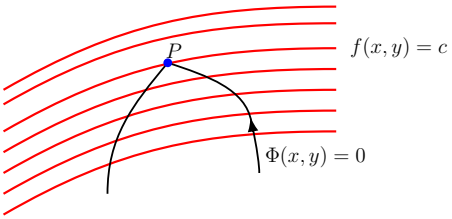


Figura 4.2.2: Curvas de nivel de $f(x,y)$ (en rojo) y curva $\Phi(x,y)=0$ (negro). En el punto P donde se alcanza el extremo $\Phi(x,y)$ no es diferenciable (tiene un pico).

Antes de continuar conviene observar que el método anterior falla si la curva Φ tiene picos pues puede ocurrir que el extremo se alcance justo en ese punto tal y como se muestra en la figura 4.2.2.

Pasemos a enunciar el problema general: Sean las funcione $\Phi_k : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, $k=1,\dots,m$, dos veces diferenciables y sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ una función dos veces diferenciable $f(x_1,x_2,\dots,x_m,x_{m+1},\dots,x_n)$ cuyas n variables satisfacen las ecuaciones

$$\text{(4.2.1)}\qquad \begin{cases} \Phi_1(x_1,x_2,\cdots,x_n)=0, \\ \Phi_2(x_1,x_2,\cdots,x_n)=0, \\ \qquad\qquad\vdots \\ \Phi_m(x_1,x_2,\cdots,x_n)=0, \end{cases}$$

i.e., que las variables no son independientes. Las ecuaciones anteriores se suelen denominar *ecuaciones de ligadura*. Por simplicidad supondremos que todas las ecuaciones de ligadura son independientes, o sea, ninguna de las Φ_k se puede escribir a partir de las demás (véase el apartado §3.6) y que están bien definidas en A. Sea $x_0 \in A$ tal que sea solución de (4.2.1). Supongamos también que el siguiente jacobiano es no nulo en todo un entorno[1] de x_0

$$\text{(4.2.2)}\qquad \det J_\Phi := \det \begin{pmatrix} \dfrac{\partial \Phi_1(x)}{\partial x_1} & \dfrac{\partial \Phi_1(x)}{\partial x_2} & \cdots & \dfrac{\partial \Phi_1(x)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial \Phi_m(x)}{\partial x_1} & \dfrac{\partial \Phi_m(x)}{\partial x_2} & \cdots & \dfrac{\partial \Phi_m(x)}{\partial x_m} \end{pmatrix} \neq 0.$$

Bajo las condiciones anteriores se tiene el siguiente teorema:

[1] Según el Teorema del rango 3.6.9 esta condición es suficiente para la independencia funcional de las ecuaciones de ligadura (4.2.1).

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Teorema 4.2.1 *Sea la función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ una función de clase $C^{(1)}(A)$ cuyas n variables satisfacen las ecuaciones de ligadura (4.2.1) y sea $x_0 \in A$ solución de (4.2.1) un extremo de f. Dicho extremo se suele denominar extremo condicionado de f por las ecuaciones de ligadura (4.2.1). Entonces existen m constantes λ_1, λ_2, ..., λ_m reales tales que la función $L : \mathbb{R}^{n+m} \mapsto \mathbb{R}$, que se denomina función de Lagrange,*

$$L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = f(x_1, \ldots, x_n) + \lambda_1 \Phi_1(x_1, \ldots, x_n) + \cdots + \lambda_m \Phi_m(x_1, \ldots, x_n) \tag{4.2.3}$$

tiene un punto crítico en x_0.

Está claro que el teorema anterior nos da condiciones necesarias pero no suficientes de que $x = a$ sea un extremo condicionado de f. Entonces ¿cómo encontrar condiciones suficientes?

Para ello, nótese que como el sistema (4.2.1) tiene solución en el punto $a \in A$, y el jacobiano $\det J_\Phi$ (4.2.2) es distinto de cero, entonces por el Teorema de la función implícita el sistema (4.2.1) es resoluble en las variables $x_1, \ldots, x_m$, es decir, en un entorno de $a \in A$ existen las funciones $x_k = g_k(x_{m+1}, \ldots, x_n)$, $k = 1, 2, \ldots, m$ tales que

$$\Phi_k(g_1(x_{m+1}, \ldots, x_n), \ldots, g_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n) = 0$$

son identidades en el entorno de $a \in A$. O sea, en las condiciones dadas el problema del cálculo de un extremo condicionado se puede transformar en el de un extremo libre (sin ecuaciones de ligadura) sustituyendo las funciones x_k, $k = 1, \ldots, m$ así obtenidas en la expresión de f, i.e, encontrando los extremos de la función

$$F(x_{m+1}, \ldots, x_n) := f(g_1(x_{m+1}, \ldots, x_n), \ldots, g_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n).$$

Para decidir si efectivamente tenemos en el punto crítico $x0a$ un extremo habría que aplicar el Teorema 4.1.7 a la función F anterior lo cual es complicado (si no imposible) en la mayoría de los casos pues, en general, no es posible obtener la solución analítica del sistema (4.2.1).

