Cálculo Diferencial

Daniel Azagra Rueda

La Resistencia del

DEPARTAMENTO DE ANÁLISIS MATEMÁTICO

FACULTAD DE CIENCIAS MATEMÁTICAS

UNIVERSIDAD COMPLUTENSE DE MADRID

Revisado: Agosto de 2019

ISBN-13: 978-84-611-6378-6

Índice general

Parte I

Conceptos métricos y topológicos en $\mathbb{R}^n$

Capítulo 1

Métricas, normas y productos escalares

El alumno de esta asignatura ya ha realizado un curso anual de álgebra lineal y está familiarizado con los espacios vectoriales de dimensión finita, aplicaciones lineales, bases, matrices y determinantes. Ha adquirido por tanto una comprensión suficiente de la estructura lineal y afín de los espacios $\mathbb{R}^n$, aunque quizás no esté tan familiarizado con la estructura métrica de los mismos. En este primer capítulo resumimos las propiedades más importantes de los espacios $\mathbb{R}^n$ que tienen que ver con ángulos, ortogonalidad de vectores, y distancias entre puntos (u otros subconjuntos) del espacio.

Comenzamos recordando la definición de producto escalar, ortogonalidad de vectores y subespacios, y demostrando aquellas de sus propiedades que serán más útiles de cara al estudio de esta asignatura, como la desigualdad de Cauchy-Schwarz. A partir del producto escalar introducimos la norma y la distancia euclideas. Esto da pie a abstraer algunas de las propiedades esenciales de esta métrica euclidea y llegar así al concepto de norma en un espacio vectorial. Continuamos con un breve estudio de las propiedades generales de las normas y algunos ejemplos.

Asociada a cada norma tenemos una distancia en el espacio $\mathbb{R}^n$ que lo dota de una estructura de espacio métrico, pero esta distancia no es una métrica arbitraria en $\mathbb{R}^n$, sino que tiene la propiedad de ser invariante por traslaciones: ésta es una de las ventajas principales de usar normas en lugar de distancias. Concluimos estudiando brevemente varios ejemplos concretos de espacios métricos, así como algunos conceptos generales que tienen que ver con distancias.

Definición 1.1. En el espacio vectorial $\mathbb{R}^n$ se define el producto escalar euclideo de dos vectores $x = (x_1, ..., x_n)$ e $y = (y_1, ..., y_n) \in \mathbb{R}^n$ como

$$\langle x, y\rangle = \sum_{j=1}^{n} x_j y_j.$$

Geométricamente, este número corresponde al coseno del ángulo α que forman los vectores x y y multiplicado por las longitudes de dichos vectores. En particular, los vectores x e y son perpendiculares si y sólo si $\langle x, y\rangle = 0$.

Las propiedades más importantes del producto escalar quedan resumidas en la proposición siguiente, cuya demostración es una mera comprobación que queda a cargo del lector.

Proposición 1.1. *Para cada $x, y, z \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$, tenemos que*

1. $\langle x, x\rangle \geq 0$, *y además* $\langle x, x\rangle = 0$ *si y sólo si* $x = 0$*;*
2. $\langle x, y\rangle = \langle y, x\rangle$*;*

3. $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$;

4. $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$.

Esto, en el lenguaje del álgebra lineal, simplemente significa que $\langle\cdot,\cdot\rangle$ es una forma bilineal simétrica definida positiva en $\mathbb{R}^n$. Por supuesto ésta no es la única forma bilineal con tales propiedades en $\mathbb{R}^n$. Por ejemplo, para cualesquiera números positivos $\lambda_1, ..., \lambda_n$, la aplicación $(x, y) \mapsto \sum_{j=1}^n \lambda_j x_j y_j$ tiene las mismas propiedades. De hecho, como el alumno de esta asignatura sabe diagonalizar matrices y encontrar bases ortonormales de formas cuadráticas en $\mathbb{R}^n$, no debe tener dificultades en admitir que *esencialmente*, todas las formas bilineales definidas positivas son de esta forma, en el sentido siguiente: B es una forma bilineal simétrica definida positiva en $\mathbb{R}^n$ si y sólo si existen números positivos $\lambda_1, ..., \lambda_n$ y un isomorfismo lineal $A : \mathbb{R}^n \to \mathbb{R}^n$ con $\det A = 1$ tal que $B(x, y) = J(Ax, Ay)$, donde $J(x, y) = \sum_{j=1}^n \lambda_j x_j y_j$.

Definición 1.2. Se dice que dos vectores $x, y \in \mathbb{R}^n$ son ortogonales (o perpendiculares) si $\langle x, y\rangle = 0$, y en este caso escribimos $x \perp y$. Si A es un subconjunto de $\mathbb{R}^n$, se define su complemento ortogonal como

$$A^\perp = \{x \in \mathbb{R}^n : \forall a \in A \;\; x \perp a\}.$$

Algunas propiedades de esta operación conjuntista son:

1. Si $A \subseteq B$ entonces $B^\perp \subseteq A^\perp$.

2. $A^\perp$ es un subespacio vectorial de $\mathbb{R}^n$.

3. $[A]^\perp = A^\perp$, donde $[A]$ denota la envoltura lineal de A.

4. Para todo subespacio W de $\mathbb{R}^n$ se tiene que $\dim W + \dim W^\perp = n$, y $\mathbb{R}^n = W \oplus W^\perp$.

5. $[A] = (A^\perp)^\perp$.

Definición 1.3. En $\mathbb{R}^n$ se define la norma euclidea, asociada al producto escalar $\langle\cdot,\cdot\rangle$, como

$$\|x\| = \sqrt{\langle x, x\rangle} = \Big(\sum_{j=1}^n x_j^2\Big)^{1/2}$$

para cada $x \in \mathbb{R}^n$.

Proposición 1.2. *La norma euclidea tiene las siguientes propiedades, para todos los* $x, y, z \in \mathbb{R}^n$*,* $\lambda \in \mathbb{R}$*:*

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

2. $\|\lambda x\| = |\lambda|\|x\|$;

3. (desigualdad triangular) $\|x + y\| \leq \|x\| + \|y\|$;

4. (desigualdad de Cauchy-Schwarz) $|\langle x, y\rangle| \leq \|x\|\|y\|$.

Demostración. La demostración de las propiedades 1 y 2 es inmediata. La propiedad triangular no es obvia pero se deduce fácilmente de la desigualdad de Cauchy-Schwarz, que a su vez se prueba como sigue. Por la bilinealidad del producto escalar, y por la propiedad 1, para todo $\alpha \in \mathbb{R}$ tenemos que

$$0 \leq \langle x + \alpha y, x + \alpha y \rangle = \|x\|^2 + 2\alpha \langle x, y \rangle + \alpha^2 \|y\|^2;$$

entonces, suponiento $y \neq 0$ (la desigualdad es obvia en este caso) y tomando

$$\alpha = -\frac{\langle x, y \rangle}{\|y\|^2}$$

(que no es otra cosa que el valor de α que minimiza el segundo miembro de la desigualdad de arriba, visto como una función de α), obtenemos que

$$\|x\|^2 - \frac{\langle x, y \rangle^2}{\|y\|^2} \geq 0,$$

lo que nos da la desigualdad deseada. La desigualdad triangular se deduce ahora directamente de:

$$\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = \big(\|x\| + \|y\|\big)^2.$$

□

La norma euclidea es el instrumento que usaremos casi siempre a lo largo del curso para medir longitudes de vectores (o distancias entre puntos).

Definición 1.4. Dado un vector x de $\mathbb{R}^n$, definimos su longitud como el número $\|x\|$. Si a, b son puntos de $\mathbb{R}^n$, definimos la distancia entre a y b como la longitud del segmento que los une, esto es, $\|b - a\|$.

Las propiedades 1, 2 y 3 de la Proposición anterior son lo suficientemente útiles e importantes como para abstraer la clase de funciones que las satisfacen y darles un nombre.

Definición 1.5. Se dice que una aplicación $\|\cdot\| : \mathbb{R}^n \to [0, \infty)$ es una *norma* en el espacio vectorial $\mathbb{R}^n$ si satisface las siguientes propiedades:

1. $\|x\| > 0$ si $x \neq 0$
2. $\|x + y\| \leq \|x\| + \|y\|$
3. $\|\lambda x\| = |\lambda| \|x\|$

para todo $x, y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.

Como ejemplos típicos de normas en $\mathbb{R}^n$ pueden citarse la norma euclidea definida más arriba, la norma del máximo,

$$\|x\|_\infty = \max\{|x_i| : i = 1, ..., n\},$$

y más en general las normas p, definidas por

$$\|x\|_p = \big(\sum_{j=1}^{n} |x_j|^p\big)^{1/p}$$

para $1 \leq p < \infty$.[1]

En la teoría de este curso usaremos solamente las normas $\|\cdot\|_2$, $\|\cdot\|_1$ y $\|\cdot\|_\infty$ y, sobre todo, la euclidea, que es la que nos da la distancia natural en $\mathbb{R}^n$.

El alumno podría preguntarse qué interés tiene el considerar normas y distancias que no son naturales en un curso elemental, como no sea el de abultar la teoría o poner a prueba su capacidad de entendimiento. Una de las razones más importantes es que, incluso a este nivel, constituyen un valioso instrumento técnico: en diversas demostraciones se realizarán acotaciones que pueden resultar evidentes para las normas $\|\cdot\|_1$, o $\|\cdot\|_\infty$, o incluso para alguna diferente de estas, pero no tanto para la norma euclidea, y después, usando el hecho de que todas las normas en $\mathbb{R}^n$ son equivalentes (en el sentido precisado más abajo, ver la Definición 1.8 y el Teorema 1.3) podrán obtenerse las desigualdades buscadas para la norma euclidea. Además, la norma $\|\cdot\|_\infty$, pese a no ser la que da la distancia natural en $\mathbb{R}^k$ (o más bien precisamente por ello) es la más adecuada para el tratamiento del producto cartesiano de conjuntos de $\mathbb{R}^n$ y $\mathbb{R}^n$ identificado como un subconjunto de $\mathbb{R}^{n+m}$ (ver el problema 1.10).

Definición 1.6. Asociada a cada norma $\|\cdot\|$ en $\mathbb{R}^n$ podemos considerar, para cada $x \in \mathbb{R}^n$ y cada $r > 0$, la bola abierta de centro x y radio r, definida como

$$B_{\|\cdot\|}(x,r) = B(x,r) := \{y \in \mathbb{R}^n : \|y-x\| < r\},$$

y la bola cerrada de centro x y radio r,

$$\overline{B}_{\|\cdot\|}(x,r) = \overline{B}(x,r) := \{y \in \mathbb{R}^n : \|y-x\| \leq r\}.$$

Esta terminología adquirirá mayor transparencia en el Capítulo 3, donde en particular se verá que las bolas abiertas son subconjuntos abiertos de $\mathbb{R}^n$, y las bolas cerradas son cerrados. También definimos la esfera de centro x y radio r como la diferencia entre la bola cerrada y la abierta con el mismo centro y radio, es decir,

$$S(x,r) := \{y \in \mathbb{R}^n : \|y-x\| = r\}.$$

Definición 1.7. Se dirá que un conjunto $A \subset \mathbb{R}^n$ está acotado cuando existe una bola (con radio suficientemente grande) que lo contiene. Equivalentemente, A está acotado en $\mathbb{R}^n$ si el subconjunto $\{\|a\| : a \in A\}$ de $\mathbb{R}$ está acotado en $\mathbb{R}$.

Ejemplo 1.1. Hacer un dibujo de $\overline{B}_{\|\cdot\|}(0,1)$, donde $\|\cdot\|$ es cada una de las normas $\|\cdot\|_1$, $\|\cdot\|_2$, y $\|\cdot\|_\infty$ del plano $\mathbb{R}^2$, comparando unas con otras.

Conviene observar que la definición de norma 1.5 tiene sentido en cualquier espacio vectorial X (no necesariamente isomorfo a $\mathbb{R}^n$). Por ejemplo, en $X = C[0,1]$, el espacio vectorial de todas las funciones continuas $f : [0,1] \to \mathbb{R}$, puede definirse la norma

$$\|f\|_\infty = \sup\{|f(t)| : t \in [0,1]\} = \text{máx}\{|f(t)| : t \in [0,1]\}$$

(compruébese que esta fórmula define efectivamente una norma en X). En este caso X no es isomorfo a ningún espacio $\mathbb{R}^n$ porque X es un espacio vectorial de dimensión infinita (no admite ninguna base finita).

[1]En el siguiente capítulo recordaremos algunas desigualdades importantes que permiten probar fácilmente que estas funciones son, en efecto, normas.

Definición 1.8. Se dice que dos normas $\|\cdot\|$ y ρ de un espacio vectorial X son equivalentes si existen $M, m > 0$ tales que
$$m\|x\| \leq \rho(x) \leq M\|x\|$$
para todo $x \in X$. Geométricamente esto significa que
$$B_\rho(x, mr) \subseteq B_{\|\cdot\|}(x, r) \subseteq B_\rho(x, Mr)$$
para cada $x \in X$, $r > 0$.

Ejemplo 1.2. Demostrar que las normas $\|\cdot\|_1$, $\|\cdot\|_2$ y $\|\cdot\|_\infty$ de $\mathbb{R}^2$ son equivalentes. Generalizar el resultado a $\mathbb{R}^n$, observando que los números M y m que dan la equivalencia de las normas dependen de la dimensión n.

El siguiente resultado es muy importante ya que nos permite realizar acotaciones con respecto a una norma cualquiera en $\mathbb{R}^n$ y, automáticamente, obtener otra acotación del mismo tipo para la norma euclidea o cualquier otra que se desee.

Teorema 1.3. *Todas las normas de $\mathbb{R}^n$ son equivalentes.*

En la demostración de este resultado se usan argumentos de continuidad, por lo que la aplazamos hasta el capítulo 5. Por supuesto, habrá que tener cuidado de no usar este teorema en la prueba de los resultados en los que a su vez descansa dicha demostración (esencialmente, que una función continua en un conjunto compacto alzanza un máximo absoluto), pero como veremos no hay ningún problema. De hecho puede suponerse que todo el desarrollo teórico de la asignatura hasta el capítulo 5 se refiere exclusivamente a una o varias normas equivalentes fijas (por ejemplo la euclidea, o bien $\|\cdot\|_1$ o $\|\cdot\|_\infty$, que ya sabemos que son equivalentes por el Ejemplo 1.2), y llegado a ese punto y demostrado este teorema, extender todos los resultados al caso en que la norma considerada sea cualquier otra.

Se deduce inmediatamente, por ejemplo, que el hecho de que un subconjunto A de $\mathbb{R}^n$ esté acotado no depende de la norma en cuestión.

Como curiosidad mencionaremos que este resultado no es cierto en espacios vectoriales normados de dimensión infinita: en todo tal espacio (por ejemplo en $C[0, 1]$, ver el ejercicio 1.12) existen siempre normas que no son equivalentes, y por tanto también conjuntos que son acotados para unas normas pero no para otras. Por tanto, un espacio de vectorial X es de dimensión finita si y sólo si todas las normas en X son equivalentes.

Ya hemos dejado ver que toda norma $\|\cdot\|$ lleva aparejada una noción de distancia entre los puntos de un espacio vectorial X, sin más que definir la distancia entre dos puntos $x, y \in X$ como $d(x, y) = \|x - y\|$. Las siguientes propiedades se demuestran inmediatamente a partir de la definición y de las propiedades de las normas. Para todos $x, y, z, a \in X$ se tiene que:

1. $d(x, y) > 0$ si $x \neq y$, y $d(x, x) = 0$;

2. $d(x, y) = d(y, x)$;

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

Estas propiedades son importantes y dan lugar, a su vez, a una nueva noción de proximidad entre puntos que tiene sentido, incluso, en conjuntos que no tienen presupuesta ninguna estructura lineal.

Definición 1.9. Se dice que (X, d) es un espacio métrico si X es un conjunto no vacío con una una función, llamada distancia (o métrica, $d : X \times X \to [0, \infty)$ que tiene las propiedades 1, 2 y 3 anteriores.

Por ejemplo, en el conjunto $\mathbb{Z}$ puede definirse la distancia $d(n, m) = |n - m|$ de modo que $(\mathbb{Z}, d)$ es un espacio métrico.

Definición 1.10. Si (X, d) es un espacio métrico, definimos la bola abierta de centro $x \in X$ y radio $r > 0$ como $B(x, r) = \{y \in X : d(y, x) < r\}$, y la bola cerrada por $\overline{B}(x, r) = \{y \in X : d(y, x) \leq r\}$. Si A es un subconjunto no vacío de X, se define la distancia de un punto $x \in X$ al conjunto A por

$$d(x, A) = \text{ínf}\{d(x, a) : a \in A\}.$$

Para A, B subconjuntos no vacíos de X, la distancia entre A y B se define por

$$d(A, B) = \text{ínf}\{d(x, y) : x \in A, y \in B\}.$$

Llamamos diámetro de A al supremo de las distancias entre dos puntos cualesquiera de A, es decir,

$$\text{diam}(A) := \sup\{d(x, y) : x, y \in A\}$$

Se dice que un subconjunto A de X está acotado si existe una bola que lo contiene, lo que equivale a que $\text{diam}(A) < \infty$.

Como en el caso de normas, podemos definir *distancias equivalentes* en un espacio métrico. Si X es un conjunto en el que tenemos definidas dos distancias d_1 y d_2, diremos que d_1 y d_2 son métricas equivalentes si para cada $x \in X$ y $r > 0$, existen $r_1, r_2 > 0$ tales que $B_{d_1}(x, r_1) \subseteq B_{d_2}(x, r)$, y también $B_{d_2}(x, r_2) \subseteq B_{d_1}(x, r)$. Es importante observar que los números r_1, r_2, incluso para un r fijo, en general dependen del punto x. Si no existe tal dependencia del punto x (es decir, si dado $r > 0$ existen r_1, r_2 tales que los contenidos anteriores se cumplen para todo $x \in X$), diremos que d_1 y d_2 son *uniformemente equivalentes*.[2] Esto es lo mismo que decir que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que:

$$d_1(x, y) < \delta \implies d_2(x, y) < \varepsilon \text{ para todos } x, y \in X,$$

y que

$$d_2(x, y) < \delta \implies d_1(x, y) < \varepsilon \text{ para todos } x, y \in X.$$

Por otra parte diremos que d_1 y d_2 son *fuertemente equivalentes* si existen constantes $\alpha, \beta > 0$ tales que, para todos $x, y \in X$, se tiene

$$\alpha\, d_1(x, y) \leq d_2(x, y) \leq \beta\, d_1(x, y).$$

Equivalentemente, para todo $r > 0$ y todo $x \in X$, se tiene $B_{d_1}(x, r/\beta) \subseteq B_{d_2}(x, r)$, y $B_{d_2}(x, \alpha r) \subseteq B_{d_1}(x, r)$. En particular dos métricas fuertemente equivalentes son uniformemente equivalentes, y por tanto también equivalentes.

Ejemplo 1.3. En el conjunto $X = \mathbb{R}$ puede definirse la función

$$d(x, y) = |\text{arctg}(x) - \text{arctg}(y)|$$

de modo que $(\mathbb{R}, d)$ es un espacio métrico en el que todo subconjunto (¡incluido el propio $\mathbb{R}$!) está acotado.

[2]Cuando lleguemos al capítulo 5, resultará evidente que esta condición también equivale a decir que la aplicación $\mathbb{I} : (X, d_1) \to (X, d_2)$ definida por $\mathbb{I}(x) = x$ es uniformemente continua y tiene inversa uniformemente continua.

Si (X, d_X) e (Y, d_Y) son espacios métricos, puede dotarse al producto cartesiano $X \times Y$ de una estructura de espacio métrico de varias maneras. Por ejemplo, puede definirse

$$d_\infty\big((x,y),(z,w)\big) = \text{máx}\{d_X(x,z)d_Y(y,w)\}.$$

Es inmediato comprobar que d_∞ es una distancia en $X \times Y$. También podríamos haber definido otras distancias más o menos naturales en $X \times Y$. Por ejemplo

$$d_1\big((x,y),(z,w)\big) = d_X(x,z) + d_Y(y,w),$$

o incluso

$$d_2\big((x,y),(z,w)\big) = \big(d_X(x,z)^2 + d_Y(y,w)^2\big)^{1/2}.$$

Aunque en este contexto la métrica d_∞ es la más usual y la que tomaremos como definición de métrica producto, salvo mención explícita en sentido contrario, puede comprobarse que las tres distancias son uniformemente equivalentes (ver el Problema 1.22).

Observación 1.4. Es natural preguntarse si es cierto un teorema análogo al 1.3 para el caso de espacios métricos. La respuesta es que no: existen distancias en $\mathbb{R}$ que no son equivalentes a la usual (inducida por el valor absoluto). Una manera de verlo es considerar en $\mathbb{R}$ la *distancia discreta* D definida por $D(x,y) = 1$ si $x \neq y$, y $D(x,y) = 0$ si $x = y$. Es fácil comprobar que D es una distancia que no es equivalente a la usual en $\mathbb{R}$ (la bola de centro 0 y radio $1/2$ para la distancia D se reduce al punto 0, y por tanto no contiene ninguna bola para la distancia usual).

1.1. Problemas

Problema 1.1. Si $\|\cdot\|$ es una norma en un espacio vectorial X, demostrar que

$$\big|\|x\| - \|y\|\big| \leq \|x - y\|$$

para todo $x, y \in X$. ¿Cuál es el análogo de esta propiedad para distancias?

Problema 1.2. Comprobar que las siguientes expresiones definen normas en $\mathbb{R}^2$:

1. $\|(x,y)\|_a = \frac{1}{3}(|x| + |y|) + \frac{2}{3}(x^2 + y^2)^{1/2}$;
2. $\|(x,y)\|_b = (x^2 + xy + y^2)^{1/2}$

Indicación: para el segundo caso, puede ser aconsejable completar cuadrados dentro de la raiz y observar que $\|(x,y)\|_b = \|T(x,y)\|_2$, donde T es un automorfismo lineal de $\mathbb{R}^2$.

Problema 1.3. Si T es un automorfismo lineal de $\mathbb{R}^n$ y $\|\cdot\|$ es una norma cualquiera de $\mathbb{R}^n$, demostrar que la expresión $\rho(x) = \|T(x)\|$ define una norma en $\mathbb{R}^n$.

Problema 1.4. Si $\|\cdot\|$ es la norma euclidea de $\mathbb{R}^n$, demostrar que:

1. (Ley del paralelogramo) $2\|x\|^2 + 2\|y\|^2 = \|x+y\|^2 + \|x-y\|^2$;
2. $\|x+y\|\|x-y\| \leq \|x\|^2 + \|y\|^2$;
3. (Identidad de polarización) $4\langle x, y\rangle = \|x+y\|^2 - \|x-y\|^2$.

Problema 1.5. Si $\|\cdot\|$ es una norma en $\mathbb{R}^n$ que cumple la ley del paralelogramo, probar que proviene de un producto escalar. *Indicación:* definir el candidato a producto escalar a partir de la identidad de polarización.

Problema 1.6. En la desigualdad de Cauchy-Schwarz, demostrar que la igualdad se tiene si y sólo si los vectores x e y son proporcionales.

Problema 1.7. Si $A = \{(x, y, z, w) \in \mathbb{R}^4 : x + y - z = 0, y + z - w = 0\}$, hallar $A^\perp$.

Problema 1.8. Probar que si $v_1, ..., v_n$ son vectores no nulos y ortogonales dos a dos (es decir, $v_i \perp v_j$ si $i \neq j$), entonces los vectores $v_1, ..., v_n$ son linealmente independientes.

Problema 1.9. En el espacio vectorial $C[0, 1]$ de todas las funciones continuas en el intervalo $[0, 1]$, definamos

$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt.$$

Comprobar que esta fórmula define un producto escalar en $C[0, 1]$. Exhibir dos funciones no nulas y ortogonales entre sí. Después construir una sucesión de funciones no nulas (f_n) en $C[0, 1]$ tales que $f_n \perp f_m$ si $n \neq m$. Concluir que el espacio vectorial $C[0, 1]$ tiene dimensión infinita.

Problema 1.10. Identificando $\mathbb{R}^n \times \mathbb{R}^m$ con $\mathbb{R}^{n+m}$ y usando la norma $\|\cdot\|_\infty$ en estos espacios, pruébese que para cada $(x, y) \in \mathbb{R}^{n+m}$ y cada $r > 0$ se tiene que

$$B_\infty(x, r) \times B_\infty(y, r) = B_\infty\big((x, y), r\big).$$

Problema 1.11. Dibujar una función continua f cualquiera en el intervalo $[0, 1]$. Después, dibujar la bola cerrada de centro f y radio $r > 0$ en el espacio normado $(C[0, 1], \|\cdot\|_\infty)$.

Problema 1.12. Sea $\|\cdot\|_2$ la norma definida a través del producto escalar del problema 1.9 en $C[0, 1]$, esto es,

$$\|f\|_2 = \Big(\int_0^1 f(t)^2 dt\Big)^{1/2},$$

y consideremos también la norma ya definida anteriormente en este espacio,

$$\|f\|_\infty = \sup\{|f(t)| : t \in [0, 1]\}.$$

Probar que las normas $\|\cdot\|_2$ y $\|\cdot\|_\infty$ *no* son equivalentes en $C[0, 1]$ (y de paso nótese que esto proporciona otra demostración de que $C[0, 1]$ tiene dimensión infinita, vía el Teorema 1.3). *Indicación:* construir una sucesión de funciones $(f_n) \subset C[0, 1]$ que esté acotada para la norma $\|\cdot\|_2$ pero no para $\|\cdot\|_\infty$.

Problema 1.13. Sea E un espacio vectorial cualquiera de dimensión finita. Construir una norma en E.

Problema 1.14. Construir una métrica en el conjunto $X = \mathbb{N} \times \mathbb{N}$ de modo que todos los subconjuntos de X estén acotados.

Problema 1.15. Estudiar si los siguientes subconjuntos de $\mathbb{R}^n$ están o no acotados:

$A = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1, z = x^2 - y^2\}$;

$B = \{(x, y) \in \mathbb{R}^2 : y = \sin \frac{1}{x}, x \neq 0\}$;

$$C = \{(x, y, z) \in \mathbb{R}^3 : x^2 - y^2 = 1, z + x + y = 0\}.$$

Problema 1.16. Estudiar cuáles de las siguientes expresiones definen métricas en X:

1. $d(x, y) = |x^2 - y^2|$, $X = \mathbb{R}$;
2. $d(n, m) = \frac{1}{n} + \frac{1}{m}$ si $n \neq m$, y $d(n, m) = 0$ si $n = m$, $X = \mathbb{N}$;
3. $X = \mathbb{R}^{\mathbb{N}}$, el conjunto de sucesiones de números reales,

$$d(x, y) = \sum_{n=1}^{\infty} \frac{|x_n - y_n|}{2^n\big(1 + |x_n - y_n|\big)}.$$

Problema 1.17. Demostrar que la función distancia de un espacio métrico X satisface que

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

para todo $x, y, z \in X$

Problema 1.18. Más en general, si A es un subconjunto no vacío de un espacio métrico X, demostrar que

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

para todo $x, y \in X$. Se dice entonces que la función $x \in X \mapsto d(x, A)$ es 1-Lipschitz.

Problema 1.19. Si $\|\cdot\|_1$ y $\|\cdot\|_2$ son dos normas en un espacio vectorial X, denotemos por d_1 y d_2 las distancias que dichas normas inducen. Probar que las normas $\|\cdot\|_1$ y $\|\cdot\|_2$ son equivalentes si y sólo si las distancias d_1 y d_2 son equivalentes.

Problema 1.20. Probar que la unión finita de subconjuntos acotados de X es acotada en X. En particular todo conjunto finito es acotado.

Problema 1.21. Se dice que dos espacios métricos (X_1, d_1) y (X_2, d_2) son *isométricos* si existe una biyección $f : X_1 \to X_2$ tal que $d(f(x), f(y)) = d(x, y)$ para todo $x, y \in X_1$. Si f no es biyectiva se habla de una inyección isométrica de X_1 en X_2. Dos espacios isométricos son indistinguibles desde el punto de vista métrico, y por tanto pueden identificarse cuando ello resulta conveniente, en problemas de naturaleza métrica, claro está.

Probar que $(\mathbb{C}, |\cdot|)$ y $(\mathbb{R}^2, \|\cdot\|_2)$ son isométricos. Demostrar también que $\mathbb{R}^n \ni x \to (x, 0) \in \mathbb{R}^{n+m}$ es una inyección isométrica cuando se consideran las distancias generadas por las normas $\|\cdot\|_1$, $\|\cdot\|_2$ o $\|\cdot\|_\infty$ (¿será esto cierto para cualquier norma?). Dar otros ejemplos de espacios isométricos.

Problema 1.22. Comprobar que las distancias d_∞, d_1 y d_2 definidas en el espacio producto $X \times Y$ son fuertememente equivalentes.

Problema 1.23. Demostrar que dos normas en un espacio vectorial son equivalentes si y sólo si las distancias que definen son fuertememente equivalentes.

Capítulo 2

Algunas desigualdades importantes

En este capítulo demostraremos algunas desigualdades clásicas debidas a Jensen, Young, Hölder y Minkowski, que resultarán útiles en esta asignatura y en muchas otras.

Comenzamos con una propiedad de las funciones convexas que nos servirá para deducir la desigualdad de Jensen.

Proposición 2.1. *Si* $h : (c, d) \subseteq \mathbb{R} \to \mathbb{R}$ *es convexa entonces para todo* $x_0 \in (c, d)$ *existe* $\xi \in \mathbb{R}$ *tal que* $h(x) \geq h(x_0) + \xi(x - x_0)$ *para todo* $x \in (c, d)$.

Demostración. Como la función h es convexa, se tiene, para todo $x, y, t, s \in (c, d)$ con $s < t < x_0 < x < y$, que

$$\frac{f(x_0) - f(s)}{x_0 - s} \leq \frac{f(x_0) - f(t)}{x_0 - t} \leq \frac{f(x) - f(t)}{x - t} \leq \frac{f(x) - f(x_0)}{x - x_0} \leq \frac{f(y) - f(x_0)}{y - x_0}.$$

Estas desigualdades implican que la función $(d, x_0) \ni x \mapsto \frac{f(x)-f(x_0)}{x-x_0}$ es creciente y es mayor o igual que la función $(c, x_0) \ni t \mapsto \frac{f(x_0)-f(t)}{x_0-t}$, que también es creciente. Se sigue que existen los límites $\lim_{x \to x_0^-} \frac{h(x)-h(x_0)}{x-x_0}$ y $\lim_{x \to x_0^+} \frac{h(x)-h(x_0)}{x-x_0}$, y que

$$\lim_{x \to x_0^-} \frac{h(x) - h(x_0)}{x - x_0} \leq \lim_{x \to x_0^+} \frac{h(x) - h(x_0)}{x - x_0}.$$

El primero de los límites es la pendiente de una recta tangente por la izquierda en x_0 a la gráfica de h, y el segundo es la pendiente de una recta tangente por la derecha en x_0 a la misma gráfica. Si tomamos ξ como cualquier número entre $\lim_{x \to x_0^-} \frac{h(x)-h(x_0)}{x-x_0}$ y $\lim_{x \to x_0^+} \frac{h(x)-h(x_0)}{x-x_0}$, tenemos que la recta de pendiente ξ que pasa por el punto $(x_0, h(x_0))$ toca a la gráfica de h en $(x_0, h(x_0))$ y queda por debajo de dicha gráfica. Es decir, tenemos que $h(x) \geq h(x_0) + \xi(x - x_0)$ para todo $x \in (c, d)$. □

Teorema 2.2 (Desigualdad de Jensen)**.** *Probar que si* f *es una función integrable en* (a, b) *y* $h : \mathbb{R} \to \mathbb{R}$ *es convexa entonces se tiene que*

$$h\left(\frac{1}{b - a} \int_a^b f(t)dt\right) \leq \frac{1}{b - a} \int_a^b h(f(t))dt. \tag{2.1}$$

Demostración. Consideremos $x_0 = \frac{1}{b-a} \int_a^b f(s)ds$. Por la proposición anterior sabemos que existe $\xi \in \mathbb{R}$ tal que $h(x) \geq h(x_0) + \xi(x - x_0)$ para todo $x \in \mathbb{R}$. Tomando x de la forma $x = f(t)$ obtenemos entonces

$$h(f(t)) \geq h\left(\frac{1}{b - a} \int_a^b f(t)dt\right) + \xi\left(f(t) - \frac{1}{b - a} \int_a^b f(s)ds\right)$$

para todo $t \in (a, b)$. Integrando ahora en (a, b) deducimos que

$$\int_a^b h(f(t))dt \geq (b-a)h\left(\frac{1}{b-a}\int_a^b f(t)dt\right) + \int_a^b \xi\left(f(t) - \frac{1}{b-a}\int_a^b f(s)ds\right)dt$$
$$= (b-a)h\left(\frac{1}{b-a}\int_a^b f(t)dt\right),$$

de donde se desprende la desigualdad deseada. □

Proposición 2.3 (Desigualdad de Young)**.** *Dados $1 < p, q < \infty$ tales que $\frac{1}{p} + \frac{1}{q} = 1$, se tiene que para cualesquiera $x, y > 0$,*

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

Demostración. Observemos que la función $\mathbb{R} \ni x \mapsto e^x \in \mathbb{R}$ es convexa. Entonces, como $1/p + 1/q = 1$, tenemos

$$xy = e^{\frac{\log(x^p)}{p} + \frac{\log(y^q)}{q}} \leq \frac{1}{p}e^{\log(x^p)} + \frac{1}{q}e^{\log(y^q)} = \frac{x^p}{p} + \frac{y^q}{q}.$$

□

Teorema 2.4 (Desigualdad de Hölder)**.** *Sean $f, g : (a, b) \to \mathbb{R}$ funciones que son continuas salvo en una cantidad finita de puntos, y p, q números reales tales que $1 < p, q < \infty$, y $1 = \frac{1}{p} + \frac{1}{q}$. Supongamos que las integrales $\int_a^b |f(t)|^p dt$ y $\int_a^b |g(t)|^q dt$ son finitas. Entonces se tiene que*

$$\int_a^b |f(t)g(t)|dt \leq \left(\int_a^b |f(t)|^p dt\right)^{\frac{1}{p}} \left(\int_a^b |g(t)|^q dt\right)^{\frac{1}{q}}. \tag{2.2}$$

Demostración. Si alguna de las integrales $\int_a^b |f(t)|^p dt$ o $\int_a^b |g(t)|^q dt$ es cero entonces necesariamente la correspondiente función es cero salvo a lo sumo en una cantidad finita de puntos, de lo que se deduce que $t \to f(t)g(t)$ también lo es, y por tanto la desigualdad resulta trivial.

Ahora, si $\int_a^b |f(t)|^p dt \neq 0 \neq \int_a^b |g(t)|^q dt$, podemos usar la desigualdad de Young con $x := |f(t)|/\left(\int_a^b |f(s)|^p ds\right)^{1/p}$, $y := |g(t)|/\left(\int_a^b |g(s)|^q ds\right)^{1/q}$, obteniendo que

$$\frac{|f(t)g(t)|}{\left(\int_a^b |f(s)|^p ds\right)^{1/p}\left(\int_a^b |g(s)|^q ds\right)^{1/q}} \leq \frac{1}{p}\frac{|f(t)|^p}{\int_a^b |f(s)|^p ds} + \frac{1}{q}\frac{|f(t)|^q}{\int_a^b |g(s)|^q ds},$$

lo que al integrar en (a, b) nos da

$$\int_a^b \frac{|f(t)g(t)|}{\left(\int_a^b |f(s)|^p ds\right)^{1/p}\left(\int_a^b |g(s)|^q ds\right)^{1/q}} \leq \frac{1}{p} + \frac{1}{q} = 1,$$

de donde se deduce la desigualdad del enunciado. □

Teorema 2.5 (Desigualdad de Minkowski)**.** *Sean $f, g : (a, b) \to \mathbb{R}$ funciones que son continuas salvo en una cantidad finita de puntos, y p un número tal que $1 \leq p < \infty$. Supongamos que las integrales $\int_a^b |f|^p$ y $\int_a^b |g|^p$ son finitas. Entonces se tiene que*

$$\left(\int_a^b |f+g|^p\right)^{1/p} \leq \left(\int_a^b |f|^p\right)^{1/p} + \left(\int_a^b |g|^p\right)^{1/p}.$$

Demostración. Como en el caso $p = 1$ la desigualdad es trivial, podemos supponer $p > 1$. Sea $q = p/(p-1)$ (de modo que se tiene $\frac{1}{p} + \frac{1}{q} = 1$). Usando la desigualdad de Hölder tenemos

$$\begin{aligned}
&\int_a^b |f+g|^p \leq \int_a^b |f+g|^{p-1}\left(|f|+|g|\right) = \int_a^b |f+g|^{p-1}|f| + \int_a^b |f+g|^{p-1}|g| \\
&\leq \int_a^b |f+g|^{p-1}|f| + \int_a^b |f+g|^{p-1}|g| = \int_a^b |f+g|^{p/q}|f| + \int_a^b |f+g|^{p/q}|g| \\
&\leq \left(\int_a^b |f+g|^p\right)^{1/q}\left(\int_a^b |f|^p\right)^{1/p} + \left(\int_a^b |f+g|^p\right)^{1/q}\left(\int_a^b |g|^p\right)^{1/p} \\
&= \left(\int_a^b |f+g|^p\right)^{(p-1)/p}\left(\left(\int_a^b |f|^p\right)^{1/p} + \left(\int_a^b |g|^p\right)^{1/p}\right) \\
&= \left(\int_a^b |f+g|^p\right)^{-1/p}\left(\left(\int_a^b |f|^p\right)^{1/p} + \left(\int_a^b |g|^p\right)^{1/p}\right),
\end{aligned}$$

de donde se desprende la desigualdad del enunciado. □

Con las mismas demostraciones, o bien aplicando los resultados anteriores a funciones escalonadas adecuadamente definidas, se deducen las versiones discretas de las desigualdades de Hölder y Minkowski.

Teorema 2.6. *Sean $p \in (1, \infty)$, y $q = p(p-1)$. Para cualesquiera vectores $a = (a_1, ..., a_n), b = (b_1, ..., b_n) \in \mathbb{R}^n$ se tiene que:*

$$\sum_{j=1}^n |a_j b_j| \leq \left(\sum_{j=1}^n |a_j|^p\right)^{1/p}\left(\sum_{j=1}^n |b_j|^q\right)^{1/q},$$

y que

$$\left(\sum_{j=1}^n |a_j + b_j|^p\right)^{1/p} \leq \left(\sum_{j=1}^n |a_j|^p\right)^{1/p} + \left(\sum_{j=1}^n |b_j|^p\right)^{1/p}.$$

También, pasando al límite cuando $n \to \infty$, o volviendo a aplicar las desigualdades integrales con funciones escalonadas adecuadamente definidas, se deducen las correspondientes desigualdades para sumas infinitas

Teorema 2.7. *Sean $p \in (1, \infty)$, y $q = p(p-1)$. Para cualesquiera sucesiones de números reales o complejos $(a_j)_{j\in\mathbb{N}}$, $(b_j)_{j\in\mathbb{N}}$ se tiene que:*

$$\sum_{j=1}^\infty |a_j b_j| \leq \left(\sum_{j=1}^\infty |a_j|^p\right)^{1/p}\left(\sum_{j=1}^\infty |b_j|^q\right)^{1/q},$$

y que

$$\left(\sum_{j=1}^\infty |a_j + b_j|^p\right)^{1/p} \leq \left(\sum_{j=1}^\infty |a_j|^p\right)^{1/p} + \left(\sum_{j=1}^\infty |b_j|^p\right)^{1/p}.$$

Capítulo 3

Conceptos topológicos básicos

A lo largo de este capítulo supondremos que X denota el espacio $\mathbb{R}^n$ o, más en general, cualquier espacio métrico. En el caso de que $(X, \|\cdot\|)$ sea un espacio normado, d designará la distancia asociada a la norma $\|\cdot\|$. Todos los resultados y definiciones de este capítulo son, por tanto, válidos en cualquier espacio métrico. Aunque la gran mayoría de los ejemplos y ejercicios se referirán al espacio $\mathbb{R}^n$ (y de hecho se aconsejará al alumno que para fijar ideas piense siempre en los casos $X = \mathbb{R}^2$ o $X = \mathbb{R}^3$), por simplicidad, y para evitar repeticiones en ulteriores desarrollos teóricos (como por ejemplo al definir abiertos y cerrados relativos a un subconjunto de $\mathbb{R}^n$), resulta conveniente definir los conceptos topológicos básicos en cualquier espacio métrico (X, d). Además la teoría de espacios métricos que desarrollaremos en este capítulo y los siguientes será de utilidad al alumno en futuras asignaturas de Análisis del grado.

Definición 3.1. Se dice que un subconjunto U de X es abierto si, para cada punto $x \in U$, existe $r > 0$ tal que $B(x, r) \subseteq U$. En otras palabras, U es abierto si es unión de bolas abiertas.

Se dice que un subconjunto F de X es cerrado si su complementario $X \setminus F$ es abierto.

Ejemplo 3.1. Toda bola abierta de X es un abierto, y toda bola cerrada es un cerrado.

Es importante observar que el r de la definición depende de cada punto x, y en general se hará cada vez más pequeño cuanto más próximo esté x de la *frontera* de U. Considérese, por ejemplo, el caso de una bola abierta, o el del complementario de una recta en $\mathbb{R}^2$.

Proposición 3.1. *La unión de una familia cualquiera de subconjuntos abiertos de X es un abierto. La intersección de una familia* finita *de abiertos es un abierto.*

Demostración. Puesto que un conjunto es abierto si y sólo si es unión (arbitraria) de bolas abiertas, la primera propiedad es evidente. Para demostrar la segunda, sean $A_1, ..., A_m$ subconjuntos abiertos de un esacio métrico (X, d). Dado $x \in \bigcap_{j=1}^m A_j$, se tiene que $x \in A_j$ para cada $j = 1, ..., m$. Puesto que los conjuntos A_j son abiertos, para cada $j \in \{1, ..., n\}$ existe $r_j > 0$ tal que $B(x, r_j) \subseteq A_j$. Definamos $r = \min\{r_1 ..., r_m\}$. Es claro que $r > 0$ y $B(x, r) \subseteq B(x, r_j) \subseteq A_j$ para cada $j \in \{1, ..., m\}$, luego también $B(x, r) \subseteq \bigcap_{j=1}^m A_j$. □

La condición de finitud en la segunda parte de la proposición anterior es esencial en general: considérese el ejemplo de una familia de bolas abiertas de centro el origen del plano $\mathbb{R}^2$ y de radio $1/n$, con $n \in \mathbb{N}$, cuya intersección es $\{0\}$, que no es abierto (¿por qué?). Lo mismo cabe observar de la segunda parte de la proposición siguiente.

Proposición 3.2. *La intersección de una familia cualquiera de subconjuntos cerrados de X es un cerrado. La unión de una familia* finita *de cerrados es un cerrado.*

Demostración. Se deduce de la proposición anterior tomando complementarios. Los detalles se dejan como ejercicio para el alumno. □

Debe observarse también que tanto el propio X como el subconjunto vacío $\emptyset$ son a la vez abiertos y cerrados de X.

Proposición 3.3. *Consideremos un conjunto X con dos métricas d_1 y d_2. Las distancias d_1 y d_2 son equivalentes si y sólo si definen los mismos conjuntos abiertos.*

Demostración. Supongamos primero que d_1 y d_2 son distancias equivalentes en X. Sea A un conjunto abierto en (X, d_1). Entonces para cada $x \in A$ existe $r_x^1 > 0$ tal que $B_{d_1}(x, r_x^1) \subseteq A$. Como las distancias d_1, d_2 son equivalentes, para cada una de las bolas $B_{d_1}(x, r_x^1)$ (en la distancia d_1) existe un $r_x^2 > 0$ tal que la bola $B_{d_2}(x, r_x^2)$ (en la distancia d_2) está contenida en $B_{d_1}(x, r_x^1)$. Por tanto tenemos $B_{d_2}(x, r_x^2) \subseteq B_{d_1}(x, r_x^1) \subseteq A$ para cada $x \in A$, lo que implica que A es abierto en (X, d_2). Análogamente se ve que si A es abierto en (X, d_2) entonces A es abierto en (X, d_1). Por consiguiente los espacios métricos (X, d_1) y (X, d_2) poseen los mismos abiertos.

Recíprocamente, supongamos que (X, d_1) y (X, d_2) tienen los mismos abiertos. Como una bola abierta $B_{d_1}(x, r_1)$ es un conjunto abierto en (X, d_1), y por tanto también en (X, d_2), y $x \in B_{d_1}(x, r_1)$, existe $r_2 > 0$ tal que $B_{d_2}(x, r_2) \subseteq B_{d_1}(x, r_1)$. Análogamente, para cada bola $B_{d_2}(x, s_2)$ existe $s_1 > 0$ tal que $B_{d_1}(x, s_1) \subseteq B_{d_2}(x, s_2)$. Consiguientemente las distancias d_1 y d_2 son equivalentes. □

Como en $\mathbb{R}^n$ todas las normas son equivalentes, y por el ejercicio 1.19 sabemos que dos normas son equivalentes si y sólo si inducen distancias equivalentes, se deduce inmediatamente lo siguiente.

Corolario 3.4. *La definición de conjunto abierto en $\mathbb{R}^n$ no depende de la norma empleada, es decir, diferentes normas en $\mathbb{R}^n$ dan lugar a los mismos conjuntos abiertos (y también los mismos cerrados).*

Definición 3.2. Se dice que un subconjunto W de X es un entorno de un punto $x \in X$ si existe un abierto U de X tal que $x \in U \subseteq W$. O lo que es lo mismo, si existe $r_x > 0$ tal que $B(x, r_x) \subseteq W$.

Ejercicio 3.1. Probar que un conjunto es abierto si y sólo si es entorno de cada uno de sus puntos.

Ejercicio 3.2. Demostrar que la intersección de una familia *finita* de entornos de un punto x sigue siendo entorno de x.

Aunque en este curso no se hablará de espacios topológicos, merece la pena reseñar que siempre que X es un conjunto no vacío en el que se ha destacado una familia τ de subconjuntos, llamados por definición *abiertos*, que satisfacen la Proposición 3.1, y de modo que tanto X como $\emptyset$ están en la familia τ, se dice que (X, τ) es un espacio topológico, y de τ se dice que es una topología en X.

Pasamos ahora a estudiar los conceptos de adherencia e interior.

Definición 3.3. Se dice que un punto x es adherente a un subconjunto A de X si todo entorno de x corta a A, es decir, si para todo $r > 0$ existe $y \in A$ tal que $d(y, x) < r$. Se define la adherencia de A, y se denota $\overline{A}$, como el conjunto de todos los puntos adherentes a A.

Por ejemplo, $(0,0)$ es un punto adherente del conjunto $A = \{(x,y) \in \mathbb{R}^2 : y = 1/n, n \in \mathbb{N}\}$ en el plano.

Ejercicio 3.3. Probar que, para todo $A \subseteq X$, su adherencia $\overline{A}$ es un cerrado.

Definición 3.4. Se dice que un punto x es interior a un subconjunto A de X si A es entorno de x, es decir, si existe $r > 0$ tal que $B(x,r) \subseteq A$. Se define el interior de A, y se denota $\text{int}(A)$, como el conjunto de todos los puntos interiores a A.

Por ejemplo, $(1/2, 0)$ es un punto interior de $\{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 1\}$.

Proposición 3.5. *Sean A, B subconjuntos de X. Se tiene que:*

1. *$\mathit{int}(A) \subseteq A$.*
2. *Si $A \subseteq B$ entonces $\mathit{int}(A) \subseteq \mathit{int}(B)$.*
3. *$A \subseteq \overline{A}$.*
4. *Si $A \subseteq B$ entonces $\overline{A} \subseteq \overline{B}$.*
5. *El interior de A es el mayor abierto contenido en A, esto es,*
$$\mathit{int}(A) = \bigcup \{U \subset X : U \ \mathit{abierto}\ , \ U \subseteq A\}.$$
6. *$X \setminus \overline{A} = \mathit{int}(X \setminus A)$.*
7. *$\overline{X \setminus A} = X \setminus \mathit{int}(A)$.*
8. *La adherencia de A es el menor cerrado que contiene a A, es decir,*
$$\overline{A} = \bigcap \{F \subset X : F \ \mathit{cerrado}\ , \ A \subseteq F\}.$$

Demostración. (1), (2), (3) y (4) son evidentes a partir de la definición de interior y de adherencia. Para probar (5), veamos primero que $\text{int}(A)$ es abierto para todo conjunto A. Dado cualquier $x \in \text{int}(A)$, existe $r_x > 0$ tal que $B(x,r) \subseteq A$. Como la bola $B(x, r_x)$ es abierta, para cada $y \in B(x, r_x)$ tal que $B(y, s_y) \subseteq B(x,r) \subseteq A$. Por tanto $y \in \text{int}(A)$ para cada $y \in B(x, r_x)$. Es decir, $B(x, r_x) \subseteq \text{int}(A)$, para cada $x \in \text{int}(A)$. Luego
$$\text{int}(A) \subseteq \bigcup_{x \in \text{int}(A)} B(x, r_x) \subseteq \text{int}(A),$$
y por tanto
$$\text{int}(A) = \bigcup_{x \in \text{int}(A)} B(x, r_x)$$
es abierto por ser unión de bolas abiertas. Ahora, como $\text{int}(A)$ es abierto y está contenido en A, es evidente que
$$\text{int}(A) \subseteq \bigcup \{U \subset X : U \ \text{abierto}\ , \ U \subseteq A\}.$$
Por otro lado, si U es abierto, entonces por las definiciones de conjunto abierto y de interior es evidente que $U \subseteq \text{int}(U)$, y por tanto que $\text{int}(U) = U$. Luego, para todo abierto U con $U \subseteq A$, se tiene $U = \text{int}(U) \subseteq \text{int}(A)$, y por tanto
$$\text{int}(A) \supseteq \bigcup \{U \subset X : U \ \text{abierto}\ , \ U \subseteq A\}.$$

Esto prueba (5).

Las demostraciones de (6) y (7) son inmediatas. Finalmente (8) se demuestra fácilmente a partir de (5), tomando complementarios y usando (6) y (7). Los detalles quedan a cargo del lector. $\square$

Corolario 3.6. *Un conjunto A es abierto si y sólo si $A = int(A)$. Un conjunto C es cerrado si y sólo si $\overline{C} = C$. Además,*

1. *$int(int(A)) = int(A)$;*
2. *$\overline{\overline{C}} = \overline{C}$.*

Demostración. En la demostración de la proposición anterior ya se ha visto que si A es abierto entonces $A = \text{int}(A)$; también se ha avisto que el interior de un conjunto es siempre abierto. Por tanto es también claro que si $A = \text{int}(A)$ entonces A es abierto. Considerando $A = X \setminus C$ y aplicando (6) y (7) de la proposición anterior y el hecho recién demostrado que A es abierto si y sólo si $A = \text{int}(A)$, deducimos que C es cerrado si y sólo si $\overline{C} = C$.

Finalmente, como $\text{int}(A)$ es siempre abierto, se deduce que $\text{int}(\text{int}(A)) = \text{int}(A)$. Análogamente, $\overline{\overline{C}} = \overline{C}$. $\square$

Definamos ahora la frontera y la acumulación de un conjunto.

Definición 3.5. Se define la frontera de un subconjunto A de X como

$$\partial A := \overline{A} \setminus \text{int}(A) = \overline{A} \cap \overline{X \setminus A}.$$

Es decir, $x \in \partial A$ si para todo $r > 0$ se tiene que $B(x, r) \cap A \neq \emptyset \neq B(x, r) \cap (X \setminus A)$.

Por ejemplo, si $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ y $B = \{(x, y) \in \mathbb{R}^2 : x, y \in \mathbb{Q}\}$ entonces $\partial A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, y $\partial B = \mathbb{R}^2$.

Conviene notar que, como consecuencia inmediata de la definición, la frontera de un conjunto es siempre un cerrado.

Definición 3.6. Se dice que x es un *punto de acumulación* de un subconjunto A de X si todo entorno de x contiene un punto de A distinto de x, es decir, si para todo $r > 0$ se tiene $B(x, r) \cap (A \setminus \{x\}) \neq \emptyset$. Denotaremos el conjunto de todos los puntos de acumulación de A por A'.

Si $x \in A$ y x no es punto de acumulación de A, se dice que x es un punto aislado de A.

Por ejemplo, el origen es punto de acumulación del conjunto $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1/n^2, n \in \mathbb{N}\}$, que no tiene puntos aislados. Por otro lado, el conjunto $\mathbb{N} \times \mathbb{N}$ no tiene ningún punto de acumulación en $\mathbb{R}^2$ y todos sus puntos son aislados. Por último, todos los puntos del conjunto del plano $C = \{(n, n + 1/2m) : n, m \in \mathbb{N}\}$ son aislados, pero tiene infinitos puntos de acumulación (a saber, $C' = \{(n, n) : n \in \mathbb{N}\}$).

Proposición 3.7. *Para todo $A \subset X$ se tiene que $\overline{A} = A \cup A'$. De hecho la adherencia de A es la unión de la acumulación de A con todos los puntos aislados de A.*

En particular se tiene que un conjunto A es cerrado si y sólo si contiene todos sus puntos de acumulación, es decir, si y solo si $A' \subseteq A$.

Demostración. Por la definición es evidente que todo punto de acumulación de A es un punto de adherencia de A, y que todo punto de A es adherente a A. Luego se tiene $A \cup A' \subseteq \overline{A}$. Por otro lado, si $x \in \overline{A} \setminus A'$ entonces para algún $r > 0$ se tiene $B(x, r) \cap (A \setminus \{x\}) = \emptyset$ y $B(x, r) \cap A \neq \emptyset$, luego necesariamente $x \in A$. Hemos probado así que $\overline{A} = A \cup A'$. Como los puntos de A que no están en A' son por definición puntos aislados de A, también se tiene que a adherencia de A es la unión de la acumulación de A con todos los puntos aislados de A. $\square$

Definición 3.7. Si (X, d) es un espacio métrico e Y es un subconjunto no vacío de X, es evidente que la distancia d de X, al restringirla a Y, sigue teniendo las propiedades $1, 2$ y 3 de la definición 1.9. Luego (Y, d) tiene una estructura natural de espacio métrico. Se dice que (Y, d) es un *subespacio métrico* de (X, d). Por consiguiente todas las definiciones y resultados anteriores se aplican también al espacio (Y, d). En este contexto, si un subconjunto A de Y es abierto en (Y, d), diremos que A es un *abierto relativo* a Y. Esto, en general, no quiere decir que A sea abierto en X, como prueba el ejemplo siguiente.

Ejemplo 3.2. Sean $X = \mathbb{R}$, $Y = [0, 5)$, $A = [0, 3)$. El conjunto A es un abierto relativo a Y, pero no es abierto en X.

Análogamente, se dice que un conjunto C es cerrado relativo a Y si C es cerrado en (Y, d). Nuevamente, esto no implica en general que C sea un cerrado de X (¿por qué?). Podemos definir de igual modo interior, adherencia, frontera y acumulación relativa de un subconjunto de Y.

El siguiente resultado caracteriza los abiertos y cerrados relativos de un subespacio métrico Y de X.

