Análisis III

Joaquín M. Ortega Aramburu

Septiembre de 1999
Actualizado en julio de 2001

Índice General

1 Continuidad en el espacio euclídeo **5**

1.1 El espacio euclídeo R^n 5

1.1.1 Distancias y límites de sucesiones 5

1.1.2 Ejercicios 9

1.2 Conceptos topológicos en R^n 10

1.2.1 Interior, exterior y frontera. Conjuntos abiertos. 10

1.2.2 Puntos adherentes, de acumulación y aislados. Conjuntos cerrados. . . . 12

1.2.3 Conjuntos compactos por sucesiones y conjuntos compactos por recubrimientos 14

1.2.4 Ejercicios 16

1.3 Límites de funciones 17

1.3.1 Conceptos y propiedades fundamentales 17

1.3.2 Ejercicios 22

1.4 Funciones continuas 23

1.4.1 Funciones continuas en un punto 23

1.4.2 Funciones continuas en un conjunto 26

1.4.3 Curvas y conexión 29

1.4.4 Ejercicios 30

1.5 Funciones uniformemente continuas. 32

1.5.1 Conceptos y propiedades fundamentales 32

1.5.2 Ejercicios 33

1.6 Nota histórica 34

1.7 Apéndice. Espacios métricos 35

2 Cálculo diferencial en varias variables **39**

2.1 Derivadas direccionales. Diferencial 39

2.1.1 Conceptos y propiedades fundamentales 39

2.1.2 Derivadas de orden superior al primero 49

2.1.3 Ejercicios 51

2.2 Diferenciales de orden superior al primero. Formula de Taylor. 52

2.2.1 Aplicaciones multilineales y diferenciales de orden superior al primero. . . 52

2.2.2 La fórmula de Taylor 56

2.2.3 Extremos de una función 58

2.2.4 Ejercicios 62
2.3 El teorema de la función inversa e implícita 63
2.3.1 El teorema de la función inversa 63
2.3.2 El problema de la dependencia funcional 68
2.3.3 El teorema de la función implícita 69
2.3.4 Subvariedades diferenciables en R^n 73
2.3.5 Extremos condicionados. Los multiplicadores de Lagrange 77
2.3.6 Ejercicios 80
2.4 Nota histórica 82
2.5 Derivaciones y diferenciales. Una introducción algebraica. 84

Capítulo 1

Continuidad en el espacio euclídeo

1.1 El espacio euclídeo R^n

1.1.1 Distancias y límites de sucesiones

Denotaremos por R^n el espacio vectorial de las n-plas de números reales. Si $x = (x_1, \ldots, x_n)$ e $y = (y_1, \ldots, y_n)$ se define su producto escalar mediante la expresión

$$x \cdot y = \sum_{i=1}^{n} x_i y_i.$$

Este producto es lineal en cada factor, conmutativo y definido positivo. Esta última propiedad significa que $x \cdot x \geq 0$ y que sólo es cero si $x = 0$.

Veamos una importante desigualdad que a su vez permitirá deducir una de las propiedades básicas de la distancia euclídea.

Teorema 1.1. Desigualdad de Schwarz

Para cada $x, y \in R^n$ se cumple

$$|x \cdot y| \leq (x \cdot x)^{\frac{1}{2}} (y \cdot y)^{\frac{1}{2}}.$$

Demostración. Si uno de los vectores es 0 el resultado es trivial. Supongamos que ambos son no nulos. Para cada $\lambda \in R$ se cumple

$$(x + \lambda y) \cdot (x + \lambda y) \geq 0.$$

Es decir

$$x \cdot x + 2\lambda x \cdot y + \lambda^2 y \cdot y \geq 0.$$

Se trata de un polinomio de segundo grado en λ que no toma valores negativos y por tanto su discriminante será negativo o nulo

$$(x \cdot y)^2 - (x \cdot x)(y \cdot y) \leq 0.$$

□

Definición 1.1. *Se define la norma euclídea de los elementos de R^n como*

$$\|x\| = (x \cdot x)^{\frac{1}{2}}.$$

La distancia euclídea entre dos puntos de R^n es

$$d(x, y) = \|x - y\|.$$

Obsérvese que utilizamos la misma notación d para las aplicaciones distancia en los diversos espacios euclídeos R^n independientemente de su dimensión. Esto no dará lugar a confusiones.

Las propiedades esenciales de la norma las resumimos en el siguiente teorema.

Teorema 1.2. *La norma euclídea cumple las siguientes propiedades*

1. $\|x\| \geq 0$, $\|x\| = 0$ *si y sólo si* $x = 0$.

2. $\|\lambda x\| = |\lambda| \|x\|$ *para* $\lambda \in R$, $x \in R^n$.

3. $\|x + y\| \leq \|x\| + \|y\|$ *para cada* $x, y \in R^n$.

Esta última desigualdad recibe el nombre de desigualdad triangular.

Demostración. Las propiedades 1 y 2 son consecuencia inmediata de la definición de norma. La propiedad 3 se deducirá de la desigualdad de Schwarz:

$$\|x + y\|^2 = (x + y) \cdot (x + y) \leq x \cdot x + y \cdot y + 2(x \cdot x)^{\frac{1}{2}} (y \cdot y)^{\frac{1}{2}} = (\|x\| + \|y\|)^2.$$

□

Definición 1.2. *Una aplicación de R^n en R que cumpla las propiedades 1,2,3 de la proposición anterior se dice que es una norma.*

Ejemplo 1.1. *Además de la euclídea otros ejemplos de normas son*

$$\|x\|_\infty = \sup_{k=1,\ldots,n} |x_k|$$
$$\|x\|_1 = \sum_{k=1}^n |x_k|.$$

Es inmediato comprobar que estas aplicaciones cumplen las tres propiedades de norma.

Como consecuencia de las propiedades de la norma se obtienen las siguientes propiedades para la distancia euclídea.

Teorema 1.3. *La aplicación distancia d tiene las siguientes propiedades*

1. $d(x, y) \geq 0$. $d(x, y) = 0$ *si y sólo si* $x = y$.

2. $d(x, y) = d(y, x)$ *para cada* $x, y \in R^n$.

3. Se cumple la llamada desigualdad triangular, $d(x, y) \leq d(x, z) + d(z, y)$ *para cada* $x, y, z \in R^n$.

Demostración. Son consecuencia inmediata de las propiedades de la norma. Veamos, por ejemplo, la desigualdad triangular

$$d(x, y) = \|x - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y).$$

□

Obsérvese que las propiedades de la distancia son consecuencia de las propiedades de la norma. Por tanto, cada norma dará lugar a una distancia que tendrá las propiedades de la proposición anterior. Así, las nomas $\|\|_\infty$ y $\|\|_1$ darán lugar a sendas nociones de distancia

$$\begin{array}{c} d_\infty(x,y) = \|x-y\|_\infty \\ d_1(x,y) = \|x-y\|_1 \end{array} .$$

La noción de distancia permite definir la noción de entorno esférico y de límite en R^n.

Definición 1.3. *Se llama entorno esférico o bola de centro a y radio $\varepsilon > 0$ al conjunto*

$$\mathcal{E}(a,\varepsilon) = \{x \in R^n;\, d(a,x) < \varepsilon\}.$$

Cada distancia define entornos esféricos distintos asociados a la misma.

Ejemplo 1.2. *El conjunto $\{x \in R^n;\, d_\infty(a,x) < \varepsilon\}$ consiste en los puntos x tales que $|x_k - a_k| < \varepsilon$ para $k = 1, \ldots, n$, es decir, se trata de un cubo centrado en a y de arista 2ε.*

Podemos definir la noción de límite de una sucesión de elementos de R^n. Escribiremos los diferentes elementos de una sucesión mediante un superíndice para no confundirlos con las coordenadas de los elementos de R^n, que hemos denotado mediante subíndices.

Definición 1.4. *Una sucesión $\{a^i\}$, $a^i \in R^n$ se dice que tiene por límite $a \in R^n$ si para cada $\varepsilon > 0$, existe i_0 tal que $a^i \in \mathcal{E}(a,\varepsilon)$ para cada $i > i_0$. Lo escribiremos $\lim\{a^i\} = a$ o simplemente $\{a^i\} \to a$.*

Distancias distintas pueden dar lugar a nociones distintas de límites. Sin embargo, en los ejemplos anteriores, la distancia euclídea y las distancias d_∞ o d_1 dan lugar a la misma noción de límite. Esto es consecuencia del siguiente lema.

Lema 1.4. *Para cada $x, y \in R^n$ se tiene*

$$d_\infty(x,y) \le d(x,y) \le d_1(x,y) \le n d_\infty(x,y).$$

Demostración. La primera desigualdad se deduce de $|x_j - y_j| \le \left(\sum_k (x_k - y_k)^2\right)^{1/2}$. La segunda se deduce de $\sum_k (x_k - y_k)^2 \le \left(\sum_k |x_k - y_k|\right)^2$. La tercera desigualdad se obtiene acotando $|x_k - y_k|$ por $\sup_{j=1,\ldots,n} |x_j - y_j|$

$$\sum_k |x_k - y_k| \le n \sup_{j=1,\ldots,n} |x_j - y_j|.$$

□

Teorema 1.5. *Las tres distancias $d,\ d_\infty,\ d_1$ dan la misma noción de límite de una sucesión.*

Demostración. Si $\lim\{a^i\} = a$ respecto a la distancia d_∞ se tiene que para cada $\varepsilon > 0$, existe i_0 tal que para $i > i_0$, $d_\infty(a^i, a) < \frac{\varepsilon}{n}$. Se cumplirá entonces $d_1(a^i, a) < \varepsilon$ y $d(a^i, a) < \varepsilon$ como consecuencia del lema. Por lo tanto $\lim\{a^i\} = a$ respecto a d_1 y respecto a d.

Recíprocamente, del mismo lema se deduce que el límite respecto a la distancia d o d_1 implica el límite respecto a la distancia d_∞.

□

Teorema 1.6. *Una sucesión $\{a^i\}$ tiene por límite a si y sólo si cada una de las sucesiones coordenadas $\{a_k^i\}$, $k = 1, \ldots, n$ tiene por límite a_k, la correspondiente coordenada del límite.*

Demostración. Es consecuencia de las desigualdades del lema

$$|x_j - y_j| \leq d(x, y) \leq \sum_{k=1}^{n} |x_k - y_k|, \; j = 1, \ldots, n.$$

Si $\{a^i\} \to a$, la primera desigualdad da que $\{a_k^i\} \to a_k$, $k = 1, \ldots, n$. Recíprocamente, si para cada $k = 1, \ldots, n$ $\{a_k^i\} \to a_k$, la segunda desigualdad da $\{a^i\} \to a$.

□

Ejemplo 1.3. *La sucesión* $\left\{\left(\frac{2n+1}{n}, \frac{\sin n}{n}\right)\right\}$ *tiene por límite* $(2, 0)$ *puesto que* $\left\{\frac{2n+1}{n}\right\} \to 2$ *y* $\left\{\frac{\sin n}{n}\right\} \to 0$.

Esta propiedad permite probar propiedades de los límites de sucesiones de elementos de R^n a partir de sus análogas para límites de números reales. Así obtendremos los siguientes corolarios.

Corolario 1.7. *Si una sucesión de puntos de R^n tiene límite, éste es único.*

Demostración. Supongamos que $\{a^i\} \to a$ y que $\{a^i\} \to b$. La sucesión de las coordenadas $k-$ésimas de $\{a^i\}$ tendrá por límite tanto a_k como b_k. Por la unicidad de los límites de sucesiones de números reales $a_k = b_k$ para cada k y, por tanto, $a = b$.

□

Si definimos una suma de sucesiones y un producto de sucesiones por escalares en la forma $\{a^i\} + \{b^i\} = \{a^i + b^i\}$, $\lambda \{a^i\} = \{\lambda a^i\}$ se tiene la siguiente proposición.

Corolario 1.8. *Sean dos sucesiones de elementos de R^n que tengan límite $\{a^i\} \to a$ y $\{b^i\} \to b$ y sea $\lambda \in R$. Se tiene*

$$\left\{a^i\right\} + \left\{b^i\right\} \to a + b, \quad \lambda \left\{a^i\right\} \to \lambda a.$$

Demostración. Como en el corolario anterior, es suficiente considerar las correspondientes sucesiones de coordenadas y tener en cuenta las propiedades análogas para las sucesiones de números reales.

□

Corolario 1.9. *Sea $\{a^i\}$ una sucesión de elementos de R^n que tenga por límite a. Toda sucesión parcial $\{a^{i_j}\}$ tiene el mismo límite.*

Demostración. Cada una de las sucesiones coordenadas $\left\{a_k^{i_j}\right\}$ por ser una sucesión parcial de $\{a_k^i\}$ tendrá por límite a_k. Por lo tanto $\{a^{i_j}\}$ tendrá por límite a.

□

Es natural definir en R^n el concepto de sucesión de Cauchy.

Definición 1.5. *Una sucesión $\{a^i\}$, $a^i \in R^n$ se dice de Cauchy si para cada $\varepsilon > 0$ existe i_0 tal que si $i, j > i_0$, $d(a^i, a^j) < \varepsilon$.*

Razonando como en el teorema 1.6 es fácil ver $\{a^i\}$ es de Cauchy si y sólo si todas las sucesiones coordenadas $\{a_k^i\}$ son sucesiones de Cauchy de números reales y que, por tanto, la sucesión es de Cauchy si y sólo si tiene límite.

1.1.2 Ejercicios

1. Halla el límite de las sucesiones en R^2:
$$\left(n \sin \frac{1}{n}, \frac{1}{\sqrt{n}(\sqrt{n+1}-\sqrt{n})}\right), \qquad \left(\frac{e^{1/n}-1}{1/n}, \frac{3n}{5n+6}\right)$$

2. ¿Qué consecuencias se pueden obtener de las desigualdades del lema 1.4 en términos de las bolas correspondientes a las métricas d, d_1 y d_2?

3. Dado un conjunto E de R^n se define $d(x,E) = \inf_{y \in E} d(x,y)$. Sea
$$A = \{(x_1,x_2) \in R^2;\ x_1 \geq 0,\ 0 \leq x_2 < x_1\}.$$
Explicita los conjuntos
   (a) $\{x \in R^2;\ d(x,A) = 0\}$.
   (b) $\cup_{a \in A} \mathcal{E}(a, 2d(a, R^2 \setminus A)$.

4. Da una norma en R^2 tal que las bolas correspondientes sean elipses cuya proporción entre los ejes esté prefijada.

5. Prueba las siguientes relaciones, para cada $a, b, a^i \in R^n$
$$\|a\| - \|b\| \leq \|a-b\|$$
$$\left\|\sum_{i=1}^m a^i\right\| \leq \sum_{i=1}^m \left\|a^i\right\|$$
$$\|a+b\|^2 + \|a-b\|^2 = 2\left(\|a\|^2 + \|b\|^2\right)$$
$$\|a+b\|\,\|a-b\| \leq \|a\|^2 + \|b\|^2.$$

6. Utilizando el ejercicio anterior prueba que si $\{a^i\}$, $a^i \in R^n$, es tal que $\sum_i \|a^i\| < +\infty$ entonces existe el límite de la sucesión de las sumas parciales de la serie $\sum_i a^i$.

7. Prueba directamente a partir de la definición que si una sucesión de elementos de R^n tiene límite, éste es único, sin utilizar la proposición equivalente para las sucesiones de números reales.

8. Sea T una aplicación de R^n en R^n tal que existe un $k < 1$ tal que para todo $x, y \in R^n$, $\|T(x) - T(y)\| \leq k\|x-y\|$. Sea $a \in R^n$. Prueba que el conjunto $\{T^i(a);\ i \in N\}$ es acotado, es decir, que existe K tal que $\|T^i(x)\| \leq K$. Indicación: Considera la igualdad $T^i(a) - a = \sum_{j=1}^i \left(T^j(a) - T^{j-1}(a)\right)$.

1.2 Conceptos topológicos en R^n

Como primera aproximación se puede decir que la topología es el estudio de las propiedades que permanecen invariantes en un espacio a través de aplicaciones biyectivas del espacio consigo mismo tales que, tanto ellas como sus inversas, transformen puntos próximos en puntos próximos. Esto implica poder hablar de proximidad. En los espacios euclídeos puede hacerse a partir de la noción de distancia, y de aplicación que transforma puntos próximos a uno dado en puntos arbitrariamente próximos a su imagen, es decir, de aplicación continua. En esta sección trataremos de los conceptos básicos y de algunas de estas propiedades topológicas en el contexto de los espacios euclídeos.

1.2.1 Interior, exterior y frontera. Conjuntos abiertos.

Definición 1.6. *Sea A un subconjunto de R^n. Un punto $a \in R^n$ se dice que es interior (resp. exterior) al conjunto si existe un entorno esférico, de centro a, $\mathcal{E}(a, \varepsilon)$ contenido en el conjunto A (resp. contenido en $R^n - A$). Un punto que no es interior ni exterior, es decir, que para cada $\mathcal{E}(a, \varepsilon)$ existen puntos del conjunto A y de su complementario se dice que pertenece a la frontera de A. El conjunto de los puntos interiores a A lo denotaremos por A°.*

Obviamente se verifica que $A^\circ \subset A$. Un tipo particularmente importante de conjuntos son aquellos que coinciden con su interior.

Definición 1.7. *Un conjunto A se dice abierto si todo punto de A es interior a A, es decir si $A^\circ = A$.*

Ejemplo 1.4. *1. Los entornos esféricos $\mathcal{E}(a, \varepsilon)$ son conjuntos abiertos. En efecto, veamos que todo punto $x \in \mathcal{E}(a, \varepsilon)$ es interior al conjunto. Si llamamos $\eta = d(x, a) < \varepsilon$ tendremos que $\mathcal{E}(x, \varepsilon - \eta) \subset \mathcal{E}(a, \varepsilon)$ puesto que si $y \in \mathcal{E}(x, \varepsilon - \eta)$*

$$d(y, a) \leq d(y, x) + d(x, a) < \varepsilon - \eta + \eta = \varepsilon$$

y por tanto $y \in \mathcal{E}(a, \varepsilon)$.

2. Los conjuntos formados por un número finito de puntos no son abiertos.

Las propiedades esenciales de los conjuntos abiertos se resumen en el siguiente teorema.

Teorema 1.10. *1. Los conjuntos R^n y ϕ son conjuntos abiertos.*
2. Si A_i es una colección de conjuntos abiertos, $\cup_i A_i$ es un conjunto abierto.
3. Si $A_1, \ldots, A_m$ es un número finito de conjuntos abiertos, $\cap_{i=1}^m A_i$ es un abierto.

Demostración. 1. Es obvio que todos los puntos de R^n son interiores y que el conjunto vacío ϕ tiene todos sus puntos (no existe ninguno) interiores.

2. Si $a \in \cup_i A_i$ para algún j se tiene que $a \in A_j$ y, por ser éste abierto, existe $\varepsilon > 0$ tal que $\mathcal{E}(a, \varepsilon) \subset A_j \subset \cup_i A_i$. Por tanto a es interior a $\cup_i A_i$.

3. Sea $a \in \cap_{i=1}^{m} A_i$. Para cada i, $a \in A_i$ y, por ser estos conjuntos abiertos, existirán $\varepsilon_i > 0$ tales que $\mathcal{E}(a, \varepsilon_i) \subset A_i$. Tomando $\varepsilon = \min\{\varepsilon_1, \ldots, \varepsilon_m\} > 0$ tendremos $\mathcal{E}(a, \varepsilon) \subset \cap_{i=1}^{m} A_i$. Por tanto a es interior al conjunto intersección.

□

Obsérvese que, en general, la intersección de infinitos conjuntos abiertos no es un conjunto abierto.

Ejemplo 1.5. *1. Los intervalos $\left(-\frac{1}{n}, \frac{1}{n}\right)$ son subconjuntos abiertos de R, pero su intersección es el conjunto $\{0\}$ que no es un conjunto abierto.*

2. Los conjuntos abiertos de R tienen una descripción especialmente simple. Los intervalos abiertos, las semirrectas abiertas, R y ϕ son conjuntos abiertos. Todo conjunto abierto de R es una unión finita o numerable de conjuntos disjuntos dos a dos del tipo anterior. En efecto, si A es un abierto no vacío y $a \in A$ se considera $M = \{x \in R,\ [a, x] \subset A\}$. Si M no está acotado, $[a, +\infty)$ está contenido en A. Si está acotado y $S = \sup M$, $[a, S)$ está contenido en A y $S \notin A$. Análogamente se considera el conjunto

$$N = \{x \in R,\ [x, a] \subset A\},$$

distinguiéndose el caso acotado y no acotado. De esta forma se llega fácilmente a la conclusión de que A es unión de intervalos abiertos, semirrectas abiertas o todo R disjuntos dos a dos. Tomando un número racional en cada una de estas componentes y sabiendo que los números racionales forman un conjunto numerable se concluye la finitud o numerabilidad del conjunto de componentes.

Los conjuntos de la forma $\mathcal{E}(a, \varepsilon)$ son conjuntos abiertos que contienen el punto a y les hemos llamado entornos esféricos de a. A un conjunto abierto que contiene un punto a se le llama también un entorno abierto de a.

Ejemplo 1.6. *El concepto de límite de una sucesión puede expresarse en términos de conjuntos abiertos o de entornos. Es fácil verificar que $\{a^i\} \to a$ si y sólo si para cada entorno abierto U de a existe i_0 tal que $a^i \in U$ si $i > i_0$. En efecto, supongamos que $\{a^i\} \to a$. Puesto que $a \in U$ y éste es abierto, existirá $\varepsilon > 0$ tal que $\mathcal{E}(a, \varepsilon) \subset U$. De la definición de límite de un sucesión, existe i_0 tal que si $i > i_0,$ $a^i \in \mathcal{E}(a, \varepsilon) \subset U$. Por tanto vale la condición en términos de entornos. En sentido contrario es trivial ya que $\mathcal{E}(a, \varepsilon)$ es un entorno abierto de a.*

Una propiedad que se cumple para los puntos de un entorno de un punto determinado se le suele denominar una propiedad local. Así, se suele decir que una función definida en un conjunto tiene un máximo local en a si existe un entorno de a (siempre se podrá suponer esférico) tal que la función restringida a este entorno toma su máximo en a. Más adelante veremos en detalle éste y otros conceptos en que es importante considerar propiedades referidas únicamente a un cierto entorno de un punto.

1.2.2 Puntos adherentes, de acumulación y aislados. Conjuntos cerrados.

Definición 1.8. *Sea A un subconjunto de R^n. Un punto $p \in R^n$ se dice que es adherente a A si para cada $\varepsilon > 0$, $\mathcal{E}(p, \varepsilon) \cap A \neq \phi$. Un punto p se dice que es de acumulación de A si en cada entorno de centro p existen elementos de A distintos de p. El conjunto de los puntos adherentes a A se denota por $\overline{A}$ y se llama adherencia de A. El conjunto de puntos de acumulación de A lo denotaremos por A' y se llama conjunto derivado.*

Ejemplo 1.7. 1. *Si $A = (a, b) \times (c, d) \subset R^2$ se tiene $\overline{A} = [a, b] \times [c, d]$.*

2. *Si $A = (1, 2) \times (1, 3) \cup \{(0, 0)\}$ se tiene $\overline{A} = [1, 2] \times [1, 3] \cup \{(0, 0)\}$ y $A' = [1, 2] \times [1, 3]$.*

Teorema 1.11. *Un punto p es de acumulación de A si y sólo si para cada $\varepsilon > 0$ existen en $\mathcal{E}(p, \varepsilon)$ infinitos elementos de A.*

Demostración. En efecto, si se cumple esta condición desde luego se trata de un punto de acumulación. Recíprocamente, sea p de acumulación. En cada $\mathcal{E}(p, \varepsilon)$ deberán existir infinitos puntos de A. De lo contrario, supongamos que existiese uno que contuviese un número finito $a^1, \ldots, a^m$. Tendríamos que $\eta = \min\{d(a^i, p), i = 1, \ldots, m \ a^i \neq p\}$ sería un número positivo y en $\mathcal{E}(p, \eta)$ no existiría ningún elemento de A distinto de p.

□

Es consecuencia inmediata de las definiciones que todo punto de A es adherente a A. Sin embargo no tiene por qué ser de acumulación de A.

Definición 1.9. *Un elemento a de A que no sea de acumulación se dice que es un punto aislado.*

Significa que existe $\varepsilon > 0$ tal que $\mathcal{E}(a, \varepsilon) \cap A = \{a\}$. Los puntos de un conjunto A son entonces o de acumulación de A o aislados. El concepto de punto adherente o de acumulación puede darse en términos de límites de sucesiones. Esto es lo que expresa el siguiente teorema.

Teorema 1.12. *Sea A un subconjunto de R^n. Un punto p es adherente a A si y sólo si existe una sucesión de puntos de A que tiene por límite p. El punto es de acumulación si y sólo si existe una sucesión de puntos de A, distintos del propio punto p, cuyo límite es p.*

Demostración. Sea $p \in \overline{A}$. Para cada $m \in N$, $\mathcal{E}\left(p, \frac{1}{m}\right) \cap A \neq \phi$. Si tomamos $a^m \in \mathcal{E}\left(p, \frac{1}{m}\right) \cap A$ tendremos que $\{a^m\} \to p$. El recíproco es trivial. Si existe una sucesión de elementos de A que tienen por límite p, este punto es adherente a A. El mismo argumento prueba la segunda parte del teorema.

□

Hemos visto que los puntos adherentes de un conjunto son puntos límites de puntos del conjunto. Son de especial interés aquellos conjuntos que coinciden con su adherencia.

Definición 1.10. *Un subconjunto C de R^n se dice que es cerrado si todo punto adherente a C es del conjunto. Equivale a decir que para cada sucesión de puntos de C que tenga límite, este límite es de C.*

Ejemplo 1.8. *El subconjunto* $[a, b]$ *de* R *es cerrado. El subconjunto* $[a, +\infty)$ *es cerrado. El subconjunto de* R *formado por los elementos de la forma* $\frac{1}{m}$ *para* $m \in N$ *unión con* $\{0\}$ *es cerrado.*

La proposición siguiente resume algunas de las propiedades esenciales de los conjuntos cerrados.

Teorema 1.13. *1. Los conjuntos* R^n *y* ϕ *son cerrados.*

2. Si C_i *es una colección de cerrados,* $\cap_i C_i$ *es cerrado.*

3. Si $C_1, \ldots, C_m$ *es una colección finita de cerrados,* $C_1 \cup \ldots \cup C_m$ *es cerrado.*

Demostración. 1. Es una consecuencia inmediata de la definición.

2. Sea p un punto de $\overline{\cap_i C_i}$. Para cada $\varepsilon > 0$, $\mathcal{E}(p, \varepsilon) \cap (\cap_i C_i) \neq \phi$. Por tanto, para cada i, $\mathcal{E}(p, \varepsilon) \cap C_i \neq \phi$ y por ser C_i cerrado, $p \in C_i$. Luego $p \in \cap_i C_i$ y este conjunto es cerrado.

3. Sea $p \in \overline{C_1 \cup \ldots \cup C_m}$. Veamos que $p \in C_1 \cup \ldots \cup C_m$. Si no fuese así, p no pertenecería a ningún C_i que, por ser cerrado coincide con $\overline{C_i}$. Existirían entonces $\varepsilon_i > 0$ tal que $\mathcal{E}(p, \varepsilon_i) \cap C_i = \phi$, $i = 1, \ldots, n$. Tomando $\varepsilon = \min\{\varepsilon_1, \ldots, \varepsilon_m\}$, será $\varepsilon > 0$ y tendremos $\mathcal{E}(p, \varepsilon) \cap (C_1 \cup \ldots \cup C_m) = \phi$, en contradicción con la hipótesis de ser $p \in \overline{C_1 \cup \ldots \cup C_m}$. Así pues, $C_1 \cup \ldots \cup C_m$ coincide con su adherencia.

□

La relación entre conjuntos cerrados y conjuntos abiertos se expresa en la siguiente proposición.

Teorema 1.14. *Un subconjunto de* R^n *es abierto si y sólo si su complementario es cerrado.*

Demostración. Sea A un conjunto abierto. Veamos que su complementario $C = R^n - A$ es cerrado. Consideremos $p \in \overline{C}$, veamos que $p \in C$ o, lo que es equivalente $p \notin A$. Si perteneciese, por ser A abierto existiría $\mathcal{E}(p, \varepsilon) \subset A$, es decir $\mathcal{E}(p, \varepsilon) \cap C = \phi$ con lo que p no pertenecería a $\overline{C}$.

Recíprocamente, partamos ahora de un conjunto C cerrado. Se trata de ver que su complementario $A = R^n - C$ es abierto. Sea entonces $a \in A$. Tendremos que $a \notin C = \overline{C}$ y por tanto existirá $\varepsilon > 0$ tal que $\mathcal{E}(a, \varepsilon) \cap C = \phi$. Esto equivale a decir que $\mathcal{E}(a, \varepsilon) \subset A$ y a es interior a A.

□

Esta proposición permite deducir las propiedades de los conjuntos abiertos a partir de la de los cerrados o la de los cerrados a partir de los abiertos. Por ejemplo, veamos cómo se puede obtener la propiedad 2 del teorema 1.13 sobre cerrados a partir de la propiedad 2 del teorema 1.10 sobre abiertos.

Sean C_i una colección de cerrados, sus complementarios $R^n - C_i$ son abiertos. De acuerdo con la proposición 1.10, $\cup_i (R^n - C_i)$ es un abierto y su complementario será un cerrado. Este complementario es $R^n - (\cup_i (R^n - C_i)) = \cap_i C_i$ y por tanto vale 2 de la proposición 1.13.

Cuando estudiemos las aplicaciones contínuas, una propiedad importante que utilizaremos en varias ocasiones es que para cada sucesión de un conjunto, exista una sucesión parcial que tenga límite un punto del propio conjunto. Los conjuntos que cumplen esta propiedad serán el objeto del próximo apartado.

1.2.3 Conjuntos compactos por sucesiones y conjuntos compactos por recubrimientos

Definición 1.11. *Un subconjunto K de R^n se dice que es compacto por sucesiones si para cada sucesión de elementos de K existe una sucesión parcial que tiene por límite un punto del propio conjunto K.*

Ejemplo 1.9. *1 Todo el espacio R^n no es un conjunto compacto por sucesiones. Podemos construir una sucesión tal que $d\left(a^i, a^j\right) \geq 1$ para cada i, j. Ninguna sucesión parcial de ésta podrá tener límite pues no será de Cauchy.*

2. $[a, b]$ es un subconjunto de R que es un compacto por sucesiones. En efecto, dada una sucesión, $\left\{a^i\right\}, a^i \in [a, b]$, por el teorema de Bolzano Weierstrass, existe una sucesión parcial convergente. El límite, por ser $[a, b]$ cerrado, será del propio conjunto.

Se trata ahora de dar una descripción de los conjuntos compactos por sucesiones.

Definición 1.12. *Un subconjunto de R^n se dice acotado si está contenido en un conjunto de la forma $[a_1, b_1] \times \ldots \times [a_n, b_n]$. Es equivalente a decir que está contenido en un entorno esférico, es decir, existe un k tal que $\|a\| \leq k$ para todo a del conjunto.*

Teorema 1.15. *Un subconjunto K de R^n es compacto por sucesiones si y sólo si es cerrado y acotado.*

Demostración. Supongamos en primer lugar que K es cerrado y acotado. Tendremos que K estará contenido en un conjunto de la forma $[a_1, b_1] \times \ldots \times [a_n, b_n]$. Sea ahora $\left\{c^i\right\}$ una sucesión de elementos de K. Consideremos la sucesión de las primeras coordenadas $\left\{c_1^i\right\}$. Sus elementos pertenecerán al intervalo $[a_1, b_1]$. Por el teorema de Bolzano Weierstrass existirá una sucesión parcial convergente $\left\{c_1^{i_j}\right\}$. Consideremos la sucesión de puntos de K correspondiente a esta sucesión parcial $\left\{c^{i_j}\right\}$. Tomemos a continuación la sucesión de las segundas coordenadas de ésta $\left\{c_2^{i_j}\right\}$. Aplicando el teorema de Bolzano Weierstrass, en el intervalo $[a_2, b_2]$ existirá una sucesión parcial de $\left\{c_2^{i_j}\right\}$ que será convergente. Sea la correspondiente sucesión de puntos $\left\{c^{i_{j_k}}\right\}$. Será una sucesión parcial de la dada tal que las sucesiones de las primeras y de las segundas coordenadas seràn convergentes. Tomaremos ahora la sucesión de las terceras coordenadas de está sucesión y continuaremos el proceso. Aplicando n veces el teorema de Bolzano Weierstrass llegamos a probar la existencia de una sucesión parcial de la dada tal que todas las sucesiones de sus coordenadas son convergentes y que, por tanto, es convergente como sucesión de puntos. El límite de esta sucesión, por ser K cerrado, pertenecerá al conjunto K. Hemos probado que el conjunto K es compacto por sucesiones.