Ahora bien, como los extremos de f con las ligaduras (4.2.1) son los mismos que los de la función de Lagrange L (4.2.3) por tanto la idea es encontrar los puntos críticos de L a partir de sus derivadas parciales, donde ahora las constantes indeterminadas λ_k, $k = 1, \ldots, m$ se consideran variables independientes. Eso nos conduce a un sistema de $n + m$ ecuaciones, a saber

$$\frac{\partial L}{\partial x_k} = 0, \quad k = 1, \ldots, n, \qquad \frac{\partial L}{\partial \lambda_i} = 0, \quad i = 1, \ldots, m.$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Nótese que las ecuaciones $\frac{\partial L}{\partial \lambda_i} = 0$ se transforman en las ecuaciones de ligadura, que sabemos de antemano que han de cumplirse. Este sistema nos proporciona cierta cantidad de puntos críticos. Supongamos que $a = (x_0, \lambda_0) \in \mathbb{R}^{n+m}$ es uno de dichos puntos críticos. Para saber si dicho punto crítico es un extremo hemos de calcular la segunda diferencial de L en dicho punto (ver (4.1.1))

$$d^2L(a) = \sum_{i,j=1}^{n} \frac{\partial^2 L(a)}{\partial x_i \partial x_j} dx_i dx_j + \sum_{i=1}^{n} \frac{\partial^2 L(a)}{\partial \lambda_i \partial \lambda_j} d\lambda_i d\lambda_j + \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{\partial^2 L(a)}{\partial x_i \partial \lambda_j} dx_i d\lambda_j.$$

Dado que $\frac{\partial L}{\partial \lambda_i} = \Phi_i(x_1, \cdots, x_n) = 0$ es una identidad, entonces todas las derivadas de orden dos se anulan, i.e., $\frac{\partial^2 L(a)}{\partial^2 \lambda_i} = \frac{\partial^2 L(a)}{\partial x_i \partial \lambda_j} = 0$, por lo que se tiene que

$$d^2L(a) = d^2 f(x_0)\Big|_{\Phi_i(x_0)=0,\ i=1,\ldots,m}. \tag{4.2.4}$$

Lo anterior nos dice que calcular la segunda diferencial de L en $a = (x_0, \lambda_0)$ es equivalente a calcular la segunda diferencial de f en x_0, pero teniendo en cuenta que las diferenciales dx_k de las variables x_k, $k = 1, \ldots, n$ no son independientes. Veamos que efectivamente esto es posible.

Tomando diferenciales en ambos lados de (4.2.1) obtenemos el siguiente sistema lineal de ecuaciones respecto a las variables $dx_1, \ldots, dx_m$

$$d\Phi_i(x_0) = \sum_{k=1}^{n} \frac{\partial \Phi_i}{\partial x_k}(x_0) dx_k = 0, \quad k = 1, \ldots, m. \tag{4.2.5}$$

El determinante del sistema anterior no es más que el jacobiano (4.2.2) que es distinto de cero en un entorno del punto x_0, por lo tanto, existen ciertas funciones lineales $g_j : \mathbb{R}^{n-m} \mapsto \mathbb{R}$, $j = 1, \ldots, m$ tales que $dx_j = g_j(dx_{m+1}, \ldots, dx_n) = A_{j,1}dx_{m+1} + \cdots + A_{j,n-m}dx_n$, $j = 1, \ldots, m$. Es decir podemos encontrar explícitamente los valores de las diferenciales $dx_1, \ldots dx_m$ de las variables $x_1, \ldots, x_m$ como función de $dx_{m+1}, \ldots, dx_n$. Sustituyendo los valores de las diferenciales $dx_1, \ldots, dx_m$ en la expresión de la segunda diferencial (4.2.4) obtenemos la expresión de la segunda diferencial de L en las variables independientes. Estudiando el signo de dicha forma cuadrática resultante tal y como se indica en el Teorema 4.1.10 podremos decidir si el punto x_0 es un extremo o no de f bajo las condiciones de ligadura (4.2.1).