Proposición 3.8. *Sea Y un subconjunto no vacío de un espacio métrico (X, d). Entonces*

1. *A es un abierto relativo a Y si y sólo si existe G abierto en X tal que $A = G \cap Y$.*
2. *C es un cerrado relativo a Y si y sólo si existe F cerrado en X tal que $C = F \cap Y$.*

Demostración. En efecto, si $A = G \cap Y$, con G abierto en X, es evidente que A es abierto en Y. Recíprocamente, si A es abierto relativo a Y, para cada $y \in A$ existe $r_y > 0$ tal que $B_Y(y, r_y) \subset A$. Definamos

$$G = \bigcup_{y \in A} B_X(y, r_y),$$

entonces es claro que G es abierto en X y $G \cap Y = A$. El caso de cerrados puede tratarse tomando complementarios. $\square$

Corolario 3.9. *Si Y es abierto en X y A es un abierto relativo a Y, entonces A es también abierto en X. Análogamente, si Y es cerrado en X y C es cerrado relativo a Y, entonces C es cerrado en X.*

Es bien sabido que el conjunto $\mathbb{Q}$ es denso en $\mathbb{R}$, es decir, en todo intervalo abierto de números reales siempre hay algún número racional, o bien $\overline{\mathbb{Q}} = \mathbb{R}$. La noción de densidad puede generalizarse a cualquier espacio métrico.

Definición 3.8. Se dice que un conjunto D de X es denso si $\overline{D} = X$. Cuando en un espacio métrico X existe un subconjunto numerable que es denso, se dice que X es *separable*.

Utilizando el hecho de que $\overline{\mathbb{Q}} = \mathbb{R}$, no es difícil ver que $\mathbb{Q}^n$ es denso en $\mathbb{R}^n$ (problema 3.18). En particular se deduce que $\mathbb{R}^n$ es separable. El siguiente teorema nos revela una de las propiedades más interesantes y utilizadas de los espacios separables.

Teorema 3.10. *Si X es un espacio métrico separable, entonces todo recubrimiento por abiertos de X tiene un subrecubrimiento numerable. Es decir, si*

$$X \subseteq \bigcup_{\alpha \in I} U_\alpha,$$

donde $\{U_\alpha\}_{\alpha\in I}$ es una familia cualquiera de abiertos, entonces existe una sucesión (α_n) de índices de I tales que

$$X \subseteq \bigcup_{n=1}^{\infty} U_{\alpha_n}.$$

Demostración. Sea $D = \{x_n : n \in \mathbb{N}\}$ un subconjunto denso de X. Para cada x_n definamos

$$R_n := \sup\{r > 0 \,|\, \exists \alpha \in I : B(x_n, r) \subseteq U_\alpha\}.$$

Como el conjunto es no vacío (ya que la unión de los U_α recubre X), $R_n \in (0, +\infty]$ está bien definido.

Para cada $n \in \mathbb{N}$ podemos escoger entonces $\alpha_n \in I$ y $r_n > 0$ tales que

$$B(x_n, r_n) \subseteq U_{\alpha_n} \text{ y}$$
$$r_n > \tfrac{R_n}{2} \text{ si } R_n < +\infty, \;\; r_n = 1 \text{ cuando } R_n = +\infty.$$

Veamos que $X \subseteq \bigcup_{n=1}^{\infty} U_{\alpha_n}$. Sea $x \in X$, entonces existe $\alpha \in I$ con $x \in U_\alpha$ y, como U_α es abierto, podemos encontrar $r \in (0,1)$ tal que $B(x, 3r) \subseteq U_\alpha$. Ahora, por la densidad de D existe $n \in \mathbb{N}$ tal que $d(x, x_n) < r$. Entonces

$$x \in B(x_n, r) \subset B(x_n, 2r) \subset B(x, 3r) \subseteq U_\alpha.$$

Por la definición de R_n esto implica que $2r \leq R_n$. Luego

$$r \leq \min\{\frac{R_n}{2}, 1\} \leq r_n,$$

y de esto se sigue que $x \in B(x_n, r) \subset B(x_n, r_n) \subseteq U_{\alpha_n}$, es decir, $x \in U_{\alpha_n}$. □

Finalizamos el capítulo con una observación importante. Por el Corolario 3.4 sabemos que la definición de abierto no depende de la norma que consideremos en $\mathbb{R}^n$. Por otro lado, resulta evidente que todos los conceptos que se han definido en este capítulo pueden formularse exclusivamente en términos de abiertos. Por consiguiente ni el interior, ni la adherencia, ni la acumulación, ni la frontera, ni la densidad de un subconjunto de $\mathbb{R}^n$ dependen de la norma considerada en $\mathbb{R}^n$.

3.1. Problemas

Problema 3.1. Determinar si los siguientes subconjuntos de $\mathbb{R}^n$ son abiertos o cerrados. Hallar también su interior, adherencia, frontera y puntos de acumulación:

$A = \{(x, y) \in \mathbb{R}^2 : 0 < x \leq 1\}$;

$B = \{(x, y) \in \mathbb{R}^2 : xy > 1\}$;

$C = \{(x, y, z) \in \mathbb{R}^3 : y^2 + z^2 \leq 1, 0 \leq x \leq 1\}$;

$D = \{(x, y) \in \mathbb{R}^2 : y \leq x^3, x \in \mathbb{Q}\}$;

$E = \{(x, y) \in \mathbb{R}^2 : y = x/n, n \in \mathbb{N}\}$;

$F = \{(x, y) \in \mathbb{R}^2 : y = rx, r \in \mathbb{Q}\}$;

$G = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1, z \geq x^2 + y^2\}$;

$H = \{(x, y, z, w) \in \mathbb{R}^4 : x + y - x + w = 5\}$;

$I = \{(x, y) \in \mathbb{R}^2 : y = \sin(1/x), x \neq 0\}$;

$J = \{(x, y) \in \mathbb{R}^2 : y = f(x)\}$, donde $f : \mathbb{R} \to \mathbb{R}$ es una función continua.

Problema 3.2. Sea A un subconjunto abierto de $\mathbb{R}^n$, y $B \subset \mathbb{R}^n$ un subconjunto cualquiera no vacío. Probar que el conjunto

$$A + B := \{x + y \in \mathbb{R}^n : x \in A, y \in B\}$$

es un abierto de $\mathbb{R}^n$.

Problema 3.3. Sea A un subconjunto cualquiera no vacío de X, $r > 0$, y definamos $B = \{x \in X : d(x, y) < r \text{ para algún } y \in A\}$. Probar que B es abierto.

Problema 3.4. Demostrar que la esfera $S(x, r) = \{y \in X : d(y, x) = r\}$ de un espacio métrico es un conjunto cerrado.

Problema 3.5. Demostrar que cualquier hiperplano en $\mathbb{R}^n$ es cerrado y tiene interior vacío.

Problema 3.6. Demostrar que si $A \subseteq B$ entonces int$(A) \subseteq$ int(B), y que $\overline{A} \subseteq \overline{B}$. Se dice por esto que la adherencia y el interior son operadores monótonos.

Problema 3.7. ¿Es verdad que int$(A) \cup$ int$(B) =$ int$(A \cup B)$?

Problema 3.8. Probar que int$(A \cap B) =$ int$(A) \cap$ int(B).

Problema 3.9. Demostrar que $\overline{A \cup B} = \overline{A} \cup \overline{B}$. ¿Es cierto que $\overline{A \cap B} = \overline{A} \cap \overline{B}$?

Problema 3.10. Probar que cualquier conjunto finito es cerrado.

Problema 3.11. Demostrar que $x \in A'$ si y sólo si cada entorno de x contiene infinitos puntos de A.

Problema 3.12. Sea A un subconjunto acotado superiormente de $\mathbb{R}$. Probar que $\sup A \in \overline{A}$.

Problema 3.13. Identificando $\mathbb{R}^n \times \mathbb{R}^m$ con $\mathbb{R}^{n+m}$, demostrar que un subconjunto A de $\mathbb{R}^{n+m}$ es abierto si y sólo si para cada $(x, y) \in A$ existen U abierto de $\mathbb{R}^n$ y V abierto de $\mathbb{R}^m$ tales que $x \in U$, $y \in V$, y $U \times V \subset A$. *Indicación:* usar la norma $\|\cdot\|_\infty$ y recordar el problema 1.10.

Problema 3.14. Demostrar que si A es abierto en $\mathbb{R}^n$ y B es abierto en $\mathbb{R}^m$ entonces el producto cartesiano $U \times V$ es abierto en $\mathbb{R}^{n+m}$.

Problema 3.15. Probar que si $A \subseteq B$ entonces $A' \subseteq B'$.

Problema 3.16. Probar que $(A \cup B)' = A' \cup B'$.

Problema 3.17. Probar que A' es siempre cerrado.

Problema 3.18. Probar que $\mathbb{Q}^n$ es denso en $\mathbb{R}^n$.

Problema 3.19. Demostrar que cualquier abierto de $\mathbb{R}^n$ es unión *numerable* de bolas abiertas. *Indicación:* considerar bolas con centros cuyas coordenadas son todas racionales y los radios también.

Problema 3.20. Generalizar el resultado anterior. Demostrar que en un espacio métrico separable cualquier abierto es unión numerable de bolas abiertas.

Problema 3.21. Si D es denso en (X, d) y U es un abierto de X, probar que $D \cap U$ es denso en (U, d). ¿Es esto cierto si U no es abierto?

Problema 3.22. Demostrar que la acotación de un conjunto no puede caracterizarse en términos de abiertos (es decir, no es un concepto topológico, sino puramente métrico). *Indicación:* ver el ejemplo 1.3.

Capítulo 4

Sucesiones, completitud y compacidad

En este capítulo, y por las mismas razones apuntadas en el precedente, continuaremos suponiendo que X es un espacio métrico, aunque se recomendará al alumno que tenga siempre presente el caso en el que X es $\mathbb{R}^n$ o un subconjunto de $\mathbb{R}^n$.

Recordemos que una sucesión (x_k) en un conjunto X no es más que una aplicación $k \mapsto x_k$ de $\mathbb{N}$ en X. Por otro lado, se dice que una sucesión (y_j) es una subsucesión de (x_k) si existe una inyección creciente $j \mapsto k_j$ de $\mathbb{N}$ en $\mathbb{N}$ tal que $y_j = x_{k_j}$ para todo j, y en este caso se denota $(y_j) = (x_{k_j})$. Recordemos que, si H es un subconjunto infinito de $\mathbb{N}$, existe una sola inyección creciente j de $\mathbb{N}$ en $\mathbb{N}$ tal que $j(\mathbb{N}) = H$. Por tanto, una subsucesión de (x_k) queda perfectamente determinada por el conjunto H de aquellos índices k que son términos de la subsucesión.

La definición siguiente no es otra cosa que la extensión natural de la noción de sucesión convergente en $\mathbb{R}$ al caso más general de espacios métricos.

Definición 4.1. Se dice que una sucesión (x_n) es convergente a un punto x de (X, d) (al que se llamará límite de dicha sucesión, y se denotará $x = \lim_{k\to\infty} x_k$) si para todo $\varepsilon > 0$ existe $k_0 \in \mathbb{N}$ tal que si $k \geq k_0$ entonces $d(x_k, x) \leq \varepsilon$.

También se escribirá $x_n \to x$ para denotar $x = \lim_{n\to\infty} x_n$.

Por supuesto, en el caso de $\mathbb{R}^n$ o de un espacio normado cualquiera, esto es lo mismo que decir que para todo $\varepsilon > 0$ existe $k_0 \in \mathbb{N}$ tal que si $k \geq k_0$ entonces $\|x_k - x\| \leq \varepsilon$. Otra formulación equivalente (válida en cualquier espacio métrico) es la siguiente: $\lim_{k\to\infty} x_k = x$ si y sólo si para todo entorno V de x existe $k_0 \in \mathbb{N}$ tal que si $k \geq k_0$ entonces $x_k \in V$.

Por ejemplo, es fácil ver que la sucesión $\{\big(n\sin(1/n), n^2/(n^3+5n^2+7)\big)\}$ converge a $(1,0)$ en $\mathbb{R}^2$.

Proposición 4.1. *El límite de una sucesión (x_k) de X, si existe, es único.*

Demostración. Supongamos que x e y son límites de una sucesión (x_k), y que $x \neq y$. Sea $\varepsilon := d(x,y)/3$, que es un número estrictamente positivo por ser $x \neq y$. Puesto que (x_k) converge a x, existe $k_1 \in \mathbb{N}$ tal que si $k \geq k_1$ entonces $d(x_k, x) \leq \varepsilon$. Como (x_k) también converge a y, existe $k_2 \in \mathbb{N}$ tal que si $k \geq k_2$ entonces $d(x_k, y) \leq \varepsilon$. Luego, para todo $k \geq \max\{k_1, k_2\}$ tendremos que

$$3\varepsilon = d(x,y) \leq d(x, x_k) + d(x_k, y) \leq \varepsilon + \varepsilon = 2\varepsilon,$$

lo cual es absurdo. $\square$

Ejercicio 4.1. Si (x_n) es convergente a x entonces cualquier subsucesión (x_{k_j}) de (x_k) también converge a x.

La siguiente proposición es evidente, pero no por eso deja de ser menos útil.

Proposición 4.2. *Una sucesión (x_n) converge a un punto x en (X,d) si y sólo si $d(x_n,x)$ converge a 0 en $\mathbb{R}$.*

Como en el caso de las sucesiones de $\mathbb{R}$, a efectos de determinar si una sucesión tiene o no límite sólo importa lo que ocurra con las colas de la sucesión.

Observación 4.3. Si (x_k), (y_k) son dos sucesiones de X tales que existe $m\in\mathbb{N}$ de modo que $x_k=y_k$ para todo $k\geq m$, entonces (x_k) converge (a un punto x) si y sólo si (y_k) converge (al punto x).

También es muy fácil ver que en un espacio normado las sucesiones convergentes son las mismas para todas las normas que sean equivalentes entre sí.

Observación 4.4. Si $\|\cdot\|_1$ y $\|\cdot\|_2$ son dos normas equivalentes en un espacio vectorial X, entonces toda sucesión convergente en $(X,\|\cdot\|_1)$ es convergente en $(X,\|\cdot\|_2)$, y tiene el mismo límite.

La proposición siguiente nos proporciona una manera muy fácil de calcular límites de sucesiones en $\mathbb{R}^n$: basta usar las técnicas ya conocidas en $\mathbb{R}$ para calcular los límites de las sucesiones coordenadas; el vector cuyas componentes son dichos límites será el límite de la sucesión de vectores de $\mathbb{R}^n$.

De aquí en adelante, cuando consideremos sucesiones de $\mathbb{R}^n$ utilizaremos la notación siguiente. Si (x_k) es una sucesión de $\mathbb{R}^n$, escribiremos cada vector con sus n componentes

$$x_k=(x_k^1,x_k^2,...,x_k^n).$$

A las n sucesiones (x_k^1), ..., (x_k^n) las llamaremos sucesiones coordenadas de (x_k), y un punto $x\in\mathbb{R}^n$ lo denotaremos también

$$x=(x^1,...,x^k).$$

Con esta notación se tiene lo siguiente.

Proposición 4.5. *La sucesión (x_k) converge a x en $\mathbb{R}^n$ si y sólo si, para cada $j=1,...,n$, la sucesión coordenada (x_k^j) converge a x_j. Es decir,*

$$\lim_{k\to\infty}x_k=x \iff \forall j=1,...,n \ \lim_{k\to\infty}x_k^j=x_j.$$

Demostración. Como sabemos que las normas $\|\cdot\|_1$, $\|\cdot\|_2$ y $\|\cdot\|_\infty$ son equivalentes en $\mathbb{R}^n$, basta que demostremos este resultado para una de estas normas. Lo más práctico en este caso será usar la norma $\|\cdot\|_\infty$. Supongamos primero que (x_k) es convergente a x; es decir, que $\|x_k-x\|_\infty\to 0$. Entonces, puesto que, para cada $j\in\{1,...,n\}$ es

$$|x_k^j-x^j|\leq\max_{1\leq i\leq n}|x_k^i-x^i|=\|x_k-x\|_\infty,$$

es evidente que también tendremos $\lim_{k\to\infty}|x_k^j-x^j|=0$, es decir, $\lim_{k\to\infty}x_k^j=x^j$.

Recíprocamente, supongamos que $\lim_{k\to\infty}x_k^j=x^j$ para cada $j\in\{1,...,n\}$; es decir, $\lim_{k\to\infty}|x_k^j-x^j|=0$ para cada $j\in\{1,...,n\}$. Por tanto, dado $\varepsilon>0$, para cada $j\in\{1,...,n\}$ existe $k_j\in\mathbb{N}$ tal que $|x_k^j-x^j|\leq\varepsilon$ para todo $k\leq k_j$. Tomando $k_0:=\max\{k_1,...,k_n\}$ tenemos entonces que

$$|x_k^j-x^j|\leq\varepsilon \text{ para todo } j\in\{1,...,n\}, \text{ y para todo } k\leq k_0,$$

lo que significa que

$$\|x_k-x\|_\infty\leq\varepsilon \text{ para todo } k\geq k_0,$$

y esto prueba que $\lim_{k\to\infty}x_k=x$ en $(\mathbb{R}^n,\|\cdot\|_\infty)$. $\square$

Conviene señalar que la convergencia de un sucesión es un concepto topológico, es decir, puede formularse exclusivamente en términos de conjuntos abiertos. En efecto, es evidente que $x_n \to x$ en X si y sólo si para todo entorno abierto U de x existe $n_0 = n_0(U) \in \mathbb{N}$ tal que $x_n \in U$ si $n \geq n_0$.

Y también a la inversa, todos los conceptos topológicos estudiados en el capıtulo anterior podrían haberse definido en términos de sucesiones. En efecto, la proposición siguiente caracteriza la adherencia, la frontera y la acumulación de un conjunto en términos de límites de sucesiones de dicho conjunto o su complementario. Una vez definida la adherencia $\overline{A}$ de un conjunto A, podemos llamar cerrados a los conjuntos C tales que $\overline{C} = C$, y abiertos a sus complementarios.

Proposición 4.6. *Sea A un subconjunto de (X, d), y x un punto de X. Entonces:*

1) $x \in \overline{A}$ *si y sólo si existe una sucesión* $(x_k) \subset A$ *tal que* $\lim_{k\to\infty} x_k = x$ *en* X*;*

2) $x \in A'$ *si y sólo si existe* $(x_k) \subset A \setminus \{x\}$ *tal que* $\lim_{k\to\infty} x_k = x$ *en* X*;*

3) $x \in \partial A$ *si y sólo si existen sucesiones* $(x_k) \subset A$ *e* $(y_k) \subset X \setminus A$ *tales que* $x = \lim_{k\to\infty} x_k = \lim_{k\to\infty} y_k = x$ *en* X*.*

Demostración. Demostremos (1). Si $x \in \overline{A}$ entonces para cada $r > 0$ sabemos que $A \cap B(x, r) \neq \emptyset$. Considerando r de la forma $r = 1/k$, $k \in \mathbb{N}$, obtenemos una sucesión (x_k) tal que $x_k \in A \cap B(x, 1/k)$ par cada $k \in \mathbb{N}$, lo cual implica que $(x_k) \subseteq A$ y que $\lim_{k\to\infty} x_k = x$. Recíprocamente, si existe una sucesión (x_k) con estas dos propiedades, entonces dado cualquier $r > 0$, como $\lim_{k\to\infty} x_k = x$, existe $k_r \in \mathbb{N}$ tal que $x_k\, inB(x, r)$ para todo $k \geq k_r$, y como además $(x_k) \subseteq A$ deducimos que $x_k \in A \cap B(x, r)$ para todo $k \geq k_r$, y en particular que $A \cap B(x, r) \neq \emptyset$. Como esto es cierto para todo $r > 0$ concluimos que $x \in \overline{A}$. Las demostraciones de (2) y (3) son similares y quedan a cargo del lector. □

En particular se tiene un corolario que resulta a veces muy útil para ver si un conjunto es cerrado.

Corolario 4.7. *Un conjunto A de X es cerrado si y sólo si, para toda sucesión $(x_k) \subset A$ y todo $x \in X$ tales que* $\lim_{k\to\infty} x_k = x$*, se tiene que $x \in A$.*

Demostración. Supongamos primero que A es cerrado, es decir que $A = \overline{A}$. Dada una sucesión $(x_k) \subset A$ y un punto $x \in X$ tales que $\lim_{k\to\infty} x_k = x$, por la proposición anterior tenemos que $x \in \overline{A}$. Como por hipótesis $A = \overline{A}$, se deduce que $x \in A$.

Recíprocamente, supongamos que A tiene la propiedad del enunciado, y probemos que $A = \overline{A}$, para lo cual basta demostrar que $\overline{A} \subseteq A$. Dado $x \in \overline{A}$, por la proposición anterior existe una sucesión $(x_k) \subset A$ tal que $\lim_{k\to\infty} x_k = x$. Entonces, usando la hiótesis, deducimos que $x \in A$. □

El siguiente resultado caracteriza los puntos de acumulación de las sucesiones de X.

Proposición 4.8. *Un punto x es de acumulación de una sucesión (x_k) si y sólo si existe una subsucesión (x_{k_j}) de (x_k) con $x_{k_j} \neq x$ para todo j y* $x = \lim_{j\to\infty} x_{k_j}$*.*

Demostración. Denotemos $A = \{x_k : k \in \mathbb{N}\}$, y supongamos que $x \in A'$, es decir, que $B(x, r) \cap (A \setminus \{x\}) \neq \emptyset$ para todo $r > 0$. Comenzamos tomando $r = 1$ para encontrar $x_{k_1} \in B(x, 1) \cap (A \setminus \{x\})$. Ahora consideramos

$$r_1 = \frac{1}{2} \min\{d(x_k, x) : k \leq k_1, x_k \neq x\},$$

que es un número estrictamente positivo, y usamos que $B(x,r_1)\cap(A\setminus\{x\})\neq\emptyset$ para encontrar x_{k_2} tal que $x_{k_2}\in B(x,r_1)\cap(A\setminus\{x\})$. Obsérvese que, por la definición de r_1, tenemos $d(x_{k_2},x)\leq 1/2$ y también $x_{k_2}\neq x_k$ para todo $k\leq k_1$, lo cual implica que $k_2>k_1$. Proseguimos considerando

$$r_2=\frac{1}{3}\min\{d(x_k,x):k\leq k_2, x_k\neq x\},$$

y usamos que $B(x,r_2)\cap(A\setminus\{x\})\neq\emptyset$ para encontrar x_{k_3} tal que $x_{k_3}\in B(x,r_2)\cap(A\setminus\{x\})$. Ahora tenemos $d(x_{k_3},x)\leq 1/3$ y también $x_{k_3}\neq x_k$ para todo $k\leq k_2$, lo que impone que $k_3>k_2$. Continuando este proceso con un argumento inductivo obtenemos una colección $(x_{k_j})_{j\in\mathbb{N}}$ de puntos de A con las propiedades de que $d(x_{k_j},x)\leq 1/j$ y $k_{j+1}>k_j$ para todo $j\in\mathbb{N}$. Esto implica que $(x_{k_j})_{j\in\mathbb{N}}$ es subsucesión de $(x_k)_{k\in\mathbb{N}}$, y que $\lim_{j\to\infty}x_{k_j}=x$.

El recíproco es consecuencia inmediata de la Proposición 4.6. □

Recordemos que si (x_k) es una sucesión acotada de $\mathbb{R}$ entonces (x_k) tiene una subsucesión convergente. Esto es lo que nos dice el teorema de Bolzano-Weierstrass. Este teorema puede extenderse al espacio $\mathbb{R}^n$, como veremos hacia el final de este capítulo cuando estudiemos la noción de compacidad, aunque el lector inquieto puede intentar una demostración directa (hágase primero en $\mathbb{R}^2$ aplicando dos veces el teorema de Bolzano-Weierstrass en $\mathbb{R}$, y después empléese un argumento inductivo).

Pasamos ahora a estudiar la noción de completitud de un espacio métrico.

Definición 4.2. Se dice que (x_k) es una sucesión de Cauchy en (X,d) si para todo $\varepsilon>0$ existe $k_0\in\mathbb{N}$ tal que si $k,j\geq k_0$ entonces $d(x_k,x_j)\leq\varepsilon$.

Se dice que el espacio (X,d) es completo si toda sucesión de Cauchy en X converge.

Se dice que un espacio vectorial normado $(E,\|\cdot\|)$ es completo si como espacio métrico es completo para la distancia inducida por su norma. A los espacios normados completos se les llama también *espacios de Banach*.

Por ejemplo, es ya sabido que $(\mathbb{R},|\cdot|)$ es completo. Sin embargo $(\mathbb{Q},|\cdot|)$ no lo es (por ejemplo la sucesión de números racionales $(1+1/n)^n$ es de Cauchy pero no converge en $\mathbb{Q}$).

Proposición 4.9. *Toda sucesión convergente es de Cauchy.*

Demostración. Si (x_k) es una sucesión convergente en un espacio métrico (X,d), dado $\varepsilon>0$ existe $k_\varepsilon>0$ tal que $d(x_k,x)\leq\varepsilon/2$ para todo $k\geq k_\varepsilon$. Por tanto, para todos $k,j\geq k_\varepsilon$ tenemos

$$d(x_k,x_j)\leq d(x_k,x)+d(x,x_j)\leq\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

Esto prueba que (x_k) es de Cauchy en (X,d). □

El recíproco no es verdad en general, como prueba el ejemplo anterior.

La noción de sucesión de Cauchy es estable por paso a subsucesiones.

Proposición 4.10. *Toda subsucesión de una sucesión de Cauchy es también de Cauchy.*

Demostración. Sea (x_k) de Cauchy en (X,d), y sea (x_{k_j}) una subsucesión de (x_k). Nótese que al ser (x_{k_j}) subsucesión de (x_k) se tiene que $j\mapsto k_j$ es una inyección creciente de N en $\mathbb{N}$, y en particular $k_j\geq j$ para todo $j\in\mathbb{N}$. Dado $\varepsilon>0$, como (x_k) es de Cauchy, existe $n\in\mathbb{N}$ tal que $d(x_k,x_\ell)\leq\varepsilon$ si $k,\ell\geq n$. Como $k_j\geq j$ para todo j, tenemos entonces que si $i,j\geq n$ entonces $k_i\geq i\geq n$ y $k_j\geq j\geq n$, luego $d(x_{k_i},x_{k_j})\leq\varepsilon$. Esto prueba que (x_{k_j}) es de Cauchy en (X,d). □

Como en el caso de sucesiones convergentes, se tiene lo siguiente.

Proposición 4.11. *Una sucesión es de Cauchy en $\mathbb{R}^n$ si y sólo si sus sucesiones coordenadas son de Cauchy en $\mathbb{R}$.*

Demostración. Queda como ejercicio para el lector, al que se recomienda inspirarse en la prueba de la Proposición 4.5. □

Utilizando la completitud de $\mathbb{R}$ y las Proposiciones 4.5 y 4.11, es fácil deducir el siguiente teorema.

Teorema 4.12. *El espacio $\mathbb{R}^n$ (con cualquiera de sus normas) es completo.*

La demostración resulta inmediata si se utiliza la norma $\|\cdot\|_\infty$. El hecho de que el resultado es cierto para todas las demás normas es entonces consecuencia del Teorema 1.3 (aunque, por supuesto, para casos concretos como las normas $\|\cdot\|_2$ o $\|\cdot\|_1$, no es necesario emplear este teorema todavía sin demostrar; basta usar el ejercicio 1.2).

Sin embargo, si a $\mathbb{R}^n$ se le dota de una estructura de espacio métrico con una distancia que *no* deriva de una norma, el teorema anterior no es cierto en general. Por ejemplo, si $X = \mathbb{R}$ con la distancia

$$d(x, y) = |\text{arctg}(x) - \text{arctg}(y)|$$

entonces la sucesión $(n)_{n\in\mathbb{N}}$ es de Cauchy en X, pero no converge a ningún punto de X. Y sin embargo esta distancia es equivalente a la usual de $\mathbb{R}$, ya que, como es fácil ver, define los mismos conjuntos abiertos. Esta observación nos indica en particular que la completitud (de hecho la definición de sucesión de Cauchy) no es un concepto topológico, es decir, no puede definirse exclusivamente en términos de conjuntos abiertos, sino que es una noción propiamente métrica, es decir que depende de la distancia y no se conserva al tomar distancias equivalentes. Sin embargo se tiene lo siguiente.

Proposición 4.13. *Sean d_1 y d_2 distancias en un conjunto X, y supongamos que son unifor-mememente equivalentes. Entonces los espacios métricos (X, d_1) y (X, d_2) tienen las mismas sucesiones de Cauchy.*

Demostración. Reescribamos la definición de sucesion de Cauchy en términos de bolas: una sucesión (x_n) es de Cauchy en un espacio métrico (X, d) si para cada $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que si $n, m \geq n_0$ entonces $x_m \in B_d(x_n, \varepsilon)$ (o, lo que es lo mismo, $x_n \in B_d(x_m, \varepsilon)$).

Supongamos que las distancias d_1 y d_2 son uniformemente equivalentes (ver la Definición 1.10). Es decir, para cada $\varepsilon > 0$ existen $\varepsilon_1, \varepsilon_2 > 0$ tales que, para todo $x \in X$, se tiene $B_{d_1}(x, \varepsilon_1) \subseteq B_{d_2}(x, \varepsilon)$, y $B_{d_2}(x, \varepsilon_2) \subseteq B_{d_1}(x, \varepsilon)$. Sea (x_n) una sucesión de Cauchy en (X, d_1). Dado $\varepsilon > 0$, sea ε_1 como antes. Al ser (x_n) de Cauchy en (X, d_1), podemos encontrar $n_0 \in \mathbb{N}$ tal que si $n, m \geq n_0$ entonces $x_m \in B_{d_1}(x_n, \varepsilon_1)$, y como $B_{d_1}(x, \varepsilon_1) \subseteq B_{d_2}(x, \varepsilon)$ obtenemos también que $x_m \in B_{d_2}(x, \varepsilon)$. Esto prueba que (x_n) es de Cauchy en (X, d_2). Análogamente se ve que cualquier sucesión de Cauchy en (X, d_2) también es de Cauchy en (X, d_1). □

Conviene observar que el recíproco de la proposición anterior no es cierto; ver el Problema 4.17.

La siguiente proposición muestra que los subconjuntos cerrados de un espacio métrico completo son a su vez completos con la distancia inducida.

Proposición 4.14. *Sea (X, d) un espacio métrico completo, y sea Y un cerrado de X. Entonces (Y, d) es completo.*

Demostración. Sea (y_n) una sucesión de Cauchy en (Y, d). Entonces (y_n) también es de Cauchy en (X, d), y como X es completo, existe $=\in X$ tal que $\lim_{n\to\infty} d(y_n, y) = 0$. Luego $y \in \overline{Y}$, y puesto que Y es cerrado en X, deducimos que $y \in Y$. Por tanto (y_n) converge a y en (Y, d). $\square$

Por otra parte, existe un recíproco fuerte de este resultado, en el sentido siguiente.

Proposición 4.15. *Si Y es un subconjunto de un espacio métrico (X, d) y el subespacio métrico (Y, d) es completo, entonces Y es cerrado en X.*

Demostración. Sea $\bar{y} \in \overline{Y}$, es decir, existe (y_n) sucesión en Y tal que $d(y_n, \bar{y}) \to 0$. Entonces (y_n) es de Cauchy en (X, d), y por tanto también en (Y, d). Como por hipótesis este espacio es completo, se tiene que existe $y \in Y$ tal que $d(y_n, y) \to 0$. Es decir, la sucesión (y_n) converge a y, y también a $\bar{y}$. Por la unicidad del límite, $\bar{y} = y$, y por tanto $\bar{y} \in Y$. Esto prueba que $\overline{Y} \subseteq Y$, es decir, que Y es cerrado en X. $\square$

El siguiente lema es a veces empleado para probar que una sucesión es convergente, y se utilizará de hecho más adelante en este mismo capítulo para probar que todo espacio métrico compacto es completo, y también en la demostración del Teorema 4.17.

Lema 4.16. *Si (x_k) es una sucesión de Cauchy en X que posee una subsucesión convergente (x_{k_j}) (digamos a un punto $x \in X$) entonces (x_k) es convergente (también a x).*

Demostración. Dado $\varepsilon > 0$, como (x_k) es de Cauchy, existe $m_\varepsilon \in \mathbb{N}$ tal que si $k, \ell \geq m_\varepsilon$ entonces $d(x_\ell, x_k) < \varepsilon/2$. Por otra parte, como (x_{k_j}) converge a x, existe $j_0 \in \mathbb{N}$ tal que si $j \geq j_0$ entonces $d(x_{k_j}, x) < \varepsilon/2$. Fijemos ahora $j \geq j_0$ suficientemente grande de modo que $k_j \geq m_\varepsilon$; esto puede hacerse porque (x_{k_j}) es subsucesión de (x_k), y en particular $\lim_{j\to\infty} k_j = \infty$. Entonces tenemos que $d(x_\ell, x) \leq d(x_\ell, x_{k_j}) + d(x_{k_j}, x) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$ para todo $\ell \geq m_\varepsilon$. $\square$

Por supuesto el lema anterior sólo tiene interés si no se sabe que el espacio X sea completo.

El siguiente teorema proporciona una caracterización bastante útil de la completitud de un espacio métrico.

Teorema 4.17 (Cantor). *Las afirmaciones siguientes son equivalentes para un espacio métrico X.*

1. *X es completo.*
2. *Para toda sucesión (C_n) de subconjuntos cerrados no vacíos de X tales que $C_{n+1} \subseteq C_n$ para todo n y con*
$$\lim_{n\to\infty} \operatorname{diam}(C_n) = 0,$$
se tiene que $\bigcap_{n=1}^{\infty} C_n \neq \emptyset$ (y de hecho esta intersección es un único punto).

Demostración. (1) $\implies$ (2): Elijamos para cada $n \in \mathbb{N}$ un punto $x_n \in C_n$. Para todo $k \geq n$ se tiene que $x_k \in C_k \subseteq C_n \ni x_n$, luego
$$d(x_n, x_k) \leq \operatorname{diam}(C_n),$$
y como $\lim_{n\to\infty} \operatorname{diam}(C_n) = 0$ se deduce inmediatamente que (x_n) es una sucesión de Cauchy. Como X es completo, existe
$$x = \lim_{n\to\infty} x_n.$$

Por otro lado, como $x_k \in C_n$ para $k \geq n$ y C_n es cerrado, se tiene que $x \in C_n$. Por tanto $x \in \bigcap_{n=1}^{\infty} C_n$. De hecho la intersección se reduce al punto x, ya que si hubiera otro punto y, el diámetro de los C_n sería mayor o igual que $d(x, y) > 0$, y no podría converger a cero.
$(2) \implies (1)$: Sea (x_n) una sucesión de Cauchy en X. Definamos

$$C_n = \overline{\{x_k : k \geq n\}}$$

para cada $n \in \mathbb{N}$. Es claro que los C_n son cerrados y $C_{n+1} \subset C_n$ para todo n. Además, dado $\varepsilon > 0$, como (x_n) es de Cauchy, debe existir $n_0 \in \mathbb{N}$ tal que si $k \geq n \geq n_0$ entonces $d(x_k, x_n) \leq \varepsilon/2$. En particular se tiene que $C_n \subseteq \overline{B}(x_{n_0}, \varepsilon/2)$, luego $\operatorname{diam}(C_n) \leq \varepsilon$ si $n \geq n_0$. Esto prueba que

$$\lim_{n\to\infty} \operatorname{diam}(C_n) = 0.$$

Por tanto existe $x \in \bigcap_{n=1}^{\infty} C_n$. Además, $d(x_n, x) \leq \operatorname{diam}(C_n) \to 0$, luego (x_n) converge a x. □

El siguiente teorema es un resultado muy importante que será utilizado, entre otras cosas, para demostrar el teorema de la función inversa en el ámbito de las funciones diferenciables. Otro tipo de aplicación importante de este resultado se presenta en el problema 4.19.

Se dice que $f : X \to X$ es una aplicación contractiva si es L-Lipschitz con $L < 1$, es decir, si existe L con $0 \leq L < 1$ tal que

$$d\big(f(x), f(y)\big) \leq L d(x, y)$$

para todo $x, y \in X$. Recordemos también que un punto fijo de una aplicación f es un punto a tal que $f(a) = a$.

Teorema 4.18 (del punto fijo para aplicaciones contractivas)**.** *Sea X un espacio métrico completo y $f : X \to X$ una aplicación contractiva. Entonces f tiene un único punto fijo a. Además, para todo $x_0 \in X$, la sucesión definida por recurrencia por $x_{n+1} = f(x_n)$ converge al punto fijo a de f.*

Demostración. Si x_0 es un punto cualquiera de X, pongamos $x_1 = f(x_0)$, y para cada $n \in \mathbb{N}$ definamos $x_{n+1} = f(x_n)$. Como f es L-Lipschitz para algún $L < 1$, tenemos que

$$d(x_{n+1}, x_{n+2}) = d\big(f(x_n), f(x_{n+1})\big) \leq L d(x_n, x_{n+1}).$$

Por iteración de esta desigualdad se deduce que también

$$d(x_n, x_{n+1}) \leq L^n d(x_0, x_1).$$

Luego, para $k > n$, tenemos que

$$d(x_n, x_k) \leq \sum_{j=n}^{k-1} d(x_j, x_{j+1}) \leq \sum_{j=n}^{k-1} L^j d(x_0, x_1) =$$
$$= \frac{L^n - L^k}{1 - L} d(x_0, x_1) \leq L^n \frac{d(x_0, x_1)}{1 - L},$$

y como $\lim_{n\to\infty} L^n = 0$ esto implica que (x_n) es de Cauchy. Como X es completo, (x_n) converge a un punto $a \in X$. Además tenemos que

$$d\big(f(x_n), f(a)\big) \leq L d(x_n, a) \to 0,$$

luego $\lim_{n\to\infty} f(x_n) = f(a)$. Por tanto,

$$a = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim f(x_n) = f(a),$$

y así resulta que a es un punto fijo de f. Por último, a es el único punto fijo de f: si hubiera otro punto fijo distinto, digamos b, entonces tendríamos

$$d(a,b) = d(f(a), f(b)) \leq Ld(a,b) < d(a,b),$$

lo que es absurdo. $\square$

Finalmente estudiaremos la noción de compacidad en un espacio métrico, y especialmente la caracterización de los compactos de $(\mathbb{R}^n, \|\cdot\|)$ como aquellos subconjuntos que son cerrados y acotados.

Definición 4.3. Se dice que un espacio métrico (X, d) es compacto si toda sucesión de X tiene una subsucesión convergente.

Se dice que un subconjunto K de (X, d) es compacto si el subespacio métrico (K, d) es compacto, lo que equivale a decir que toda sucesión $(x_n) \subseteq K$ tiene una subsucesión que converge en X a algún punto $x \in K$.

Así, por ejemplo, es inmediato ver que $\mathbb{N}$ y $\{1/n^2 : n \in \mathbb{N}\}$ *no* son compactos en $\mathbb{R}$, mientras que cualquier intervalo cerrado y acotado $[a, b]$ de $\mathbb{R}$ sí es compacto. De hecho se tiene la siguiente caracterización de los compactos de $\mathbb{R}$.

Teorema 4.19. *Un subconjunto K de $(\mathbb{R}, |\cdot|)$ es compacto si y sólo si es cerrado y acotado.*

Demostración. Supongamos que K es cerrado y acotado. Sea $(x_n) \subset K$. Por ser K acotado, esta sucesión está acotada, luego, por el teorema de Bolzano-Weierstrass para sucesiones, (x_n) tiene una subsucesión (x_{n_j}) convergente en $\mathbb{R}$, digamos, a un punto $x \in \mathbb{R}$. En particular x es un punto de $\overline{K} = K$. Por tanto (x_{n_j}) converge a $x \in K$ en K.

El recíproco es siempre verdad para un espacio métrico (es decir, no depende de la estructura particular de $\mathbb{R}$), como los dos resultados que vemos a continuación prueban. $\square$

Teorema 4.20. *Si K es un compacto de un espacio métrico (X, d) entonces (K, d) es completo, y en particular K es cerrado en X.*

Demostración. Sea (x_n) de Cauchy en K. Como K es compacto, tiene una subsucesión convergente a un punto $x \in K$. Por el Lema 4.16, (x_n) converge a x. Luego K es completo. Que K es cerrado se deduce aplicando la Proposición 4.15. $\square$

Definición 4.4. Se dice que un subconjunto A de un espacio métrico X es totalmente acotado (o precompacto, según los textos) si, para todo $\varepsilon > 0$, A puede cubrirse por una colección finita de subconjuntos de X de diámetro menor que ε. Esto es lo mismo que decir que, para cada $\varepsilon > 0$, A está contenido en una unión finita de bolas de radio menor que ε.

Obsérvese que cualquier subconjunto de un totalmente acotado es también totalmente acotado.

Ejercicio 4.2. Probar que si X es totalmente acotado entonces es acotado.

Teorema 4.21. *Si K es un compacto de X entonces K es totalmente acotado (y en particular acotado).*

Demostración. Fijemos $\varepsilon > 0$. Supongamos que K no estuviera contenido en la unión de ninguna cantidad finita de bolas de radio ε. Escojamos $x_1 \in K$, y consideremos $B(x_1, \varepsilon)$. Como esta bola no recubre K, entonces podemos escoger $x_2 \in K \setminus B(x_1, \varepsilon)$; en particular $d(x_2, x_1) \geq \varepsilon$. Por inducción se construye una sucesión $(x_n) \subseteq K$ con la propiedad de que

$$x_{n+1} \in K \setminus \bigcup_{j=1}^{n} B(x_j, \varepsilon)$$

para cada $n \in \mathbb{N}$. En particular vemos que $d(x_n, x_m) \geq \varepsilon$ para todo $n > m$. Por tanto ninguna subsucesión de (x_n) puede ser de Cauchy, y en particular (x_n) no tiene ninguna subsucesión convergente, lo que contradice que K sea compacto. $\square$

Corolario 4.22. *Todo espacio métrico compacto es separable.*

Demostración. Sea X compacto. Por el teorema anterior, X es totalmente acotado. Por tanto, para cada $k \in \mathbb{N}$, existe una colección finita de bolas $B(z_j^k, \frac{1}{n})$, $j = 1, ..., N_k$ tales que

$$X \subseteq \bigcup_{j}^{N_k} B(z_j^k, \frac{1}{k}).$$

Definamos $D_k = \{z_j^k : j = 1, ..., N_k\}$, y $D = \bigcup_{k=1}^{\infty} D_k$. Veamos que D es denso en X y habremos acabado. Sea x un punto de X y U un abierto cualquiera que contiene a x. Existe por tanto $r_x > 0$ tal que $B(x, r_x) \subset U$. Sea $k \in \mathbb{N}$ suficientemente grande para que $1/k < r_x$. Como las bolas $B(z_j^k, 1/k)$ recubren X, debe existir un j tal que $x \in B(z_j^k, 1/k)$, es decir, $d(x, z_j^k) < 1/k$, o bien $D \ni z_j^k \in B(x, 1/k) \subset U$. $\square$

A continuación veremos que los subconjuntos cerrados de un compacto heredan la propiedad de compacidad.

Proposición 4.23. *Si $C \subseteq K$, C es cerrado y K es compacto, entonces C también es compacto.*

Demostración. Sea $(x_n) \subseteq C$. Como $C \subseteq K$ y K es compacto, existe (x_{n_j}) subsucesión de (x_n) que converge a un punto $x \in K$. Luego $x \in \overline{C}$, y como C es cerrado, $x \in C$. Así (x_{n_j}) converge a x en C. $\square$

El siguiente resultado muestra que, si X es compacto, no es preciso pedir que el diámetro de los C_n tienda a 0 en el Teorema 4.17 para concluir que la intersección es no vacía (aunque en este caso, claro está, puede que contenga más de un punto).

Teorema 4.24 (Cantor)**.** *Sea X un espacio compacto, y sea (C_n) una sucesión de cerrados no vacíos encajados de X (es decir, $C_{n+1} \subseteq C_n$). Entonces*

$$\bigcap_{n=1}^{\infty} C_n \neq \emptyset.$$

Demostración. Para cada $n \in \mathbb{N}$ elijamos $x_n \in C_n$. Como X es compacto, la sucesión (x_n) tiene una subsucesión convergente, digamos $\lim_{j\to\infty} x_{n_j} = x \in X$. Para cada $m \in \mathbb{N}$ se tiene que $x_{n_j} \in C_m$ si $j \geq k_m$ para algún k_m suficientemente grande, y como C_m es cerrado se concluye que $x \in C_m$ para todo m, de donde $x \in \bigcap_{n=1}^{\infty} C_n$. $\square$

La siguiente caracterización de la compacidad es muy importante. De hecho a veces se toma (2) como definición de espacio métrico compacto.

Teorema 4.25 (Borel-Lebesgue)**.** *Sea K un subconjunto de un espacio métrico X. Las siguientes afirmaciones son equivalentes.*

(1) K es compacto.

(2) Todo recubrimiento de K por abiertos de X tiene un subrecubrimiento finito.

(3) Todo subconjunto infinito de K tiene algún punto de acumulación.

Es decir, K es compacto si y sólo si para toda familia $\mathcal{F} = \{U_\alpha\}_{\alpha\in I}$ de abiertos de X tal que

$$K \subseteq \bigcup_{\alpha\in I} U_\alpha$$

existe una subfamilia finita $\{U_{\alpha_1}, ..., U_{\alpha_n}\}$ de $\mathcal{F}$ tal que

$$K \subseteq \bigcup_{j=1}^{n} U_{\alpha_j}.$$

Demostración. $(1) \implies (2)$: Sea $\mathcal{F}$ una familia de abiertos de X que recubre K. Gracias al Teorema 3.10 sabemos que existe un subrecubrimiento numerable, es decir, podemos suponer que tenemos

$$K \subseteq \bigcup_{n=1}^{\infty} U_n,$$

donde (U_n) es una sucesión (de abiertos de X) contenida en $\mathcal{F}$. Definamos $C_n = K \setminus \bigcup_{j=1}^{n} U_j$, que es un cerrado de K para cada $n \in \mathbb{N}$. Obviamente es $C_{n+1} \subseteq C_n \subseteq K$ para todo $n \in \mathbb{N}$. Si no existiera ningún subrecubrimiento finito, entonces tendríamos que $C_n \neq \emptyset$ para todo n. Por el Teorema de Cantor 4.24 concluiríamos que

$$\bigcap_{n=1}^{\infty} C_n \neq \emptyset,$$

o lo que es lo mismo, K no estaría contenido en

$$K \setminus \bigcap_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} (K \setminus C_n) = K \cap \Big(\bigcup_{j=1}^{\infty} U_j\Big),$$

lo que es absurdo.
$(2) \implies (3)$: Sea A un subconjunto infinito de K, y supongamos que A no tiene ningún punto de acumulación en K. Escojamos $x_1 \in A$. Como $x_1 \notin A'$, existe $r_1 > 0$ tal que $B(x_1, r_1) \cap (A \setminus \{x_1\}) = \emptyset$. Como A es infinito, podemos escoger ahora $x_2 \in A$ tal que $x_2 \neq x_1$ y $x_2 \notin A'$. Por tanto existe $r_2 > 0$ tal que $B(x_2, r_2) \cap (A \setminus \{x_2\}) = \emptyset$. Continuando de este modo y usando inducción podemos construir dos sucesiones (x_n) de puntos distintos de A y (r_n) de números positivos con la propiedad de que

$$B(x_n, r_n) \cap A = \{x_n\}$$

para todo $n \in \mathbb{N}$. Además, el conjunto $C := \{x_n : n \in \mathbb{N}\}$ es cerrado (ya que $C \subseteq A$, luego $C' \subseteq A' = \emptyset$, es decir, $C' = \emptyset$, y por tanto $\overline{C} = C \cup C' = C$). Consideremos entonces el siguiente recubrimiento de K por abiertos de X:

$$K \subseteq (X \setminus C) \cup \bigcup_{n=1}^{\infty} B(x_n, r_n).$$

Este recubrimiento de K no tiene ningún subrecubrimiento finito, ya que ello implicaría la existencia de algún $n \in \mathbb{N}$ tal que $A \ni x_j \in B(x_n, r_n)$ para infinitos $j \neq n$. Esto contradice (2).
$(3) \implies (1)$: Sea (x_n) una sucesión de K. Si $C := \{x_n : n \in \mathbb{N}\}$ es un conjunto finito, entonces evidentemente (x_n) debe repetir infinitas veces uno de los elementos de C, lo que da una subsucesión constante y en particular convergente. Podemos suponer pues que C es infinito. Entonces, por (3), existe $x \in C' \cap K$. Y por la Proposición 4.8 esto significa que existe una subsucesión de (x_n) que converge a x en K. $\square$

El siguiente resultado puede verse como un refinamiento del teorema de Cantor. Se dice que una familia $\mathcal{F}$ de conjuntos tiene la propiedad de intersección finita si, siempre que se tome una cantidad finita de miembros $F_1, ..., F_n \in \mathcal{F}$, se obtiene que $\bigcap_{j=1}^n F_j \neq \emptyset$.

Teorema 4.26. *Sea K un espacio métrico. Las afirmaciones siguientes son equivalentes.*

(1) K es compacto.

(2) Para toda familia $\mathcal{F}$ de conjuntos cerrados de K con la propiedad de intersección finita, se tiene que
$$\bigcap\{F : F \in \mathcal{F}\} \neq \emptyset.$$

Demostración. $(1) \implies (2)$: Supongamos que fuera $\bigcap\{F : F \in \mathcal{F}\} = \emptyset$. Entonces $\mathcal{G} = \{K \setminus F : F \in \mathcal{F}\}$ sería una familia de abiertos de K tal que $K = \bigcup\{G : G \in \mathcal{G}\}$. Luego, por ser K compacto, y usando el teorema anterior, existen $F_1, ..., F_n \in \mathcal{F}$ tales que $K = \bigcup_{j=1}^n \big(K \setminus F_j\big) = K \setminus \bigcap_{j=1}^n F_j$. Por tanto ha de ser $\bigcap_{j=1}^n F_j = \emptyset$, lo que contradice la hipótesis en (2).
$(2) \implies (1)$: Sea $\{G : g \in \mathcal{G}\}$ un recubrimiento por abiertos de K. Definamos $\mathcal{F} = \{K \setminus G : G \in \mathcal{G}\}$. Se tiene que $\bigcap\{F : F \in \mathcal{F}\} = \emptyset$,. Luego $\mathcal{F}$ no puede tener la propiedad de intersección finita. Existen pues $G_1, ..., G_n \in \mathcal{G}$ tales que los $F_j = K \setminus G_j$ satisfacen $F_1 \cap ... \cap F_n = \emptyset$, es decir, $K = G_1 \cup ... \cup G_n$. Utilizando el teorema anterior se concluye que K es compacto. $\square$

Antes de pasar a estudiar los compactos de $\mathbb{R}^n$ vamos a establecer una última caracterización de los compactos de un espacio métrico general que precisamente nos va a proporcionar una demostración bastante simple del hecho que, en $\mathbb{R}^n$, un conjunto es compacto si y sólo si es cerrado y acotado.

Teorema 4.27. *Un espacio métrico X es compacto si y sólo si es totalmente acotado y completo.*

Demostración. Ya sabemos (ver Teoremas 4.20 y 4.21) que si X es compacto entonces es completo y totalmente acotado. Recíprocamente, supongamos que X es totalmente acotado y completo, y veamos que toda sucesión (x_n) en X tiene una subsucesión convergente. Para $\varepsilon = 1$, como X es totalmente acotado, existe una cantidad finita de subconjuntos de X de diámetro menor o igual a 1 que recubren X. Al menos uno de estos subconjuntos debe contener infinitos términos de la sucesión; llamémoslo X_1, y escojamos $n_1 \in \mathbb{N}$ tal que $x_{n_1} \in X_1$. Ahora, para $\varepsilon = 1/2$, como X_1 es totalmente acotado, puede recubrirse por una cantidad finita de subconjuntos suyos de diámetro menor o igual que $1/2$, y al menos uno de éstos contendrá infinitos términos de (x_n); llamemos X_2 a uno de éstos y escojamos $n_2 > n_1$ tal que $x_{n_2} \in X_2$. Por inducción se construye así una subsucesión (x_{n_k}) de (x_n) tal que, si $j > k$, $x_{n_j}, x_{n_k} \in X_k$, y como $\text{diam}(X_k) \leq 1/k$ también $d((x_{n_j}, x_{n_k}) \leq 1/k$ para $j > k$, lo que implica que (x_{n_k}) es de Cauchy, y por tanto convergente, gracias a la completitud de X. $\square$

Veamos ahora la anunciada caracterización de los compactos de $\mathbb{R}^n$.

Teorema 4.28. *Sea K un subconjunto de $\mathbb{R}^n$. Entonces, K es compacto si y sólo si K es cerrado y acotado.*

Demostración. Sólo hay que probar que si $K \subset \mathbb{R}^n$ es cerrado y acotado entonces es compacto. Para ellos usaremos el teorema anterior. Como $\mathbb{R}^n$ es completo y K es cerrado, K también es completo (gracias a la Proposición 4.14). Entonces basta probar que cualquier acotado de $\mathbb{R}^n$ es de hecho totalmente acotado, para lo cual a su vez es suficiente probar que cualquier cubo de la forma $C = [-M, M] \times ... \times [-M, M]$ en $\mathbb{R}^n$, con $M > 0$, está totalmente acotado. Pero esto es evidente: dado $\varepsilon > 0$, escojamos $k \in \mathbb{N}$ tal que $2\sqrt{n}M/k \leq \varepsilon$, dividamos el intervalo $[-M, M]$ en k subintervalos contiguos $I_1, ..., I_k$ de igual longitud $2M/k$, y definamos

$$C_{(j_1,...,j_n)} = I_{j_1} \times ... \times I_{j_n}$$

para cada $(j_1, ..., j_n) \in \{1, 2, ..., k\}^n$. Entonces es claro que los k^n cubos así definidos recubren C y tienen diámetro (respecto la norma euclidea) igual a $2\sqrt{n}M/k \leq \varepsilon$. □

Otra manera de probar el teorema anterior se indica en los ejercicios 4.24 y 4.25.

Concluimos con un enunciado que resume todas las propiedades que caracterizan la compacidad vistas hasta ahora.

Teorema 4.29. *Sea X un espacio métrico y K un subconjunto suyo. Las siguientes propiedades son todas equivalentes y cualquiera de ellas puede tomarse como definición de compacto.*

1. *Toda sucesión de K tiene una subsucesión convergente a un punto de K.*

2. *Todo recubrimiento de K por abiertos de X tiene un subrecubrimiento finito.*

3. *Toda familia de cerrados de K con la propiedad de intersección finita tiene intersección no vacía.*

4. *Todo subconjunto infinito de K tiene al menos un punto de acumulación en K.*

5. *K es totalmente acotado y completo.*

Además, en el caso $X = \mathbb{R}^n$, todas las propiedades equivalen a decir que K es cerrado y acotado.

4.1. Problemas

Problema 4.1. Estudiar la convergencia de las sucesiones (x_n) de $\mathbb{R}^2$ y $\mathbb{R}^3$ definidas por las siguientes expresiones, calculando el límite de las mismas cuando exista. Si no existe, estudiar si alguna subsucesión suya converge.

i) $x_n = (e^{3+5n-n^2}, \text{arctg} n)$ en $\mathbb{R}^2$;

ii) $x_n = (e^{3+5n-n^2}, n^2)$ en $\mathbb{R}^2$;

iii) $x_n = (\sin(n^2 + 7), 1/n^3)$ en $\mathbb{R}^2$;

iv) $x_n = (\sin^{3n}(n^2 + \sqrt{2}), (-1)^n, e^{-5n})$ en $\mathbb{R}^3$.

Problema 4.2. Sea (x_n) una sucesión de $\mathbb{R}^3$ tal que

$$\|x_n - x_k\| \leq \frac{1}{n} + \frac{1}{k}$$

para todo $n, k \in \mathbb{N}$. Probar que (x_n) es convergente.

Problema 4.3. Sea (x_n) una sucesión de $\mathbb{R}^3$ tal que

$$\|x_{n+1} - x_n\| \leq \frac{1}{n^2}$$

para todo $n \in \mathbb{N}$. Probar que (x_n) es convergente.