Partimos ahora de un conjunto K compacto por sucesiones. Se trata de ver que es cerrado y acotado.

Veamos que K es cerrado. Sea $\{c^i\} \to c$ con $c^i \in K$. Debemos probar que $c \in K$. Por ser K compacto por sucesiones existirá una sucesión parcial que tiene por límite un punto de K. Toda sucesión parcial de $\{c^i\}$ tiene por límite c y por tanto $c \in K$.

Veamos que K es acotado. Si no lo fuese existiría $c^1 \in K$ tal que $d(O, c^1) \geq 1$ donde O es el origen de coordenadas. Existiría $c^2 \in K$ tal que $d(O, c^2) \geq d(O, c^1) + 1$. Por recurrencia, existiría $c^i \in K$ tal que $d(O, c^i) \geq d(O, c^{i-1}) + 1$. Tendríamos entonces que para $k > 0$

$$\begin{aligned} d\left(c^{i+k}, c^i\right) &\geq d\left(O, c^{i+k}\right) - d\left(O, c^i\right) \\ &= \sum_{j=i}^{i+k-1} \left(d\left(O, c^{j+1}\right) - d\left(O, c^j\right)\right) \geq k \geq 1. \end{aligned}$$

Esta sucesión de elementos de K no podría tener ninguna sucesión parcial convergente. Así pues K es acotado.

□

Veamos un concepto relacionado con el anterior que aparece con naturalidad cuando se estudian la funciones uniformemente continuas y otros problemas. Introduzcamos para ello la definición de recubrimiento abierto.

Definición 1.13. *Un recubrimiento abierto de un conjunto K es una colección de abiertos A_i tales que $K \subset \cup_{i \in I} A_i$.*

Sea $\{A_i,\ i \in I\}$, un recubrimiento de K. Un subrecubrimiento o recubrimiento parcial del dado es una colección de conjuntos A_i para $i \in J \subset I$ tales que $K \subset \cup_{i \in J} A_i$. El subrecubrimiento se dice finito si está formado por un número finito de conjuntos.

Definición 1.14. *Un subconjunto K de R^n se dice que es compacto, o compacto por recubrimientos, si para cada recubrimiento abierto de K existe un subrecubrimiento parcial y finito.*

Ejemplo 1.10. *R no es un conjunto compacto. En efecto, la colección $(-n, n)$, $n \in N$ es un recubrimiento abierto que no admite ningún subrecubrimiento parcial y finito.*

Teorema 1.16. *Un subconjunto K de R^n es compacto si y sólo si es compacto por sucesiones.*

Demostración. Supongamos que K sea compacto por sucesiones. Sea $A_i, i \in I$ un recubrimiento abierto de este conjunto. El primer paso es probar que existe $\varepsilon > 0$ tal que para cada $p \in K$ existe un abierto del recubrimiento A_j tal que $\mathcal{E}(p, \varepsilon) \subset A_j$. Lo demostraremos por reducción al absurdo. Supongamos que no existiese un tal ε. Para cada $m \in N$ existiría p^m de K, tal que $\mathcal{E}\left(p^m, \frac{1}{m}\right)$ no estaría contenido en ningún abierto del recubrimiento. Puesto que K es compacto por sucesiones existiría una sucesión parcial de $\{p^m\}$ que tendría por límite un punto de K. Sea $\{p^{m_j}\} \to a \in K$. El punto a pertenecerá a un cierto A_j. Por ser este conjunto abierto existirá $\eta > 0$ tal que $\mathcal{E}(a, \eta) \subset A_j$. Consideremos un $m_j \in N$ tal que $d(a, p^{m_j}) < \frac{\eta}{2}$ y al tiempo $\frac{1}{m_j} < \frac{\eta}{2}$. Tendríamos entonces que $\mathcal{E}\left(p^{m_j}, \frac{1}{m_j}\right) \subset \mathcal{E}(a, \eta) \subset A_j$. Contradicción con la propiedad que cumplían los puntos p^m.

El segundo paso consiste en probar que dado un $\delta > 0$ podemos recubrir K mediante un número finito de bolas de radio δ centradas en puntos de K. En efecto, tomemos un punto $p^1 \in K$

y consideremos $\mathcal{E}(p^1,\delta)$. Si K está contenido en $\mathcal{E}(p^1,\delta)$, hemos terminado. Si no es así consideremos un punto $p^2 \in K$ y $p^2 \notin \mathcal{E}(p^1,\delta)$. Si $K \subset \mathcal{E}(p^1,\delta) \cup \mathcal{E}(p^2,\delta)$ habríamos acabado. Si no es así consideremos $p^3 \in K$ y $p^3 \notin \mathcal{E}(p^1,\delta) \cup \mathcal{E}(p^2,\delta)$. El proceso no puede prolongarse indefinidamente ya que de lo contrario tendríamos una sucesión $\{p^m\}$ de elementos de K tales que $d(p^m,p^r) \geq \delta$ y, puesto que toda sucesión convergente es de Cauchy, no podría existir ninguna sucesión parcial convergente, en contradicción con la hipótesis sobre K.

El tercer paso consiste en aplicar los dos anteriores para obtener un recubrimiento parcial y finito del dado. Por el primer paso existe $\varepsilon > 0$ tal que para cada $p \in K$ existe A_j tal que $\mathcal{E}(p,\varepsilon) \subset A_j$. Apliquemos a este ε el segundo paso. Sea $K \subset \mathcal{E}(p^1,\varepsilon) \cup \ldots \cup \mathcal{E}(p^m,\varepsilon)$. Para cada $i = 1,\ldots,m$ existe A_{j_i} tal que $\mathcal{E}(p^j,\varepsilon) \subset A_{j_i}$. Tendremos que $A_{j_1},\ldots,A_{j_m}$ es un recubrimiento parcial y finito del dado.

Veamos ahora el teorema recíproco. Suponemos que K es compacto y queremos ver que es compacto por sucesiones. Sea $\{p^m\}$ una sucesión de puntos de K. Se trata de probar que una sucesión parcial tiene por límite un punto de K. Si no fuese así, ningún $a \in K$ sería límite de una sucesión parcial. Luego, fijado a, existiría $\varepsilon > 0$ y un natural m_a tal que p^m no pertenecería a $\mathcal{E}(a,\varepsilon)$ para $m > m_a$. De lo contrario existiría $p^{i_1} \in \mathcal{E}(a,1)$, a continuación un $p^{i_2} \in \mathcal{E}\left(a,\frac{1}{2}\right)$ con $i_2 > i_1$ y, en general, $p^{i_n} \in \mathcal{E}\left(a,\frac{1}{n}\right)$ con $i_n > i_{n-1}$. La sucesión p^{i_j} tendría por límite a. Si hacemos lo propio con cada punto de K tendremos un recubrimiento abierto formado por los conjuntos $\mathcal{E}(a,\varepsilon)$. Existirá un recubrimiento parcial y finito, es decir tal que

$$K \subset \mathcal{E}\left(a^1,\varepsilon\right) \cup \ldots \cup \mathcal{E}(a^r,\varepsilon).$$

Tomando $m > \max\{m_{a^1},\ldots,m_{a^r}\}$ tendremos que $p^m \notin K$. Hemos llegado a contradicción.

□

1.2.4 Ejercicios

1. Halla la adherencia, el interior y la frontera de los siguientes conjuntos:

   a) $\{(x,y) \in R^2;\ 0 < x \leq 1,\ 2 \leq y < 3\}$.

   b) $\{(x,y) \in R^2;\ x^2 + y^2 < 2\}$.

   c) Q^n en R^n.

   d) $\{(x,y) \in R^2;\ y = 0\}$.

2. Averigua cuales de los siguientes conjuntos son abiertos, cerrados o de ninguno de estos dos tipos.

   a) En R los números de la forma $\frac{1}{n} + \frac{1}{m}$ para $n,m \in N$.

   b) $\{(x,y) \in R^2;\ y \geq x^2\}$.

   c) $\{(x,y) \in R^2;\ x^2 + y^2 = 4\}$.

   d) $\{x \in R^n;\ \|x\| \leq 1\}$.

3. Prueba que la adherencia del interior de $\{x \in R^n;\ d(x,a) \leq 1\}$ es el mismo conjunto.

4. Da un ejemplo de un conjunto cerrado que la adherencia del interior no coincida con el conjunto.

5. Prueba las siguientes relaciones:

   a) $(A^\circ)^\circ = A^\circ$.

   b) $A^\circ = R^n - \overline{(R^n - A)}$.

   c) $(R^n - A)^\circ = R^n - \overline{A}$.

6. Prueba que si $A \subset R^n$ el conjunto $B = \{x \in R^n; d(x, A) \leq 1\}$ es un conjunto cerrado. Dar un ejemplo de un conjunto A en que B° no coincide con $\{x \in R^n; d(x, A) < 1\}$.

7. Da un ejemplo de dos conjuntos A y B de interior vacío y tal que el interior de la unión sea $\{(x, y) \in R^2;\ 0 < x < 1\}$.

8. Averigua cuales de los siguientes conjuntos son compactos.

   a) $\{(x, y) \in R^2;\ y = x^2,\ 0 \leq x \leq 1\}$.

   b) $\{(x, y, z) \in R^3;\ x^2 + 3y^2 + 2z^2 \leq 4\}$.

   c) $\{(x, y) \in R^2;\ xy = 1\}$.

   d) $\{(x, y) \in R^2;\ 0 < x^2 + y^2 \leq 2\}$.

9. Da un ejemplo de un recubrimiento abierto de $(0, 1)$ tal que no admite ningún recubrimiento parcial y finito.

10. Sea $A \subset K$ con K un conjunto compacto. Prueba que la frontera de A es un conjunto compacto.

11. Prueba que la unión de un número finito de conjuntos compactos es compacto. Da un ejemplo de una unión de infinitos compactos que sea un conjunto acotado pero no compacto.

12. Prueba que si A es un conjunto abierto de R^n entonces el interior de la frontera de A es vacío

13. Prueba que si K_1 y K_2 son dos compactos de R^{n_1} y R^{n_2} respectivamente, entonces $K_1 \times K_2$ es un compacto de $R^{n_1+n_2}$.

1.3 Límites de funciones

1.3.1 Conceptos y propiedades fundamentales

Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m. La idea intuitiva de límite de f cuando x tiende a un punto $a \in R^n$ es el de la existencia de un $l \in R^m$ tal que que los valores de $f(x)$ estén arbitrariamente próximos a l siempre que se tome $x \in D,\ x \neq a$

suficientemente próximo a a. Para poder asegurar que existen puntos de D tan próximos como queramos a a pero distintos de a supondremos que éste es un punto de acumulación de D. Veamos que la definición de límite puede darse alternativamente en términos de entornos o en términos de límites de sucesiones.

Definición 1.15. *I Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m y sea $a \in R^n$ un punto de acumulación de D. Diremos que el límite de $f(x)$ cuando x tiende a a es $l \in R^m$ si para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si $0 < d(x,a) < \delta$, $x \in D$, se cumple $d(f(x), l) < \varepsilon$. Lo escribiremos $\lim_{x\to a} f(x) = l$, o bien $f(x) \to_{x\to a} l$.*

II Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m y sea $a \in R^n$ un punto de acumulación de D. Diremos que el límite de $f(x)$ cuando x tiende a a es $l \in R^m$ si para cada sucesión $\{x^i\} \to a$, $x^i \in D$, $x^i \neq a$, se cumple $\{f(x^i)\} \to l$.

Teorema 1.17. *Las dos definiciones anteriores son equivalentes.*

Demostración. Supongamos que se cumple I y consideremos una sucesión $\{x^i\} \to a$, $x^i \in D$, $x^i \neq a$. Queremos probar que $\{f(x^i)\} \to l$. Sea $\varepsilon > 0$ y consideremos el δ dado por la definición I. Se tiene que $0 < d(x^i, a) < \delta$ para $i > i_0$ y por tanto $d(f(x^i), l) < \varepsilon$. La sucesión $\{f(x^i)\}$ tiene por límite l y vale la definición II.

Supongamos ahora que vale II. Si no fuese cierto I, existiría $\varepsilon > 0$ tal que para cada $i \in N$, existiría $x^i \in D$ tal que $0 < d(x^i, a) < \frac{1}{i}$ pero $d(f(x^i), l) \geq \varepsilon$. Tendríamos entonces que la sucesión $\{x^i\} \to a$, $x^i \in D$, $x^i \neq a$, pero no se cumpliría $\{f(x^i)\} \to l$ y por tanto no valdría II. $\square$

Obsérvese que en la noción de límite no aparece el valor de la función en el punto a en el que estamos calculando el límite. No es relevante si la función está o no definida en a y si lo está cuál es su valor. Por ello si la definición I se quiere expresar en términos de entornos es cómodo dar una notación específica para un entorno al que se le excluye su centro. Así escribiremos $\mathcal{E}^*(a,\varepsilon) = \mathcal{E}(a,\varepsilon) - \{a\}$. Con esta notación la definición I puede enunciarse de la siguiente forma.

Definición 1.16. *Se dice que $\lim_{x\to a} f(x) = l$ si para cada $\mathcal{E}(l,\varepsilon)$, existe $\delta > 0$ tal que $f(D \cap \mathcal{E}^*(a,\delta)) \subset \mathcal{E}(l,\varepsilon)$.*

Ejemplo 1.11. *1. Consideremos la función definida en $R^2 - \{(0,0)\}$ a valores en R mediante $f(x,y) = \frac{xy}{x^2+y^2}$. Veamos que no existe $\lim_{(x,y)\to(0,0)} f(x,y)$. En efecto, tomemos la sucesión $\left(\frac{1}{m}, 0\right)$ que tiene por límite $(0,0)$. La sucesión de las imágenes es $\left\{f\left(\frac{1}{m},0\right)\right\} = \{0\}$ que tiene por límite 0. Sin embargo podemos tomar otras sucesiones, como $\left(\frac{1}{m},\frac{1}{m}\right)$, que también tiene por límite $(0,0)$ pero que la sucesión de las imágenes $\left\{f\left(\frac{1}{m},\frac{1}{m}\right)\right\} = \left\{\frac{1}{2}\right\}$ tiene un límite distinto del anterior.*

2. Consideremos la función definida en $R^2 - \{(0,0)\}$ a valores en R mediante $f(x,y) = \frac{xy^2}{x^2+y^2}$. Veamos que $\lim_{(x,y)\to(0,0)} f(x,y) = 0$. En efecto, sea $\varepsilon > 0$. Escribamos el pun-

to (x, y) en la forma $x = \rho \cos \varphi$, $y = \rho \sin \varphi$. Tendremos $f(x, y) = \frac{\rho^3 \cos \varphi \sin^2 \varphi}{\rho^2} = \rho \cos \varphi \sin^2 \varphi$. De esta forma si $d((x, y), (0, 0)) = \rho < \delta = \varepsilon$ se obtendrá $|f(x, y)| < \varepsilon$.

La siguiente proposición permite reducir el estudio de límites de funciones a valores en R^m al de funciones a valores números reales.

Teorema 1.18. *Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m y sea $a \in R^n$ un punto de acumulación de D. Llamemos $f_1, \ldots, f_m$ a las m funciones coordenadas de f. Entonces $\lim_{x \to a} f(x) = l$ si y sólo si para cada $i = 1, \ldots, m$ se cumple $\lim_{x \to a} f_i(x) = l_i$.*

Demostración. Es suficiente considerar la definición II de límite y tener en cuenta que una sucesión tiene límite si y sólo si la sucesión de cada una de las coordenadas tiene por límite la correspondiente coordenada del límite.

□

Veamos algunas propiedades elementales de los límites. Tanto los enunciados como sus demostraciones son enteramente análogos a los de las funciones de una variable.

Teorema 1.19. *Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m y sea a un punto de acumulación de D. Si existe $\lim_{x \to a} f(x)$, éste es único.*

Demostración. Supongamos que $\lim_{x \to a} f(x) = l$ y que $\lim_{x \to a} f(x) = r$. Consideremos una sucesión $\{x^i\} \to a$, $x^i \in D$, $x^i \neq a$. Tendremos que $\{f(x^i)\} \to l$ y, al tiempo, $\{f(x^i)\} \to r$. Dada la unicidad de los límites de las sucesiones, $l = r$.

□

Teorema 1.20. *Sean f y g dos funciones definidas en un conjunto $D \subset R^n$ a valores en R^m y sea a un punto de acumulación de D. Si existen $\lim_{x \to a} f(x) = l_1$, $\lim_{x \to a} g(x) = l_2$, entonces*

$$\lim_{x \to a} (f + g)(x) = l_1 + l_2.$$

Si $m = 1$

$$\lim_{x \to a} (fg)(x) = l_1 l_2.$$

Si, además $g(x)$ es no nulo para todo x y $l_2 \neq 0$,

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{l_1}{l_2}.$$

Demostración. Se reduce, como en la proposición anterior, a las propiedades análogas para los límites de sucesiones. Veamos por ejemplo la primera propiedad. Consideremos una sucesión $\{x^i\} \to a$, $x^i \in D$, $x^i \neq a$. Tendremos $\{f(x^i)\} \to l_1$ y $\{g(x^i)\} \to l_2$. De aquí que $\{(f + g)(x^i)\} \to l_1 + l_2$. Puesto que esto vale para toda sucesión $\{x^i\} \to a$, $x^i \in D$, $x^i \neq a$ tendremos que $\lim_{x \to a} (f + g)(x) = l_1 + l_2$.

□

Teorema 1.21. *1. Sean f y g dos funciones definidas en un conjunto $D \subset R^n$ a valores en R y sea a un punto de acumulación de D. Supongamos que para todo $x \in D$ se cumple que $f(x) \leq g(x)$ y que existen $\lim_{x \to a} f(x) = l_1$, $\lim_{x \to a} g(x) = l_2$, entonces $l_1 \leq l_2$.*

2. Sean f, g, h tres funciones definidas en $D \subset R^n$ a valores en R , tales que para cada $x \in D$ se cumple $f(x) \leq h(x) \leq g(x)$ y que $\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = l$. Entonces $\lim_{x \to a} h(x) = l$.

Demostración. Veamos 1. Sea $\{x^i\} \to a$, $x^i \in D$, $x^i \neq a$. Tendremos $f(x^i) \leq g(x^i)$ y por tanto $l_1 \leq l_2$. De una forma análoga se prueba 2.

□

Para las funciones de varias variables a valores en R pueden darse, como en el caso de funciones de una sola variable, la noción de que una función tenga por límite $+\infty, -\infty$ o ∞. Por ejemplo sea f una función definida en $D \subset R^n$ y a un punto de acumulación de D. Se dirá que el límite de la función en a es $+\infty$ si para cada k existe $\epsilon > 0$ tal que $f(x) > k$ para cada $x \in D \cap \mathcal{E}(a, \epsilon)$ $x \neq a$. Valen las mismas reglas de cálculo con las mismas demostraciones y se tienen los mismos casos de indeterminación. Nos remitimos para todo ello al estudio de las funciones de una variable.

También ahora se puede utilizar la notación de los infinitésimos. Una función f definida en $D \subset R^n$ a valores en R se dice que es un infinitésimo en un punto a de acumulación de D si $\lim_{x \to a} f(x) = 0$. En el capítulo siguiente utilizaremos la notación clásica de "o pequeña" para establecer que un infinitésimo es de orden superior a otro. Recordemos ahora sólo la siguiente definición.

Definición 1.17. *Dados dos infinitésimos f, g en un mismo punto a, se dice que $f = o_{x \to a}(g)$ (f es un o pequeña de g) o bien que f es un infinitésimo de orden superior a g si para cada $\varepsilon > 0$, existe $\delta > 0$ tal que si $0 < d(x, a) < \delta$, se cumple $|f(x)| < \varepsilon |g(x)|$. Equivale a decir, si $g(x)$ es no nulo, que $\lim_{x \to a} \frac{f(x)}{g(x)} = 0$.*

Ejemplo 1.12. *Sean $f(x, y) = x^3 + y^3 + x^4$ y $g(x, y) = x^2 + y^2$. Se cumple que $f = o_{(x,y) \to (0,0)}(g)$. En efecto, si escribimos $x = \rho \cos \lambda$, $y = \rho \sin \lambda$ tendremos $g(x, y) = \rho^2$ y $f(x, y) = \rho^3 \cos^3 \lambda + \rho^3 \sin^3 \lambda + \rho^4 \cos^4 \lambda$. De aquí que para cada $\epsilon > 0$, se cumple $|f| \leq \epsilon |g|$ para ρ suficientemente pequeño.*

Una noción importante es la de límite según un subconjunto.

Definición 1.18. *Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m. Sea $E \subset D$ y sea a un punto de acumulación de E. Diremos que el límite de f cuando x tiende a a según el subconjunto E es l si para cada sucesión $\{x^i\} \to a$, $x^i \in E$, $x^i \neq a$, se cumple $\{f(x^i)\} \to l$. Lo escribiremos $\lim_{x \to a \ x \in E} f(x) = l$.*

Obsérvese que decir que una función tiene límite l según un subconjunto E equivale a decir que la función restricción al conjunto E tiene por límite l. Por tanto todas las propiedades que se han visto para los límites valdrán para estos límites según subconjuntos. En particular, podría haberse dado la definición de límite según un subconjunto en términos de entornos en lugar de definirlo a partir de los límites de sucesiones. Se cumplirá que $\lim_{x\to a\ x\in E} f(x) = l$ si para cada $\mathcal{E}(l,\varepsilon)$ existe $\delta > 0$ tal que $f(\mathcal{E}^*(a,\delta)\cap E) \subset \mathcal{E}(l,\varepsilon)$.

Un caso de particular interés es aquel en que el conjunto E está formado por los puntos de D que están en una recta de ecuación $x = a + \lambda u$. Se llama límite direccional en la dirección definida por el vector u.

Ejemplo 1.13. *Consideremos la función definida en $R^2 - \{(0,0)\}$ a valores en R mediante $f(x,y) = \frac{xy}{x^2+y^2}$. Hemos visto (ejemplo 1de 1.11) que no existe $\lim_{(x,y)\to(0,0)} f(x,y)$. Veamos que existen los límites direccionales. Tendremos*

$$\lim_{(x,y)\to(0,0)\ (x,y)=\lambda(u_1,u_2)} f(x,y) = \lim_{\lambda\to 0} \frac{\lambda^2 u_1 u_2}{\lambda^2 (u_1)^2 + \lambda^2 (u_2)^2} = \frac{u_1 u_2}{(u_1)^2 + (u_2)^2}.$$

La proposición siguiente da las relaciones básicas entre el límite y los límites según subconjuntos.

Teorema 1.22. *1. Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m. Sea E un subconjunto de D y a un punto de acumulación de E. Si existe $\lim_{x\to a} f(x)$, entonces existe $\lim_{x\to a\ x\in E} f(x)$ y coincide con él.*

2. Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m. Sean E_i un número finito de subconjuntos de D tales que $D = \cup_i E_i$. Si existe para cada i, $\lim_{x\to a\ x\in E_i} f(x)$ y todos ellos toman el mismo valor l, entonces existe $\lim_{x\to a} f(x)$ y vale l.

Demostración. El primer apartado es consecuencia inmediata de las definiciones. Probemos el segundo apartado. Dado $\varepsilon > 0$ existen $\delta_i > 0$ tales que si $0 < d(x,a) < \delta_i$, $x \in E_i$, se cumple $d(f(x), l) < \varepsilon$. Consideremos $\delta = \min\{\delta_i\}$ Es un número positivo por ser el mínimo de un número finito de números positivos. Tendremos que si $0 < d(x,a) < \delta$, $x \in D$, se cumplirá $d(f(x), l) < \varepsilon$.

□

Obsérvese que no podemos sustituir en el segundo apartado un número finito de subconjuntos E_i por una colección cualquiera de subconjuntos. El ejemplo 1.13 da una función para la que existen los límite direccionales pero no existe el límite. En este ejemplo los límites direccionales son distintos. Aunque los límites coincidan tampoco se puede asegurar la existencia del límite.

Ejemplo 1.14. *La función definida en $R^2 - \{(0,0)\}$ mediante $f(x,y) = \frac{x^3}{x^2-y^2}$ si $x \neq \pm y$, $f(x,\pm x) = 0$ tiene límites direccionales en $(0,0)$. Si $u_1 \neq \pm u_2$ se cumple*

$$\lim_{(x,y)\to(0,0)\ (x,y)=\lambda(u_1,u_2)} f(x,y) = \lim_{\lambda\to 0} \frac{\lambda^3 u_1}{\lambda^2 (u_1)^2 - \lambda^2 (u_2)^2} = 0.$$

Si $u_1 = \pm u_2$ el límite direccional es cero ya que la función sobre los puntos de la forma $(x,\pm x)$ toma el valor cero.

Hemos visto que existen todos los límites direccionales y coinciden. Sin embargo, no existe el límite. De lo contrario el límite según cualquier subconjunto debiera coincidir con él y por *tanto con* 0. *Sin embargo el límite según la parábola* $E = \{(x, y)\,;\ y = x + x^2\}$ *da*

$$\lim_{(x,y)\to(0,0)\ y=x+x^2} f(x, y) = \lim_{x\to 0} \frac{x^3}{x^2 - (x + x^2)^2} = -\frac{1}{2}.$$

1.3.2 Ejercicios

1. Calcula el límite en el punto $(0, 0)$ o prueba su no existencia en los siguientes casos:

   a) $f(x, y) = \frac{x^2y^2}{x^2+y^2}$ definida en $R^2 - \{(0, 0)\}$.

   b) $f(x, y) = \frac{x^4}{x^2+y^2}$ definida en $R^2 - \{(0, 0)\}$.

   c) $f(x, y) = \frac{(x+y)(x-y)^3}{(x-y)^2+(x+y)^2}$ definida en $R^2 - \{(0, 0)\}$.

   d) $f(x, y) = \frac{1-\cos(xy)}{x^2y^2}$ definida en $x \neq 0,\ y \neq 0$.

   e) $f(x, y) = \frac{\sin xy}{(x^2+y^2)^{\frac{1}{2}}}$ definida en $R^2 - \{(0, 0)\}$.

   f) $f(x, y) = \frac{\sin x \cdot \sin y}{x-y}$ definida en $x \neq y$.

   g) $f(x, y) = \frac{|x|-|y|}{|x|+|y|}$ definida en $R^2 - \{(0, 0)\}$.

2. Halla el límite en el punto $(0, 0)$ de la función $f(x, y) = \frac{x^a y^b}{x^2+y^2-2xy}$ definida para $x \neq y$ según los diversos valores de a, b positivos.

3. Halla los límites, si existen, según los diferentes subconjuntos en los siguientes casos:

   a) $f(x, y) = \frac{e^{\left(x^2\right)}-1}{y}$ definida para $y \neq 0$ según el subconjunto $y = x^2$ en el punto $(0, 0)$.

   b) $f(x, y) = \frac{x-\sin y}{x^3}$ definida en $x \neq 0$ según el subconjunto $y = 2x^2$ en el punto $(0, 0)$.

   c) $f(x, y) = \frac{1-\cos x}{e^y-1-x}$ definida en $e^y - x - 1 \neq 0$ en el punto $(0, 0)$.

4. a) Sea una función f definida en un conjunto $D \subset R^2$ a valores en R tal que exista $\lim_{(x,y)\to(a,b)} f(x, y) = l$. Supongamos que para cada x existe $\lim_{y\to b} f(x, y)$ y para cada y existe $\lim_{x\to a} f(x, y)$. Prueba que

   $$\lim_{x\to a}\left(\lim_{y\to b} f(x, y)\right) = \lim_{y\to b}\left(\lim_{x\to a} f(x, y)\right) = l.$$

   b) Aplica la proposición anterior para probar que la función $f(x, y) = \frac{x^2-y^2}{x^2+y^2}$ no tiene límite en $(0, 0)$.

   c) Comprueba que la función $f(x, y) = \frac{xy}{x^2+y^2}$ tiene límites iterados que coinciden y, sin embargo, no tiene límite en $(0, 0)$.

5. Halla, en los diferentes casos, los valores de $m \in N$ tales que:

   a) $\tan^4 xy = o_{(x,y)\to(0,0)}\left(\left(x^2+y^2\right)^m\right)$.

   b) $1 - e^{x^2y^2} = o_{(x,y)\to(0,0)}\left(\left(x^2+y^2\right)^m\right)$.

6. Sea $f : R^n \to R^n$ y un punto $a \in R^n$ tal que $\lim_{x\to a} f(x) = l$ y $f(x) \neq l$ para $x \neq a$. Sea $g : R^n \to R$ tal que $\lim_{x\to l} g(x) = L$. Prueba que $\lim_{x\to a} g \circ f = L$. Da un ejemplo que muestre que no ocurre esto en general si suprimimos la hipótesis de que $f(x) \neq l$ para $x \neq a$.

7. Halla valores de α y β para que se cumpla

$$x^3 + y^3 - 1 - \alpha(x-1) - \beta y = o_{(x,y)\to(0,0)}\left(\left((x-1)^2 + y^2\right)^{\frac{1}{2}}\right).$$

8. Sea f una función definida en $D \subset R^n$ a valores en R y a un punto de acumulación de D. Prueba que si para cada sucesión $\{x^i\} \to a$, $x^i \in D$, $x^i \neq a$, la sucesión $\{f(x^i)\}$ tiene límite, entonces existe $\lim_{x\to a} f(x)$. Indicación: Prueba que todos los límites de las sucesiones $\{f(x^i)\}$ coinciden.

9. Prueba el siguiente criterio de Cauchy para la existencia de límite de una función. Sea f una función definida en $D \subset R^n$ a valores en R y a un punto de acumulación de D. Prueba que existe $\lim_{x\to a} f(x)$ si y sólo si para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si $d(x,a) < \delta$, $d(y,a) < \delta$, $x, y \in D$, $x, y \neq a$ se cumple $|f(x) - f(y)| < \varepsilon$. Indicación: Si existe el límite se verifica directamente la condición. Si se cumple la condición comprueba que estamos en las condiciones del ejercicio anterior.

1.4 Funciones continuas

1.4.1 Funciones continuas en un punto

Sea un punto del dominio de definición de una función. La idea intuitiva de función continua en este punto, es la de una función que a pequeñas variaciones de la variable corresponden variaciones arbitrariamente pequeñas de la función. Si la función está definida en D y el punto considerado es aislado en D no existen de hecho pequeñas variaciones de la variable y la función siempre se considera continua en estos puntos. El concepto es significativo cuando el punto es de acumulación de D. En este caso equivale a decir que el límite coincide con el valor de la función en el punto. Tendremos las siguientes definiciones.

Definición 1.19. *1. Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m y $a \in D$. Si a es un punto de acumulación de D, existe $\lim_{x\to a} f(x)$ y coincide con $f(a)$ diremos que la función es continua en a. Si a es aislado diremos siempre que f es continua en a.*

2. *Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m y $a \in D$. Diremos que f es continua en a si para cada $\varepsilon > 0$ existe $\delta > 0$ tal que*

$$f\left(D \cap \mathcal{E}\left(a, \delta\right)\right) \subset \mathcal{E}\left(f\left(a\right), \varepsilon\right).$$

3. *Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m y $a \in D$. Diremos que f es continua en a si para cada sucesión $\{x^i\} \to a,\ x^i \in D$ se cumple $\{f(x^i)\} \to f(a)$.*

Teorema 1.23. *Las tres definiciones anteriores son equivalentes.*

Demostración. El caso en que el punto a es aislado es trivial ya que la única sucesión que cumple $\{x^i\} \to a,\ x^i \in D$ es la que es constante igual a a a partir de un término y 1 y 2 no imponen ninguna condición sobre la función en este caso.

Consideremos ahora que a es un punto de acumulación de D. La prueba de la equivalencia entre las definiciones es prácticamente la misma que la de la equivalencia entre las definiciones de límite que ya hemos dado (Teorema 1.17) y por ello no repetiremos los detalles. Las únicas diferencias son que al ser el límite $f(a)$ en la condición 2 no es preciso evitar la imagen del punto a y en la condición 3 no imponemos que los puntos de la sucesión x^i sean distintos de a. □

Ejemplo 1.15. 1. *La función $f(x,y) = \frac{x^4}{x^2+y^2}$ si $(x,y) \neq (0,0)$ y $f(0,0) = 0$ es continua en $(0,0)$. En efecto, dado $\varepsilon > 0$ existe $\delta = \varepsilon^{\frac{1}{2}}$ tal que si $d((x,y),(0,0)) < \delta$ se cumple si $(x,y) \neq (0,0)$, $\left|\frac{x^4}{x^2+y^2}\right| < \rho^2 < \delta^2 = \varepsilon$ y $|f(0,0)| = 0$.*

2. *Sea $f(x,y) = p(x,y)$ donde p es un polinomio. Es continua en todos los puntos. En efecto, sea $\{(x^i, y^i)\} \to (a_1, a_2)$. Se tendrá que $\{x^i\} \to a_1$ y $\{y^i\} \to a_2$ y, por tanto, $p(x^i, y^i) \to p(a_1, a_2)$.*

3. *En el ejemplo 1.14 se ha dado una función que no es continua en $(0,0)$ ya que no existe el límite en este punto.*

4. *El espacio $R^n \times R^n$ puede identificarse en forma natural con R^{2n}. Veamos que la función definida en $R^n \times R^n$ a valores en R^n mediante $f(x,y) = x + y$ es continua en todos los puntos. Obsérvese que ahora x e y indican puntos de R^n. En efecto, si $\{(x^i, y^i)\} \to (x,y)$ se tiene que $x^i + y^i \to x + y$. Para comprobarlo es suficiente ver que para cada una de las componentes coordenadas $x_k^i + y_k^i \to x_k + y_k$. Esto a su vez se deduce de las propiedades de las sucesiones de números reales.*

Obsérvese que la condición de continuidad de una función en un punto es de tipo local, es decir, el que una función sea o no continua en un punto depende del comportamiento de la función únicamente en un entorno del punto.