Mostremos como funciona este método con unos ejemplos:

Ejemplo 4.2.2 Encontrar los extremos de la función $f(x, y) = x^2 + y^2$ con la condición de ligadura $(x-3)^2 + (y-4)^2 = 10^2$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Nótese que en este caso tenemos que la condición de ligadura es una circunferencia S en $\mathbb{R}^2$. Como toda circunferencia es un conjunto compacto y f es continua, entonces el Teorema de Weierstrass 2.1.2 establece que f alcanza en S su máximo y mínimo absolutos. Entonces, en dichos puntos, por la condición necesaria de extremo $df(x) = 0$. Para resolver el problema escribimos la función de Lagrange:

$$L(x, y, \lambda) = x^2 + y^2 + \lambda((x-3)^2 + (y-4)^2 - 10^2),$$

y calculamos sus puntos críticos:

$$\frac{\partial L}{\partial x} = 2x(\lambda+1) - 6\lambda = 0, \quad \frac{\partial L}{\partial y} = 2y(\lambda+1) - 8\lambda = 0,$$

$$\frac{\partial L}{\partial \lambda} = (x-3)^2 + (y-4)^2 - 10^2 = 0.$$

La resolución del sistema nos conduce a dos puntos críticos:

1. $\lambda = -1/2$, $A = (-3, -4)$,
2. $\lambda = -3/2$, $B = (9, 12)$.

Dado que sólo tenemos dos, y $f(-3,-4) = 25$ y $f(9,12) = 225$, entonces el primero ha de ser un mínimo y el segundo un máximo. Comprobémoslo calculando la segunda diferencial:

$$d^2L = 2(\lambda+1)dx^2 + 2(\lambda+1)dy^2.$$

De la ecuación de ligadura obtenemos $(2x-6)dx + (2y-8)dy = 0$ que evaluada en ambos puntos nos conduce a la siguiente relación entre los diferenciales dx y dy: $dy = -3/4dx$ que al sustituirlo en la expresión anterior para d^2L nos conduce a $d^2L = \frac{13}{2}(\lambda+1)dx^2$. Así, efectivamente, en el punto A tenemos $d^2L > 0$ luego hay un mínimo y en el B como $d^2L < 0$ tenemos un máximo. □

Ejemplo 4.2.3 Encontrar el máximo y mínimo absolutos de la función $f(x, y) = x^2 + y^2 - x - y + 1$ en la región S definida por $x^2 + y^2 \leq 1$.

Ante todo notemos que como S es cerrada y acotada, el Teorema de Weierstrass 2.1.2 nos asegura que han de existir los extremos absolutos. Vamos a encontrarlos.

Está claro que para resolver este problema hay que separarlo en dos problemas complementarios e independientes. El primero es uno de extremos libres sobre en interior del círculo $x^2 + y^2 < 1$ y el otro, un problema de extremos condicionados sobre la frontera $x^2 + y^2 = 1$. Como ya hemos dicho, dado que la región donde está

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

definida f es un compacto, sabemos que f debe alcanzar su máximo y mínimo absolutos.

Comenzamos con el problema de extremos libres. La condición necesaria de extremo nos da

$$\frac{\partial f}{\partial x} = 2x - 1 = 0, \quad \frac{\partial f}{\partial y} = 2y - 1 = 0$$

de donde obtenemos un único punto crítico $P = (1/2, 1/2)$ que además está en el interior del círculo $x^2 + y^2 < 1$. Como

$$D^2 f(1/2, 1/2) = 2(dx^2 + dy^2) > 0,$$

el punto P es un mínimo local. Además $f(1/2, 1/2) = 1/2$.

Pasemos a ver que ocurre en la frontera ∂S. Para ello escribimos la función de Lagrange:

$$L(x, y, \lambda) = x^2 + y^2 - x - y + 1 + \lambda(x^2 + y^2 - 1),$$

y calculamos sus puntos críticos:

$$\frac{\partial L}{\partial x} = 2x(\lambda + 1) - 1 = 0, \quad \frac{\partial L}{\partial y} = 2y(\lambda + 1) - 1 = 0, \quad \frac{\partial L}{\partial \lambda} = x^2 + y^2 - 1 = 0.$$

Está claro que $(0, 0)$ queda excluido, así como el valor $\lambda = -1$. De las dos primeras ecuaciones obtenemos x e y en función de λ y sustituyendo el resultado nos quedan los puntos:

1. $\lambda = -1 + \sqrt{2}/2$, $A = (\sqrt{2}/2, \sqrt{2}/2)$,
2. $\lambda = -1 - \sqrt{2}/2$, $B = (-\sqrt{2}/2, -\sqrt{2}/2)$.