Problema 4.4. Definir cuándo una serie de vectores de $\mathbb{R}^n$ es convergente.

Problema 4.5. Diremos que una serie de vectores $\sum_{k=1}^{\infty} x_k$ de $\mathbb{R}^n$ es *absolutamente convergente* si la serie de números reales $\sum_{k=1}^{\infty} \|x_k\|$ es convergente en $\mathbb{R}$. Probar que toda serie absolutamente convergente es convergente en $\mathbb{R}^n$.

Problema 4.6. Sea (x_n) una sucesión en un espacio métrico X. Para cada $m \in \mathbb{N}$ definamos $Q_m = \{x_n : n \geq m\}$. Probar que (x_n) es de Cauchy si y sólo si $\lim_{m\to\infty} \operatorname{diam}(Q_m) = 0$.

Problema 4.7. Mostrar con un ejemplo que el teorema de Cantor no es cierto si $\operatorname{diam}(C_n)$ no converge a 0 o si X no es compacto.

Problema 4.8. Si $x_n \to x$ en un espacio normado X, demostrar que $\|x_n\| \to \|x\|$.

Problema 4.9. Más en general, demostrar que si $x_n \to x$ entonces $d(x_n, z) \to d(x, z)$.

Problema 4.10. Demostrar que el espacio $(C[0,1], \|\cdot\|_\infty)$ es completo.

Problema 4.11. Demostrar que $C[0,1]$, con la norma proveniente del producto escalar definido en el ejercicio 1.9, no es completo.

Problema 4.12 (Compleción de un espacio métrico)**.** Si tenemos un espacio métrico (X, d) que no es completo, podemos extenderlo hasta hacerlo completo como sigue. Sea Z el conjunto de todas las sucesiones de Cauchy en X. Definamos la relación $(x_n)\mathcal{R}(y_n) \iff \lim_{n\to\infty} d(x_n, y_n) = 0$. Pruébese que $\mathcal{R}$ es una relación de equivalencia en Z. Consideremos el cociente de X por esta relación de equivalencia, y llamémoslo $\hat{X}$. Si $(x_n) \in Z$, denotemos $[x_n]$ la clase de equivalencia representada por (x_n). Si $\hat{x} = [x_n], \hat{y} = [y_n] \in \hat{X}$, definamos

$$\hat{d}(\hat{x}, \hat{y}) = \lim_{n\to\infty} d(x_n, y_n).$$

Probar que $\hat{d}$ define una distancia en $\hat{X}$. A continuación demostrar que el espacio $(\hat{X}, \hat{d})$ es completo. A este espacio se le llama la compleción de X. El espacio X puede identificarse con un subconjunto de $\hat{X}$ mediante la inyección de X en $\hat{X}$ que a cada $x \in X$ le asigna la clase de equivalencia de la sucesión constante cuyos términos son todos x, es decir,

$$X \ni x \mapsto [(x, x, x, ...)] \ni \hat{X}.$$

Pruébese que esta aplicación es una inyección isométrica de X en $\hat{X}$ (ver el problema 1.21). Pruébese también que con esta identificación X es denso en $\hat{X}$.

Problema 4.13. Continuando con el problema anterior, probar que si el espacio X es ya de por sí completo, entonces esta inyección isométrica de X en $\hat{X}$ es sobreyectiva y por tanto X y $\hat{X}$ son isométricos.

Problema 4.14. Sean ahora X e Y dos espacios métricos tales que Y es completo, y supongamos que existe una inyección isométrica $f : X \to Y$ de modo que $f(X)$ es denso en Y. Demostrar que $\hat{X}$ e Y son isométricos. Esto indica que la compleción de un espacio métrico definida en el problema 4.12 es única salvo isometrías.

Problema 4.15. Utilizando el problema anterior, probar que $\hat{\mathbb{Q}} = \mathbb{R}$.

Problema 4.16. Si $A = \{(x, y) \in \mathbb{R}^2 : 0 < x^2 + y^2 < 1\}$, ¿qué es $\hat{A}$?

Problema 4.17. Sea $X = [0, \infty) \subset \mathbb{R}$, y definamos $d_1(x, y) = |x - y|$, $d_2(x, y) = |x^2 - y^2|$.

1. Probar que d_1 y d_2 definen son distancias en X.

2. Probar que los espacios métricos (X, d_1) y (X, d_2) tienen las mismas sucesiones de Cauchy.

3. Probar que las distancias d_1 y d_2 no son uniformemente equivalentes en X. *Indicación:* encontrar dos sucesiones (x_n), (y_n) en X tales que $d_1(x_n, y_n) \to 0$ pero $d_2(x_n, y_n) \not\to 0$.

Problema 4.18. Poner un ejemplo de una aplicación $f : \mathbb{R} \to \mathbb{R}$ que satisfaga $|f(x) - f(y)| \leq |x - y|$ para todo $x, y \in \mathbb{R}$ y que no tenga ningún punto fijo. Así, el Teorema del punto fijo 4.18 no es cierto para aplicaciones L-Lipschitz con $L \geq 1$.

Problema 4.19. En muchos problemas derivados de la física y la biología se consideran ecuaciones integrales de la forma

$$f(x) = A(x) + \int_0^x k(x, y) f(y) dy, \qquad (*)$$

donde A y k son funciones dadas, y se pretende encontrar una función f que satisfaga este ecuación. Este tipo de problemas tiene que ver también con las ecuaciones diferenciales; por ejemplo, $f(x) = ae^x$ es solución de $df/dx = f(x)$, que puede reescribirse en la forma

$$f(x) = a + \int_0^x f(y) dy.$$

Usar el Teorema 4.18, aplicado al espacio $X = (C[0, 1], \|\cdot\|_\infty)$, para probar que si $A \in C[0, 1]$ y $k : [0, 1] \times [0, 1] \to \mathbb{R}$ es una función continua (véase el capítulo siguiente) y tal que

$$\sup\{\int_0^x |k(x, y)| dy : x \in [0, 1]\} < 1,$$

entonces la ecuación integral $(*)$ tiene una única solución para $f \in C[0, 1]$. Hacer un estudio análogo de la ecuación integral

$$f(x) = A(x) + \int_0^1 k(x, y) f(y) dy. \qquad (**)$$

Problema 4.20. Utilizando el teorema del punto fijo para aplicaciones contractivas, demostrar que el sistema de ecuaciones

$$\begin{cases} x = \frac{1}{4}\sin(x+y) + \frac{1}{3}, \\ y = -\frac{1}{4}\sin(x-y) + \frac{1}{5} \end{cases}$$

tiene solución única para $|x| \leq 1$, $|y| \leq 1$.

Problema 4.21. Demostrar que la esfera de $\mathbb{R}^n$ (para cualquier norma) es un compacto.

Problema 4.22. Demostrar que existe una sucesión (n_k) de números naturales tales que existen los límites $\lim_{k\to\infty}\cos(n_k)$, $\lim_{k\to\infty}\sin(n_k^2)$ y $\lim_{k\to\infty}\cos(n_k^3)$.

Problema 4.23. Determinar cuáles de los siguientes conjuntos son compactos.

$$A = \{(x,y) \in \mathbb{R}^2 : |x| \leq 1\}$$

$$B = \{(x,y) \in \mathbb{R}^2 : x^4 + y^6 \leq 5\}$$

$$C = \{(x,y)\} \in \mathbb{R}^2 : x \in \mathbb{Q}, x^4 + y^6 \leq 5\}$$

$$D = \{(x,y,z) \in \mathbb{R}^3 : z \geq x^2 + y^2\}$$

$$E = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^4 + z^6 \leq 3\}.$$

Problema 4.24. Demostrar como sigue que cualquier cerrado acotado K de $\mathbb{R}^2$ es compacto. Dada $\{(x_k, y_k)\}_{k\in\mathbb{N}} \subset K$, como (x_k) está acotada (¿por qué?) en $\mathbb{R}$, puede aplicarse el teorema de Bolzano-Weierstrass para obtener una subsucesión (x_{k_j}) convergente a un punto $x \in \mathbb{R}$. Aplicar ahora otra vez este teorema a la sucesión $(y_{k_j}) \subset \mathbb{R}$ para obtener una subsucesión que $(y_{k_{j_l}})$ que converge a un punto $y \in \mathbb{R}$. Entonces $(x_{k_{j_l}}, y_{k_{j_l}})$ converge a (x,y). Para terminar la demostración, probar que $(x,y) \in K$.

Problema 4.25. Usando inducción sobre n, generalizar el argumento del problema anterior para probar que un subconjunto de $\mathbb{R}^n$ es compacto si y sólo si es cerrado y acotado.

Problema 4.26. Si (x_n) es una sucesión que converge a x en un espacio métrico X, demostrar que el conjunto $K := \{x_n : n \in \mathbb{N}\} \cup \{x\}$ es compacto.

Problema 4.27 (Conjunto de Cantor)**.** Sean $F_1 = [0,1]$, $F_2 = [0,1/3] \cup [2/3,1]$, $F_3 = [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1]$, y en general sea F_{n+1} el cerrado que se obtiene al quitar de cada uno de los intervalos que componen F_n el tercio medio. Definamos $C = \bigcap_{n=1}^{\infty} F_n$, el conjunto de Cantor. Demostrar que:

1. C es compacto y no vacío;
2. C tiene interior vacío;
3. C no tiene puntos aislados;
4. C es infinito no numerable.

Problema 4.28. Probar que todo subconjunto finito de un espacio métrico es compacto.

Problema 4.29. Probar que la frontera de un conjunto acotado en $\mathbb{R}^n$ es siempre compacta.

Problema 4.30. Probar que si K es un compacto de X entonces K tiene a lo sumo una cantidad finita de puntos aislados.

Capítulo 5

Límites, continuidad y continuidad uniforme

A lo largo de este capítulo, X volverá a denotar un espacio métrico (por ejemplo un subconjunto cualquiera del espacio $\mathbb{R}^n$ con la distancia inducida por una norma). Comenzaremos con las definiciones de límite de una función en un punto y de continuidad en un punto, las propiedades más importantes de estas nociones y los métodos más comunes para determinar su existencia y calcularlos en su caso. Seguiremos con los teoremas más importantes que enlazan los conceptos de continuidad global y compacidad, como el de que las funciones continuas alcanzan siempre máximos y mínimos absolutos sobre los conjuntos compactos. Realizaremos un estudio análogo respecto de la continuidad uniforme, aplazando los resultados que ligan conexión y continuidad hasta el capítulo siguiente. Finalmente demostraremos que todas las normas en $\mathbb{R}^n$ son equivalentes, y estudiaremos la definición y propiedades de la norma de una aplicación lineal.

Definición 5.1. Sean X e Y espacios métricos, sea a un punto de acumulación de un subconjunto A de X, y sea $f : A \to Y$ una aplicación. Diremos que

$$\lim_{x\to a} f(x) = b$$

si para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que si $x \in A$ y $0 < d(x, a) < \delta$ entonces $d(f(x), b) < \varepsilon$.

Equivalentemente, $\lim_{x\to a} f(x) = b$ si y sólo si para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que si $x \in A \cap B(x, \delta) \setminus \{a\}$ entonces $f(x) \in B(b, \varepsilon)$.

Por ejemplo, dada $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $f(x, y) = xy \sin \frac{1}{x^2+y^2}$ si $(x, y) \neq (0, 0)$, y de cualquier otra manera para $(x, y) = (0, 0)$, es inmediato comprobar que

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

El siguiente resultado caracteriza la existencia del límite de una función en en un punto mediante límites de sucesiones.

Proposición 5.1. *Sean X e Y espacios métricos, a un punto de acumulación de un subconjunto A de X, y $f : A \to Y$ una aplicación. Entonces existe*

$$\lim_{x\to a} f(x) = b$$

si y sólo si, para cada sucesión $(x_n) \subseteq A \setminus \{a\}$ convergente al punto a en X, la sucesión $(f(x_n))$ converge a b en Y.

Demostración. Supongamos primero que existe $\lim_{x\to a} f(x) =: b$, y sea (x_n) una sucesión contenida en $A \setminus \{a\}$ tal que $\lim_{n\to\infty} x_n = a$. Dado $\varepsilon > 0$, puesto que $\lim_{x\to a} f(x) = b$, existe $\delta > 0$ tal que si $x \in A \cap B(a,\delta) \setminus \{a\}$ entonces $f(x) \in B(b,\varepsilon)$. Por otro lado, como $\lim_{n\to\infty} x_n = a$ y $(x_n) \subseteq A \setminus \{a\}$, existe $n_0 \in \mathbb{N}$ tal que si $n \geq n_0$ entonces $x_n \in A \cap B(a,\delta) \setminus \{a\}$. Luego, si $n \geq n_0$ también se tendrá $f(x_n) \in B(b,\varepsilon)$. Esto prueba que $\lim_{n\to\infty} f(x_n) = b$.

Recíprocamente, supongamos que no se tiene que $\lim_{x\to\infty} f(x) = b$; entonces existe $\varepsilon_0 > 0$ tal que para todo $\delta > 0$ existe $x_\delta \in A \cap B(a,\delta) \setminus \{a\}$ tal que $d(f(x_\delta), b) \geq \varepsilon_0$. Tomando δ de la forma $\delta_n = 1/n$ para cada $n \in \mathbb{N}$, obtenemos una sucesión $(x_n) \subseteq A \setminus \{a\}$ para la que se tiene $d(x_n, a) < 1/n$ y $d(f(x_n), b) \geq \varepsilon_0$ para todo $n \in \mathbb{N}$. Esto implica que $(x_n) \subseteq A \setminus \{a\}$, $\lim_{n\to\infty} x_n = a$, y sin embargo no se cumple que $\lim_{n\to\infty} f(x_n) = b$. $\square$

Esta Proposición resulta especialmente útil a la hora de probar que un límite no existe: bastará encontrar dos sucesiones convergentes al mismo punto de modo que la función, a lo largo de estas sucesiones, converge a puntos diferentes (o no converge). Por ejemplo, el límite

$$\lim_{x\to 0} \sin \frac{1}{x^2}$$

no existe, ya que si tomamos $f(x) = \sin \frac{1}{x^2}$, $(x_n) = (1/\sqrt{\pi n}))$, $(y_n) = (1/\sqrt{\pi/2 + 2\pi n}))$, se tiene que ambas sucesiones convergen a 0, pero $f(x_n) \to 0$ mientras que $f(y_n) \to 1$.

Cuando $X = \mathbb{R}^n$ o $Y = \mathbb{R}$, podemos añadir las nociones de límites infinitos y límites en el infinito.

Definición 5.2. Sea $f : A \subseteq X \to \mathbb{R}$ una función, $a \in A'$. Se dice que $\lim_{x\to a} f(x) = +\infty$ si para todo $M > 0$ existe $\delta > 0$ tal que, si $x \in A$ y $0 < d(x,a) < \delta$, entonces $f(x) > M$.

Si A es un subconjunto *no acotado* de $\mathbb{R}^n$, y $f : A \to \mathbb{R}$, diremos que $\lim_{x\to\infty} f(x) = b \in \mathbb{R}$ si para cada $\varepsilon > 0$ existe $N > 0$ tal que, si $x \in A$ y $\|x\| > N$, entonces $|f(x) - b| < \varepsilon$.

Finalmente, diremos que $\lim_{x\to\infty} f(x) = +\infty$ si para todo $M > 0$ existe $N > 0$ tal que si $x \in A$ y $\|x\| > N$ entonces $f(x) > M$.

Análogamente pueden definirse $\lim_{x\to a} f(x) = -\infty$ y $\lim_{x\to\infty} f(x) = -\infty$.

Ejercicio 5.1. Probar que:

$$\lim_{(x,y)\to(0,0)} \frac{1}{x^2 + |y|} = +\infty, \quad \lim_{(x,y)\to\infty} \frac{1}{x^2 + |y|} = 0, \quad \lim_{(x,y)\to\infty} x^2 - x + y^4 = \infty.$$

Ejercicio 5.2. Siguiendo el modelo de la Proposición 5.1, establecer caracterizaciones, mediante sucesiones, de los conceptos de límites infinitos y límites en el infinito. Por ejemplo, probar que para una función $f : A \subseteq X \to \mathbb{R}$ y un punto $a \in A'$ se tiene que $\lim_{x\to a} f(x) = +\infty$ si y sólo si para toda sucesión $(x_k) \subseteq A \setminus \{a\}$ que converja a a se cumple que $\lim_{k\to\infty} f(x_k) = +\infty$.

Paralelamente a la teoría de límites abordaremos el concepto de continuidad de una función en un punto.

Definición 5.3. Sean X, Y espacios métrico, $A \subseteq X$, $f : A \to Y$ una aplicación. Se dice que x es continua en $a \in A$ si, o bien a es un punto aislado de A, o bien

$$\lim_{x\to a} f(x) = f(a)$$

cuando a es punto de acumulación de A. Esto equivale a decir que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que si $d(x,a) < \delta$ entonces $d(f(x), f(a)) < \varepsilon$.

Por ejemplo, es inmediato ver que la función $f(x, y) = x^2 y$ es continua en $(0, 0)$. Obsérvese también que toda función definida sobre un conjunto de puntos aislados es, por definición, continua.

A continuación estudiamos la estabilidad de sumas, productos y cocientes respecto de la operación de paso al límite.

Proposición 5.2. *Sean $f, g : A \subseteq X \to \mathbb{R}$ dos funciones, $a \in A'$. Supongamos que existen los límites*

$$\lim_{x \to a} f(x) = l \text{ y } \lim_{x \to a} g(x) = m.$$

Entonces:

1. *Existe* $\lim_{x \to a} (f + g)(x) = l + m$*;*
2. *Existe* $\lim_{x \to a} (f \cdot g)(x) = l \cdot m$*;*
3. *Si* $m \neq 0$*, existe también* $\lim_{x \to a} \frac{f}{g}(x) = \frac{l}{m}$.

Demostración. Puede usarse la carecterización del límite por sucesiones, en combinación con los resultados análogos bien conocidos para sucesiones de $\mathbb{R}$, para deducir este resultado de manera completamente trivial. Sin embargo, con el fin de familiarizarnos con la importante noción de límite, daremos una demostración autocontenida que requiere un esfuerzo algo mayor.

(1): Dado $\varepsilon > 0$, puesto que $\lim_{x \to a} f(x) = l$, existe $\delta_1 > 0$ tal que si $x \in A \cap B(a, \delta_1) \setminus \{a\}$ entonces $|f(x) - l| < \varepsilon/2$. Análocamente, como $\lim_{x \to a} f(x) = l$, existe $\delta_2 > 0$ tal que si $x \in A \cap B(a, \delta_2) \setminus \{a\}$ entonces $|g(x) - m| < \varepsilon/2$. Tomemos $\delta := \min\{\delta_1, \delta_2\}$. Entonces, para todo $x \in A \cap B(a, \delta) \setminus \{a\}$ se tiene que

$$|f(x) + g(x) - (l + m)| \leq f(x) - l| + |g(x) - m| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

(2). Dado $\varepsilon > 0$, como $\lim_{x \to a} f(x) = l$, existe δ_1 tal que si $x \in A \cap B(a, \delta_1) \setminus \{a\}$ entonces se tiene

$$|f(x) - l| < \min\left\{1, \frac{\varepsilon}{2\,(|m| + 1)}\right\},$$

lo que en particular implica que

$$|f(x)| \leq |l| + 1.$$

Por otra parte, como $\lim_{x \to a} f(x) = l$, existe δ_2 tal que si $x \in A \cap B(a, \delta_2) \setminus \{a\}$ entonces se tiene

$$|g(x) - m| < \frac{\varepsilon}{2\,(|l| + 1)}.$$

Por tanto, definiendo $\delta := \min\{\delta_1, \delta_2\}$ tenemos que, si $x \in A \cap B(a, \delta) \setminus \{a\}$ entonces

$$\begin{aligned} &|f(x)g(x) - lm| = |f(x)g(x) - f(x)m + f(x)m - lm| \\ &\leq |f(x)|\,|g(x) - m| + |m|\,|f(x) - l| \\ &< (|l| + 1)\,\frac{\varepsilon}{2\,(|l| + 1)} + |m| \frac{\varepsilon}{2\,(|m| + 1)} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

(3): En vista de (2), basta probar (3) en el caso particular de la función constante $f = 1$. Dado $\varepsilon > 0$, como $\lim_{x \to a} g(x) = m \neq 0$, existe $\delta > 0$ tal que

$$|g(x) - m| < \min\left\{\frac{|m|}{2}, \frac{\varepsilon |m|^2}{2}\right\},$$

lo que implica que $|g(x)| \geq |m|/2$. Entonces, si $x \in A \cap B(a,\delta) \setminus \{a\}$ se tiene que

$$\left|\frac{1}{g(x)} - \frac{1}{m}\right| = \left|\frac{m - g(x)}{mg(x)}\right| \leq \frac{2}{|m|^2}\,|m - g(x)| < \frac{2}{|m|^2}\frac{\varepsilon|m|^2}{2} = \varepsilon.$$

□

Como consecuencia inmediata puede deducirse un enunciado análogo para la continuidad en un punto, es decir: la suma, producto y cociente (cuando el límite de la función denominador no es cero) de funciones continuas en un punto son continuas en ese punto.

Observación 5.3. La Proposición 5.2 puede extenderse al caso de límites infinitos o en el infinito para funciones $f, g : A \subseteq \mathbb{R}^n \to \mathbb{R}$, es decir, el mismo enunciado sigue siendo cierto si ponemos $a = \infty$ o $l, m \in [-\infty, \infty]$, con las siguientes convenciones:

1. $+\infty + \infty = +\infty$,
2. $-\infty + (-\infty) = -\infty$,
3. $r + \infty = +\infty$ para todo $r \in \mathbb{R}$,
4. $r - \infty = -\infty$ para todo $r \in \mathbb{R}$,
5. $r/+\infty = 0 = r/-\infty$ para todo $r \in \mathbb{R}$,
6. $r\infty = \text{signo}(r)\infty$ para todo $r \in \mathbb{R}$ con $r \neq 0$,

quedando el resultado indeterminado en los demás casos. También pueden demostrarse otros enunciados más generales en esta línea, como que si f está acotada y $\lim_{x\to a} g(x) = \infty$ entonces $\lim_{x\to a} f(x)/g(x) = 0$, o que si $\lim_{x\to a} g(x) = 0$, $g(x) > 0$ y $f(x) > \alpha$ para todo x y cierto $\alpha > 0$ entonces $\lim_{x\to a} f(x)/g(x) = +\infty$.

Además, la misma demostración de las partes (1) y (2) de la Proposición 5.2 sirve en el caso de que f y g tomen valores en $\mathbb{R}^d$, cambiando producto por producto escalar. El lector queda invitado a comprobar la veracidad de todas estas afirmaciones.

Respecto a la composición, tenemos lo siguiente.

Proposición 5.4. *Sean $f : X \to Y$, $g : Y \to Z$ aplicaciones entre espacios métricos. Supongamos que f es continua en $a \in X$ y que g es continua en $f(a) \in Y$. Entonces la composición $g \circ f : X \to Z$ es continua en a.*

Demostración: Fijemos $\varepsilon > 0$. Por ser g continua en $f(a)$, existe $\delta' > 0$ tal que si $d(y, f(a)) < \delta'$ entonces $d(g(y), g(f(a))) < \varepsilon$. Ahora, por ser f continua en a, existe a su vez $\delta > 0$ tal que si $d(x,a) < \delta$ entonces $d(f(x), f(a)) < \delta'$. Por tanto, para todo $x \in X$ con $d(x,a) < \delta$ tendremos que $d(g(f(x)), g(f(a))) < \varepsilon$. □

Corolario 5.5. *Sean $f : A \subseteq X \to Y$, $g : Y \to Z$, $a \in A'$. Supongamos que $\lim_{x\to a} f(x) = b$, y que g es continua en $b \in Y$. Entonces, existe*

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

Demostración. Basta aplicar la Proposición anterior a las funciones g y $\tilde{f} : A \cup \{a\} \to Y$ definida por
$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in A, x \neq a, \\ b & \text{si } x = a. \end{cases}$$
□

Así es inmediato calcular por ejemplo
$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^2+y^7)}{x^2+y^7} = 1.$$

Sin embargo, cuando la función g no es continua en el valor del límite de f, este corolario no es cierto en general. Consideremos por ejemplo el caso de $g(x) = 1$ si $x \neq 0$, $g(0) = 0$, $f(x) = x$. Entonces $\lim_{x\to 0} g(f(x)) = 1 \neq 0 = g(0) = g(\lim_{x\to 0} f(x))$. En este caso el problema está en que, aunque g sí tiene límite en 0, su valor no coincide con el que toma g en 0. No obstante esta clase de dificultad tiene fácil arreglo: podríamos redefinir g en 0 como el valor del límite de g en 0, y así estaríamos en condiciones de aplicar el resultado a la nueva función. Este arreglo es el que proporciona la demostración del siguiente Corolario.

Corolario 5.6. *Sean $f : A \subseteq X \to Y$, $g : Y \to Z$, $a \in A'$. Supongamos que existe $\lim_{x\to a} f(x) = b$, que $f(x) \neq b$ para todo $x \neq a$ en algún entorno de a, y que existe $\lim_{y\to b} g(y) = c$. Entonces, existe*
$$\lim_{x\to a} g(f(x)) = c.$$

Este resultado puede formularse también para límites infinitos o en el infinito; ver el ejercicio 5.4.

Como en el caso de sucesiones, cuando tenemos una función $f : X \to \mathbb{R}^m$, para calcular el límite de f en un punto es suficiente hallar los límites de las m funciones coordenadas asociadas a f. Si X es un espacio métrico y $f : A \subseteq X \to \mathbb{R}^m$, para cada $x \in A$ el vector $f(x) \in \mathbb{R}^m$ lo denotaremos
$$f(x) = (f_1(x), ..., f_m(x)),$$
y a las m funciones $f_j : A \to \mathbb{R}$ definidas por $x \mapsto f_j(x)$ las llamaremos funciones coordenadas de f. Con esta notación tenemos lo siguiente.

Proposición 5.7. *Sean $f : A \subseteq X \to \mathbb{R}^m$, $a \in A'$. Entonces, existe $\lim_{x\to a} f(x)$ si y sólo si existen los límites $\lim_{x\to a} f_j(x)$ para cada función coordenada f_j, $j = 1, ..., m$. Además, en este caso,*
$$\lim_{x\to a} f(x) = \big(\lim_{x\to a} f_1(x), ..., \lim_{x\to a} f_m(x)\big).$$

Demostración. Queda a cargo del lector (por ejemplo, puede usarse la Proposición 5.1 en combinación con la Proposición 4.5 para hacer la tarea prácticamente trivial. □

Veamos a continuación unos cuantos ejemplos de técnicas para determinar la existencia del límite de una función, y calcularlo. Gracias a la Proposición anterior, en la inmensa mayoría de los ejemplos y ejercicios de este curso podremos suponer que la función toma valores en $\mathbb{R}$. De hecho, en lo que sigue y hasta nuevo aviso, asumiremos que A es un subconjunto de $\mathbb{R}^n$, y $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, con $a \in A'$. Supongamos que deseamos saber si existe el límite
$$\lim_{x\to a} f(x),$$

y en caso afirmativo cuál es su valor. Si la respuesta no es evidente a primera vista, he aquí una serie de pistas que en muchos casos podrán ayudarnos a encontrar la respuesta.

1. Lo primero que el alumno debe preguntarse es si la función f es continua en a, hecho que podría ser conocido o inmediato. En caso afirmativo obviamente tendremos que existe $\lim_{x\to a} = f(a)$. Por ejemplo, esto es lo que sucede con

$$\lim_{(x,y)\to(3,0)} e^{x^2-\sin^2 y} = e^9,$$

ya que la función es continua al estar definida mediante suma y composición de funciones continuas en todos los puntos.

2. Si la función no es continua, el siguiente punto a examinar será si puede aplicarse la Proposición 5.2 o los Corolarios 5.5 o 5.6, junto con hechos ya conocidos, para asegurar la existencia de límite y calcularlo. Por ejemplo, puesto que $\lim_{(x,y)\to(0,0)} x^2+y^4 = 0$ y que $\lim_{(x,y)\to(0,0)} \cos(x+y) = 1$, se tiene que

$$\lim_{(x,y)\to(0,0)} \frac{\cos(x+y)}{x^2+y^4} = +\infty.$$

También, puesto que $\lim_{t\to\infty} \frac{\log t}{t} = 0$, se tiene que

$$\lim_{(x,y)\to\infty} \frac{\log(|x|+y^2)}{|x|+y^2} = 0.$$

3. Puede ocurrir que la Proposición 5.2 no sea directamente aplicable porque nos conduzca a una indeterminación, pero en algunos casos una sencilla manipulación algebraica puede eliminar este problema. Consideremos por ejemplo

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

Los límites del numerador y del denominador som ambos cero, por lo que no podemos aplicar directamente este resultado. Sin embargo se tiene

$$\frac{4x^2-y^2}{2x+y} = \frac{(2x-y)(2x+y)}{2x+y} = 2x-y,$$

luego el límite es claramente cero. Otros ejemplos de límites que pueden resolverse de este modo son

$$\lim_{(x,y)\to(0,0)} \frac{x^4y^2+x^2y^4}{x^2+y^2}, \quad \lim_{(x,y)\to\infty} \frac{\sqrt{|x|+|y|}-\sqrt{|x|-|y|}}{|y|+1}.$$

4. Las técnicas de acotación son también muy importantes en estos problemas. Consideremos por ejemplo la función

$$f(x,y) = \begin{cases} x+y & \text{si } x>y, \\ x^2-y^2 & \text{si } x\le y. \end{cases}$$

Entonces es claro que, para (x,y) con $|x|<1, |y|<1$, se tiene $|f(x,y)| \le |x|+|y|$, y como $\lim_{(x,y)\to(0,0)} |x|+|y| = 0$, se deduce que

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

Aquí estamos haciendo uso del *teorema del bocadillo*, es decir, si una función está entre otras dos que tienen el mismo límite en un punto, entonces la función tiene el mismo límite. Se invita al alumno a demostrar este hecho en el problema 5.2. También hemos usado el carácter local de la operación de paso al límite, esto es, si existe un entorno U de a tal que $f(x) = g(x)$ para todo $x \in U$, entonces $\lim_{x\to a} f(x) = \lim_{x\to a} g(x)$ (véase el ejercicio 5.3).

Otros ejemplos de límites que pueden calcularse mediante acotaciones muy sencillas son

$$\lim_{(x,y)\to\infty} \frac{\sin(x^2+y^3)}{x^2+y^2} = 0, \qquad \lim_{(x,y)\to(0,0)} (x^2+y)\cos(x+y^2) = 0.$$

5. Si la combinación de los métodos anteriores no ha dado resultado satisfactorio, quizás sea el momento de comenzar a sospechar que el límite no existe. Si este es verdaderamente el caso también hay una serie de pistas que nos pueden ayudar a demostrar que el límite no existe. Una primera posibilidad, para funciones de $\mathbb{R}^2$ o $\mathbb{R}^3$ en $\mathbb{R}$, es la de calcular los límites iterados y ver si son iguales (caso de que existan). Si existen los límites iterados y son distintos, es decir,

$$\lim_{x\to x_0}\Big(\lim_{y\to y_0} f(x,y)\Big) \neq \lim_{y\to y_0}\Big(\lim_{x\to x_0} f(x,y)\Big),$$

entonces no puede existir $\lim_{(x,y)\to(x_0,y_0)} f(x,y)$ (ver el problema 5.7). Por ejemplo, el límite

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

no existe, ya que los límites iterados son 1 y -1.

6. También puede ocurrir que los límites iterados existan y sean todos iguales sin que ello implique la existencia del límite, así que el criterio anterior no es ni mucho menos concluyente. En efecto, si

$$f(x,y) = \frac{|x|+|y|}{\sqrt{x^2+y^2}},$$

los límites iterados existen y son ambos iguales a uno, y sin embargo no existe $\lim_{(x,y)\to(0,0)} f(x,y)$. Supongamos que existiera, y llamémoslo L. Si nos acercamos al origen por rectas de pendiente $\lambda \in \mathbb{R}$, esto es, si hacemos $y = \lambda x$ y sustituimos en la función, la expresión $f(x,\lambda x)$ debería tender al límite L cuando $x \to 0$. Sin embargo se tiene

$$\lim_{x\to 0} f(x,\lambda x) = \frac{1+|\lambda|}{\sqrt{1+\lambda^2}},$$

que claramente es distinto para cada λ. Luego no puede existir el límite. Otro ejemplo en el que este criterio es aplicable es

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

en este caso el límite no existe ya que al aproximarnos al origen por la recta $x = 0$ el límite es cero, mientras que a lo largo de la recta $x = y$ da uno. Para la justificación teórica de este tipo de criterios véase el problema 5.8.

7. Otra manera de ver que no existe el límite anterior es hacer un cambio a coordenadas polares, es decir, poner $x = \rho\cos\theta$, $y = \rho\sin\theta$, y hacer tender ρ a cero. Si el resultado depende de θ entonces el límite no puede existir. De hecho puede demostrarse (ver el ejercicio 5.5) que para toda funtión $f : \mathbb{R}^2 \to \mathbb{R}$, se tiene que $\lim_{(x,y)\to(0,0)} f(x,y) = L$ si y solo si $\lim_{\rho\to 0} f(\rho\cos\theta,\rho\sin\theta) = L$ uniformemente en θ (es decir, si y solo si para todo $\varepsilon > 0$ existe

$\delta > 0$ tal que si $0 < \rho < \delta$ entonces $|f(\rho\cos\theta, \rho\sin\theta) - L| < \varepsilon$ para todo $\theta \in [0, 2\pi]$). En el caso de la función

$$f(x,y) = \frac{|x| + |y|}{\sqrt{x^2 + y^2}}$$

tenemos que $\lim_{\rho\to 0} f(\rho\cos\theta, \rho\sin\theta) = |\cos\theta| + |\sin\theta|$, que depende claramente de θ, luego el límite $\lim_{(x,y)\to(0,0)} f(x,y)$ no existe. También puede aplicarse este criterio, por ejemplo, para ver que no existe

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

En este caso existe el límite a lo largo de cualquier recta $y = \lambda x$, pero no existe el límite. En efecto, sea (ρ_n) una sucesión de números positivos con $\lim_{n\to\infty} \rho_n = 0$. Se tiene que $f(\rho\cos\theta, \rho\sin\theta) = \tan\theta\sin\rho^2$, y como $\lim_{\theta\to\pi/2} \tan\theta = +\infty$, para cada $n \in \mathbb{N}$ podemos encontrar θ_n próximo a $\pi/2$ de modo que

$$\tan\theta_n \geq \frac{1}{\sin^2\rho_n^2};$$

entonces es evidente que

$$f(\rho_n\cos\theta_n, \rho_n\sin\theta_n) \geq \frac{1}{\sin\rho_n^2} \to \infty.$$

Por otro lado, para $\theta = 0$, es claro que $\lim_{\rho\to 0} \tan\theta\sin\rho^2 = 0$. Por consiguiente no puede existir $\lim_{\rho\to 0} \tan\theta\sin\rho^2$ uniformemente en θ.

8. A propósito de límites iterados, conviene advertir que puede darse el caso de que no existan, o no exista uno de ellos, y sin embargo sí que exista el límite global (por tanto el criterio del ejercicio 5.7 solo sirve cuando los tres límites existen). Consideremos por ejemplo la función

$$f(x,y) = \begin{cases} y & \text{si } x \in \mathbb{Q}, \\ 0 & \text{si } x \in \mathbb{R}\setminus\mathbb{Q}. \end{cases}$$

Se tiene que no existe $\lim_{x\to 0} f(x,y)$ para ningún $y \neq 0$, luego tampoco existe el límite $\lim_{y\to 0}(\lim_{x\to 0} f(x,y))$. Sin embargo el otro límite iterado sí existe, y también existe el límite global

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

ya que $|f(x,y)| \leq |y| \to 0$ cuando $(x,y) \to (0,0)$.

9. A veces el criterio de calcular límites por rectas que pasan por el punto donde se toma el límite no es concluyente. En efecto puede ocurrir que todos los límites a lo largo de rectas existan y sean iguales, y que sin embargo el límite no exista. En muchos de estos casos el aproximarnos por curvas continuas más generales puede dar buen resultado para ver que no existe el límite. Como ejemplo consideremos la función

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

Es fácil ver que los límites iterados existen y son ambos cero, y también existen los límites a lo largo de rectas y son todos iguales: $\lim_{x\to 0} f(x, \lambda x) = 0$ para cualquier λ. Sin embargo el límite de f en $(0,0)$ no puede existir, ya que si consideramos las curvas continuas $\alpha, \beta : \mathbb{R} \to \mathbb{R}^2$ definidas por

$$\alpha(t) = (t^2, t), \quad \text{y} \quad \beta(t) = (0, t)$$

(cuyas ecuaciones implícitas son $x = y^2$ y $x = 0$ respectivamente) entonces tenemos que $\lim_{t\to 0} \alpha(t) = (0,0) = \lim_{t\to 0} \beta(t)$, y sin embargo

$$\lim_{t\to 0} f(\alpha(t)) = 1/2 \neq 0 = \lim_{t\to 0} f(\beta(t)).$$

Ver el problema 5.8 para la justificación teórica del criterio que estamos empleando.

10. Finalmente, un criterio que nunca falla a la hora de demostrar que un límite no existe, y que suele dar resultados más rápidos y por lo menos igual de efectivos que todos los anteriores (al menos si se tienen las ideas claras sobre la razón de por qué no existe el límite), es el de las sucesiones. En virtud de la Proposicion 5.1 basta encontrar dos sucesiones que converjan al punto donde se toma el limite y a lo largo de las cuales la funcion converge a puntos diferentes (o no converge). Para ilustrar este criterio, sea $f : \mathbb{R}^2 \to \mathbb{R}$ la función definida por

$$f(x,y) = \begin{cases} 1 & \text{si } x - y \in \mathbb{Q}, \\ 0 & \text{en caso contrario.} \end{cases}$$

Entonces es claro que $\lim_{n\to\infty} f(1/n, 0) = 1 \neq 0 = \lim_{n\to\infty} f(\sqrt{2}/n, 0)$ y por tanto no existe $\lim_{(x,y)\to(0,0)} f(x,y)$.

Pasamos ahora a estudiar el comportamiento global de las funciones que son continuas en todos los puntos. En lo que sigue X e Y volverán a denotar espacios métricos arbitrarios.

Definición 5.4. Se dice que una función $f : X \to Y$ es continua en un subconjunto A de X si f es continua en x para cada $x \in A$. Si $f : X \to Y$ es continua en todo X, diremos simplemente que f es continua.

Por ejemplo, usando los resultados ya vistos sobre límites y continuidad en un punto, es inmediato ver que las funciones $S, P : \mathbb{R}^2 \to \mathbb{R}$ definidas por $S(x,y) = x + y$, $P(x,y) = xy$ son continuas en todo $\mathbb{R}^2$, y que la función $Q(x,y) = x/y$ es continua en $\mathbb{R}^2 \setminus \{y = 0\}$. También que la composición de funciones continuas es continua, y que una función $f = (f_1, ..., f_m) : X \to \mathbb{R}^m$ es continua si y sólo si lo son las funciones f_i para todo $i = 1, ..., n$. Se deduce entonces el siguiente:

Corolario 5.8. *La suma, el producto, el cociente y la composición de funciones continuas son continuas (siempre que estén bien definidas).*

Es evidente que si $f : X \to Y$ es continua en $A \subseteq X$ entonces $f_{|A} : A \to Y$ es también continua. El recíproco no es cierto en general, es decir, puede ocurrir perfectamente que $f_{|A}$ sea continua sin que ello implique que f es continua en A (piénsese en el caso en que $A = \{a\}$ sea un punto y f discontinua en a). Sin embargo sí es cierto cuando A es abierto.

Proposición 5.9. *Sean $f : X \to Y$ una aplicación, A un subconjunto abierto de X, y supongamos que $f_{|A} : A \to Y$ es continua. Entonces f es continua en A.*

Demostración. La demostración es muy fácil y queda a cargo del lector. □

Un ejemplo de aplicación de este resultado es el siguiente. Consideremos la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

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

Sea $A = \mathbb{R}^2 \setminus \{0\}$. Como A es abierto y $f_{|A}(x, y) = x \cos \frac{1}{x^2+y^2}$ es continua al ser composición de funciones continuas, se tiene por la proposición anterior que f es continua en A. Además $|f(x, y)| \leq |x| \to 0 = f(0, 0)$ cuando $(x, y) \to (0, 0) =$, luego f es también continua en $(0, 0)$ y por tanto $f : \mathbb{R}^2 \to \mathbb{R}$ es continua.

El siguiente teorema caracteriza la continuidad global de una función: las funciones continuas son aquellas tales que las imágenes inversas de abiertos son abiertas (y lo mismo para cerrados).

Teorema 5.10. *Sea $f : X \to Y$. Las siguientes afirmaciones son equivalentes:*

1. *f es continua.*
2. *Para todo subconjunto abierto A de Y, $f^{-1}(A)$ es abierto en X.*
3. *Para todo subconjunto cerrado C de Y, $f^{-1}(C)$ es cerrado en X.*

Demostración. $(1) \implies (2)$: Sea $x \in f^{-1}(A)$. Como A es abierto y $f(x) \in A$, existe $\varepsilon > 0$ tal que $B(f(x), \varepsilon) \subseteq A$. Ahora, como f es continua en x existe $\delta > 0$ tal que $f(B(x, \delta)) \subseteq B(f(x), \varepsilon) \subseteq A$, y por tanto $B(x, \delta) \subseteq f^{-1}(A)$.
$(2 \implies (1)$: Dados $x \in X$, $\varepsilon > 0$, como $f^{-1}(B(f(x), \varepsilon))$ es un entorno abierto de x, existe $\delta > 0$ tal que $B(x, \delta) \subseteq f^{-1}(B(f(x), \varepsilon))$, lo que implica $f(B(x, \delta)) \subseteq B(f(x), \varepsilon)$.
$(2) \iff (3)$: Basta tener en cuenta que $f^{-1}(X \setminus A) = X \setminus f^{-1}(A)$ y que los complementarios de los abiertos son cerrados y viceversa. □

Una aplicación evidente pero muy útil en la práctica del teorema anterior es que, para toda función continua $f : X \to \mathbb{R}$, los conjuntos $\{x \in X : f(x) > 0\}$ y $\{x \in X : f(x) < 0\}$ son abiertos, y los conjuntos $\{x \in X : f(x) \geq 0\}$, $\{x \in X : f(x) \leq 0\}$ y $\{x \in X : f(x) = 0\}$ son cerrados. Así, ejercicios de relativa enjundia calculística (como por ejemplo probar que $\{(x, y) \in \mathbb{R}^2 : x^2 + \sin y^3 < 1\}$ es abierto en $\mathbb{R}^2$) cuando sólo disponíamos de las herramientas del capítulo 3, resultan ahora insignificantes.

La siguiente proposición resulta provechosa para verificar que ciertas funciones definidas a trozos son continuas. En el caso en que el espacio esté descompuesto como unión finita de cerrados y la restricción a cada uno de éstos de la función en cuestión sea continua, se podrá deducir la continuidad de la función en todo el dominio.

Proposición 5.11. *Sea $f : X \to Y$, supongamos que $X = C_1 \cup ... \cup C_n$, donde cada C_j es cerrado, y que $f_{|C_j} : C_j \to Y$ es continua para cada $j = 1, ..., n$. Entonces f es continua en X.*

Demostración. Sea D un cerrado de Y. Se tiene que

$$f^{-1}(D) = \bigcup_{j=1}^{m} f^{-1}(D) \cap C_j = \bigcup_{j=1}^{m} f_{|C_j}^{-1}(D).$$

Como cada $f_{|C_j}^{-1}(D)$ es un cerrado relativo a C_j (por ser $f_{|C_j} : C_j \to Y$ continua), y cada C_j es asimismo cerrado, resulta que $f_{|C_j}^{-1}(D)$ es cerrado en X para cada $j = 1, ..., n$, y por consiguiente, al ser unión finita de cerrados, $f^{-1}(D)$ es cerrado en X. Por el teorema anterior esto prueba que f es continua. □

Como ejemplos de aplicación de este teorema puede considerarse la función $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} \sin x & \text{si } x \leq 0, \\ x & \text{si } 0 \leq x \leq 1, \\ \frac{1}{x} & \text{si } x \geq 1, \end{cases}$$

o la función $g : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$g(x, y) = \begin{cases} e^{-x^2-y^2} & \text{si } x^2 + y^2 \leq 1, \\ \frac{1}{e}\cos(x^2 + y^2 - 1) & \text{si } x^2 + y^2 \geq 1. \end{cases}$$

La función f está bien definida y es evidentemente continua cuando se restringe a los cerrados $(-\infty, 0]$, $[0, 1]$ y $[1, \infty)$, que recubren $\mathbb{R}$, y por tanto f es continua en $\mathbb{R}$. Lo mismo sucede con la función g y los cerrados $C_1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ y $C_2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \geq 1\}$ cuya unión es $\mathbb{R}^2$.

Una propiedad fundamental de las funciones continuas es que conservan la compacidad.

Teorema 5.12. *Sea $f : X \to Y$ una aplicación continua, y sea $K \subseteq X$ un subconjunto compacto de X. Entonces $f(K)$ es compacto.*

Demostración. Sea $(y_n) = (f(x_n)) \subseteq f(K)$ una sucesión de $f(K)$. Como $(x_n) \subseteq K$ y K es compacto, (x_n) tiene una subsucesión (x_{n_j}) que converge a un punto $x_0 \in K$. Entonces, por ser f continua, $(y_{n_j}) = (f(x_{n_j}))$ converge a $f(x_0) \in f(K)$. □

En el caso $Y = \mathbb{R}$ puede deducirse que toda función continua definida sobre un compacto y con valores reales alcanza siempre un máximo y un minimo absolutos sobre el compacto. Recuérdese que $f : X \to \mathbb{R}$ alcanza un máximo absoluto sobre $A \subseteq X$ si existe $x_0 \in A$ tal que $f(x) \leq f(x_0)$ para todo $x \in A$. Análogamente, f tiene un mínimo absoluto en A si existe $x_1 \in A$ tal que $f(x_1) \leq f(x)$ para todo $x \in A$.

Teorema 5.13. *Sea $f : X \to \mathbb{R}$ continua, y $K \subseteq X$ compacto (no vacío). Entonces existen $x_1, x_2 \in K$ tales que $f(x_1) \leq f(x) \leq f(x_2)$ para todo $x \in K$, es decir, f alcanza un máximo y un mínimo absolutos en K. En particular f está acotada en K.*

Demostración. Por el teorema anterior $f(K)$ es compacto, y en particular cerrado y acotado. Sean $\alpha = \inf f(K), \beta = \sup f(K)$, que existen por ser $f(K)$ acotado. Como $\sup f(K), \inf f(K) \in \overline{f(K)}$, y $f(K)$ es cerrado, se tiene que $\alpha, \beta \in f(K)$ lo que significa que existen $x_1, x_2 \in K$ con

$$f(x_1) = \alpha \leq f(x) \leq \beta = f(x_2)$$

para todo $x \in K$. □

La condición de que K sea compacto es esencial en estos teoremas, como es fácil ver. De hecho, puede demostrarse que si *toda* función continua $f : X \to \mathbb{R}$ alcanza un máximo absoluto en X entonces X es compacto; consultar los problemas 5.20, 5.21 y 5.22.

Otra consecuencia importante del teorema de conservación de la compacidad por aplicaciones continuas es que para toda aplicación continua e inyectiva definida en un compacto, su inversa es también continua.

Teorema 5.14. *Sean K un espacio métrico compacto y $f : K \to Y$ inyectiva y continua. Entonces la aplicación inversa $f^{-1} : f(K) \subseteq Y \to K$ es también continua.*

Demostración. Sea C un cerrado de K; hay que ver que $(f^{-1})^{-1}(C) = f(C)$ es cerrado en $f(K)$. Como C es cerrado en K y K es compacto, se tiene que C es conpacto, luego por el Teorema 5.12 resulta que $f(C)$ es compacto, y por tanto cerrado en $f(K)$. □

Definición 5.5. Se dice que una aplicación $f : X \to Y$ es un *homeomorfismo* entre dos espacios métricos X e Y si f es biyectiva y tanto f como $f^{-1} : Y \to X$ son continuas. Se dice que dos espacios métricos son homeomorfos si existe un homeomorfismo entre ellos.

Como consecuencia inmediata de esta definición y del Teorema anterior obtenemos lo siguiente.

Corolario 5.15. *Sea $f : X \to Y$ una aplicación biyectiva y continua, y supongamos que X es compacto. Entonces f es un homeomorfismo.*

Dos espacios homeomorfos son indistinguibles desde el punto de vista topológico, y en particular tienen las mismas propiedades de compacidad, separabilidad, o conexión (como veremos en al capítulo siguiente). En general, toda propiedad que pueda definirse exclusivamente mediante conjuntos abiertos o cerrados y sin hacer referencia explícita a la distancia, será una propiedad topológica y por tanto será conservada por homeomorfismos.

No obstante hay otras propiedades estrictamente métricas que no se conservan por homeomorfismos, como por ejemplo la completitud. En efecto, si consideramos en $\mathbb{R}$ la distancia

$$d'(x, y) = |\text{arctg}(x) - \text{arctg}(y)|,$$

y por d denotamos la distancia usual de $\mathbb{R}$ definida por el valor absoluto, entonces es fácil ver que d y d' son distancias equivalentes, es decir definen los mismos abiertos en $\mathbb{R}$, o lo que es lo mismo, la identidad es un homeomorfismo entre $(\mathbb{R}, d)$ y $(\mathbb{R}, d')$. Y sin embargo $(\mathbb{R}, d)$ es completo, mientras que $(\mathbb{R}, d')$ no lo es.

Entramos ya en la parte final de este capítulo, dedicada a analizar el concepto de continuidad uniforme y su relación con las nociones hasta ahora estudiadas.

Definición 5.6. Sean X e Y espacios métricos, y sea $f : X \to Y$. Se dice que f es uniformemente continua en $A \subseteq X$ si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que si $x, y \in A$ y $d(x, y) < \delta$ entonces $d(f(x), f(y)) < \varepsilon$.

Por ejemplo, es claro que $f(x) = x$ es uniformemente continua en $\mathbb{R}$. La diferencia obvia con la definición de continuidad es que aquí el δ no depende de cada punto x o y, sino solamente de ε. Es también evidente que si f es uniformemente continua en A entonces $f_{|A} : A \to Y$ es continua (lo cual, nótese bien, no quiere decir que $f : X \to Y$ sea continua en A, como ya sabemos). En particular, si $f : X \to Y$ es uniformemente continua en X entonces $f : X \to Y$ es continua. El recíproco no es cierto, como prueban los siguientes ejemplos.

Ejercicio 5.3. Las funciones $f : \mathbb{R} \to \mathbb{R}$ y $g : (0, 1) \to \mathbb{R}$ definidas por $f(x) = x^2$ y $g(x) = 1/x$ son continuas, pero no uniformemente continuas.

Como en el caso de continuidad, existe una caracterización muy útil de la continuidad uniforme en términos de sucesiones

Proposición 5.16. *Una aplicación $f : X \to Y$ es uniformemente continua si y sólo si para cada par de sucesiones $(x_n), (y_n) \subseteq X$ con $d(x_n, y_n) \to 0$ se tiene que $d(f(x_n), f(y_n)) \to 0$.*

Demostración. Sean $f : X \to Y$ sea uniformemente continua, y (x_n), (y_n) dos sucesiones tales que $d(x_n, y_n) \to 0$. Dado $\varepsilon > 0$, por ser f uniformemente continua existe $\delta > 0$ tal que si $d(x, y) < \delta$ entonces $d(f(x), f(y)) < \varepsilon$. Y puesto que $d(x_n, y_n) \to 0$, existe $n_0 \in \mathbb{N}$ tal que si $n \geq n_0$ entonces $d(x_n, y_n) < \delta$. Por tanto también $d(f(x_n), f(y_n)) < \varepsilon$ si $n \geq n_0$. Esto prueba que $\lim_{n\to\infty} d(f(x_n), f(y_n)) = 0$.

Recíprocamente, si $f : X \to Y$ no es uniformemente continua, existe $\varepsilon_0 > 0$ tal que para todo $\delta > 0$ existen $x_\delta, y_\delta \in X$ tales que $d(x_\delta, y_\delta) < \delta$ pero $d(f(x_\delta, f(y_\delta)) \geq \varepsilon_0$. Tomando δ de la forma $\delta = 1/n$, $n \in \mathbb{N}$, obtenemos dos sucesiones (x_n), (y_n) en X tales que $d(x_n, y_n) < 1/n$ pero $d(f(x_n), f(y_n)) \geq \varepsilon_0$, lo que implica que $\lim_{n\to\infty} d(x_n, y_n) = 0$ y sin embargo no se tiene que $\lim_{n\to\infty} d(f(x_n), f(y_n)) = 0$. $\square$

Este resultado resulta especialmente indicado en la práctica para ver que una determinada función *no* es uniformemente continua: basta encontrar dos sucesiones tales que la distancia entre sus términos tiende a cero pero la distancia entre los términos de sus sucesiones imágenes no tiende a cero. Se invita al lector a repetir el Ejercicio 5.3 usando esta caracterización de la continuidad uniforme en vez de la definición.

En general puede parecer difícil determinar si una función es uniformemente continua, pero hay dos criterios positivos que resultan enormemente efectivos en la práctica. Uno nos lo proporciona el siguiente teorema, que afirma que toda función continua definida sobre un compacto es uniformemente continua en ese compacto. El otro, como veremos después, es el hecho de que toda función Lipschitz (en particular toda función con derivada acotada) es uniformemente continua.

Teorema 5.17. *Sea $f : X \to Y$ continua, y supongamos que X es compacto. Entonces f es uniformemente continua en X.*

Demostración. Sean $(x_n), (y_n)$ sucesiones en X tales que $d(x_n, y_n) \to 0$, y veamos que $d(f(x_n), f(y_n)) \to 0$. Supongamos que no fuera así. Entonces existirían $\varepsilon > 0$ y subsucesiones (x_{n_j}) e (y_{n_j}) de (x_n) e (y_n), respectivamente, tales que

$$d(f(x_{n_j}), f(y_{n_j})) \geq \varepsilon \qquad (*)$$

para todo $j \in \mathbb{N}$. Como X es compacto, existe una subsucesión de (x_{n_j}) que converge a algún $x \in X$. Para no sobrecargar la notación seguiremos denotando a esta subsucesión por (x_{n_j}). Análogamente, la subsucesión de (y_{n_j}) correspondiente a esta subsucesión tiene a su vez una subsucesión que converge a algún $y \in Y$. Por las mismas razones apuntadas antes seguiremos denotando esta subsucesión por (y_{n_j}). Como $d(x_{n_j}, y_{n_j}) \to 0$, se tiene que $d(x, y) = 0$, es decir, $x = y$. Pero entonces, al ser f continua en $x = y$, y puesto que $x_{n_j} \to x$ e $y_{n_j} \to y$, se tiene que $f(x_{n_j}) \to f(x)$ y $f(y_{n_j}) \to f(y) = f(x)$, luego

$$d(f(x_{n_j}), f(y_{n_j})) \to d(f(x), f(y)) = 0,$$

lo cual contradice $(*)$.