Las propiedades algebraicas que hemos dado sobre los límites de funciones (teorema 1.20) dan inmediatamente el siguiente teorema.

Teorema 1.24. *Sean f y g dos funciones definidas en un conjunto $D \subset R^n$ a valores en R^m continuas en un punto a. Sea λ un número real. Entonces las funciones $f+g$ y λf son continuas en a. Si $m=1$ el producto fg es continuo en a. Si, además, $g(x)$ no se anula en ningún punto la función $\frac{f}{g}$ es continua en a.*

Ejemplo 1.16. *La función $f(x,y) = \frac{p(x,y)}{q(x,y)}$ donde p y q son polinomios está definida en el conjunto de los puntos donde no se anula el denominador. En todos estos puntos la función es continua por ser cociente de funciones continuas.*

Veamos el comportamiento de la continuidad respecto a la composición de funciones.

Teorema 1.25. *Sea f una función definida en $D \subset R^n$ a valores en R^m. Sea g una función definida en un conjunto $E \subset R^m$ que contenga a $f(D)$. Supongamos que f es continua en un punto $a \in D$ y que g es continua en $b = f(a)$. Entonces la función compuesta $g \circ f$ es continua en a.*

Demostración. Por ser g continua en $f(a)$ para cada $\varepsilon > 0$ existe $\delta > 0$ tal que

$$g(\mathcal{E}(f(a),\delta) \cap E) \subset \mathcal{E}(g(f(a)),\varepsilon).$$

Por la continuidad de f en a existirá $\eta > 0$ tal que

$$f(\mathcal{E}(a,\eta) \cap D) \subset \mathcal{E}(f(a),\delta).$$

De las dos relaciones obtenemos que

$$g(f(\mathcal{E}(a,\eta) \cap D)) \subset \mathcal{E}(g(f(a)),\varepsilon)$$

y, por tanto, obtenemos la continuidad de $g \circ f$ en a.

□

Ejemplo 1.17.

1. *Si f es una función de una variable, continua en todos los puntos, y $p(x,y)$ es un polinomio, la función $f(p(x,y))$ es continua en todos los puntos.*

2. *El teorema 1.24 puede obtenerse como consecuencia de la proposición anterior. Por ejemplo, veamos que la suma de dos funciones continuas f y g en un punto es una función continua en dicho punto. Hemos visto que la función h definida en $R^m \times R^m$ a valores en R^m mediante $h(x,y) = x+y$ es una función continua en todos los puntos. La función de R^n en $R^m \times R^m$ que asigna a x el punto $(f(x), g(x))$ es continua en el punto considerado. Por tanto, la función compuesta $f(x)+g(x)$ es una función continua en dicho punto.*

1.4.2 Funciones continuas en un conjunto

Hemos visto hasta el momento propiedades de la continuidad de las funciones en un punto. Se trata ahora de estudiar propiedades que se deducen de la continuidad de una función en todos los puntos del dominio de definición.

Definición 1.20. *Sea D un subconjunto de R^n y f una función definida en D a valores en R^m. Se dice que es continua en D si es continua en todos los puntos del conjunto D.*

Veamos un teorema que da una condición equivalente a la continuidad en un conjunto en términos de los conjuntos abiertos o de los conjuntos cerrados.

Teorema 1.26. *Sea D un subconjunto de R^n y f una función definida en D a valores en R^m. Son equivalentes las siguientes proposiciones*

1. f es continua en D.
2. Para cada abierto $A \subset R^m$ existe un abierto U de R^n tal que $f^{-1}(A) = D \cap U$.
3. Para cada cerrado $C \subset R^m$ existe un cerrado F de R^n tal que $f^{-1}(C) = D \cap F$.

Los conjuntos de la forma $D \cap U$ con U un abierto de R^n se denominan abiertos relativos a D. Entonces 2 se puede enunciar diciendo que la antiimagen de todo abierto debe ser un abierto relativo a D. Obsérvese que si D es un conjunto abierto, los abiertos relativos a D son los abiertos contenidos en D. En otro caso puede haber otros conjuntos. Por ejemplo, si $D = [0, 1]$ un conjunto de la forma $\left[0, \frac{1}{2}\right)$ es un abierto relativo a D pues es la intersección de un abierto como $(-1, 1/2)$ con D. En el caso en que D sea abierto la condición 2 es, simplemente, que la antiimagen de todo abierto es un abierto. Pueden hacerse análogas consideraciones para los conjuntos de la forma $D \cap F$ con F cerrado que reciben el nombre de cerrados relativos a D. Ahora si D es cerrado los cerrados relativos a D son los cerrados contenidos en D.

Demostración. Veamos que 1 implica 2. Sea A un abierto de R^m y sea $p \in D$ con $f(p) \in A$. Existirá $\varepsilon > 0$ tal que $\mathcal{E}(f(p), \varepsilon) \subset A$. Por ser f continua en p sea $\delta_p > 0$ tal que

$$f(\mathcal{E}(p, \delta_p) \cap D) \subset \mathcal{E}(f(p), \varepsilon) \subset A.$$

Haciendo lo mismo con todos los puntos de A tendremos que el conjunto $U = \cup_{p,\ f(p) \in A} \mathcal{E}(p, \delta_p)$ es un abierto tal que $f^{-1}(A) = D \cap U$.

Veamos que 2 implica 1. Sea $p \in D$. Se trata de ver que f es continua en este punto. Dado $\varepsilon > 0$, el conjunto $\mathcal{E}(f(p), \varepsilon)$ es un conjunto abierto. Por la condición 2 existirá un abierto U de R^n tal que

$$f^{-1}(\mathcal{E}(f(p), \varepsilon)) = U \cap D.$$

Por ser U abierto existirá $\delta > 0$ tal que $\mathcal{E}(p, \delta) \subset U$. De ambas relaciones se deduce

$$f(\mathcal{E}(p, \delta) \cap D) \subset \mathcal{E}(f(p), \varepsilon)$$

y, por tanto, la continuidad de f en p.

Veamos que 2 implica 3. Sea C un cerrado de R^m. Su complementario será un abierto $A = R^m - C$ al que podremos aplicar la condición 2. Existirá un abierto U de R^n tal que $f^{-1}(A) = D \cap U$. Tendremos entonces

$$f^{-1}(C) = D - f^{-1}(A) = D - D \cap U = D \cap (R^n - U).$$

Puesto que $R^n - U$ es un conjunto cerrado, vale la condición 3.

La demostración de que 3 implica 2 se hace como en el caso anterior teniendo en cuenta que pasamos de abiertos a cerrados por paso al complementario.

$\square$

Ejemplo 1.18. *Veamos que el conjunto*

$$U = \left\{(x, y) \in R^2;\ p(x, y) < k\right\}$$

es un conjunto abierto si p es un polinomio. En efecto, la función $p(x, y)$ es continua en R^2 y el conjunto U es la antiimagen por p del conjunto abierto $(-\infty, k)$.

Definición 1.21. *Una aplicación biyectiva de un conjunto $D \subset R^n$ en un conjunto $E \subset R^m$ que sea continua con aplicación inversa continua se llama un homeomorfismo de D en E.*

La topología en R^n es el estudio de las propiedades que permanecen invariantes por los homeomorfismos de R^n consigo mismo. Así, la noción de sucesión que tiene límite o sucesión convergente es un concepto topológico. Una sucesión tiene límite si y sólo si su imagen por un homeomorfismo también la tiene. El concepto de punto interior de un conjunto también lo es. Un punto es interior a un conjunto si y sólo si su transformado por un homeomorfismo es interior al conjunto transformado. En cambio, por ejemplo, la propiedad de que un subconjunto sea una circunferencia no permanece invariante por un homeomorfismo. Es una propiedad de tipo métrico, sólo es invariante por las transformaciones que conserven la distancia, y no es una propiedad de tipo topológico que son las estamos interesados en este momento.

Veamos ahora algunas de las propiedades básicas de las funciones continuas en conjuntos compactos. La primera permitirá afirmar que la compacidad es un concepto topológico.

Teorema 1.27. *Sea f una función definida en un conjunto compacto K de R^n a valores en R^m y continua en K. Entonces $f(K)$ es un conjunto compacto.*

Demostración. Dada la equivalencia entre conjuntos compactos y compactos por sucesiones (teorema 1.16) podemos probar cualquiera de las dos propiedades. Probemos por ejemplo la última. Sea una sucesión de elementos de $f(K)$. Será de la forma $\{f(x^i)\}$ para alguna sucesión $\{x^i\}$ de elementos de K. Por ser este conjunto compacto por sucesiones existe una sucesión parcial convergente a un punto de K. Sea $\{x^{i_j}\} \to p \in K$. Por la continuidad de la función en p, tendremos $\{f(x^{i_j})\} \to f(p) \in f(K)$. Por tanto, $f(K)$ es compacto por sucesiones.

$\square$

Puede darse una demostración alternativa utilizando la propiedad de compacidad por recubrimientos y la caracterización de las funciones continuas en términos de las antiimágenes de los abiertos. Veamos brevemente como puede hacerse. Sea $A_i, \ i \in I$, un recubrimiento abierto de $f(K)$. Se tiene que $f^{-1}(A_i) = K \cap U_i$ para ciertos abiertos U_i de R^n. La colección de $U_i \ i \in I$ forma un recubrimiento abierto de K. Sea $U_i \ i \in J \subset I$ un recubrimiento parcial y finito de K. La colección de $A_i \ i \in J$ constituirá un recubrimiento del conjunto $f(K)$ parcial y finito del dado. Esto da la compacidad de este conjunto.

Corolario 1.28. *La imagen por una función continua de un conjunto cerrado y acotado es un conjunto cerrado y acotado.*

Demostración. Es suficiente tener en cuenta el teorema anterior y la caracterización de los compactos como los cerrados y acotados (teorema 1.15).

□

Debe observarse que la imagen de un conjunto acotado por una aplicación continua en este conjunto no tiene por qué serlo. No obstante, si la función continua está definida no sólo en el conjunto sino en todo el espacio R^n, entonces la imagen de un acotado B es un acotado ya que $\overline{B}$ sería compacto y por tanto su imagen sería un compacto. En particular sería un conjunto acotado y $f(B)$, por estar contenido en $f\left(\overline{B}\right)$ estará acotado.

La imagen por una aplicación continua de un conjunto cerrado no tiene por qué ser cerrado.

Ejemplo 1.19.

1. *Consideremos la función continua de una variable* $f(x) = \frac{1}{x^2}$ *definida en* $(0,1)$. *Su recorrido es* $(1, +\infty)$ *que no es un acotado a pesar de serlo* $(0,1)$.

2. *La imagen por la función continua* $f(x) = \frac{1}{1+x^2}$ *del cerrado* R *es el conjunto* $(0,1]$ *que no es cerrado.*

Máximos y mínimos de funciones continuas

Un problema importante es saber si dada una función definida en $D \subset R^n$ a valores reales, existe un punto $a \in D$ que cumple que $f(x) \leq f(a)$ (resp. $f(x) \geq f(a)$) para todo $x \in D$. Un punto con esta propiedad se dice que es un máximo (resp. mínimo) absoluto de la función o simplemente máximo (resp. mínimo). Es bien conocido que una función continua no tiene por qué tener máximos o mínimos. Por ejemplo, la función de una variable real $f(x) = \frac{1}{x}$ definida y continua en $R - \{0\}$ no tiene máximo ni mínimo. Sin embargo si el dominio de definición es compacto si que podemos asegurar la existencia de máximo y de mínimo.

Teorema 1.29. *Sea f una función definida en un conjunto compacto K de R^n a valores reales y continua en K. Entonces f tiene un máximo y un mínimo.*

Demostración. Probaremos la existencia de máximo. La existencia de mínimo puede hacerse en forma análoga o bien aplicar la existencia de un máximo a la función $-f$.

Hemos visto que el conjunto imagen $f(K)$ es un conjunto de números reales compacto (teorema 1.27) y por tanto cerrado y acotado (teorema 1.15). Existirá en particular el supremo S de este conjunto. Lo que queremos probar es que este supremo es el máximo, es decir que es alcanzado en algún punto. Por ser S el supremo de $f(K)$ será adherente al conjunto ya que para cada $\varepsilon > 0$, $S - \varepsilon$ no es cota superior de $f(K)$ y existirán elementos de este conjunto en $(S - \varepsilon, S + \varepsilon)$. Por ser el conjunto $f(K)$ cerrado, $S \in f(K)$ y por tanto S es la imagen de algún punto de K.

$\square$

Corolario 1.30. *Sea f una función definida en un conjunto cerrado y acotado K de R^n a valores reales y continua en K. Entonces f tiene un máximo y un mínimo.*

Continuidad de la función inversa

El siguiente teorema es útil para establecer la continuidad de las funciones inversas de ciertas funciones continuas.

Teorema 1.31. *Sea f una función inyectiva, definida y continua en un conjunto compacto K de R^n, a valores en R^m. Entonces la aplicación inversa f^{-1} definida en $f(K)$ es también continua.*

Demostración. Llamemos g a la aplicación inversa f^{-1}. El dominio de definición de g es $f(K)$. Por ser compacto es en particular cerrado. Puesto que los cerrados relativos a éste conjunto son cerrados deberemos probar que la antiimagen por g de todo cerrado F de R^n es un cerrado. Ahora bien $g^{-1}(F) = g^{-1}(F \cap K) = f(F \cap K)$. Puesto que $F \cap K$ es un cerrado y acotado será un compacto y su imagen por f será, en consecuencia, compacto. En particular $g^{-1}(F)$ será un cerrado.

$\square$

Ejemplo 1.20. *Consideremos la función inyectiva definida en $[0, \pi] \times [1, 2]$ que a (φ, ρ) le corresponde $(\rho \cos \varphi, \rho \sin \varphi)$. Es continua y definida en un compacto, por tanto la aplicación inversa será continua.*

1.4.3 Curvas y conexión

Definición 1.22. *Un arco de curva, o simplemente curva, en un conjunto D de R^n es una aplicación continua $\gamma : [0, 1] \to D$. Si $\gamma(0) = a$ y $\gamma(1) = b$ se dice que a y b son el origen y el extremo respectivamente del arco de curva.*

Definición 1.23. *Un conjunto D se dice que es arcoconexo si para cada dos puntos de éste existe un arco de curva que tiene uno como origen y el otro como extremo.*

Ejemplo 1.21. *Un conjunto tal que con cada dos puntos contiene el intervalo que los une se dice que es un convexo. Entonces todo convexo es un arcoconexo. Por ejemplo, todo entorno circular es un conjunto arcoconexo.*

Consideremos A un conjunto abierto de R^n y a un punto de A. Se trata de ver cómo es el conjunto de puntos que pueden unirse a a mediante un arco de curva en A, es decir, aquellos puntos b para los que existe un arco de curva en A tal que a es el origen y b es el extremo.

Teorema 1.32. *Sea A un abierto de R^n y a un punto de A. El conjunto de los puntos que pueden unirse a a mediante un arco de curva en A es un conjunto abierto. El conjunto de los puntos que no se pueden unir a a es también un conjunto abierto.*

Demostración. Sea B el subconjunto de A de los puntos que pueden unirse a a mediante un arco de curva y sea $b \in B$. Existirá $\mathcal{E}(b, \varepsilon) \subset A$. Veamos que también $\mathcal{E}(b, \varepsilon)$ está contenido en B con lo que este conjunto será abierto. En efecto, si γ es el arco que une a con b y $c \in \mathcal{E}(b, \varepsilon)$, un arco de curva que une a y c puede construirse uniendo el arco γ con un segmento que une b y c . Explícitamente, mediante un arco de la forma $\gamma_1(t) = \gamma(2t)$ para $0 \leq t \leq 1/2$ y $\gamma_1(t) = b + 2(t - 1/2)(c - b)$ para $1/2 \leq t \leq 1$.

Análogamente, sea D el conjunto de los puntos de A que no pueden unirse a a mediante un arco en A y sea $d \in D$. Supongamos que $\mathcal{E}(d, \varepsilon) \subset A$. Ningún punto de $\mathcal{E}(d, \varepsilon)$ podrá unirse con a mediante un arco en A ya que de lo contrario uniendo éste, como antes, con el segmento de origen este punto y extremo d obtendríamos un arco que une a y d y llegaríamos a contradicción. □

Observamos, en consecuencia que, dado un abierto y un punto del mismo tenemos una partición del mismo en dos conjuntos abiertos disjuntos. Los que pueden unirse al punto mediante un arco en el conjunto y los que no pueden hacerlo. Si un abierto no es arcoconexo ambos conjuntos para un cierto punto serán no vacíos. Esto sugiere la siguiente definición.

Definición 1.24. *Un conjunto D de R^n se dice que no es conexo si existe una partición del mismo en dos abiertos relativos a D no vacíos.*

Si se trata de un conjunto abierto no será conexo si existe una partición en dos abiertos no vacíos. Será conexo cuando no exista esta partición. Veamos que para los conjuntos abiertos este concepto y el de arcoconexión son equivalentes.

Teorema 1.33. *Un conjunto abierto A es conexo si y sólo si es arcoconexo.*

Demostración. Hemos visto ya que si A no es arcoconexo existe una partición del conjunto en dos abiertos no vacíos y por tanto no es conexo. Recíprocamente, supongamos ahora que A es arcoconexo. Deberá ser conexo ya que de lo contrario tendríamos $A = A_1 \cup A_2$ con A_i disjuntos y no vacíos. Sean $a \in A_1$ y $b \in A_2$. Consideremos γ un arco en A de origen a y extremo b. Se tendrá que $\gamma^{-1}(A_1)$ y $\gamma^{-1}(A_2)$ serán una partición del intervalo $[0, 1]$ en abiertos no vacíos relativos a este conjunto. Esto no puede ocurrir. Si por ejemplo $0 \in \gamma^{-1}(A_1)$ el conjunto de los puntos x tales que $[0, x] \subset \gamma^{-1}(A_1)$ tendrá un supremo que no estará ni en $\gamma^{-1}(A_1)$ ni en $\gamma^{-1}(A_2)$ como no es difícil de comprobar. □

1.4.4 Ejercicios

1. Estudia los puntos de continuidad y de discontinuidad de las siguientes funciones:
   a) $f(x, y) = \frac{x^2+y^2}{x-y}$ si $x \neq y$, y $f(x, x) = 0$.
   b) $f(x, y) = \frac{y}{x} \sin(x^2 + y^2)$ si $x \neq 0$, $f(0, y) = 0$.
   c) $f(x, y) = x \sin \frac{1}{x^2+y^2}$ si $(x, y) \neq (0, 0)$, $f(0, 0) = 0$.
   d) $f(x, y) = x \cdot g(xy)$ donde $g(z) = 0$ si $z \in Q$ y $g(z) = 1$ si $z \notin Q$.

e) $f(x,y) = g(xy)$ definida para $x > 0,\, y > 0$ con $g(z) = 0$ si $z \notin Q$ y $g(z) = \frac{1}{m}$ si $z = \frac{n}{m}$ donde esta fracción es irreducible.

2. Di si existen y en su caso da un ejemplo:

   a) Una función continua exhaustiva de $(0,1)$ en $[-1,1]$.

   b) Una función continua exhaustiva de $(0,1) \times [0,1)$ en $[-1,1] \times (-1,0]$.

   c) Una función continua exhaustiva de $[0,1] \times [0,1]$ en $[-1,1] \times (0,1]$.

   d) Una función continua exhaustiva de $[-1,1] \times [-1,1]$ en

   $\{(x,y) \in R^2;\, x^2 + y^2 \geq 1\}$.

3. Averiguar cuales de los siguientes conjuntos son compactos: a) $\{(x,y) \in R^2; xy = 1\}$.

   b) $\{(x,y,z) \in R^3;\, 6x^2 + 7y^2 + 3z^2 \leq 1\}$.

4. Una función $f : R^n \to R$ se dice que cumple una condición de Lipschitz de orden c en un punto $a \in R^n$ si existe un entorno de a y una constante K tal que para toda x de este entorno $|f(x) - f(a)| \leq K \|x - a\|^c$. Prueba que una tal función es continua en a. Da un ejemplo, para $n = 1$, de una función continua en a pero que no cumple ninguna condición de Lipschitz en a.

5. Sea $A \subset R^n$. Prueba que la función $f : R^n \to R$ definida por

$$f(x) = \inf_{y \in A} d(x,y)$$

   es una función continua. Se denotará $f(x) = d(x,A)$.

6. Sea K un compacto de R^n y $a \notin K$. Prueba que existe un punto $b \in K$ tal que $d(a,b) = d(a,K)$. ¿Puede decirse lo mismo para un cerrado de R^n? ¿Y para un conjunto cualquiera?

7. Sean K_1 y K_2 dos compactos disjuntos de R^n. Prueba que existen $a_1 \in K_1$ y $a_2 \in K_2$ tales que $d(a_1,a_2) = \inf\{d(x,y);\, x \in K_1, y \in K_2\}$.

8. Consideremos la aplicación $\pi : R^2 \to R$ tal que $\pi(x,y) = x$. Prueba que la imagen de un abierto es un abierto. Comprueba que la imagen por π de un cerrado no es necesariamente cerrado.

9. Sea $f : R \to R$ continua. Prueba que el gráfico $G = \inf\{(x,y) \in R^2; y = f(x)\}$ es un conjunto cerrado. Da un ejemplo de una función en que su gráfico sea cerrado y que sin embargo no sea continua.

10. Prueba que $f : R^n \to R^m$ es continua si y sólo si para cada $B \subset R^m$ se cumple

$$f^{-1}(B^\circ) \subset \left(f^{-1}(B)\right)^\circ.$$

11. Sea f es una función continua definida en un arcoconexo D y dos puntos $a, b \in D$ tales que $f(a)$ y $f(b)$ tienen signos distintos. Prueba que existe algún punto $x \in D$ tal que $f(x) = 0$.

12. Da un ejemplo de un abierto arcoconexo que no sea convexo.

13. Prueba que la imagen por una función continua de un conjunto conexo es un conjunto conexo.

1.5 Funciones uniformemente continuas.

1.5.1 Conceptos y propiedades fundamentales

Sea f una función continua en un conjunto $D \subset R^n$ a valores en R^m. Esto significa que es continua en cada punto de D. Es decir para cada $a \in D$ y cada $\varepsilon > 0$ existe $\delta > 0$ tal que si $d(x, a) < \delta$, $x \in D$, entonces $d(f(x), f(a)) < \varepsilon$. Obsérvese que δ depende desde luego de ε pero también depende del punto $a \in D$ que se esté considerando. Es interesante la propiedad de que δ pueda tomarse únicamente dependiente de ε y no del punto a. Son las llamadas funciones uniformemente continuas.

Definición 1.25. *Sea f una función definida en un conjunto $D \subset R^n$ a valores en R^m. Se dice que es uniformemente continua si para cada $\varepsilon > 0$, existe $\delta > 0$ tal que si $x, y \in D$, $d(x, y) < \delta$ entonces $d(f(x), f(y)) < \varepsilon$.*

Este concepto, aunque en un principio no parezca demasiado diferente del de función continua, si que lo es. Digamos, por ejemplo, que no puede darse una caracterización análoga a la que dimos para las funciones continuas en términos de conjuntos abiertos. De hecho, el concepto de función uniformemente continua no es un concepto topológico. Si se compone una función uniformemente continua con un homeomorfismo no se obtiene en general una función uniformemente continua.

Ejemplo 1.22. *1. La función $\sin x$ definida en R es una función uniformemente continua ya que, aplicando el teorema del valor medio*

$$|\sin x - \sin y| = |(x - y)\cos\lambda| \leq |x - y|.$$

De aquí que dado $\varepsilon > 0$, tomando $\delta = \varepsilon$, si $|x - y| < \delta$, se cumple $|\sin x - \sin y| < \varepsilon$.

2. la función $y = x^2$ definida en R no es uniformemente continua ya que para $\varepsilon > 0$ no existe ningún $\delta > 0$ tal que $|x - z| < \delta$ implique $|x^2 - z^2| < \varepsilon$. En efecto, tomando $|x - z| = \frac{\delta}{2}$

$$\left|x^2 - z^2\right| = |x - z|\,|x + z| = \frac{\delta}{2}\,|x + z|.$$

Si x y z son suficientemente grandes, aunque $|x - z| < \delta$ se cumplirá $|x^2 - z^2| \geq \varepsilon$. Hemos probado la no continuidad uniforme de la función. Obsérvese que, sin embargo, la función es continua.

Un caso especialmente importante en que la continuidad implica la continuidad uniforme es el siguiente.

Teorema 1.34. *Sea f una función continua definida en un conjunto compacto $K \subset R^n$ a valores en R^m. Entonces la función es uniformemente continua en K.*

Demostración. Sea $\varepsilon > 0$. En cada $a \in K$ la función es continua. Existe entonces δ_a tal que si $x \in K$ y $d(x,a) < 2\delta_a$ se cumple $d(f(x), f(a)) < \frac{\varepsilon}{2}$. Haciendo lo propio con cada punto $a \in K$ tendremos que los conjuntos $\mathcal{E}(a, \delta_a)$ forman un recubrimiento abierto de K. Puesto que este es un conjunto compacto existirá un recubrimiento parcial y finito $K \subset \mathcal{E}(a^1, \delta_{a^1}) \cup \ldots \cup \mathcal{E}(a^r, \delta_{a^r})$. Consideremos $\delta = \min\{\delta_{a^1}, \ldots, \delta_{a^r}\}$. Tomemos $x, y \in K$ tales que $d(x,y) < \delta$. Supongamos que $x \in \mathcal{E}(a^i, \delta_{ai})$. Se cumplirá

$$d\left(y, a^i\right) \leq d(y,x) + d\left(x, a^i\right) < \delta + \delta_{a^i} \leq 2\delta_{a^i}.$$

Por lo tanto

$$d\left(f(x), f\left(a^i\right)\right) < \frac{\varepsilon}{2} \text{ y } d\left(f(y), f\left(a^i\right)\right) < \frac{\varepsilon}{2}$$

y, en consecuencia,

$$d(f(x), f(y)) < \varepsilon.$$

□

Ejemplo 1.23. *La función $y = x^2$ es uniformemente continua en cada intervalo.*

1.5.2 Ejercicios

1. Prueba que toda función derivable de R en R, con derivada acotada, es uniformemente continua.

2. Prueba que la función $f(x) = \sin\frac{1}{x}$ es una función continua pero no uniformemente continua en $(0,1)$.

3. Sea f una función uniformemente continua en un conjunto $D \subset R^n$ a valores en R^m. Sea g uniformemente continua en $f(D)$ a valores en R. Prueba que la función compuesta $g \circ f$ es uniformemente continua en D.

4. Prueba que una función uniformemente continua transforma sucesiones de Cauchy en sucesiones de Cauchy.

5. Prueba que toda función uniformemente continua $f : D \to R$ se puede extender a una función uniformemente continua en $\overline{D}$. Indicación: Si $\{x^i\} \to a,\ x^i \in D$, prueba que $\{f(x^i)\}$ tiene límite verificando que es de Cauchy. Este límite será la extensión de f a a.

1.6 Nota histórica

Entre los iniciadores del rigor en el Análisis cabe destacar a Bolzano y Cauchy. Bolzano (1781-1848) dió la primera definición (1817) de función continua en un intervalo como una función f tal que la diferencia $f(x+w) - f(x)$ puede ser tan pequeña como se desee tomando w suficientemente pequeño. Probó la continuidad de los polinomios. Estableció el teorema de que una función continua en un intervalo que toma valores de signos opuestos en los extremos se anula en algún punto utilizando el criterio de convergencia de Cauchy. Bolzano en 1834 aclaró la diferencia entre continuidad y derivabilidad dando incluso una función continua en un intervalo cerrado no derivable en ningún punto. Sin embargo, sus trabajos no fueron conocidos hasta finales del siglo XIX y la primera publicación de una función continua tal que en todo intervalo tenía infinitos puntos en que no era derivable fue debida a Riemann en 1868.

Cauchy (1789-1857) en sus Lecciones sobre el cálculo infinitesimal (1823) desarrolló el cálculo diferencial e integral sobre la base del concepto de límite. Definió el concepto de límite, el de infinitesimal, clarificando el concepto de infinitesimal de Leibniz, y el de función continua. Sus definiciones, al estilo de Bolzano se basan en consideraciones puramente aritméticas.

Weierstrass (1815-1897) dio la actual definición de función continua en términos de ε y δ. Sus trabajos, si bien anteriores, fueron conocidos con ocasión de sus clases en la Universidad de Berlin a partir de 1859. Utilizó la idea de Bolzano de subdivisión sucesiva de intervalos para probar que si tenemos un conjunto acotado con infinitos puntos existe uno tal que en todo entorno de él existen infinitos puntos del conjunto. Es el que actualmente se conoce como teorema de Bolzano -Weierstrass. Cauchy había utilizado, sin probarlo, la existencia de un mínimo para toda función continua definida en un intervalo. Weierstrass probó la existencia de mínimo y de máximo para funciones continuas de una y de varias variables.

Heine (1821-1881) definió el concepto de continuidad uniforme. Probó que toda función continua en un intervalo cerrado es uniformemente continua. En su método utilizó el teorema de que para cada recubrimiento por intervalos abiertos de un intervalo cerrado existe un recubrimiento parcial y finito. Este teorema, cuando el recubrimiento de partida es numerable, fue probado por Borel (1871-1956). La primera publicación de su generalización a una colección cualquiera se debe a Cousin (1867-1933) en 1895.

El estudio de la topología como rama independiente probablemente proviene de Riemann (1826-1886). La idea de la topología como la geometría de los invariantes por las transformaciones continuas, aparece ya explicitada por Klein (1849-1925) en el Programa de Erlangen en 1872. Antes de poder proceder a este estudio fue necesario formalizar la teoría de los números reales y la teoría de conjuntos sobre el plano y el espacio.

Cantor (1845-1918), el creador de la teoría de conjuntos, fue el primero en definir los conceptos de punto de acumulación, punto aislado, cerrado y abierto en la recta y en el espacio n-dimensional. Obtuvo la estructura de estos conjuntos en la recta.

Hacia finales del siglo XIX se amplían estos conceptos sobre conjuntos de puntos a conjuntos de curvas o de funciones, fundamentalmente en el cálculo de variaciones, en la teoría de variable compleja y en el estudio de ecuaciones funcionales. Por otra parte las teorías axiomáticas van teniendo mayor relevancia sobre todo a partir de los trabajos de axiomatización de la geometría.

Frechet (1878-1973) trató de unificar la teoría de Cantor y el tratamiento de funciones como espacios de puntos. Introdujo los espacios métricos y definió los conjuntos compactos como los que actualmente llamamos relativamente compactos (de adherencia compacta).

Hausdorff (1868-1942) utilizó la noción de entorno, que había ya sido usada por Hilbert en una axiomatización de la geometría del plano, para dar una teoría axiomática de lo que hoy se entiende por topología general.

1.7 Apéndice. Espacios métricos

Todos los conceptos que hemos visto pueden darse en un contexto más general. De hecho se ha partido de la distancia euclídea pero en pocas ocasiones, aunque si en alguna, hemos utilizado la forma especial de esta distancia o las propiedades específicas de trabajar en el espacio R^n. Se trata ahora de ver como una noción de distancia general lleva a similares conceptos a los que hemos dado y de dar algún ejemplo que pueda apuntar el interés de estas nociones.

Definición 1.26. *Sea M un conjunto y d una aplicación de $M \times M$ en R que cumpla las siguientes propiedades:*

1.$d(x,y) \geq 0$, $d(x,y) = 0$ si y sólo si $x = y$.

2. Para cada $x, y \in M$, $d(x,y) = d(y,x)$.

3. Para cada $x, y, z \in M$, $d(x,y) \leq d(x,z) + d(z,y)$. (Desigualdad triangular).

Entonces a la aplicación d se le llama una distancia en M y a este conjunto un espacio métrico.