Nótese que la frontera del dominio es una circunferencia $x^2 + y^2 = 1$, es decir, un conjunto cerrado y acotado (compacto), y como f es continua ha de alcanzar en dicho conjunto sus mínimo y máximo absolutos. Como $f(-\sqrt{2}/2, -\sqrt{2}/2) = \sqrt{2} + 2$ y $f(\sqrt{2}/2, \sqrt{2}/2) = 2 - \sqrt{2}$, entonces necesariamente en A alcanzara su un mínimo absoluto (siempre sobre la circunferencia) y en B, el máximo absoluto.

Vamos a comprobar lo anterior usando la segunda diferencial. De la ecuación de ligadura se sigue que $xdx + ydy = 0$, que en los puntos A y B nos conducen a la misma relación entre los diferenciales $dy = -dx$. La segunda diferencial de L se escribe, por tanto,

$$d^2 L = 2(\lambda + 1)dx^2 + 2(\lambda + 1)dy^2 = 4(\lambda + 1)dx^2,$$

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

que es positiva en A y negativa en B, luego, si nos restringimos a la circunferencia, efectivamente en $A = (\sqrt{2}/2, \sqrt{2}/2)$ hay un mínimo local y en $B = (-\sqrt{2}/2, -\sqrt{2}/2)$ un máximo local, que al ser los únicos extremos se convierten en extremos absolutos sobre la circunferencia. Si además tenemos en cuenta que en el interior del dominio tenemos un mínimo local que se alcanza en P y $f(1/2, 1/2) = 1/2$ y que en A, $f(\sqrt{2}/2, \sqrt{2}/2) = 2 - \sqrt{2} < 1/2$, entonces en el punto $B(-\sqrt{2}/2, -\sqrt{2}/2)$ se ha de alcanzar, necesariamente, el máximo absoluto de f y en $P = (1/2, 1/2)$ el mínimo absoluto. En la figura 4.2.3 (gráfica izquierda) están representados los puntos P (negro), A (rojo) y B (azul).

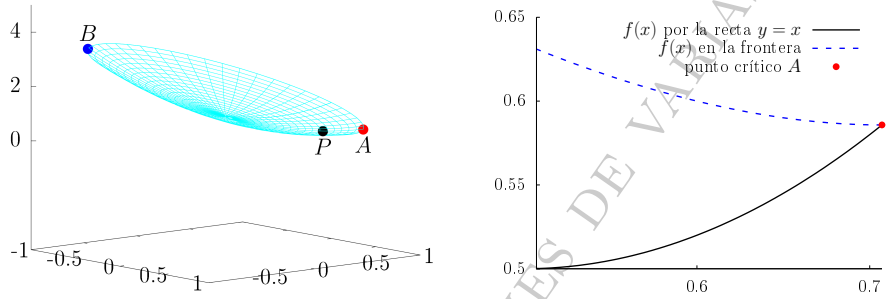


Figura 4.2.3: Gráfica de la función $f(x, y) = x^2+y^2-x-y+1$ en la región definida por $x^2 + y^2 \leq 1$ (izquierda) y valores de la función $f(x, y) = x^2 + y^2 - x - y + 1$ en el entorno del punto A tomando distintas direcciones.

Para decidir si en el punto $A = (\sqrt{2}/2, \sqrt{2}/2)$ hay realmente un mínimo local sobre todo el conjunto S (y no sólo sobre la frontera de S) hay que estudiar que ocurre localmente alrededor de dicho punto. Por un lado sabemos que si nos acercamos a A por la frontera dicho punto es mínimo, es decir, el valor más pequeño de f en la frontera es $f(\sqrt{2}/2, \sqrt{2}/2) = 2 - \sqrt{2}$. Dichos valores están representados en la figura 4.2.3 (gráfica de la derecha) mediante la linea discontinua azul. Vamos ahora a acercarnos por alguna trayectoria del interior del recinto, por ejemplo por la recta $y = x$ con $x < \sqrt{2}/2$.