Si el lector siente vértigo al considerar tantas subsucesiones de subsucesiones de subsucesiones, puede estudiar la siguiente demostración alternativa, quizás más técnica, aunque con menos engorros de notación. Consideremos $\varepsilon > 0$ dado. Para cada $x \in X$, por ser f continua en x, existe $\delta_x > 0$ tal que si $y \in B(x, 2\delta_x)$ entonces $d(f(y), f(x)) < \varepsilon/2$; en particular, si $y, z \in B(x, 2\delta_x)$ entonces

$$d(f(z), f(y)) \leq d(f(z), f(x)) + d(f(x), f(y)) < \varepsilon/2 + \varepsilon/2 = \varepsilon. \qquad (**)$$

Consideremos ahora el recubrimiento abierto

$$X = \bigcup_{x \in X} B(x, \delta_x).$$

Por ser X compacto existen $x_1, ..., x_n \in X$ tales que

$$X = \bigcup_{n=1}^{n} B(x_j, \delta_j),$$

donde denotamos $\delta_j = \delta_{x_j}$. Definamos entonces $\delta = \min\{\delta_1, ..., \delta_n\}$. Afirmamos que si $d(z, y) < \delta$ entonces $d(f(z), f(y)) < \varepsilon$. En efecto, fijados $z, y \in X$ con $d(z, y) < \delta$, como $X = \bigcup_{j=1}^{n} B(x_j, \delta_j)$, existe $i = i_z \in \{1, ..., n\}$ tal que $z \in B(x_i, \delta_i)$. Ahora, como $d(z, y) < \delta \leq \delta_i \geq d(z, x_i)$, se deduce que $d(y, x_i) \leq d(y, z) + d(z, x_i) < 2\delta_i$. Así pues $z, y \in B(x_i, 2\delta_i)$, lo que gracias a $(**)$ garantiza que $d(f(z), f(y)) < \varepsilon$. □

Definición 5.7. Sea $f : X \to Y$ una aplicación entre dos espacios métricos. Se dice que f es Lipschitz si existe $L \geq 0$ tal que

$$d(f(x), f(y)) \leq L d(x, y)$$

para todo $x, y \in X$. También se dice en este caso que f es L-Lipschitz, o que L es una constante de Lipschitz para f. Por *la* constante de Lipschitz de f suele entenderse la menor de las L posibles, es decir,

$$\text{Lip}(f) = \inf\{L \geq 0 : d(f(x), f(y)) \leq L d(x, y) \ \forall x, y \in X\}.$$

Es fácil ver utilizando el teorema del valor medio que si $f : I \subseteq \mathbb{R} \to \mathbb{R}$ es una función derivable tal que f' está acotada en I, digamos $|f'(t)| \leq C$ para todo $t \in I$, entonces f es C-Lipschitz en I. Por ejemplo, la función $f : \mathbb{R} \to \mathbb{R}$, $f(t) = \cos t$, es 1-Lipschitz en $\mathbb{R}$.

Este resultado se generaliza, como veremos más adelante, a funciones definidas entre subconjuntos de $\mathbb{R}^n$ y $\mathbb{R}^m$, en el sentido de que si A es un subconjunto convexo de $\mathbb{R}^n$ y $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable y tiene derivada acotada en A (lo que equivale a decir que todas las derivadas parciales de primer orden de f están acotadas en A, entonces f es Lipschitz en A). Recíprocamente, es fácil demostrar que si f es Lipschitz en A y es diferenciable en A, entonces la derivada de f está acotada en A. Por tanto, dentro de la categoría de las funciones derivables, las funciones Lipschitz son exactamente las que tienen derivada acotada. Pero hay que ser cautos, porque una función Lipschitz puede muy bien no ser derivable en algunos puntos (considérese el ejemplo del valor absoluto en $\mathbb{R}$, o más en general el de una norma en $\mathbb{R}^n$, que nunca son derivables en el origen, y sin embargo son obviamente funciones 1-Lipschitz). No obstante, un teorema debido a Rademacher afirma que todas las funciones Lipschitz son derivables en muchos puntos (de hecho en la mayoria de ellos, o en *casi todo punto*, aunque ahora mismo no estamos en condiciones de precisar qué quieren decir estas expresiones, ni mucho menos de demostrar este resultado).

Una de las propiedades importantes de las funciones Lipschitz, y la razón fundamental por la cual sacamos a colación esta definición en este momento, es que toda función Lipschitz es también uniformemente continua.

Proposición 5.18. *Si $f : X \to Y$ es Lipschitz en $A \subseteq X$, entonces f es uniformemente continua en A.*

Demostración. Sea $C > 0$ una constante de Lipschitz para f en A. Dado $\varepsilon > 0$, tomemos $\delta = \varepsilon/C > 0$. Entonces, si $x, y \in A$ y $d(x, y) \leq \delta$, se tiene que $d(f(x), f(y)) \leq Cd(x, y) \leq C\delta = \varepsilon$. □

Conviene advertir que el recíproco de esta Proposición no es cierto, pues hay funciones uniformemente continuas que no son Lipschitz. Por ejemplo, la función $f : [-1, 1] \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} x \sin\left(\frac{1}{x}\right) & \text{si }, x \neq 0 \\ 0 & \text{si } x = 0, \end{cases}$$

es continua y, como $[-1, 1]$ es compacto, también uniformemente continua en $[-1, 1]$. Sin embargo f no es Lipschitz en $[-1, 1]$, puesto que ni siquiera lo es en $(0, 1)$ ya que su derivada, que existe y está definida por

$$f'(x) = \sin\left(\frac{1}{x}\right) - \frac{1}{x} \cos\left(\frac{1}{x}\right)$$

en este subintervalo, no está acotada.

Una propiedad interesante de las aplicaciones uniformemente continuas es que transforman sucesiones de Cauchy en sucesiones de Caychy.

Proposición 5.19. *Sea $f : X \to Y$ uniformemente continua, y supongamos que (x_n) es de Cauchy en X. Entonces $(f(x_n))$ es de Cauchy en Y.*

Demostración. Dado $\varepsilon > 0$, la continuidad uniforme de f nos permite encontrar $\delta_\varepsilon > 0$ tal que si $d(x, y) < \delta_\varepsilon$ entonces $d(f(x), f(y)) < \varepsilon$. Como (x_n) es de Cauchy en X, existe $n_\varepsilon \in \mathbb{N}$ tal que si $n, m \geq n_\varepsilon$ entonces $d(x_n, x_m) < \delta_\varepsilon$. Por tanto también $d(f(x_n), f(x_m)) < \varepsilon$ si $n, m \geq n_\varepsilon$. Esto prueba que $(f(x_n))$ es de Cauchy en Y. □

La proposición anterior no es cierta, en general, si f no es uniformemente continua, como prueba el ejemplo de $f : (0, 1) \to \mathbb{R}$, $f(t) = 1/t$, $(t_n) = (1/n)$.

Otra propiedad importante de las funciones uniformemente continuas es que siempre poseen extensiones únicas a la adherencia de sus dominios, al menos cuando toman sus valores en espacios métricos completos.

Teorema 5.20. *Sean X, Y espacios métricos, $f : A \subseteq X \to Y$ uniformemente continua, y supongamos que Y es completo. Entonces existe una única extensión $\overline{f} : \overline{A} \to Y$ uniformemente continua de f a $\overline{A}$.*

Demostración. Sea $x \in \overline{A}$, y tomemos $(x_n) \subseteq A$ tal que $x_n \to x$. Como (x_n) es de Cauchy (por ser convergente) y f es uniformemente continua, tenemos que $(f(x_n))$ es de Cauchy en Y, y como Y es completo deducimos que $(f(x_n))$ converge a un punto y_x de Y. Además, si tomamos otra sucesión $(z_n) \subseteq A$ con $z_n \to x$, también se tiene que $f(z_n) \to y_x$. En efecto,

$$d(f(z_n), y_x) \leq d(f(z_n), f(x_n)) + d(f(x_n), y_x) \to 0 + 0 = 0,$$

ya que $d(z_n, x_n) \to d(x, x) = 0$, f es uniformemente continua y $f(x_n) \to y_x$.

Esto prueba que para cada $x \in \overline{A}$ existe un único $y_x \in Y$ tal que para toda sucesión $(x_n) \subseteq A$ con $\lim_{n\to\infty} x_n = x$ es $y_x = \lim_{n\to\infty} f(x_n)$. O lo que es lo mismo, para cada $x \in X$ existe $\lim_{z\in A, z\to x} f(z) = y_x$. Por tanto la aplicación

$$\overline{f} : \overline{A} \to Y, \ \ x \to \overline{f}(x) := \lim_{z\in A, z\to x} f(z)$$

está bien definida.

Veamos que $\overline{f} : \overline{A} \to Y$ es uniformemente continua. Fijemos $\varepsilon > 0$. Como f es uniformemente continua en A, existe $\delta > 0$ tal que si $x', y' \in A$ y $d(x', y') \leq 3\delta$ entonces

$$d(f(x'), f(y')) \leq \varepsilon/3. \qquad (*)$$

Tomemos ahora $x, y \in \overline{A}$ con $d(x, y) \leq \delta$. Como $\lim_{z\in A, z\to x} f(z) = \overline{f}(x)$ y $\lim_{z\in A, z\to y} f(z) = \overline{f}(y)$, y A es denso en $\overline{A}$, se sigue que existen $x', y' \in A$ tales que

$$d(x, x') \leq \delta,\ d(y, y') \leq \delta,\ d(\overline{f}(x), f(x')) \leq \varepsilon/3 \text{ y } d(\overline{f}(y), f(y')) \leq \varepsilon/3.$$

Entonces tenemos que

$$d(x', y') \leq d(x', x) + d(x, y) + d(y, y') \leq \delta + \delta + \delta = 3\delta,$$

y por tanto, teniendo en cuenta $(*)$,

$$d(\overline{f}(x), \overline{f}(y)) \leq d(\overline{f}(x), f(x')) + d(f(x'), f(y')) + d(f(y'), \overline{f}(y)) \leq \\ \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.$$

Esto prueba que $\overline{f}$ es uniformemente continua en $\overline{A}$.

Finalmente, $\overline{f}$ es la única extensión uniformemente continua de f a $\overline{A}$. En efecto, como A es denso en $\overline{A}$ y $\overline{f}_{|A} = f$, esto se sigue del hecho de que dos funciones continuas que coincidan en un subconjunto denso de su dominio han de ser iguales; ver el problema 5.10. □

Veamos a continuación una serie de pistas y ejemplos que muestran cómo pueden emplearse los resultados expuestos hasta ahora para determinar si una función es o no es uniformemente continua en un subconjunto dado de su dominio.

1. Lo primero es determinar si f es continua, ya que se trata de una condición necesaria para ser uniformemente continua. Por ejemplo, la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

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

es discontinua en $(0, 0)$ y por tanto no es uniformemente continua en ningún subconjunto de $\mathbb{R}^2$ que contenga el origen.

2. Si f es continua y su dominio es un compacto, entonces por el Teorema 5.17 sabemos que f es uniformemente continua. Por ejemplo, la función $f : B \subseteq \mathbb{R}^2 \to \mathbb{R}$ definida por

$$f(x, y) = \begin{cases} |x + y| \sin \frac{1}{x^2+y^2} & \text{si } (x, y) \neq (0, 0), \\ 0 & \text{si } (x, y) = (0, 0), \end{cases}$$

donde B es la bola unidad cerrada de $\mathbb{R}^2$, es continua, y como B es compacto, se sigue que f es uniformemente continua.

3. Si f está definida como suma o composición de funciones que ya sabemos que son uniformemente continuas, entonces f es uniformemente continua. El producto de funciones uniformemente continuas, sin embargo, no tiene por qué ser uniformemente continuo (aunque sí lo es si las funciones involucradas en el producto están acotadas); ver los problemas 5.42, 5.43 y 5.44. Así por ejemplo puede comprobarse que $f(x, y) = |\sin x \cos y|$ es uniformemente continua en $\mathbb{R}^2$.

4. El siguiente test puede ser comprobar si f es derivable y, en caso afirmativo, si su derivada está acotada en un convexo que contenga el conjunto en cuestión (para la definición de conjunto convexo véase el capítulo siguiente). En tal caso la función será Lipschitz, y por tanto uniformemente continua. Por ejemplo, consideremos la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

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

Esta función tiene derivadas parciales continuas en $\mathbb{R}^2$, a saber,

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

por lo que es diferenciable en $\mathbb{R}^2$ (ver el Capítulo 9), y como estas derivadas parciales están acotadas en todo $\mathbb{R}^2$, se sigue que f es Lipschitz, y en particular uniformemente continua, como veremos en el Capítulo 10. Si el alumno siente escrúpulos hacia la utilización en la práctica de resultados teóricos que todavía no han sido establecidos, puede limitarse a usar este resultado solamente en el caso de funciones de variable real hasta que llegue al Capítulo 10.

Cuando la función no es derivable o su derivada no es acotada, este criterio no es concluyente, ya que como hemos visto más arriba existen funciones Lipschitz que no son derivables, y también existen funciones uniformemente continuas que no son Lipschitz.

5. Otra manera de ver que una función es Lipschitz (y por tanto también uniformemente continua) es verificar que está definida mediante sumas o composiciones de funciones Lipschitz. En efecto, por el problema 5.25 la suma de dos funciones de Lipschitz es de Lipschitz, y lo mismo ocurre con la composición. En general el producto de dos funciones Lipschitz no es Lipschitz (considérese $x \to x^2 = x$, que es producto de la identidad consigo misma), pero sí lo es si ambas funciones están acotadas (ver problema 5.26). Así, para la función f del ejemplo anterior, resulta que f es Lipschitz, ya que es composición de la función

$$h(t) = \frac{1}{1 + t^2},$$

que es Lipschitz en todo $\mathbb{R}$ por tener derivada acotada, con la función norma euclidea, $(x, y) \to \|(x, y)\|$, que como bien sabemos es 1-Lipschitz. Otros ejemplos de funciones en los que puede verse que son Lipschitz con este tipo de criterios son $g : \mathbb{R}^2 \to \mathbb{R}$, $h : B \to \mathbb{R}$, donde B es la bola unidad cerrada de $\mathbb{R}^3$ y

$$g(x, y) = \sin |x + y| + \cos^2 |x - y|, \ h(x, y, z) = (x^2 + y^2 + z^2)e^{-|x+y-z|}.$$

Nótese también que, por la manera en que está involucrado el valor absoluto en esta definición, las funciones tienen puntos de no diferenciabilidad, por lo que el criterio de 3 no es aplicable.

6. Aunque el dominio A de la función f no sea un compacto, si ésta es continua el Teorema 5.17 puede seguir siendo aplicable, ya que tal vez f pueda extenderse con continuidad a una función $\overline{f}$ definida sobre un compacto K que contenga a A, con lo que $\overline{f}$ será uniformemente continua en K y por tanto $\overline{f}_{|A} = f$ también será uniformemente continua en A. Por ejemplo, consideremos el conjunto $A = \{(x, y) \in \mathbb{R}^2 : 0 < y < x < 1\}$ y la función $f : A \to \mathbb{R}$ definida por

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

Puede resultar curioso constatar que esta expresión no tiene límite en el origen, como ya vimos en el punto **7** de los criterios para la existencia de límite expuestos más arriba, y en particular ni ella ni ninguna extensión suya puede ser continua en $(0,0)$, entendida como función de $\mathbb{R}^2$ en $\mathbb{R}$. Sin embargo, al restringirnos al dominio A, sí es posible definir una extesión $\overline{f}:\overline{A}\to\mathbb{R}$ de f que es continua en $\overline{A}=\{(x,y)\in\mathbb{R}^2 : 0\leq y\leq x\leq 1\}\ni(0,0)$ (entendida como función de $\overline{A}$ en $\mathbb{R}$, claro está). En efecto, si $0<y\leq x\leq 1$ se tiene que

$$|\frac{y}{x}\sin(x^2+y^2)|\leq\sin(x^2+y^2)\to 0$$

cuando $(x,y)\to 0$, y por tanto

$$\lim_{(x,y)\in\overline{A},(x,y)\to(0,0)}\frac{y}{x}\sin(x^2+y^2)=0,$$

de modo que la función $\overline{A}\to\mathbb{R}$ definida por

$$\overline{f}(x,y)=\begin{cases}\frac{y}{x}\sin(x^2+y^2) & \text{si } (x,y)\in\overline{A}, (x,y)\neq(0,0),\\ 0 & \text{si } (x,y)=(0,0)\end{cases}$$

en continua en $\overline{A}$ y por consiguiente, al ser $\overline{A}$ compacto, $\overline{f}$ es uniformemente continua en $\overline{A}$, y en particular $f=\overline{f}_{|A}$ es también uniformemente continua en A.

7. También puede emplearse la táctica de las extensiones (uniformemente) continuas en la dirección inversa, es decir, para demostrar que una función f no puede ser uniformemente continua en un conjunto A. En efecto, si f es uniformemente continua en A, el Teorema 5.20 nos garantiza que f tiene una (única) extensión uniformemente continua $\overline{f}$ definida sobre $\overline{A}$. En particular deberá existir el límite

$$\lim_{x\in A,x\to x_0}f(x)$$

para cada $x_0\in A$. Así que para ver que f *no* es uniformemente continua en A basta encontrar un punto $x_0\in\overline{A}$ de modo que no exista

$$\lim_{x\in A,x\to x_0}f(x).$$

Por ejemplo, consideremos $A=\{(x,y)\in\mathbb{R}^2 : 0<x<1, 0<y<1\}$, y $f:A\to\mathbb{R}^2$ definida por

$$f(x,y)=\Big(\frac{xy}{x^2+y^2},\frac{\sin(x^2+y^2)}{x^2+y^2}\Big).$$

Tenemos que $(0,0)\in\overline{A}$, y es fácil ver que no existe

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

(considéresen los límites a lo largo de las semirrectas $y=\lambda x$, $x>0$, con $0<\lambda<\infty$, que son todos diferentes). Por tanto f no puede ser uniformemente continua en A.

8. Otro criterio negativo indirecto lo proporciona el problema 5.48: si una función $f:A\subseteq\mathbb{R}^n\to\mathbb{R}^m$ es uniformemente continua entonces es acotada sobre los acotados de A. Por tanto, si detectamos un conjunto acotado sobre el cual la función dada no es acotada, podemos concluir que ésta no es uniformemente continua. Este criterio puede emplearse, por ejemplo, para ver que la función

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

no es uniformemente continua en la bola unidad abierta $B = \{(x, y, z) \in \mathbb{R}^3 : x^2+y^2+z^2 < 1\}$. Ver también los problemas 5.49 y 5.50.

9. A la inversa, un criterio positivo nos lo proporciona el problema 5.38: si una funcion continua $f : \mathbb{R}^n \to \mathbb{R}$ tiene limite finito en el infinito entonces es uniformemente continua. Por ejemplo,

$$g(x, y, z) = \frac{x^2 - y^2 + z}{1 + x^4 + y^6 + z^8}$$

es uniformemente continua por este motivo.

10. Finalmente, como en el caso de límites, uno de los criterios más efectivos cuando se tienen las ideas claras sobre por qué puede no ser uniformemente continua una función determinada es la proporcionada por la Proposición 5.16: basta encontrar dos sucesiones cuya distancia tiende a cero pero la distancia entre sus imágenes no. Consideremos, por ejemplo, la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

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

Sean $(x_n, y_n) = (0, \sqrt{2n\pi})$, $(z_n, w_n) = (0, \sqrt{2n\pi} + \sqrt{\pi/32n})$. Entonces

$$\|(x_n, y_n) - (z_n, w_n)\| = \sqrt{\frac{\pi}{32n}} \to 0,$$

y sin embargo

$$|f(x_n, y_n) - f(z_n, w_n)| = |\sin(2n\pi) - \sin(2n\pi + \frac{\pi}{2} + \frac{\pi}{32n})| =$$
$$|\sin(\frac{\pi}{2} + \frac{\pi}{32n})| \to \sin\frac{\pi}{2} = 1 \neq 0,$$

luego f no es uniformemente continua en $\mathbb{R}^2$.

Para concluir, demostraremos el resultado, ya anunciado en el Capítulo 1, de que todas las normas en $\mathbb{R}^n$ son equivalentes. Puede considerarse que toda la teoría sobre conceptos topológicos y métricos desarrollada hasta ahora se ha referido exclusivamente a una cualquiera de las normas $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ (o a las tres), y que el resultado siguiente permite extender todos los resultados precedentes al caso de una norma arbitraria en $\mathbb{R}^n$.

Teorema 5.21. *Sea E un espacio vectorial de dimensión finita. Entonces todas las normas en E son equivalentes.*

Demostración. Sea $n = \dim(E)$, y elijamos $\{e_1, ..., e_n\}$ una base cualquiera de E. Cada vector $x \in E$ tiene por tanto una expresión única como combinación lineal de $e_1, ..., e_n$, digamos

$$x = \sum_{i=1}^{n} x_i e_i,$$

y podemos definir entonces

$$\|x\|_\infty = \max\{|x_1|, ..., |x_n|\}.$$

Es inmediato comprobar que $\|\cdot\|_\infty$ es una norma en E. Consideremos ahora otra norma cualquiera $\|\cdot\|$ en E, y veamos que $\|\cdot\|_\infty$ y $\|\cdot\|$ son equivalentes.

Para cada $x \in E$ tenemos que

$$\|x\| = \|\sum_{i=1}^{n} x_i e_i\| \leq \sum_{i=1}^{n} |x_i|\|e_i\| \leq \sum_{i=1}^{n} \|x\|_\infty \|e_i\| = M\|x\|_\infty, \qquad (*)$$

donde $M := \sum_{i=1}^{n} \|e_i\|$ obviamente no depende de x. Por tanto

$$\|x-y\| \leq M\|x-y\|_\infty$$

para todo $x, y \in E$, es decir, $\|\cdot\|$ es Lipschitz, y en particular continua, en $(E, \|\cdot\|_\infty)$. Por otro lado, la esfera unidad $S = \{x \in E : \|x\|_\infty\}$ es compacta en $(E, \|\cdot\|_\infty)$. Entonces, por el Teorema 5.13, existe $x_0 \in S$ tal que $\|x_0\| = \inf\{\|x\| : x \in S\} := m$. Además $m = \|x_0\| > 0$ ya que $x_0 \in S$ y $0 \notin S$. Esto significa que

$$\|\frac{x}{\|x\|_\infty}\| \leq m$$

para todo $x \in E \setminus \{0\}$, o lo que es lo mismo,

$$m\|x\|_\infty \leq \|x\| \tag{$**$}$$

para todo $x \in E$. Juntando $(*)$ y $(**)$ obtenemos que $m\|x\|_\infty \leq \|x\| \leq M\|x\|_\infty$ para todo $x \in E$, es decir, $\|\cdot\|$ y $\|\cdot\|_\infty$ son equivalentes. □

Para terminar haremos un breve estudio de la continuidad de las aplicaciones lineales y sus normas.

Proposición 5.22. *Sean E y F espacios vectoriales normados, y sea $T : E \to F$ una aplicación lineal. Son equivalentes:*

1. *T es continua;*
2. *T es continua en 0;*
3. *Existe $M \geq 0$ tal que $\|T(x)\| \leq M\|x\|$ para todo $x \in E$;*
4. *$\sup\{\|T(x)\| : x \in E, \|x\| \leq 1\} < +\infty$;*
5. *T es Lipschitz.*

Demostración. $(1) \implies (2)$: es trivial.

$(2) \implies (3)$. Como f es continua en 0, existe $\delta > 0$ tal que si $\|x\| \leq \delta$ entonces $\|T(x)\| \leq 1$. Luego, para todo $x \in E \setminus \{0\}$ se tiene que

$$\|T(x)\| = \left\| T\left(\frac{\|x\|}{\delta}\frac{\delta x}{\|x\|}\right)\right\| = \left\|\frac{\|x\|}{\delta} T\left(\frac{\delta x}{\|x\|}\right)\right\| = \frac{\|x\|}{\delta}\left\| T\left(\frac{\delta x}{\|x\|}\right)\right\| \leq \frac{\|x\|}{\delta} = M\|x\|,$$

donde $M = 1/\delta$. Y si $x = 0$ entonces $\|T(x)\| \leq M\|x\|$ trivialmente.

$(3) \implies (4)$: es trivial

$(4) \implies (5)$: denotemos $\alpha := \sup\{\|T(x)\| : x \in E, \|x\| \leq 1\}$. Se tiene, para todo $x \neq y$,

$$\begin{aligned} &\|T(x) - T(y)\| = \|T(x-y)\| = \left\| T\left(\|x-y\|\frac{x-y}{\|x-y\|}\right)\right\| \\ &= \|x-y\| \left\| T\left(\frac{x-y}{\|x-y\|}\right)\right\| \leq \|x-y\|\alpha, \end{aligned}$$

y en particular T es α-Lipschitz.

$(5) \implies (1)$: es trivial. □

En realidad, en los espacios que manejaremos en este curso estas condiciones se cumplen siempre, ya que toda aplicación lineal entre espacios de dimensión finita es continua.

Proposición 5.23. *Sean E y F espacios vectoriales normados, y supongamos que dim$(E) < \infty$. Entonces toda aplicación lineal $T : E \to F$ es continua.*

Demostración. Sea $\{e_1, ..., e_n\}$ una base de E. Para todo $x = \sum_{i=1}^n x_i e_i \in E$, si consideramos la norma $\|x\|_\infty = \sup\{|x_i| : 1 = 1, ..., n\}$, se tiene que

$$\|T(x)\|_F = \|\sum_{i=1}^n x_i T(e_i)\|_F \leq \sum_{i=1}^n |x_i| \|T(e_i)\|_F \leq C\|x\|_\infty,$$

donde $C = \sum_{i=1}^n \|T(e_i)\|_F$. Por otra parte, como todas las normas son equivalentes en E, existe $C' > 0$ tal que $\|x\|_\infty \leq C'\|x\|_E$ para todo $x \in E$. Por tanto se obtiene que

$$\|T(x)\|_F \leq M\|x\|_E$$

para todo $x \in E$, donde $M = CC'$. □

Puede demostrarse que en todo espacio vectorial normado de dimensión infinita existen siempre formas lineales no continuas (ver problema 5.54). Por tanto la Proposición anterior caracteriza los espacios vectoriales E de dimensión finita.

Es fácil ver que para cualquier aplicación continua $T : E \to F$ se tiene que

$$\sup\{\|T(x)\| : \|x\| \leq 1\} = \text{ínf}\{M \leq 0 : \|T(x)\| \leq M\|x\| \text{ para todo } x \in E\}.$$

A este número se le llama *norma* de la aplicación lineal T, y puede comprobarse que es efectivamente una norma en el espacio vectorial de las aplicaciones lineales continuas de E en F, que denotaremos $\mathcal{L}(E, F)$. Así, se define

$$\|T\| = \sup\{\|T(x)\| : x \in E, \|x\| \leq 1\}$$

para todo $T \in \mathcal{L}(E, F)$.

El espacio de las aplicaciones lineales entre $\mathbb{R}^n$ y $\mathbb{R}^m$, $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, es naturalmente isomorfo al espacio de las matrices $n \times m$, que a su vez denotaremos por $\mathcal{M}_{n\times m}$ (y que es obviamente isomorfo a $\mathbb{R}^{nm}$).

Nótese que, como consecuencia inmediata de la Proposición 5.23, todo isomorfismo lineal entre dos espacios vectoriales de dimensión finita es también un homeomorfismo, cualesquiera que sean las normas consideradas.

5.1. Problemas

Problema 5.1. Determinar si los siguientes límites existen o no, y en caso afirmativo hallar su valor:

$$\begin{aligned}
&\lim_{(x,y)\to(0,0)} \frac{e^{xy}}{x+1}, \quad \lim_{(x,y)\to(0,0)} \frac{x^2y}{x^2+y^2},\\
&\lim_{(x,y)\to(0,0)} \frac{\sin(x^2+y^2)}{x^2+y^2}, \quad \lim_{(x,y)\to(0,0)} \frac{|x|+|y|}{\sqrt{x^2+y^2}},\\
&\lim_{(x,y)\to(0,0)} \frac{e^{xy}-1}{x}, \quad \lim_{(x,y)\to(0,0)} \frac{x|y|}{\sqrt{x^2+y^2}},\\
&\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2}, \quad \lim_{(x,y)\to(0,0)} \frac{x^2y^2}{x^2y^2+(x-y)^2},\\
&\lim_{(x,y)\to(0,0)} (x+y)\sin\frac{1}{x}\sin\frac{1}{y}, \quad \lim_{(x,y)\to(0,0)} \frac{1}{x}\sin(xy),\\
&\lim_{(x,y)\to(0,0)} x^2\sqrt{y}\log(x^2+y^2), \quad \lim_{(x,y)\to(0,0)} \frac{y}{x}\sin(x^2+y^2).
\end{aligned}$$

Problema 5.2 (Teorema del bocadillo)**.** Demostrar que si $f, g, h : A \subseteq X \to \mathbb{R}$ son funciones tales que $f(x) \leq g(x) \leq h(x)$ para todo $x \in A$ y existen $\lim_{x\to a} f(x) = \lim_{x\to a} h(x)$, entonces también existe $\lim_{x\to a} g(x)$, y además

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

Problema 5.3. Sean $f, g : A \subseteq X \to Y$ tales que $f(x) = g(x)$ para todo x de un entorno U de $a \in A'$. Supongamos que existe $\lim_{x\to a} g(x)$. Demostrar que entonces también existe $\lim_{x\to a} f(x)$, y

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

Deducir que la continuidad de una función f en un punto a también es una propiedad local.

Problema 5.4. Si $\lim_{x\to a} f(x) = \infty$ y $\lim_{t\to\infty} g(t) = L$, probar que $\lim_{x\to a} g(f(x)) = L$.

Problema 5.5. Sea $f : \mathbb{R}^2 \to \mathbb{R}$. Demostrar que $\lim_{(x,y)\to(0,0)} f(x,y) = L$ si y solo si $\lim_{\rho\to 0} f(\rho\cos\theta, \rho\sin\theta) = L$ uniformemente en θ; es decir, si y solo si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que si $0 < \rho < \delta$ entonces $|f(\rho\cos\theta, \rho\sin\theta) - L| \leq \varepsilon$ para todo $\theta \in [0, 2\pi]$.

Problema 5.6. Aplicando el ejercicio anterior, probar que si $|f(\rho\cos\theta, \rho\sin\theta)| \leq F(\rho)G(\theta)$, donde $\lim_{\rho\to 0} F(\rho) = 0$ y G es una función acotada, entonces $\lim_{(x,y)\to(0,0)} f(x,y) = 0$.

Problema 5.7. Sea $f : \mathbb{R}^2 \to \mathbb{R}$, $(x_0, y_0) \in \mathbb{R}^2$. Supongamos que existen los límites

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y), \ \lim_{x\to x_0}\Big(\lim_{y\to y_0} f(x,y)\Big), \ y \ \lim_{y\to y_0}\Big(\lim_{x\to x_0} f(x,y)\Big):$$

Probar que entonces los tres límites son iguales.

Problema 5.8. Sea B una bola de centro a en un espacio vectorial normado E, y sea $f : B \to Y$, donde Y es un espacio métrico. Demostrar que $\lim_{x\to a} f(x) = b$ si y solo si para cada curva continua $\gamma : [0,1] \to B$ tal que $\gamma(0) = a$ y $\gamma(t) \neq a$ para $t \neq 0$ se tiene que

$$\lim_{t\to 0^+} f(\gamma(t)) = b.$$

Indicación: para ver que la condición sobre las curvas implica la existencia de límite, supóngase que éste no existe o que no es b, entonces existen $\varepsilon_0 > 0$ y una sucesión (x_n) en B tales que $\lim_{n\to\infty} x_n = a$ y sin embargo $d(f(x_n), b) \geq \varepsilon_0$ para todo $n \in \mathbb{N}$. Defínase entonces una curva continua $\gamma : [0,1] \to B$ tal que $\gamma(0) = a$ y $\gamma(1/n) = x_n$ para todo $n \in \mathbb{N}$. Como curiosidad e indicio de que el resultado no es cierto en todo espacio métrico X, piénsese cuáles han sido las ventajas derivadas de poder trabajar en una bola de un espacio normado.

Problema 5.9. Estudiar la existencia (y calcular en su caso) los límites siguientes:

$$\lim_{(x,y)\to\infty} \frac{\cos(x+y)}{e^x + e^y}, \quad \lim_{(x,y)\to\infty} \frac{x+y}{e^{x^2+y^2}}$$
$$\lim_{(x,y)\to\infty} \frac{(3-x^2)\cos(y-x^3)}{x^2+y^4}, \quad \lim_{(x,y)\to\infty} \frac{(y-x^2)\log(1+x^2+y^2)}{1+\sin^2 x}.$$

Problema 5.10. Probar que si $f, g : X \to Y$ son continuas y existe un subconjunto D denso en X tal que $f(x) = g(x)$ para todo $x \in D$, entonces $f = g$.

Problema 5.11. Sea $(E, \|\cdot\|)$ un espacio vectorial normado. Probar que $\|\cdot\| : E \to \mathbb{R}$ es continua.

Problema 5.12. Poner un ejemplo de una función continua $f : \mathbb{R}^2 \to \mathbb{R}^2$ y un abierto $U \subseteq \mathbb{R}^2$ tales que $f(U)$ no es abierto. Encontrar también $g : \mathbb{R}^2 \to \mathbb{R}^2$ continua y $C \subseteq \mathbb{R}^2$ tal que C es cerrado pero $f(C)$ no lo es.

Problema 5.13. Se dice que una aplicación $f : X \to Y$ es abierta si $f(U)$ es abierto en Y para todo abierto U de X. Demostrar que las proyecciones $P_1 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ y $P_2 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ definidas por $P_1(x,y) = x$, $P_2(x,y) = y$, son aplicaciones abiertas.

Problema 5.14. *Poner un ejemplo de función continua $f : \mathbb{R}^2 \to \mathbb{R}$ y $K \subset \mathbb{R}$ compacto tales que $f^{-1}(K)$ no es compacto.*

Problema 5.15. Sea $f : X \to \mathbb{R}$ continua en x_0 con $f(x_0) > 0$. Probar que existe $\delta > 0$ tal que $f(x) > 0$ para todo $x \in B(x_0, r)$.

Problema 5.16. Dar un ejemplo de una función separadamente continua que no es continua, es decir, una función $f : \mathbb{R}^2 \to \mathbb{R}$ tal que para cada $y \in \mathbb{R}$ la función $\mathbb{R} \ni x \mapsto f_y(x) = f(x,y)$ es continua, y para cada $x \in \mathbb{R}$ la función $\mathbb{R} \ni y \mapsto f_x(y) = f(x,y)$ es también continua, pero $f : \mathbb{R}^2 \to \mathbb{R}$ no es continua.

Problema 5.17. Probar que si $f : \mathbb{R}^n \to \mathbb{R}^m$ es continua y $B \subset \mathbb{R}^n$ es acotado entonces $f(B)$ es acotado en $\mathbb{R}^m$. Observar que esto es cierto gracias a que el dominio de la función es todo $\mathbb{R}^n$, pero no es cierto en general si sustituimos $\mathbb{R}^n$ por un subconjunto abierto suyo.

Problema 5.18. Poner un ejemplo que muestre que el Teorema 5.14 falla si X no es compacto. *Indicación:* Si se desea un ejemplo en donde X sea un subconjunto de $\mathbb{R}^n$, hay que tomar $n \geq 2$.

Problema 5.19. Encontrar una función $f : \mathbb{R}^2 \to \mathbb{R}$ que sea continua a lo largo de cualquier recta que pase por el origen y que sin embargo no sea continua en $(0,0)$.

Problema 5.20. Dar un ejemplo de una función (discontinua) definida sobre un compacto que no alcance ni máximo ni mínimo.

Problema 5.21. Dar un ejemplo de función continua definida sobre un conjunto de $\mathbb{R}^n$ que no alcance ni máximo ni mínimo. Más en general, demostrar que si A es un subconjunto de $\mathbb{R}^n$ que tiene la propiedad de que *toda* función continua $f : A \to \mathbb{R}$ alcanza un máximo absoluto en A, entonces A es compacto.
Indicación: Basta ver que A es cerrado y acotado. Si A no fuera cerrado, tómese $x_0 \in \overline{A} \setminus A$ y considérese la función $A \ni x \mapsto f(x) = 1/d(x, x_0)$. Si A no fuera acotado, considérese $A \ni x \mapsto g(x) = \|x\|$.

Problema 5.22. Aún con más generalidad, demuéstrese que un espacio métrico X es compacto si y sólo si toda función continua $f : X \to \mathbb{R}$ alcanza un máximo absoluto en X.
Indicación: Hay que ver que X es completo y totalmente acotado. Si X no fuera completo, considérese su compleción $\hat{X}$ (ver el problema 4.12), y tómense $\hat{x} \in \hat{X} \setminus X$ y $X \ni x \mapsto f(x) = \hat{d}(x, \hat{x})$. Si X no fuera totalmente acotado, existirían $\varepsilon > 0$ y una sucesión $(x_n) \subseteq X$ tal que $d(x_n, x_m) \geq 3\varepsilon$ si $n \neq m$; entonces puede construirse una función cuya gráfica es una unión de conos de altura cada vez más grande, con base de radio ε y centrados en los puntos x_n. Con más precisión, sea $h : \mathbb{R} \to \mathbb{R}$ la función definida por $h(t) = 1 - |t|$ si $|t| \leq 1$, y $h(t) = 0$ si $|t| \geq 1$, y considérese

$$f(x) = \sum_{n=1}^{\infty} nh\big(\frac{d(x, x_n)}{\varepsilon}\big).$$

Problema 5.23. Sea $f : X \to \mathbb{R}$ una función continua, y $\alpha : [0,1] \to X$ una curva continua en X. Probar que f alcanza un máximo y un mínimo sobre α.

Problema 5.24. Probar que no existe ninguna aplicación continua y sobreyectiva de $[0,1]$ en $(0,1)$.

Problema 5.25. Probar que si $f, g : X \to \mathbb{R}^m$ son Lipschitz entonces $f + g$ también es Lipschitz. Demostrar también que si $F : X \to Y$, $G : Y \to Z$ son Lipschitz entonces $G \circ F : X \to Z$ también es Lipschitz. Estimar las constantes de Lipschitz de $f + g$ y de $G \circ F$ en función de las de f, g, F y G.

Problema 5.26. Probar que si $f, g : X \to \mathbb{R}$ son Lipschitz y acotadas en X entonces su producto $fg : X \to \mathbb{R}$ también es Lipschitz en X.

Problema 5.27. Sean $a, b \in \mathbb{R}$ con $a < b$. Probar que:

1. $[a, b)$ y $(a, b]$ son homeomorfos;
2. (a, b) es homeomorfo a $\mathbb{R}$;
3. $[a, b]$ es homeomorfo a $[0, 1]$.

Problema 5.28. En los siguientes casos, estudiar si $f : \mathbb{R} \to f(\mathbb{R}) \subset \mathbb{R}^2$ es un homeomorfismo sobre su imagen:

1. $f(t) = (t^2 - t, t^2 + 2t)$;
2. $f(t) = \big(\frac{1}{1+t^2}, \frac{t}{1+t^2}\big)$.

Problema 5.29. Sean $A = [-1, 1] \times [-1, 1]$, $f : \mathbb{R}^2 \to \mathbb{R}^2$,

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

Determinar $f(A)$, y estudiar si $f_{|A} : A \to f(A)$ es un homeomorfismo.

Problema 5.30. Sea $f : X \to Y$ una aplicación continua. Demostrar que entonces su gráfica $G_f = \{(x, y) \in X \times Y : y = f(x)\}$ es un conjunto cerrado en $X \times Y$.

Problema 5.31. Demostrar que el recíproco del problema anterior no es cierto en general, exhibiendo una función $f : [-1, 1] \to \mathbb{R}$ discontinua y con gráfica cerrada en $\mathbb{R}^2$.

Problema 5.32. Demostrar que, sin embargo, cuando Y es compacto, entonces el hecho de que G_f sea cerrado en $X \times Y$ equivale a decir que $f : X \to Y$ es continua. Deducir de aquí que si $f : X \to \mathbb{R}^m$ tiene gráfica cerrada y es acotada, entonces f es continua.

Problema 5.33. Demostrar que $f : X \to Y$ es continua si y sólo si para todo $A \subseteq X$ se tiene que

$$f(\overline{A}) \subseteq \overline{f(A)}.$$

Problema 5.34. Encontrar una función $f : A \to \mathbb{R}$ y un subconjunto B denso de A tal que $f_{|B}$ es continua, pero f no es continua en ningún punto.

Problema 5.35. Demostrar que la función distancia a un conjunto $A \subseteq X$,

$$x \mapsto d(x, A) = \operatorname{ínf}\{d(x, a) : a \in A\},$$

es 1-Lipschitz en X. Probar también que

$$\overline{A} = \{x \in X : d(x, A) = 0\}.$$

Problema 5.36. Demostrar que si A es cerrado y K es compacto en X tales que $K \cap A = \emptyset$, entonces

$$d(K, A) = \operatorname{ínf}\{d(x, a) : x \in K, a \in A\} > 0,$$

y éste ínfimo se alcanza, es decir existe $x_0 \in K$ tal que $0 < d(x_0, A) \leq d(x, A)$ para todo $x \in K$. En particular se obtiene que, si A es un cerrado y B es un compacto de X, entonces

$$d(A, B) = 0 \iff A \cap B \neq \emptyset.$$

Problema 5.37. El problema anterior no es cierto en general si uno de los dos conjuntos no es compacto. Poner un ejemplo de dos cerrados disjuntos $A, B \subseteq \mathbb{R}^2$ tales que $d(A, B) = 0$.

Problema 5.38. Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función continua tal que existe

$$\operatorname{lím}_{x \to \infty} f(x)$$

y es finito. Probar que f es uniformemente continua en $\mathbb{R}$.

Problema 5.39. Sea E un espacio vectorial euclideo y $\langle \cdot, \cdot \rangle$ su producto escalar. Demostrar que $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{R}$ es continua, pero no uniformemente continua.

Problema 5.40. Estudiar si las funciones siguientes son uniformemente continuas o no en los dominios indicados.

1. $f : \mathbb{R} \to \mathbb{R}$, $f(t) = t \sin t$;
2. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = xe^y + e^{-y}$;
3. $f : \mathbb{R} \to \mathbb{R}^2$, $f(t) = \big(1/(1 + t^2), t/(1 + t^4)\big)$;

4. $f : [0,1] \to \mathbb{R}$, $f(t) = \sum_{n=1}^{\infty} \frac{\cos(nt)}{n^2}$.

Problema 5.41. Sea $f : \mathbb{R} \to \mathbb{R}$ continua y acotada. ¿Es f necesariamente uniformemente continua?

Problema 5.42. Probar que si $f, g : X \to \mathbb{R}$ son uniformemente continuas entonces su suma $f + g : X \to \mathbb{R}$ es uniformemente continua.

Problema 5.43. Demostrar que si $f : X \to Y$ es uniformemente continua y $g : Y \to Z$ es uniformemente continua entonces la composición $g \circ f : X \to Y$ es uniformemente continua.

Problema 5.44. Probar que el producto de dos funciones uniformemente continuas y acotadas es uniformemente continuo. Dar un ejemplo que pruebe que esto no es cierto en general si las funciones no son acotadas.

Problema 5.45. Demostrar que $f : X \to Y$ *no* es uniformemente continua si y sólo si existen $\varepsilon > 0$ y sucesiones $(x_n), (y_n) \subseteq X$ tales que $d(x_n, y_n) \leq 1/n$ y $d(f(x_n), f(y_n)) \geq \varepsilon$ para todo $n \in \mathbb{N}$.

Problema 5.46. Demostrar que $f : X \to \mathbb{R}^m$ es continua (respectivamente uniformemente continua) si y sólo si las funciones coordenadas $f_j : X \to \mathbb{R}$ son continuas (resp. uniformemente continuas) para todo $j = 1, ..., m$.

Problema 5.47. Sea X un espacio métrico compacto, y $f : X \to X$ una inyección isométrica (es decir, $d(f(x), f(y)) = d(x, y)$). Probar que f es también sobreyectiva. *Indicación:* si $x_0 \in X \setminus f(X)$, ver que existe $c > 0$ tal que $d(x, y) \geq c$ para todo $y \in f(X)$, y usar la sucesión definida por $x_1 = x_0$, $x_n = f(x_{n-1})$ si $n \geq 2$, para contradecir la compacidad de X.

Problema 5.48. Sea $f : A \subseteq \mathbb{R}^n \to Y$ una función uniformemente continua, donde Y es un espacio métrico completo. Probar que f está acotada sobre cada subconjunto acotado de A. *Indicación:* combinar los Teoremas 5.20 y 5.17.

Problema 5.49. Se dice que un subconjunto C de un espacio vectorial normado E es *convexo* si para todos $x, y \in C$, el segmento que une x con y, es decir $[x, y] = \{tx + (1-t)y : t \in [0,1]\}$ está contenido en C. Probar que si $f : C \subseteq E \to Y$ es uniformemente continua entonces f es acotada sobre cada subconjunto acotado de C.

Indicación: Fijar un punto c_0 y un número $R > 0$, y pruébese que f está acotada en $B(c_0, R) \cap C$ como sigue. Tómense $\varepsilon = 1$ y $\delta > 0$ tal que si $x, y \in C$ y $\|x - y\| \leq \delta$ entonces $d(f(x), f(y)) \leq 1$. Tomemos $N \in \mathbb{N}$ tal que $R/N \leq \delta$. Ahora, dado $x \in C \cap B(c_0, R)$, nótese que puede dividirse el segmento $[c_0, x] \subseteq C \cap B(c_0, R)$ en N subsegmentos contiguos $[x_{j-1}, x_j]$ de igual longitud (menor o igual que δ). Aplíquese ahora la propiedad de continuidad uniforme para deducir que $f(x) \in B(f(c_0), N)$ para todo $x \in C \cap B(c_0, R)$.

Problema 5.50. En el problema anterior es importante que C sea convexo. Probar con un ejemplo que si C no es convexo el resultado es falso en general.

Indicación: El contraejemplo *no* puede estar en $\mathbb{R}^n$, como indica el problema 5.48. Considérese $X = (C[0,1], \|\cdot\|_\infty)$ y constrúyase una sucesión de funciones $(f_n) \subset C[0,1]$ tal que $\|f_n\|_\infty \leq 1$ para todo n y $\|f_n - f_m\|_\infty \geq 1$ si $n \neq m$. Sea $A = \bigcup_{n=1}^{\infty} B(f_n, 1/3)$. El conjunto A es abierto y $\overline{A} = \bigcup_{n=1}^{\infty} \overline{B}(f_n, 1/3)$. Sea $C = \overline{A}$, y para cada $f \in C$ defínase $F(f) = n_f$, donde n_f es el único n tal que $f \in \overline{B}(f_n, 1/3)$. Probar que C es acotado, y que $F : C \subset C[0,1] \to \mathbb{R}$ es uniformemente continua pero no acotada.

Problema 5.51. Probar que si E es un espacio vectorial normado (o más en general un convexo en un espacio normado) y $f : E \to Y$ es uniformemente continua, entonces, para cada $\Delta > 0$ existe $L_\Delta > 0$ tal que si $\|x - y\| \geq \Delta$ entonces $d(f(x), f(y)) \leq L_\Delta \|x - y\|$. Este hecho se resume a veces diciendo que las funciones uniformemente continuas sobre convexos son *Lipschitz a grandes distancias*. *Indicación:* inspirarse en la demotración del problema 5.49.

Problema 5.52. En los siguientes casos estudiar la compacidad de A y de $\overline{f(A)}$, así como la continuidad uniforme de f en A.

1. $A = \{(x, y) \in \mathbb{R}^2 : y < x^2 \leq 4x - y^2\}$, $f(x, y) = \left(\frac{x^2 y}{x^2+y^2}, \frac{\sin(x^2+y^2)}{x^2+y^2}\right)$;
2. $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 2y, y > 0\}$, $f(x, y) = \left(\frac{y^2}{x^2+y^2}, \frac{xy}{x^2+y^2}\right)$;
3. $A = \{(x, y) \in \mathbb{R}^2 : 0 < x^2 + y^2 < r\}$, y $f(x, y) = \frac{\cos(x^2+y^2)-1}{\sin(x^2+y^2)}$, donde $r \in \{1, \pi\}$ (dos casos).

Problema 5.53. Estudiar la continuidad uniforme de las aplicaciones $f, g : A \to \mathbb{R}$, donde $A = \{(x, y) \in \mathbb{R}^2 : |y|(1 + x^2) \leq x\}$, y

$$f(x, y) = xy, \;\; g(x, y) = |x|^{1/2} y.$$

Problema 5.54. Sea E un espacio vectorial de dimensión infinita. Demostrar que existe $T : E \to \mathbb{R}$ lineal no continua. *Indicación:* Sea $\{e_i\}_{i \in I}$ una base algebraica infinita de E, y escojamos $(i_n) \subset I$ una sucesión de índices distintos de I. Defínase $T(e_i) = n$ si $i = i_n$ para algún $n \in \mathbb{N}$, y $T(e_i) = 0$ en otro caso, y extiéndase T a todo E por linealidad.

Problema 5.55. El espacio de las aplicaciones (formas) lineales de $\mathbb{R}^n$ sobre $\mathbb{R}$ puede identificarse al de las matrices de una fila y n columnas, que a su vez es obviamente identificable a $\mathbb{R}^n$. Utilizando esta identificación, si $T \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ es una forma lineal cuya matriz es $(a_1, ..., a_n)$, calcúlese la norma de T, es decir,

$$\|T\| = \sup\{T(x) : x \in \mathbb{R}^n, \|x\| \leq 1\},$$

en los casos en que en $\mathbb{R}^n$ se consideran las normas $\|\cdot\|_1$, $\|\cdot\|_2$ y $\|\cdot\|_\infty$.

Problema 5.56. Para cada vector $v \in \mathbb{R}^n$, probar que

$$T(x) = \langle x, v \rangle$$

define una forma lineal $T : \mathbb{R}^n \to \mathbb{R}$. Recíprocamente, probar que toda forma lineal en $\mathbb{R}^n$ es $\langle \cdot, v \rangle$ para algún $v \in \mathbb{R}^n$.

Problema 5.57. Probar que la función determinante $\det : \mathcal{M}_{n \times n} \to \mathcal{M}_{n \times n}$ es continua.

Problema 5.58. Probar que si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una aplicación lineal cuya matriz (a_{ij}) respecto de las bases canónicas satisface que $|a_{ij}| \leq M$ para todo i, j, entonces $\|T\| \leq \sqrt{nm} M$, donde la norma de T se supone definida respecto de las normas euclideas en $\mathbb{R}^n$ y $\mathbb{R}^m$.

Problema 5.59. Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ lineal. Probar que T es inyectiva si y sólo si existe $m > 0$ tal que

$$\|T(x)\| \geq m\|x\|$$

para todo $x \in \mathbb{R}^n$. *Indicación*: tener en cuenta que si T es inyectiva entonces $\rho(x) = \|T(x)\|$ define una norma en $\mathbb{R}^n$.

Capítulo 6

Conexión y convexidad

En este capítulo estudiaremos brevemente las nociones de conexión y conexión por caminos, así como las relaciones de éstas con la continuidad de funciones y otros conceptos ya expuestos. También estudiaremos algunas propiedades de los subconjuntos convexos de $\mathbb{R}^n$.

Intuitivamente puede decirse que un espacio es conexo cuando es *de una sola pieza*, o cuando no puede descomponerse en dos o más trozos *separados* unos de otros.

Definición 6.1. Se dice que un espacio métrico X es conexo si los únicos subconjuntos suyos que son a la vez abiertos y cerrados son $\emptyset$ y el propio X. Si X no es conexo entonces se dice que X es desconexo. Es evidente que esta definición puede reformularse de varias maneras. De hecho las siguientes afirmaciones son equivalentes:

1. X es desconexo.
2. Existen A_1, A_2 subconjuntos abiertos no vacíos de X tales que $X = A_1 \cup A_2$ y $A_1 \cap A_2 = \emptyset$.
3. Existen C_1, C_2 subconjuntos cerrados no vacíos de X tales que $X = C_1 \cup C_2$ y $C_1 \cap C_2 = \emptyset$.

Diremos que un subconjunto Y de un espacio métrico (X, d) es conexo si el subespacio métrico (Y, d) es conexo. Teniendo en cuenta la caracterización de cerrados y abiertos relativos a un subespacio métrico, es claro que, para cualquier subconjunto Y de X, las siguientes afirmaciones son equivalentes:

1. Y es desconexo en X.
2. Existen abiertos G_1, G_2 de X tales que $Y \subseteq G_1 \cup G_2$, $G_1 \cap G_2 \cap Y = \emptyset$, $G_1 \cap Y \neq \emptyset \neq G_2 \cap Y$.
3. Existen cerrados F_1, F_2 de X tales que $Y \subseteq F_1 \cup F_2$, $F_1 \cap F_2 \cap Y = \emptyset$, $F_1 \cap Y \neq \emptyset \neq F_2 \cap Y$.

Es también claro que si $Z \subseteq Y \subseteq X$ entonces Z es conexo en Y si y sólo si Z es conexo en X. Así pues, la conexión de un subespacio métrico no depende del superespacio que se desee considerar.

Por ejemplo, es evidente que $X = [0, 1] \cup [2, 3]$ es desconexo en $\mathbb{R}$. Quizás pueda resultar menos obvio, aunque es verdad, que cualquier intervalo de $\mathbb{R}$ es un espacio conexo, y de hecho éstos son los únicos subconjuntos conexos de $\mathbb{R}$.

Teorema 6.1. *Un subconjunto A de $\mathbb{R}$ es conexo en $(\mathbb{R}, |\cdot|)$ si y sólo si A es un intervalo.*

Demostración. ($\Rightarrow$): Si A no es un intervalo, existen $a, b, c \in \mathbb{R}$ tales que $a < b < c$, $a, c \in A$, pero $b \notin A$. Entonces tenemos que

$$A \subseteq (-\infty, b) \cup (b, +\infty),$$

y la intersección de A con cada uno de estos dos intervalos abiertos es no vacía, luego A es desconexo.
($\Leftarrow$): Supongamos ahora que A es un intervalo. Si A fuera desconexo existirían G_1, G_2 abiertos de $\mathbb{R}$ y $a_i \in G_i$, $i = 1, 2$, tales que $A \subseteq G_1 \cup G_2$, $a_i \in G_i \cap A$, y $A \cap G_1 \cap G_2 = \emptyset$. Definamos la función $f : A \to \mathbb{R}$ por

$$f(x) = \begin{cases} 1 & \text{si } x \in G_1 \cap A, \\ -1 & \text{si } x \in G_2 \cap A. \end{cases}$$

Para cada abierto G de $\mathbb{R}$ se tiene que

$$f^{-1}(G) = \begin{cases} G_1 \cap A & \text{si } 1 \in G, -1 \notin G, \\ G_2 \cap A & \text{si } 1 \notin G, -1 \in G, \\ A & \text{si } 1, -1 \in G, \end{cases}$$

y en todo caso es claro que $f^{-1}(G)$ es abierto en A. Por tanto $f : A \to \mathbb{R}$ es continua. Pero $f(a_1) = 1 > 0 > -1 = f(a_2)$, $a_1, a_2 \in A$, A es un intervalo y f no toma el valor 0, lo cual contradice el teorema de Bolzano. □

Al igual que ocurre con la compacidad, la imagen continua de un espacio conexo es conexa.

Teorema 6.2. *Sea $f : X \to Y$ una aplicación continua, y supongamos que X es conexo. Entonces $f(X)$ es conexo en Y. En particular, si f es además sobreyectiva entonces Y es conexo.*

Demostración. Podemos suponer $Y = f(X)$, y que f es sobreyectiva, puesto que f es continua de X en $f(X)$. Si $Y = f(X)$ fuera desconexo, tendríamos que $Y = G_1 \cup G_2$, con los G_i abiertos no vacíos y disjuntos. Pongamos $A_i = f^{-1}(G_i)$, con $i = 1, 2$. Entonces es claro que los A_i son abiertos (por ser f continua), no vacíos, y disjuntos, y $X = A_1 \cup A_2$, luego X sería desconexo. □

Corolario 6.3. *Si X e Y son homeomorfos entonces X es conexo si y sólo si Y es conexo.*

Veamos algunos ejemplos de cómo puede aplicarse el Teorema 6.2.