Ejemplo 1.24. *1. Sea M un subconjunto de R^n y d la distancia euclídea definida sobre los puntos de M. Se trata de un espacio métrico.*

2. Consideremos el conjunto de las sucesiones de números reales $\{a^i\}$ tales que $\sum_i |a^i|^2 < +\infty$. Veamos que se trata de un espacio vectorial con la suma de sucesiones y producto por escalares habitual. En efecto, si $\sum_i |a^i|^2 < +\infty$ y $\sum_i |b^i|^2 < +\infty$, tendremos

$$\sum_i \left|a^i + b^i\right|^2 \leq 2\left(\sum_i \left|a^i\right|^2 + \sum_i \left|b^i\right|^2\right) < +\infty.$$

Por tanto, la sucesión $\{a^i + b^i\}$ pertenece al espacio. También

$$\sum_i \left|\lambda a^i\right|^2 = |\lambda|^2 \sum_i \left|a^i\right|^2 < +\infty.$$

Es pues un espacio vectorial, de dimensión infinita, que se suele denotar por l^2. Puede definirse en este espacio un producto escalar mediante

$$\left\{a^i\right\} \cdot \left\{b^i\right\} = \sum_i a^i b^i.$$

Es fácil ver que la serie es convergente ya que tenemos la estimación

$$\left|a^i b^i\right| \leq \frac{1}{2}\left(\left|a^i\right|^2 + \left|b^i\right|^2\right)$$

y las series $\sum_i |a^i|^2$ *y* $\sum_i |b^i|^2$ *son convergentes. Es ya trivial verificar que este producto escalar es lineal en cada variable, simétrico y definido positivo. Por tanto vale, como hicimos en los espacios euclídeos, la desigualdad de Schwarz y esto nos permite definir una norma*

$$\left\| \left\{ a^i \right\} \right\| = \left(\left\{ a^i \right\} \cdot \left\{ a^i \right\} \right)^{\frac{1}{2}}.$$

Finalmente se define una distancia como

$$d\left(\left\{ a^i \right\}, \left\{ b^i \right\} \right) = \left\| \left\{ a^i \right\} - \left\{ b^i \right\} \right\|.$$

Las demostraciones son las mismas que para los espacios euclídeos.

3. *Consideremos el espacio* $\mathcal{C}$ *de las funciones continuas definidas en un intervalo* $[a, b]$ *a valores números reales. Se trata de un espacio vectorial en el cual se puede definir una norma mediante*

$$\|f\| = \sup_{x \in [a,b]} |f(x)|.$$

Las propiedades de norma se pueden verificar directamente (ahora no proviene de un producto escalar). Esto permite definir una distancia mediante

$$d(f, g) = \|f - g\|.$$

En un espacio métrico se definen entornos circulares, puntos adherentes, aislados, de acumulación, límites de sucesiones, conjuntos abiertos y cerrados en la misma forma y con las mismas propiedades que vimos para los espacios euclídeos.

Ejemplo 1.25. 1. *En el primer ejemplo en que* M *es un subconjunto de* R^n *los entornos circulares de un punto de* M *son la intersección de los entornos circulares en el espacio euclídeo con* M*. De aquí se deduce que los conjuntos abiertos de* M *resultan ser lo que habíamos llamado abiertos relativos a* M*.*

2. *Veamos qué significa en el espacio* $\mathcal{C}$ *que una sucesión* $\{f^i\}$ *sea convergente a* f*. Para cada* $\varepsilon > 0$ *debe existir* i_0 *tal que para* $i > i_0$*, se cumple* $d(f^i, f) < \varepsilon$*. Esto significa que*

$$\sup_{x \in [a,b]} \left| f^i(x) - f(x) \right| < \varepsilon.$$

Es lo mismo que decir que la sucesión f^i *converge uniformemente hacia* f*.*

Pueden darse los conceptos de compacto y compacto por sucesiones y también aquí son equivalentes. Sin embargo la equivalencia con los cerrados y acotados (definidos también en su forma natural) no es en general cierta. No podemos ahora aplicar como se hizo para los espacios euclídeos el teorema de Bolzano-Weierstrass.

Los conceptos de límite de una función definida en un espacio métrico y de función continua son enteramente análogos a los ya dados tanto en términos de sucesiones como de entornos.

Ejemplo 1.26. 1. *Definamos la siguiente función sobre l^2. Consideremos un punto $\{b^i\} \in l^2$. Consideremos la función definida en l^2 mediante $f(\{x^i\}) = \sum_i x^i b^i$. Es una aplicación continua ya que*

$$\left|f\left(\left\{x^i\right\}\right) - f\left(\left\{a^i\right\}\right)\right| = \left|\sum_i \left(x^i - a^i\right) b^i\right| \leq \left\|\left\{x^i - a^i\right\}\right\| \left\|\left\{b^i\right\}\right\|.$$

Podemos considerar $\{b^i\}$ no nulo, ya que de lo contrario es trivial. Dado $\varepsilon > 0$, se considera $\delta = \varepsilon \left\|\{b^i\}\right\|^{-1}$. Tendremos que si $\|\{x^i\} - \{a^i\}\| < \delta$ se tiene

$$\left|f\left(\left\{x^i\right\}\right) - f\left(\left\{a^i\right\}\right)\right| < \varepsilon$$

y por tanto f es continua.

2. *Consideremos en $\mathcal{C}$ la función F definida mediante*

$$F(f) = \int_a^b f(x)\, dx.$$

Se trata de una función continua ya que dado $\varepsilon > 0$ tomando $d(f,g) < \delta = \varepsilon (b-a)^{-1}$ tendremos

$$|F(f) - F(g)| = \left|\int_a^b f(x)\, dx - \int_a^b g(x)\, dx\right| \leq \int_a^b |f(x) - g(x)|\, dx < \varepsilon.$$

Ejercicios

1. Prueba que si M es un espacio métrico, la frontera de $\mathcal{E}(a,r)$ es la esfera de centro a y radio r, es decir $\{x \in M;\ d(x,a) = r\}$.

2. Prueba que si una sucesión en un espacio métrico tiene límite, éste es único.

3. En un conjunto se define una aplicación d mediante $d(x,y) = 1$ si $x \neq y$ y 0 si $x = y$. Prueba que se trata de un espacio métrico. Describe los conjuntos abiertos y los conjuntos compactos.

4. Sea M un espacio con una distancia d. Prueba que $q(x,y) = \frac{d(x,y)}{1+d(x,y)}$ es otra distancia en M tal que $\{a_i\} \to a$ respecto a la distancia d si y sólo si $\{a_i\}$ tiene por limite a respecto a la distancia q.

5. En el espacio l^2 consideremos los elementos $e_i = (0, \ldots, \overset{i)}{1}, 0, \ldots)$. Prueba que el conjunto de estos elementos es acotado pero que la sucesión $\{e_i\}$ no admite ninguna parcial convergente. (Obsérvese la diferencia con respecto al espacio euclídeo R^n).

Capítulo 2

Cálculo diferencial en varias variables

2.1 Derivadas direccionales. Diferencial

2.1.1 Conceptos y propiedades fundamentales

Al tratar de introducir la noción de derivada para una función de varias variables se plantea un problema que no se presentaba con las funciones de una variable. Si queremos definir, como en éstas, un cociente incremental $(f(x+h)-f(x))/h$, la primera dificultad que se presenta es que ahora h es un elemento de R^n y este cociente no tiene sentido. No podemos dividir el incremento de la función por un elemento de R^n. Lo que si puede hacerse es considerar únicamente incrementos a lo largo de una dirección determinada por un vector unitario v y dividir por el parámetro correspondiente a este incremento. Esto nos llevará al concepto de derivada direccional.

Definición 2.1. *Sea f una función definida en un entorno de a en R^n. Sea v un vector unitario de R^n. Si existe el limite*

$$\lim_{t\to 0}\frac{f(a+tv)-f(a)}{t}$$

le llamaremos derivada direccional de la función en el punto a, correspondiente a la dirección v. Designaremos este valor por $D_v f(a)$.

En el caso particular en que $v = e_i = (0,0,\ldots,\overset{i)}{1},\ldots,0)$ se le llama derivada parcial respecto la variable x_i y se escribe $D_i f(a)$ o también $\frac{\partial f}{\partial x_i}(a)$. Es decir, si $a=(a_1,\ldots,a_n)$

$$D_i f(a)=\frac{\partial f}{\partial x_i}(a)=\lim_{t\to 0}\frac{f(a_1,\ldots,a_{i-1},a_i+t,a_{i+1},\ldots a_n)-f(a_1,\ldots,a_n)}{t}$$

Observemos que la derivada direccional $D_v f(a)$ no es otra cosa que la derivada de la función de una variable $h(t)=f(a+tv)$ en $t=0$. En ocasiones se considera también la definición de D_v sin restringirse a que el vector v tenga norma 1. Se habla de la derivación respecto al vector v. Si no decimos lo contrario, nosotros supondremos que v es unitario.

Para funciones de una variable real se define la función derivada como la función que asigna a cada punto el valor de la derivada. Análogamente podemos definir las funciones derivadas parciales.

Definición 2.2. *Se denomina función derivada parcial respecto a la variable x_i y se escribe $D_i f$ o $\frac{\partial f}{\partial x_i}$ a la función que asigna a cada punto a el valor $D_i f\,(a)$.*

La derivada direccional de la función f en a en la dirección dada por v no es otra cosa que la derivada de la función de una variable real $g(t) = f(a + tv)$ en $t = 0$. Obsérvese que al estar definida f en un entorno de a, la función g lo está en un entorno de 0. Las propiedades de las derivadas direccionales son entonces análogas a las de las derivadas de las funciones de una variable. En particular, para calcular la derivada parcial respecto una variable se consideran todas las otras como constantes y se aplican las reglas habituales.

Ejemplo 2.1. *Sea $f(x, y) = x^2 + xy - 5y^2$. Las funciones derivadas parciales serán $\frac{\partial f}{\partial x} = 2x + y$, $\frac{\partial f}{\partial y} = x - 10y$. La derivada direccional en $(0, 0)$ en la dirección dada por el vector $u = (\frac{1}{2}, \frac{\sqrt{3}}{2})$ será la derivada en $t = 0$ de la función $(\frac{1}{2}t)^2 + \frac{1}{2}t\frac{\sqrt{3}}{2}t - 5\left(\frac{\sqrt{3}}{2}t\right)^2$, es decir 0.*

Puede darse otra noción que extiende la de derivada para funciones de una variable. Para éstas, la existencia de derivada equivale a la existencia de recta tangente al gráfico de la función. Es decir, la función $y = f(x)$ de R en R es derivable en a si existe una aplicación lineal de R en R que a h le hace corresponder mh, que cumple

$$f(a + h) - f(a) - mh = o_{h\to 0}(h).$$

La constante m es la derivada de la función en a y la imagen de h por la aplicación lineal se denota por $df_a(h) = mh$. La recta tangente al gráfico de la función en el punto $(a, f(a))$ es $y = f(a) + m(x - a)$.

Una función se dice que es diferenciable en un punto si se puede aproximar por una función afín, es decir, por la suma del valor que toma la función en el punto mas una aplicación lineal. Concretamente, tendremos la siguiente definición.

Definición 2.3. *Sea $y = f(x)$ una función definida en un abierto D de R^n, con valores en R, $a \in D$. Se dice que f es diferenciable en a si existe una aplicación lineal df_a de R^n en R tal que*

$$f(a + h) - f(a) - df_a(h) = o_{h\to 0}(\|h\|). \tag{2.1}$$

La aplicación lineal df_a que, como veremos, es única, se denomina diferencial de f en el punto a.

La relación 2.1 si $x = a + h$ se escribirá en la forma

$$f(x) = f(a) + df_a(x - a) + o_{x\to a}\left(\|x - a\|\right).$$

En ocasiones también la escribiremos

$$f(x) = f(a) + df_a(x - a) + g(x)\,\|x - a\|$$

donde $g(x) \rightarrow_{x\rightarrow a} 0$.

Por ejemplo, si $n = 2$ la diferenciabilidad de f en a significa que existe un plano en R^3 que pasa por el punto $(a_1, a_2, f(a_1, a_2))$ y que tiene con la función un contacto de orden superior al primero. Este plano se denomina plano tangente al gráfico de la función $y = f(x_1, x_2)$ en el punto $(a_1, a_2, f(a_1, a_2))$. Explícitamente, la condición de diferenciabilidad se expresa diciendo que existen constantes α y β tales que

$$\lim_{x\rightarrow a} \frac{f(x_1, x_2) - (f(a_1, a_2) + \alpha(x_1 - a_1) + \beta(x_2 - a_2))}{\|x - a\|} = 0.$$

Obsérvese que las constantes α y β definen la aplicación lineal $df_a(h) = \alpha h_1 + \beta h_2$ y que la ecuación del plano tangente a $y = f(x_1, x_2)$ en el punto a es $y = f(a_1, a_2) + \alpha(x_1 - a_1) + \beta(x_2 - a_2)$.

El ejemplo más sencillo de aplicación diferenciable es una aplicación lineal $T : R^n \rightarrow R$. Se tiene $T(a+h) = T(a)+T(h)$. Por lo tanto es válido 2.1 con $o_{h\rightarrow 0}(\|h\|) \equiv 0$, y T es diferenciable con diferencial la propia aplicación T.

Veamos una primera relación entre la diferencial de una función y las derivadas direccionales en un punto.

Teorema 2.1. *Sea $y = f(x)$ una función definida en un abierto D de R^n, con valores en R, $a \in D$. Si f es diferenciable en a, entonces f tiene derivada direccionales en a y se cumple*

$$D_u f(a) = df_a(u).$$

En particular, las derivadas direccionales están determinadas por la diferencial.

Demostración. De la definición de diferenciabilidad se tiene

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

Los límites direccionales serán también cero, es decir, para cada $u \in R^n$

$$\lim_{t\rightarrow 0} \frac{f(a + tu) - f(a) - df_a(tu)}{\|tu\|} = 0$$

que es equivalente a

$$\lim_{t\rightarrow 0} \frac{f(a + tu) - f(a) - tdf_a(u)}{t} = 0.$$

Esta relación puede escribirse

$$\lim_{t\rightarrow 0} \frac{f(a + tu) - f(a)}{t} = df_a(u)$$

es decir, la derivada direccional de f en a según la dirección u es $df_a(u)$.

□

El teorema implica que la diferencial, si existe, es única puesto que queda determinada por las derivadas direccionales que son únicas.

Es importante observar que en una función diferenciable las derivadas parciales determinan todas las derivadas direccionales. En efecto, si $u = u_1 e_1 + \ldots . + u_n e_n$

$$D_u f(a) = df_a(u) = u_1 df_a(e_1) + \ldots + u_n df_a(e_n) = u_1 D_1 f(a) + \ldots + u_n D_n f(a).$$

Tendremos entonces que la diferenciabilidad de f en a se podrá expresar en la forma

$$f(x) = f(a) + \sum_{i=1}^{n} D_i f(a)(x_i - a_i) + o_{x\to a}(\|x - a\|)$$

Veamos una forma de interpretar la expresión de la derivada direccional en términos de las derivadas parciales. Precisaremos de la siguiente definición

Definición 2.4. *Se llama gradiente de f en a y se denota por $\nabla f(a)$ al vector*

$$(D_1 f(a), \ldots, D_n f(a)).$$

La expresión $D_u f(a) = \sum_{i=1}^{n} D_i f(a) u_i$ puede pensarse como el producto escalar de u por $\nabla f(a)$. Esto permite dar una interesante interpretación del vector gradiente. Supongamos que existe la diferencial de una función en un punto y es no nula. Nos preguntamos por el valor máximo de $|D_u f(a)|$ al variar u. Puesto que

$$|D_u f(a)| = |u \cdot \nabla f(a)| = \|u\| \, \|\nabla f\,(a)\| \, |\cos(u, \nabla f(a))|$$

el módulo de la derivada direccional alcanzará el valor máximo en la dirección dada por el gradiente de f en a.

Debe observarse que la existencia de derivadas direccionales en un punto no implica la diferenciabilidad de la función.

Ejemplo 2.2.

1. *La función $f(x,y) = \frac{x^2 y}{x^2+y^2}$ si $(x,y) \neq (0,0)$, $\quad f(0,0) = 0$ tiene derivadas direccionales pues*
$$D_u f(0,0) = \lim_{t\to 0} \frac{u_1^2 t^2 u_2 t}{t^3} = u_1^2 u_2.$$
En particular, las derivadas parciales son 0. La función no puede ser diferenciable ya que entonces todas las derivadas direccionales deberían ser 0, lo que no es así como acabamos de ver.

2. *La función $f(x,y) = \frac{x^2 y^2}{x^2+y^2}$ si $(x,y) \neq (0,0)$, $\quad f(0,0) = 0$ tiene derivadas parciales en $(0,0)$ que valen 0. La función es diferenciable en este punto pues $f(x,y) = o_{(x,y)\to(0,0)}\left(\|(x,y)\|\right)$. En efecto, si escribimos $x = \rho\cos\theta$, $y = \rho\sin\theta$, tendremos $\frac{x^2y^2}{x^2+y^2} \leq \rho^2 = o_{\rho\to 0}\left(\rho\right)$.*

Más adelante veremos que la existencia de las funciones derivadas parciales no sólo en el punto considerado sino en todo un entorno y su continuidad en el punto implica la diferenciabilidad en este punto.

Establezcamos antes la continuidad de las funciones diferenciables.

Teorema 2.2. *Sea f es una función definida en un entorno de un punto a y diferenciable en él. Entonces f es continua en este punto.*

Demostración. Obsérvese que si T es una aplicación lineal de R^n en R existe una constante C tal que $|T(h)| \leq C \|h\|$. En efecto, si $C = \max_{\|u\| \leq 1} |T(u)|$ se cumple que $\left| T\left(\frac{h}{\|h\|}\right)\right| \leq C$. Tendremos entonces que, aplicando esta desigualdad a df_a se obtiene

$$|f(a+h) - f(a)| \leq o_{h\to 0}(\|h\|) + \|df_a(h)\| \leq o_{h\to 0}(\|h\|) + C\|h\|.$$

El límite del segundo miembro cuando $h \to 0$ es 0. Esto implica la continuidad de f en a.

□

La existencia de todas las derivadas direccionales no implica la continuidad.

Ejemplo 2.3. *La función $f(x,y) = x^2/y$ si $y \neq 0$, $f(x,0) = 0$ no es continua en $(0,0)$ pero tiene derivadas en cualquier dirección. La no continuidad de la función puede establecerse observando que el límite en $(0,0)$ según el subconjunto $y = mx^2$ es $1/m$ que varía al tomar diversos valores de m. La existencia de las derivadas direccionales sigue directamente de su definición.*

La existencia de derivadas direccionales expresa únicamente una propiedad referente a la restricción de la función a cada una de las rectas que pasa por el punto considerado. La diferenciabilidad de una función expresa mejor el comportamiento en todo un entorno del punto. Una primera indicación es el teorema que hemos visto según el cual la diferenciabilidad implica la continuidad. La existencia de derivadas direccionales en el punto ya hemos visto que no implica la continuidad. El teorema siguiente da una condición suficiente de diferenciabilidad.

Teorema 2.3. *Sea $y = f(x)$ una función definida en un abierto $D \subset R^n$ a valores en R, $a \in D$. Si existen todas las derivadas parciales de f en un entorno de a y son continuas en este punto, entonces f es diferenciable en a.*

Demostración. Se trata de ver que

$$f(a+h) - f(a) = \sum_{i=1}^{n} D_i f(a) h_i + o_{h\to 0}(\|h\|)$$

En efecto, consideremos un entorno $E(a,\delta) \subset D$ en el que existen $D_i f$, $i = 1, \ldots, n$. Escribamos

$$f(a+h) - f(a) = \sum_{i=1}^{n} \Big(f(a+h^i) - f(a+h^{i-1}) \Big)$$

donde $h^0 = 0$, $h^1 = h_1 e_1$, $h^2 = h_1 e_1 + h_2 e_2, \ldots, h^n = h_1 e_1 + \ldots + h_n e_n$.

Aplicando el teorema del valor medio en cada variable, existirán $c_i = a + h_1 e_1 + \ldots + h_{i-1}e_{i-1} + h_i \lambda_i e_i$ con $0 < \lambda_i < 1$ tales que $f(a+h^i) - f(a+h^{i-1}) = D_i f(c_i) h_i$. Se obtendrá entonces

$$f(a+h) - f(a) = \sum_{i=1}^{n} D_i f(a) h_i + \sum_{i=1}^{n} \left(D_i f(c_i) - D_i f(a)\right) h_i$$

Ahora bien, $\sum_{i=1}^{n} \left(D_i f(c_i) - D_i f(a)\right) h_i$ es $o_{h\to 0}\left(\|h\|\right)$ ya que las funciones $D_i f$ son continuas en a y $|h_i| \leq \|h\|$.

□

Definición 2.5. *Se dice que una función f definida en un abierto de de R^n a valores en R es de clase $\mathcal{C}^1$ si tiene derivadas parciales en el dominio de definición y son continuas.*

Del teorema anterior se deduce que toda función de la clase $\mathcal{C}^1$ es diferenciable en todos los puntos.

Ejemplo 2.4.

1. *La función $z = (x+y)\sin(x^2+y)$ es diferenciable en todos sus puntos ya que es de clase $\mathcal{C}^1$.*

2. *La función $f\left(x,y\right) = \left(x+y\right)^2 \sin \frac{1}{\left(x^2+y^2\right)^{\frac{1}{2}}}$ si $\left(x,y\right) \neq \left(0,0\right)$ es diferenciable en el origen. En efecto, tenemos que $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$, y la condición de diferenciabilidad será*

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

Sin embargo la función no es de clase $\mathcal{C}^1$ ya que

$$\frac{\partial f}{\partial x} = 2\left(x+y\right) \sin \frac{1}{\left(x^2+y^2\right)^{\frac{1}{2}}} - \left(x+y\right)^2 \frac{x}{\left(x^2+y^2\right)^{\frac{3}{2}}} \cos \frac{1}{\left(x^2+y^2\right)^{\frac{1}{2}}}$$

si $(x,y) \neq (0,0)$. El límite de esta función para $(x,y) \to (0,0)$ no existe. Es suficiente comprobar que el límite según el subconjunto $y = mx$ no existe.

Las nociones de función diferenciable y de derivada direccional se pueden dar también para funciones a valores en R^m. Será conveniente extender la noción de "o" a funciones a valores vectoriales.

Definición 2.6. *Una función $g(h)$ definida en un entorno del origen de R^n a valores en R^m se dice que es un $o_{h\to 0}(\|h\|)$ si para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que si $0 < \|h\| < \delta$ entonces $\|g(h)\| < \varepsilon\|h\|$.*

Tendremos ahora

Definición 2.7. *Sea $y = f(x)$ una función definida en un abierto D de R^n, con valores en R^m, $a \in D$. Se dice que f es diferenciable en a si existe una aplicación lineal df_a de R^n en R^m tal que*

$$f(a+h) - f(a) - df_a(h) = o_{h\to 0}(\|h\|).$$

Análogamente, la misma definición que hemos dado de derivada direccional para funciones a valores en R sirve para funciones valoradas en R^m. Ahora las derivadas direccionales en un punto serán elementos de R^m. El estudio de estas funciones puede reducirse al de las funciones valoradas en R como muestra el siguiente teorema.

Teorema 2.4. *Sea $f = (f_1, \ldots, f_m)$ una función definida en una abierto D de R^n, a valores en R^m. La función f es diferenciable en $a \in D$ si y sólo si lo son las funciones $f_1, \ldots, f_m$. Se cumple, en este caso, que las m componentes de df_a son $df_{1\ a}, \ldots df_{m\ a}$. Análogamente, f tiene derivada direccional en a respecto a la dirección u si y sólo si existen las correspondientes derivadas direccionales de cada una de las funciones $f_1, \ldots, f_m$ y se tiene $D_u f(a) = (D_u f_1(a), \ldots, D_u f_m(a))$.*

Demostración. Si f es diferenciable en a significa que existe una aplicación lineal $df_a : R^n \to R^m$ tal que $f(a + h) - f(a) - df_a(h) = o_{h\to 0}(\|h\|)$. Si llamamos $(df_a)_i$ a cada una de las componentes de la aplicación df_a tendremos que $f_i(a+h) - f_i(a) - (df_a)_i(h) = o_{h\to 0}(\|h\|)$ para $i = 1, \ldots, m$. Esto nos dice que cada f_i es diferenciable y que su diferencial es la componente i-ésima de df_a. El razonamiento puede invertirse en el sentido de que si todas las funciones coordenadas f_i son diferenciables se sigue la diferenciabilidad de f. Análogamente se prueba la afirmación sobre las derivadas direccionales. De hecho el teorema es una simple consecuencia de que la existencia de un límite de una función a valores en R^m, equivale a la existencia del correspondiente límite de cada una de las funciones coordenadas.

□

La expresión en coordenadas de la diferencial de una función a valores en R^m se deduce entonces de la expresión en coordenadas de cada una de las $df_{i\ a}$. Se expresará como una matriz de n columnas y m filas. La fila i-ésima se obtendrá calculando las imágenes de los vectores de la base canónica $e_1, \ldots, e_n$, es decir $df_{i\ a}(e_j) = D_j f_i(a)$. La matriz se denomina matriz jacobiana de f en el punto a, y será

$$\begin{pmatrix} D_1 f_1(a) & \ldots & D_n f_1(a) \\ . & . & . \\ . & . & . \\ D_1 f_m(a) & \ldots & D_n f_m(a) \end{pmatrix}$$

Veamos algunas propiedades elementales de las funciones diferenciables.

Teorema 2.5. *Sean f y g funciones definidas en un abierto D de R^n, a valores en R^m, diferenciables en un punto $a \in D$, $\lambda \in R$. Entonces $f + g$ y λf son diferenciables en a y se cumple $d(f + g)_a = df_a + dg_a$, $d(\lambda f)_a = \lambda df_a$. Si $m = 1$, fg es diferenciable en a y $d(fg)_a = f(a)dg_a + g(a)df_a$. Si, además, g es no nula, entonces f/g es diferenciable y $d(f/g)_a = (g(a)df_a - f(a)dg_a)/g(a)^2$.*

Demostración. Si f y g son diferenciables en a se tiene

$$f(a + h) = f(a) + df_a(h) + o_{h\to 0}(\|h\|) \tag{2.2}$$

$$g(a+h) = g(a) + dg_a(h) + o_{h\to 0}(\|h\|). \tag{2.3}$$

Sumando 2.2 y 2.3 miembro a miembro se tiene

$$(f+g)(a+h) = (f+g)(a) + (df_a + dg_a)(h) + o_{h\to 0}(\|h\|).$$

Esta expresión nos dice que $f+g$ es diferenciable en a y que su diferencial es $df_a + dg_a$.

Multiplicando 2.2 por λ obtenemos la diferenciabilidad de λf

$$\lambda f(a+h) = \lambda f(a) + \lambda df_a(h) + o_{h\to 0}(\|h\|).$$

Sea ahora $m = 1$. Multipliquemos miembro a miembro las expresiones 2.2 y 2.3. Obtendremos

$$(fg)(a+h) = (fg)(a) + f(a)dg_a(h) + g(a)df_a(h) + o_{h\to 0}(\|h\|)$$

ya que $|df_a(h)dg_a(h)| \leq c\,\|h\|^2$, $df_a(h)o_{h\to 0}(\|h\|) = o_{h\to 0}(\|h\|)$ y $o_{h\to 0}(\|h\|)o_{h\to 0}(\|h\|) = o_{h\to 0}(\|h\|)$.

Probemos la regla de diferenciación del cociente. Consideremos, en primer lugar, que $f = 1$. Si g es no nula tendremos que

$$\frac{1}{g(a+h)} - \frac{1}{g(a)} - \left(\frac{-dg_a(h)}{g(a)^2}\right) = \frac{-dg_a(h) + o_{h\to 0}(\|h\|)}{g(a)g(a+h)} - \left(\frac{-dg_a(h)}{g(a)^2}\right) = o_{h\to 0}\left(\|h\|\right)$$

ya que g es continua en a y se tiene que

$$\begin{gathered} g(a)\left(-dg_a(h) + o_{h\to 0}(\|h\|)\right) + g(a+h)dg_a(h) = \\ g(a)\left(-dg_a(h) + o_{h\to 0}(\|h\|)\right) + \left(g\left(a\right) + dg_a\left(h\right) + o_{h\to 0}\left(\|h\|\right)\right) dg_a\left(h\right) = \\ g(a)o_{h\to 0}(\|h\|) + dg_a(h)^2 + o_{h\to 0}(\|h\|)dg_a(h) = o_{h\to 0}(\|h\|). \end{gathered}$$

Por último, la diferenciación del cociente $\frac{f}{g}$ se deduce de la regla de diferenciación del producto y del caso anterior, expresando la función como el producto de f y de $\frac{1}{g}$.

□

Las propiedades sobre las diferenciales tienen sus equivalentes para su expresión en coordenadas, es decir, para las matrices jacobianas. Así la matriz jacobiana de la suma de dos funciones es la suma de las matrices jacobianas o la matriz jacobiana del producto de una función por un escalar es el producto de dicho escalar por la matriz jacobiana de la función.

El conjunto de las aplicaciones lineales de R^n en R tiene estructura de espacio vectorial de dimensión n. A veces es conveniente, en lugar de expresar una aplicación lineal como una matriz, hacerlo como una combinación lineal de una base de dicho espacio vectorial. Una tal base está formada por las aplicaciones coordenadas $(x_1, \ldots x_n) \to x_i$. Estas aplicaciones, por ser lineales, sabemos que son diferenciables en todo punto y que su diferencial coincide con la propia aplicación. Así, dx_i coincidirá con la aplicación coordenada x_i, es decir, $dx_i(x) = x_i$.

Toda aplicación lineal de R^n en R se expresará como una combinación lineal de $dx_1, \ldots, dx_n$. En particular si f es diferenciable en a tendremos

$$df_a = \sum_{i=1}^{n} \lambda_i dx_{i\ a}.$$

Las coordenadas λ_i se pueden obtener aplicando ambos miembros al vector e_j. Tendremos $\left(\sum \lambda_i dx_{i\ a}\right)(e_j) = \lambda_j$ y $df_a\,(e_j) = D_j f(a)$. De aquí la expresión

$$df_a = \sum_{i=1}^{n} D_i f\,(a)\, dx_{i\ a}.$$

Veamos ahora que la composición de funciones diferenciables es también diferenciable.

Teorema 2.6. Regla de la Cadena.

Sea f una función definida en un abierto $D \subset R^n$ a valores en R^m y sea g una función definida en un abierto $D_1 \supset f(D)$ a valores en R^k. Supongamos que f es diferenciable en un punto a y que g lo es en el punto $b = f(a)$. Entonces $g \circ f$ es diferenciable en a y $d(g \circ f)_a = dg_{f(a)} \circ df_a$.

Demostración. Tendremos

$$f(a+h) = f(a) + df_a(h) + o_{h\to 0}(\|h\|)$$

$$g(b+t) = g(b) + dg_b(t) + o_{t\to 0}(\|t\|). \tag{2.4}$$

De donde

$$g\left(f\,(a+h)\right) = g\left(f(a) + df_a(h) + o_{h\to 0}(\|h\|)\right)$$

Aplicando 2.4 tomando $t = df_a(h) + o_{h\to 0}(\|h\|)$ tendremos

$$\begin{gathered} g\left(f\,(a+h)\right) = g\left(f\,(a)\right) + dg_b\left(df_a(h) + o_{h\to 0}(\|h\|)\right) + \\ o_{df_a(h)+o_{h\to 0}(\|h\|)\to 0}\left(\|df_a(h) + o_{h\to 0}(\|h\|)\|\right). \end{gathered}$$

Puesto que dg_b es lineal tendremos

$$\begin{gathered} g\left(f\,(a+h)\right) = g\left(f\,(a)\right) + dg_b\left(df_a(h)\right) + dg_b\left(o_{h\to 0}(\|h\|)\right) + \\ o_{df_a(h)+o_{h\to 0}(\|h\|)\to 0}\left(\|df_a(h) + o_{h\to 0}(\|h\|)\|\right). \end{gathered}$$

El teorema quedará demostrado si probamos que

$$dg_b\left(o_{h\to 0}(\|h\|)\right) + o_{df_a(h)+o_{h\to 0}(\|h\|)\to 0}\left(\|df_a(h) + o_{h\to 0}(\|h\|)\|\right) = o_{h\to 0}(\|h\|).$$

En efecto, sabemos que $\|dg_b\left(o_{h\to 0}(\|h\|)\right)\| \le c\,\|o_{h\to 0}(\|h\|)\| = o_{h\to 0}(\|h\|)$. Por otro lado, dado $\varepsilon > 0$, existe $\delta > 0$ tal que si $\|df_a(h) + o_{h\to 0}(\|h\|)\| < \delta$, se cumple

$$\left\|o_{df_a(h)+o_{h\to 0}(\|h\|)\to 0}\left(\|df_a(h) + o_{h\to 0}(\|h\|)\|\right)\right\| < \varepsilon\,\|df_a(h) + o_{h\to 0}(\|h\|)\|$$

Ahora bien,

$$\|df_a(h) + o_{h\to 0}(\|h\|)\| \leq C\,\|h\|$$

si $\|h\| < \delta_1$. Tendremos entonces que si $\|h\| <$Min$(\delta_1, \delta/C)$

$$\left\| o_{df_a(h)+o_{h\to 0}(\|h\|)\to 0}\left(\|df_a(h) + o_{h\to 0}(\|h\|)\|\right)\right\| < \varepsilon C\,\|h\|$$

y por tanto $o_{df_a(h)+o_{h\to 0}(\|h\|)\to 0}\left(\|df_a(h) + o_{h\to 0}(\|h\|)\|\right) = o_{h\to 0}\left(\|h\|\right).$

□

El teorema anterior nos dice que la diferencial de la función compuesta es la composición de las diferenciales de las funciones. Sabemos que la expresión en coordenadas de la diferencial es la matriz jacobiana. Tendremos entonces que la matriz jacobiana de una función compuesta es el producto de las matrices jacobianas en los puntos correspondientes.