Es fácil comprobar que en este caso $f(x, x) = 2x^2 - 2x + 1$ y que en el punto $x = \sqrt{2}/2$ es, como función de una variable, creciente, es decir, por dicha trayectoria la $f(x, y) < f(\sqrt{2}/2, \sqrt{2}/2)$, i.e., si f tiene un máximo tal y como se muestra en la figura 4.2.3 (gráfica de la derecha, línea continua negra). El análisis anterior muestra, por tanto, que el punto $A = (\sqrt{2}/2, \sqrt{2}/2)$ no es ningún extremo de nuestra función f.

Para terminar vamos a mostrar en la figura 4.2.4 las dos maneras de acercarnos a A antes mencionadas. En la gráfica de la izquierda se observa que, efectivamente, en $A = (\sqrt{2}/2, \sqrt{2}/2)$ f tiene un mínimo local si nos restringimos a la frontera

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

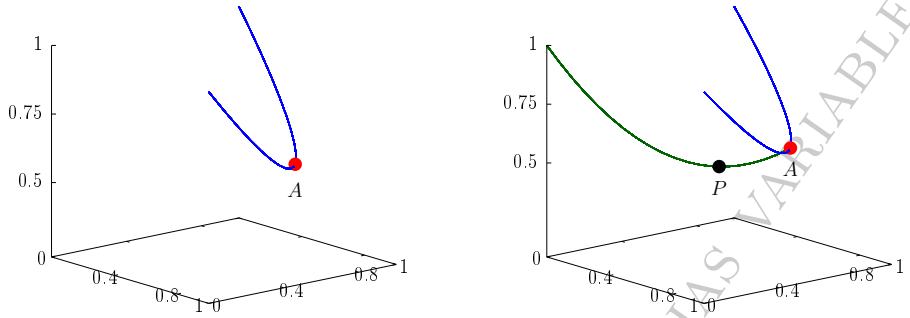


Figura 4.2.4: A la izquierda se muestra la gráfica de $f(x,y) = x^2+y^2-x-y+1$ en un entorno de $A = (\sqrt{2}/2, \sqrt{2}/2)$ (en rojo) cuando nos restringimos a la frontera del dominio. En este caso se ve que A es un mínimo de f por la frontera del dominio. A la derecha se muestran en verde los valores de f si nos acercamos a A por la trayectoria $y = x$ (por en interior del dominio) que prueba que A no es un mínimo local $f(P) < f(A)$.

∂S del dominio S. A la derecha se muestra, sin embargo, que dicho punto A no es un mínimo local en S ya que al acercarnos a A por la trayectoria $y = x$ en el interior de la región S la curva es creciente teniendo su mínimo en el punto P, que como vimos es el mínimo global. □

4.2.1. Problemas propuestos

Resolver, utilizando el método de los multiplicadores de Lagrange, los siguientes problemas:

1. Encontrar los puntos de mayor y menor distancia al origen de la curva (hipérbola) definida por $x^2 + 3xy + y^2 - 4 = 0$.
2. Encontrar la mínima distancia al origen de la curva $(x-3)^2+(y-4)^2 = a^2$ con $a = 1$ y $a = 10$.
3. Encontrar los extremos de la función $f(x,y) = xy$ sobre la curva $x^2+y^2 = 1$.
4. Encontrar los puntos críticos y los extremos absolutos (si los tiene) de la función $f(x,y) = x^2 + y^2 - x - y + 1$ en la región $x^2 + y^2 \leq 2$.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

5. Encontrar los puntos críticos y los posibles extremos de la función $f(x,y,z) = x^2+y^2+z^2-2x+4y-3$ en el dominio definido por las desigualdades $x^2+y^2+z^2 \leq 9$ y $x \geq 0$.

6. Encuentra los puntos más cercanos y alejados del elipsoide $4x^2+y^2+(z-2)^2 = 1$ al punto $(-1,0,2)$.

7. Sea la curva definida por la intersección del paraboloide $z = x^2+y^2$ y el plano $x+y+z = 12$. Encontrar el punto más alto y el más bajo. Encontrar los puntos de mayor y menor distancia de la curva al punto $(0,0,0)$.

8. Encontrar los puntos de mayor y menor distancia al origen de la curva definida por las ecuaciones $x^2+y^2-z^2-xy-1 = 0$ y $x^2+y^2-1 = 0$.

9. Encuentra los extremos absolutos de la función $f : D \setminus \{(0,0)\} \to \mathbb{R}$, $f(x,y) = \log\left(x^2+y^2\right)$ donde D es la región definida por $\{(x,y) : x \geq 1, (x-1)^2+y^2 \leq 4\}$.

10. Sea la curva definida por $x^2+y^2+z^2 = 11$ y el plano $x+y+z = 3$. Encontrar los extremos de $f(x,y,z) = z^2+2x+2y+20$ sobre dicha curva.