Ejemplo 6.1. La circunferencia unidad $C = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ es un subconjunto conexo de $\mathbb{R}^2$. En efecto, como $\mathbb{R}$ es conexo (por ser un intervalo) y la aplicación $f : \mathbb{R} \to C$ definida por

$$f(t) = (\cos t, \sin t)$$

es continua y sobreyectiva, se deduce que C es conexo.

Análogamente puede demostrarse que una recta en $\mathbb{R}^n$ es conexa, o que, más en general, la imagen de cualquier curva continua en un espacio métrico es conexa.

Por otra parte, para cualquier función continua $f : X \to Y$, se tiene que X es conexo si y sólo si la gráfica de f, $G_f = \{(x, y) \in X \times Y : y = f(x)\}$, es conexo en $X \times Y$. En efecto, basta tener en cuenta que la función $x \mapsto (x, f(x))$ es continua de X en $X \times Y$, y que la proyección $P : X \times Y \to X$ definida por $P(x, y) = x$ es también continua.

El siguiente resultado nos proporciona una caracterización bastante útil de los espacios conexos en términos de funciones continuas.

Proposición 6.4. *Un espacio métrico X es desconexo si y sólo si existe una función continua y sobreyectiva $f : X \to \{0,1\}$.*
Por tanto X es conexo si y sólo si para toda función continua $f : X \to \{0,1\}$ se tiene que o bien $f \equiv 0$ o bien $f \equiv 1$ (es decir que f es constante).

Demostración. $(\Rightarrow)$: Si X es desconexo existen A_1, A_2 abiertos no vacíos y disjuntos tales que $X = A_1 \cup A_2$. Definamos $f : X \to \{0,1\}$ por

$$f(x) = \begin{cases} 1 & \text{si } x \in A_1, \\ 0 & \text{si } x \in A_2. \end{cases}$$

Es claro que f es continua (por ser localmente constante) y sobreyectiva.
$(\Leftarrow)$: Recíprocamente, si existe $f : X \to \{0,1\}$ continua y sobreyectiva, entonces, definiendo $C_1 = f^{-1}(1), C_2 = f^{-1}(0)$, obtenemos una partición $X = C_1 \cup C_2$ en cerrados disjuntos no vacíos. $\square$

El siguiente teorema es la extensión natural del teorema de Bolzano a los espacios métricos conexos.

Teorema 6.5. *Sea X un espacio métrixo conexo, y $f : X \to \mathbb{R}$ continua. Supongamos que existen $x, y \in X$ $c \in \mathbb{R}$ tales que $f(x) < c < f(y)$. Entonces existe $z \in X$ tal que $f(z) = c$.*

Demostración. Como X es conexo y f es continua, resulta que $f(X)$ es conexo en $\mathbb{R}$, es decir, un intervalo. Entonces, puesto que $f(x), f(y) \in f(X)$ y $f(x) < c < f(y)$, se concluye que $c \in f(X)$. $\square$

El recíproco de este teorema es, obviamente, también cierto teniendo en cuenta la Proposición anterior: si X es un espacio métrico tal que toda función continua tiene la propiedad del enunciado, entonces X es conexo.

Las siguientes propiedades de los conjuntos conexos resultarán bastante provechosas.

Proposición 6.6. *Sea X un espacio métrico, sea $\mathcal{F} = \{C_i\}_{i\in I}$ una familia de subconjuntos conexos de X.*

1. *Supongamos que $\bigcap_{i\in I} C_i \neq \emptyset$. Entonces $\bigcup_{i\in I} C_i$ es conexo.*

2. *Supongamos que Y es un subconjunto conexo de X tal que $Y \cap C_i \neq \emptyset$ para todo $i \in I$. Entonces $Y \cup \bigcup_{i\in I} C_i$ es conexo.*

Demostración. (1): Por la Proposición 6.4, basta ver que toda función continua $f : \bigcup_{i\in I} C_i \to \{0,1\}$ es constante. Sea $y_0 \in \bigcap_{i\in I} C_i$. Para cada $i \in I$, como C_i es conexo y $f_{|C_i} : C_i \to \{0,1\}$ es continua, se tiene que $f_{|C_i}(y) = f(y_0)$ para todo $y \in C_i$. Por tanto $f(y) = f(y_0)$ para todo $y \in \bigcup_{i\in I} C_i$, es decir, f es constante.
(2): Basta aplicar (1) a la familia $\mathcal{F}' = \{Y \cup C_i\}_{i\in I}$, cuya intersección es no vacía porque contiene Y, y cuyos miembros son conexos otra vez gracias a (1) (puesto que lo son C_i e Y, que tienen intersección no vacía para cada $i \in I$). $\square$

Así, por ejemplo, se tiene que los conjuntos $\{(x,y) \in \mathbb{R}^2 : y = qx, q \in \mathbb{Q}\}$ y $\{(x,y) \in \mathbb{R}^2 : x \in \mathbb{Z}\} \cup \{(x,y) \in \mathbb{R}^2 : y = x\}$ son conexos.

A continuación vemos otra aplicación interesante de esta proposición: el hecho de que el producto cartesiano de dos espacios conexos es asimismo conexo.

Proposición 6.7. *Sean X e Y espacios métricos conexos. Entonces $X \times Y$ es conexo.*

Demostración. El producto es vacío, luego conexo, si uno de los factores es vacío. Podemos suponer pues que existen $a \in X, b \in Y$. La aplicación $x \mapsto (x, b)$ es claramente un homeomorfismo de X sobre el subespacio métrico $X \times \{b\}$ de $X \times Y$, y por tanto $X \times \{b\}$ es conexo. Análogamente, $\{x\} \times Y$ es conexo para todo $x \in X$. Entonces el subconjunto

$$H_x = \big(X \times \{b\}\big) \cup \big(\{x\} \times Y\big)$$

de $X \times Y$, siendo unión de dos conexos con el punto (x, b) en común, es conexo, y contiene el punto (a, b) para todo $x \in X$. Por tanto deducimos que

$$X \times Y = \bigcup_{x \in X} H_x$$

es conexo. □

Por otra parte, si a un conjunto conexo le añadimos puntos de su adherencia, el conjunto resultante sigue siendo conexo.

Proposición 6.8. *Si A es un subconjunto conexo de X y $A \subseteq B \subseteq \overline{A}$, entonces B también es conexo. En particular la adherencia de un conjunto conexo es conexa.*

Demostración. Sea $f : B \to \{0, 1\}$ una función continua. Como $f_{|A} : A \to 0, 1$ es continua y A es conexa, $f_{|A}$ es constante, digamos $f_{|A} = 1$. Entonces, si $x \in B$, existe $(x_n) \subseteq A$ tal que $x_n \to x$. Así $f(x_n) = 1$ para todo n, y como f es continua se tiene que $f(x) = \lim_{n\to\infty} f(x_n) = 1$. Luego f es constante en B. □

Un espacio métrico puede no ser conexo, pero siempre podrá escribirse como unión de subconjuntos conexos, a los cuales llamaremos componentes conexas.

Definición 6.2. Se dice que C es una componente conexa de un subconjunto A de un espacio métrico si C es conexo y no existe ningún subconjunto conexo propio de A que contenga estrictamente a C.

Proposición 6.9. *Sea A un subconjunto de un espacio métrico X. Entonces:*

1. *Cada componente conexa de A es un cerrado relativo a A.*
2. *A es unión disjunta de sus componentes conexas.*

Demostración. (1): Sea C una componente conexa de A. Como $C \subseteq \overline{C} \cap A \subseteq \overline{C}$, por la Proposición 6.8 tenemos que $\overline{C} \cap A$ es conexo, y contiene a C, luego $C = \overline{C} \cap A$ y C es cerrado relativo a A.
(2): Para cada punto $x \in A$, sea $\{C_i^x\}_{i \in I_x}$ la familia de todos los conjuntos conexos que contienen a x. Esta familia es no vacía ya que $\{x\}$ es conexo. Por la Proposición 6.6 se tiene que

$$A_x = \bigcup_{i \in I_x} C_i^x$$

es conexo, y evidentemente es el mayor conexo que contiene a x, luego A_x es componente conexa de A. Puesto que $x \in A_x$ para todo $x \in A$, es obvio también que

$$A = \bigcup_{x \in A} A_x.$$

Además, para todos $x, y \in A$, o bien es $A_x \cap A_y = \emptyset$ o bien $A_x = A_y$. En efecto, si $A_x \cap A_y \neq \emptyset$ entonces, al ser A_x y A_y conexos con intersección no vacía, su unión $A_x \cup A_y$ es conexa, y por maximalidad de A_x y A_y se concluye que $A_x = A_x \cup A_y = A_y$. Por tanto A es unión disjunta de sus componentes conexas. $\square$

Observación 6.10. *Las componentes conexas de un conjunto en general* no *son abiertos relativos a ese conjunto. Por ejemplo, si $A = \mathbb{Q}$ entonces las componentes conexas de A son todos los puntos $\{r\}$, con $r \in \mathbb{Q}$, que son cerrados, pero no abiertos relativos a $\mathbb{Q}$.*

Sin embargo, cuando A es un abierto de $\mathbb{R}^n$, sus componentes conexas sí son abiertos.

Proposición 6.11. *Si A es un subconjunto abierto de un espacio vectorial normado entonces todas las componentes conexas de A son abiertos.*

Demostración. Sea C una componente conexa de A, y $x \in C$. Como A es abierto existe $r > 0$ tal que $B(x, r) \subseteq A$. Ahora bien, $B(x, r)$ es conexo (ver el ejercicio 6.5), y como C es el mayor conexo que contiene a x, se concluye que $B(x, r) \subseteq C$. $\square$

En la demostración anterior hemos utilizado el hecho de que las bolas de un espacio normado son conjuntos conexos. La manera más fácil de probar esto es a través de la noción de conexión por caminos, que pasamos a estudiar ahora.

Definición 6.3. Se dice que un espacio métrico X es conexo por caminos si para cada $x, y \in X$ existe un camino continuo que une x con y (es decir, existe una aplicación continua $\gamma : I \to X$, donde $I = [a, b]$ es un intervalo de $\mathbb{R}$, tal que $\gamma(a) = x$ y $\gamma(b) = y$).

Se dice que un subconjunto A de (X, d) es conexo por caminos si el subespacio métrico (A, d) es conexo, lo cual equivale a decir que para cada $x, y \in A$ existe $\gamma : I \to X$ continua $\gamma(a) = x$, $\gamma(b) = y$ y $\gamma(t) \in A$ para todo $t \in I$.

Es inmediato ver que siempre puede suponerse $[a, b] = [0, 1]$.

Por ejemplo, cualquier espacio vectorial normado es conexo por caminos (dados dos puntos basta considerar el segmento que los une). Lo mismo ocurre con cualquier bola de un espacio vectorial normado (y más en general, como veremos después, con los conjuntos llamados convexos, que son precisamente los que tienen la propiedad de que el segmento que une dos cualesquiera de sus puntos siempre está incluido en el conjunto).

Al igual que ocurre en el caso de los conjuntos conexos, la imagen continua de un conjunto conexo por caminos es conexa por caminos.

Proposición 6.12. *Si X es conexo por caminos y $f : X \to Y$ es continua entonces $f(X)$ es conexo por caminos en Y.*

Demostración. Sean $f(x), f(y) \in f(X)$. Como X es conexo por caminos existe un camino continuo $\gamma : [a, b] \to X$ tal que $\gamma(a) = x$, $\gamma(b) = y$. Entonces $f \circ \gamma : [a, b] \to f(X)$ es un camino continuo que une $f(x)$ con $f(y)$ en $f(X)$. $\square$

Veamos ahora cuáles son las relaciones que hay entre la noción de conexión por caminos y la de conexión.

Proposición 6.13. *Si X es conexo por caminos entonces X es conexo.*

Demostración. Sea $f : X \to \{0, 1\}$ continua, y veamos que f es constante. Elijamos $x_0, x \in X$. Como X es conexo por caminos existe $\gamma : [0, 1] \to X$ continuo tal que $f(0) = x_0$, y $f(1) = x$. Como $f \circ \gamma : [0, 1] \to \{0, 1\}$ es continua y $[0, 1]$ es conexo, $f \circ \gamma$ debe ser constante, es decir, $f(x_0) = f(\gamma(0)) = f(\gamma(1)) = f(x)$. Así $f(x) = f(x_0)$ para todo $x \in X$, es decir, f es constante. □

Sin embargo el recíproco no es cierto: existen espacios que son conexos pero no conexos por caminos.

Ejemplo 6.2. Sea $X = \{(x, y) \in \mathbb{R}^2 : y = \sin(1/x), 0 < x \leq 1\} \cup \{(0, 0)\}$. Entonces X es conexo, pero no es conexo por caminos en $\mathbb{R}^2$.

Demostración. Si llamamos $A = \{(x, y) \in \mathbb{R}^2 : y = \sin x, 0 < x \leq 1\}$, vemos que A es conexo (de hecho conexo por caminos) puesto que $A = f((0, 1])$, donde $f(t) = (t, \sin(1/t))$ es continua de $(0, 1]$ en $\mathbb{R}^2$, y el intervalo semiabierto $(0, 1]$ es conexo. Entonces, por la Proposición 6.8, $X = A \cup \{(0, 0)\}$ es conexo, ya que

$$A \subset X \subset \overline{A} = \{(x, y) : y = \sin(1/x), 0 < x \leq 1\} \cup \{(x, y) : x = 0, -1 \leq y \leq 1\}.$$

Sin embargo X no es conexo por caminos. En efecto, supongamos que exista $\gamma : [0, 1] \to A$ continuo tal que $\gamma(0) = (0, 0)$ y $\gamma(1) = (1, \sin 1)$. El camino γ debe ser sobreyectivo, ya que de lo contrario existiría $x_0 \in (0, 1)$ tal que $\gamma(t) \neq (x_0, \sin(1/x_0))$ para todo $t \in [0, 1]$, y así los abiertos disjuntos

$$G_1 = \{(x, y) : x < x_0\},\ G_2 = \{(x, y) : x > x_0\}$$

separarían $\gamma[0, 1]$, con lo que $\gamma[0, 1]$ no sería conexo. Entonces $X = \gamma([0, 1])$, y como $[0, 1]$ es compacto y γ es continua, X debe ser compacto, lo que es absurdo, pues X no es cerrado en $\mathbb{R}^2$. □

Observación 6.14. El mismo argumento probaría que el conjunto

$$Y = \{(x, y) \in \mathbb{R}^2 : y = \sin(1/x), 0 < x \leq 1\} \cup \{(x_0, y_0)\}$$

tampoco es conexo por caminos, cualquiera que sea $(x_0, y_0) \in \mathbb{R}^2$ con $x_0 = 0$.

Podría pensarse que la razón por la que X no es conexo por caminos es que $X \neq \overline{A}$, pero que si añadimos a A los puntos de su adherencia (haciéndolo así compacto) tal vez el conjunto seguiría siendo conexo por caminos. Esto no es así, como vemos a continuación.

Ejemplo 6.3. El conjunto $Z = \{(x, y) : y = \sin(1/x), 0 < x \leq 1\} \cup \{(x, y) : x = 0, -1 \leq y \leq 1\}$ es conexo, pero no es conexo por caminos.

Supongamos que lo fuera, y sea $\gamma : [0, 1] \to Z$ un camino continuo tal que $\gamma(0) = (0, 0)$ y $\gamma(1) = (1, \sin 1)$. Denotemos $\gamma(t) = (\gamma_1(t), \gamma_2(t))$, y definamos

$$t_0 = \sup\{t \in [0, 1] : \gamma_1(t) = 0\},\ t_1 = 1,\ \text{ y } \beta = \gamma|_{[t_0, t_1]}$$

Entonces $\beta : [t_0, t_1] \to Z$ es un camino continuo que une $\beta(t_0) = (x_0, y_0)$ con $(1, \sin 1)$, y además, por continuidad de γ se tiene que $x_0 = \beta_1(t_0) = \gamma_1(t_0) = 0$. Por otra parte, como en el ejemplo anterior, se ve fácilmente que para todo $(x, y) \in Z$ con $x > 0$ existe $t \in (0, 1]$ tal que $\gamma(t) = (x, y)$. Así, el conjunto

$$\beta([t_0, t_1]) = \{(x, y) \in \mathbb{R}^2 : y = \sin(1/x), 0 < x \leq 1\} \cup \{(x_0, y_0)\}$$

es conexo por caminos (al ser imagen continua de un intervalo), lo cual contradice la observación anterior. □

Sin embargo, vamos a probar que para todo conjunto abierto de un espacio normado las nociones de conexión y de conexión por caminos sí son equivalentes. De hecho veremos que son también equivalentes a la de conexión por poligonales. Si x, y son dos puntos de un espacio normado E, denotaremos el segmento de recta que une x con y por $[x, y]$, es decir,

$$[x, y] = \{(1-t)x + ty : 0 \leq t \leq 1\}.$$

Diremos que $\gamma : [a, b] \to E$ es un camino poligonal si γ es continuo y existen puntos $x_0, x_1, ..., x_n \in E$ tales que

$$\gamma([a, b]) = [x_0, x_1] \cup ... \cup [x_{n-1}, x_n],$$

con $\gamma(a) = x_0$ y $\gamma(b) = x_n$.

Definición 6.4. Se dice que un conjunto A de un espacio normado es conexo por poligonales si para todo $x, y \in A$ existe un camino poligonal γ que une x con y. Equivalentemente, si para todo $x, y \in A$ existen $x_0, x_1, ..., x_n \in A$ tales que $x_0 = x$, $x_n = y$, y además

$$[x_0, x_1] \cup [x_1, x_2] \cup ... \cup [x_{n-1}, x_n] \subseteq A.$$

Teorema 6.15. *Sea A un subconjunto* abierto *de un espacio normado E. Las siguientes afirmaciones son equivalentes:*

1. *A es conexo.*

2. *A es conexo por caminos.*

3. *A es conexo por poligonales.*

Demostración. (3) $\implies$ (2) es evidente, y (2) $\implies$ (1) ya lo sabemos. Probemos que (1) $\implies$ (3). Fijemos un punto $x_0 \in A$. Queremos probar que para todo $x \in A$ existe una poligonal que une x_0 con x dentro de A. Para ello consideremos el conjunto

$$G = \{y \in A : \text{ existe una poligonal que une } x_0 \text{ con } x \text{ dentro de } A\}.$$

Por un lado, se tiene que G es abierto. En efecto, si $y \in G$, al ser A abierto existe $r > 0$ tal que $B(y, r) \subseteq A$, y como $y \in A$ existen $x_1, ..., x_n$ tales que

$$[x_0, x_1] \cup ... \cup [x_{n-1}, x_n] \subset A,$$

con $x_n = y$. Sea ahora $z \in B(y, r)$. Como $B(y, r)$ es convexo (ejercicio 6.5) se tiene que $[y, z] \subset B(y, r) \subseteq A$, y por tanto

$$[x_0, x_1] \cup ... \cup [x_{n-1}, y] \cup [y, z] \subset A,$$

lo cual indica que $z \in G$. Por tanto $B(y, r) \subseteq G$.

Por otro lado, $A \setminus G$ también es abierto (es decir G es un cerrado relativo de A). En efecto, si $x \in A \setminus G$, al ser A abierto podemos tomar $r > 0$ tal que $B(x, r) \subset A$. Entonces $B(x, r) \subset A \setminus G$, ya que si existiera $y \in G \cap B(x, r)$ entonces podríamos encontrar $x_1, ..., x_{n-1}$ tales que

$$[x_0, x_1] \cup ... \cup [x_{n-1}, y] \subset A,$$

y como $[y,x] \subset B(x,r) \subseteq A$ también tendríamos

$$[x_0, x_1] \cup ... \cup [x_{n-1}, y] \cup [y, x] \subset A,$$

es decir, $x \in G$, lo que es absurdo. Esto prueba que $A \setminus G$ es abierto.

Ahora, como $\emptyset \neq G \subseteq A$, G es cerrado y abierto de A, y A es conexo, se deduce que $G = A$. Esto significa que todo punto $x \in A$ puede unirse a x_0 mediante una poligonal dentro de A. $\square$

Conviene notar el papel fundamental que la convexidad de las bolas en el espacio normado E ha jugado en la demostración anterior (así como en la de la Proposición 6.11). La parte final de este capítulo la dedicaremos a hacer un breve estudio de las nociones de conjunto convexo y de función convexa.

Definición 6.5. Sea E un espacio vectorial normado. Diremos que un subconjunto C de E es convexo si para todo $x, y \in C$, el segmento que une x con y está contenido en C, es decir, si $(1-t)x + ty \in C$ para todo $t \in [0,1]$.

Por ejemplo, como ya hemos señalado (ejercicio 6.5), cualquier bola en un espacio normado E es un conjunto convexo. También lo es, obviamente, cualquier subespacio afín de E.

Proposición 6.16. *Si C es convexo entonces C es conexo por caminos.*

Demostración. Si $x, y \in C$, basta tener en cuenta que la aplicación $\gamma(t) = (1-t)x + ty$, $t \in [0,1]$, es continua. $\square$

El recíproco es falso en general. Una circunferencia es por ejemplo un subconjunto conexo por caminos pero no convexo.

Definición 6.6. Sea D un subconjunto convexo de un espacio normado E. Se dice que una función $f : D \to \mathbb{R}$ es convexa si para cada $x, y \in D$ y cada $t \in [0,1]$ se tiene que $f\big(tx + (1-t)y\big) \leq tf(x) + (1-t)f(y)$.

Es fácil ver que $f : D \to \mathbb{R}$ es convexa si y sólo si su epigrafo $\text{epi}(f) := \{(x,y) \in D \times \mathbb{R} : y \geq f(x)\}$ es convexo en $E \times \mathbb{R}$ (ejercicio 6.21).

A continuación veremos que toda función convexa es continua (de hecho localmente Lipschitz) si y sólo si es localmente acotada (se dice que una función $f : D \to \mathbb{R}$ es localmente acotada si para todo $x \in D$ existe $r > 0$ tal que f está acotada en $D \cap B(x,r)$).

Proposición 6.17. *Sea $f : D \to \mathbb{R}$ convexa, y supongamos que f es localmente acotada. Entonces f es localmente Lipschitz, es decir, para cada $z \in D$ existen $K, r > 0$ tal que si $x, y \in D \cap B(z,r)$ entonces*

$$|f(x) - f(y)| \leq K\|x - y\|.$$

En particular f es continua en D.

Demostración. En el argumento que sigue podemos suponer que todas las bolas que aparecen están contenidas en D; si no fuera así bastaría tomar su intersección con D, teniendo en cuenta que la única propiedad (no específicamente métrica) de las bolas que se usa es la convexidad. Fijemos $p \in D$. Por hipótesis existe $r > 0$ tal que f es acotada en $B(p,r)$. Sean $M = \sup\{f(x) : x \in B(p, 2r)\}$, y $m = \inf\{f(x) : x \in B(p, 2r)\}$. Vamos a ver que f es

K-Lispchitz en la bola $B(p,r)$, donde $K=(M-m)/r$. En efecto, tomemos $x,y\in B(p,r)$. Sea $\gamma:\mathbb{R}\to E$ definida por

$$\gamma(t)=x+\frac{t}{\|y-x\|}(y-x),$$

de modo que $\gamma(0)=x, \gamma(\|x-y\|)=y$. Definamos $t_1=0$, $t_2=\|y-x\|$, $t_3=t_1-r$, $t_4=t_2+r$. Se tiene que $\gamma(t_i)\in B(p,2r)$ para $i=1,2,3,4$, $\gamma(t_1)=x$, y $\gamma(t_2)=y$. Si definimos $g:[t_3,t_4]\to\mathbb{R}$ por

$$g(t)=f(\gamma(t))=f\left(x+\frac{t}{\|y-x\|}(y-x)\right),$$

resulta que g es una función convexa de una variable real, y por tanto, como $t_3<t_1<t_2<t_4$, se verifica que

$$\begin{gathered}-\frac{(M-m)}{r}\leq\frac{g(t_1)-g(t_3)}{t_1-t_3}\leq\frac{g(t_2)-g(t_1)}{t_2-t_1}=\\ \frac{f(y)-f(x)}{\|y-x\|}\leq\frac{g(t_4)-g(t_2)}{t_4-t_2}\leq\frac{M-m}{r},\end{gathered}$$

luego $|f(x)-f(y)|\leq K\|x-y\|$, donde $K=(M-m)/r$. □

En el problema 6.28 se indica cómo deducir de aquí que toda función convexa $f:\mathbb{R}^n\to\mathbb{R}$ es continua. Esto no es cierto en espacios normados de dimensión infinita: toda forma lineal es convexa, pero en tales espacios existen siempre aplicaciones lineales no continuas (ver el ejercicio 5.54).

A continuación vamos a probar que si tenemos dos convexos cerrados y disjuntos de $\mathbb{R}^n$, y uno de ellos es compacto, entonces pueden separarse por un hiperplano, en el sentido de que uno de los semiespacios cerrados delimitados por el hiperplano contiene a uno de los convexos, y el otro convexo queda fuera de este semiespacio.

Recordemos que se define un hiperplano afín en $\mathbb{R}^n$ como un subespacio afín de codimensión uno de $\mathbb{R}^n$. Equivalentemente, H es un hiperplano afín si y sólo si

$$H=\{x\in\mathbb{R}^n : f(x)=\alpha\},$$

donde $f:\mathbb{R}^n\to\mathbb{R}$ es una aplicación lineal y $\alpha\in\mathbb{R}$. Un semiespacio cerrado es cualquier conjunto de la forma

$$\{x\in\mathbb{R}^n : f(x)\leq\alpha\},$$

y un semiespacio abierto es

$$\{x\in\mathbb{R}^n : f(x)<\alpha\},$$

donde f,α son como antes.

Teorema 6.18. *Sean C,K conjuntos convexos y cerrados disjuntos de $\mathbb{R}^n$, y supongamos que K es compacto. Entonces existe un hiperplano afín H de $\mathbb{R}^n$ que separa C y K. Es decir, existen $f:\mathbb{R}^n\to\mathbb{R}$ forma lineal, y $\alpha\in\mathbb{R}$ tales que*

$$f(z)\leq\alpha<f(x)$$

para todos $x\in K$, $z\in C$.

Nota: Este resultado es un caso particular del llamado Teorema de Hahn-Banach (y que es válido incluso en espacios normados de dimensión infinita, lo cual no demostraremos aquí).

Demostración. Como C y K son cerrados disjuntos y K es compacto, sabemos que $0 < d(K,C) = \inf\{d(x,C) : x \in K\}$, y que este ínfimo se alcanza: existe $x_0 \in K$ tal que $d(x_0,C) = d(K,C)$ (ver el problema 5.36). Si C es acotado entonces es compacto, y $d(x_0,C)$ también se alcanza. Si C no es acotado entonces podemos considerar $D = C \cap \overline{B}(x_0, d(K,C)+1)$, que es compacto, y se tiene que existe $c_0 \in D$ tal que $d(x_0,c_0) = d(x_0,D)$. Por otro lado es claro que

$$d(x_0,D) = d(x_0,C).$$

Por tanto tenemos que $d(x_0,c_0) = d(K,C)$.

Sea ahora $f : \mathbb{R}^n \to \mathbb{R}$ la forma lineal definida por

$$f(x) = \langle x, x_0 - c_0 \rangle.$$

Definamos $H := \{x \in \mathbb{R}^n : f(x) = \alpha\}$, donde $\alpha := f(c_0)$. El hiperplano H es perpendicular al vector $v := x_0 - c_0$. Vamos a ver que H separa los convexos K y C. Más precisamente, probaremos que

$$f(x) > \alpha \geq f(z)$$

para todos $x \in K$, $z \in C$.

Veamos primero que si $z \in C$ entonces $f(z) \leq \alpha$. Supongamos por el contrario que fuera $f(z) > \alpha = f(c_0)$. El punto z no puede estar en la recta $\mathcal{R}$ que pasa por x_0 y c_0: si estuviera en esta recta, como $f(c_0) < f(x_0)$ y $f(c_0) < f(z)$, el punto z debería quedar en la semirrecta $\mathcal{R} \cap \{x : f(x) > f(c_0)\}$; de hecho $z \in (c_0, x_0]$ (ya que si $f(z) > f(x_0)$ entonces $x_0 \in [c_0, z] \subseteq C$, lo que es absurdo), pero entonces $d(K,C) \leq d(x_0,z) < d(x_0,c_0) = d(K,C)$, una contradicción. Consideremos ahora el plano $\mathcal{P}$ determinado por los puntos x_0, c_0, z. En este plano, la recta $\mathcal{R}$ es perpendicular al la recta $\mathcal{Q} := \mathcal{P} \cap \{x : f(x) = \alpha\}$, y a su vez esta recta es tangente en el punto c_0 a la circunferencia $\mathcal{C}$ de centro x_0 y radio $r = d(x_0,c_0)$, mientras que el punto z queda en el semiplano $\mathcal{S} := \{x \in \mathcal{P} : f(x) > \alpha\}$, el mismo semiplano en el que está x_0. Como C es convexo, el segmento $(c_0, z]$ está en $C \cap \mathcal{S}$. Y como $r = d(x_0,c_0) = d(K,C)$, ningún punto del segmento $(c_0,z]$ puede estar dentro de la circunferencia $\mathcal{C}$. Pero esta situación geométrica es imposible. En efecto, traduciendo la situación a $\mathbb{R}^2$, es decir, identificando $c_0 = (0,0)$, $x_0 = (0,r)$, $\mathcal{R} = \{(x,y) : x = 0\}$, $\mathcal{Q} = \{(x,y) : y = 0\}$, $\mathcal{S} = \{(x,y) : y > 0\}$, $\mathcal{C} = \{(x,y) : x^2 + (y-r)^2 = r^2\}$, $z = (x,y)$ con $y > 0$, los puntos $tz \in (0,z]$ deben quedar dentro del círculo delimitado por $\mathcal{C}$ si $t > 0$ es suficientemente pequeño (problema 6.30).

De igual manera (la situación es simétrica) puede probarse que $f(x) \geq f(x_0) > f(c_0) = \alpha$ para todo $x \in K$. □

Corolario 6.19. *Si C es un convexo cerrado de $\mathbb{R}^n$ y $x \in \mathbb{R}^n \setminus C$, existen una forma lineal $f : \mathbb{R}^n \to \mathbb{R}$ y un número $\alpha \in \mathbb{R}$ tales que*

$$f(z) \leq \alpha < f(x)$$

para todo $z \in C$.

Corolario 6.20. *Todo convexo cerrado de $\mathbb{R}^n$ es la intersección de todos los semiespacios cerrados que lo contienen.*

Demostración. Sea C convexo cerrado de $\mathbb{R}^n$. Es evidente que C está contenido en la intersección de todos los semiespacios que lo contienen. Recíprocamente, si un punto x está en esta intersección, entonces debe estar en C: de lo contrario, por el corolario anterior existiría un semiespacio cerrado que contendría a C y no a x. □

Corolario 6.21. *Un cerrado de $\mathbb{R}^n$ es convexo si y sólo si C es intersección de semiespacios cerrados.*

Demostración. ($\Rightarrow$): es el corolario anterior.
($\Leftarrow$): Todo semiespacio es convexo, y la intersección de convexos es convexo (problema 6.31). $\square$

6.1. Problemas

Problema 6.1. En los siguientes casos, estúdiese si M y $M \setminus \{(0,0)\}$ son conexos, y si no lo son determínese sus componentes conexas:

1. $M = \{(x,y) : x^2 + (y+1)^2 \leq 1\} \cup \{(0,y) : 0 \leq y < 1\}$
2. $M = \{(x,y) : x \in \mathbb{Q}, 0 < y \leq 1\} \cup \{(x,y) : x \notin \mathbb{Q}, -1 \leq y < 0\} \cup \mathbb{R} \times \{1\} \cup \mathbb{R} \times \{-1\} \cup \{(0,0)\}$
3. $M = \{(x,y) \in \mathbb{R}^2 : y = \alpha x,\ x \in \mathbb{R},\ \alpha \in \mathbb{R} \setminus \mathbb{Q}\}$

Problema 6.2. Estúdiese si son conexos o conexos por caminos los conjuntos:

1. $\mathbb{R}^2 \setminus M$, donde M es numerable.
2. $\{(x,y) : 1 < 4x^2 + 9y^2 < 9; x \neq \frac{1}{n}, n \in \mathbb{N}, n > 1\} \cup \{(x,y) : 1 = 4x^2 + 9y^2\} \subseteq \mathbb{R}^2$
3. $\bigcup_{n=1}^{\infty} \{(x,y) : y = \frac{x}{n},\ 0 \leq x \leq 1\} \cup \{(\frac{1}{2}, 0)\}$

Problema 6.3. Pruébese que $A \subset \mathbb{R}^n$ es conexo si, y sólo si, cada subconjunto suyo no vacío tiene frontera no vacía.

Problema 6.4. Sea X un espacio métrico. Supongamos que existen caminos continuos γ_1, que une x con y, y γ_2, que une y con z. Probar que existe un camino continuo γ que une x con z.

Problema 6.5. Demostrar que cualquier bola en un espacio normado es un conjunto convexo, es decir, el segmento que une dos cualesquiera de sus puntos está dentro de ella.

Problema 6.6. Sea E un espacio vectorial normado de dimensión por lo menos dos. Demostrar que:

1. $E \setminus \{0\}$ es conexo por caminos.
2. $S(x,r) = \{y \in E : \|y - x\| = r\}$ es conexo por caminos, para cada $x \in E$, $r > 0$.

Problema 6.7. Demostrar que $\mathbb{R}^n \setminus \overline{B}(0,1)$ es conexo si y sólo si $n \geq 2$.

Problema 6.8. Probar que las componentes conexas de $\mathbb{Q}$ son sus puntos.

Problema 6.9. Sea A un subconjunto de un espacio métrico, y sean $x \in A$, $y \in X \setminus A$, y $\gamma : [0,1] \to X$ un camino continuo que une x con y. Demostrar que existe $t \in [0,1]$ tal que $\gamma(t) \in \partial A$.

Problema 6.10. Se dice que un espacio métrico es localmente conexo por caminos si cada uno de sus puntos tiene un entorno que es conexo por caminos. Demostrar que un espacio métrico es conexo por caminos si y sólo si es conexo y localmente conexo por caminos.

Problema 6.11. Probar que si X e Y son conexos por caminos entonces $X \times Y$ es conexo por caminos.

Problema 6.12. Encontrar dos subconjuntos A, B y un punto x_0 de $\mathbb{R}^2$ tales que $A \cup B$ es desconexo pero $A \cup B \cup \{x_0\}$ es conexo por caminos.

Problema 6.13. Se dice que un conjunto es totalmente desconexo si sus únicas componentes conexas son los puntos (ver el problema 6.8). Probar que el conjunto de Cantor (ver el problema 4.27) es totalmente desconexo.

Problema 6.14. Sean X, Y espacios métricos, $f : X \to Y$ continua, y supongamos que X es conexo. Probar que la gráfica de f, $G_f = \{(x, y) \in X \times Y : y = f(x), x \in X\}$ es un conjunto conexo.

Problema 6.15. ¿Es válido el resultado anterior si sustinuímos *conexo* por *conexo por caminos*?

Problema 6.16. Sean X un espacio métrico, $f : X \to \mathbb{R}$ una función. ¿Es cierto que si la gráfica G_f es un conjunto cerrado y conexo entonces f es continua? *Indicación:* considerar $X = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, la circunferencia unidad.

Problema 6.17. Probar que si $f : \mathbb{R} \to \mathbb{R}$ es una función cuya gráfica es un conjunto cerrado y conexo entonces f es continua. ¿Es esto cierto si solamente pedimos que la gráfica sea conexa?

Problema 6.18. Demostrar que $\mathbb{R}$ y $\mathbb{R}^2$ no son homeomorfos.

Problema 6.19. Demostrar que la circunferencia unidad $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ no es homeomorfa a ningún intervalo de $\mathbb{R}$.

Problema 6.20. Sea C un subconjunto convexo de un espacio normado. Probar que si $x \in \operatorname{int}(C)$, $y \in C$, entonces $(1-t)x + ty \in \operatorname{int}(C)$ para $0 \leq t < 1$. Deducir que $\operatorname{int}(C)$ es convexo.

Problema 6.21. Probar que $f : D \to \mathbb{R}$ es convexa si y sólo si su epigrafo $\operatorname{epi}(f) := \{(x, y) \in D \times \mathbb{R} : y \geq f(x)\}$ es convexo en $E \times \mathbb{R}$.

Problema 6.22. Si A es un subconjunto de un espacio normado E, definimos la envoltura convexa de A como

$$\operatorname{co}(A) := \{\sum t_i a_i : a_i \in A, t_i \geq 0, \sum t_i = 1\},$$

donde todas las sumas son finitas. Demostrar que $\operatorname{co}(A)$ es convexo, y que es el menor conjunto convexo que contiene a A (es decir la intersección de todos los convexos que contienen a A). Deducir que C es convexo si y sólo si $\operatorname{co}(C) = C$.

Problema 6.23. Definamos

$$\overline{\operatorname{co}}(A) = \overline{\operatorname{co}(A)}.$$

Demostrar que $\overline{\operatorname{co}}(A)$ es el menor convexo cerrado que contiene a A.

Problema 6.24. Sea $T : E \to F$ una aplicación lineal entre espacios normados. Probar que si C es convexo en E entonces $T(C)$ es convexo en F.

Problema 6.25. Si $T : E \to F$ es lineal y D es convexo en F, ¿es $T^{-1}(D)$ convexo en E?

Problema 6.26. Demostrar que $f : D \to \mathbb{R}$ es convexa si y sólo si

$$f(\sum t_i x_i) \leq \sum t_i f(x_i)$$

siempre que $x_i \in D$, $t_i \in [0,1]$, $\sum t_i = 1$, donde las sumas son siempre finitas.

Problema 6.27. Si $A = \{(1,1),(1,-1),(-1,1),(-1,-1)\}$, ¿qué es $\mathrm{co}(A)$? Generalizar este resultado a $\mathbb{R}^n$.

Problema 6.28. Demostrar que si $f : \mathbb{R}^n \to \mathbb{R}$ es convexa entonces f está acotada en los acotados de $\mathbb{R}^n$. Utilizar la Proposición 6.17 para concluir que f es continua (de hecho Lipschitz sobre cada acotado de $\mathbb{R}^n$). *Indicación:* combinar los problemas 6.26 y 6.27.

Problema 6.29. Sea $f : \mathbb{R} \to \mathbb{R}$ convexa, y supongamos que f no es constante. Demostrar que entonces $\lim_{x\to+\infty} f(x) = +\infty$ o $\lim_{x\to-\infty} f(x) = +\infty$.

Problema 6.30. Consideremos en $\mathbb{R}^2$ la circunferencia $\mathcal{C} = \{(x,y) : x^2 + (y-r)^2 = r^2\}$, y sea $z = (x,y)$ con $y > 0$. Probar que la recta $y = 0$ es tangente a $\mathcal{C}$ en el origen, y que los puntos tz quedan en el interior del círculo delimitado por $\mathcal{C}$ para $t > 0$ suficientemente pequeño.

Problema 6.31. Demostrar que la interseccion arbitraria de conjuntos convexos es convexo. ¿Es cierto que la unión de convexos es convexo?

Problema 6.32. Demostrar que todo cerrado convexo C de $\mathbb{R}^n$ es intersección *numerable* de semiespacios cerrados, es decir, puede escribirse

$$C = \bigcap_{k=1}^{\infty} \{x \in \mathbb{R}^n : f_k(x) \leq \alpha_k\},$$

donde $f_k : \mathbb{R}^n \to \mathbb{R}$ son formas lineales, y $\alpha_k \in \mathbb{R}$, para todo $k \in \mathbb{N}$. *Indicación:* Combinar el Corolario 6.20 y el Teorema 3.10.

Problema 6.33. Sea C un convexo en $\mathbb{R}^n$. Demostrar que C es acotado si y sólo si C no contiene ninguna semirecta. *Indicación:* Puede suponerse que C es cerrado, y que $0 \in C$ ¿por qué? Razonar por reducción al absurdo: si C no estuviera acotado, existiría una sucesión $(x_k) \subset C$ con $\|x_k\| \to \infty$; defínase $z_k = x_k/\|x_k\|$, extráigase una subsucesión de (x_k) que converja a un punto $z_0 \in C$ con $\|z\| = 1$, y pruébese que $\{tz : t \geq 0\} \subset C$.

Problema 6.34. ¿Es cierto el resultado del problema anterior para los convexos de un espacio normado de dimensión infinita?

Capítulo 7

Algunos temas más avanzados de topología

En este capítulo consideraremos algunos resultados importantes que, si bien en su mayoría tienen demostraciones que no podremos exponer en este curso (por requerir, por ejemplo, teoría de integración de funciones de varias variables) creemos que tienen enunciados lo suficientemente sencillos para que el alumno pueda comprenderlos y apreciarlos en este momento. Comenzamos por varios teoremas fundamentales sobre topología en $\mathbb{R}^n$, como son el teorema del punto fijo de Brouwer, el teorema de invariancia del dominio y el teorema de la curva de Jordan. Continuamos con un estudio del espacio de las funciones continuas sobre un espacio métrico, para terminar con los teoremas de Stone-Weierstrass y Ascoli-Arzelà, sin olvidar el teorema de Riesz que caracteriza los espacios vectoriales de dimensión finita como aquéllos en los que los cerrados y acotados son compactos.

Teorema 7.1 (del punto fijo de Brouwer)**.** *Sea B la bola unidad cerrada de $\mathbb{R}^n$, y $f : B \to B$ una aplicación continua. Entonces f tiene un punto fijo.*

Recordemos que se dice que $f : B \to B$ tiene un punto fijo si existe $x_0 \in B$ tal que $f(x_0) = x_0$. Debe observarse que, a diferencia del teorema del punto fijo para aplicaciones contractivas ya visto, la función f no tiene por qué ser contractiva, ni siquiera Lipschitz (aunque forzosamente será uniformemente continua, por ser B compacto). Puede consultarse una demostración relativamente elemental de este teorema en el Apéndice (el alumno estará preparado para comprender esta demostración en cuanto haya terminado esta asignatura y haya comenzado el estudio de la teoría de integración de funciones de varias variables).

Una formulación equivalente de este teorema del punto fijo es que no existe ninguna *retracción* continua $g : B \to S$, es decir, no existe ninguna aplicación continua g de la bola en la esfera unidad tal que $g(x) = x$ para todo $x \in S$; ver los ejercicios 7.3 y 7.4.

El siguiente resultado, que puede parecer intuitivamente obvio, se conoce como teorema de la curva de Jordan. Se dice que C es una curva cerrada simple en un espacio métrico X si existe una curva continua $\gamma : [0,1] \to X$ tal que $\gamma[0,1] = C$, $\gamma(0) = \gamma(1)$, y $\gamma(t) \neq \gamma(s)$ para todo $s, t \in (0,1)$ con $s \neq t$. Es decir, la curva empieza y acaba en el mismo punto (es cerrada) y no se corta a sí misma en ningún otro punto que $\gamma(0) = \gamma(1)$. Esto es equivalente a decir que C es homeomorfa a la circunferencia unidad $S^1 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ (ver ejercicio 7.5).

Teorema 7.2 (de la curva de Jordan)**.** *Sea C una curva cerrada simple en $\mathbb{R}^2$. Entonces $\mathbb{R}^2 \setminus C$ tiene exactamente dos componentes conexas (una de ellas acotada y la otra no acotada) cuya frontera común es C.*

La demostración de este resultado es mucho más difícil de lo que su inocente enunciado puede hacer suponer, y está fuera del alcance de este curso. Lo mismo sucede, por supuesto, con la siguiente generalización al espacio $\mathbb{R}^n$. Denotaremos $S^n = \{x \in \mathbb{R}^{n+1} : \|x\| = 1\}$, la esfera unidad de $\mathbb{R}^{n+1}$.

Teorema 7.3 (de la curva de Jordan generalizado)**.** *Si $A \subset \mathbb{R}^n$ es homeomorfo a S^{n-1}, donde $n \geq 2$, entonces $\mathbb{R}^n \setminus A$ tiene exactamente dos componentes conexas cuya frontera común es A.*

Puede deducirse que una de estas componentes conexas ha de ser forzosamente acotada (la que queda *dentro* de A), y la otra no acotada (la que queda *fuera* de A); ver el problema 7.6.

El siguiente teorema es también debido a Brouwer, aunque suele citarse simplemente como teorema de invariancia del dominio. Puede resumirse diciendo que los homeomorfismos entre subconjuntos abiertos de espacios vectoriales de dimensión finita conservan la dimensión.

Teorema 7.4 (de Brouwer de invariancia del dominio)**.** *Sea U un subconjunto abierto de $\mathbb{R}^n$, y $f : U \to \mathbb{R}^n$ inyectiva y continua. Entonces $f(U)$ es abierto en $\mathbb{R}^n$.*

Aunque el teorema de invariancia del dominio quizás sea menos intuitivo que el de la curva de Jordan generalizado, puede deducirse con relativa facilidad de éste si además asumimos como cierto el siguiente hecho: *si X es homeomorfo a la bola unidad cerrada B de $\mathbb{R}^n$, entonces $\mathbb{R}^n \setminus X$ es conexo* (para $n \geq 2$). En efecto, para cada $x \in U$, escojamos $r_x > 0$ con $\overline{B}(x, r_x) \subset U$, y definamos $B_x = B(x, r_x)$, y $S_x = S(x, r_x) = \partial B(x, r_x)$. Por la asunción que estamos haciendo, $\mathbb{R}^n \setminus f(\overline{B_x})$ es conexo. Por otro lado, $f(B_x) = f(\overline{B_x}) \setminus f(S_x)$ es conexo por ser imagen continua de un conexo. Tenemos entonces

$$\mathbb{R}^n \setminus f(S_x) = \big(\mathbb{R}^n \setminus f(\overline{B_x})\big) \cup f(B_x),$$

donde la unión es disjunta, y por el teorema de la curva de Jordan generalizado sabemos que $\mathbb{R}^n \setminus f(S_x)$ tiene dos componentes conexas. Entonces forzosamente $\mathbb{R}^n \setminus f(\overline{B_x})$ y $f(B_x)$ son las dos componentes conexas de $\mathbb{R}^n \setminus f(S_x)$, y como a su vez este conjunto es abierto en $\mathbb{R}^n$, se deduce que $f(B_x)$ es abierto en $\mathbb{R}^n$ (ver la Proposición 6.11). Se concluye entonces que

$$f(U) = \bigcup_{x \in U} f(B_x),$$

al ser unión de abiertos, es abierto en $\mathbb{R}^n$. El caso $n = 1$ es por supuesto bien conocido de un primer curso de análisis de variable real.

Del teorema de invariancia del dominio se deduce inmediatamente la siguiente propiedad de los homeomorfismos entre subconjuntos de espacios vectoriales de dimensión finita.

Corolario 7.5. *Sean $A, B \subseteq \mathbb{R}^n$, y $h : A \to B$ un homeomorfismo. Entonces h lleva puntos interiores de A en puntos interiores de B, y puntos frontera en puntos frontera.*

Este resultado no es cierto cuando A y B son subconjuntos de un espacio normado de dimensión infinita. Lo mismo ocurre con el teorema del punto fijo de Brouwer y el de la curva de Jordan generalizado: sólo son válidos en espacios normados de dimensión finita. Ver los problemas 7.11 y 7.12.

Antes de presentar los teoremas de Stone-Weierstrass y de Arzelà-Ascoli haremos un breve estudio del espacio de las funciones continuas sobre un espacio métrico y de las nociones de convergencia puntual y uniforme de sucesiones de funciones.

Definición 7.1. Sean X, Y espacios métricos. Denotaremos por $C(X,Y)$ el conjunto de todas las funciones continuas $f : X \to Y$, y por $C_b(X,Y)$ el subconjunto de todas las funciones continuas y *acotadas* de X en Y. En este subconjunto definiremos la distancia entre dos funciones $f, g \in C_b(X,Y)$ como

$$d_\infty(f,g) = \sup\{d_Y(f(x), g(x)) : x \in X\}.$$

Es inmediato ver que d define una distancia en $C_b(X,Y)$.

Cuando Y es un espacio vectorial normado, $C_b(X,Y)$ y $C(X,Y)$ resultan ser también espacios vectoriales. Al espacio $C_b(X,Y)$ podemos dotarlo de la norma

$$\|f\|_\infty = \sup\{\|f(x)\|_Y : x \in X\},$$

que genera precisamente la distancia d_∞ anteriormente definida en $C_b(X,Y)$.

Cuando $Y = \mathbb{R}$, el espacio Y suele omitirse en esta notacion. Así $C_b(X,\mathbb{R})$ suele escribirse simplemente $C_b(X)$. Cuando X es compacto, todas las funciones continuas de X en Y son acotadas, y se tiene que $C(X,Y) = C_b(X,Y)$.

Definición 7.2. Si (f_n) es una sucesión de funciones de X en Y (no necesariamente continuas), diremos que (f_n) converge puntualmente a f si la sucesión $(f_n(x))$ de Y converge a $f(x)$ en el espacio métrico Y para cada $x \in X$. Si $A \subseteq X$, diremos que (f_n) converge uniformemente a f sobre A si para todo $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que si $n \geq n_0$ entonces $d_Y(f_n(x), f(x)) \leq \varepsilon$ para todo $x \in A$.

Cuando $Y = F$ es un espacio vectorial normado, diremos que una serie de funciones $\sum_{n=1}^\infty g_n$ converge uniformemente en X (resp. puntualmente) si la sucesión (f_n) de las sumas parciales $f_n = \sum_{j=1}^n g_j$ de la serie converge uniformemente (resp. puntualmente) en X.

Como en el caso ya conocido por el alumno en que X es un intervalo de $\mathbb{R}$, tenemos que el límite uniforme de funciones continuas es continuo.

Proposición 7.6. *Si (f_n) es una sucesión de de funciones continuas de X en Y que converge uniformemente a $f : X \to Y$ sobre X, entonces f es continua.*

Sin embargo el límite puntual de funciones continuas puede no ser continuo; ver el problema 7.13.

El siguiente resultado relaciona la convergencia uniforme de funciones y el espacio métrico $C_b(X,Y)$.

Proposición 7.7. *Una sucesión (f_n) de $C_b(X,Y)$ converge uniformemente a $f \in C_b(X,Y)$ si y sólo si $f_n \to f$ en el espacio métrico $(C_b(X,Y), d_\infty)$.*

Definición 7.3. Se dice que una sucesión (f_n) de funciones de X en Y es uniformente de Cauchy si para todo $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que si $n, m \geq n_0$ entonces $d(f_n(x), f_m(x)) \leq \varepsilon$ para todo $x \in X$.

Cuando Y es completo, es fácil ver que toda tal sucesión es convergente uniformemente en X. También es inmediato que si $(f_n) \subset C_b(X,Y)$, entonces (f_n) es uniformemente de Cauchy si y sólo si (f_n) es de Cauchy en $C_b(X,Y)$.

Por supuesto las definiciones y resultados anteriores son también válidos para series de funciones, pues basta aplicarlos a la sucesión de las sumas parciales. Se invita al alumno a que enuncie y demuestre la versión para series.

Como cabe esperar a la vista de todo esto, hay también una estrecha relación entre la completitud de Y y la de $C_b(X,Y)$.

Proposición 7.8. *Si Y es completo entonces $C_b(X,Y)$ es también completo. Si Y es un espacio de Banach, también lo es $(C_b(X,Y), \|\cdot\|_\infty)$.*

Para probar este resultado basta tener en cuenta dos hechos básicos: el primero, que una sucesión de Cauchy en $C_b(X,Y)$ convergerá puntualmente, siempre que Y sea completo, a una función $f : X \to Y$ acotada; el segundo, que el límite uniforme de funciones continuas es continuo, como ya hemos visto. □

Pasamos ahora a estudiar el teorema de Arzelà-Ascoli, que nos proporciona una caracterización de los subconjuntos compactos de $C(K,Y)$, donde K es un espacio métrico compacto. En este sentido es análogo al teorema de Heine-Borel que caracteriza la compacidad de los subconjuntos de $\mathbb{R}^n$. En la mayoría de los espacios del tipo $C_b(X,Y)$ el teorema de Heine-Borel es falso, como sucede más en general en todo espacio vectorial normado de dimensión infinita. En efecto, el siguiente teorema, debido a Riesz, implica que los únicos espacios vectoriales normados en los que los cerrados y acotados son compactos son los de dimensión finita.

Teorema 7.9 (Riesz)**.** *Sean E un espacio vectorial normado, B_E su bola unidad cerrada, y supongamos que B_E es compacto. Entonces E es de dimensión finita.*

Para demostrar esto se utiliza el siguiente lema.

Lema 7.10. *Sean E un espacio vectorial normado y $W \subset E$ un subespacio vectorial cerrado propio. Entonces, para cada $\varepsilon > 0$ existe $u \in E$ tal que $\|u\| = 1$ y $d(u,W) \geq 1-\varepsilon$.*

Demostración. Elijamos $v \in E \setminus W$. Como W es cerrado tenemos que $d(v,W) > 0$, y podemos escoger $w \in W$ tal que

$$d(v,W) \leq \|v-w\| \leq \frac{d(v,W)}{1-\varepsilon}.$$

Entonces

$$u := \frac{v-w}{\|v-w\|}$$

nos resuelve el problema. En efecto, para todo $x \in W$ tenemos que

$$\|u-x\| = \|\frac{v-w}{\|v-w\|} - x\| = \frac{\|v-(w+\|v-w\|x)\|}{\|v-w\|} \geq \frac{d(v,W)}{\|v-w\|} \geq 1-\varepsilon,$$

puesto que $w + \|v-w\|x \in W$, y así deducimos que $d(u,W) \geq 1-\varepsilon$. □

Para demostrar el teorema de Riesz, supongamos que E tiene dimensión infinita. Entonces, para cada $n \in \mathbb{N}$ existe un subespacio vectorial de E de dimensión finita, n, y de modo que $E_n \subset E_{n+1}$ para todo $n \in \mathbb{N}$. Usando el lema anterior podemos construir entonces una sucesión de vectores (u_n) tales que $u_n \in E_n \setminus E_{n-1}$, $\|u_n\| = 1$, $d(u_n, E_{n-1}) \geq 1/2$ para todo n. En particular se tendrá $\|u_n - u_m\| \geq 1/2$ si $n \neq m$, y es evidente entonces que (u_n) no puede tener ninguna subsucesión convergente y en conclusión que B_E no es compacta. □

Para enunciar el teorema de Arzelà-Ascoli, que como ya hemos avanzado caracteriza los subconjuntos compactos de $C(K,Y)$, necesitamos la noción de *equicontinuidad*.

Definición 7.4. Se dice que una familia $\mathcal{F}$ de funciones de $C_b(X,Y)$ es equicontinua si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que para todo $x, y \in X$ y para toda $f \in \mathcal{F}$ se tiene que $d(f(x), f(y)) \leq \varepsilon$ siempre que $d(x,y) \leq \delta$.

Es decir, en la definición usual de continuidad uniforme se pide además que el mismo δ sirve a la vez para todas las funciones de $\mathcal{F}$.

Diremos también que $\mathcal{F}$ es *puntualmente compacto* si $\mathcal{F}_x := \{f(x) : f \in \mathcal{F}\}$ es compacto en Y para cada $x \in X$.