Ejemplo 2.5.

1. *Sea f una función de R en R^n y g de R^n en R, ambas diferenciables. La función compuesta $g \circ f$ es una función de una variable real a valores reales. Su diferencial vendrá dada por una matriz de una fila y una columna que es simplemente su derivada. Su expresión, como consecuencia de la regla de la cadena, será la siguiente:*

$$\frac{d\,(g\circ f)}{dt}(t_0) = \begin{pmatrix} \frac{\partial g}{\partial x_1} & \cdots & \frac{\partial g}{\partial x_n} \end{pmatrix}_{f(t_0)} \begin{pmatrix} \frac{\partial f_1}{\partial t} \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial f_n}{\partial t} \end{pmatrix}_{t_0} = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}\left(f\left(t_0\right)\right)\frac{\partial f_i}{\partial t}\left(t_0\right).$$

2. *Si f es una función de R^n en R^m y g es una función de R^m en R, ambas diferenciables, tendremos que, operando como antes*

$$\frac{\partial\,(g\circ f)}{\partial t_j}\left(t_0\right) = \sum_{i=1}^{m} \frac{\partial g}{\partial x_i}\left(f\left(t_0\right)\right)\frac{\partial f_i}{\partial t_j}\left(t_0\right).$$

3. *Sea f es una función de R^n en R^n, diferenciable en a y g otra función de R^n en R^n diferenciable en $b = f(a)$ y tal que g es la inversa de f. Entonces $dg_b = df_a^{-1}$. En efecto, $dg_b \circ df_a = d(g\circ f)_a = I$. Análogamente $df_a \circ dg_b = I$.*

Veamos cómo actúa la diferencial de una aplicación sobre los vectores tangentes a las curvas. Una curva de clase $\mathcal{C}^r$ en R^n es una aplicación γ de un intervalo de R en R^n cuyas funciones coordenadas γ_i son de esta clase. El vector $\gamma'\left(t\right)$=$\lim_{h\to 0}\frac{\gamma(t+h)-\gamma(t)}{h}$ es un vector tangente a la curva en el punto $\gamma\left(t\right)$.

Teorema 2.7. *Sea f una aplicación definida en el entorno de un punto p en R^n a valores en R^m diferenciable en p. Sea $v = \gamma'\left(0\right)$ el vector tangente a una curva γ tal que $\gamma\left(0\right) = p$. Se tiene que $df_p\left(v\right)$ es el vector tangente a la curva imagen $f\circ\gamma$ en $t = 0$.*

Demostración. El vector tangente a $f \circ \gamma$ en $t = 0$ tendrá por componente $i-$ésima

$$\frac{d}{dt}\left(f_i \circ \gamma\right)(0) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}\left(\gamma\left(0\right)\right)\gamma_j'\left(0\right) = df_{i\ p}\left(v\right).$$

Por lo tanto $\frac{d}{dt}\left(f \circ \gamma\right)(0) = df_p\left(v\right).$

□

Veamos una aplicación de la regla de la cadena.

Teorema 2.8. *Sea f una función de clase $\mathcal{C}^1$ definida en un abierto conexo A y tal que $df_x = 0$ para cada $x \in A$. Entonces la función es constante.*

Demostración. Consideremos un punto $a \in A$. Todo punto b de A puede unirse con a mediante un arco de curva $\gamma : [0,1] \to A$ diferenciable a trozos, es decir que existe una partición de $[0,1]$ en un número finito de intervalos tal que sobre cada uno de ellos γ es diferenciable. Esto se deduce como en el teorema 1.33. Consideremos la función $h\left(t\right) = f\left(\gamma(t)\right)$. En cada uno de los intervalos en que γ es diferenciable h será derivable con derivada idénticamente cero. Por lo tanto h será constante en cada uno de estos intervalos y en consecuencia constante en $[0,1]$. En particular $h\left(0\right) = h\left(1\right)$ y, por tanto, $f\left(a\right) = f\left(b\right)$.

□

2.1.2 Derivadas de orden superior al primero

Consideremos una función $f : R^n \to R$ que admita funciones derivadas parciales $\frac{\partial f}{\partial x_i}$. Supongamos que, a su vez, éstas admitan derivadas parciales. Se presenta el problema de saber si las derivadas iteradas dependen o no del orden de derivación. Un ejemplo probará que si no añadimos hipótesis el resultado puede depender del orden en que se efectúe la derivación.

Ejemplo 2.6.

1. *Sea f la función definida en R^2 mediante $f(x,y) = \frac{xy\left(x^2-y^2\right)}{x^2+y^2}$ si $(x,y) \neq (0,0)$ y $f(0,0) = 0$. Se cumple $\frac{\partial f}{\partial x}\left(0,y\right) = -y$ y por tanto $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0,0) = -1$. Por otro lado $\frac{\partial f}{\partial y}\left(x,0\right) = x$ y por tanto $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0) = 1$. No coinciden en este caso las dos derivadas iteradas.*

2. *Sea f la función definida en R^2 mediante $f\left(x,y\right) = \cos xy$. Tendremos $\frac{\partial f}{\partial x} = -y\sin xy$ y $\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0,0) = 0$. Análogamente $\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0) = 0$. En este caso coinciden las derivadas iteradas.*

Veamos que una simple hipótesis suplementaria sobre las funciones derivadas parciales de la función permitirá asegurar la igualdad de las derivadas iteradas independientemente del orden de derivación.

Teorema 2.9. *Sea f una función definida en un entorno de un punto $a \in R^n$ a valores en R, que admite funciones derivadas parciales en este entorno y son diferenciables en a. Entonces para cada i,j se cumple $D_i\left(D_j f\right)(a) = D_j\left(D_i f\right)(a)$.*

Demostración. Sabemos que si una función g es diferenciable en un punto a y dejamos fijas unas variables, es decir, componemos la función g con la función $(x_1, \ldots x_i, \ldots, x_j, \ldots x_n) \to (a_1, \ldots x_i, \ldots, x_j, \ldots a_n)$ obtenemos una función diferenciable. Las hipótesis del teorema se cumplirán entonces por la función de dos variables reales $f(a_1, \ldots x_i, \ldots, x_j, \ldots a_n)$ que se obtiene dejando fijas todas las variables excepto la x_i y la x_j. Si la conclusión del teorema es válida para esta función también lo será para la función de partida ya que ambas tienen las mismas derivadas parciales respecto a las variables i y j en el punto a. Podemos pues suponer, sin que ello limite la validez del teorema, que la función de partida depende únicamente de dos variables x, y. Supondremos también, para simplificar las notaciones, que $a = (0,0)$ y que las hipótesis sobre las funciones se cumplen en un entorno circular de este punto. Consideremos la función

$$F(h) = f(h,h) - f(h,0) - f(0,h) + f(0,0).$$

Veamos que $\frac{F(h)}{h^2}$ tiene por límite cuando $h \to 0$ tanto a $D_1(D_2 f)(0,0)$ como a $D_2(D_1 f)(0,0)$.

Sea la función $G(x) = f(x,h) - f(x,0)$. Por el teorema del valor medio existe c, comprendido entre 0 y h, tal que

$$F(h) = G(h) - G(0) = hG'(c) = h(D_1 f(c,h) - D_1 f(c,0)).$$

Puesto que $D_1 f$ es diferenciable en el origen tendremos

$$F(h) = h^2 D_2(D_1 f)(0,0) + h\left(\sqrt{c^2+h^2}F_1(c,h) - |c|\, F_2(c)\right)$$

donde $F_1(c,h)$ y $F_2(c)$ tienen por límite cero cuando $(c,h) \to (0,0)$ y c$\to 0$ respectivamente. Esto es así si $h \to 0$. Tendremos entonces

$$\lim_{h\to 0} \frac{F(h)}{h^2} = D_2(D_1 f)(0,0).$$

Análogamente, empezando con la función $f(h,y) - f(0,y)$, se obtiene que el límite es $D_1(D_2 f)(0,0)$.

$\square$

Corolario 2.10. *Sea f una función definida en un entorno de un punto $a \in R^n$ a valores en R, que admite funciones derivadas parciales segundas y que éstas son continuas en a. Entonces para cada i, j se cumple $D_i(D_j f)(a) = D_j(D_i f)(a)$.*

Demostración. Basta aplicar el teorema anterior y el criterio de diferenciabilidad del teorema 2.3.

$\square$

Una función definida en R^n se dice de clase $\mathcal{C}^2$ si admite derivadas parciales de segundo orden y son todas ellas continuas. Para estas funciones se cumple entonces que $D_i(D_j f) = D_j(D_i f)$.

Ejemplo 2.7. *No existe ninguna función f de R^2 en R de clase $\mathcal{C}^2$ tal que $D_1 f = 3y$, $D_2 f = 5x$. En efecto, $D_2(D_1 f) = 3$ mientras que $D_1(D_2 f) = 5$.*

2.1.3 Ejercicios

1. Calcula, donde existan, las funciones derivadas parciales de las funciones:

   $f(x,y) = x^{2y}$.

   $f(x,y) = \int_1^{x+y} g(t)\,dt$, donde g es una función continua de R en R.

   $f(x,y) = \int_x^{y} g(t)\,dt$, donde g es una función continua de R en R.

   $f(x,y) = (\sin(x\cos y), y^x)$.

2. Calcula las derivadas direccionales, en el caso en que existan, así como la diferenciabilidad en $(0,0)$ de las funciones:

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

   b) $f(x,y) = |x+y|$.

   c) $f(x,y) = |xy|^{\frac{1}{2}}$.

3. Encuentra el plano tangente a la superficie $z = 4x^2y - 2xy$ en el punto $(0,1,0)$.

4. Da la matriz jacobiana en (x,y) de la función $g \circ f$ donde $f(x,y) = (x^2, e^{x+y}, x-y)$, $g(u,v,z) = (u - v^2, e^z)$.

5. Calcula las diferenciales de las siguientes funciones, a valores en R, en cada punto en que existan:

   $f(x,y) = g(x+y)$ con g derivable.

   $f(x,y) = g(x^y, y^x)$ con g diferenciable en R^2.

   $f(x,y) = g(x, h(x,y))$ con g y h diferenciables en R^2.

   $f(x,y) = \int_1^{xy} g(t)\,dt$ con g continua en R.

6. Expresa las derivadas parciales de segundo orden de las funciones

   (a) $F(x,y) = f(xy, x+y)$ con $f \in \mathcal{C}^2$.

   (b) $F(x,y) = h(x, g(xy))$ con h y g de clase $\mathcal{C}^2$

7. Prueba que toda función diferenciable de R^n en R con $\|df_x\| \leq K$ para todo x, es uniformemente continua.

8. Sea $f : R^2 \to R$ tal que $D_1 f \equiv 0$. Prueba que f no depende de la primera variable. Aplica esta observación para encontrar la expresión general de una función f definida en R^2 a valores en R, de clase $\mathcal{C}^1$ tal que $D_1 f = 3x$. De estre éstas halla una tal que $D_2 f = 6$.

9. Sea $g : R^2 \to R^2$ diferenciable en $(0,0)$ tal que $g \circ f$ es la aplicación identidad en un entorno de $(0,0)$ y $f(x,y) = (e^x \sin y, (e^x - 1)\cos y)$. Calcula la diferencial de g en $(0,0)$.

10. Sea $f(x,y) = \frac{xy^2}{x^2+y^4}$ si $(x,y) \neq (0,0)$, $f(0,0) = 0$. Prueba que tiene derivadas direccionales en todos los puntos pero no es continua en $(0,0)$.

11. Prueba que la función $f(x,y) = \frac{|x|y}{(x^2+y^2)^{\frac{1}{2}}}$ si $(x,y) \neq (0,0)$, $f(0,0) = 0$ es continua y tiene derivadas parciales en $(0,0)$ pero no es diferenciable en este punto.

12. Sean f_1, f_2 dos funciones definidas en R^2 a valores en R tales que en un punto a existen los límites
$$\lim_{x\to a} \frac{f_i(x) - f_i(a)}{x_i - a_i}, i = 1, 2.$$
(Obsérvese que es una condición más fuerte que la existencia de $\frac{\partial f_i}{\partial x_i}(a)$). Prueba que la función $f(x) = f_1(x) + f_2(x)$ es diferenciable en a.

13. Determina, según los valores de m, la continuidad, diferenciabilidad y la pertenencia a la clase $\mathcal{C}^1$ de las funciones:

    a) $f(x,y) = (x+y)^m \sin\left(x^2+y^2\right)^{-\frac{1}{2}}$ si $(x,y) \neq (0,0)$, $f(0,0) = 0$.

    b) $f(x,y) = \frac{y^m}{x^2+y^2}$ si $(x,y) \neq (0,0)$, $f(0,0) = 0$.

14. Prueba que una función f de R^n en R^m tal que $\|f(x)\| \leq C\|x\|^2$ para C constante, es diferenciable en el origen.

15. Sea f una función diferenciable en $\mathcal{E}(a,r) \subset R^n$ a valores en R. Prueba que si $df_x = 0$ para todo $x \in \mathcal{E}(a,r)$, la función es constante.

16. Sea $f : R^n \to R^n$ diferenciable en un punto $a \in R^n$ con df_a inyectiva. Probar que $\inf_{\|h\|=1} \|df_a(h)\| > 0$. Deduce que existe un entorno U de a tal que si $x \in U$, $x \neq a$ se cumple $f(x) \neq f(a)$.

17. Una función f definida en un abierto A de R^n se dice que es homogénea de grado p si $f(\lambda x) = \lambda^p f(x)$ para $\lambda \in R,\ x, \lambda x \in A$. Prueba que si una tal función f es diferenciable se cumple
$$x \cdot \nabla f(x) = pf(x)\,. \text{ (Teorema de Euler).}$$
Indicación: Calcula $g'(1)$ para $g(\lambda) = f(\lambda x)$.

2.2 Diferenciales de orden superior al primero. Formula de Taylor.

2.2.1 Aplicaciones multilineales y diferenciales de orden superior al primero.

Sabemos que el conjunto de las aplicaciones lineales de R^n en R, $\mathcal{L}(R^n, R)$ es un espacio vectorial de dimensión n. Si $e_1, \ldots, e_n$ es una base de R^n, una base de $\mathcal{L}(R^n, R)$ está formada por

$w_1, \ldots w_n$ donde $w_i(e_j) = \delta_{ij}$ que es 0 si $i \neq j$ y 1 si $i = j$. Es la llamada base dual de la primera. El conjunto de las aplicaciones bilineales de R^n en R se denotará por $\mathcal{L}^2(R^n, R)$. Sus elementos son aplicaciones de $R^n \times R^n$ en R tales que son lineales en cada una de las variables. El conjunto $\mathcal{L}^2(R^n, R)$ tiene en forma natural estructura de espacio vectorial. Su dimensión es n^2. Una base está formada por $w_i \otimes w_j$ $i, j = 1, \ldots, n$, donde $w_i \otimes w_j(u, v) = w_i(u) w_j(v)$. Análogamente el espacio de las aplicaciones $k-$ lineales $\mathcal{L}^k(R^n, R)$ es un espacio vectorial de dimensión n^k. Una base estará formada por $w_{i_1} \otimes \cdots \otimes w_{i_k}$, $i_1, \ldots, i_k = 1, \ldots, n$ donde

$$w_{i_1} \otimes \cdots \otimes w_{i_k}(u_1, \ldots, u_k) = w_{i_1}(u_1) \cdots w_{i_k}(u_k).$$

Una aplicación $k-$lineal T se dice simétrica si $T(u_{i_1}, \ldots u_{i_k}) = T(u_1, \ldots, u_k)$ para cualquier permutación $i_{1,} \ldots, i_k$ de $1, \ldots, k$. Las aplicaciones multilineales que utilizaremos serán simétricas y la expresión que aparecerá más frecuentemente es la imagen de k vectores coincidentes. Cabe reseñar, aunque es un hecho que no lo utilizaremos en forma explícita, que estas imágenes determinan la aplicación multilineal. Veámoslo únicamente para el caso de las aplicaciones bilineales mediante una fórmula explícita. De la relación

$$T(u+v, u+v) = T(u,u) + T(u,v) + T(v,u) + T(v,v)$$

se sigue

$$T(u,v) = \frac{1}{2}\left(T(u+v, u+v) - T(u,u) - T(v,v)\right). \tag{2.5}$$

Si expresamos un vector x en coordenadas mediante $x = \sum x_i e_i$ y T es la aplicación bilineal, se tiene que $T(x,x) = \sum x_i x_j T(e_i, e_j)$. Se trata de un polinomio homogéneo de grado dos en las variables $x_1, \ldots, x_n$. Es pues equivalente el conocer la forma bilineal simétrica o el polinomio homogéneo de grado dos asociado. Si no da lugar a confusión, por ejemplo si un subíndice determina el grado del polinomio homogéneo, se sustituirá $T_k(x, \ldots, x)$ simplemente por $T_k(x)$. Se pueden dar fórmulas análogas a la 2.5 para aplicaciones $k-$lineales.

Ejemplo 2.8. *Consideremos una aplicación bilineal $T_2 = \sum_{i,j=1}^n a_{ij} w_i \otimes w_j$. Esta aplicación es simétrica si y sólo si para cada i, j $T_2(e_i, e_j) = T_2(e_j, e_i)$, es decir, $a_{ij} = a_{ji}$. Su polinomio homogéneo asociado es $T_2(x) = \sum a_{ij} x_i x_j$. Para una aplicación 3-lineal simétrica $\sum a_{ijk} w_i \otimes w_j \otimes w_k$ su polinomio homogéneo asociado es $\sum a_{ijk} x_i x_j x_k$.*

Veamos como los espacios de las aplicaciones multilineales de orden k se pueden identificar con las aplicaciones lineales de R^n en el espacio de las aplicaciones multilineales de orden $k-1$. Es decir $\mathcal{L}^k(R^n, R) \simeq \mathcal{L}\left(R^n, \mathcal{L}^{k-1}(R^n, R)\right)$. Basta asociar a cada elemento $T \in \mathcal{L}^k(R^n, R)$ la aplicación de $\mathcal{L}\left(R^n, \mathcal{L}^{k-1}(R^n, R)\right)$ que envía $u \in R^n$ al elemento de $\mathcal{L}^{k-1}(R^n, R)$ que a $v_1, \ldots, v_k$ le asocia $T(u, v_1, \ldots, v_k)$. Es simple comprobar que se trata de un isomorfismo.

Ejemplo 2.9. *La aplicación bilineal $T_2 = \sum_{i,j=1}^n a_{ij} w_i \otimes w_j$ se puede identificar con la aplicación de $\mathcal{L}(R^n, \mathcal{L}(R^n, R))$ que asocia a un elemento $u \in R^n$ el elemento*

$$\sum_{i,j=1}^n a_{ij} w_i(u) w_j$$

perteneciente a $\mathcal{L}(R^n, R)$.

Estamos ya en condiciones de definir las diferenciales de orden superior al primero.

Definición 2.8. *Sea f una función definida en un abierto D de R^n a valores en R, diferenciable en todos sus puntos. La función diferencial df es una aplicación de D en $\mathcal{L}(R^n, R)$ que asocia a cada punto $x \in R^n$, $df_x = \sum \frac{\partial f}{\partial x_i}(x)\, dx_{i\ x}$. Si esta función a valores en $\mathcal{L}(R^n, R) \simeq R^n$ es a su vez diferenciable en un punto a diremos que la función es diferenciable hasta el orden 2 en a. Equivale a decir que la aplicación $x \to \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\right)$ es diferenciable en a. La diferencial será una aplicación lineal de R^n en $\mathcal{L}(R^n, R)$ que podrá identificarse a un elemento de $\mathcal{L}^2(R^n, R)$. Concretamente si una base de $\mathcal{L}^2(R^n, R)$ es $dx_{i\ a} \otimes dx_{j\ a}$ la diferencial segunda es un elemento de la forma $T_2 = \sum_{i,j=1}^n \lambda_{ij} dx_{i\ a} \otimes dx_{j\ a}$.*

La coordenada λ_{ij} se podrá expresar como $\lambda_{ij} = d\left(\frac{\partial f}{\partial x_j}\right)_a (e_i) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)$. Resumiendo la diferencial segunda de la función f en el punto a es simplemente la aplicación bilineal

$$\sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, dx_{i\ a} \otimes dx_{j\ a}.$$

Por el teorema de las derivadas parciales iteradas (teorema 2.9) esta aplicación bilineal es simétrica y tendrá como polinomio homogéneo asociado

$$\sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\, x_i x_j$$

que se denotará también por $d^2 f_a(x)$.

Definición 2.9. *Se define una función diferenciable en a hasta el orden k como aquella que es diferenciable de orden $k-1$ en un entorno de a y tal que la función $x \to d^{k-1} f_x$ es diferenciable en el punto a. La diferencial de orden k en a será la aplicación $k-$lineal simétrica $\sum_{i_1,\ldots,i_k=1}^n \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(a)\, dx_{i_1\ a} \otimes \ldots \otimes dx_{i_k\ a}$ y su polinomio homogéneo asociado es*

$$\sum_{i_1,\ldots,i_k=1}^n \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(a)\, x_{i_1} \cdots x_{i_k}.$$

La primera la escribiremos $d^k f_a$ y el segundo $d^k f_a(x)$.

Todas las consideraciones que se han hecho para funciones a valores en R pueden repetirse con variaciones mínimas para funciones a valores en R^m. Ahora las aplicaciones multilineales también estarán valoradas en R^m y el espacio de estas aplicaciones se escribirá $\mathcal{L}^k(R^n, R^m)$.

La definición de funciones de clase $\mathcal{C}^1$ se extiende en forma natural a funciones de clase $\mathcal{C}^r$.

Definición 2.10. *Una función f definida en un abierto $A \subset R^n$ a valores en R^m se dice de clase $\mathcal{C}^r$ si existen todas las derivadas parciales de orden menor o igual a r y son continuas. Al conjunto de estas funciones se le designará por $\mathcal{C}^r(A)$.*

La definición anterior puede sustituirse por la definición recurrente de que la aplicación de A en $\mathcal{L}(R^n; R^m)$, $x \to df_x$ sea de clase $\mathcal{C}^{r-1}$.

La función se dice que es de clase $\mathcal{C}^\infty$ si es de clase $\mathcal{C}^r$ para todo r.

La regla de la cadena puede extenderse a diferenciales de orden superior al primero. Nosotros estamos interesados en un caso particular que vendrá dado por el siguiente teorema.

Teorema 2.11. *Sea h la aplicación de R en R^n definida por $h(t) = a + t(b-a)$ y f una función k veces diferenciable en un entorno de a, de R^n, a valores en R. Entonces la función compuesta $g(t) = f(a+t(b-a))$ es derivable hasta el orden k en un entorno de 0 y se cumple*

$$g^{(k)}(t) = d^k f_{a+t(b-a)}(b-a) = \sum_{i_1,\ldots,i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1}\ldots\partial x_{i_k}}(a+t(b-a))(b_{i_1}-a_{i_1})\ldots(b_{i_k}-a_{i_k}).$$

Si f es diferenciable k veces tan sólo en a entonces g es diferenciable en 0 hasta el orden k y vale la fórmula anterior en $t=0$.

Demostración. Veamos la demostración por inducción sobre k. Para $k=1$ es simplemente la regla habitual de la cadena pues h es diferenciable.

Supongamos que el teorema es cierto para $k-1$. Tenemos

$$g^{(k-1)}(t) = \sum_{i_1,\ldots,i_{k-1}=1}^{n} \frac{\partial^{k-1} f}{\partial x_{i_1}\ldots\partial x_{i_{k-1}}}(a+t(b-a))(b_{i_1}-a_{i_1})\cdots(b_{i_{k-1}}-a_{i_{k-1}}).$$

Aplicando la regla de la cadena a la función h compuesta con cada una de las funciones

$$\frac{\partial^{k-1} f}{\partial x_{i_1}\ldots\partial x_{i_{k-1}}}$$

tendremos

$$\begin{array}{c} g^{(k)}(t) = \sum_{i_1,\ldots,i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1}\ldots\partial x_{i_k}}(a+t(b-a))(b_{i_1}-a_{i_1})\cdots(b_{i_k}-a_{i_k}) = \\ d^k f_{a+t(b-a)}(b-a)\,. \end{array}.$$

La demostración se repite si la hipótesis de diferenciabilidad se da únicamente en el punto a.

□

También precisaremos de la siguiente proposición

Teorema 2.12. *La composición de funciones de clase $\mathcal{C}^r$ es del mismo tipo.*

Demostración. Basta aplicar reiteradamente la regla de la cadena para obtener las expresiones de las derivadas parciales de la función compuesta en términos de las derivadas parciales de ambas funciones. De ellas se sigue la existencia y continuidad de las de orden inferior o igual a r.

□

No daremos estas expresiones de las derivadas sucesivas de la función compuesta explícitamente pues son de fácil cálculo pero de compleja notación. Aconsejamos al lector que las explicite en algún caso particular no demasiado largo.

2.2.2 La fórmula de Taylor

Es bien conocido que si una función de una variable es derivable hasta el orden m en un punto a existe un único polinomio de grado menor o igual que m, $p_m(x)$, tal que

$$f(x) = p_m(x) + o_{x\to a}(|x-a|^m)$$

El polinomio $p_m(x)$ recibe el nombre de polinomio de Taylor de orden m en el punto a y tiene la expresión

$$p_m(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \ldots + \frac{f^{(m)}(a)}{m!}(x-a)^m.$$

Se trata de dar un teorema análogo para funciones de varias variables.

Teorema 2.13. *Sea D un abierto de R^n y f una función definida en D a valores en R, f diferenciable hasta el orden m en un punto a de D. Existe entonces un único polinomio $p_m(x)$ de grado menor o igual a m tal que*

$$f(x) = p_m(x) + o_{x\to a}(\|x-a\|^m),\ x \in R^n.$$

El polinomio $p_m(x)$ tiene la expresión

$$p_m(x) = \sum_{k=0}^{m} \frac{1}{k!} d^k f_a\,(x-a)\,.$$

Este polinomio se denomina polinomio de Taylor de orden m de la función f en el punto a .

Demostración. Se trata de probar que

$$\lim_{x\to a} \frac{f(x) - \sum_{k=0}^{m} \frac{1}{k!} d^k f_a\,(x-a)}{\|x-a\|^m} = 0$$

Consideremos $b \in R^n$, con $\|b-a\| = 1$, $x = a + t(b-a)$.

Sea g la función de una variable real

$$g(t) = f(a + t(b-a))$$

definida para valores suficientemente pequeños de $|t|$. Se tiene que g es derivable hasta el orden m en 0. Su derivada de orden k , para $k < m$, es $g^{(k)}(t) = d^k f_{a+t(b-a)}(b-a)$. Obsérvese que g es una función que depende de b. Puesto que $d^k f_a\,(x-a) = g^{(k)}\,(0)\,t^k$ se tratará de probar que

$$\lim_{t\to 0} \frac{g(t) - \sum_{k=0}^{m} \frac{1}{k!} g^{(k)}(0) t^k}{t^m} = 0,$$

donde el límite debe entenderse uniforme en el punto $b \in R^n$ considerado, ya que $x - a = t(b-a)$ y $\|x-a\| = |t|$. Esto significa que debemos probar que para cada $\varepsilon > 0$, existe un $\delta > 0$, independiente de b, tal que si $|t| < \delta$ se cumple que $\left|\frac{g(t) - \sum_{k=0}^{m} \frac{1}{k!} g^{(k)}(0) t^k}{t^m}\right| < \epsilon$. Obsérvese que reiteradas aplicaciones de la regla de l'Höspital nos llevan al cálculo del siguiente límite

$$\lim_{t\to 0} \frac{g^{(m-1)}(t) - g^{(m-1)}(0) - g^{(m)}(0)t}{t} = 0$$

Este límite podrá ser escrito, nuevamente en términos de la función f como:

$$\lim_{t\to 0} \frac{d^{m-1}f_{a+t(b-a)}(b-a) - d^{m-1}f_a(b-a) - d^m f_a(b-a)t}{t} = 0.$$

Puesto que $d^{m-1}f_{a+t(b-a)}(b-a) = \sum \frac{\partial^{m-1}f(a+t(b-a))}{\partial x_{i_1}\dots\partial x_{i_{m-1}}}(b_{i_1}-a_{i_1})\dots(b_{i_{m-1}}-a_{i_{m-1}})$ equivaldrá a probar que si $h(x) = \frac{\partial^{m-1}f(x)}{\partial x_{i_1}\dots\partial x_{i_{m-1}}}$

$$\lim_{t\to 0} \frac{h(a+t(b-a)) - h(a) - dh_a(b-a)t}{t} = 0$$

uniformemente en b. Pero esto es precisamente la expresión de la diferenciabilidad de h en el punto a que era la hipótesis del teorema. Esto prueba la existencia del polinomio de Taylor.

Veamos la unicidad del polinomio. Si $p(x)$ y $q(x)$ fuesen dos polinomios de grado menor o igual a m que cumpliesen la condición del límite, tendríamos que el polinomio diferencia $r(x) = p(x) - q(x)$ cumpliría

$$\lim_{x\to a} \frac{r(x)}{\|x-a\|^m} = 0$$

Escribiendo, como antes, $x - a = t(b-a)$, tendremos que $r(a+t(b-a))$ será un polinomio en una variable t , de grado menor o igual a m, tal que $\lim_{t\to 0} \frac{r(a+t(b-a))}{t^m} = 0$. Esto implica que este polinomio es idénticamente nulo. Esto es válido para todo b y, por lo tanto, $r(x)$ es idénticamente nulo.

$\square$

Ejemplo 2.10. *Calculemos el polinomio de Taylor de grado 5 de la función* $\sin(x+y)$ *en el punto* $(0,0)$. *Se puede utilizar directamente la expresión del polinomio de Taylor. Otra forma de operar es la siguiente. Sabemos que* $\sin t = t - \frac{t^3}{3!} + \frac{t^5}{5!} + o_{t\to 0}(t^5)$. *Tendremos entonces*

$$\sin(x+y) = x + y - \frac{(x+y)^3}{3!} + \frac{(x+y)^5}{5!} + o_{|x+y|\to 0}(|x+y|^5)$$

Ahora bien, puesto que $|x+y| \le 2\,\|(x,y)\|$, *se sigue que la función* $o_{|x+y|\to 0}(|x+y|^5)$ *es también* $o_{(x,y)\to(0,0)}(\|(x,y)\|^5)$. *De aquí que el polinomio buscado es* $x + y - \frac{(x+y)^3}{3!} + \frac{(x+y)^5}{5!}$.

Término complementario en la fórmula de Taylor.

Si suponemos que la función f antes considerada, es diferenciable hasta el orden $m+1$ en un entorno del punto a se puede dar una información más explícita sobre la diferencia entre la función y el polinomio de Taylor de orden m. Se tiene el siguiente teorema:

Teorema 2.14. *Sea f una función definida en $B(a,\varepsilon) \subset R^n$, a valores en R, diferenciable hasta el orden $m+1$ en todos los puntos de $B(a,\varepsilon)$. Para cada $x \in B(a,\varepsilon)$, existe $0 < \theta < 1$ tal que*

$$f(x) = \sum_{k=0}^{m} \frac{1}{k!} d^k f_{(a)}(x-a) + \frac{1}{(m+1)!} d^{m+1} f_{(a+\theta(x-a))}(x-a)$$

Demostración. Consideremos la función $g(t) = f(a+t(x-a))$. Estará definida en el intervalo $[0,1]$ y será derivable hasta el orden $m+1$ en él. Podremos aplicar la formula de Taylor en 0, hasta le orden m, con el término complementario de Lagrange. Tomemos $t = 1$. Existirá $0 < \theta < 1$ tal que

$$g(1) = \sum_{k=0}^{m} \frac{1}{k!} g^{(k)}(0) + \frac{1}{(m+1)!} g^{(m+1)}(\theta) \quad .$$

La expresión de las derivadas de la función g en términos de la función f nos da el teorema. □

Corolario 2.15. Teorema del valor medio. *Sea f una función a valores en R, diferenciable en todos los puntos de un abierto convexo U, es decir, tal que si $a, b \in U$, el intervalo de extremos a y b está contenido en U. Existirá entonces $0 < \theta < 1$ tal que*

$$f(b) - f(a) = df_{a+\theta(b-a)}(b-a).$$

Obsérvese que la función debe ser a valores en R. Se trata de un caso particular del teorema anterior. Si la función de partida fuese a valores en un espacio R^m con $m > 1$ para cada función coordenada tendríamos un punto intermedio que en principio podrían ser distintos.