11. Encontrar los extremos de la función $f(x,y) = x+y$ sobre la curva $4y^2+x^2-1 = 0$.

12. Encuentra los puntos del elipsoide $x^2+2y^2+4z^2 = 9$ más cercanos y alejados del punto $(0,0,3)$.

13. Encontrar los puntos de mayor y menor distancia al origen de la curva definida por $y^2-8y+x^2-6x-75 = 0$.

14. Encuentra todos los puntos críticos de la función $f(x,y) = (x-1)^2+y^2$ definida sobre la región $D = \left\{(x,y) : y \geq -x, y^2 \leq 2x, (x-2)^2+y^2 \leq 4\right\}$ y decide si son extremos relativos. Encuentra, caso que los tenga, los extremos absolutos en D.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Bibliografía

Para saber más sobre funciones de varias variables se recomienda consultar los textos [1, 2, 3, 5, 6, 9]. Para las secciones 1.2-1.6 consultar [4, 7]. Para saber más de de funciones de una variable ver [5, 8].

[1] T.M. Apostol *Análisis Matemático*, 2^a edición. Reverté, Barcelona 1976.

[2] J. de Burgos, *Cálculo infinitesimal de varias variables*. McGraw-Hill, Madrid, 2002.

[3] R. Courant y F. John, *Introducción al Cálculo y al Análisis Matemático*, Vol. I y II. Limusa, México, 1976 y 1978.

[4] Kreyszig, E., *Introductory Functional Analysis with Applications*. Wiley Classics Library Edition, Nueva York, 1989.

[5] Kudriavtsev L.D., *Curso de Analisis Matematico*, Vol. I y II. MIR, Moscú, 1983.

[6] J. Marsden, A.J. Tromba y A. Weinstein, *Basic multivariate calculus*. Springer, New York, 1993.

[7] B.P. Rynne y M.A. Youngson. *Linear Functional Analysis* (2th Edition). Springer, London, 2008.

[8] V.A. Zorich, *Mathematical Analysis I*. Springer-Verlag, Berlín, 2015

[9] V.A. Zorich, *Mathematical Analysis II*. Springer-Verlag, Berlín, 2016.

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Diferenciación de funciones de varias variables

Índice alfabético

aplicación lineal, 31
    acotada, 32

cambio de variables, 96
Condición necesaria de extremo relativo, 104
Condición suficiente de diferenciabilidad, 56, 58
Condición suficiente de extremo relativo, 105
    Criterio de Sylvester, 108
conjunto acotado, 3
conjuntos numerables, 2
Criterio de Cauchy para las sucesiones, 4

dependencia funcional, 99
    Teorema del rango, 102
derivada de una función, 6
derivada direccional, 44
derivada parcial, 43
derivadas de orden superior, 62
desigualdad de Cauchy-Schwarz, 35
diferencial de una función, 45
    representación matricial, 48
diferenciales de orden superior, 68

espacio compacto, 30
espacio completo, 25
espacio euclídeo, 35
espacio métrico, 16
espacio normado, 18
    de dimensión finita, 26
espacio separable, 26
espacio vectorial, 13
extremos condicionados, 111
    Condición necesaria, 114
    Condición suficiente, 115
extremos libres de una función, 103
extremos locales, 7

función continua, 5, 39
función diferenciable, 45

gradiente de una función, 48

infinitésimos equivalentes, 10

Lema de Fermat, 8
límite
    de una sucesión, 3
    de una función de una variable, 5
    de una función de varias variable, 39

normas equivalentes, 28

o pequeña $o(x)$, 11

plano tangente a una función, 50
polinomio de Taylor, 10
punto de acumulación, 2

Regla de la cadena, 55

subsucesión, 4
sucesión de Cauchy, 4, 25
sucesión numérica, 1

Teorema de Bolzano, 6

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES

Teorema de Bolzano-Weierstrass, 4, 29
Teorema de Bonnet, 66
Teorema de Heffter-Young, 67
Teorema de la función implícita, 83
Teorema de la función inversa, 93
Teorema de Schwarz, 64
Teorema de sistemas de funciones implícitas, 89
Teorema de Taylor, 10, 75
Teorema de Weierstrass para funciones continuas, 40
Teorema del valor medio, 59
Teorema del valor medio de Lagrange, 8
Teorema local de Taylor, 75

DIFERENCIACIÓN DE FUNCIONES DE VARIAS VARIABLES