Teorema 7.11. *Sea K un espacio métrico compacto, Y un espacio métrico cualquiera y $\mathcal{F}$ un subconjunto de $C(K,Y)$. Entonces $\mathcal{F}$ es compacto en $C(K,Y)$ si y sólo si $\mathcal{F}$ es cerrado, equicontinuo y puntualmente compacto.*

Demostración. $(\Rightarrow)$: Si $\mathcal{F}$ es compacto entonces, para cada $x \in K$, como la función $\pi_x : \mathcal{F} \to Y$, $\pi_x(f) = f(x)$ es continua, se tiene que que $\mathcal{F}_x = \pi_x(\mathcal{F})$ es compacto en Y; es decir $\mathcal{F}$ es puntualmente compacto. Por otra parte, es inmediato ver que la aplicación $\pi : \mathcal{F} \times K \to Y$, definida por $\pi(f,x) = f(x)$, es continua, y por tanto, como $\mathcal{F} \times K$ es compacto, también uniformemente continua. Esto significa que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, para todos $x, y \in K$, $f, g \in \mathcal{F}$, si $d(x,y) \leq \delta$ y $d(f,g) \leq \delta$, entonces $d(f(x), g(y)) \leq \varepsilon$. En particular, poniendo $f = g$, obtenemos que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que, para todos $x, y \in K$, si $d(x,y) \leq \delta$ entonces $d(f(x), g(y)) \leq \varepsilon$; es decir, $\mathcal{F}$ es equicontinuo.
$(\Leftarrow)$: Ésta es la parte más difícil y la más importante. Si (f_n) es una sucesión de $\mathcal{F}$, veremos que (f_n) tiene una subsucesión convergente a una función f en $\mathcal{F}$. Como $\mathcal{F}$ es puntualmente compacto esto puede hacerse en cada punto $x \in K$. Usando la equicontinuidad de $\mathcal{F}$, la compacidad de K y un argumento de diagonalización, veremos que puede obtenerse convergencia uniforme.

Como K es compacto es totalmente acotado, y así, para cada $n \in \mathbb{N}$ podemos encontrar un subconjunto finito C_n de K tal que cualquier punto de K está a distancia menor que $1/n$ de algún punto de C_n. Definamos

$$C = \bigcup_{n=1}^{\infty} C_n,$$

y puesto que C es numerable escribamos

$$C = \{x_n : n \in \mathbb{N}\}.$$

Como $\mathcal{F}_{x_1}$ es compacto, la sucesión $(f_n(x_1))$ tiene una subsucesión convergente en Y, que denotaremos

$$f_{11}(x_1), f_{12}(x_1), ..., f_{1n}(x_1), ...$$

A su vez, como $(f_{1k})_{k\in\mathbb{N}} \subseteq \mathcal{F}$ y $\mathcal{F}_{x_2}$ es compacto, la sucesión $(f_{1k}(x_2))_{k\in\mathbb{N}}$ tiene una subsucesión convergente en Y, que denotamos

$$f_{21}(x_2), f_{22}(x_2), ..., f_{2n}(x_2), ...$$

Continuando el proceso de esta manera, obtenemos una sucesión de subsucesiones de (f_n), a saber, $(f_{nk})_{k\in\mathbb{N}}$, $n \in \mathbb{N}$, de tal manera que $(f_{(n+1)k})_{k\in\mathbb{N}}$ es subsucesión de $(f_{nk})_{k\in\mathbb{N}}$ para cada $n \in \mathbb{N}$, y la sucesión $(f_{nk}(x_n))_{k\in\mathbb{N}}$ converge en Y para cada $n \in \mathbb{N}$.

Definamos ahora $g_n = f_{nn}$ para cada $n \in \mathbb{N}$ (ésta es la diagonal de la matriz infinita $(f_{nk})_{n,k\in\mathbb{N}}$). Es claro que para todo $m \in \mathbb{N}$ la sucesión $(g_n)_{n\geq m}$ es una subsucesión de la subsucesión $(f_{mk})_{k\in\mathbb{N}}$, para todo $m \in \mathbb{N}$, y por tanto la sucesión $(g_n(x))_{n\in\mathbb{N}}$ converge en Y para todo $x \in C$.

Probaremos que de hecho la sucesión $(g_n(x))$ converge en todo punto $x \in K$, y que la convergencia es uniforme (en particular la función límite $x \mapsto f(x) := \lim_{n\to\infty} g_n(x)$ será continua), lo cual demostrará que (g_n) converge a f en $C(K,Y)$ y, como $\mathcal{F}$ es cerrado, también que $g_n \to f$ en $\mathcal{F}$.

Sea $\varepsilon > 0$, y elijamos $\delta > 0$ como en la definición de equicontinuidad de $\mathcal{F}$. Sea $C_\delta = \{y_1, ..., y_k\}$ un subconjunto finito de C tal que todo punto de K esté a distancia menor que δ de algún punto de C_δ (es decir, tómese $C_\delta = C_{n_\delta}$ con $n_\delta > 1/\delta$). Puesto que las sucesiones

$$(g_n(y_1))_{n\in\mathbb{N}}, (g_n(y_2))_{n\in\mathbb{N}}, ..., (g_n(y_k))_{n\in\mathbb{N}}$$

son todas convergentes, existe $N \in \mathbb{N}$ tal que si $n, m \geq N$ entonces

$$d(g_n(y_j), g_m(y_j)) \leq \varepsilon/3$$

para todo $j = 1, 2, ..., k$. Ahora, para cada $x \in K$, podemos tomar $y_x \in C_\delta$ tal que $d(x, y_x) \leq \delta$, luego por la hipótesis de equicontinuidad se tiene que

$$d(g_n(x), g_n(y_x)) \leq \varepsilon/3$$

para todo $n \in \mathbb{N}$. Por tanto, si $n, m \geq N$, tenemos

$$d(g_n(x), g_m(x)) \leq d(g_n(x), g_n(y_x)) + d(g_n(y_x), g_m(y_x)) + d(g_m(y_x), g_m(x)) \leq \varepsilon.$$

Así hemos visto que para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que si $n, m \geq N$ entonces

$$d(g_n(x), g_m(x)) \leq \varepsilon \qquad (*)$$

para todo $x \in K$. En particular $(g_n(x))$ es de Cauchy en Y para cada $x \in K$, y como $(g_n(x)) \subset \mathcal{F}_x$ y $\mathcal{F}_x$ es completo (por ser compacto), se deduce que $(g_n(x))$ converge para cada $x \in K$. Si denotamos

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

fijando $n \in \mathbb{N}$ y tomando límite cuando $m \to \infty$ en $(*)$, deducimos que para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que si $n \geq N$ entonces

$$d(g_n(x), f(x)) \leq \varepsilon \qquad (*)$$

para todo $x \in K$, lo cual significa que (g_n) converge a f uniformemente en K, y en particular $f \in C(K, Y)$ y $g_n \to f$ en $C(K, Y)$. Por último, como $\mathcal{F}$ es cerrado en $C(K, Y)$, se tiene que $f \in \mathcal{F}$. $\square$

Corolario 7.12. *Sea K un espacio métrico compacto, y $\mathcal{B}$ un subconjunto equicontinuo y puntualmente acotado de $C(K, \mathbb{R}^n)$. Entonces toda sucesión de $\mathcal{B}$ tiene una subsucesión uniformemente convergente a una función de $C(K, \mathbb{R}^n)$.*

Como es natural, decir que $\mathcal{B}$ es puntualmente acotado significa que $\mathcal{B}_x := \{f(x) : f \in \mathcal{B}\}$ es acotado en $\mathbb{R}^n$ para cada $x \in K$.

Ejercicio 7.1. Sea (f_n) una sucesión de funciones continuas $f_n : [a, b] \to \mathbb{R}$ uniformemente acotada, y para cada $n \in \mathbb{N}$ definamos

$$F_n(x) = \int_a^x f_n(t)dt, \;\; x \in [a, b].$$

Probar que (F_n) tiene una subsucesión que converge uniformemente.

Concluimos este capítulo con un estudio del teorema de Stone-Weierstrass, que es uno de los resultados más importantes en el contexto del espacio de las funciones continuas sobre un compacto, y que tiene multitud de aplicaciones.

El origen de este teorema es un resultado probado por Weierstrass en 1885: toda función continua sobre un compacto de $\mathbb{R}^n$ es límite uniforme de una sucesión de polinomios en n indeterminadas y con coeficientes reales.

Teorema 7.13 (Weierstrass). *Sean K un subconjunto compacto de $\mathbb{R}^n$, y $f : K \to \mathbb{R}$ continua. Entonces, para cada $\varepsilon > 0$ existe $p : \mathbb{R}^n \to \mathbb{R}$ polinomio tal que*

$$|f(x) - p(x)| \leq \varepsilon$$

para todo $x \in K$.

Aunque Weierstrass dio una demostración directa de este resultado, en 1937 Stone descubrió una generalización muy útil del teorema, que vale no solamente para compactos de $\mathbb{R}^n$, sino para más familias de funciones, aparte de los polinomios.

Teorema 7.14 (Stone-Weierstrass). *Sea K un espacio métrico compacto, y $\mathcal{B}$ un subconjunto de $C(K, \mathbb{R})$ que satisfaga:*

1. *$\mathcal{B}$ es un álgebra de funciones, es decir, para todas $f, g \in \mathcal{B}$, $\alpha \in \mathbb{R}$ se tiene que $f + g$, fg y αf están en $\mathcal{B}$.*

2. *La función constante $x \mapsto 1$ está en $\mathcal{B}$.*

3. *$\mathcal{B}$ separa puntos, esto es, para todos $x, y \in K$ con $x \neq y$, existe $f \in \mathcal{B}$ tal que $f(x) \neq f(y)$.*

Entonces $\mathcal{B}$ es densa en $C(K, \mathbb{R})$; es decir, para toda $f \in C(K, \mathbb{R})$ y todo $\varepsilon > 0$ existe $\varphi \in \mathcal{B}$ tal que $|f(x) - \varphi(x)| \leq \varepsilon$ para todo $x \in K$.

Para deducir el Teorema 7.13 basta observar que $\mathcal{P} = \{p_{|_K} : p : \mathbb{R}^n \to \mathbb{R} \text{ polinomio}\}$ satisface las tres condiciones del enunciado del teorema de Stone-Weierstrass (las dos primeras son evidentes, y la tercera es un sencillo ejercicio).
Demostración del Teorema 7.14: Usaremos la notación siguiente:

$$(f \vee g)(x) = \text{máx}\{f(x), g(x)\}, \text{ y } (f \wedge g)(x) = \text{mín}\{f(x), g(x)\}.$$

Sea $\overline{\mathcal{B}}$ la adherencia de $\mathcal{B}$. Por la continuidad de las operaciones de suma y multiplicación, es claro que $\overline{\mathcal{B}}$ es también un álgebra, y por otra parte es evidente que $\overline{\mathcal{B}}$ satisface también las otras propiedades de $\mathcal{B}$ puesto que contiene a este conjunto.

El primer paso de la demostración es probar que si $f \in \overline{\mathcal{B}}$ entonces $|f| \in \overline{\mathcal{B}}$. Para ver esto, usamos el siguiente caso particular del teorema de Weierstrass.

Lema 7.15. *Existe una sucesión de polinomios $p_n : \mathbb{R} \to \mathbb{R}$ tal que para todo $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ de modo que si $n \geq n_0$ entonces $|p_n(t) - \sqrt{t}| \leq \varepsilon$ para todo $t \in [0, 1]$.*

Una tal sucesión (p_n) puede definirse inductivamente como sigue. Pongamos $p_0(t) = 0$, y para $n \geq 1$, sea

$$p_n(t) = p_{n-1}(t) + \frac{1}{2}[t - p_{n-1}(t)^2].$$

Puede probarse por inducción (ver el ejercicio 7.17) que

$$0 \leq \sqrt{t} - p_n(t) \leq \frac{2\sqrt{t}}{2 + n\sqrt{t}}$$

para todo $t \in [0, 1]$, $n \in \mathbb{N}$, lo que obviamente demuestra el lema.

Ahora, si $f \in \overline{\mathcal{B}}$, existe $b > 0$ tal que $|f(x)| \leq b$ para todo $x \in K$. Por el lema anterior, para cada $\varepsilon > 0$ podemos encontrar $n_0 \in \mathbb{N}$ tal que $|p_n(t) - \sqrt{t}| \leq \varepsilon/b$ para todo $t \in [0,1]$, $n \geq n_0$. En particular

$$\left| p_{n_0}\Big(\frac{f^2(x)}{b^2}\Big) - \sqrt{\frac{f^2(x)}{b^2}} \right| \leq \frac{\varepsilon}{b},$$

y por tanto

$$\left| b p_{n_0}(f^2(x)/b^2) - |f(x)| \right| \leq \varepsilon$$

para todo $x \in K$. Como para cada $n \in \mathbb{N}$ la función $x \mapsto b p_n(f^2(x)/b^2)$ está en $\overline{\mathcal{B}}$ (porque la composición de un polinomio con funciones del álgebra siempre está en el álgebra, al ser sumas y productos de funciones del álgebra), se deduce que $|f| \in \overline{\mathcal{B}}$ siempre que $f \in \overline{\mathcal{B}}$.

Teniendo en cuenta este hecho y las fórmulas

$$\text{máx}\{f, g\} = \frac{f + g + |f - g|}{2}, \text{ y } \text{mín}\{f, g\} = \frac{f + g - |f - g|}{2},$$

deducimos que si $f, g \in \overline{\mathcal{B}}$ entonces $f \vee g$ y $f \wedge g$ están también en $\overline{\mathcal{B}}$.

Veamos ahora que $\overline{\mathcal{B}} = C(K, \mathbb{R})$. Puesto que $\overline{\overline{\mathcal{B}}} = \overline{\mathcal{B}}$, es suficiente probar que para cada $h \in C(K, \mathbb{R})$ y $\varepsilon > 0$ existe $f \in \overline{\mathcal{B}}$ tal que

$$|f(x) - h(x)| \leq \varepsilon$$

para todo $x \in K$.

Para cada $x_1, x_2 \in K$ con $x_1 \neq x_2$, elijamos $g \in \mathcal{B}$ tal que $g(x_1) \neq g(x_2)$, y pongamos

$$f_{x_1 x_2}(x) = \alpha g(x) + \beta,$$

donde

$$\alpha = \frac{h(x_1) - h(x_2)}{g(x_1) - g(x_2)}, \text{ y } \beta = \frac{g(x_1)h(x_2) - h(x_1)g(x_2)}{g(x_1) - g(x_2)}.$$

Es obvio que $f_{x_1 x_2} \in \mathcal{B}$, y los números α y β han sido escogidos para que

$$f_{x_1 x_2}(x_1) = h(x_1) \text{ y } f_{x_1 x_2}(x_2) = h(x_2).$$

Fijemos $x \in K$. Para cada $y \in K$, por continuidad de h, existe un entorno abierto U_y de y tal que $f_{yx}(z) > h(z) - \varepsilon$ si $z \in U_y$. Como K es compacto, podemos escribir

$$K = \bigcup_{j=1}^{n} U_{y_j}$$

para ciertos $y_1, ..., y_n \in K$. Definamos

$$f_x = f_{y_1 x} \vee ... \vee f_{y_n x}.$$

Se tiene que $f_x \in \overline{\mathcal{B}}$,

$$f_x(x) = h(x), \text{ y } f_x(z) > h(z) - \varepsilon$$

para todo $z \in K$. Por tanto, por continuidad de h y f_x, existe V_x entorno abierto de x tal que

$$f_x(y) < h(y) + \varepsilon$$

para todo $y \in V_x$. Por compacidad de K podemos escribir también

$$K = \bigcup_{j=1}^{m} V_{x_j}$$

para ciertos $x_1, ..., x_m \in K$. Definamos por fin

$$f = f_{x_1} \wedge ... \wedge f_{x_m}.$$

Nuevamente se tiene que $f \in \overline{\mathcal{B}}$. Además

$$f(z) > h(z) - \varepsilon$$

para todo $z \in K$, puesto que $f_{x_j}(z) > h(z) - \varepsilon$ para todo $z \in K$, $j = 1, ..., m$. Por otra parte, para cada $y \in K$ existe $j = j(y) \in \{1, ..., m\}$ tal que $y \in V_{x_j}$, y así $f(y) \leq f_{x_j}(y) < h(y) + \varepsilon$, luego también es

$$f(y) < h(y) + \varepsilon$$

para todo $y \in K$. Combinando las desigualdades anteriores concluimos que $|f(z) - h(z)| < \varepsilon$ para todo $z \in K$. $\Box$

7.1. Problemas

Problema 7.1. Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ una aplicación continua y acotada. Definamos $g : \mathbb{R}^n \to \mathbb{R}^n$ por

$$g(x) = x + f(x).$$

Demostrar que g es sobreyectiva. *Indicación:* Fijado $y \in \mathbb{R}^n$, como $f(\mathbb{R}^n)$ está acotado existe $R > 0$ tal que $\{y - f(x) : x \in \mathbb{R}^n\} \subset \overline{B}(0, R)$. Aplicar el teorema del punto fijo de Brouwer a la función

$$h(x) = y - f(x)$$

en la bola $\overline{B}(0, R)$.

Problema 7.2. Probar que existe una retracción continua de $\mathbb{R}^n \setminus \{0\}$ en S^{n-1}, es decir, existe $r : \mathbb{R}^n \setminus \{0\} \to S^{n-1}$ continua tal que $r(x) = x$ si $x \in S^{n-1}$.

Problema 7.3. Sean B y S la bola unidad cerrada y la esfera unidad cerrada de un espacio normado X. Sea $f : B \to B$ una aplicación continua, y supongamos que f no tiene ningún punto fijo. Probar que entonces existe $g : B \to S$ continua tal que $g(x) = x$ para todo $x \in S$. *Indicación:* definir $g(x)$ como aquél de los puntos intersección de la recta que pasa por x y $f(x)$ con la esfera S que esté más cerca de x que de $f(x)$.

Problema 7.4. Recíprocamente, supóngase que existe $g : B \to S$ continua tal que $g(x) = x$ para todo $x \in S$. Probar que entonces existe $f : B \to B$ continua sin puntos fijos. *Indicación:* Definir $f(x) = -g(x)$.

Problema 7.5. Demostrar que C es una curva cerrada simple si y sólo si C es homeomorfa a la circunferencia S^1.

Problema 7.6. En el teorema 7.3, probar que una de las componentes conexas de $\mathbb{R}^n \setminus A$ debe ser acotada, y la otra no acotada. *Indicación:* La parte menos obvia es que una debe ser acotada. Llamemos U y V a las componentes conexas de $\mathbb{R}^n \setminus A$. Escojamos una bola B tal que $A \subset B$. Si U y V son ambas no acotadas, tendremos que $\mathbb{R}^n \setminus B$ podrá recubrirse por la unión de los abiertos U y V, que son disjuntos y tienen intersección no vacía con $\mathbb{R}^n \setminus B$.

Problema 7.7. Si A y B son subconjuntos homeomorfos de $\mathbb{R}^n$ y A es cerrado, ¿se sigue que B es cerrado?

Problema 7.8. Demostrar, sin usar el teorema de invariancia del dominio, que si $\mathbb{R}^n$ es homeomorfo a $\mathbb{R}$ entonces $n = 1$.

Problema 7.9. Usando el teorema de invariancia del dominio, demostrar que $\mathbb{R}^n$ y $\mathbb{R}^m$ son homeomorfos si y sólo si $n = m$.

Problema 7.10. Si $f : B_1 \to B_2$ es un homeomorfismo entre dos bolas cerradas de $\mathbb{R}^n$, probar que f transforma ∂B_1 en ∂B_2.

Problema 7.11. En el espacio $(C[0,1], \|\cdot\|_\infty)$, consideremos la aplicación $T : C[0,1] \to C[0,1]$ definida por

$$T(f)(s) = \begin{cases} 2sf(0) & \text{si } 0 \leq s \leq \frac{1}{2}, \\ f(2s-1) & \text{si } \frac{1}{2} \leq s \leq 1. \end{cases}$$

para cada $s \in [0,1]$. Definamos $X = C[0,1]$, $Y = T(C[0,1])$, y sean S_X y S_Y las esferas de centro 0 y radio 1 en X e Y respectivamente. Probar que:

1. T es lineal.
2. T es una isometría sobre su imagen, es decir que $\|T(f) - T(g)\|_\infty = \|f - g\|_\infty$. En particular $T : X \to Y$ es un homeomorfismo.
3. Y es un subespacio cerrado propio de X.
4. Y tiene interior vacío en X. En particular vemos que el teorema de invariancia del dominio falla en $X = C[0,1]$, pues X es abierto en este espacio y homeomorfo a Y, que no es abierto en X.
5. Si denotamos $S_Y = T(S_X) = \{f \in Y : \|f\|_\infty = 1\}$, se tiene que $X \setminus S_X$ tiene dos componentes conexas, mientras que $X \setminus T(S_X) = X \setminus S_Y$ es conexo (de hecho conexo por caminos). Por tanto el teorema de la curva de Jordan generalizado también falla en X.

Problema 7.12. Probar que el teorema del punto fijo de Brouwer es falso en espacios normados de dimensión infinita. *Indicación*: Sean $X = (C[0,1], \|\cdot\|_\infty)$, $Y = \{f \in X : f(1) = 0\}$. Probar que Y es un subespacio vectorial cerrado de X. Definir una aplicación $T : Y \to Y$ como sigue: para cada $f \in C[0,1]$, sea $T(f)$ la función continua definida por $\Phi(f)(s) = 1 - \|f\|_\infty$ si $0 \leq s \leq 1/3$, $\Phi(f)(s) = f(3s-2)$ si $2/3 \leq s \leq 1$, y $\Phi(f)$ afín en $[1/3, 2/3]$. Demostrar que $T(B_Y) \subseteq B_Y$, donde B_Y es la bola unidad cerrada de Y, que T es continua, y que T no tiene ningún punto fijo.

Problema 7.13. Considerar la sucesión de funciones (f_n) definida por $f_n(x) = x^n$, $x \in [0,1]$, $n \in \mathbb{N}$. Probar que f_n converge puntualmente a una función que no es continua en $[0,1]$.

Problema 7.14. Demostrar el criterio M de Weierstrass para la convergencia uniforme de series de funciones: si (g_n) es una sucesión de funciones de X en E, donde X es un espacio métrico y E es un espacio de Banach, tales que

$$\|g_k(x)\| \leq M_k$$

para todos $k \in \mathbb{N}$, $x \in X$, y la serie $\sum_{k=1}^{\infty} M_k$ es convergente, entonces la serie

$$\sum_{k=1}^{\infty} g_k$$

converge uniformemente (y absolutamente) en X. *Indicación:* Probar que la sucesión de las sumas parciales es de Cauchy uniformemente.

Problema 7.15. Sea $\mathcal{B}$ un subconjunto cerrado, acotado y equicontinuo de $C[0,1]$, y definamos $I : \mathcal{B} \to \mathbb{R}$ por

$$I(f) = \int_0^1 f(x)dx.$$

Probar que existe $f_0 \in \mathcal{B}$ que maximiza el valor de I sobre $\mathcal{B}$.

Problema 7.16. Sea (f_n) una sucesión de funciones continuas en $[0,1]$ y derivables en $(0,1)$ tales que $|f_n(x)| \leq 100$ para todo $x \in [0,1]$ y $|f_n'(t)| \leq 100$ para todo $t \in (0,1)$. Probar que f_n tiene una subsucesión que converge uniformemente en $[0,1]$ a una función continua.

Problema 7.17. Definamos $p_0(t) = 0$, y para $n \geq 1$, sea

$$p_n(t) = p_{n-1}(t) + \frac{1}{2}[t - p_{n-1}(t)^2].$$

Demostrar por inducción que

$$0 \leq \sqrt{t} - p_n(t) \leq \frac{2\sqrt{t}}{2 + n\sqrt{t}}$$

para todo $t \in [0,1]$, $n \in \mathbb{N}$.

Problema 7.18. Consideremos el conjunto de todas las funciones $\varphi : [0,1] \to \mathbb{R}$ de la forma

$$\varphi(t) = \sum_{j=1}^{n} a_j e^{b_j t},$$

donde $a_j, b_j \in \mathbb{R}$, $n \in \mathbb{N}$. ¿Es este conjunto denso en el espacio $C[0,1]$?

Problema 7.19. Cualquier suma finita de senos y cosenos de la forma

$$a_0 + \sum_{k=1}^{n} (a_k \cos(k\theta) + b_k \sin(k\theta))$$

se dice que es un *polinomio trigonométrico*. Demostrar que si $f : \mathbb{R} \to \mathbb{R}$ es una función continua y periódica de período 2π, para todo $\varepsilon > 0$ existe un polinomio trigonométrico q tal que $|q(\theta) - f(\theta)| < \varepsilon$ para todo $\theta \in \mathbb{R}$. Es decir, los polinomios trigonométricos son densos en el espacio de las funciones continuas y periódicas de período 2π. Éste es un resultado probado por Weierstrass en 1885. *Indicación:* Consideremos la circunferencia unidad $\mathbb{T} = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. Existe una identificación natural entre las funciones $g : \mathbb{T} \to \mathbb{R}$ y las funciones periódicas $f : R \to \mathbb{R}$ de período 2π (basta identificar cada número θ con el punto $(\cos\theta, \sin\theta)$ de $\mathbb{T}$). Aplicar el teorema de Stone-Weierstrass en el espacio $C(\mathbb{T})$.

Parte II

Cálculo diferencial en $\mathbb{R}^n$

Capítulo 8

Derivadas direccionales, derivadas parciales, gradientes, tangente a la gráfica de una función

Después de todos los preparativos topológicos y el estudio de las funciones continuas y de sus propiedades, podemos ya comenzar en este capítulo la teoría de diferenciación de funciones de varias variables. En este capítulo presentaremos la definición de función diferenciable, y la interpretación geométrica de la diferencial de la función en un punto como la aplicación lineal que proporciona una mejor aproximación a la función entorno a ese punto, y que permite calcular el plano tangente a su gráfica en dicho punto. También explicaremos cómo calcular la matriz jacobiana o el vector gradiente de una función que se supone diferenciable en términos de las derivadas parciales de sus funciones componentes.

La definición de derivada de una función $f : \mathbb{R} \to \mathbb{R}$ se extiende inmediatamente a funciones de una variable real con valores vectoriales.

Definición 8.1. Sea $f : (a, b) \subseteq \mathbb{R} \to \mathbb{R}^n$ una curva. Se define la derivada de f en $t \in (a, b)$ como

$$f'(t) := \lim_{h \to 0} \frac{f(t+h) - f(t)}{h}$$

cuando este límite existe, en cuyo caso se dice que f es derivable en t, y a $f'(t)$ se la llama derivada de f en t. También suele denotarse

$$f'(t) = \frac{df(t)}{dt}.$$

Si visualizamos f como una curva en el espacio $\mathbb{R}^n$, y $f'(t_0) \neq 0$, la derivada $f'(t_0)$ puede interpretarse geométricamente como el vector tangente a la curva $f(a, b)$ en el punto $f(t_0)$. La ecuación paramétrica de la recta tangente será pues

$$\{f(t_0) + \lambda f'(t_0),\ \lambda \in \mathbb{R}\}.$$

Nótese que, cuando f es derivable en t_0, el error cometido al aproximar un punto $f(t_0 + h)$ de la curva f por un punto $f(t_0) + f'(t_0)h$ en la recta tangente a f en el punto $f(t_0)$ es

$$R(h) = \frac{f(t_0 + h) - f(t_0) - f'(t_0)h}{h},$$

cantidad que tiende a 0 cuando h tiende a 0. Esto indica que dicha recta tangente es una buena aproximación lineal (de hecho la única buena) a la curva descrita por f en el punto $f(t_0)$.

Si visualizamos $f(t)$ como el vector de posición de una partícula en el instante t, la interpretación física de la derivada $f'(t_0)$ será la de velocidad instantánea de dicha partícula en el instante $t = t_0$. Esto tiene perfecto sentido si tenemos en cuenta que

$$\frac{f(t_0 + h) - f(t_0)}{h}$$

es el vector velocidad media del móvil entre los tiempos t_0 y $t_0 + h$, y que cuando la diferencia entre los tiempos en que se toma esta velocidad media es cada vez más pequeña, debemos estar cada vez más cerca de lo que intuitivamente puede pensarse que es la velocidad del móvil en un instante determinado.

Análogamente, si $f'(t)$ existe para todo $t \in (a, b)$ y consideramos

$$f''(t_0) := \text{lím}\, h \to 0 \frac{f'(t_0 + h) - f'(t_0)}{h}$$

este límite, al que se llama derivada segunda de f en t_0, si existe, puede interpretarse físicamente como el vector aceleración instantánea del móvil en el instante t_0. Así, la aceleración de un móvil no es más que la velocidad con que varía su vector velocidad.

Por otro lado, si una curva $f(t)$ está parametrizada de modo que $\|f'(t)\| = 1$ para todo t (es decir se mueve con un vector de velocidad que podrá variar en su dirección pero no en su longitud, que permanece constante igual a la unidad), entonces la derivada segunda $f''(t)$ es un vector normal a la curva (es decir $\langle f''(t), f'(t)\rangle = 0$ para todo t), cuya norma $\|f''(t)\|$ se interpreta geométricamente como la *curvatura* de f en el punto $f(t)$. Por ejemplo, si $f(t) = (R\cos(t/R), R\,\text{sen}(t/R))$ (es decir f recorre una circunferencia de radio R con velocidad constante igual a uno) entonces $\|f''(t)\| = 1/R$ para todo t (la circunferencia tiene una curvatura igual al inverso de su radio: cuanto menor es éste mayor es la curvatura).

La noción de derivada para una función de varias variables reales, esto es, definida en un subconjunto abierto U de $\mathbb{R}^n$, es algo más sutil. Para empezar, si $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, el cociente

$$\frac{f(x + h) - f(x)}{h}$$

no tiene ningún sentido, puesto que h no es ya un número, sino un vector. Por otro lado la interpretación geométrica y física de la derivada de f, sea la que fuere, tampoco puede ser la misma, ya que la gráfica de f no es una curva, sino una superficie (en el caso $n = 2$) o una hipersuperficie (en el caso general).

Lo que sí podemos hacer es cortar la gráfica $G_f := \{(x, f(x)) : x \in U\} \subset \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1}$ (que puede visualizarse como una superficie de dimensión n en $\mathbb{R}^{n+1}$) por planos perpendiculares al hiperplano $\mathbb{R}^n \times \{0\}$ y determinados por una dirección $v \in \mathbb{R}^n$, es decir, planos de la forma $\{(x + tv, s) \in \mathbb{R}^n \times \mathbb{R} : t, s \in \mathbb{R}\}$, obteniendo así secciones de la gráfica de f que corresponden a las gráficas de las funciones

$$(-\varepsilon, \varepsilon) \subset \mathbb{R} \ni t \to f(x + tv) \in \mathbb{R},$$

cuyas derivadas sí que sabemos definir. Llegamos así al concepto de derivada direccional de una función de varias variables.

Definición 8.2. Sean $U \subseteq \mathbb{R}^n$ abierto, $a \in U$, $v \in \mathbb{R}^n$, $f : U \to \mathbb{R}^m$ una función. Se llama derivada direccional de f en a según el vector v a

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

cuando este límite existe.

Si v es uno de los vectores de la base canónica $\{e_1, e_2, ..., e_n\}$ de $\mathbb{R}^n$ entonces llamamos a $D_{e_i}f(a)$ derivada parcial i-ésima de f en el punto a, y se denota

$$\frac{\partial f}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(x)_{|x=a} = D_{e_i}f(a) = D_i f(a).$$

Así,

$$\frac{\partial f}{\partial x_i}(x) = \lim_{t \to 0} \frac{f(x_1, ..., x_{i-1}, x_i + t, x_{i+1}, ..., x_n) - f(x_1, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n)}{t}$$

cuando este límite existe. Por tanto, hallar $\frac{\partial f}{\partial x_i}$ no es otra cosa que derivar la expresión que define la función f respecto de la variable x_i solamente, considerando el resto de las x_j, $j \neq i$, como constantes.

Ejemplo 8.1. Calcular la derivada direccional $D_v f(a)$, donde $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \sin(x^2 + y^2)$, $a = (1, 3)$, $v = (1, 2)$. Calcular también las derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$ en el punto $(1, 3)$. Comprobar que

$$D_v f(a) = \begin{pmatrix} \frac{\partial f}{\partial x}(a) & \frac{\partial f}{\partial y}(a) \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Las derivadas direccionales no siempre existen. Por ejemplo, si $f(x, y) = \sqrt{|xy|}$, puede comprobarse que

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

y sin embargo $D_{(1,1)}f(0, 0)$ no existe. Este ejemplo muestra también que no basta con que existan las derivadas parciales para que la función tenga derivadas direccionales en toda dirección (que a su vez es un requisito indispensable para que la función sea diferenciable, como veremos).

Ejemplo 8.2. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

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

y consideremos $w = (u, v) \in \mathbb{R}^2$. Comprobar que

$$D_w f(0, 0) = \frac{u|v|}{\sqrt{u^2 + v^2}}$$

(y en particular $\frac{\partial f}{\partial x}(0, 0) = 0 = \frac{\partial f}{\partial y}(0, 0)$).

Si bien para $n \geq 2$ no tiene sentido la pregunta de cuál es la recta tangente a la gráfica de una función $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, puesto que G_f es una (hiper)superficie de dimensión n en $\mathbb{R}^{n+1}$, sí que podemos preguntarnos cuál es el subespacio afín de dimensión n que mejor aproxima la gráfica de f en un punto $(x, f(x))$. Para fijar ideas supongamos que, dada $f : \mathbb{R}^2 \to \mathbb{R}$, deseamos hallar el plano tangente a la gráfica de f en un punto $(x_0, y_0, f(x_0, y_0))$, si es que tal plano existe. Las intersecciones de los planos $x = x_0$ e $y = y_0$ con G_f determinan dos curvas paramétricas φ_1, φ_2 en $\mathbb{R}^3$, definidas por

$$\varphi_1(t) = (t, y_0, f(t, y_0)), \text{ y } \varphi_2(t) = (x_0, t, f(x_0, t)).$$

Si hallamos los vectores tangentes a estas curvas en $(x_0, y_0, f(x_0, y_0))$ obtenemos

$$\varphi_1'(x_0) = (1, 0, \frac{\partial f}{\partial x}(x_0, y_0)), \text{ y } \varphi_2'(y_0) = (0, 1, \frac{\partial f}{\partial y}(x_0, y_0)).$$

Si existe el plano tangente $\mathcal{P}$ que buscamos, debería estar generado por estos dos vectores (que al ser tangentes a curvas contenidas en G_f, deberían ser también tangentes a G_f). Nótese que estos vectores $\varphi_1'(x_0), \varphi_2'(y_0)$ son siempre linealmente independientes, luego ciertamente generan un plano en $\mathbb{R}^3$. Por tanto la ecuación implícita de $\mathcal{P}$ debería ser

$$z - f(x_0, y_0) = (x - x_0)\frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0)\frac{\partial f}{\partial y}(x_0, y_0). \tag{$*$}$$

El problema con este esquema es que las derivadas parciales $\frac{\partial f}{\partial x}(x_0, y_0)$ y $\frac{\partial f}{\partial y}(x_0, y_0)$ (o lo que es casi lo mismo, los vectores tangentes $\varphi_1'(x_0), \varphi_2'(y_0)$) pueden existir sin que exista de hecho el plano $\mathcal{P}$ que buscamos. En efecto, si f es la función del ejemplo 8.2, ambas derivadas parciales son cero, luego, de existir $\mathcal{P}$, debería ser $\mathcal{P} = \{z = 0\}$. Por otro lado, la curva paramétrica

$$\gamma(t) = (t, t, f(t, t)),$$

como está contenida en G_f, debe tener la propiedad de que $\gamma'(0) \in \mathcal{P}$. Pero

$$\gamma'(0) = (1, 1, D_{(1,1)}f(0, 0)) = (1, 1, \frac{1}{\sqrt{2}}) \notin \{z = 0\}.$$

Por tanto la superficie G_f no admite un plano tangente en el punto $(0, 0, 0)$. En particular vemos que no es posible dar una definición de plano tangente en $(x_0, y_0, f(x_0, y_0))$ pidiendo solo que existan las derivadas parciales de f en (x_0, y_0) (tampoco ni siquiera pidiendo que existan todas las derivadas direccionales en (x_0, y_0)).

Podría entonces sugerirse que el plano tangente $\mathcal{P}$ debería existir siempre que *todas* las derivadas direccionales $D_{(v_1,v_2)}f(x_0, y_0)$ existan y que *además* los vectores tangentes que ellas generan, a saber,

$$(v_1, v_2, D_{(v_1,v_2)}f(x_0, y_0)),$$

que son precisamente los vectores tangentes en $(x_0, y_0, f(x_0, y_0))$ a las curvas

$$\gamma_v(t) = (x_0 + tv_1, y_0 + tv_2, f(x_0 + tv_1, y_0 + tv_2)),$$

estén todos ellos contenidos en un mismo plano. A primera vista esta opción puede parecer una definición satisfactoria de la existencia de $\mathcal{P}$, que vendría entonces determinado por la ecuación $(*)$. Pero, mirándolo más despacio, todavía cabe preguntarse si esta exigencia nos garantizará que si $\gamma(t)$ es una curva diferenciable contenida en G_f, con $\gamma(0) = (x_0, y_0, f(x_0, y_0))$, entonces su recta tangente en este punto, es decir $\gamma(0) + t\gamma'(0)$, estará contenido en $\mathcal{P}$ (cualquier definición de plano tangente que pretenda satisfacer a un geómetra debería tener esta propiedad). Desgraciadamente esto no es así, como prueba el siguiente ejemplo.

Ejemplo 8.3. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ la función definida por

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

Entonces, si ponemos

$$\gamma_v(t) = (tv_1, tv_2, f(tv_1, tv_2)),$$

se comprueba sin dificultad que $\gamma_v'(0) = (v_1, v_2, 0)$ para todo $v \in \mathbb{R}^2$. En particular, todas las derivadas direccionales existen y son cero, y los vectores tangentes a las curvas γ_v en $(0,0,0)$ están contenidos todos ellos en el plano $z = 0$. Sin embargo, si consideramos la curva

$$\gamma(t) = (t^2, t, f(t^2, t)),$$

que claramente es derivable y está contenida en G_f, obtenemos que su vector tangente en el origen es

$$\gamma'(0) = (0, 1, \frac{1}{2}),$$

que no pertenece al plano $z = 0$.

Para salir de este atolladero es por tanto necesario imponer que siempre que γ sea una curva contenida en G_f y con $\gamma(0) = (x_0, y_0, f(x_0, y_0))$, su vector tangente en $\gamma(0)$, si existe, debe estar en un mismo plano fijo $\mathcal{P}$, cuya ecuación deberá ser de la forma $(*)$, es decir,

$$(x_0, y_0, f(x_0, y_0)) + \gamma'(0) \in \mathcal{P} = \{z - f(x_0, y_0) = \alpha(x - x_0) + \beta(y - y_0)\}$$

para toda curva de la forma

$$\gamma(t) = (\gamma_1(t), \gamma_2(t), f(\gamma_1(t), \gamma_2(t)))$$

(es decir, contenida en G_f), que sea derivable en 0, y tal que $\gamma(0) = (x_0, y_0, f(x_0, y_0))$. Esto significa que

$$\lim_{t\to 0} \frac{f(\gamma_1(t), \gamma_2(t)) - f(x_0, y_0)}{t} = \alpha \lim_{t\to 0} \frac{\gamma_1(t) - x_0}{t} + \beta \lim_{t\to 0} \frac{\gamma_1(t) - y_0}{t},$$

o lo que es igual,

$$\lim_{t\to 0} \frac{f(\gamma_1(t), \gamma_2(t)) - f(x_0, y_0) - (\alpha(\gamma_1(t) - x_0) + \beta(\gamma_1(t) - y_0))}{t} = 0.$$

Si denotamos $\varphi_1(t) = \gamma_1(t) - \gamma_1(0)$ y $\varphi_2(t) = \gamma_2(t) - \gamma_2(0)$, y ponemos

$$L(x, y) = \alpha x + \beta y,$$

la forma lineal $L : \mathbb{R}^2 \to \mathbb{R}$ cuya matriz es $(\alpha\ \beta)$, lo que estamos pidiendo es que

$$\lim_{t\to 0} \frac{f(x_0 + \varphi_1(t), y_0 + \varphi_2(t)) - f(x_0, y_0) - L(\varphi_1(t), \varphi_2(t))}{t} = 0$$

para toda curva $\varphi : \mathbb{R} \to \mathbb{R}^2$ que sea derivable en 0 y satisfaga $\varphi(0) = (0, 0)$. Supongamos además que $\gamma'(0) \neq 0$. Entonces tenemos

$$\lim_{t\to 0^+} \frac{t}{\|\varphi(t)\|} = \frac{1}{\| \lim_{t\to 0^+} \frac{\varphi(t)}{t} \|} = \frac{1}{\|\varphi'(0)\|} > 0,$$

luego, multiplicando esta igualdad por la anterior, obtenemos que

$$\lim_{t\to 0^+} \frac{f(x_0 + \varphi_1(t), y_0 + \varphi_2(t)) - f(x_0, y_0) - L(\varphi_1(t), \varphi_2(t))}{\|\varphi(t)\|} = 0$$

para toda curva $\varphi : \mathbb{R} \to \mathbb{R}^2$ tal que $\gamma'(0) \neq 0$ y $\gamma(0) = (0, 0)$.

De aquí a pedir que

$$\lim_{h\to 0} \frac{f(x_0+h_1, y_0+h_2) - f(x_0, y_0) - L(h)}{\|h\|} = 0$$

hay sólo un pequeño salto (sobre todo teniendo presente el problema 5.8). Precisamente el problema 8.13 se ocupa de formalizar este salto: para cualquier función $g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ y cualquier punto $x_0 \in U$ (donde U es abierto), se tiene que existe

$$\lim_{x\to x_0} g(x) = l$$

si y sólo si para toda curva $\varphi : \mathbb{R} \to U$ tal que $\varphi'(0) \neq 0$ y $\varphi(0) = x_0$ es

$$\lim_{t\to 0^+} g(\varphi(t)) = l.$$

Por tanto, si queremos que exista el plano tangente a la gráfica G_f de f en el punto (x_0, y_0) (entendido como un plano que contenga todos los vectores tangentes en $(x_0, y_0, f(x_0, y_0))$ a curvas diferenciables contenidas en la gráfica que pasen por este punto), debe existir una aplicación lineal $L : \mathbb{R}^2 \to \mathbb{R}$ tal que

$$\lim_{h\to 0} \frac{f(x_0+h_1, y_0+h_2) - f(x_0, y_0) - L(h)}{\|h\|} = 0.$$

Hemos llegado así a la definición de función diferenciable.

Definición 8.3. Sean U un abierto de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$, $x_0 \in U$. Se dice que la función f es diferenciable en x_0 si existe una aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^m$ tal que

$$\lim_{h\to 0} \frac{1}{\|h\|} \left(f(x_0+h) - f(x_0) - L(h)\right) = 0.$$

Si existe una aplicación lineal L con estas características, es única. De hecho se tiene lo siguiente.

Observación 8.1. *Supongamos que f es diferenciable en a, siendo $L : \mathbb{R}^n \to \mathbb{R}^m$ una aplicación lineal que satisface la definición anterior. Entonces:*

1. *L es la única aplicación lineal con esta propiedad.*

2. *Para todo $v \in \mathbb{R}^n$ existe la derivada direccional $D_v f(a)$, y*

$$D_v f(a) = L(v) = \sum_{i=1}^{n} v_i \frac{\partial f}{\partial x_i}(a).$$

Demostración. Supongamos $v \neq 0$. Como

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

se deduce en particular, tomando $h = tv$ y haciendo $t \to 0$, que

$$0 = \lim_{t\to 0} \frac{f(a+tv) - f(a) - tL(v)}{t} =$$
$$\lim_{t\to 0} \frac{f(a+tv) - f(a)}{t} - L(v) = D_v f(a) - L(v),$$

luego $L(v) = D_v f(a)$ para todo $v \in \mathbb{R}^n \setminus \{0\}$. Por otro lado, si $v = 0$ entonces es inmediato que $D_0 f(a) = 0$, y como L es lineal también $L(0) = 0$. En todo caso $L(v) = D_v f(a)$, y por la unicidad de $D_v f(a)$ (consecuencia a su vez de la unicidad del límite que la define), se deduce que L es única. Además, si $v = (v_1, ..., v_n)$, se tiene que

$$L(v) = L(\sum_{j=1}^n v_j e_j) = \sum_{j=1}^n v_j L(e_j) = \sum_{j=1}^n v_j D_{e_j} f(a) = \sum_{j=1}^n v_j \frac{\partial f}{\partial x_j}(a).$$

$\square$

Definición 8.4. Si $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $a \in U$, a la única aplicación lineal $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ que satisface

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

la llamaremos *diferencial de f en a*, y la denotaremos $L = Df(a)$. Así se tiene que

$$Df(a)(v) = D_v f(a)$$

para todo $v \in \mathbb{R}^n$.

Si f toma valores en $\mathbb{R}$, es decir, $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, y suponiendo que f sea diferenciable e $a \in U$, definiremos el plano tangente a la gráfica G_f de f en el punto $(a, f(a)) \in U \times \mathbb{R}$ por

$$\mathcal{T}G_{f\,(a,f(a))} = \{(x, z) \in \mathbb{R}^n \times \mathbb{R} : z - f(a) = \sum_{j=1}^n (x_j - a_j) \frac{\partial f}{\partial x_j}(a)\}.$$

En el caso de una función $z = f(x, y)$, la ecuación de este plano no es más que la ecuación $(*)$ propuesta más arriba. Puede también probarse que este plano es, en efecto, lo que buscábamos: la unión de todas las tangentes a curvas diferenciables contenidas en la gráfica de f en el punto dado. De hecho se tiene lo siguiente:

Teorema 8.2. *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable en un punto $p \in U$. Denotemos la gráfica de f por G_f. Entonces*

$$\begin{aligned}
&\{(x, z) \in \mathbb{R}^n \times \mathbb{R} : z - f(a) = \sum_{j=1}^n (x_j - a_j) \frac{\partial f}{\partial x_j}(a)\} = \\
&\{(p, f(p)) + (v, Df(p)(v)) : v \in \mathbb{R}^n\} = \\
&\{(p, f(p)) + \gamma'(0) \mid \gamma : (-1, 1) \to G_f \textit{ diferenciable }, \gamma(0) = (p, f(p))\}.
\end{aligned}$$

Por tanto cualquiera de estos conjuntos puede tomarse como definición de plano tangente a G_f en $(p, f(p))$. El primero de ellos nos proporciona la ecuación implícita del plano tangente, y el segundo las ecuaciones paramétricas. Nótese también que los vectores

$$\{\, (e_1, \frac{\partial f}{\partial x_1}),\ (e_2, \frac{\partial f}{\partial x_1}),\ ...,\ (e_n, \frac{\partial f}{\partial x_n})\,\}$$

forman una base del hiperplano vectorial H asociado al hiperplano afín $\mathcal{T}G_{f\,(p,f(p))} = (p, f(p)) + H$.

Demostración. La primera igualdad se comprueba fácilmente considerando $v = x - p$ y usando el hecho ya probado de que $Df(p)(v) = D_v f(p) = \sum_{j=1}^{n} v_j \frac{\partial f}{\partial x_j}(p)$. Probar el contenido del segundo conjunto en el tercero es también fácil considerando $\gamma(t) = (p + tv, f(p + tv))$.

Veamos que el último conjunto está contenido en el segundo. Dado $\gamma : (-1, 1) \to G_f$ diferenciable con $\gamma(0) = (p, f(p))$, podemos escribir $\gamma(t) = (\varphi(t), f(\varphi(t)))$, donde $\varphi : (-1, 1) \to U$ es diferenciable y $\varphi(0) = p$. Como f es diferenciable en p, φ es continua y $\varphi(0) = p$, se tiene que

$$\lim_{t\to 0} \frac{f(\varphi(t)) - f(p) - Df(p)(\varphi(t) - p)}{\|\varphi(t) - p\|} = 0,$$

lo que multiplicado por $\lim_{t\to 0} \|\frac{\varphi(t)-p}{t}\| = \|\varphi'(0)\|$ nos da

$$\lim_{t\to 0} \frac{f(\varphi(t)) - f(p) - Df(p)(\varphi(t) - p)}{t} = 0.$$

Por otro lado, al ser $Df(p)$ una aplicación lineal (y continua) y existir $\varphi'(0)$, tenemos

$$\lim_{t\to 0} \frac{Df(p)(\varphi(t) - p)}{t} = Df(p) \left(\lim_{t\to 0} \frac{\varphi(t) - p}{t} \right) = Df(p)(\varphi'(0)).$$

Sumando las dos últimas ecuaciones obtenemos

$$\lim_{t\to 0} \frac{f(\varphi(t)) - f(p)}{t} = Df(p)(\varphi'(0)),$$

y por tanto $\gamma'(0) = (\varphi'(0), Df(p)(\varphi'(0)))$, con lo que queda claro que el punto $(p, f(p)) + \gamma'(0)$ está en el segundo conjunto. $\square$

Si f es diferenciable en todos los puntos de U, diremos que f es diferenciable en U. En este caso la aplicación

$$U \ni x \mapsto Df(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$$

está bien definida, y a aplicación $Df : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ se la llama diferencial de f (sin hacer referencia a punto alguno).

Ejercicio 8.1. Sea $f(x, y) = x^2 + y^3 + 2x - y + 5$. Demostrar que f es diferenciable en $(0, 0)$, y hallar la ecuación del plano tangente a su gráfica en el punto $(0, 0, 5)$.

Para concluir la discusión que ha precedido la definición de función diferenciable, diremos que en los problemas 8.13, 8.14 y 8.15 se indica cómo probar que una función $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en x_0 si y sólo si para toda curva diferenciable $\gamma : (-1, 1) \to U$ con $\gamma(0) = x_0$ y $\gamma'(0) \neq 0$ existe

$$\frac{df(\gamma(t))}{dt}(0) = \lim_{t\to 0} \frac{f(\gamma(t)) - f(x_0)}{t}$$

y los vectores tangentes a las curvas $t \mapsto \varphi(t) = (\gamma(t), f(\gamma(t))) \in U \times \mathbb{R}$ en el punto $(x_0, f(x_0))$ están contenidos en un mismo hiperplano de $\mathbb{R}^{n+1}$.

Conviene observar asimismo que la aplicación $x \mapsto f(a) + Df(a)(x - a)$ nos proporciona la mejor aproximación afín a la función f entorno al punto a, en el sentido de que si llamamos $R(x - a)$ a la diferencia entre $f(x)$ y $f(a) + Df(x - a)$, entonces se tiene que

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

y además $x \mapsto f(a) + Df(a)(x - a)$ es la única aplicación afín que tiene esta propiedad. Como no podía ser de otra manera, cuando f es lineal, f es diferenciable en todos los puntos, y su diferencial es ella misma (problema 8.7).

Ejemplo 8.4. Estudiar si las funciones $f, g : \mathbb{R}^2 \to \mathbb{R}$ definidas por

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

y

$$g(x,y) = \begin{cases} \frac{x|y|}{\sqrt{x^2+y^2}} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

son o no son diferenciables en el origen.

También es preciso subrayar, aunque sea casi obvio, que la diferenciabilidad es una propiedad local de las funciones: si f y g son dos funciones que coinciden en un entorno de un punto a, y una de ellas es diferenciable en a, entonces la otra también lo es, y $Df(a) = Dg(a)$. Ver el problema 8.9.

Proposición 8.3. *Sea $f : (\alpha, \beta) \subseteq \mathbb{R} \to \mathbb{R}^m$ una función, $t_0 \in (\alpha, \beta)$, y denotemos $f = (f_1, ..., f_m)$, donde f_j son las funciones coordenadas de f. Las siguientes afirmaciones son equivalentes:*

1. *f es diferenciable en t_0.*
2. *f es derivable en t_0.*
3. *f_j es derivable en t_0 para todo $j = 1, 2, ..., m$.*

Además,

$$f'(t_0) = (f_1'(t_0), f_2'(t_0), ..., f_m'(t_0)).$$

Demostración. es fácil y queda a cargo del lector. □

La observación 8.1, combinada con la proposición anterior, nos proporciona una manera de hallar la matriz de la diferencial de una función $f : U \subseteq U \to \mathbb{R}^m$ en un punto $a \in U$ respecto de las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$. En efecto, la matriz de $Df(a)$ es

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & ... & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & ... & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & ... & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$$

y a esta matriz se la llama *matriz jacobiana* de f, y se denota por $Jf(a)$ o $f'(a)$. Así, mientras que $Df(a)$ representa la aplicación lineal de $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ que satisface la condición de diferenciabilidad de f en a, $f'(a)$ denota la matriz de esta aplicación lineal respecto de las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$.

Cuando la función f toma valores reales hay solamente una función coordenada y entonces tenemos que la matriz $f'(a)$ es una matriz de una única fila y con n columnas, que identificaremos con un vector. Así, en el caso de $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, se tendrá

$$f'(a) = \left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), ..., \frac{\partial f}{\partial x_n}(a) \right),$$

y a este vector se le llama *vector gradiente* de f en a, y suele denotarse también

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

Ejemplo 8.5. Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$ definida por

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

Asumiendo que f es diferenciable en cualquier punto $(x, y, z) \in \mathbb{R}^3$, calcular la matriz jacobiana de f en (x, y, z), y hallar $Df(x, y, z)(u, v, w)$ para cualquier $(u, v, w) \in \mathbb{R}^3$.

Concluiremos este capítulo con la observación de que una función $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable si y sólo si lo son sus m funciones coordenadas.

Proposición 8.4. *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$. La función $f = (f_1, ..., f_m)$ es diferenciable en $a \in U$ si y sólo si f_j es diferenciable en a para cada $j = 1, 2, ..., n$, y en este caso*

$$Df(a)(v) = (Df_1(a)(v), ..., Df_m(a)(v))$$

para cada $v \in \mathbb{R}^n$.

Demostración. es fácil y queda a cargo del lector. □

8.1. Problemas

Problema 8.1. Calcúlense las derivadas direccionales de las funciones siguientes en los puntos y direcciones indicados:

1. $f(x, y) = x \arctan(\frac{x}{y})$ en el punto $(1, 1)$; dirección $(1/\sqrt{2}, 1/\sqrt{2})$.
2. $f(x, y) = e^x \cos \pi y$ en el punto $(0, -1)$; dirección $(\frac{-1}{\sqrt{5}}, \frac{2}{\sqrt{5}})$.
3. $f(x, y, z) = e^x + yz$, en el punto (1,1,1); dirección $(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$.

Problema 8.2. Estudiar la existencia de derivadas direccionales en el origen para las funciones siguientes.

1. $f(x, y) = \sqrt{xy}$;
2. $f(x, y) = \max\{|x|, |y|\}$.

Problema 8.3. Asumiendo la diferenciabilidad de todas ellas, calcúlese la matriz jacobiana en el origen de las funciones siguientes:

$$\begin{array}{cc} f(x,y) = (e^x + \sin y, x^2 \cos y) & g(x,y) = (\sin x + \log(1+y^2), \cos(xy)) \\ h(x,y,z) = e^{x^2+y^2+z^3} & \phi(t) = (\tan t, \sin t, e^t) \\ \varphi(u,v) = (u^2v^3, u^3v, u^4v^2) & \psi(x,y,z) = (x^4 y \cos z, xe^z). \end{array}$$

Problema 8.4. Sea $f(x, y) = x^{x^{x^y}} + (\sin \pi x) \arctan(\arctan(\sin\cos(xy) - \log(x + y))))$. Calcúlese $\frac{\partial f}{\partial y}(1, y)$.

Problema 8.5. Calcúlese la derivada direccional de la función

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

en el punto (1,3,2) según la dirección del vector (1,-1,0). ¿En qué dirección es máxima la derivada direccional? ¿Cuál es el valor de la derivada direccional máxima?

Problema 8.6. Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función constante. Probar que f es diferenciable, con $Df(x) = 0$ para todo $x \in U$.

Problema 8.7. Sea $L : \mathbb{R}^n \to \mathbb{R}^m$ una aplicación lineal. Probar que L es diferenciable en todo punto, y que $DL(x) = L$ para todo $x \in \mathbb{R}^n$.