Es interesante observar que, en las condiciones del corolario anterior, si $\|df_x\| \leq M$ para $x \in U$ tendremos que

$$\|f(b) - f(a)\| \leq M \|b-a\| \,. \tag{2.6}$$

Definición 2.11. *Una función que cumple la condición 2.6 para todos los puntos a y b de un conjunto se dice de Lipschitz en este conjunto.*

Hemos visto que toda función diferenciable en un convexo con diferencial acotada es de este tipo. Obsérvese que de aquí se deduce inmediatamente que, en las condiciones anteriores, si la diferencial de la función es cero en todos sus puntos, la función es constante.

2.2.3 Extremos de una función

Definición 2.12. *Sea A un abierto de R^n y $f : A \to R$. Se dice que f tiene un máximo (mínimo) local en un punto $a \in A$ si existe $\varepsilon > 0$ tal que $f(x) \leq f(a)$ (resp. $f(x) \geq f(a)$) para $\|x-a\| < \varepsilon$. En ambos casos se dice que la función tiene un extremo local en a.*

En ocasiones, si $f(x) < f(a)$ (resp. $f(x) > f(a)$) para $0 < \|x-a\| < \varepsilon$ se dice que el máximo (resp. mínimo) es estricto.

Es bien conocido que si A es un abierto de R y f es derivable una condición necesaria para que la función tenga un máximo o mínimo en a es que la derivada sea nula en este punto. El teorema análogo para funciones de varias variables es el siguiente.

Teorema 2.16. *Sea A un abierto de R^n y $f : A \to R$ una función diferenciable. Si f tiene en $a \in A$ un extremo local se cumple que $df_a = 0$.*

Demostración. Sea $u \in R^n$. Consideremos la función de una variable real λ, definida en un entorno de 0, mediante la función $g(\lambda) = f(a + \lambda u)$. Será una función derivable con un extremo en $\lambda = 0$ y por lo tanto $g'(0) = df_a(u) = 0$.

$\square$

La condición necesaria de extremo $df_a = 0$ desde luego no es suficiente. Esto es bien conocido para funciones de una variable real. Por ejemplo $y = x^3$ tiene derivada cero en el origen pero no tiene extremos locales. Ocurre lo mismo para funciones de varias variables. Considérese los siguientes ejemplos

Ejemplo 2.11.

1. *La función $f(x, y) = (x + y)(x - y)$ es de clase C^∞ , tiene sus derivada parciales en el origen nulas y no tiene un extremo en él. Existen puntos en cualquier entorno del origen en que la función toma valores positivos y otros en que toma valores negativos.*

2. *La función $f(x, y) = (y - 2x^2)(y - 3x^3)$ es tal que su restricción a cada recta que pasa por el origen tiene un mínimo local en este punto. Sin embargo, en todo entorno del origen existen puntos en que la función toma valores negativos y por tanto la función no tiene extremo local en el origen.*

Definición 2.13. *Sea A un abierto de R^n y $f : A \to R$ una función diferenciable. Si $df_a = 0$ para un punto $a \in A$ se dice que a es un punto estacionario para la función f.*

Se trata ahora de dar condiciones suficientes para que un punto estacionario de una función sea un extremo local. En el siguiente teorema se dan unas condiciones en términos de la diferencial segunda de la función. Es un teorema análogo al que nos establece que una función de una variable real con $f'(a) = 0$ y $f''(a) \neq 0$ tiene un extremo en a.

Teorema 2.17. *Sea A un abierto de R^n y $f : A \to R$ una función diferenciable de orden 2 en un punto $a \in A$ que es estacionario para la función f. Se tiene*

a) Si $d^2f_a(u) > 0$ para todo $u \neq 0$, f tiene un mínimo local en a.

b) Si $d^2f_a(u) < 0$ para todo $u \neq 0$, f tiene un máximo local en a.

c) Si $d^2f_a(u)$ toma valores positivos y negativos al variar u, f no tiene en a ni máximo ni mínimo local. En este caso se dice que el punto es de ensilladura.

Demostración. Veamos el apartado a). Si $d^2f_a(u) > 0$ para $u \neq 0$ se tiene que $M = \min_{\|v\|=1} d^2f_a(v) > 0$ y, por tanto para cada u se tiene $d^2f_a(u/\|u\|) \geq M$, luego $d^2f_a(u) \geq M\|u\|^2$.

Puesto que a es un punto estacionario la fórmula de Taylor nos dará

$$f(a + u) - f(a) = \frac{1}{2}d^2f_a(u) + g(u)\|u\|^2 , \text{ con } g(u) \to_{u\to 0} 0.$$

Sea $0 < \varepsilon < \frac{M}{4}$. Existe $\delta > 0$ tal que $|g(u)| < \varepsilon$ si $\|u\| < \delta$. Por tanto

$$f(a+u) - f(a) \geq \frac{1}{2}M \|u\|^2 - \varepsilon \|u\|^2 \geq \frac{M}{4} \|u\|^2$$

si $\|u\| < \delta$ y, por lo tanto, f tiene un mínimo local en a.

El apartado b) se demuestra análogamente. Veamos c). Sean u_1 y u_2 tales que $d^2 f_a(u_1) > 0$ y $d^2 f_a(u_2) < 0$. Tendremos

$$f(a+\rho u_i) - f(a) = \frac{1}{2}(d^2 f_a(u_i) + g_i(\rho) \|u_i\|^2)\rho^2 \text{ con } g_i(\rho) \to 0 \text{ si } \rho \to 0,$$

y $f(a+\rho u_i) - f(a)$ tendrá el mismo signo que $d^2 f_a(u_i)$ para valores pequeños de ρ.

□

La forma cuadrática $d^2 f_a(u) = \sum \frac{\partial^2 f}{\partial x_i \partial x_j}(a) u_i u_j$ juega, en consecuencia, un papel relevante en las condiciones de extremo. La condición $d^2 f_a(u) > 0$ (resp.< 0) si $u \neq 0$ se expresa diciendo que la forma cuadrática es definida positiva (resp. negativa). No obstante si la forma es degenerada, es decir, para algunos valores $u \neq 0,\ d^2 f_a(u) = 0$, esta forma no determina la existencia o no de extremos. No obstante el apartado c) implica que si a es un extremo y $d^2 f_a$ es no idénticamente nula, entonces debe ser $d^2 f_a(u) \geq 0$ para todo u o bien $d^2 f_a(u) \leq 0$ para todo u.

Ejemplo 2.12. 1. *La función $f(x,y) = -(x^2+y^2) + x^3$ tiene en el origen un punto estacionario y $d^2 f_{(0,0)}(x,y) = -(x^2+y^2)$ es definida negativa y por tanto se trata de un máximo*

2. *La función $f(x,y) = -x^2$, tiene en el origen un punto estacionario. Su diferencial segunda en este punto es una forma negativa pero degenerada. En este caso la función tiene un máximo, aunque no estricto ya que en todos los puntos de la forma $(0,y)$ toma el valor 0. La función $g(x,y) = -x^2 - y^4$ está en una situación similar a la anterior pero ahora el máximo es estricto. Sin embargo, la función $h(x,y) = -x^2 + y^4$ no tiene máximo ni mínimo ya que su restricción a la recta $y = 0$ tiene un máximo, mientras que la restricción a $x = 0$ tiene un mínimo.*

Es importante pues conocer cuando una forma cuadrática simétrica $\sum a_{ij} x_i x_j$, $a_{ij} \in R$ es definida positiva o definida negativa. Si denominamos Δ_k al determinante de la matriz obtenida de la (a_{ij}) suprimiendo las n-k últimas filas y columnas se tiene el siguiente criterio cuya demostración puede verse en los textos de álgebra multilineal (véase por ejemplo Prop. 2.3 cap XI en Algebra lineal y Geometría, Manuel Castellet, Irene Llerena):

Teorema 2.18. *La forma cuadrática simétrica $\sum a_{ij} x_i x_j$, $a_{ij} \in R$ es definida positiva (resp. negativa) si y sólo si $\Delta_1 > 0, \Delta_2 > 0, \ldots \Delta_n > 0$ (resp. $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \ldots$).*

En dimensión dos puede darse una demostración elemental del teorema anterior. Veámosla. Una forma de orden dos en R^2 se puede expresar como

$$ax^2 + by^2 + 2cxy.$$

Si a o b son ceros, la forma no es trivialmente ni definida positiva ni negativa. Fuera de la forma idénticamente nula tendremos que en las notaciones del teorema anterior $\Delta_2 = ab - c^2 < 0$.

Podemos suponer entonces que a y b son no nulos. Si dividimos por $y^2 \neq 0$ la forma tendremos

$$a\left(\frac{x}{y}\right)^2 + b + 2c\frac{x}{y}.$$

Es un polinomio de segundo grado en $\frac{x}{y}$. Es ya evidente que toma valores estrictamente positivos si $-\Delta_2 = c^2 - ab < 0$ y $\Delta_1 = a > 0$. Toma valores estrictamente negativos si $-\Delta_2 = c^2 - ab < 0$ y $\Delta_1 = a < 0$. Para $y = 0$, $x \neq 0$ la forma toma el valor ax^2 que tiene por signo el de a. Todo ello da el teorema en dimensión dos.

Ejemplo 2.13.

1. *Calculemos los puntos estacionarios de la función*

$$f(x, y) = 2x^3 - 3x^2 + 2y^3 + 3y^2.$$

Se tiene que

$$\frac{\partial f}{\partial x} = 6x^2 - 6x$$
$$\frac{\partial f}{\partial y} = 6y^2 + 6y$$

Los puntos estacionarios son los ceros de las funciones anteriores, lo que nos da como soluciones $(0,0), (1,0), (0,-1), (1,-1)$. *Los polinomios homogéneos asociados a las diferenciales segunda en estos puntos darán respectivamente* $-6x^2+6y^2,\ 6x^2+6y^2,\ -6x^2-6y^2,\ \ -6x^2-6y^2$. *Tendremos entonces que* $(1,0)$ *será un mínimo,* $(0,-1)$ *un máximo y los puntos* $(0,0)$ *y* $(1,-1)$ *serán de ensilladura.*

2. *Se trata ahora de hallar los máximos y mínimos absolutos de la función*

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

sobre $M = \left\{(x, y, z) \in R^3; x^2 + y^2 + z^2 \leq 2^2,\ z \geq 1\right\}$.

La función es continua y M *es compacto. Existirán entonces máximos y mínimos de la función. Si estos puntos están en el interior de* M *deberán ser puntos estacionarios de* f *y, por tanto, cumplirán*

$$\frac{\partial f}{\partial x} = 2x + 1 = 0$$
$$\frac{\partial f}{\partial y} = 2y + 1 = 0$$
$$\frac{\partial f}{\partial z} = 2z = 0.$$

La solución del sistema es $(-\frac{1}{2}, -\frac{1}{2}, 0)$. *Este punto no está en el interior de* M *y, por tanto, los extremos de la función se hallarán en la frontera de* M. *Es decir, sobre los puntos en que* $z = 1$ *y* $x^2 + y^2 \leq 3$, *o bien los puntos en que* $x^2 + y^2 + z^2 = 2^2,\ z \geq 1$.

En el primer caso se debe estudiar la función de dos variables reales $g(x, y) = x^2 + y^2 + x + y + 1$ *definida en los puntos de* R^2 *tales que* $x^2 + y^2 \leq 3$. *Nuevamente en los puntos del interior de este dominio deberá cumplirse*

$$\frac{\partial g}{\partial x} = 2x + 1 = 0$$
$$\frac{\partial g}{\partial y} = 2y + 1 = 0.$$

La solución es $\left(-\frac{1}{2}, -\frac{1}{2}\right)$ *, que está en el interior del dominio. Consideremos ahora la función* g *restringida a* $x^2 + y^2 = 3$. *Si escribimos* $x = \sqrt{3}\cos t,\ \ y = \sqrt{3}\sin t$ *deberemos calcular los extremos de la función* $h(t) = \sqrt{3}\cos t + \sqrt{3}\sin t + 4$ *en* $[0, 2\pi]$ *. Será suficiente considerar el extremo* $t = 0$ *y los puntos de* $(0, 2\pi)$ *en que* $h'(t) = -\sqrt{3}\sin t + \sqrt{3}\cos t = 0$, *es decir,* $t = \frac{\pi}{4}$ *y* $t = \frac{5\pi}{4}$.

Finalmente sea la restricción de f *a* $x^2 + y^2 + z^2 = 2^2,\ \ z \geq 1$ *. Deberemos considerar la función* $u(x, y) = 4 + x + y$ *sobre* $4 - x^2 - y^2 \geq 1$. *No existen puntos estacionarios y la frontera ha sido considerada anteriormente.*

En resumen, los posibles extremos serán

$$\left(-\frac{1}{2}, -\frac{1}{2}, 1\right),\ (\sqrt{3}, 0, 1),\ \left(\frac{\sqrt{6}}{2}, \frac{\sqrt{6}}{2}, 1\right) \text{ y } \left(\frac{-\sqrt{6}}{2}, \frac{-\sqrt{6}}{2}, 1\right).$$

Calculando los valores de la función en estos puntos concluimos que $\left(\frac{\sqrt{6}}{2}, \frac{\sqrt{6}}{2}, 1\right)$ *es el mínimo y que* $\left(\frac{-\sqrt{6}}{2}, \frac{-\sqrt{6}}{2}, 1\right)$ *es el máximo.*

2.2.4 Ejercicios

1. Da la fórmula de Taylor hasta el orden 2 con el término complementario para las funciones:

   a) $f(x, y) = x^2y^2 + xy$ en el punto $(1, 1)$.

   b) $f(x, y) = \sin(x^2 + y^2)$ en el punto $(0, 0)$.

   c) $f(x, y) = \sin x + \cos y$ en el punto $(0, 0)$.

   d) $f(x, y) = e^{x+y}$ en el punto $(0, 0)$.

2. Calcula
$$\lim_{(x,y)\to(1,0)} \frac{\ln(x + e^y) - \frac{1}{2}(x + y - 1) - \ln 2}{\sqrt{(x-1)^2 + y^2}}.$$
   Indicación: Da la fórmula de Taylor de $f(x, y) = \ln(x + e^y)$ en $(1, 0)$.

3. Halla un polinomio $p(x, y)$ tal que la función cos$(x + y) - p(x, y)$ tenga con la función e^{x-y} un contacto de orden superior al segundo en el punto $(0, 0)$, es decir, que la diferencia sea $o_{(x,y)\to(0,0)}(x^2 + y^2)$. ¿Existe un polinomio único con esta condición?

4. Halla los extremos relativos, si existen, de las siguientes funciones:

   a) $f(x, y) = 2x^3 - 3x^2 + 2y^3 + 3y^2$.

   b) $f(x, y) = y^2 - x^3$

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

   d) $f(x, y) = x^3 + 3xy^2 - 3x^2 - 3y^2$.

5. Sea $f(x) = (\sin x, \cos x)$, definida en $[0, 2\pi]$. ¿Existe algún $t \in (0, 2\pi)$ tal que $f(2\pi) - f(0) = 2\pi f'(t)$?

6. Consideremos $a_1, \ldots, a_n \in R^n$. Prueba que la función $f(x) = \sum_{i=1}^{n} \|x - a_i\|^2$, definida en R^n tiene un mínimo en $\frac{a_1 + \ldots + a_n}{n}$.

7. Prueba que $f(x, y) = (y - x^2)(y - 3x^2)$ no tiene un extremo local en $(0, 0)$ pero que su restricción a cada recta que pasa por el origen si lo tiene.

8. Halla los máximos y mínimos absolutos de la función $f(x, y) = 2 + \sin x + \sin y + \sin(x + y)$ en el conjunto

$$K = \left\{(x, y) \in R^2;\ 0 \leq x,\ 0 \leq y,\ x + y \leq 2\pi\right\}.$$

9. Consideremos la función definida en R^2 por $f(x, y) = (x^2 + 2y^2)\, e^{-\left(x^2 + y^2\right)}$. Prueba la existencia de extremos absolutos y calcúlalos.

2.3 El teorema de la función inversa e implícita

2.3.1 El teorema de la función inversa

Si T es una aplicación lineal de R^n en R^n es bien conocido que esta aplicación tiene inversa si y solamente si $\det T \neq 0$. Se trata de dar un teorema que permita conocer si una aplicación con cierta regularidad admite función inversa del mismo tipo. Si suponemos que las funciones son diferenciables se puede obtener fácilmente una condición necesaria. Sea g la función inversa de f, tendremos que $g \circ f$ es la aplicación identidad. Supuestas ambas diferenciables deducimos de la regla de la cadena que para cada punto a se tiene $dg_{f(a)} \circ df_a = I$. En particular la aplicación df_a es un isomorfismo lineal. El teorema de la función inversa dirá que esta condición es suficiente para que localmente exista la función inversa.

Teorema 2.19. *Sea f una función de clase $\mathcal{C}^r$, $r \geq 1$, definida en un abierto A de R^n con valores en R^n. Supongamos que en un punto $a \in A$ la aplicación df_a es un isomorfismo. Entonces existe un abierto $U \subset A$, entorno del punto a, tal que*

a) La aplicación f es inyectiva en U.

b) La imagen $f(U)$ es un conjunto abierto de R^n.

c) La aplicación inversa de la aplicación biyectiva f de U en $f(U)$ es de clase $\mathcal{C}^r$.

Observaciones

1. La condición de que df_a sea un isomorfismo equivale simplemente a decir que $\det df_a \neq 0$. Esto es, que el determinante de la matriz jacobiana sea no nulo

$$\begin{vmatrix} D_1 f_1(a) & \dots & D_n f_1(a) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ D_1 f_n(a) & \dots & D_n f_n(a) \end{vmatrix} \neq 0$$

A este determinante se le denomina determinante jacobiano y se escribe $\frac{\partial(f_1,\dots,f_n)}{\partial(x_1,\dots,x_n)}(a)$.

2. La existencia de la función inversa es sólo de tipo local, existe un entorno del punto considerado en que la función f restringida a este entorno admite inversa. Posteriormente veremos mediante algún ejemplo que, aunque $\det df$ sea no nulo en todos los puntos, no se puede asegurar la existencia de una inversa global. Esta es una situación diferente de lo que ocurre si suponemos que la función es de una variable y está definida en un intervalo. En este caso, supuesta la función de clase C^1 con derivada no nula en todos los puntos, es entonces estrictamente monótona y existe la función inversa definida globalmente.

3. Como veremos, el teorema no da una forma de explicitar la función inversa. Simplemente asegura su existencia y prueba que es de clase $\mathcal{C}^r$. No obstante la diferencial y las derivadas sucesivas de la función inversa pueden explicitarse como consecuencia de la regla de la cadena. Ya veremos ejemplos al respecto.

Demostración. Veamos en primer lugar que existe un entorno de a en el que f es inyectiva. Sea $\varepsilon > 0$ tal que $\mathcal{E}(a,\varepsilon) \subset A$. La función f será inyectiva en $\mathcal{E}(a,\varepsilon)$ si para cada dos puntos $x, y \in \mathcal{E}(a,\varepsilon)$ la condición $f(x) = f(y)$ implica $x = y$. Consideremos dos tales puntos. Apliquemos el teorema del valor medio (2.15) a cada una de las funciones coordenadas f_i. Existirán z_i pertenecientes al intervalo de extremos x e y tales que

$$0 = f_i(x) - f_i(y) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(z_i)(y_j - x_j) \quad i = 1, \dots, n. \tag{2.7}$$

Puesto que por hipótesis $\det\left(\frac{\partial f_i}{\partial x_j}(a)\right) \neq 0$ y las funciones derivadas parciales de las f_i son continuas, existirá $\epsilon > 0$ tal que para cada $z_i \in \mathcal{E}(a,\varepsilon)$, $\det\left(\frac{\partial f_i}{\partial x_j}(z_i)\right) \neq 0$. Tendremos ahora que el sistema lineal homogéneo 2.7 tendrá como única solución $x = y$. Hemos probado el apartado a).

Probemos b). Sea $0 < \delta < \epsilon$. La función f sobre el compacto $\overline{\mathcal{E}(a,\delta)}$ es inyectiva y continua y por tanto se trata de un homeomorfismo. Consideremos $S = \{x \in R^n; \|x - a\| = \delta\}$. Sea $m = \min\{\|f(x) - f(a)\|;\ x \in S\}$. Por la inyectividad de f se tiene $m > 0$. Veamos que $f(\mathcal{E}(a,\delta)) \supset \mathcal{E}\left(f(a), \frac{m}{2}\right)$. En efecto, sea $y \in \mathcal{E}\left(f(a), \frac{m}{2}\right)$. Se trata de ver que la función $d^2(x) = \|f(x) - y\|^2$ toma el valor cero en algún punto de $\mathcal{E}(a,\delta)$. Desde luego alcanza el mínimo en algún punto del compacto $\overline{\mathcal{E}(a,\delta)}$ puesto que la función es continua. Este mínimo no se puede alcanzar sobre la frontera S ya que para $x \in S$, $\|f(x) - y\| \geq \|f(x) - f(a)\| - \|f(a) - y\| \geq m - \frac{m}{2} = \frac{m}{2}$ mientras que la imagen del punto a cumple $\|f(a) - y\| < \frac{m}{2}$.

Sea entonces un punto $z \in \mathcal{E}(a,\delta)$ donde d^2 alcanza un mínimo. Puesto que esta función es diferenciable deberá tener su diferencial cero en este punto. Tendremos

$$\sum_{i=1}^{n} (f_i(z) - y_i) \frac{\partial f_i}{\partial x_j}(z) = 0 \quad j = 1, \ldots, n.$$

Puesto que $\det\left(\frac{\partial f_i}{\partial x_j}(z)\right) \neq 0$ se sigue que $f(z) = y$ y por tanto hemos probado que

$$f(\mathcal{E}(a,\delta)) \supset \mathcal{E}\left(f(a), \frac{m}{2}\right).$$

Si llamamos U a la antiimagen por f en $\mathcal{E}(a,\delta)$ de $\mathcal{E}\left(f(a), \frac{m}{2}\right)$ la aplicación f será biyectiva de U en $\mathcal{E}\left(f(a), \frac{m}{2}\right)$. Hemos probado b).

Se trata de ver que la aplicación f^{-1} definida en $f(U)$, a valores en U es de clase C^r. Sea ahora $x \in U$ y $y = f(x)$. Veamos que f^{-1} es diferenciable en y. Sabemos que, como consecuencia de la regla de la cadena, si f^{-1} es diferenciable en y su diferencial deberá ser $(df_x)^{-1}$. Veamos que, efectivamente, es diferenciable. Sea $z = f(u)$, $u \in U$. Tendremos

$$\begin{gathered}\left\|f^{-1}(z) - f^{-1}(y) - (df_x)^{-1}(z-y)\right\| = \\ \left\|u - x - (df_x)^{-1}(df_x(u-x) + o_{u\to x}(\|u-x\|))\right\| = \left\|(df_x)^{-1}(o_{u\to x}(\|u-x\|))\right\| \leq \\ \left\|(df_x)^{-1}\right\| \|o_{u\to x}(\|u-x\|)\|.\end{gathered} \tag{2.8}$$

Se trata de probar que $f^{-1}(z) - f^{-1}(y) - (df_x)^{-1}(z-y)$ es de la forma $o_{z\to y}(\|z-y\|)$. Deberemos entonces acotar superiormente $\|u-x\|$ por $k\|z-y\|$. Tendremos

$$\begin{gathered}\|z-y\| = \|f(u) - f(x)\| \geq \|df_x(u-x)\| - \|o_{u\to x}(\|u-x\|)\| \geq \\ \frac{1}{\left\|(df_x)^{-1}\right\|}\|u-x\| - \|o_{u\to x}(\|u-x\|)\|.\end{gathered}$$

Dado $0 < \xi < \frac{1}{2\left\|(df_x)^{-1}\right\|}$ existe $\eta > 0$ tal que si $\|u-x\| < \eta$ se cumple $\|o_{u\to x}(\|u-x\|)\| < \xi\|u-x\|$ y por tanto para estos elementos u vale

$$\|z-y\| \geq \frac{1}{2\left\|(df_x)^{-1}\right\|}\|u-x\|. \tag{2.9}$$

Por otro lado, puesto que f es un homeomorfismo, existe $\delta_1 > 0$ tal que si $\|z-y\| < \delta_1$, se tiene $\|u-x\| < \eta$ y de 2.8 y 2.9

$$\begin{gathered}\left\|f^{-1}(z) - f^{-1}(y) - (df_x)^{-1}(z-y)\right\| \leq \left\|(df_x)^{-1}\right\| \xi \|u-x\| \leq \\ \xi\left\|(df_x)^{-1}\right\| 2\left\|(df_x)^{-1}\right\| \|z-y\|.\end{gathered}$$

Esto demuestra la diferenciabilidad de f^{-1} y que su diferencial en y es $\left(df_{f^{-1}(y)}\right)^{-1}$. La función df^{-1} puede obtenerse entonces como composición de las aplicaciones $y \to f^{-1}(y)$, la aplicación definida en U, tal que $x \to df_x$ y, por último, la aplicación definida en los isomorfismos lineales de R^n, $T \to T^{-1}$. Las tres aplicaciones son continuas y por tanto lo es la composición. Hemos probado que f^{-1} es de clase $\mathcal{C}^1$. Supongamos ahora que f es de clase $\mathcal{C}^r$, $r > 1$. Probemos por inducción que f^{-1} es de la misma clase. Por hipótesis de inducción la primera función $y \to f^{-1}(y)$ será de clase $\mathcal{C}^{r-1}$. La segunda aplicación $x \to df_x$ es también de clase $\mathcal{C}^{r-1}$ por la hipótesis sobre f. La tercera aplicación $T \to T^{-1}$ es de clase $\mathcal{C}^\infty$. La composición $y \to \left(df_{f^{-1}(y)}\right)^{-1} = df^{-1}_y$ será de clase $\mathcal{C}^{r-1}$ y por tanto f será de clase $\mathcal{C}^r$.

□

Corolario 2.20. *Sea A un abierto de R^n y f una función de clase $\mathcal{C}^1$ definida en A, a valores en R^n y con determinante jacobiano no nulo en todos los puntos de A. Entonces la aplicación es abierta, es decir, la imagen de todo conjunto abierto es un conjunto abierto.*

Ejemplo 2.14.

1. *Ya hemos dicho que el teorema no da, salvo en dimensión uno, la existencia de una inversa global. Por ejemplo, la función $f(x, y) = (e^x \cos y, e^x \sin y)$ tiene por determinante jacobiano e^{2x} que es no nulo en todos los puntos. Sin embargo, globalmente, en R^2, la aplicación no es inyectiva ya que $f(x, y) = f(x, y + 2k\pi)$ para cualquier k entero.*

2. *Consideremos la función definida en R^2 a valores en R^2 mediante $x = u^2 - v^2$, $y = uv$. El determinante jacobiano es $2(u^2 + v^2)$ que es no nulo salvo en el origen. El teorema permite asegurar que para cada punto distinto del $(0, 0)$ existe un entorno en el que la función es invertible con inversa de clase $\mathcal{C}^\infty$ como lo es la función de partida. Calculemos las derivadas parciales de esta función inversa, es decir de u y v respecto de x e y. La composición de la función y de su inversa será la aplicación identidad. Es decir*

$$\begin{gathered} x = u(x, y)^2 - v(x, y)^2 \\ y = u(x, y)\, v(x, y). \end{gathered}$$

Si derivamos con respecto a x tendremos

$$\begin{gathered} 1 = 2u\frac{\partial u}{\partial x} - 2v\frac{\partial v}{\partial x} \\ 0 = v\frac{\partial u}{\partial x} + u\frac{\partial v}{\partial x}. \end{gathered}$$

De aquí podemos despejar $\frac{\partial u}{\partial x}$ y $\frac{\partial v}{\partial x}$. Tendremos $\frac{\partial u}{\partial x} = \frac{u}{2(u^2+v^2)}$, $\frac{\partial v}{\partial x} = \frac{-v}{2(u^2+v^2)}$. Se calculan análogamente las derivadas parciales respecto a la variable y. Obsérvese que en estos cálculos no hemos necesitado explicitar la función inversa. Otra forma equivalente de calcular las anteriores derivadas se puede hacer calculando la matriz inversa de la matriz jacobiana de las funciones x, y respecto a las variables u, v.

Sistemas de coordenadas

Ya hemos visto que una aplicación de clase $\mathcal{C}^1$ definida en un abierto de R^n, a valores en R^n, aunque tenga su determinante jacobiano no nulo en todos sus puntos no tiene por qué ser biyectiva en todo el dominio de definición. En el caso en que sea biyectiva todo punto queda determinado por su imagen. Es por ello que recibe el nombre de sistema de coordenadas diferenciable. En este caso la aplicación inversa será también de clase $\mathcal{C}^1$.

Definición 2.14. *Sea A un abierto de R^n y $f : A \to R^n$ de clase $\mathcal{C}^1$. Supongamos que f sea biyectiva de A en su imagen y con jacobiano no nulo en todos sus puntos. Las funciones $y_i = f_i(x_1, \ldots, x_n)$ diremos que constituyen un sistema de coordenadas o, también, que se trata de un cambio de coordenadas diferenciable.*

El teorema de la función inversa nos dice que una función de clase $\mathcal{C}^1$ con determinante jacobiano no nulo constituye un sistema de coordenadas local, es decir, para cada punto existe un entorno en el que la aplicación es un sistema de coordenadas.

Observese que si $f_1, \ldots, f_n$ es un sistema de coordenadas en un abierto U, su imagen es un abierto $f(U)$ (corolario 2.20) y que, a su vez, la aplicación inversa es un sistema de coordenadas en dicho abierto.

Definición 2.15. *Una aplicación $f = (f_1, \ldots, f_n)$ que establece una biyección de clase $\mathcal{C}^1$ de un abierto en el abierto imagen, y cuya aplicación inversa es de clase $\mathcal{C}^1$ se dice que es un difeomorfismo de clase $\mathcal{C}^1$ entre los dos abiertos.*

Las funciones componentes de un difeormorfismo constituyen entonces un sistema de coordenadas diferenciable. Análogamente se definen difeomorfismos de clase $\mathcal{C}^r$.

Ejemplo 2.15. *1.* **Coordenadas polares**

Consideremos la aplicación definida en $(0, +\infty) \times (-\pi, \pi)$, a valores en R^2 que a (ρ, φ) hace corresponder

$$x_1 = \rho \cos \varphi$$
$$x_2 = \rho \sin \varphi.$$

Es una aplicación biyectiva de clase $\mathcal{C}^1$ de $(0, +\infty) \times (-\pi, \pi)$ en su imagen $A = R^2 - \{(x_1, 0)\,;\ x_1 \leq 0\}$. El determinante jacobiano es ρ, que es no nulo. La aplicación inversa, es decir, las funciones ρ y φ con dominio de definición en A son las llamadas coordenadas polares. Se denominan respectivamente módulo y argumento.

2. **Coordenadas cilíndricas**

Consideremos la aplicación definida en $(0, +\infty) \times (-\pi, \pi) \times R$, a valores en R^3 tal que a (ρ, φ, z) le corresponde

$$x_1 = \rho \cos \varphi$$
$$x_2 = \rho \sin \varphi$$
$$x_3 = z.$$

Es una aplicación biyectiva de clase $\mathcal{C}^1$ de $(0, +\infty) \times (-\pi, \pi) \times R$ en su imagen $A = R^3 - \{(x_1, 0, x_3)\,;\ x_1 \leq 0\}$. La aplicación inversa, es decir las funciones ρ, φ, z definidas en A se denominan coordenadas cilíndricas. Es una simple comprobación verificar que el determinante jacobiano es distinto de cero.

3. **Coordenadas esféricas**

Consideremos la aplicación de clase $\mathcal{C}^1$ y determinante jacobiano no nulo definida en

$$(0,+\infty)\times\left(-\frac{\pi}{2},\frac{\pi}{2}\right)\times(0,2\pi)$$

que a (r,θ,φ) le corresponde

$$\begin{aligned} x_1 &= r\cos\theta\cos\varphi \\ x_2 &= r\cos\theta\sin\varphi \\ x_3 &= r\sin\theta. \end{aligned}$$

Se trata de una aplicación biyectiva de $(0,+\infty)\times\left(-\frac{\pi}{2},\frac{\pi}{2}\right)\times(0,2\pi)$ en su imagen $A = R^3-\{(x_1,0,x_3)\,;\ x_1\geq 0\}$. La aplicación inversa, es decir las funciones r,θ,φ definidas en A se llaman coordenadas esféricas. Corresponden a la idea geométrica del módulo, longitud y latitud respectivamente.

Pueden darse variantes de estas coordenadas tomando, por ejemplo, como dominio de la longitud el intervalo $(-\pi,\pi)$ en lugar del $(0,2\pi)$.

2.3.2 El problema de la dependencia funcional

Consideremos A un abierto de R^m y $\varphi=(\varphi_1,\ldots,\varphi_n)$ una aplicación de A en R^n de clase $\mathcal{C}^r$, r≥ 1. Cada función f definida en R^n define una función $f\circ\varphi$ sobre R^m. El problema que nos planteamos es el caracterizar este conjunto de funciones, es decir conocer cuándo una función g definida sobre A es de la forma $f\circ\varphi$. En este caso se dice que g depende funcionalmente de $\varphi_1,\ldots,\varphi_n$. El teorema siguiente resuelve el problema localmente.

Teorema 2.21. *Consideremos A un abierto de R^m y $\varphi=(\varphi_1,\ldots,\varphi_n)$ una aplicación de A en R^n de clase $\mathcal{C}^r$, r≥ 1 tal que $d\varphi_{1\,x},\ldots,d\varphi_{n\,x}$ son linealmente independientes para cada $x\in A$. Entonces la aplicación φ tiene por imagen un abierto de R^n. Además, la condición necesaria y suficiente para que una función g definida en A, de clase $\mathcal{C}^r$, sea tal que para cada $p\in A$ exista una función f de la misma clase definida en un entorno de $\varphi(p)$, tal que $g=f\circ\varphi$ es que dg_x sea combinación lineal de $d\varphi_{1\,x},\ldots,d\varphi_{n\,x}$ para cada $x\in A$.*

Demostración. Veamos que $\varphi(A)$ es un abierto. Sea $p\in A$. Obsérvese en primer lugar que la hipótesis de independencia de $d\varphi_{1\,x},\ldots,d\varphi_{n\,x}$ implica que $m\geq n$ ya que suponemos que existen n elementos linealmente independientes del espacio dual de R^m. Supongamos por ejemplo que $\frac{\partial(\varphi_1,\ldots,\varphi_n)}{\partial(x_1,\ldots,x_n)}(p)\neq 0$. Definamos $\Phi: A\to R^m$ mediante

$$\Phi(x_1,\ldots,x_m)=(\varphi_1(x),\ldots,\varphi_n(x),x_{n+1},\ldots,x_m).$$

Tendremos que $d\Phi_p\neq 0$ por lo que estamos en las condiciones del teorema de la función inversa. Existirán entornos U de p y $V\times V_1$ de

$$(\varphi_1(p),\ldots,\varphi_n(p),p_{n+1},\ldots,p_m)$$

en $R^n\times R^{m-n}$ tales que $\Phi: U\to V\times V_1$ es un difeomorfismo de clase $\mathcal{C}^r$. En particular $\varphi(U)=V$ que es un abierto de R^n que contiene p y está contenido en $\varphi(A)$. Este conjunto es pues abierto.

Consideremos ahora una función g como la del enunciado del teorema tal que dg_x sea combinación lineal de $d\varphi_{1\ x},\ldots,d\varphi_{n\ x}$ para cada $x\in A$. Sea $F:V\times V_1\to R$ definida por $F=g\circ\Phi^{-1}$. Tendremos entonces $g=F\circ\Phi$. es decir

$$g\left(x\right)=F\left(\varphi_1\left(x\right),\ldots,\varphi_n\left(x\right),x_{n+1},\ldots,x_m\right).$$

Sabemos que para x en un entorno de p, $d\varphi_{1\ x},\ldots,d\varphi_{n\ x},dx_{n+1\ x},\ldots,dx_{m\ x}$ forman una base de $\mathcal{L}\left(R^m,R\right)$. Por otro lado, por la regla de la cadena, tenemos que

$$dg_x=\sum_{i=1}^{n}\frac{\partial F}{\partial x_i}d\varphi_{i\ x}+\sum_{j=n+1}^{m}\frac{\partial F}{\partial x_j}dx_{j\ x}.$$

Puesto dg_x es combinación lineal de $d\varphi_{1\ x},\ldots,d\varphi_{n\ x}$ se cumplirá

$$\frac{\partial F}{\partial x_j}\left(\varphi_1\left(x\right),\ldots,\varphi_n\left(x\right),x_{n+1},\ldots,x_m\right)=0\quad j=n+1,\ldots,m.$$

en todos los puntos de $V\times V_1,$ luego F no depende de las $m-n$ últimas variables, es decir

$$F\left(x_1,\ldots,x_n,x_{n+1},\ldots,x_m\right)=f\left(x_1,\ldots,x_n\right)$$

para una función $f\in\mathcal{C}^r\left(V\right)$. Esta función cumplirá como queríamos $g=f\circ\varphi$.

Por otra parte, es consecuencia de la regla de la cadena que si una función g es localmente de la forma $g=f\circ\varphi$, se cumple que dg_x es combinación lineal de $d\varphi_{1\ x},\ldots,d\varphi_{n\ x}$.

□

En el caso particular en que $n=m$ el teorema es consecuencia trivial del teorema de la función inversa. En este caso $\varphi_1,\ldots,\varphi_n$ es un cambio de coordenadas local y toda función g es del tipo $f\left(\varphi_1,\ldots,\varphi_n\right)$.

Ejemplo 2.16. *Consideremos las funciones $f(x,y)=2xy+2x+1$, $g(x,y)=x^2y^2+2x^2y+x^2-1$. Veamos si g depende funcionalmente de f. Las diferenciales de estas funciones son $df=\left(2y+2\right)dx+2xdy$, $dg=\left(2xy^2+4xy+2x\right)dx+\left(2x^2y+2x^2\right)dy$. El determinante de sus coeficientes es nulo y por tanto son linealmente dependientes. Así, para cada punto en que df sea no nulo, es decir distinto de $\left(0,-1\right)$, el teorema nos dice que existe un entorno en que g depende funcionalmente de f. Existe entonces una función de una variable G que cumple $g=G\circ f$ en dicho entorno.*

2.3.3 El teorema de la función implícita

Es evidente que el gráfico de una función $y=g(x)$ coincide con los ceros de la función de dos variables $(x,y)\to y-g(x)$. Se trata ahora del problema inverso, consideremos los ceros de una función f de dos variables, es decir, el conjunto de los pares (x,y) tales que $f(x,y)=0$. El problema que nos planteamos es saber si este conjunto es el gráfico de una función $y=g(x)$ o, lo que es equivalente, si para cada valor de la variable x existe un único valor de la variable y que cumple la ecuación $f(x,y)=0$. En el caso de que así sea se desean saber propiedades de esta función. Empecemos considerando algunos ejemplos.

Ejemplo 2.17. 1. *Sea la ecuación $x^2+y^2=1$. Los valores de la variable x para los que existe algún y que cumpla la ecuación son los del intervalo $[-1,1]$. Para todos estos valores salvo los de los extremos existen dos valores de y que cumplen la ecuación. Entonces la ecuación no define globalmente y como función de x. Naturalmente que si considerásemos no todos los valores de $(x,y) \in [-1,1] \times R$ sino únicamente los valores en que la variable y tomase valores no negativos, ahora a cada valor de la variable x le correspondería un único valor de la variable y, concretamente $y = +\sqrt{1-x^2}$. Obsérvese que esta función es derivable en todos los puntos del dominio de definición salvo en los extremos.*

2. *La ecuación $x - y^3 = 0$ define en todo R la variable y como función de x en el sentido de que para cada valor de la variable x existe un único valor de la variable y que cumple la ecuación. Se escribirá ésta como $y = x^{\frac{1}{3}}$. Esta función es derivable en todos los puntos salvo en $x = 0$.*

El teorema de la función implícita trata de dar condiciones para que una ecuación o un sistema de ecuaciones definan unas variables como función de otras y de que estas funciones tengan el tipo de regularidad de las ecuaciones de partida. El resultado será un teorema de existencia local. Para que los ceros de la función de partida sean el gráfico de una función deberemos restringirnos a un entorno de un punto que cumpla la ecuación.

Consideremos en primer lugar un caso bien conocido. Supongamos que tenemos un sistema homogéneo de n ecuaciones lineales y $k+n$ incógnitas

$$f_i(t_1,\ldots,t_k,x_1,\ldots,x_n) = 0 \quad i = 1,\ldots,n.$$

La condición necesaria y suficiente para que estas ecuaciones definan unívocamente las x_i como funciones de las variables $t_1,\ldots,t_k$ es que el determinante de los coeficientes de las variables $x_1,\ldots,x_n$ sea no nulo. Esta condición se puede escribir en la forma $\frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)} \neq 0$. Obsérvese que en este caso las funciones de $t_1,\ldots,t_k$ están definidas globalmente, en todo el espacio R^k. Esta situación no se dará en general en las condiciones que ahora consideraremos.

Supondremos ahora que las funciones definidas en $R^k \times R^n$ no son necesariamente lineales. Escribiremos el espacio donde están definidas las funciones en la forma $R^k \times R^n$ para hacer notar que en el teorema jugarán un papel distinto unas variables de otras. Las segundas serán las que resultarán ser funciones de las primeras. Denotaremos por t a los elementos de R^k y por x a los de R^n. Estamos ya en condiciones de enunciar el teorema de la función implícita.

Teorema 2.22. *Sea D un abierto de $R^k \times R^n$ y f una función de clase $\mathcal{C}^r$, $r \geq 1$, definida en D a valores en R^n. Consideremos un punto $(a,b) \in R^k \times R^n$ tal que $f(a,b) = 0$ y $\left.\frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)}\right|_{(a,b)} \neq 0$. Entonces existe un entorno A de a en R^k, un entorno U de (a,b) en $R^k \times R^n$ y una función $g : A \to R^n$ tal que el gráfico de la función g coincide con los ceros de la función f en U. La función g es de clase $\mathcal{C}^r$. Esta función se dice definida implícitamente por la función f.*

Demostración. Consideremos la función $F : D \to R^k \times R^n$ mediante

$$F(t,x) = (t, f(t,x)).$$

Se cumple $F(a,b) = (a,0)$, es de clase $\mathcal{C}^r$, y $\frac{\partial(F_1,\ldots,F_{n+k})}{\partial(t_1,\ldots,t_k,x_1,..,x_n)}_{(a,b)} = \frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)}_{(a,b)} \neq 0$. Estamos entonces en las condiciones del teorema de la función inversa. Existirán entornos U y V de (a,b) y de $(a,0)$ respectivamente en $R^k \times R^n$ tales que F es una aplicación biyectiva de U en V, cuya inversa G es de clase $\mathcal{C}^r$. Podemos además suponer, restringiendo si es preciso las funciones, que el entorno V es de la forma $A \times B_0$, donde A es un entorno de a en R^k y B_0 es un entorno de 0 en R^n.

LLamemos v y u a las dos componentes de la función G, es decir,

$$\begin{aligned} u(F(t,x)) &= t \\ v(F(t,x)) &= x. \end{aligned}$$

Definamos la función $g : A \to R^n$ mediante $g(t) = v(t,0)$. Desde luego es de clase $\mathcal{C}^r$. Veamos que el gráfico de g coincide con los ceros de f en U. Sea un punto del gráfico de g, $(t, g(t))$ con $t \in A$. Se cumplirá

$$f(t, g(t)) = f(t, v(t,0)) = f(G(t,0)) = 0.$$

Recíprocamente, si tenemos un punto $(t,x) \in U$ tal que $f(t,x) = 0$, tendremos $F(t,x) = (t,0)$ con $t \in A$. Se cumplirá

$$g(t) = v(t,0) = v(F(t,x)) = x$$

y por tanto (t,x) es del gráfico de g.

$\square$

Es interesante observar que la aplicación definida en $A \subset R^k$ a valores en $R^k \times R^n$ que asigna a t el punto $(t, g(t))$ es una aplicación biyectiva y continua de A en el conjunto de los puntos de U que son ceros de la función f. Si nos restringimos a un abierto con adherencia contenida en A se tratará de un homeomorfismo con su imagen.

Obsérvese que el punto crucial de la demostración del teorema de la función implícita que hemos presentado es la utilización del teorema de la función inversa. Inversamente, se puede obtener el teorema de la función inversa a partir del teorema de la función implícita. De hecho el teorema de la función inversa puede pensarse como un caso particular del teorema de la función implícita. En efecto supongamos que la función $f : R^n \to R^n$, $x \to f(x)$ está en las condiciones del teorema de la función inversa. La función de $R^n \times R^n$ en R^n que asigna a (t,x) el valor $t - f(x)$ estará en las condiciones del teorema de la función implícita y la ecuación $t - f(x) = 0$ definirá las variables x como función de las t, lo que nos da el teorema de la función inversa para la función f.

Cálculo de la diferencial de una función definida implícitamente

Sea D un abierto de $R^k \times R^n$ y f una función definida en D a valores en R^n que esté en las condiciones del teorema de la función implícita. Sea $x = g(t)$ la función implícita definida en un entorno A de un punto $a \in R^k$. Veamos que puede obtenerse dg_a sin conocerse explícitamente la función g. Sabemos que la función composición

$$\begin{array}{ccccc} A & \to & D & \to & R^n \\ t & \to & (t, g(t)) & \to & f(t, g(t)) \end{array}$$

es idénticamente cero. Su diferencial será la composición de las diferenciales y será la aplicación cero. En lenguaje matricial tendremos

$$\begin{pmatrix} \frac{\partial f_1}{\partial t_1}, & . & . & ,\frac{\partial f_1}{\partial t_k} & \frac{\partial f_1}{\partial x_1}, & . & . & ,\frac{\partial f_1}{\partial x_n} \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ \frac{\partial f_n}{\partial t_1}, & . & . & ,\frac{\partial f_n}{\partial t_k} & \frac{\partial f_n}{\partial x_1}, & . & . & ,\frac{\partial f_n}{\partial x_n} \end{pmatrix}_{(t,g(t))} \begin{pmatrix} 1 & 0 & . & . & 0 \\ 0 & 1 & 0 & . & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & . & 1 \\ \frac{\partial g_1}{\partial t_1}, & . & . & . & ,\frac{\partial g_1}{\partial t_k} \\ . & . & . & . & . \\ . & . & . & . & . \\ \frac{\partial g_n}{\partial t_1}, & & & & ,\frac{\partial g_n}{\partial t_k} \end{pmatrix}_t$$

$$= \begin{pmatrix} 0, & . & ,0 \\ . & . & . \\ 0, & . & ,0 \end{pmatrix}.$$

Puede escribirse como n sistemas de ecuaciones cada uno de ellos de n ecuaciones con n incógnitas que, debido a la condición $\frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)}_{(t,g(t))} \neq 0$ permite calcular $\frac{\partial g_i}{\partial tj}_t$ para $i = 1, \ldots, n$, $j = 1, \ldots, k$.

Obsérvese que el sistema que ha servido para obtener $\frac{\partial g_i}{\partial tj}_t$ para un j concreto y para $i = 1, \ldots, n$ se ha obtenido, simplemente, derivando por la regla de la cadena respecto a la variable t_j las ecuaciones $f_i(t, g(t)) \equiv 0$ para $i = 1, \ldots, n$. Si las ecuaciones que se obtienen se derivan nuevamente respecto a otra variable t_k se pueden obtener las derivadas segundas de las funciones g_i. Análogamente pueden obtenerse las derivadas sucesivas.

Ejemplo 2.18. *1. Consideremos la ecuación $xyz - \sin(z-2) + x + y = 0$. Define en un entorno de $(0,0,2)$, z como función implícita de x e y. En efecto, si llamamos $f(x,y,z) = xyz - \sin(z-2) + x + y$ se tiene $\frac{\partial f}{\partial z} = xy - \cos(z-2)$, que en el punto $(0,0,2)$ es no nula. Sea entonces $z = g(x,y)$ la función definida implícitamente en un entorno de $(0,0)$. Calculemos sus derivadas parciales. Derivemos respecto a x el primer miembro de la identidad*

$$xyg(x,y) - \sin(g(x,y) - 2) + x + y \equiv 0 \tag{2.10}$$

tendremos

$$yz + xy\frac{\partial g}{\partial x} - \cos(z-2)\frac{\partial g}{\partial x} + 1 = 0. \tag{2.11}$$

Por lo tanto

$$\frac{\partial g}{\partial x} = \frac{-1 - yz}{xy - \cos(z-2)}.$$

Análogamente

$$\frac{\partial g}{\partial y} = \frac{-1 - xz}{xy - \cos(z-2)}.$$

Si quisiéramos obtener derivadas sucesivas de la función g deberíamos derivar la relación 2.11 sucesivamente o bien la que se obtiene de 2.10 derivando respecto a la variable y.

Por ejemplo, para calcular $\frac{\partial^2 g}{\partial x^2}$*, derivaremos 2.11 respecto a la variable* x *teniendo en cuenta que* $z = g(x, y)$ *. Obtendremos*

$$y\frac{\partial g}{\partial x} + y\frac{\partial g}{\partial x} + xy\frac{\partial^2 g}{\partial x^2} + \sin(z-2)\left(\frac{\partial g}{\partial x}\right)^2 - \cos(z-2)\frac{\partial^2 g}{\partial x^2} = 0.$$

De donde

$$\frac{\partial^2 g}{\partial x^2} = \frac{-\sin(z-2)\left(\frac{\partial g}{\partial x}\right)^2 - y\frac{\partial g}{\partial x} - y\frac{\partial g}{\partial x}}{xy - \cos(z-2)}.$$

Obsérvese que $\frac{\partial g}{\partial x}$ *ya ha sido obtenida anteriormente y que el denominador siempre es* $\frac{\partial f}{\partial z}$ *que es no nulo en un entorno del punto considerado. Reiteramos, por último, que pueden obtenerse las derivadas de la función* g *sin conocer explícitamente esta función. Esto permite estudiar propiedades de la misma. Por ejemplo, en nuestro caso* $\frac{\partial g}{\partial x}(0,0) = \frac{\partial g}{\partial x}(0,0) = 1$*. La función* g *no podrá tener un extremo relativo en este punto.*

2. *Consideremos el sistema de ecuaciones*

$$\begin{array}{l} xu^6 + y^2v^3 = 1 \\ xy^3 + v^2u^2 = 0 \end{array} . \tag{2.12}$$

Veamos que en un entorno de $(x, y, u, v) = (0, 1, 0, 1)$ *define* x, y *como función de* u,v*. En efecto, si llamamos* f_1 *y* f_2 *a las dos funciones que definen el sistema de ecuaciones tenemos*

$$\frac{\partial(f_1, f_2)}{\partial(x, y)} = \begin{vmatrix} u^6 & 2yv^3 \\ y^3 & 3y^2x \end{vmatrix} = 3y^2xu^6 - 2y^4v^3 \tag{2.13}$$

que es no nulo en $(0, 1, 0, 1)$ *. Tendremos entonces que el sistema define las funciones* $x = g_1(u, v)$ *y* $y = g_2(u, v)$ *. Si queremos obtener, por ejemplo, las derivadas de estas funciones con respecto a la variable* u*, derivaremos respecto a esta variable en 2.12 teniendo en cuenta que* $x = g_1(u, v)$ *y* $y = g_2(u, v)$ *.*

$$\begin{array}{l} \frac{\partial g_1}{\partial u}u^6 + x6u^5 + 2y\frac{\partial g_2}{\partial u}v^3 = 0 \\ \frac{\partial g_1}{\partial u}y^3 + 3xy^2\frac{\partial g_2}{\partial u} + 2v^2u = 0 \end{array} .$$

De este sistema pueden despejarse $\frac{\partial g_1}{\partial u}$ *y* $\frac{\partial g_2}{\partial u}$ *ya que el determinante de los coeficientes coincide con el 2.13. En forma análoga pueden obtenerse las derivadas sucesivas de las funciones* g_1 *y* g_2*.*

Veamos ahora el concepto de subvariedad diferenciable. En él los teoremas de la función implícita e inversa juegan un papel fundamental.

2.3.4 Subvariedades diferenciables en R^n

Consideremos la esfera de R^3 de centro el origen de coordenadas y radio unidad, esto es, el conjunto de los puntos (x, y, z) tales que $f(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$. Se tiene $\frac{\partial f}{\partial x} =$

$2x,\ \frac{\partial f}{\partial y} = 2y,\ \frac{\partial f}{\partial z} = 2z$. En cada punto de la esfera al menos una de estas derivadas parciales será distinta de cero. De esta forma en cada punto de la esfera estaremos en las condiciones del teorema de la función implícita y una de las variables se podrá expresar localmente como función de las otras dos. Supongamos, por ejemplo, que $z = g(x, y)$ para (x, y) variando en un abierto V de R^2. La aplicación de V en R^3 que asigna al punto (x, y) el punto $(x, y, g(x, y))$ es un homeomorfismo de un abierto de R^2 en los puntos de la esfera que están en un entorno de uno de sus puntos. La esfera puede pensarse que es, en este sentido, localmente como un abierto de R^2. Muchos de los conceptos que pueden darse para un espacio euclídeo como el concepto de dimensión o la noción de función diferenciable pueden extenderse en forma natural a conjuntos como la esfera. Otros, como la estructura de espacio vectorial, se han perdido. Estos subconjuntos de R^n cuya definición precisa ahora daremos se dice que son una subvariedad de R^n. Puede darse el concepto de variedad diferenciable en general, generalización natural del que ahora daremos pero en el que nosotros no entraremos.

Definición 2.16. *Dado un subconjunto M de R^n se dice que es una subvariedad diferenciable de R^n de dimensión k y de clase $\mathcal{C}^r$ si para cada punto $p \in M$ existe un entorno U de p y $n - k$ funciones $f_{k+1}, \ldots, f_n \in \mathcal{C}^r(U)$ con diferenciales en p linealmente independientes tales que $M \cap U = \{x \in U;\ f_{k+1}(x) = \ldots = f_n(x) = 0\}$.*

Ejemplo 2.19. *El conjunto de los puntos de R^3 que cumplen $x^2 + y^2 + z^2 - 1 = 0,\ x + y + z = 1$ forma una subvariedad de R^3 de dimensión 1 y de clase $\mathcal{C}^\infty$.*

Hemos introducido el concepto de subvariedad diferenciable de dimensión k como un subconjunto que localmente son los ceros de $n-k$ funciones con diferenciales independientes. Estas funciones pueden variar al pasar de un punto a otro. En el ejemplo de la esfera con el que empezamos se trata de una única función y es la misma para todos los puntos. El primer teorema nos dirá que una subvariedad diferenciable de dimensión k puede pensarse como si localmente fuese un abierto de R^k.

Teorema 2.23. *Sea M una subvariedad diferenciable de R^n de dimensión k y de clase $\mathcal{C}^r$. Para cada punto $p \in M$ existe un entorno U del mismo, un abierto V de R^k y $g_1, \ldots, g_n \in \mathcal{C}^r(V)$, con diferenciales de rango k, tales que establecen un homeomorfismo de V en $U \cap M$.*

Demostración. Sabemos que en un entorno de p los puntos de M son los ceros de $n - k$ funciones $f_{k+1}, \ldots, f_n$ de clase $\mathcal{C}^r$ con diferenciales linealmente independientes. Supongamos, por ejemplo, que

$$\frac{\partial(f_{k+1}, \ldots, f_n)}{\partial(x_{k+1}, \ldots, x_n)}(p) \neq 0.$$

Por el teorema de la función implícita existen entornos U de $p = (p_1, \ldots, p_n)$ en R^n y V de $(p_1, \ldots, p_k)$ en R^k y funciones $g_{k+1}, \ldots, g_n \in \mathcal{C}^r(V)$ tales que la aplicación de V en $U \cap M$ que asigna a $(x_1, \ldots, x_k)$ el punto $(x_1, \ldots, x_k, g_{k+1}(x_1, \ldots, x_k), \ldots, g_n(x_1, \ldots, x_k))$ es un homeomorfismo. Las funciones que establecen el teorema son entonces $g_i(x_1, \ldots, x_k) = x_i$ para $i = 1, \ldots, k$ junto a las ya definidas $g_{k+1}, \ldots, g_n$. Obviamente sus diferenciales son de rango k ya que las diferenciales de las primeras k funciones coinciden con ellas mismas y son obviamente independientes.

$\square$

Definición 2.17. *Sea M una subvariedad diferenciable de R^n de dimensión k y de clase $\mathcal{C}^r$, $p \in M$ y U un entorno del mismo. Sea V un abierto de R^k y $g_1, \ldots, g_n \in \mathcal{C}^r(V)$, con diferenciales de rango k, tales que establecen un homeomorfismo de V en $U \cap M$. Se dice que constituyen una parametrización de $U \cap M$. La aplicación inversa estará formada por k funciones definidas en $U \cap M$ a valores en R y se dice que son un sistema de coordenadas en $U \cap M$.*

El teorema nos asegura la existencia local de una parametrización para la variedad. Si en $U \cap M$ existe una parametrización, cada uno de sus puntos queda determinado por su correspondiente elemento en el conjunto de los parámetros $V \subset R^k$. Es por ello que es natural llamar sistema de coordenadas a la correspondencia que asocia a cada elemento de $U \cap M$ su punto en el conjunto de los parámetros.

Ejemplo 2.20. *Consideremos la subvariedad de R^3 definida por $f(x,y,z) = x^2+y^2+z^2-1 = 0$. Sea $p = (0,0,1)$ un punto de la misma. Puesto que $\frac{\partial f}{\partial z}(0,0,1) \neq 0$, una parametrización en un entorno de p vendrá dada por $(x,y) \to \left(x, y, +\left(1-x^2-y^2\right)^{1/2}\right)$. Se escribirá habitualmente $x = r,\ y = t,\ z = +\left(1-r^2-t^2\right)^{1/2}$. Un sistema de coordenadas en un entorno de p vendráà dado por las funciones primera y segunda coordenada. Es decir, por las funciones que a cada punto (x,y,z) de la esfera le asignan x o y respectivamente.*

El teorema siguiente nos expresa que en cierto sentido si partimos de una "parametrización", su imagen es una subvariedad.

Teorema 2.24. *Sea A un abierto de R^k y sean $g_1, \ldots, g_n \in \mathcal{C}^r(A)$ con diferenciales de rango k. Para cada punto $a \in A$ existe un entorno $U \subset A$ tal que $g = (g_1, \ldots, g_n)$ sobre U es un homeomorfismo con su imagen $g(U) \subset R^n$. Esta es una subvariedad diferenciable de R^n de dimensión k.*

Demostración. LLamemos $(t_1, \ldots, t_k)$ a las coordenadas en R^k. Supongamos que

$$\frac{\partial(g_1, \ldots, g_k)}{\partial(t_1, \ldots, t_k)}(a) \neq 0.$$

Podemos aplicar el teorema de la función inversa a las funciones $g_1, \ldots, g_k$. Existirá entonces un entorno U de a tal que éstas darán un difeomorfismo con su imagen. Llamemos a esta imagen U_1. Será un abierto de R^k y en él estará definida la aplicación inversa $t_1 = h_1(x_1, \ldots, x_k), \ldots, t_k = h_k(x_1, \ldots, x_k)$.

Consideremos las funciones definidas en $U_1 \times R^{n-k} \subset R^n$ mediante

$$\begin{gathered}
f_{k+1}(x_1, \ldots, x_n) = x_{k+1} - g_{k+1}(h_1(x_1, \ldots, x_k), \ldots, h_k(x_1, \ldots, x_k)) \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots. \\
f_n(x_1, \ldots, x_n) = x_n - g_n(h_1(x_1, \ldots, x_k), \ldots, h_k(x_1, \ldots, x_k))
\end{gathered}$$

Los ceros de estas funciones coinciden con el conjunto imagen $g(U)$. Son $n-k$ funciones de clase $\mathcal{C}^r$ con diferenciales linealmente independientes pues $\frac{\partial(f_{k+1}, \ldots, f_n)}{\partial(x_{k+1}, \ldots, x_n)} \neq 0$ y por tanto se trata de una subvariedad de dimensión k. La función g sobre U es inyectiva, continua y con inversa $t_1 = h_1(x_1,\ldots,x_k),\ldots,t_k = h_k(x_1,\ldots,x_k)$ que es continua. Se trata pues de un homeomorfismo con su imagen.

□

Ejemplo 2.21. *1. La aplicación $x = 2t,\ y = 3t^2,\ z = t^2 + t^3$ define localmente una subvariedad diferenciable de dimensión 1 en R^3.*

2. La aplicación $x = \cos t \cos u,\ y = \cos t \sin u,\ z = \sin t$ para $0 < t < \frac{\pi}{2}$ $0 < u < \frac{\pi}{2}$ define una subvariedad diferenciable de dimensión 2 en R^3.

En las hipótesis del teorema no puede suprimirse la condición de que las diferenciales sean de rango la dimensión del espacio de los parámetros. Por ejemplo la aplicación de R en R^2 definida por las funciones $x = t^3$, $y = t^2$ tienen sus diferenciales en el origen nulas. Dejamos como ejercicio el comprobar que el conjunto imagen tiene en el origen de coordenadas una cúspide y que en un entorno de este punto no puede ser el conjunto de los ceros de una función diferenciable con diferencial no nula.

También debe observarse que en el teorema no puede afirmarse que la imagen de todo el abierto sea una subvariedad diferenciable, incluso aun suponiendo la aplicación inyectiva. Por ejemplo la función definida en $\left(-\frac{\pi}{4},\frac{\pi}{2}\right)$ a valores en R^2, definida por $x = t\left(\frac{\pi}{2} - t\right),\ y = t\cos t$ no tiene por imagen una subvariedad diferenciable. Todo entorno de $(0,0)$ tiene puntos que son imágenes por la aplicación de un entorno de 0 en R, pero también imágenes de un entorno de $\frac{\pi}{2}$. Por ello no pueden ser los ceros de una función diferenciable con diferencial no nula.

Vectores tangentes a una subvariedad

Recordemos que una curva de clase $\mathcal{C}^r$ en R^n es una aplicación γ de un intervalo de R en R^n de esta clase y que el vector $\gamma'(t)$=$\lim_{h\to 0}\frac{\gamma(t+h)-\gamma(t)}{h}$ es un vector tangente a la curva en el punto $\gamma(t)$. Se trata ahora de definir lo que se entiende por vector tangente a una variedad.

Definición 2.18. *Sea M una subvariedad diferenciable de R^n, $p \in M$ y γ una curva de clase $\mathcal{C}^r$ tal que $\gamma(0) = p$ y tal que la imagen de un entorno de 0 está contenida en M. El vector tangente $\gamma'(0)$ diremos que es un vector tangente a la subvariedad en el punto p. El conjunto de estos vectores tangentes se denominará $T_p(M)$.*

Veamos una caracterización de los vectores tangente a una variedad en términos de una parametrización.

Teorema 2.25. *Sea Φ una parametrización de un entorno de un punto p de la subvariedad M. Supongamos que el dominio de definición de Φ es un entorno de $O = (0,\ldots,0)$ en R^k y que $\Phi(O) = p$. Entonces Img$d\Phi_O = T_p(M)$.*

Demostración. Sea $v \in R^k$, $d\Phi_O(v)$ es el vector tangente a la curva $t \to \Phi(tv)$ para $t = 0$ ya que éste es el vector de R^n cuya componente i-ésima es $\sum_{j=1}^{k}\frac{\partial\Phi_i}{\partial x_j}(O)\,v_j = d\Phi_{i\,O}(v)$. Por tanto pertenece a $T_p(M)$.

Recíprocamente, sea γ una curva con $\gamma(0) = p$ contenida en M. Sea $\gamma'(0)$ su vector tangente. La curva $t \to \Phi^{-1}(\gamma(t))$ es de R^k, la imagen de 0 es O. LLamemos v a su vector tangente en O.

Es decir $v = \frac{d}{dt}\left(\Phi^{-1}\left(\gamma\left(t\right)\right)\right)(0)$. Tendremos entonces que la imagen por $d\Phi_O$ de v será $\gamma'(0)$. En efecto, $d\Phi_O(v) = \frac{d}{dt}\left(\Phi\left(\Phi^{-1}\left(\gamma\left(t\right)\right)\right)\right)(0) = \gamma'(0)$.

□

Como consecuencia se obtiene que Img$d\Phi_O$ no depende de la particular parametrización y que $T_p(M)$, por ser la imagen por una aplicación lineal de rango k, es un espacio vectorial de dimensión k.

El teorema siguiente nos caracteriza $T_p(M)$ en términos de las funciones que definen M como sus ceros.