Problema 8.8. Sea $L : \mathbb{R}^n \to \mathbb{R}^m$ una aplicación lineal, y sea $h : \mathbb{R}^n \to \mathbb{R}^m$ una aplicación tal que, para algún $M \geq 0$,

$$\|h(x)\| \leq M\|x\|^2$$

para todo $x \in \mathbb{R}^n$. Defínase $f(x) = L(x) + h(x)$. Demostrar que f es diferenciable en 0. ¿Cuál es $Df(0)$?

Problema 8.9. Sean $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ dos funciones que coinciden en un entorno de un punto $a \in U$ (es decir existe $r > 0$ tal que $f(x) = g(x)$ para todo $x \in B(a, r)$). Supongamos que g es diferenciable en $a \in U$. Probar que entonces f es también diferenciable en a, y $Df(a) = Dg(a)$. *Indicación:* recordar el problema 5.3.

Problema 8.10. Sea f una aplicación entre dos subconjuntos abiertos de dos espacios normados de dimensión finita. Demostrar que el hecho de que f sea o no diferenciable en un punto a, y en su caso la diferencial $Df(a)$, no dependen de las normas consideradas en estos espacios. *Indicación:* Usar que en cualquier espacio de dimensión finita todas las normas son equivalentes.

Problema 8.11. Estudiar la continuidad, existencia de derivadas direccionales y diferenciabilidad en $(0, 0)$ de las siguientes funciones.

$$\begin{cases} \frac{xy^2}{x^2+y^4} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0); \end{cases} \qquad \begin{cases} \frac{x|y|}{(x^2+y^2)^{1/2}} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0); \end{cases}$$

$$\begin{cases} x\sin(\frac{1}{x^2+y^2}) & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0); \end{cases} \qquad \begin{cases} (x^2+y^2)\sin\frac{1}{y^2} & \text{si } y \neq 0, \\ 0 & \text{si } y = 0; \end{cases}$$

$$\begin{cases} \frac{xy^3}{x^2+y^6} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0); \end{cases} \qquad \begin{cases} x+y & \text{si} xy = 0, \\ 0 & \text{si } xy \neq 0. \end{cases}$$

Problema 8.12. Si $B : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ es una aplicación bilineal, probar que B es diferenciable en todo punto (x, y) de $\mathbb{R}^n \times \mathbb{R}^m$, y que

$$DB(x,y)(h,k) = B(x,k) + B(h,y)$$

para todo $(h, k) \in \mathbb{R}^n \times \mathbb{R}^m$. *Indicación:* Probar primero que existe $M \geq 0$ tal que $\|B(x, y)\| \leq M\|x\|\|y\|$ para todo $x \in \mathbb{R}^n, y \in \mathbb{R}^m$.

Problema 8.13. Sea U un abierto de $\mathbb{R}^n$, $x_0 \in U$, $g : U \to \mathbb{R}$. Demostrar que existe

$$\lim_{x \to x_0} g(x) = l$$

si y sólo si para toda curva $\varphi : (-1, 1) \subset \mathbb{R} \to U$ tal que $\varphi'(0) \neq 0$ y $\varphi(0) = x_0$ existe

$$\lim_{t \to 0^+} g(\varphi(t)) = l.$$

Indicación: Por el problema 5.8 ya sabemos que si existe el límite, también existen todos los límites a lo largo de curvas continuas que pasan por x_0, y son todos iguales. Para probar el recíproco, supóngase que existen $\varepsilon > 0$ y una sucesión (x_k) que tiende a $x_0 = 0$ y tal que $\|g(x_k) - l\| \geq \varepsilon$ para todo k. Definamos $z_k = x_k/\|x_k\|$ para cada k. La sucesión (z_k) está en la esfera unidad. Por tanto, tomando una subsucesión si fuera preciso, podemos suponer que existe z_0 tal que $\lim_{k\to\infty} z_k = z_0$. Tomando otra vez subsucesiones si fuera preciso, podemos suponer también que $\|x_{k+1}\| < \|x_k\|$ para todo k. El objetivo es construir una curva continua $\gamma, \varphi : (-1,1) \to U$, diferenciable en 0 y con $\gamma'(0) \neq 0$, $\gamma(0) = 0$, tal que $\lim_{t\to 0} g(\gamma(t)) \neq l$, llegando así a una contradicción. Para construir γ, póngase

$$\gamma(t) = \frac{t}{\|(t_k - t)x_{k+1} + (t - t_{k+1})x_k\|}\left[(t_k - t)x_{k+1} + (t - t_{n+1})x_k\right]$$

si $t \in [t_{k+1}, t_k]$, donde $t_k = \|x_k\|$, y defínase

$$\gamma(t) = tz_0$$

si $t \leq 0$. Pruébese que $\gamma(t_k) = x_k$ para todo k, y por tanto $\lim_{k\to\infty} g(\gamma(t_k)) \neq l$. Demuéstrese también que

$$\gamma'(0) = \lim_{t\to 0} \frac{\gamma(t)}{t} = z_0 \neq 0.$$

Problema 8.14. Piénsese cómo podría modificarse la demostración anterior para obtener el siguiente resultado. Si U es un abierto de $\mathbb{R}^n$, $g : U \to \mathbb{R}$, $x_0 \in U$, entonces existe $\lim_{x\to x_0} g(x) = l$ si y sólo si para todas las curvas diferenciables $\gamma : (-1,1) \to U$ con $\gamma'(0) \neq 0$ y $\gamma(0) = x_0$ existen los límites $\lim_{t\to 0} g(\gamma(t))$ y son todos iguales a l.

Problema 8.15. Deducir de los problemas anteriores, y de la discusión del principio de este capítulo, que una función $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ es diferenciable en x_0 si, y sólo si, para toda curva diferenciable $\gamma : (-1,1) \to U$ con $\gamma(0) = x_0$ y $\gamma'(0) \neq 0$ existe

$$\frac{df(\gamma(t))}{dt}(0) = \lim_{t\to 0} \frac{f(\gamma(t)) - f(x_0)}{t}$$

y los vectores tangentes a las curvas $t \mapsto \varphi(t) = (\gamma(t), f(\gamma(t))) \in U \times \mathbb{R}$ en el punto $(x_0, f(x_0))$ están, todos ellos, contenidos en un mismo hiperplano de $\mathbb{R}^{n+1}$.

Problema 8.16. Sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ una función par (es decir $f(x) = f(-x)$) diferenciable en el origen. Calcúlese $Df(0)$.

Problema 8.17. ¿Son tangentes las gráficas de las funciones $f(x,y) = x^2 + y^2$ y $f(x,y) = x^2 - y^2 + xy^3$ en el punto $(0,0)$?

Problema 8.18. Una partícula sigue la trayectoria $\gamma(t) = (t^2, t^3 - 4t, 0)$, y se sale por la tangente en el instante $t = 2$. ¿Cuál es su velocidad en este momento? Calcular su posición en el instante $t = 3$.

Problema 8.19. Supóngase que $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ tiene un máximo o un mínimo local en un punto $a \in U$, y que existe la derivada direccional $D_v f(a)$ en una dirección $v \in \mathbb{R}^n$. Probar que $D_v f(a) = 0$.

Problema 8.20. Sea $f(r,\theta) = (r\cos\theta, r\sin\theta)$. Calcular la matriz jacobiana de f asumiendo que es diferenciable en todo punto (r,θ).

Problema 8.21. Demostrar que si $T : \mathbb{R}^m \to \mathbb{R}^p$ es lineal y $f : \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en a entonces $L \circ f$ es diferenciable en a y

$$D(L \circ f)(a) = L \circ Df(a).$$

Problema 8.22. Sea $f : \mathbb{R}^3 \to \mathbb{R}$ definida por

$$f(x,y,z) = \begin{cases} \frac{xyz}{(x^2+y^2+z^2)^\alpha} & \text{si } (x,y,z) \neq (0,0,0), \\ 0 & \text{si } (x,y,z) = (0,0,0). \end{cases}$$

Demostrar que f es diferenciable en $(0,0,0)$ si y sólo si $\alpha < 1$.

Problema 8.23. Sea $f : \mathbb{R}^n \to \mathbb{R}^m$ diferenciable en a, y definamos $g(x) = f(x+a)$. Probar que g es diferenciable en 0 y que $Dg(0) = Df(a)$.

Capítulo 9

Funciones diferenciables

En este capítulo estudiamos las propiedades elementales de las funciones diferenciables (tales como su comportamiento respecto de la suma, multiplicación, cociente y composición). Se discuten también condiciones suficientes para asegurar la diferenciabilidad de una función, y criterios para descartarla. Como consecuencia de la regla de la cadena se presentan asimismo algunos resultados de carácter geométrico acerca de la relación entre las superficies o curvas de nivel de una función de valores reales y su gradiente.

Una primera observación que debe hacerse acerca de las funciones diferenciables es que son continuas.

Proposición 9.1. *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función diferenciable en $a \in U$. Entonces existen $r, K > 0$ tales que*

$$\|f(x) - f(a)\| \leq K\|x - a\|$$

para todo $x \in B(a, r)$. En particular f es continua en a.

Nota: Esto *no* quiere decir que f sea Lipschitz en un entorno de a; de hecho f podría ser continua sólo en a y discontinua en todos los demás puntos; ver el problema 9.1.

Demostración. Denotemos $L = Df(a)$, y sea $\|L\|$ la norma de esta aplicación lineal. Tomando $\varepsilon = 1$ en la definición de diferenciabilidad de f en a, obtenemos que existe $\delta > 0$ tal que si $\|x - a\| \leq \delta$ entonces

$$\|f(x) - f(a) - L(x - a)\| \leq \varepsilon\|x - a\|,$$

y por tanto también

$$\|f(x) - f(a)\| \leq \|L(x - a)\| + \varepsilon\|x - a\| \leq (\|L\| + \varepsilon)\|x - a\| := K\|x - a\|$$

si $x \in B(a, \delta)$. $\square$

Ya sabemos (problemas 8.6 y 8.7) que toda función constante es diferenciable, y lo mismo ocurre con las lineales. A continuación vamos a ver varios resultados que nos permitirán ampliar muchísimo el repertorio de ejemplos genéricos de funciones diferenciables.

Proposición 9.2. *Sean $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ funciones diferenciables en un punto $a \in U$. Entonces:*

1. $f + g$ es diferenciable en a, y $D(f + g)(a) = Df(a) + Dg(a)$.

2. *Para cada $\lambda \in \mathbb{R}$, λf es diferenciable en a, con $D(\lambda f)(a) = \lambda Df(a)$.*

La demostración es un simple ejercicio, y también puede deducirse, como veremos después de la regla de la cadena.

Con el producto ocurre otro tanto: cuando se pueden multiplicar dos funciones diferenciables, su producto es diferenciable, y hay una regla similar a la ya conocida para funciones de una variable. Puede darse una demostracion directa de este hecho, pero resultaría poco económico, ya que no sería del todo inmediata, mientras que sí que puede deducirse trivialmente de la regla de la cadena, como veremos más abajo.

Pasamos por tanto ya a estudiar la regla para diferenciar una composición de funciones.

Teorema 9.3 (Regla de la cadena). *Sean U y V abiertos de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente, y $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $g : V \to \mathbb{R}^p$ aplicaciones, con $f(U) \subseteq V$. Supongamos que f es diferenciable en a, y que g es diferenciable en $b = f(a)$. Entonces $g \circ f : U \to \mathbb{R}^p$ es diferenciable en a, y además*

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

Demostración. Como f es diferenciable en a, existen $K, \delta_1 > 0$ tal que

$$\|f(x) - f(a)\| \leq K\|x - a\| \tag{1}$$

si $\|x - a\| \leq \delta_1$. Además, si ponemos $M = \|Dg(b)\|$, se tiene

$$\|Dg(b)(v)\| \leq M\|v\| \tag{2}$$

para todo $v \in \mathbb{R}^m$. Fijemos $\varepsilon > 0$. Como g es diferenciable en b, existe $\delta_2 > 0$ tal que

$$\|g(y) - g(b) - Dg(b)(y - b)\| \leq \frac{\varepsilon}{M + K}\|y - b\| \tag{3}$$

si $\|y - b\| \leq \delta_2$. Por otro lado, de la diferenciabilidad de f en a se deduce la existencia de $\delta_3 > 0$ tal que

$$\|f(x) - f(a) - Df(a)(x - a)\| \leq \frac{\varepsilon}{M + K}\|x - a\| \tag{4}$$

si $\|x - a\| \leq \delta_3$. Pongamos $\delta := \min\{\delta_1, \delta_2/K, \delta_3\}$. Entonces, si $\|x - a\| \leq \delta$ se tiene que

$$\|f(x) - f(a)\| \leq K\|x - a\| \leq K\delta \leq \delta_2,$$

luego, combinando las desigualdades anteriores (y sumando y restando el vector $Dg(b)(f(x) - f(a))$ dentro de la norma de la primera línea de las que siguen, y aplicando también la linealidad de $Dg(b)$), obtenemos que

$$\begin{aligned}
&\|g(f(x)) - g(f(a)) - Dg(f(a))(Df(a)(x - a))\| \leq \\
&\|g(f(x)) - g(f(a)) - Dg(f(a))(f(x) - f(a))\| + \\
&+\|Dg(f(a))\big(f(x) - f(a) - Df(a)(x - a)\big)\| \leq \\
&\frac{\varepsilon}{M + K}\|f(x) - f(a)\| + M\|f(x) - f(a) - Df(a)(x - a)\| \leq \\
&\frac{\varepsilon}{M + K}K\|x - a\| + M\frac{\varepsilon}{M + K}\|x - a\| = \varepsilon\|x - a\|.
\end{aligned}$$

Así, hemos visto que para todo $\varepsilon > 0$ existe $\delta > 0$ tal que si $\|x - a\| \leq \delta$ entonces

$$\|g(f(x)) - g(f(a)) - Dg(f(a))(Df(a)(x - a))\| \leq \varepsilon\|x - a\|,$$

lo cual significa que

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

es decir que $g\circ f$ es diferenciable en a, con

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

$\square$

Observación 9.4. *Teniendo en cuenta que la operación de composición de aplicaciones lineales se traduce en multiplicación de sus matrices, la igualdad $D(g\circ f)(a) = Dg(f(a))\circ Df(a)$ significa que*

$$(g\circ f)'(a) = g'(f(a))f'(a),$$

es decir

$$\begin{pmatrix} \frac{\partial(g\circ f)_1}{\partial x_1}(a) & \dots & \frac{\partial(g\circ f)_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial(g\circ f)_p}{\partial x_1}(a) & \dots & \frac{\partial(g\circ f)_p}{\partial x_n}(a) \end{pmatrix} =$$

$$\begin{pmatrix} \frac{\partial g_1}{\partial y_1}(f(a)) & \dots & \frac{\partial g_1}{\partial y_m}(f(a)) \\ \vdots & & \vdots \\ \frac{\partial g_p}{\partial y_1}(f(a)) & \dots & \frac{\partial g_p}{\partial y_m}(f(a)) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix},$$

o de manera más compacta,

$$\frac{\partial(g\circ f)_j}{\partial x_i} = \sum_{k=1}^{m} \frac{\partial g_j}{\partial y_k}(f(a))\frac{\partial f_k}{\partial x_i}(a).$$

En el caso en que $f = \gamma : I \subseteq \mathbb{R} \to V \subseteq \mathbb{R}^m$ es una curva, se obtiene que la derivada de la curva $g\circ\gamma : I \to \mathbb{R}^p$ en un punto t viene dada por

$$(g\circ\gamma)'(t) = Dg(\gamma(t))(\gamma'(t)).$$

Si además la función g toma valores reales, es decir, $p = 1$, la fórmula anterior para la derivada de $g\circ\gamma : I \to \mathbb{R}$ es

$$(g\circ\gamma)'(t) = \langle \nabla g(\gamma(t)), \gamma'(t)\rangle.$$

Si denotamos las funciones coordenadas de $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ por

$$y_j(x_1, ..., x_n), j = 1, ..., m,$$

y las de $g : \mathbb{R}^m \to \mathbb{R}^p$ por

$$z_k(y_1, ..., y_m), k = 1, ..., p,$$

la regla de la cadena puede escribirse, con la notación clásica de Leibnitz para las derivadas, como

$$\frac{\partial z_k}{\partial x_i} = \sum_{j=1}^{m} \frac{\partial z_k}{\partial y_j}\frac{\partial y_j}{\partial x_i},$$

donde $\partial z_k/\partial x_i = \partial(g\circ f)_k/\partial x_i)(a)$, $\partial z_k/\partial y_j = (\partial g_k/\partial y_j)(f(a))$, y $\partial y_j/\partial x_i = (\partial f_j/\partial x_i)(a)$.

Sean $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ dos funciones diferenciables en un punto $a \in U$. Consideremos también las funciones $\Phi = (f, g) : U \to \mathbb{R}^m \times \mathbb{R}^m$ y $S : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ definidas por

$$\Phi(x) = (f(x), g(x)), \quad \text{y} \quad S(u, v) = u + v.$$

La función Φ es diferenciable en a al serlo f y g, y $D\Phi(x) = (Df(x), Dg(x))$. La función S es lineal y por tanto diferenciable en todos los puntos, con $dS(u, v) = S$ para todo $u, v \in \mathbb{R}^m$. Si aplicamos la regla de la cadena a función $S \circ \Phi$ obtenemos que

$$\begin{aligned} &D(f+g)(a) = D(S \circ \Phi)(a) = DS(f(a), g(a)) \circ D\Phi(a) = \\ &S(D\Phi(a)) = S(Df(a), Dg(a)) = Df(a) + Dg(a), \end{aligned}$$

redemostrando así la primera parte de la Proposición 9.2.

Supongamos ahora que f y g toman valores en $\mathbb{R}$, y sea $P : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$P(x, y) = xy.$$

Es evidente que P es una forma bilineal, luego por el problema 8.12 sabemos que P es diferenciable, y que

$$DP(x, y)(h, k) = P(x, k) + P(h, y)$$

(si no se ha hecho el problema 8.12, tanto la fórmula como el hecho de la diferenciabilidad de P, pueden deducirse de que las derivadas parciales de P existen y son continuas, según el criterio de diferenciabilidad que probaremos a continuación de estos ejemplos). Aplicando esta fórmula y la regla de la cadena a la composición $P \circ \Phi$, deducimos que

$$\begin{aligned} &D(fg)(a) = D(P \circ \Phi)(a) = DP(f(a), g(a))(Df(a), Dg(a)) = \\ &f(a)Dg(a) + g(a)Df(a), \end{aligned}$$

la regla para diferenciar el producto de dos aplicaciones diferenciables.

El mismo argumento sigue siendo válido si se sustituye P por una aplicación bilineal cualquiera $B : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^p$. En particular, si ponemos $B = \langle \cdot, \cdot \rangle : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$, el producto escalar usual de $\mathbb{R}^m$, y $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, obtenemos la fórmula

$$D\big(\langle f, g\rangle\big)(a)(\cdot) = \langle f(a), Dg(a)(\cdot)\rangle + \langle g(a), Df(a)(\cdot)\rangle.$$

Sinteticemos estas conclusiones en una Proposición.

Proposición 9.5. *Sean $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ funciones diferenciables en $a \in U$. Entonces su producto fg es también diferenciable en a, y*

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

Si f y g toman valores vectoriales en $\mathbb{R}^m$ entonces su producto escalar $\langle f, g\rangle$ es diferenciable en a, y

$$D\big(\langle f, g\rangle\big)(a)(\cdot) = \langle f(a), Dg(a)(\cdot)\rangle + \langle g(a), Df(a)(\cdot)\rangle.$$

Para finalizar con las operaciones más usuales con funciones diferenciables, estudiemos qué ocurre con el cociente. Sea $V = \{(x, y) \in \mathbb{R}^2 : y \neq 0\}$, y definamos $Q : V \to \mathbb{R}$ por

$$Q(x, y) = \frac{x}{y}.$$

La función Q tiene derivadas parciales continuas en el abierto V, a saber,

$$\frac{\partial Q}{\partial x} = \frac{1}{y}, \quad \frac{\partial Q}{\partial y} = -\frac{x}{y^2},$$

y por tanto, por el criterio que veremos a continuación, Q es diferenciable en todo punto de V, con

$$\nabla Q(x,y) = (\frac{1}{y}, -\frac{x}{y^2}).$$

En particular su diferencial viene dada por

$$DQ(x,y)\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \frac{1}{y} & -\frac{x}{y^2} \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = \frac{uy - xv}{y^2}.$$

Aplicando esta fórmula y la regla de la cadena a $f/g = Q \circ (f,g)$, obtenemos que si f y g son diferenciables en a, y asumimos que $g(a) \neq 0$ (lo que implica también que $g(x) \neq 0$ para todo x en un entorno de a, al ser g continua en a por se diferenciable en este punto), entonces el cociente f/g es diferenciable en a, con

$$D(\frac{f}{g})(a) = \frac{1}{g(a)^2}\big(g(a)Df(a) - f(a)Dg(a)\big).$$

Proposición 9.6. *Sean $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciables en $a \in U$, y supongamos que $g(a) \neq 0$. Entonces f/g es diferenciable en a, y*

$$D(\frac{f}{g})(a) = \frac{1}{g(a)^2}\big(g(a)Df(a) - f(a)Dg(a)\big).$$

Probemos ahora el criterio de diferenciabilidad ya anunciado, a saber, que una función con derivadas parciales continuas es diferenciable. A pesar de que no nos da una condición necesaria (ver los ejemplos en la Observación 9.9 más abajo), sino solamente suficiente, seguramente es el más empleado en la práctica de todos cuantos criterios de diferenciabilidad sean conocidos. Por otra parte, que la mera existencia de derivadas parciales no es suficiente para asegurar la diferenciabilidad no debería sorprendernos: en efecto, aparte de que exista la matriz jacobiana o el vector gradiente, la definición de diferenciabilidad exige que la propiedad de aproximación afín sea satisfecha. La existencia de derivadas parciales nos permite asegurar que, para una función $f : \mathbb{R}^2 \to \mathbb{R}$, la gráfica de la función parcial $\mathbb{R} \ni x \to f(x, a_2)$ se aproxima bien en un entorno de (a_1, a_2) por la recta tangente de pendiente $(\partial f/\partial x)(a_1, a_2)$, y lo mismo ocurre con la función $y \to f(a_1, y)$ y la derivada parcial $(\partial f/\partial y)(a_1, a_2)$. Pero no hay razón alguna para pensar que esto pueda llevar a ninguna conclusión acerca del comportamiento local de f en cualquier otra dirección. Sin embargo, cuando las derivadas parciales son continuas, $(\partial f/\partial x)(b)$ está próxima a $(\partial f/\partial x)(a)$ cuando b está cerca de a, luego $(\partial f/\partial x)(a)$ nos da también información local sobre $f(x, b_2)$, y lo mismo ocurre con $\partial f/\partial y$. Resumiendo, tenemos buena información sobre $f(x,y)$ para (x,y) en un pequeño rectángulo alrededor de $a = (a_1, a_2)$, y esto nos permite asegurar la diferenciabilidad de f en a.

Teorema 9.7. *Sea U un abierto de $\mathbb{R}^n$, y $f : U \to \mathbb{R}^m$. Supongamos que todas las derivadas parciales $\partial f_j/\partial x_i$ existen en un entorno de un punto a, y que son continuas en a. Entonces f es diferenciable en a.*

Demostración. Podemos suponer $m = 1$, gracias a la Proposición 8.4. Escribamos

$$\begin{aligned}
&f(x) - f(a) = \\
&f(x_1, ..., x_n) - f(a_1, x_2, ..., x_n) + \\
&f(a_1, x_2, ..., x_n) - f(a_1, a_2, x_3, ..., x_n) + \\
&f(a_1, a_2, x_3, ..., x_n) - f(a_1, a_2, a_3, x_4, ..., x_n) + ... + \\
&f(a_1, a_2, ..., a_{n-1}, x_n) - f(a_1, a_2, ..., a_n).
\end{aligned}$$

Por el teorema del valor medio para funciones de una variable, se tiene que

$$f(x_1, ..., x_n) - f(a_1, x_2, ..., x_n) = \frac{\partial f}{\partial x_1}(u_1, x_2, ..., x_n)$$

para cierto $u_1 = u_1(x) \in [a_1, x_1]$ o $[x_1, a_1]$. Aplicando el teorema del valor medio de la misma manera al resto de las expresiones en cada línea de la igualdad de arriba para $f(x) - f(a)$, obtenemos que

$$\begin{aligned}
&f(x) - f(a) = \\
&\frac{\partial f}{\partial x_1}(u_1, x_2, ..., x_n)(x_1 - a_1) + \\
&\frac{\partial f}{\partial x_2}(a_1, u_2, x_3, ..., x_n)(x_2 - a_2) + ... + \\
&\frac{\partial f}{\partial x_n}(a_1, a_2, ..., a_{n-1}, u_n)(x_n - a_n).
\end{aligned}$$

Puesto que $f(a)'(x - a) = \sum_{i=1}^{n} (\partial f/\partial x_i)(a_1, ..., a_n)(x_i - a_i)$, se obtiene que

$$\begin{aligned}
&|f(x) - f(a) - Df(a)(x - a)| \leq \\
&\{\left|\frac{\partial f}{\partial x_1}(u_1, x_2, ..., x_n) - \frac{\partial f}{\partial x_1}(a_1, a_2, ..., a_n)\right| + ... + \\
&\left|\frac{\partial f}{\partial x_n}(a_1, a_2, ..., a_{n-1}, u_n) - \frac{\partial f}{\partial x_n}(a_1, a_2, ..., a_n)\right|\}\|x - a\|
\end{aligned}$$

(usando la desigualdad triangular y el hecho de que $|x_i - a_i| \leq \|x - a\|$). Como las n derivadas parciales $\partial f/\partial x_i$ son continuas y u_i está entre x_i y a_i, para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que el término entre llaves en la expresión anterior es menor que ε para todo x con $\|x - a\| < \delta$.

Así se ha visto que para todo $\varepsilon > 0$ existe $\delta > 0$ de modo que si $\|x - a\| < \delta$ entonces

$$|f(x) - f(a) - f'(a)(x - a)| < \varepsilon\|x - a\|,$$

lo cual quiere decir que f es diferenciable en a. □

Observación 9.8. Un análisis cuidadoso de la demostración anterior indica que basta con que todas las derivadas parciales, salvo quizás $\partial f/\partial x_n$ (y por tanto salvo quizás una cualquiera de ellas, si reordenamos las variables) sean continuas en a. En efecto, en el último término de la suma telescópica en la que se descompone $f(x) - f(a)$, no hace falta aplicar el teorema del valor medio y la continuidad de la última derivada parcial, basta con usar la definición de derivada parcial en este caso para que todo cuadre.

Ejemplo 9.1. Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$ definida por

$$f(x, y, z) = (x^2 + \sin(yz), xy \cos z).$$

Calcular las derivadas parciales de f y ver que son continuas en $\mathbb{R}^3$. Deducir que f es diferenciable en todo punto de $\mathbb{R}^3$.

Como en el ejemplo anterior, el Teorema 9.7 nos permite asegurar la diferenciabilidad de funciones bien definidas mediante expresiones *razonables*.

Observación 9.9. La condición de diferenciabilidad del Teorema 9.7 no es necesaria: existen funciones diferenciables en un punto ninguna de cuyas derivadas parciales son continuas. Por ejemplo, consideremos

$$f(x, y) = \begin{cases} x^2 + y^2 & \text{si } x^2 + y^2 \in \mathbb{Q}, \\ 0 & \text{en caso contrario.} \end{cases}$$

La función f satisface que $|f(x, y)| \leq x^2 + y^2$ y por tanto es diferenciable en $(0, 0)$ (ver problema 8.8). Sin embargo las derivadas parciales de f no existen en ningún punto diferente de $(0, 0)$, y por tanto mucho menos son continuas.

Por otro lado, la existencia de derivadas parciales en un punto, o en todos los puntos, incluso la existencia de derivadas direccionales en todos los puntos, por sí sola tampoco es suficiente para garantizar la diferenciabilidad de una función: recuérdese el ejemplo 8.3 del capítulo anterior,

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

Esta función tiene derivadas direccionales en todos los puntos. Sin embargo la función no es difercnciable en $(0, 0)$. En efecto, si lo fuera su diferencial debería ser 0 (ya que las derivadas parciales en el origen son ambas cero, y por tanto su vector gradiente debería ser cero), y entonces tendríamos

$$0 = \lim_{(x,y)\to(0,0)} \frac{f(x, y)}{\sqrt{x^2 + y^2}} = \lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2 + y^4};$$

pero es fácil ver que este límite no existe (considérense los límites a lo largo de las curvas $x = y^2$ por un lado y $x = y$ por otro). Obsérvese también que las derivadas parciales de f solo dejan de ser continuas en el origen. Por tanto, en $\mathbb{R}^2 \setminus \{0\}$ la función f es diferenciable.

Corolario 9.10. *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}_m$. Consideremos las siguientes condiciones.*

1. *f tiene derivadas parciales en todos los puntos de U y son todas ellas continuas en U.*

2. *f es diferenciable en U*

3. *f tiene derivadas parciales en todos los puntos de U.*

Entonces se tiene que $(1) \implies (2) \implies (3)$*, pero los conversos son en general falsos.*

Resumiendo, podría decirse que la diferenciabilidad es un concepto que ocupa un lugar impreciso entre la continuidad de las derivadas parciales y su mera existencia.

Definición 9.1. Se dice que $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es de clase C^1 en U si todas las derivadas parciales de f existen y son continuas en U.

Así toda función de clase C^1 en U es diferenciable, pero el recíproco no es cierto en general.

Estudiaremos ahora algunas otras consecuencias de la regla de la cadena, particularmente la relación geométrica entre los conjuntos de nivel de una función y su gradiente.

Definición 9.2. Sea $F : U \subseteq \mathbb{R}^n \to \mathbb{R}$. Se llama *conjunto de nivel (*r de F al conjunto

$$S_r = \{x \in U : F(x) = r\}.$$

Cuando $n = 2$ se dice que S_r es una *curva de nivel*, y si $n = 3$ a S_r se la llama superficie de nivel. Para $n \geq 4$ se habla de hipersuperficies de nivel.

Ejemplo 9.2. Representar la gráfica y las curvas o superficies de nivel de las siguientes funciones.

1. $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = x^2 + y^2$.
2. $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = xy$
3. $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x, y) = y - \sin x$
4. $F : \mathbb{R}^3 \to \mathbb{R}$, $F(x, y, z) = x^2 + y^2 - z^2$.

La relación entre curvas o superficies de nivel de F y su gradiente queda patente en el siguiente resultado.

Teorema 9.11. *Sean U un abierto de $\mathbb{R}^n$, $F : U \to \mathbb{R}$ diferenciable en U, $r \in \mathbb{R}$, $p \in S_r$. Sea $\gamma : I = (\alpha, \beta) \subseteq \mathbb{R} \to \mathbb{R}^n$ una curva diferenciable, y $t_0 \in I$ tales que $\gamma(t_0) = p$ y $\gamma(I) \subset S_r$. Entonces*

$$\gamma'(t_0) \perp \nabla F(p).$$

Es decir, el vector gradiente $\nabla F(p)$ es perpendicular a cualquier vector tangente a la superficie S_r, por lo que se dice que $\nabla F(p)$ es perpendicular a S_r en p.

Demostración. Aplicando la regla de la cadena a la composición $f = F \circ \gamma$ obtenemos que

$$f'(t) = \langle \nabla F(\gamma(t)), \gamma'(t) \rangle.$$

Por otro lado, al estar $\gamma(t)$ en S_r para todo t, se tiene que $f(t) = r$ para todo t, es decir, h es constante, y por tanto $f'(t) = 0$. Obtenemos así que

$$\langle \nabla F(\gamma(t)), \gamma'(t) \rangle = 0$$

para todo t. Poniendo $t = t_0$ se concluye que $\gamma'(t_0)$ y $\nabla F(p)$ son perpendiculares. □

Esto justifica la siguiente definición.

Definición 9.3. Si $F : U \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función diferenciable, se define el hiperplano (afín) tangente $\mathcal{T}S_{rp}$ a la (hiper)superficie de nivel S_r en un punto $p \in S_r$ tal que $\nabla F(p) \neq 0$ como el hiperplano de ecuación

$$\frac{\partial F}{\partial x_1}(p)(x_1 - p_1) + ... + \frac{\partial F}{\partial x_n}(p)(x_n - p_n) = 0,$$

Del vector gradiente $\nabla F(p)$ se dice que es un vector *normal* a la superficie S_r.

De momento no diremos nada de lo que puede pasar en el caso $\nabla F(p) = 0$, salvo que S_r podría muy bien no ser lo que uno entiende de manera intuitiva como una (hiper)superficie. Por ejemplo, si $F : \mathbb{R}^3 \to \mathbb{R}$, $F(x, y, z) = 1$, entonces S_1 es todo $\mathbb{R}^3$, y no tiene sentido hablar de planos tangentes a S_1 en ninguno de sus puntos. Para comprender mejor este tipo de fenómenos habrá que esperar a los capítulos 14 y 15.

Conviene observar que la definición anterior de plano tangente coincide con la dada en el capítulo anterior en el caso de que la superficie de nivel sea la gráfica de una función $f : W \subseteq \mathbb{R}^n \to \mathbb{R}$. En efecto, para una función tal f, definamos $F : U = W \times \mathbb{R} \subseteq \mathbb{R}^{n+1} \to \mathbb{R}$ por

$$F(x, z) = f(x) - z.$$

La función F es diferenciable en $U = W \times \mathbb{R}$ si y sólo si f lo es en W, y además

$$G_f = S_0 = \{(x, z) \in U \times \mathbb{R} : F(x, z) = 0\}.$$

Nótese también que, si $p = (a, f(a))$,

$$\nabla F(p) = \left(\frac{\partial f}{\partial x_1}(a), ..., \frac{\partial f}{\partial x_n}(a), -1 \right).$$

Según la definición anterior, el hiperplano tangente a $S_0 = G_f$ en p es el hiperplano de ecuación

$$\frac{\partial F}{\partial x_1}(p)(x_1 - p_1) + ... + \frac{\partial F}{\partial x_n}(p)(x_n - p_n) + \frac{\partial F}{\partial z}(p)(z - f(a)),$$

es decir

$$\frac{\partial f}{\partial x_1}(a)(x_1 - p_1) + ... + \frac{\partial f}{\partial x_n}(a)(x_n - p_n) - (z - f(a)) = 0,$$

la misma que en la definición dada en el capítulo 8.

En el caso particular en que $F : U \subseteq \mathbb{R}^2 \to \mathbb{R}$, si F es diferenciable en p con $\nabla F(p) \neq 0$, la ecuación de la recta tangente en $p = (x_0, y_0)$ a la curva de nivel $C_r = \{(x, y) \in U : F(x, y) = r\}$ viene dada por

$$\frac{\partial F}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial F}{\partial x}(x_0, y_0)(y - y_0) = 0,$$

y $\nabla F(p)$ es el vector normal a la curva C_r en p.

Si consideramos la gráfica de F, es decir la superficie $z = F(x, y)$ en $\mathbb{R}^3$, que podemos imaginar como un paisaje con montañas y valles, entonces las curvas de nivel de F pueden interpretarse como líneas de contorno de un *mapa* del paisaje proyectado en el plano $\mathbb{R}^2$; estas líneas son las que unen los puntos de igual altitud en el paisaje real, es decir, C_r es la línea que une todos los puntos del paisaje que están a altitud r, y $F(x, y)$ representa la altitud del punto (x, y) del mapa en el paisaje real. Imaginemos ahora que deseamos caminar sobre el paisaje en una dirección que nos permita elevarnos con la máxima rapidez, es decir, aumentar el valor de $F(x, y)$ lo más rápidamente posible. Es intuitivamente obvio que lo que debemos hacer entonces es movernos en una dirección perpendicular a las curvas de nivel, y en el sentido en que la altitud aumenta. En otras palabras, si deseamos, en un punto (x_0, y_0), elegir una dirección en la que $F(x, y)$ aumente con la máxima rapidez, la dirección elegida debería ser la del vector $\nabla F(x_0, y_0)$ (que por la discusión anterior ya sabemos que es perpendicular a las curvas de nivel). Todo esto sugiere que el siguiente resultado debería ser verdad: la dirección v en la que la derivada direccional $D_v F(x_0, y_0)$ (que mide la velocidad con la que F aumenta a lo largo de la recta de dirección v y que pasa por el punto (x_0, y_0)) es máxima es la dirección de $\nabla F(x_0)$. A continuación vemos que, en efecto, esto es así.

Teorema 9.12. *Sea $F : U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable en $a \in U$ tal que $\nabla F(a) \neq 0$. Entonces,*

$$\text{máx}\{D_v F(a) : \|v\| = 1\} = D_{v_0} F(a),$$

donde

$$v_0 = \frac{\nabla F(a)}{\|\nabla F(a)\|},$$

y la norma considerada es la euclidea.

Demostración. Usando la desigualdad de Cauchy-Schwarz, tenemos que

$$D_v F(a) = \langle \nabla F(a), v\rangle \leq \|\nabla F(a)\| \|v\|,$$

y sabemos que la igualdad se da si y sólo si v es un múltiplo escalar de $\nabla F(a)$. Por tanto, para todo $v \in \mathbb{R}^n$ con $\|v\| = 1$, tenemos

$$D_v F(a) = \langle \nabla F(a), v\rangle \leq \|\nabla F(a)\| \|v\| = \|\nabla F(a)\| = \langle \nabla F(a), v_0\rangle = D_{v_0} F(a).$$

Análogamente, la dirección de máximo decrecimiento de F vendrá dada por $w_0 = -\nabla F(a)/\|\nabla F(a)\|$. □

9.1. Problemas

Problema 9.1. Dar un ejemplo de función $f : \mathbb{R}^2 \to \mathbb{R}$ que es diferenciable en el origen pero discontinua en todos los demás puntos.

Problema 9.2. Sean $h : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

$$x = h_1(\rho, \theta) = \rho \cos\theta, \; y = h_2(\rho, \theta) = \rho \sin\theta,$$

es decir el cambio a coordenadas polares. Demostrar que f es diferenciable en $\mathbb{R}^2$, y calcular su matriz jacobiana. Sea ahora $f : \mathbb{R}^2 \to \mathbb{R}$ una función diferenciable cualquiera. Aplicar la regla de la cadena para calcular la matriz jacobiana de $f \circ h$. Con la notación clásica, comprobar que

$$\frac{\partial f}{\partial \rho} = \cos\theta \frac{\partial f}{\partial x} + \sin\theta \frac{\partial f}{\partial y}, \quad \text{y} \quad \frac{\partial f}{\partial \theta} = -\rho \sin\theta \frac{\partial f}{\partial x} + \rho \cos\theta \frac{\partial f}{\partial y}.$$

Problema 9.3. Sea $f(x, y, z) = \frac{x^2 - z^2}{x^2 + y^2 + z^2} \sin(x - y)$ si $(x, y, z) \in \mathbb{R}^3 \setminus \{0\}$. ¿Qué valor habría que dar a $f(0, 0, 0)$, para que f fuese continua en todo $\mathbb{R}^3$? ¿Es en este caso f diferenciable?, ¿y de clase C^1?

Problema 9.4. Calcúlese la diferencial en el punto $(0, 0, 0)$ de la función:

$$f(x, y, z) = \int_z^{xy^2} (t^2 + 1)e^{-t} dt$$

Problema 9.5. Sea $\gamma : \mathbb{R} \longrightarrow \mathbb{R}^n$ una curva diferenciable con $\gamma'(t) \neq 0$ para todo $t \in \mathbb{R}$. Sea $p \in \mathbb{R}^n$ un punto que no pertenezca a $\gamma(\mathbb{R})$. Supongamos que existe $q = \gamma(t_0)$ el punto de la curva más cercano a p. Demuéstrese que el vector $p - q$ es ortogonal a la curva en q.

Problema 9.6. (a) Sean $f, g : \mathbb{R} \longrightarrow \mathbb{R}^n$ dos curvas diferenciables con $f'(t) \neq 0$, $g'(t) \neq 0$, para todo $t \in \mathbb{R}$. Supongamos que existen $p = f(s_0)$ y $q = g(t_0)$ tales que

$$\|p - q\| \leq \|f(s) - g(t)\|$$

para todos $t, s \in \mathbb{R}$. Pruébese que $p - q$ es un vector ortogonal a las curvas f y g en p y q respectivamente.

(b) Aplíquese el resultado de (a) para encontrar la distancia entre las rectas $f(s) = (s, 2s, -s)$ y $g(t) = (t - 1, t - 2, 2t + 3)$.

Problema 9.7. Estúdiese la diferenciabilidad en el origen de las funciones del problema 5.1.

Problema 9.8. Sean n un número natural y $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ la función definida por:

$$f(x,y) = \begin{cases} (x+y)^n \cos \frac{1}{(x^2+y^2)^{1/2}} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

¿Cómo hay que elegir n para que: a) f sea continua?; b) f sea diferenciable?; c) f tenga derivadas parciales continuas?

Problema 9.9. Sean $g_v(t) = (v_1 t, v_2 t)$, con $v = (v_1, v_2) \neq (0,0)$, y

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

1. Pruébese que f admite derivadas parciales en $(0,0)$ y calcúlense.
2. Pruébese que $f \circ g_v$ es diferenciable.
3. Compruébese que si $v_1 \neq 0 \neq v_2$ entonces $(f \circ g_v)'(0) \neq \langle \nabla f(0,0), g_v'(0) \rangle$. ¿Por qué? ¿Contradice esto la regla de la cadena?

Problema 9.10. Usando la regla de la cadena, dar una demostración más simple que la proporcionada en el capítulo anterior del hecho que el plano tangente a la gráfica G_f de una función diferenciable $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ en un punto p es la unión de todas las rectas tangentes en p a curvas diferenciables contenidas en G_g y que pasan por p.

Problema 9.11. Sean

$$f(x,y) = \begin{cases} xy \sin \frac{x}{y} & \text{si } y \neq 0 \\ 0 & \text{si } y = 0 \end{cases} \qquad \text{y} \qquad g(x,y) = \frac{1}{\pi} e^{x+y} + \int_0^x \frac{t^2}{(1+t^4)^{1/2}}\, dt$$

1. Pruébese que $F(x,y) = (f(x,y), g(x,y))$ es diferenciable en $(0,0)$ y $(0,r)$.
2. Dedúzcase que $G = F \circ F$ es diferenciable en $(0,0)$ y calcúlese $DG(0,0)$.

Problema 9.12. Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, definida por

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

1. Demuéstrese que f no es diferenciable en (0,0).

2. Demuéstrese que, si $g : [0,1] \longrightarrow \mathbb{R}^2$ es diferenciable, entonces $h = f \circ g$ es diferenciable.

Problema 9.13 (Teorema de Euler)**.** Sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ una función diferenciable y homogenea de grado m, es decir $f(tx) = t^m f(x)$ para todo $t \in \mathbb{R}$ y para todo $x \in \mathbb{R}^n$. Pruébese que $mf(x) = Df(x)(x)$ para todo $x \in \mathbb{R}^n$.

Problema 9.14. En los siguientes casos, estúdiese si la función es diferenciable y si es de clase C^1:

$$f(x,y) = \begin{cases} \frac{x\sin^2 y}{x^2+2y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$
$$g(x,y) = \begin{cases} \frac{\log(1+x^2y^2)}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

Problema 9.15. Descríbanse las curvas de nivel de la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por $f(x,y) = \frac{2xy}{x^2+y^2}$ si $(x,y) \neq (0,0)$, y $f(0,0) = 0$.

Problema 9.16. Calcúlese $\frac{\partial h}{\partial x}$, siendo $h(x,y) = f(u(x,y), v(x,y))$, y

$$f(u,v) = \frac{u^2+v^2}{u^2-v^2}, \quad u(x,y) = e^{-x-y}, \quad v(x,y) = e^{xy}.$$

Problema 9.17. Si $F(x,y) = 2e^{-x^2} + e^{-3y^2}$ representa la altura de una montaña en el punto (x,y) de un mapa, ¿en qué dirección desde $(0,1)$ deberá comenzar uno a caminar para escalar más rápidamente?

Problema 9.18. Sea $f(x,y,z) = (e^{xz}+x-1, z+ae^y)$, y $g(u,v) = v+u\sin\frac{1}{u^2+v^2}$, $g(0,0) = 0$.

1. Estúdiese la diferenciabilidad de f y g.

2. ¿Para qué valores de a puede asegurarse, utilizando la regla de la cadena, que $g \circ f$ es diferenciable en $(0,0,0)$?

3. Calcúlese $D(g \circ f)(0,0,0)$, y hállese el valor de a que hace máxima la derivada direccional en la dirección $(0,1,0)$

Problema 9.19. En este problema D denotará el conjunto $\{(x,y) \in \mathbb{R}^2 : x^2+y^2 < 1\}$.

1. Sea $\varphi_1 : D \longrightarrow \mathbb{R}^3$ la función definida por $\varphi_1(x,y) = (x,y,1-\sqrt{x^2+y^2})$. Pruébese que no es diferenciable en $(0,0)$. Hágase un dibujo del conjunto $\varphi_1(D)$, para lo cual hay que ver que relación existe entre x, y, z cuando $(x,y,z) \in \varphi_1(D)$.

2. Lo mismo para las funciones $\varphi_2, \varphi_3 : D \longrightarrow \mathbb{R}^3$ definidas por
$$\varphi_2(x,y) = (x,y,1-(x^2+y^2)); \quad \varphi_3(x,y) = (x,y,\sqrt{1-x^2-y^2})$$

3. Estúdiese la imagen de las siguientes funciones: $F_1, F_2, F_3 : (0,2\pi) \times (0,1) \longrightarrow \mathbb{R}^3$. definidas por: $F_1(\theta,\rho) = (\rho\cos\theta, \rho\sin\theta, 1-\rho)$, $F_2(\theta,\rho) = (\rho\cos\theta, \rho\sin\theta, 1-\rho^2)$, $F_3(\theta,\rho) = (\rho\cos\theta, \rho\sin\theta, \sqrt{1-\rho^2})$. Compárese el resultado con los apartados (a) y (b).

4. Dibújense los conjuntos: $G_1(D)$ y $G_2(D)$, siendo $G_1(x,z) = (x, 1-\sqrt{x^2+z^2}, z)$ y $G_2(y,z) = (1-(z^2+y^2), y, z)$.

Problema 9.20. Consideremos la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por $f(x,y) = 1 - \sqrt{x^2+y^2}$. Dibújese su gráfica. Encuéntrense funciones tales que los conjuntos $\varphi_2(D)$ y $\varphi_3(D)$ del ejercicio anterior sean su gráfica. ¿Se pueden expresar como gráfica de una función los conjuntos $G_1(D)$ y $G_2(D)$?

Problema 9.21. Consideremos la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por $f(x,y) = 2y^2 - x(x-1)^2$.

1. Dibújense sus curvas de nivel, es decir los conjuntos $f^{-1}(\alpha)$, $\alpha \in \mathbb{R}$. (Indicación: Dibújense primero las curvas correspondientes a los niveles $\alpha = 0$ y $\alpha = 1/3$ y a continuación represéntense las correspondientes a $\alpha > 0$ y $\alpha < 0$).

2. Calcúlese el gradiente de f, $\nabla f(x,y)$. ¿En qué puntos es 0? Calcúlense $\nabla f(0,0)$, $\nabla f(\frac{4}{3},0)$, $\nabla f(2,1)$, $\nabla f(2,-1)$ y $\nabla f(\frac{1}{3}, \sqrt{\frac{2}{27}})$. Obsérvese que estos vectores son perpendiculares a las correspondientes curvas de nivel.

3. Hágase un esbozo de la gráfica de f.

Problema 9.22. Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función convexa. Supongamos que existe $x_0 \in \mathbb{R}^n$ tal que $Df(x_0) = 0$. Probar que entonces f tiene un mínimo absoluto en x_0, es decir, $f(x) \geq f(x_0)$ para todo $x \in \mathbb{R}^n$. *Indicación:* reducir el problema al caso $n = 1$.

Capítulo 10

Teorema del valor medio

Uno de los teoremas más importantes del cálculo para funciones $f : \mathbb{R} \to \mathbb{R}$ es el teorema del valor medio (que ya se ha empleado en la demostración del Teorema 9.7). Recordemos que el teorema del valor medio afirma que, si $f : [a, b] \subset \mathbb{R}$ es continua en $[a, b]$ y derivable en (a, b) entonces existe $c \in (a, b)$ tal que $f(b) - f(a) = f'(c)(b - a)$. Una de las razones de su importancia radica en que permite deducir propiedades sobre el comportamiento de la función f a partir de propiedades de su derivada. Por ejemplo, que si la derivada f' es positiva en un intervalo entonces la función f es creciente en ese intervalo.

El teorema del valor medio sigue siendo cierto para funciones de varias variables que toman valores reales, y de hecho es fácil deducirlo, usando la regla de la cadena, del teorema del valor medio para funciones de $\mathbb{R}$ en $\mathbb{R}$.

Teorema 10.1 (del valor medio)**.** *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable en un segmento $[a, b] \subset U$. Entonces existe $c \in [a, b]$ tal que*

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

Demostración. Definamos la función $\varphi : [0, 1] \to U$ por

$$\varphi(t) = (1 - t)a + tb.$$

Nótese que φ es continua, y diferenciable en $(0, 1)$, con $\varphi'(t) = b - a$. Consideremos ahora la composición $g := f \circ \varphi : [0, 1] \to \mathbb{R}$. Es claro que g es continua en $[0, 1]$ y, por la regla de la cadena, diferenciable en $(0, 1)$, con

$$g'(t) = Df(\varphi(t))(\varphi'(t)) = Df(\varphi(t))(b - a) = \langle \nabla f(\varphi(t)), b - a\rangle.$$

Podemos aplicar entonces el teorema del valor medio, ya conocido para funciones de $\mathbb{R}$ en $\mathbb{R}$, para encontrar $\theta \in (0, 1)$ tal que

$$f(b) - f(a) = g(1) - g(0) = g'(\theta) = \langle \nabla f(\varphi(\theta)), b - a\rangle,$$

de donde, poniendo $c := \varphi(\theta) \in [a, b]$, se obtiene el resultado buscado. □

Sin embargo, cuando la función f toma valores vectoriales, el teorema del valor medio, al menos en esta formulación, no es válido. Por ejemplo, consideremos $f : \mathbb{R} \to \mathbb{R}^3$ definida por

$$f(t) = (\cos t, \sin t, t),$$

una curva que describe una hélice en el espacio. Esta curva es derivable (y por ello también continua) en todo $\mathbb{R}$. Sin embargo, no existe ningún $\theta \in [0, 2\pi]$ tal que $f(2\pi) - f(0) = f'(\theta)(2\pi - 0)$. En efecto, $f(2\pi) - f(0) = (0, 0, 2\pi)$,

$$f'(t) = (-\sin t, \cos t, 1)$$

para todo t, luego $f'(t) \neq (0, 0, 2\pi)$ para todo t. Otro ejemplo de función $f : \mathbb{R} \to \mathbb{R}^2$ que no satisface el teorema del valor medio lo proporciona el ejercicio 10.1.

No obstante, sí puede demostrarse una *desigualdad del valor medio* para funciones con valores vectoriales.

Teorema 10.2 (Desigualdad del valor medio)**.** *Sea U un conjunto abierto y convexo de $\mathbb{R}^n$. Sea $f : U \to \mathbb{R}^m$ diferenciable en U. Entonces, para cada $x, y \in U$ existe $c \in [x, y]$ tal que*

$$\|f(x) - f(y)\| \leq \|Df(c)(y - x)\|,$$

donde la norma considerada es la euclidea.

Demostración. Pongamos $h = y - x$, $u = f(y) - f(x)$. Como U es convexo, el segmento $[x, y]$ está dentro de U. Definamos $g : [0, 1] \to \mathbb{R}$ por

$$g(t) = \langle u, f(x + th) \rangle.$$

Es claro que g es diferenciable, con

$$g'(t) = \langle u, Df(x + th)(h) \rangle.$$

Por el teorema del valor medio, existe $\theta \in [0, 1]$ tal que

$$g(1) - g(0) = g'(\theta),$$

es decir,

$$\langle u, f(y) - f(x) \rangle = \langle u, Df(x + \theta h)(h) \rangle.$$

Por tanto, usando la desigualdad de Cauchy-Schwarz,

$$\begin{aligned} &\|f(y) - f(x)\|^2 = \langle u, f(y) - f(x) \rangle = \\ &\langle u, Df(x + \theta h)(h) \rangle \leq \|u\| \|Df(x + \theta h)(h)\| = \\ &\|f(y) - f(x)\| \|Df(x + \theta h)(h)\|, \end{aligned}$$

de donde se obtiene el resultado cancelando $\|f(y) - f(x)\|$ y poniendo $c = x + \theta h$. □

Una consecuencia importante de la desigualdad del valor medio es que toda función con derivada acotada en un conjunto abierto y convexo es Lipschitz (y en particular uniformemente continua) en dicho conjunto.

En lo que resta de capítulo se supondrá que todas las normas en $\mathbb{R}^n$ y $\mathbb{R}^m$ son las euclideas.

Corolario 10.3. *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función diferenciable sobre un abierto convexo, y supongamos que $\|Df(x)\| \leq M$ para todo $x \in U$. Entonces*

$$\|f(x) - f(y)\| \leq M\|x - y\|$$

para todos $x, y \in U$, es decir, f es M-Lipschitz en U.

Demostración. Basta aplicar el teorema anterior, teniendo en cuenta que

$$\|Df(c)(y-x)\| \leq \|Df(c)\|\|y-x\| \leq M\|y-x\|,$$

por la definición de norma de la aplicación lineal $Df(c)$. $\square$

Observación 10.4. *En el corolario anterior es importante que el conjunto sea convexo; el resultado no es cierto para conjuntos abiertos conexos. Ver los problemas 10.8 y 10.9.*

Corolario 10.5. *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función diferenciable sobre un abierto convexo, y supongamos que*

$$|\frac{\partial f_j}{\partial x_i}(x)| \leq M$$

para todo $x \in U$. Entonces

$$\|f(x) - f(y)\| \leq \sqrt{nm}M\|x-y\|$$

para todos $x, y \in U$.

Demostración. Basta tener en cuenta que si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una aplicación lineal cuya matriz (a_{ij}) respecto de las bases canónicas satisface que $|a_{ij}| \leq M$ para todo i, j, entonces $\|T\| \leq \sqrt{nm}M$. Ver el ejercicio 5.58. $\square$

Otra consecuencia típica del teorema del valor medio es que las funciones con derivada nula son constantes, cuando su dominio es un conjunto conexo y abierto.

Corolario 10.6. *Sean $U \subseteq \mathbb{R}^n$ un abierto conexo, y $f : U \to \mathbb{R}^m$ una función diferenciable tal que $Df(x) = 0$ para todo $x \in U$. Entonces f es constante en U.*

Demostración. Cuando U es *convexo*, el resultado es consecuencia inmediata del corolario anterior, tomando $M = 0$. Si U no es convexo, pero sí conexo, podemos razonar como sigue. Fijemos $a \in U$, y definamos

$$\Omega = \{x \in U : f(x) = f(a)\}.$$

El conjunto Ω es abierto, ya que si $x \in \Omega$, es decir, $f(x) = f(a)$, como U es abierto existe $r > 0$ tal que $B(x,r) \subset U$, y como la bola $B(x,r)$ es convexa, por lo anterior se deduce que f es constante en esta bola, es decir, $f(y) = f(a)$ para todo $y \in B(x,r)$, lo que significa que $B(x,r) \subset \Omega$.