Teorema 2.26. *Sea M una $\mathcal{C}^r$ subvariedad de R^n de dimensión k definida en un entorno de un punto p como los ceros de $g_{k+1}, \ldots, g_n$, funciones con diferenciales linealmente independientes. Entonces*

$$T_p(M) = \{v \in R^n; dg_{k+1}(v) = \ldots = dg_n(v) = 0\}$$

.

Demostración. Sea $v \in T_p(M)$. Será el vector tangente a una curva γ contenida en M con $\gamma(0) = p$. Tendremos que $dg_{i\,p}(v) = dg_{i\,p}(\gamma'(0)) = \frac{d}{dt}(g_i \circ \gamma)_{t=0} = 0$ pues $g_i \circ \gamma(t) \equiv 0$. Vemos entonces que $T_p(M) \subset \{v \in R^n; dg_{k+1}(v) = \ldots = dg_n(v) = 0\}$. De la independencia de $dg_{k+1\,p}, \ldots, dg_{n\,p}$ tendremos que el espacio $\{v \in R^n; dg_{k+1}(v) = \ldots = dg_n(v) = 0\}$ es de dimensión k y por tanto coincide con $T_p(M)$. □

Ejemplo 2.22. 1. *El espacio de los vectores tangentes a $x^2 + y^2 + z^2 - 1 = 0$ en el punto $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$ está formado por los vectores $v = (v_1, v_2, v_3)$ tales que*

$$(2xdx + 2ydy + 2zdz)_{\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)}(v) = 0.$$

Es decir, los que cumplen $v_1 + v_2 + v_3 = 0$. La ecuación del plano tangente en este punto será entonces $(x - \frac{1}{\sqrt{3}}) + (y - \frac{1}{\sqrt{3}}) + (z - \frac{1}{\sqrt{3}}) = 0$.

2. *Consideremos un polinomio $p(x, y, z)$ tal que $p(0, 0, 0) = 0$ con diferencial no nula en $(0, 0, 0)$. Veamos el plano tangente a $p(x, y, z) = 0$ en el origen de coordenadas. Estará formado por los vectores (x, y, z) tales que $\frac{\partial p}{\partial x}(0, 0, 0)\,x + \frac{\partial p}{\partial y}(0, 0, 0)\,y + \frac{\partial p}{\partial z}(0, 0, 0)\,z = 0$. Se trata de la componente homogénea lineal del polinomio p.*

2.3.5 Extremos condicionados. Los multiplicadores de Lagrange

Sea M una subvariedad de dimensión k de R^n y f una aplicación de R^n en R de clase $\mathcal{C}^1$. Se trata de determinar los máximos y mínimos relativos de la función f restringida a M. Es decir, se trata de determinar puntos $p \in M$ tales que exista $\varepsilon > 0$ tal que $f(x) \leq f(p)$ (resp. $f(x) \geq f(p)$) para los $x \in M$ tales que $\|x - p\| < \varepsilon$.

Si conocemos una parametrización de M el problema puede reducirse al estudio de máximos y mínimos relativos sin ninguna restricción. En efecto, si $\Phi : U \to M$ es la parametrización, con U un abierto de R^k, el problema se reduce al cálculo de los extremos de la función $f \circ \Phi$ en U. Sin embargo con frecuencia la variedad M vendrá dada como los ceros comunes de $n-k$ funciones $g_{k+1},\ldots,g_n$ con diferenciales linealmente independientes. A estas funciones se les suele llamar, en el contexto de este problema, ligaduras. El método de los multiplicadores de Lagrange da una condición necesaria que debe cumplir un punto para que tenga un extremo condicionado a estas ligaduras.

Teorema 2.27. *Sea U un abierto de R^n, $f\in\mathcal{C}^1(U)$ y M una subvariedad de dimensión k definida en un entorno de un punto $p\in M\cap U$ como los ceros de las funciones $g_{k+1},\ldots,g_n$ cuyas diferenciales son linealmente independientes. Si f tiene un extremo relativo condicionado a las ligaduras $g_{k+1}=\ldots=g_n=0$ en el punto p, existen constantes $\lambda_{k+1},\ldots,\lambda_n$ tales que $df_p=\lambda_{k+1}dg_{k+1\ p}+\ldots+\lambda_n dg_{n\ p}$.*

Demostración. Sea $v\in T_p(M)$. Será el vector tangente a una curva $\gamma(t)$ contenida en M en $p=\gamma(0)$. La función $f(\gamma(t))$ deberá tener un extremo local en $t=0$ y por tanto su derivada será cero. Tendremos entonces que $\frac{d}{dt}f(\gamma(t))(0)=df_p(v)=0$. Se cumplirá entonces que df_p es cero sobre los ceros comunes de $dg_{k+1\ p},\ldots,dg_{n\ p}$ y por tanto será combinación lineal de éstas.

□

El método de los multiplicadores de Lagrange se basa en el teorema anterior. Si se trata de buscar los puntos en que la función f tiene un extremo local sujeto a las ligaduras $g_{k+1}=\ldots=g_n=0$ será suficiente considerar los puntos que cumplen el sistema

$$\begin{gathered} g_{k+1}=\ldots=g_n=0 \\ \frac{\partial f}{\partial x_i}=\lambda_{k+1}\frac{\partial g_{k+1}}{\partial x_i}+\ldots+\lambda_n\frac{\partial g_n}{\partial x_i},\ \ i=1,\ldots,n. \end{gathered}$$

Se trata de un sistema de $2n-k$ ecuaciones y $2n-k$ incógnitas que son el punto $(x_1,\ldots,x_n)$ y los parámetros $\lambda_{k+1},\ldots,\lambda_n$. Estos parámetros reciben el nombre de multiplicadores de Lagrange. Las soluciones del sistema dan los puntos p tales que si Φ es una parametrización local de la superficie cumplen $d(f\circ\Phi)_p=0$. Son entonces condiciones necesarias para la existencia de extremo condicionado pero no son suficientes. La suficiencia deberá buscarse a través de otras consideraciones suplementarias.

Ejemplo 2.23. 1. *Calculemos el paralelepípedo de mayor volumen inscrito en el elipsoide de ecuaciones*

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1. \tag{2.14}$$

Se debe calcular el máximo absoluto de $|xyz|$ cuando x,y,z cumplen 2.14. Este conjunto es un compacto y la función es continua por lo que se alcanzará un máximo que obviamente será un valor estrictamente positivo. No importa considerar los puntos en que una de las coordenadas es 0 ya que corresponden a paralelepípedos de volumen 0. Podemos entonces reducirnos al cálculo del máximo de la función $V=8xyz$ condicionado a la ligadura $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ para $x,y,z>0$.

Podemos aplicar el método de los multiplicadores de Lagrange con lo que el máximo deberá cumplir el sistema

$$\begin{gathered} 8yz = \lambda\frac{2x}{a^2} \\ 8zx = \lambda\frac{2y}{b^2} \\ 8xy = \lambda\frac{2z}{c^2} \\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \end{gathered}$$

Como consecuencia de las tres primeras ecuaciones

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

Estas ecuaciones junto a la cuarta del sistema nos da $x = \frac{a}{\sqrt{3}},\ y = \frac{b}{\sqrt{3}},\ z = \frac{c}{\sqrt{3}}$. *Es el único extremo y por tanto se trata del máximo.*

2. *Se trata de calcular el mínimo de la distancia del punto* $(0,3)$ *de* R^2 *a la parábola de ecuación* $y = x^2$. *Debemos calcular el mínimo de la función* $f(x,y) = x^2 + (y-3)^2$ *sujeto a esta ligadura. El método de los multiplicadores de Lagrange nos lleva a resolver el sistema*

$$\begin{gathered} 2x - \lambda 2x = 0 \\ 2(y-3) + \lambda = 0 \\ y - x^2 = 0. \end{gathered}$$

Sus soluciones son $\left(\sqrt{\frac{5}{2}}, \frac{5}{2}\right), \left(-\sqrt{\frac{5}{2}}, \frac{5}{2}\right), (0,0)$. *Si calculamos el valor de* f *en estos puntos obtenemos* $f\left(\sqrt{\frac{5}{2}}, \frac{5}{2}\right) = \frac{11}{4},\ f\left(-\sqrt{\frac{5}{2}}, \frac{5}{2}\right) = \frac{11}{4},\ f(0,0) = 9$. *Se alcanzará el mínimo en los puntos* $\left(\sqrt{\frac{5}{2}}, \frac{5}{2}\right)$ *y* $\left(-\sqrt{\frac{5}{2}}, \frac{5}{2}\right)$. *Obsérvese que, en este caso, podemos explicitar muy fácilmente una parametrización de la parábola en la forma* $x = t,\ y = t^2$ *con lo que el problema se puede reducir al cálculo de los mínimos sin ninguna ligadura para la función de una variable real* $f(t,t^2) = t^2 + (t^2-3)^2$.

3. *Calculemos el máximo y el mínimo absoluto de la función* $f(x,y,z) = x^2+y^2+z^2+x+y+z$ *sobre el conjunto*

$$K = \left\{(x,y,z) \in R^3;\ x^2+y^2+z^2 \leq 4,\ z \leq 0\right\}.$$

Se trata de un compacto y f *es una función continua que alcanzará un máximo y un mínimo. Si estos se alcanzasen en el interior serían extremos relativos. En particular* df *debería ser cero en estos puntos. Deberán cumplir*

$$\begin{gathered} \frac{\partial f}{\partial x} = 2x+1 = 0 \\ \frac{\partial f}{\partial y} = 2y+1 = 0 \\ \frac{\partial f}{\partial z} = 2z+1 = 0. \end{gathered}$$

Su solución es $\left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)$. *Se trata de un punto de* K *y es un posible extremo.*

Pudiera ser que el extremo se alcanzase sobre la frontera de K. Dicha frontera está formada por los conjuntos

$$C_1 = \{(x,y,z) \in R^3;\ x^2+y^2+z^2 < 4,\ z=0\}$$
$$C_2 = \{(x,y,z) \in R^3;\ x^2+y^2+z^2 = 4,\ z<0\}$$
$$C_3 = \{(x,y,z) \in R^3;\ x^2+y^2+z^2 = 4,\ z=0\}.$$

Veamos los posibles extremos en cada uno de estos conjuntos. Sobre C_1 la función valdrá $g(x,y) = f(x,y,0) = x^2+y^2+x+y$. Debemos calcular los extremos en $x^2+y^2<4$. Deberemos resolver

$$\frac{\partial g}{\partial x} = 2x+1 = 0$$
$$\frac{\partial g}{\partial y} = 2y+1 = 0.$$

La solución es $\left(-\frac{1}{2}, -\frac{1}{2}\right)$. Tendremos que un posible extremo será $\left(-\frac{1}{2}, -\frac{1}{2}, 0\right)$.

Para estudiar los extremos sobre C_2 veamos los extremos condicionados de la función f con la ligadura $x^2+y^2+z^2 = 4$. Apliquemos el método de los multiplicadores de Lagrange. Debemos resolver

$$2x+1 = \lambda 2x$$
$$2y+1 = \lambda 2y$$
$$2z+1 = \lambda 2z$$
$$x^2+y^2+z^2 = 4.$$

Este sistema da como soluciones $\left(\frac{2}{\sqrt{3}}, \frac{2}{\sqrt{3}}, \frac{2}{\sqrt{3}}\right)$ y $\left(-\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}\right)$. Deberemos considerar únicamente el punto cuya tercera coordenada es negativa.

Por último, sobre C_3, aplicando el método de los multiplicadores de Lagrange con dos ligaduras tendremos que resolver el sistema

$$2x+1 = \lambda 2x$$
$$2y+1 = \lambda 2y$$
$$2z+1 = \lambda 2z + \mu$$
$$x^2+y^2+z^2 = 4$$
$$z = 0.$$

Sus soluciones son $\left(\frac{2}{\sqrt{2}}, \frac{2}{\sqrt{2}}, 0\right)$, $\left(-\frac{2}{\sqrt{2}}, -\frac{2}{\sqrt{2}}, 0\right)$.

Hemos encontrado como posibles extremos $(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})$, $(-\frac{1}{2}, -\frac{1}{2}, 0)$, $(-\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}})$, $(\frac{2}{\sqrt{2}}, \frac{2}{\sqrt{2}}, 0)$, $(-\frac{2}{\sqrt{2}}, -\frac{2}{\sqrt{2}}, 0)$. Calculando el valor de f sobre cada uno de estos puntos tendremos $f(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}) = -\frac{3}{4}$, $f(-\frac{1}{2}, -\frac{1}{2}, 0) = -\frac{1}{2}$, $f(-\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}}) = 4 - \frac{6}{\sqrt{3}}$, $f(\frac{2}{\sqrt{2}}, \frac{2}{\sqrt{2}}, 0) = 4 + \frac{4}{\sqrt{2}}$, $f(-\frac{2}{\sqrt{2}}, -\frac{2}{\sqrt{2}}, 0) = 4 - \frac{4}{\sqrt{2}}$. El máximo se tomará en $(\frac{2}{\sqrt{2}}, \frac{2}{\sqrt{2}}, 0)$ y el mínimo en $(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2})$.

2.3.6 Ejercicios

1. Consideremos la aplicación definida en R^2, a valores en R^2,

$$f(x,y) = (e^x \cos y + e^x \sin y, e^x \cos y - e^x \sin y).$$

Prueba que está en las condiciones del teorema de la función inversa pero que no es biyectiva. Explicita un abierto para el que la aplicación restringida sea inyectiva.

2. Halla los puntos en que la función $f(x,y) = (x^2+y^2, xy)$ tiene inversa local. Halla la matriz jacobiana de la función inversa donde exista.

3. Sea A un abierto de R^n y $f : A \to R^n$ de clase $\mathcal{C}^1$ tal que df_x sea inyectiva para todo $x \in A$. Prueba que la imagen de todo abierto contenido en A es un abierto.

4. Sea $f : R \to R$ definida por $f(x) = \frac{x}{2} + x^2 \sin \frac{1}{x}$ si $x \neq 0$ y $f(0) = 0$. Prueba que $f'(0) \neq 0$ y que f no es localmente invertible en 0. ¿Está esto en contradicción con el teorema de la función inversa?

5. Comprueba la dependencia o independencia funcional de las siguientes funciones:

   a) $f(x,y) = 4xy + 4x + 2$, $g(x,y) = x^2y^2 + 2x^2y + x^2 - 1$

   b) $f_1(x,y,z) = x + y + z$, $f_2(x,y,z) = 3x^2 + 8z^2 - 6xy - 8zy + 10xz$, $f_3(x,y,z) = 2x - y + 3z$.

6. Prueba que la ecuación $z^2 + xz + yx^2 + y^3 = 0$ define en un entorno de $(1,0,0)$, $z = g(x,y)$. Halla el polinomio de Taylor de grado 2, de la función g en el punto $(1,0)$.

7. Prueba que las ecuaciones $xz^6 + y^2u^3 - 1 = 0$, $xy^3 + u^2z^2 = 0$ definen x, y como función implícita de z, u en un entorno de $(x,y,z,u) = (0,1,0,1)$. Sean g_1 y g_2 estas funciones. Calcula las derivadas parciales de las g_i respecto z y u en un entorno de $(0,1)$.

8. Prueba que las ecuaciones $xy = 1,\ x^2 + y^2 + z + z^2 = 2$ definen en un entorno de $(1,1,0)$ una subvariedad diferenciable de dimensión 1. Halla la recta tangente a la variedad en este punto.

9. Encuentra los puntos en que la ecuación $x^3 + y^3 - 6xy = 0$ no define localmente una variedad diferenciable.

10. Prueba que el elipsoide $M = \left\{(x,y,z) \in R^3; \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\right\}$ es una subvariedad diferenciable de R^3 de dimensión 2. Encuentra parametrizaciones para entornos de diversos puntos de manera que la unión de estos entornos sea todo M.

11. Considera la esfera S en R^3 de centro el origen de coordenadas y radio 1. Sea la llamada proyección estereográfica de R^2 en S que se define de la siguiente forma. A cada (x,y) de R^2 se le hace corresponder la intersección de S con el intervalo de extremos $P = (x,y,0)$ y $N = (0,0,1)$. Encuentra las ecuaciones de la aplicación y prueba que se trata de una parametrización de $S - \{N\}$.

12. Prueba que la ecuación $y^3 + 3x^2y - x^3 + 2x + 3y = 0$ define una función implícita $y = g(x)$ definida para todo valor de x. Prueba que es de clase $\mathcal{C}^\infty$.

13. Prueba que la ecuación $\sin z + xyz - 2x - 2y - 2x^2y^2 + 6xy = 0$ define en un entorno de $(1,1,0)$, $z = g(x,y)$. ¿Tiene la función g en el punto $(1,1)$ un extremo relativo?
14. Prueba que si $f : R^2 \to R$ es una función de clase $\mathcal{C}^1$, no puede ser inyectiva. Indicación: Si, por ejemplo $\frac{\partial f}{\partial x}(x,y) \neq 0$ en un abierto U considera la aplicación de U en R^2 definida por $F(x,y) = (f(x,y), y)$ y aplica el teorema de la función inversa.
15. Determina la distancia mínima del origen a los puntos de la curva definida por $z = x^2 + y^2$, $x + y + z = 1$.
16. Halla el máximo y el mínimo de la función $f(x,y) = x^2 + y^2 - 2\sqrt{2}x$ sobre el disco cerrado de centro $\left(\sqrt{2}, \sqrt{2}\right)$ y radio 3.
17. Halla el plano que pasa por el punto $(2,1,1)$ y que determina con los ejes de coordenadas un tetraedro de volumen mínimo. ¿Existe uno que sea de volumen máximo?
18. Calcula el máximo de la función $(x_1x_2\cdots x_n)^2$ sujeto a la condición $\sum_{i=1}^n (x_i)^2 = 1$. Deduce que para $a_1, \ldots, a_n \geq 0$ se cumple $(a_1a_2\cdots a_n)^{\frac{1}{n}} \leq \frac{a_1+\ldots+a_n}{n}$.
19. Sea $a \in R^n$ y $p > 1$. Llamemos q al número tal que $\frac{1}{p} + \frac{1}{q} = 1$ y definamos

$$\|a\|_q = \sup_{\sum_{i=1}^n |x_i|^p = 1} a \cdot x.$$

Prueba que $\|a\|_q = \left(\sum_{i=1}^n |a_i|^q\right)^{\frac{1}{q}}$.

2.4 Nota histórica

Newton (1642-1727) y Leibnitz (1646-1716) son considerados como los principales creadores del Cálculo Infinitesimal. Sistematizaron y dotaron de algoritmos y notaciones generales los problemas de cálculo diferencial e integral que se habían tratado en muchos casos particulares. Aritmetizaron el Cálculo independizándolo de los argumentos geométricos utilizados hasta esa época. A finales del siglo XVII y principios del XVIII Jacobo y Nicolas Bernuilli (1654-1705 y 1687-1759) y B. Taylor (1685-1731), entre otros, hicieron aportaciones importantes.

Tanto Newton como los Bernoulli manejaron derivadas parciales en casos particulares. No obstante el desarrollo del cálculo diferencial de varias variables se da fundamentalmente en el siglo XVIII. Este siglo es el del gran desarrollo del Cálculo. Aparecen a partir de él nuevas ramas de las matemáticas como la teoría de las ecuaciones diferenciales, la geometría diferencial y las funciones de variable compleja. La principal fuente de problemas en este siglo son las ciencias de la naturaleza, aunque también se tratan problemas que provienen del propio Cálculo como la sumación de series o la evaluación de integrales. En esta época el tratamiento del Cálculo es formal. Lo importante eran los resultados obtenidos mas que el rigor en su obtención. Se poseía una intuición formidable pero las principales dificultades radicaban en las propias nociones básicas. El propio concepto de función no estaba claro. Durante mucho tiempo se consideran únicamente funciones que, salvo puntos aislados, admitían desarrollos en serie de potencias. No hay una distinción clara entre incremento de una función y diferencial, etc.

Euler (1707-1783) fue el matemático más destacado de su época. Trabajó en la teoría de ecuaciones diferenciales, geometría diferencial, teoría de números, series, cálculo de variaciones. Aplicó las matemáticas a la mecánica, acústica, teoría del calor, etc. Euler, junto con Clairaut y D'Alembert desarrollaron la teoría de las derivadas parciales. Sus trabajos se establecieron en conexión con la teoría de ecuaciones diferenciales. En unos trabajos relativos a hidrodinámica, en 1734, probó que $\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$. Entre 1748 y 1766 trató el cambio de variable, los determinantes funcionales y otras cuestiones relativas a las derivadas parciales. Otra gran figura de la época fue Lagrange (1736-1813). A él se le debe el teorema del valor medio y la expresión del término complementario de la fórmula de Taylor que lleva su nombre. La fórmula de Taylor había sido ya establecida por Taylor en 1712, desde luego sin tratar el problema de la convergencia de la serie. Este sería estudiado mucho más tarde por Cauchy. A Lagrange también se le debe la introducción de las coordenadas esféricas para la evaluación de integrales. Resolvió el paso a la forma diagonal de las cónicas en 1759 y lo aplicó al cálculo de los máximos de funciones. Entre otros matemáticos de la época cabe destacar a Clairaut (1713-1765), que fue uno de los iniciadores de la teoría de curvas a las que representaba como intersección de dos superficies, a D'Alembert (1717-1783) cuyos trabajos respecto a las derivadas parciales se dan en sus trabajos sobre dinámica entre 1744 y 1745, a Legendre, y a MacLauren.

Jacobi (1804-1851) introdujo en 1829 los "jacobianos". Expresó los cambios de variable de las integrales múltiples en términos de estos. En 1841 dio el teorema de dependencia funcional. En el mismo trabajo expresó el jacobiano de una función compuesta como el producto de los jacobianos. El hessiano fue introducido por Hesse (1811-1874) en 1844. El vector gradiente fue introducido por Hamilton (1805-1865) hacia mediados del siglo.

En el siglo XIX se rigorizan los conceptos básicos del Cálculo. Bolzano (1781-1848) clarificó la definición de función continua en términos de que "$f(x+w) - f(x)$ puede ser tan pequeño como se quiera tomando w suficientemente pequeño". En 1817 define la derivada como el límite del cociente incremental cuando el incremento tiende a cero. Cauchy (1789-1857) considera ya la diferencial de x, dx como un valor finito cualquiera y dy como $f'(x)\,dx$. Weierstrass (1815-1897) fue el que primero reconoció que para establecer con rigor las propiedades de las funciones continuas precisaba de la teoría de los números reales. Además de la introducción de la derivada como límite del cociente incremental la define también mediante la fórmula $f(x+h) = f(x) + hf'(x) + ho_{h\to 0}(1)$. En 1861 probó la igualdad de las derivadas cruzadas si las derivadas segundas eran continuas.

Obsérvese que todo este desarrollo se hizo sin una definición explícita de espacio vectorial. La introducción axiomática de espacio vectorial sobre los números reales no se da hasta 1888 por Peano (1858-1932).

Frechet (1878-1973) en 1914 define la diferencial de un funcional sobre un espacio de funciones en la forma que la hemos introducido.

2.5 Apéndice. Derivaciones y diferenciales. Una introducción algebraica.

Hemos definido la derivada direccional de una función f en la dirección dada por un vector u como la derivada de la función $g(t) = f(a + tu)$ en $t = 0$. En este capítulo hemos considerado que u era un vector unitario, pero puede de hecho prescindirse de esta condición y en el contexto actual será cómodo no imponerla. En esta definición es preciso que los puntos de la forma $a+tu$ pertenezcan al dominio de definición de la función por lo menos para t en un entorno de 0. Si se trata de introducir derivaciones para funciones que estén definidas únicamente sobre los puntos de una subvariedad no podemos utilizar esta definición. Los puntos de la forma $a+tu$ no estarán en general en la subvariedad y por tanto no tendría sentido considerar la función g.

Si queremos definir la diferencial de una función definida en los puntos de una subvariedad nos aparece también un problema. Hemos definido la diferencial de f en a como una aplicación lineal que aproxima la función $f(x) - f(a)$. Si la función está definida en un conjunto que no es un espacio vectorial, ¿qué sentido tiene su aproximación lineal? Se trata entonces de dar definiciones que cubran las situaciones anteriores y otras en otros contextos. Las daremos para funciones definidas en el espacio euclídeo para no complicarnos con tecnicismos que no son de interés en este momento pero dando una mayor relevancia a las funciones que están involucradas y no utilizando, en la medida de lo posible, la estructura de espacio vectorial del dominio de definición de las funciones. Consideraremos en este apéndice que todas las funciones que aparecen son de clase $\mathcal{C}^\infty$.

Definición 2.19. *Sea $a \in R^n$ y γ un arco de curva definido en un entorno de 0, a valores en R^n, tal que $\gamma(0) = a$. Sea f una función definida en R^n a valores en R. Se denomina derivada de f en a respecto a la curva γ a la derivada de la función $g(t) = f(\gamma(t))$ en $t = 0$.*

Obsérvese que en esta definición de derivación respecto de γ no intervienen mas que los valores de la función sobre la imagen de γ. El resultado que se obtiene no es otra cosa que la derivada en la dirección dada por el vector tangente a γ en a. En efecto

$$g'(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\,\gamma_i'(0) = D_{\gamma'(0)} f.$$

Dos curvas que tengan el mismo vector tangente en a dan entonces la misma derivación en a. Obsérvese también que una derivación direccional no es otra cosa que la derivada según la curva $\gamma(t) = a + tu$. Entonces, considerar derivaciones a lo largo de curvas o derivaciones direccionales es algo equivalente. La ventaja de considerar derivaciones respecto de curvas es que si la función estuviese definida únicamente sobre los puntos de una subvariedad seguiría teniendo sentido esta derivación para curvas cuyas imágenes estén contenidas en la subvariedad. Las propiedades básicas de la aplicación que asigna a cada función su derivada respecto a una curva se resumen en el siguiente teorema.

Teorema 2.28. *Sea $a \in R^n$ y γ un arco de curva definido en un entorno de 0, a valores en R^n, tal que $\gamma(0) = a$. La aplicación X de $\mathcal{C}^\infty(R^n)$ en R que asigna a cada función f la derivada respecto a γ tiene las siguientes propiedades:*

1. X es una aplicación lineal, es decir

$$X(f_1+f_2) = X(f_1) + X(f_2)$$
$$X(\lambda f) = \lambda X(f) \text{ para } \lambda \in R.$$

2. Si k es una función constante $X(k) = 0$.

3. Para cada $f, g \in \mathcal{C}^\infty(R^n)$, $X(fg) = f(a)X(g) + X(f)g(a)$. (Regla de derivación del producto).

Recíprocamente, toda aplicación de $\mathcal{C}^\infty(R^n)$ en R que tenga estas propiedades es la derivación respecto de una curva.

Demostración. El que la derivación respecto de una curva tenga estas propiedades es una comprobación directa. La propiedad 3 es consecuencia de la regla de derivación del producto para funciones de una variable.

Para ver el recíproco necesitamos el siguiente lema.

Lema 2.29. *Sea $f \in \mathcal{C}^\infty(R^n)$ y $a \in R^n$. Existen funciones $g_i \in \mathcal{C}^\infty(R^n)$ tales que*

$$f(x) - f(a) = \sum_{i=1}^{n} g_i(x)(x_i - a_i).$$

Se tiene, además, que $g_i(a) = \frac{\partial f}{\partial x_i}(a)$.

Demostración. Se tiene

$$f(x) - f(a) = \int_0^1 \tfrac{d}{dt} f(a + (x-a)t)\, dt =$$
$$\sum_{i=1}^n \left(\int_0^1 \tfrac{\partial f}{\partial x_i}(a + (x-a)t)\, dt \right)(x_i - a_i).$$

Las funciones $g_i(x) = \int_0^1 \frac{\partial f}{\partial x_i}(a + (x-a)t)\, dt$ son de clase $\mathcal{C}^\infty$. Para comprobarlo es suficiente utilizar los teoremas de derivación bajo el signo integral elementales.

La relación $g_i(a) = \frac{\partial f}{\partial x_i}(a)$ es consecuencia de la expresión demostrada y de la definición de derivada parcial.

$\square$

Podemos pasar a concluir la demostración del teorema.

Apliquemos el lema a la función f y a cada una de las funciones g_i que aparecen en la descomposición. Tendremos que para ciertas $g_{ij} \in \mathcal{C}^\infty(R^n)$

$$f(x) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \sum_{i,j=1,\ldots,n} g_{ij}(x)(x_i - a_i)(x_j - a_j).$$

Aplicando ahora las propiedades 1,2,3 tendremos

$$X(f) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) X(x_i).$$

Si llamamos $c_i = X(x_i)$ se obtiene que $X(f)$ es la derivada direccional dada por el vector $(c_1, \ldots, c_n)$.

$\square$

El teorema permite tomar como derivación una aplicación de $\mathcal{C}^{\infty}(R^n)$ en R que cumpla las tres propiedades del teorema. Obsérvese que es una definición algebraica en que no aparece la noción de límite. Por otro lado el conjunto de estas aplicaciones forma un espacio vectorial con la definición natural de suma y producto por escalares. En la identificación que hemos hecho de estas derivaciones con las derivadas direccionales se ha identificado el elemento $u \in R^n$ con la derivación
$$X(f) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, u_i.$$
En particular el conjunto de las derivaciones en a es un espacio vectorial de dimensión n. Una base de este espacio está formado por las derivadas parciales en el punto a. Llamemos a este espacio $T_a(R^n)$.

Podemos definir el espacio de las diferenciales en el punto a como el espacio dual de $T_a(R^n)$. La diferencial de una función f en a será el elemento de $(T_a(R^n))'$ que asigna a la derivación X el número real $X(f)$. La escribiremos $df_a(X) = X(f)$. En particular si como función tomamos x_i tendremos $dx_{i\,a}(X) = X(x_i)$. Esta relación nos dice que $dx_{1\,a}, \ldots, dx_{n\,a}$ forman la base dual de $\frac{\partial}{\partial x_i}{}_a, \ldots, \frac{\partial}{\partial x_n}{}_a$. De esta forma podemos dar un sentido a las diferenciales de una función sin utilizar la estructura de espacio vectorial del dominio de definición de la función. No obstante, con este punto de vista, se ha perdido la intuición de pensar la diferencial como una aproximación lineal de la función $f(x) - f(a)$. Una introducción de las diferenciales que preserva la idea de que la diferencial de una función consiste en "eliminar en el infinitésimo $f(x) - f(a)$ los infinitésimos de orden superior al primero" es la siguiente.

Llamemos $m_a = \{f \in \mathcal{C}^{\infty}(R^n)\,;\, f(a) = 0\}$. Por ejemplo, si $f \in \mathcal{C}^{\infty}(R^n)$, entonces $f - f(a) \in m_a$. Denotaremos m_a^2 al conjunto de las expresiones $\sum g_i h_i$ para $g_i, f_i \in m_a$.

Teorema 2.30. *El espacio vectorial cociente* $\frac{m_a}{m_a^2}$ *es un espacio vectorial de dimensión* n.

Demostración. Los elementos de m_a son de la forma $f - f(a)$ para $f \in \mathcal{C}^{\infty}(R^n)$. Apliquemos la descomposición del lema 2.29 como en el teorema anterior.
$$f(x) - f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \sum_{i,j=1,\ldots,n} g_{ij}(x)(x_i - a_i)(x_j - a_j).$$
Tendremos que $f(x) - f(a)$ darán la misma clase que $\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i)$ en $\frac{m_a}{m_a^2}$. Por otra parte, las funciones $x_i - a_i$, $i = 1, \ldots, n$ dan clases independientes pues si
$$\sum_{i=1}^{n} \lambda_i (x_i - a_i) = \sum_{j=1}^{n} g_j h_j \quad \text{para } g_j, f_j \in m_a$$
derivando respecto x_i y dando a x el valor a obtenemos que $\lambda_i = 0$ para cada i.

□

Hemos visto además en la demostración que la clase definida por $f - f(a)$ coincide con la de $\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i)$. Si escribimos como habitualmente $h = x - a$, el incremento de la variable, tenemos que la aplicación $\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, h_i$ es lineal en h. Se trata de la única aplicación lineal de entre los representantes de la clase definida por $f - f(a)$.

Al espacio vectorial $\frac{m_a}{m_a^2}$ se le llama espacio de las diferenciales en el punto a. De una forma intuitiva, a partir de los infinitésimos en a se obtienen las diferenciales tomándolos como equivalentes si su diferencia es un infinitésimo de segundo orden. La diferencial de una función f es la clase definida por $f - f\left(a\right)$. Se escribirá df_a. Una base de las diferenciales estará formada por $dx_{1\ a}, \ldots, dx_{n\ a}$. Las derivaciones se podrán definir o bien como hemos hecho antes como las aplicaciones de $\mathcal{C}^\infty\left(R^n\right)$ en R, lineales, que cumplen la regla del producto y que se anulan sobre las constantes, o bien como el espacio dual del espacio de las diferenciales $\frac{m_a}{m_a^2}$.