Por otra parte Ω es también un cerrado relativo a U ya que $\Omega = f^{-1}(\{f(a)\})$ y f es continua en U por ser diferenciable. Además Ω es obviamente no vacío. Como U es conexo, se deduce que $\Omega = U$, es decir f es constante en U. $\square$

Observación 10.7. En el resultado anterior es esencial que el conjunto donde la derivada se anula sea abierto. Existen ejemplos de funciones $f : \mathbb{R}^2 \to \mathbb{R}$ diferenciables de clase C^1 y caminos continuos $\gamma : [0,1] \to \mathbb{R}^2$ tales que $\nabla f(\gamma(t)) = 0$ para todo $t \in [0,1]$ y sin embargo $f(\gamma(0)) = 0 < 1 = f(\gamma(1))$, es decir f no es constante sobre $\gamma[0,1]$. El primero de estos ejemplos fue encontrado por H. Whitney [15] en 1935.

10.1. Problemas

Problema 10.1. Sea $f : \mathbb{R} \to \mathbb{R}^2$ definida por $f(t) = (t^2, t^3)$. Comprobar que no existe ningún $c \in [0, 1]$ tal que $f(1) - f(0) = Df(c)(1 - 0)$.

Problema 10.2. Probar que las únicas funciones diferenciables de $\mathbb{R}^n$ en $\mathbb{R}^m$ cuya derivada es constante son las funciones afines. Es decir, si $f : \mathbb{R}^n \to \mathbb{R}^m$ satisface que

$$Df(x) = L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$$

para todo $x \in \mathbb{R}^n$, entonces

$$f(x) = L(x) + c$$

para todo $x \in \mathbb{R}^n$, donde $c \in \mathbb{R}^m$ es un vector constante.

Problema 10.3. Sea $f : \mathbb{R}^n \to \mathbb{R}$ de clase C^1, y sean $a, b \in \mathbb{R}^n$ tales que $f(a) = f(b)$. ¿Existe $c \in [a, b]$ tal que $Df(c) = 0$?

Problema 10.4. Sea $f : [0, \infty) \to \mathbb{R}$ continua, derivable en $(0, \infty)$, y tal que $f(0) = 0$ y f' es creciente. Probar que $g(x) = f(x)/x$ es creciente en $(0, \infty)$.

Problema 10.5. Sea $f : \mathbb{R} \to \mathbb{R}$ derivable. Supongamos que para todo $x \in \mathbb{R}$ se tiene $0 \leq f'(x) \leq f(x)$. Demostrar que si f se anula en algún punto, entonces $f = 0$ en todo $\mathbb{R}$. *Indicación:* Probar que $g(x) = e^{-x} f(x)$ es decreciente.

Problema 10.6. Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ una función diferenciable, donde U es un abierto convexo. Supongamos que la matriz jacobiana de f es definida positiva en U, es decir,

$$\sum_{i,j=1}^{n} \frac{\partial f_j}{\partial x_i}(x) h_i h_j > 0$$

para todos $x \in U$, $h \in \mathbb{R}^n$. Demostrar que f es inyectiva. *Indicación:* Si $a, b \in U$ y $a \neq b$, poner $h = b - a$ y aplicar el teorema del valor medio a $g(x) = \langle f(x) - f(a), h \rangle$.

Problema 10.7. Sea U un subconjunto convexo y abierto en $\mathbb{R}^2$, y sea $f : U \to \mathbb{R}$ diferenciable en U. Supongamos que

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

para todo $(x, y) \in C$. Probar que f no depende de y, es decir,

$$f(x, y_1) = f(x, y_2)$$

para todos $(x, y_1), (x, y_2) \in U$. Generalizar el resultado para funciones definidas en abiertos convexos de $\mathbb{R}^n$ para $n \geq 3$. ¿Es cierto esto si se supone solamente que U es conexo? *Indicación:* Para contestar a la última pregunta, imagínese un dominio en forma de media corona circular.

Problema 10.8. Dar un ejemplo de función diferenciable y no constante, definida en un abierto no conexo, cuya derivada sea constantemente cero.

Problema 10.9. El objeto de este problema es demostrar que la desigualdad de valor medio falla en dominios no convexos, incluso aunque éstos sean conexos por caminos. Más precisamente, vamos a construir un conjunto U abierto, conexo y acotado de $\mathbb{R}^2$, y una función $f : U \to \mathbb{R}$ de clase C^1 tal que $\|\nabla f(x, y)\| \leq 1$ para todo $(x, y) \in U$ y sin embargo f no es uniformemente continua (por tanto tampoco Lipschitz) en U. Con tal fin se recomienda seguir estas instrucciones:

1. Para cada $n \in \mathbb{N}$, pongamos

$$V_n = (\frac{1}{n} - \frac{1}{2^{n+1}}, \frac{1}{n} + \frac{1}{2^{n+1}}) \times (-2, 2);$$

para $n \in \mathbb{N}$ impar sea

$$H_n = (\frac{1}{n+1} - \frac{1}{2^{n+2}}, \frac{1}{n} + \frac{1}{2^{n+1}}) \times (2 - \frac{1}{4}, 2),$$

y para $n \in \mathbb{N}$ par definamos

$$H_n = (\frac{1}{n+1} - \frac{1}{2^{n+2}}, \frac{1}{n} + \frac{1}{2^{n+1}}) \times (-2, -2 + \frac{1}{4}).$$

Sea entonces

$$U = \bigcup_{n=1}^{\infty} V_n \cup H_n.$$

Comprobar que $V_n \cap V_m = \emptyset$ si $n \neq m$. Hágase un dibujo del conjunto U. Probar que U es abierto, conexo por caminos y acotado.

2. Construir una función $\varphi : \mathbb{R} \to \mathbb{R}$ de clase C^1 tal que $\varphi(t) = 0$ si $t \leq -1$, $\varphi(t) = 1$ si $t \geq 1$, y $|\varphi'(t)| \leq 1$ para todo t.

3. Definir ahora $f : U \to \mathbb{R}$ como sigue. Poner

$$f(x, y) = \varphi(y) \ \text{ si } \ (x, y) \in V_1.$$

Nótese que f es constante en $V_1 \cap \{y > 1\}$, donde vale 1, y también en $V_1 \cap \{y < -1\}$, donde vale 0. Sea

$$f(x, y) = 1 \ \text{ si } \ (x, y) \in H_1.$$

En V_2 sea

$$f(x, y) = \varphi(-y) \ \text{ si } \ (x, y) \in V_2.$$

Obsérvese que ahora f vale 1 en $V_2 \cap \{y > 1\}$, y es constantemente 2 en $V_2 \cap \{y < -1\}$. En general, defínase

$$f(x, y) = n \ \text{ si } \ (x, y) \in H_n,$$

$$f(x, y) = \varphi(y) + n - 1 \ \text{ si } \ (x, y) \in V_n, \ \text{ con } n \text{ impar},$$

y

$$f(x, y) = \varphi(-y) + n - 1 \ \text{ si } \ (x, y) \in V_n, \ \text{ con } n \text{ par}.$$

Demostrar que f es diferenciable de clase C^1 en U, y que

$$\|\nabla f(x, y)\| \leq 1$$

para todo $(x, y) \in U$.

4. Probar que, sin embargo, f no es uniformemente continua en U (para verlo, notar que si f fuera uniformemente continua tendría una única extensión uniformemente continua a $\overline{U}$, que es compacto, y por tanto debería estar acotada).

5. Poner más ejemplos de recintos abiertos conexos y acotados, y de funciones diferenciables que no satisfagan la desigualdad del valor medio (pensar, por ejemplo, en recintos con forma de serpiente o de espiral infinita, sobre los cuales puedan definirse funciones que van creciendo siempre con rapidez controlada a lo largo del recinto y que sin embargo no están acotadas, al tener la espiral longitud infinita).

Problema 10.10. En el ejemplo de Whitney mencionado al final de este capítulo el camino continuo γ no puede ser diferenciable. ¿Por qué?

Capítulo 11

Derivadas de orden superior. Teorema de Schwartz.

A lo largo de este capítulo estudiaremos las funciones que son diferenciables varias veces, su relación con las funciones que tienen derivadas parciales de orden superior que son continuas, llamadas de clase C^p, y las propiedades de éstas, como el teorema de Schwartz sobre la simetría de las derivadas de orden superior, y el hecho de que la suma, producto y composición de funciones de clase C^p son de clase C^p.

Recordemos que una función $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ se dice que es de clase C^1 si todas sus derivadas parciales de primer orden (es decir, $\partial f_j/\partial x_i$, $1 \leq i \leq n, 1 \leq j \leq m$) son continuas. Lo primero que debemos observar es que esto es equivalente a decir que f es diferenciable en U y además $Df : U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ es continua.

Proposición 11.1. *Sea U un abierto de $\mathbb{R}^n$, y sea $f : U \to \mathbb{R}^m$. Las siguientes afirmaciones son equivalentes.*

1. *Todas las derivadas parciales de primer orden f existen y son continuas en U.*

2. *f es diferenciable en U, y la aplicación*

$$U \ni x \to Df(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$$

es continua.

Demostración. $(1) \implies (2)$: Ya sabemos que f es diferenciable en U gracias al Teorema 9.7. Para ver la continuidad de Df, recuérdese que $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ es un espacio de dimensión nm, naturalmente isomorfo al espacio de las matrices $\mathcal{M}_{n\times m}$, que a su vez identificamos con $\mathbb{R}^{nm}$. Como todas las normas son equivalentes en $\mathbb{R}^{nm}$, todos estos isomorfismos dan lugar a homeomorfismos. Podemos por tanto identificar cualquier elemento A de $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ con su matriz (a_{ij}) respecto de las bases canónicas, y ésta con el vector $(a_{11}, ..., a_{mn}) \in \mathbb{R}^{nm}$, y considerar la norma en $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ definida por

$$\|A\| = \sup\{|a_{11}|, ..., |a_{mn}|\}.$$

Haciendo estas identificaciones, resultará que Df es continua si y sólo si la aplicación

$$U \ni x \to (\frac{\partial f_1}{\partial x_1}(x), ..., \frac{\partial f_m}{\partial x_n}(x)) \in \mathbb{R}^{nm} = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$$

es continua. Pero esto equivale a decir que las funciones componentes de esta aplicación, es decir, las funciones

$$U \ni x \to \frac{\partial f_j}{\partial x_i}(x)$$

son continuas en U para todos $1 \le i \le n, 1 \le j \le m$.
$(2) \implies (1)$: Como Df es continua si y sólo si todas sus funciones componentes son continuas, es claro, por la discusión anterior, que f tiene derivadas parciales de primer orden continuas en U. $\square$

Recordemos que una función diferenciable no tiene por qué ser de clase C^1.

Ejemplo 11.1. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

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

Comprobar que f es diferenciable en todo $\mathbb{R}^2$, pero f no es de clase C^1 en ningún entorno de $(0,0)$.

Las funciones de clase C^1 tiene una propiedad importante que puede parecer un poco técnica pero que será de gran utilidad más adelante. Para hacerse una idea de cómo puede usarse, véase el problema 11.1.

Proposición 11.2. *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función de clase $C^1(U, \mathbb{R}^m)$. Entonces, para cada $a \in U$, $\varepsilon > 0$, existe $\delta > 0$ tal que*

$$\|f(x) - f(y) - Df(a)(x-y)\| \le \varepsilon \|x - y\|$$

para todos $x, y \in B(a, \delta)$. También, en particular,

$$\|f(x) - f(y)\| \le (\|Df(a)\| + \varepsilon) \|x - y\|$$

si $x, y \in B(a, \delta)$.

Demostración. Definamos $g : U \to \mathbb{R}^m$, $g(x) = f(x) - L(x)$, donde $L = Df(a)$. Es obvio que $g \in C^1(U, \mathbb{R}^m)$. Como $Dg(a) = 0$ y, por la proposición anterior, sabemos que Dg es continua, existe $\delta > 0$ tal que $\|Dg(x)\| \le \varepsilon$ si $x \in B(a, \delta)$. Entonces, por la desigualdad del valor medio 10.2 se tiene que, para $x, y \in B(a, \delta)$

$$\|g(x) - g(y)\| \le \varepsilon \|x - y\|,$$

lo cual significa precisamente que

$$\|f(x) - f(y) - Df(a)(x-y)\| \le \varepsilon \|x - y\|$$

para todos $x, y \in B(a, \delta)$. La otra desigualdad se deduce inmediatamente de ésta usando la propiedad triangular y la definición de norma de una aplicación lineal. $\square$

Pasamos ahora a considerar derivadas parciales de segundo orden.

Definición 11.1. Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Supongamos que la derivada parcial $\partial f/\partial x_i$ existe en U. Si la función

$$\frac{\partial f}{\partial x_i} : U \to \mathbb{R}^m$$

admite derivada parcial j-ésima en $x \in U$, se dice que f tiene derivada parcial segunda en x y se denota

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

Observación 11.3. Aplicando las propiedades que ya conocemos de las funciones diferenciables, es inmediato ver que la derivada parcial de $f = (f_1, ..., f_m)$

$$\frac{\partial^2 f}{\partial x_j \partial x_i}$$

existe si y sólo si existen las derivadas parciales

$$\frac{\partial^2 f_k}{\partial x_j \partial x_i}$$

para todo $k = 1, ..., m$, y en este caso se tiene la igualdad

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \left(\frac{\partial^2 f_1}{\partial x_j \partial x_i}, ..., \frac{\partial^2 f_m}{\partial x_j \partial x_i}\right).$$

Ejemplo 11.2. Si $f : \mathbb{R}^3 \to \mathbb{R}$ está definida por

$$f(x, y, z) = x^2 yz + 9,$$

calcular todas las derivadas parciales de segundo orden de f. Comprobar que

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

La propiedad de simetría de las derivadas parciales cruzadas presente en el ejemplo anterior no se debe a una casualidad; es una propiedad fundamental de las funciones con derivadas parciales de segundo orden que son continuas.

Teorema 11.4 (Schwartz)**.** *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ tal que las derivadas parciales $\partial^2 f/\partial x_i \partial x_j$ y $\partial^2 f/\partial x_j \partial x_i$ existen y son continuas en U. Entonces*

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

Demostración. Por la Observación 11.3, podemos suponer $m = 1$. Además, dejando quietas todas las variables menos dos y aplicando la definición de derivada parcial, es obvio que podemos también suponer $n = 2$. Así, queremos ver que

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

para cada $(x, y) \in U$. Fijemos $(x, y) \in U$, y para todos $h, k \in \mathbb{R}$ suficientemente pequeños, consideremos la cantidad

$$S_{h,k} = [f(x + h, y + k) - f(x, y + k)] - [f(x + h, y) - f(x, y)].$$

Definamos para cada k la función g_k por

$$g_k(u) = f(u, y+k) - f(u, y),$$

y obsérvese que

$$S_{h,k} = g_k(x+h) - g_k(x).$$

Aplicando el teorema del valor medio, podemos escribir entonces

$$S_{h,k} = g_k'(c_{k,h})h$$

para cierto $c_{k,h}$ entre x y $x+h$. Es decir,

$$S_{h,k} = \left(\frac{\partial f}{\partial x}(c_{k,h}, y+k) - \frac{\partial f}{\partial x}(c_{k,h}, y)\right) h = \frac{\partial^2 f}{\partial y \partial x}(c_{k,h}, d_{k,h})hk \qquad (1)$$

para cierto $d_{k,h}$ entre y e $y+k$, si aplicamos otra vez el teorema del valor medio a la diferencia de las primeras derivadas parciales.

Puesto que la expresión $S_{h,k}$ es simétrica en h, k y x, y, intercambiando los dos términos del medio en la expresión que define $S_{h,k}$, podemos deducir de manera exactamente análoga que

$$S_{h,k} = \frac{\partial^2 f}{\partial x \partial y}(\tilde{c}_{k,h}, \tilde{d}_{k,h})hk \qquad (2)$$

para ciertos $\tilde{c}_{k,h}, \tilde{d}_{k,h}$ entre x y $x+h$ e y e $y+k$. Igualando (1) y (2), cancelando hk, y haciendo $(h,k) \to (0,0)$ y usando la continuidad de las derivadas parciales segundas, llegamos al resultado buscado. $\square$

Observación 11.5. Este resultado puede refinarse para obtener lo siguiente: si f es de clase C^1 y $\partial^2 f/\partial x \partial y$ existe y es continua, entonces $\partial^2 f/\partial y \partial x$ también existe, y son iguales. Ver el ejercicio 11.2.

Sin embargo, la mera existencia de las derivadas parciales cruzadas de segundo orden, no garantiza su igualdad, como prueba el ejemplo siguiente.

Ejemplo 11.3. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

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

Calcular las derivadas parciales de primer y segundo orden de f, y comprobar que

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

Diremos que $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función de clase C^2 si todas sus derivadas parciales de segundo orden existen y son continuas en U. Más en general podemos definir funciones de clase C^p para $p = 1, 2, 3, ..., \infty$.

Definición 11.2. Sean $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $p \in \mathbb{N}$. Diremos que f es de clase C^p en U, y se denota $f \in C^p(U, \mathbb{R}^m)$, si las derivadas parciales de orden p,

$$\frac{\partial^p f}{\partial x_{i_1} \partial x_{i_2} ... \partial x_{i_p}}(x) := D_{i_1} D_{i_2} ... D_{i_p} f(x)$$

existen para cada $x \in U$, y las aplicaciones

$$U \ni x \to \frac{\partial^p f}{\partial x_{i_1} \partial x_{i_2} ... \partial x_{i_p}}(x)$$

son continuas en U, para todos $i_1, i_2, ..., i_p \in \{1, 2, ..., n\}$.

Diremos que f es de clase C^∞ en U, y escribiremos $f \in C^\infty(U, \mathbb{R}^m)$, si f es de clase C^p para todo $p \in \mathbb{N}$.

Por último, $C^0(U, \mathbb{R}^m)$ denotará el espacio de las funciones continuas de U en $\mathbb{R}^m$.

Proposición 11.6. *Para todo $p \in \mathbb{N}$ se tiene que*

$$C^p(U, \mathbb{R}^m) \subset C^{p-1}(U, \mathbb{R}^m);$$

y en particular $C^\infty(U, \mathbb{R}^m) \subset C^p(U, \mathbb{R}^m)$ para todo p.

Demostración. Es consecuencia inmediata del Teorema 9.7, de la Proposición 9.1, y de la definición de derivadas parciales de orden superior. □

La demostración de la siguiente proposición es también inmediata.

Proposición 11.7. *Una función $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ está en $C^p(U, \mathbb{R}^m)$ si y sólo si $D_i f = \partial f / \partial x_i$ está en $C^{p-1}(U, \mathbb{R}^m)$ para todo $i = 1, 2, ..., n$.*

Ejemplo 11.4. Las siguientes aplicaciones son todas de clase C^∞.

1. $S : \mathbb{R}^2 \to \mathbb{R}$, $S(x, y) = x + y$.
2. $P : \mathbb{R}^2 \to \mathbb{R}$, $P(x, y) = xy$.
3. Cualquier función lineal $T : \mathbb{R}^n \to \mathbb{R}^m$.
4. Cualquier aplicación bilineal $B : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$.

De hecho las derivadas de cualquier orden mayor o igual que 3 de todas estas funciones son cero, con lo que queda claro que son todas ellas continuas.

Análogamente, cualquier aplicación multilineal $A : \mathbb{R}^{n_1} \times ... \times \mathbb{R}^{n_k} \to \mathbb{R}^m$ es de clase C^∞, ya que todas sus derivadas de orden mayor o igual que $k + 1$ se anulan.

Teorema 11.8. *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ de clase C^p con $p \geq 2$. Sean $i_1, ..., i_q \in \{1, ..., n\}$, con $q \leq p$. Entonces, para toda permutación $\{i'_1, ..., i'_q\}$ de $\{i_1, ..., i_q\}$, se tiene que*

$$\frac{\partial^q f}{\partial x_{i'_1} ... \partial x_{i'_q}}(x) = \frac{\partial^q f}{\partial x_{i_1} ... \partial x_{i_q}}(x)$$

para todo $x \in U$.

Demostración. Basta combinar el Teorema de Schwartz 11.4 con el hecho de que toda permutación es producto de transposiciones y un simple argumento de inducción. □

A continuación veremos que la combinación (por composición, suma, producto o cociente) de funciones de clase C^p es de clase C^p.

Teorema 11.9. *Sea $p \in \mathbb{N} \cup \{\infty\}$. Entonces:*

1. *La composición de funciones de clase C^p es de clase C^p. Es decir, si $f \in C^p(U, \mathbb{R}^m)$, $g \in C^p(V, \mathbb{R}^k)$ y $f(U) \subseteq V$. Entonces $g \circ f \in C^p(U, \mathbb{R}^k)$.*

2. *La suma de funciones de clase C^p es de clase C^p.*

3. *El producto de funciones de clase C^p es de clase C^p. Es decir, si $f \in C^p(U, \mathbb{R}^k)$, $\lambda \in C^p(U, \mathbb{R})$ entonces $\lambda f \in C^p(U, \mathbb{R}^k)$.*

Demostración. Se deja como ejercicio. Debe usarse inducción sobre p, y emplearse la regla de la cadena en la forma siguiente:

$$D_i(g \circ f) = \sum_{j=1}^{m} [(D_j g) \circ f] D_i f_j,$$

entendiendo también que la hipótesis de inducción no se refiere a dos funciones fijas, sino que se consideran a la vez todo el conjunto de funciones y conjuntos de dominios posibles y las tres partes del enunciado. □

Observación 11.10. El cociente de funciones de clase C^p, cuando está bien definido, es también de clase C^p. Es decir, si $f, g \in C^p(U, \mathbb{R})$ y $g(x) \neq 0$ para todo $x \in U$, entonces $f/g \in C^p(U, \mathbb{R})$. Basta tener en cuenta que

$$\frac{f}{g} = f \cdot (h \circ g),$$

donde $h : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $h(t) = 1/t$, es de clase C^∞ (ejercicio 11.3).

Vamos a ver ahora que si (f_k) es una sucesión de funciones de $C^p(U, \mathbb{R})$ tal que ella y todas las sucesiones de sus derivadas parciales convergen uniformemente en U, entonces $f = \lim_{k\to\infty} f_k$ es también de clase C^p en U.

Teorema 11.11. *Sea U abierto en $\mathbb{R}^n$, y sea (f_k) una sucesión en $C^p(U, \mathbb{R})$. Supongamos que todas las sucesiones de derivadas parciales de (f_k) convergen uniformemente en U, es decir, para cada $i_1, ..., i_q \in \{1, ..., n\}$ con $q \leq p$, la sucesión*

$$\left(\frac{\partial^q f_k}{\partial x_{i_1} ... \partial x_{i_q}} \right)_{k=1}^{\infty}$$

converge uniformemente en U, y que (f_k) converge puntualmente a una función f. Entonces

$$f = \lim_{k\to\infty} f_k \in C^p(U, \mathbb{R}),$$

y además

$$\frac{\partial^q f}{\partial x_{i_1} ... \partial x_{i_q}} = \lim_{k\to\infty} \frac{\partial^q f_k}{\partial x_{i_1} ... \partial x_{i_q}}$$

uniformemente en U, para todos $i_1, ..., i_q \in \{1, ..., n\}$ con $q \leq p$.

Demostración. Por inducción sobre p.

Veamos que el resultado es cierto para $p = 1$. Denotemos, para cada $i = 1, 2, ..., n$,

$$g_i = \lim_{k\to\infty} \frac{\partial f_k}{\partial x_i};$$

como este límite es uniforme en U y las funciones $\partial f_k/\partial x_i$ son continuas, se tiene que g_i es continua en U. Debemos comprobar que

$$\frac{\partial f}{\partial x_i}(x) = g_i(x)$$

para todo $x \in U$, $i = 1, ..., n$, con lo cual quedará demostrado que f tiene derivadas paciales continuas y por tanto es de clase C^1 en U. Fijemos pues $x \in U, i \in \{1, 2, ..., n\}$. Usando el teorema fundamental del cálculo, tenemos que

$$\begin{aligned}
&f(x+te_i) - f(x) = \lim_{k\to\infty}[f_k(x+te_i) - f_k(x)] = \\
&\lim_{k\to\infty}\int_0^1 t\frac{\partial f_k}{\partial x_i}(x+tse_i)ds = \int_0^1 \lim_{k\to\infty} t\frac{\partial f_k}{\partial x_i}(x+tse_i)ds = \\
&\int_0^1 tg_i(x+tse_i)ds = \int_0^t g_i(x+ue_i)du,
\end{aligned}$$

donde hemos podido pasar el límite dentro de la integral gracias a la convergencia uniforme de $\partial f_k/\partial x_i$ a g_i. Entonces, tomando límites cuando $t \to 0$, obtenemos que

$$\frac{\partial f}{\partial x_i}(x) := \lim_{t\to 0}\frac{f(x+te_i) - f(x)}{t} = \lim_{t\to 0}\frac{\int_0^t g_i(x+ue_i)du}{t} = g_i(x)$$

gracias otra vez al teorema fundamental del cálculo.

Veamos ahora que si el resultado es cierto para p, también lo es para $p+1$. Podemos suponer $q = p+1$ (si $q \leq p$ no hay nada que no nos diga ya la hipótesis de inducción). Sea f una función que satisface las hipótesis del enunciado para $q \leq p+1$. En este caso queremos ver que

$$\frac{\partial^{p+1} f}{\partial x_{i_1}...\partial x_{i_{p+1}}} = g_{i_1,...,i_{p+1}} := \lim_{k\to\infty}\frac{\partial^{p+1} f_k}{\partial x_{i_1}...\partial x_{i_{p+1}}},$$

y como este es un límite uniforme de funciones continuas se deducirá que las derivadas parciales de orden $p+1$ de f existen y son continuas, es decir $f \in C^{p+1}(U, \mathbb{R})$. Consideremos las funciones

$$h = \frac{\partial^p f}{\partial x_{i_2}...\partial x_{i_{p+1}}}, \ h_k = \frac{\partial^p f_k}{\partial x_{i_2}...\partial x_{i_{p+1}}}.$$

Por hipótesis tenemos que

$$\lim_{k\to\infty}\frac{\partial h_k}{\partial x_{i_1}} = \lim_{k\to\infty}\frac{\partial^{p+1} f_k}{\partial x_{i_1}...\partial x_{i_{p+1}}} = g_{i_1,...,i_{p+1}},$$

donde el límite es uniforme en U, y además h_k converge a h (uniformemente) en U, puesto que estamos suponiendo que el resultado es cierto para p.

Entonces, aplicando que el resultado es cierto para $p = 1$, deducimos que $\partial h/\partial x_1$ existe y es igual a $g_{i_1,...,i_{p+1}}$. Esto quiere decir que

$$\frac{\partial^{p+1} f}{\partial x_{i_1}...\partial x_{i_{p+1}}} := \frac{\partial h}{\partial x_{i_1}} = g_{i_1,...,i_{p+1}},$$

que es lo que queríamos. □

Este resultado puede también emplearse en forma de series: basta considerar la sucesión de las sumas parciales de la serie y aplicar el resultado anterior. Si lo combinamos con el criterio M de Weierstrass (ver el problema 7.14), obtenemos lo siguiente.

Corolario 11.12. *Sea U abierto en $\mathbb{R}^n$, y sea (g_k) una sucesión en $C^p(U,\mathbb{R})$. Supongamos que todas las series de derivadas parciales de (g_k) están uniformemente acotadas por series numéricas absolutamente convergentes, es decir, para cada $i_1,...,i_q \in \{1,...,n\}$ con $q \leq p$ se tiene que*

$$|\frac{\partial^q g_k}{\partial x_{i_1}...\partial x_{i_q}}(x)| \leq M_k^{(i_1,...,i_q)}$$

para todos $x \in U$, $k \in \mathbb{N}$. Supongamos también que cada una de estas series

$$\sum_{k=1}^{\infty} M_k^{(i_1,...,i_q)}$$

es convergente en $\mathbb{R}$, y que $\sum_{k=1}^{\infty} g_k$ converge puntualmente a una función g. Entonces

$$g = \sum_{k=1}^{\infty} g_k \in C^p(U,\mathbb{R}),$$

y además

$$\frac{\partial^q g}{\partial x_{i_1}...\partial x_{i_q}} = \sum_{k=1}^{\infty} \frac{\partial^q g_k}{\partial x_{i_1}...\partial x_{i_q}}$$

uniformemente en U, para todos $i_1,...,i_q \in \{1,...,n\}$ con $q \leq p$.

En los problemas 11.12 y 11.13 pueden encontrarse dos aplicaciones interesantes de este Corolario.

Ejercicio 11.1. Extender el teorema y corolario anteriores al caso de funciones que toman valores vectoriales en $\mathbb{R}^m$ (no hay que repetir la demostración, puede reducirse todo al caso ya conocido $m = 1$).

Por último vamos a definir qué es una función varias veces diferenciable y cuáles son sus derivadas de orden superior. Para $k = 1$ nos remitimos a la definición de diferencial dada en el Capítulo 8.3, que tiene siempre sentido para funciones $f : U \subseteq E \to F$, donde E y F son espacios normados de dimensión finita.

Si tenemos una función $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ que es diferenciable en el abierto U, podemos considerar su diferencial

$$Df : U \subseteq \mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n,\mathbb{R}),\ x \mapsto Df(x),$$

y puesto que $F = \mathcal{L}(\mathbb{R}^n,\mathbb{R})$ es un espacio normado de dimensión finita, preguntarnos si Df es a su vez diferenciable en $a \in U$ según la definición 8.3, es decir, si existe una aplicación lineal $A \in \mathcal{L}(\mathbb{R}^n, F) = \mathcal{L}(\mathbb{R}^n,\mathcal{L}(\mathbb{R}^n,\mathbb{R}))$ tal que

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

donde el límite se toma en el espacio normado $\mathcal{L}(\mathbb{R}^n,\mathbb{R})$, que es de dimensión n y por tanto isomorfo a $\mathbb{R}^n$. En este caso diremos que f es diferenciable dos veces en a, y que

$$D^2 f(a) = D(Df)(a) = A$$

es su diferencial en a, una aplicación lineal de $\mathbb{R}^n$ en $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$.

Ahora bien, el espacio $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$ es un poco difícil de manejar conceptualmente. Por ello conviene identificarlo con el espacio $\mathcal{L}^2(\mathbb{R}^n, \mathbb{R})$ de las aplicaciones bilineales de $\mathbb{R}^n$ en $\mathbb{R}$ (es decir, las aplicaciones $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ tales que $x \mapsto B(x, y)$ es lineal de $\mathbb{R}^n$ en $\mathbb{R}$ para cada $y \in \mathbb{R}^n$, y $y \mapsto B(x, y)$ es también lineal para cada $x \in \mathbb{R}^n$). En efecto, a cada elemento $A \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$ le hacemos corresponder $B \in \mathcal{L}^2(\mathbb{R}^n, \mathbb{R})$ definida por

$$B(x, y) = A(x)(y),$$

y recíprocamente, si $B \in \mathcal{L}^2(\mathbb{R}^n, \mathbb{R})$ entonces $x \to A(x) = B(x, \cdot)$ es una aplicación de $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$. Esta correspondencia es por tanto biyectiva, y lineal, de modo que podemos identificar

$$\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R})) = \mathcal{L}^2(\mathbb{R}^n, \mathbb{R}).$$

Recordemos también que una forma bilineal $B \in \mathcal{L}^2(\mathbb{R}^n, \mathbb{R})$ queda determinada por los valores $B(e_i, e_j) = b_{ij}$, donde $\{e_1, ..., e_n\}$ es la base canónica de $\mathbb{R}^n$; en efecto

$$B(x, y) = B(\sum x_i e_i, \sum y_j e_j) = \sum x_i B(e_i, \sum y_j e_j) =$$
$$\sum_i \sum_j x_i y_j B(e_i, e_j) = \sum_i \sum_j b_{ij} x_i y_j,$$

o bien, en forma matricial, $B(x, y) = x^t (b_{ij}) y$. Otra forma de expresar esto es poniendo

$$B = \sum_{i,j=1}^{n} b_{ij} e_i^* \otimes e_j^*,$$

donde $e_i^* \otimes e_j^*$ es la forma bilineal de $\mathcal{L}^2(\mathbb{R}^n, \mathbb{R})$ definida por

$$(e_i^* \otimes e_j^*)(x, y) = e_i^*(x) e_j^*(y) = x_i y_j.$$

No es difícil comprobar que $\{e_i^* \otimes e_j^* : 1 \leq i, j \leq n\}$ es una base de $\mathcal{L}^2(\mathbb{R}^n, \mathbb{R})$, que tiene dimensión n^2 (ver el problema 11.14).

Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función dos veces diferenciable en a. Nos preguntamos ahora cuál será la matriz de la forma bilineal $D^2 f(a)$ respecto de la base canónica de $\mathbb{R}^n$. Puesto que

$$Df(x) = (\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n}) = (Df(x)(e_1), ..., Df(x)(e_n)),$$

derivando otra vez tendremos que

$$D^2 f(a) = D(Df)(a) = D((\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n}))(a),$$

luego

$$D(Df)(a)(e_i) = (\frac{\partial^2 f}{\partial x_i \partial x_1}(a), ..., \frac{\partial^2 f}{\partial x_i \partial x_n}(a)),$$

y así

$$D^2 f(a)(e_i, e_j) = D(Df)(a)(e_i)(e_j) = \frac{\partial^2 f}{\partial x_i \partial x_j},$$

es decir, la matriz de $D^2f(a)$ es

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(a) & ... & \frac{\partial^2 f}{\partial x_n \partial x_1}(a) \\ \vdots & ... & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) & ... & \frac{\partial^2 f}{\partial x_n \partial x_n}(a) \end{pmatrix},$$

y tenemos que

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

Si la función f es de clase C^2, el teorema de Schwartz nos indica que esta matriz es simétrica. El siguiente resultado nos asegura que basta que f sea diferenciable dos veces en a para que la matriz de $D^2f(a)$ sea simétrica.

Teorema 11.13. *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable en U, y supongamos que*

$$\frac{\partial f}{\partial x_i}, \frac{\partial f}{\partial x_j} : U \to \mathbb{R}$$

son ambas diferenciables en $a \in U$. Entonces

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

En particular, si f es dos veces diferenciable en a, la matriz de su diferencial segunda $D^2f(a)$ es simétrica.

Demostración. Es parecida a la del teorema de Schwartz 11.4, y como en este caso podemos suponer $n = 2$.

Fijemos $(a, b) \in U$, y para todo $h \in \mathbb{R}$ suficientemente pequeño, consideremos la función

$$F(h) = f(a + h, b + h) - f(a + h, b) - f(a, b + h) + f(a, b)].$$

Si ponemos $\varphi(x) = f(x, b + h) - f(x, b)$, entonces $F(h) = \varphi(a + h) - \varphi(a)$, y aplicando el teorema del valor medio obtenemos

$$F(h) = \varphi'(t_h)h = [\frac{\partial f}{\partial x}(t_h, b + h) - \frac{\partial f}{\partial x}(t_h, b)]h,$$

para cierto t_h entre a y $a + h$. Como $\partial f/\partial x$ es diferenciable en (a, b), podemos escribir

$$\frac{\partial f}{\partial x}(t_h, b + h) = \frac{\partial f}{\partial x}(a, b) + \frac{\partial^2 f}{\partial x^2}(a, b)(t_h - a) + \frac{\partial^2 f}{\partial y \partial x}(a, b)h + \epsilon_1(h)|h|,$$

y también

$$\frac{\partial f}{\partial x}(t_h, b) = \frac{\partial f}{\partial x}(a, b) + \frac{\partial^2 f}{\partial x^2}(a, b)(t_h - a) + \epsilon_2(h)|h|,$$

donde ϵ_1, ϵ_2 son funciones de h tales que $\lim_{h\to 0} \epsilon_i(h) = 0$. Por tanto

$$F(h) = \frac{\partial^2 f}{\partial y \partial x}(a, b)h^2 + \epsilon(h)|h|^2, \tag{1}$$

donde $\epsilon(h) = \epsilon_1(h) - \epsilon_2(h)$ y en particular $\lim_{h\to 0} \epsilon(h) = 0$.

Por otra parte, si ponemos $\psi(y) = f(a+h,y) - f(a,y)$, se tiene $F(h) = \psi(b+h) - \psi(b)$, y una repetición del argumento anterior nos da que

$$F(h) = \frac{\partial^2 f}{\partial x \partial y}(a,b)h^2 + \overline{\epsilon}(h)|h|^2, \tag{2}$$

donde $\lim_{h\to 0} \overline{\epsilon}(h) = 0$. Juntando (1) y (2) obtenemos que

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

que es lo que pretendíamos demostrar. □

Ejemplo 11.5. Demostrar que la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$f(x,y) = e^x y + \cos(xy)$$

es dos veces derivable en $\mathbb{R}^2$, y calcular la matriz de $D^2 f(x,y)$. Hallar también

$$D^2 f(1,3)\left((1,5),(7,1)\right).$$

Denotaremos por $\mathcal{L}_s^2(\mathbb{R}^n, \mathbb{R})$ el espacio de las formas bilineales simétricas de $\mathbb{R}^n$ en $\mathbb{R}$, que naturalmente es un subespacio vectorial de $\mathcal{L}^2(\mathbb{R}^n, \mathbb{R})$. Así pues, el teorema anterior nos dice que $D^2 f(a) \in \mathcal{L}_s^2(\mathbb{R}^n, \mathbb{R})$ si f es dos veces diferenciable en a.

Podemos definir por inducción las diferenciales de orden superior o igual a 2 de una función $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ siguiendo el mismo método que hemos utilizado para llegar a $D^2 f(a) = D(Df)(a)$. Supongamos que f es una función k veces diferenciable en U, y consideremos la aplicación

$$U \ni x \to D^k(x) \in \mathcal{L}_s^k(\mathbb{R}^n, \mathbb{R}),$$

donde $\mathcal{L}_s^k(\mathbb{R}^n, \mathbb{R})$ es el espacio de las formas k-lineales simétricas de $\mathbb{R}^n$, es decir, el conjunto de todas las aplicaciones $A : \mathbb{R}^n \times ... \times \mathbb{R}^n \to \mathbb{R}$ (k veces) que son separadamente lineales en cada variable, y de modo que $A(x_1, ..., x_k) = A(x_{i_1}, ..., x_{i_k})$ para todos $x_1, ..., x_k \in \mathbb{R}^n$ y cualquier permutación $\{i_1, ..., i_k\}$ de $\{1, ..., k\}$.

Nótese que $E = \mathcal{L}_s^k(\mathbb{R}^n, \mathbb{R})$ es un espacio vectorial de dimensión finita (isomorfo a $\mathbb{R}^{n^k}$), y aunque no hemos especificado ninguna norma en este espacio, como todas son equivalentes podemos elegir cualquiera para aplicar la definición de función diferenciable entre $U \subseteq \mathbb{R}^n$ y E, sabiendo por el problema 8.10 que la definición no va a depender de la norma usada. Así, diremos que f es $k+1$ veces diferenciable en $a \in U$ si la aplicación $D^k f : U \to E$ es diferenciable en a, y de

$$D(D^k f)(a) := D^{k+1} f(a)$$

diremos que es la diferencial de orden $k+1$ de f en a. Obsérvese que en principio $D^{k+1} f(a) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}^k(\mathbb{R}^n, \mathbb{R}))$, pero, como en el caso de las formas bilineales, podemos identificar una aplicación $A \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}^k(\mathbb{R}^n, \mathbb{R}))$ con la aplicación $k+1$ lineal $B \in \mathcal{L}^{k+1}(\mathbb{R}^n, \mathbb{R})$ definida por

$$B(x_1, x_2, ..., x_{k+1}) = A(x_1)(x_2, ..., x_{k+1}),$$

y tener en cuenta, inversamente, que toda $B \in \mathcal{L}^{k+1}(\mathbb{R}^n, \mathbb{R})$ es de esta forma, para $A \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}^k(\mathbb{R}^n, \mathbb{R}))$ definida por

$$A(x)(\cdot, ..., \cdot) = B(x, \cdot, ..., \cdot),$$

y esta correspondencia biyectiva entre $\mathcal{L}(\mathbb{R}^n, \mathcal{L}^k(\mathbb{R}^n, \mathbb{R}))$ y $\mathcal{L}^{k+1}(\mathbb{R}^n, \mathbb{R})$ es también lineal, luego podemos identificar ambos espacios vectoriales. Además, aplicando reiteradamente el Teorema 11.13, se constata fácilmente que la aplicación $k+1$ lineal $D^{k+1}f(a)$ es simétrica, y así podemos considerar a todos los efectos que

$$D^{k+1}f(a) \in \mathcal{L}_s^{k+1}(\mathbb{R}^n, \mathbb{R}).$$

Veamos ahora cómo calcular estas derivadas de orden superior. En los casos $k = 1$ y $k = 2$ ya sabemos cómo hacerlo. En el caso general, una reiteración del argumento que nos permitió ver que

$$D^2 f(a)(e_i, e_j) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a)$$

muestra que si $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ es k veces diferenciable en $a \in U$ entonces

$$D^k f(a)(e_{i_1}, ..., e_{i_k}) = \frac{\partial^k f}{\partial x_{i_1} ... \partial x_{i_k}}(a)$$

para todos $e_{i_1}, ..., e_{i_k} \in \{e_1, ..., e_n\}$ de la base canónica de $\mathbb{R}^n$, con lo cual $D^k f(a)$ queda determinada aplicando multilinealidad. De hecho, si definimos $(e_{i_1}^* \otimes ... \otimes e_{i_k}^*) \in \mathcal{L}^k(\mathbb{R}^n, \mathbb{R})$ por

$$(e_{i_1}^* \otimes ... \otimes e_{i_k}^*)(h^1, ..., h^k) = h_{i_1}^1 ... h_{i_k}^k,$$

resulta que $\{(e_{i_1}^* \otimes ... \otimes e_{i_k}^*) : 1 \leq i_1, ..., i_k \leq n\}$ es una base de $\mathcal{L}^k(\mathbb{R}^n, \mathbb{R})$, que tiene dimensión n^k. Con respecto a esta base podemos expresar $D^k f(a)$ como sigue:

$$D^k f(a) = \sum_{i_1, ..., i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_1} ... \partial x_{i_k}}(a)(e_{i_1}^* \otimes ... \otimes e_{i_k}^*).$$

Debe observarse que todo lo que se ha hecho hasta ahora para funciones de $\mathbb{R}^n$ en $\mathbb{R}$ puede extenderse sin problemas al caso de funciones con valores vectoriales en $\mathbb{R}^m$, sin más que tener cuenta que una función con valores vectoriales es diferenciable si y sólo si cada una de sus funciones componentes es diferenciable. Para una función $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ que sea k veces diferenciable, se tendrá

$$D^k f(a) \in \mathcal{L}_s^k(\mathbb{R}^n, \mathbb{R}^m),$$

el espacio de las aplicaciones multilineales simétricas de $\mathbb{R}^n \times ... \times \mathbb{R}^n$ (k veces) en $\mathbb{R}^m$.

Finalmente, vamos a generalizar la Proposición 11.1, probando que toda función de clase C^p es p veces derivable y que su diferencial de orden k es continua.

Proposición 11.14. *Sea U un abierto de $\mathbb{R}^n$, y sea $f : U \to \mathbb{R}^m$. Las siguientes afirmaciones son equivalentes.*

1. *$f \in C^p(U, \mathbb{R}^m)$, es decir todas las derivadas parciales de orden p de f existen y son continuas en U.*

2. *f es p veces diferenciable en U, y la aplicación*

   $$U \ni x \to D^p f(x) \in \mathcal{L}_s^p(\mathbb{R}^n, \mathbb{R}^m)$$

   es continua.

Demostración. Puede suponerse $m = 1$, ¿por qué?.
$(2) \implies (1)$: Como la aplicación

$$U \ni x \to D^{k+1}f(x) \in \mathcal{L}^{k+1}(\mathbb{R}^n, \mathbb{R})$$

es continua, y también lo es

$$\mathcal{L}^{k+1}(\mathbb{R}^n, \mathbb{R}) \ni A \to A(e_{i_1}, ..., e_{i_{k+1}}) \in \mathbb{R},$$

resulta que la composición de estas dos aplicaciones,

$$U \ni x \to D^{k+1}f(x)(e_{i_1}, ..., e_{i_{k+1}}) = \frac{\partial^{k+1} f}{\partial x_{i_1} ... \partial x_{i_{k+1}}}(x) \in \mathbb{R},$$

es continua para cada $e_{i_1}, ..., e_{i_{k+1}} \in \{e_1, ..., e_n\}$, lo que es decir que f es de clase $C^p(U, \mathbb{R})$.
$(1) \implies (2)$: Por inducción sobre p. Para $p = 1$ el resultado es ya conocido: la Proposición 11.1. Supongamos que el resultado es cierto para $p = k$, y comprobemos que también lo es entonces para $p = k + 1$. Primero observemos que la hipótesis de que todas las derivadas parciales de orden $k+1$ sean continuas implica que todas las derivadas parciales de orden menor o igual que k existen y son diferenciables y en particular continuas. Luego por la hipótesis de inducción se tiene que f es k veces diferenciable, y $D^k f$ es continua en U. Tenemos que ver que $D^k f$ es diferenciable en U, y que $D(D^k f) = D^{k+1} f$ es continua. Consideremos por tanto

$$U \ni x \to D^k f(x) \in E := \mathcal{L}^k(\mathbb{R}^n, \mathbb{R}).$$

Por la discusión anterior al enunciado de este teorema sabemos que E tiene dimensión n^k y que $\{(e_{i_1}^* \otimes ... \otimes e_{i_k}^*) : 1 \leq i_1, ..., i_k \leq n\}$ es una base suya, y que podemos escribir

$$D^k f(x) = \sum_{i_1, ..., i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_1} ... \partial x_{i_k}}(x)(e_{i_1}^* \otimes ... \otimes e_{i_k}^*).$$

Es decir, podemos ver $D^k f$ como una aplicación de $\mathbb{R}^n$ en $\mathbb{R}^{n^k}$ cuyas funciones componentes son las derivadas parciales

$$x \mapsto \frac{\partial^k f}{\partial x_{i_1} ... \partial x_{i_k}}(x). \qquad (*)$$

Entonces, por la Proposición 8.4 sabemos que $D^k f$ será diferenciable con continuidad en U si y sólo si lo son estas funciones componentes. Es decir, basta demostrar que las derivadas parciales de orden k son diferenciables con continuidad. Como cada una de ellas es una función de $\mathbb{R}^n$ en $\mathbb{R}$, aplicando el caso $p = 1$ basta comprobar entonces que las derivadas parciales de orden uno de cada función de $(*)$, es decir,

$$x \mapsto \frac{\partial}{\partial x_j} \left(\frac{\partial^k f}{\partial x_{i_1} ... \partial x_{i_k}}(x) \right) = \frac{\partial^{k+1} f}{\partial x_j \partial x_{i_1} ... \partial x_{i_k}}(x), \ j = 1, ..., n,$$

son continuas, lo cual es evidente por la suposición inicial ya que se trata de derivadas parciales de orden $k + 1$. $\square$

11.1. Problemas

Problema 11.1. Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ una aplicación de clase C^1, y sea $a \in U$ tal que $Df(a)$ es un isomorfismo lineal (es decir, $\det Df(a) \neq 0$). Probar que existe $\delta > 0$ tal que $f_{|B(a,\delta)} : B(a,\delta) \to \mathbb{R}^n$ es inyectiva. *Indicación:* Combinar la Proposición 11.2 y el problema 5.59.

Problema 11.2. Sea $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ de clase C^1, y supongamos que $\partial^2 f/\partial x \partial y$ existe y es continua en U. Demostrar que entonces $\partial^2 f/\partial y \partial x$ también existe en U, y que

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Indicación: Modificar la demostración del Teorema 11.4.

Problema 11.3. Demostrar que la función $h : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $h(t) = 1/t$, es de clase C^∞.

Problema 11.4. Sea $F : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, definida por $F(x,y) = (f(x,y), g(x,y))$, en donde

$$f(x,y) = y^2 \sin(x/y) \qquad \text{si } y \neq 0; \qquad f(x,0) = 0,$$

y g es diferenciable en (0,0), tal que: $g(0,0) = 0$, $D_u g(0,0) = 1$, $D_v g(0,0) = 0$, donde u y v los vectores $(1,1)$ y $(-1,1)$, respectivamente.

1. Calcúlese $\nabla g(0,0)$.
2. Estúdiese la diferenciabilidad de F en (0,0).
3. Demuéstrese que la función $G = (F \circ F) + F$ de $\mathbb{R}^2$ en $\mathbb{R}^2$ es diferenciable y calcúlese $DG(0,0)$.
4. Demuéstrese que $\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$.

Problema 11.5. ¿Se cumple el Teorema de Schwarz sobre las parciales cruzadas para la siguiente función?

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

Problema 11.6. Sea $P : \mathbb{C} \to \mathbb{C}$ un polinomio complejo. Identificando $\mathbb{C} \equiv \mathbb{R}^2$, probar que P es una aplicación de clase $C^\infty(\mathbb{R}^2, \mathbb{R}^2)$. *Indicación:* Probar primero que la multiplicación compleja es una aplicación de clase C^∞ de $\mathbb{R}^2 \times \mathbb{R}^2$ en $\mathbb{R}^2$.

Problema 11.7. Sean $P, Q : \mathbb{C} \to \mathbb{C}$ polinomios complejos. Probar que su cociente P/Q es una aplicación de clase $C^\infty(U, \mathbb{R}^2)$, donde $U = \{z \in \mathbb{R}^2 : Q(z) \neq 0\}$. *Indicación:* Probar primero que la aplicación $(z,w) \to z/w$ es de clase C^∞ de $\mathbb{R}^2 \times \{w \in \mathbb{R}^2 : w \neq 0\}$ en $\mathbb{R}^2$.

Problema 11.8. Sea $f : \mathbb{R} \to \mathbb{R}$ definida por $f(t) = t^p$ para $t > 0$, y $f(t) = 0$ si $t \leq 0$. Probar que f es de clase C^k si $k < p$, pero no lo es si $k \geq p$.

Problema 11.9. Sea $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(t) = \begin{cases} e^{-1/t} & \text{si } t > 0, \\ 0 & \text{si } t \leq 0. \end{cases}.$$

Probar que $f \in C^\infty(\mathbb{R}, \mathbb{R})$. Si $a < b$, dibujar la gráfica de

$$g(t) = \frac{f(t-a)}{f(t-a) + f(b-t)}$$

y demostrar que $g \in C^\infty(\mathbb{R}, \mathbb{R})$.

Problema 11.10. Usando el problema anterior, definir una función $\varphi : \mathbb{R}^n \to [0,1]$ tal que $\varphi(x) = 1$ si $x \in B(a,r)$; $\varphi(x) = 0$ si y sólo si $x \notin B(a,R)$, y $\varphi \in C^\infty(\mathbb{R}^n, \mathbb{R})$, donde $0 < r < R$, $a \in \mathbb{R}^n$ son dados, y las bolas son euclideas.

Problema 11.11. Construir una función convexa $\theta : \mathbb{R} \to \mathbb{R}$ de clase $C^\infty(\mathbb{R}, \mathbb{R})$ tal que $\theta(t) = 0$ si $t \leq 0$, y $\theta(t) > 0$ si $t > 0$. *Indicación:* Usando el problema anterior puede construirse $\alpha : \mathbb{R} \to \mathbb{R}$ de clase $C^\infty(\mathbb{R}, [0,1])$ tal que $\alpha(t) > 0$ si y sólo si $t \in (0,1)$; ahora, definir $\beta(t) = \int_0^t \alpha(s)ds$, y $\theta(t) = \int_0^t \beta(s)ds$.

Problema 11.12. Demostrar que todo cerrado C de $\mathbb{R}^n$ es el conjunto de ceros de alguna función $g : \mathbb{R}^n \to [0, \infty)$ de clase C^∞. *Indicación:* Seguir los siguientes pasos:

1. Expresar $U := \mathbb{R}^n \setminus C$ como unión numerable de bolas euclideas,
$$U = \bigcup_{k=1}^{\infty} B(a_k, r_k).$$
2. Para cada bola $B(a_k, r_k)$ construir una función $\varphi_k \in C^\infty(\mathbb{R}^n, [0,1])$ tal que $\varphi_k(x) > 0$ si y sólo si $x \in B(a_k, r_k)$ (ver el problema 11.10).
3. Para cada $k, q \in \mathbb{N}$, definir
$$\alpha_k^q = \sum_{i_1,\ldots,i_q=1}^{n} \sup\{|\frac{\partial^q \varphi_k}{\partial x_{i_1} \ldots \partial x_{i_q}}(x)| : x \in \mathbb{R}^n\},$$
y comprobar que $\alpha_k^q < \infty$ (tener en cuenta que todas las derivadas parciales de φ_k son continuas y se anulan fuera de la bola $B(a_k, r_k)$).
4. Poner ahora
$$M_k = \frac{1}{2^k(1 + \sum_{i,j=1}^k \alpha_j^i)},$$
y definir $g = \sum_{k=1}^\infty M_k \varphi_k$. Usando el Corolario 11.12, probar que g es de clase C^∞ en $\mathbb{R}^n$ (para ello ver primero que
$$|\frac{\partial^q (M_k \varphi_k)}{\partial x_{i_1} \ldots \partial x_{i_q}}(x)| \leq \frac{\alpha_k^q}{2^k(1 + \sum_{i,j=1}^k \alpha_j^i)} \leq \frac{1}{2^k}$$
si $k \geq q$, para todo $x \in \mathbb{R}^n$).
5. Finalmente, comprobar que $g(x) = 0$ si y sólo si $x \in C$.

Problema 11.13. Demostrar ahora que todo cerrado *convexo* C en $\mathbb{R}^n$ es el conjunto de mínimos de alguna función *convexa* de clase C^∞ en $\mathbb{R}^n$. *Indicación:* Seguir los siguientes pasos:

1. Escribir $U = \mathbb{R}^n \setminus C$ como unión numerable de semiespacios abiertos (ver el problema 6.32), es decir, poner
$$U = \bigcup_{k=1}^{\infty} A_k,$$
con
$$A_k = \{x \in \mathbb{R}^n : f_k(x) > r_k\},$$
y donde $f_k \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$, $r_k \in \mathbb{R}$. Observar también que puede suponerse que $\|f_k\| = 1$ para todo k.

2. Sea $\theta : \mathbb{R} \to \mathbb{R}$ una función convexa de clase C^∞ tal que $\theta(t) = 0$ si $t \leq 0$ y $\theta(t) > 0$ para $t > 0$ (ver el problema 11.11). Definir $\varphi_k : \mathbb{R}^n \to [0, \infty)$ por
$$\varphi_k(x) = \theta(f_k(x) - r_k).$$
Demostrar que f_k es una función convexa, de clase C^∞, y tal que $f_k(x) > 0$ si y sólo si $x \in A_k$.

3. Para cada $k, q \in \mathbb{N}$, definir
$$\alpha_k^q = \sum_{i_1,...,i_q=1}^{n} \sup\{|\frac{\partial^q \varphi_k}{\partial x_{i_1}...\partial x_{i_q}}(x)| : x \in B(0,k)\},$$
y comprobar que $\alpha_k^q < \infty$.

4. Poner
$$M_k = \frac{1}{2^k(1 + \sum_{i,j=1}^k \alpha_j^i)},$$
y definir $g = \sum_{k=1}^\infty M_k \varphi_k$. Fijado $R > 0$, sea $k_0 \in \mathbb{N}$ con $k_0 \geq R$. Comprobar que
$$|\frac{\partial^q (M_k \varphi_k)}{\partial x_{i_1}...\partial x_{i_q}}(x)| \leq \frac{\alpha_k^q}{2^k(1 + \sum_{i,j=1}^k \alpha_j^i)} \leq \frac{1}{2^k}$$
si $k \geq \max\{q, k_0\}$, $x \in B(0, R)$, y usando el Corolario 11.12 concluir que $g \in C^\infty(B(0,R), \mathbb{R})$ para cada $R > 0$ y por tanto $g \in C^\infty(\mathbb{R}^n, \mathbb{R})$.

5. Demostrar que $g : \mathbb{R}^n \to [0, \infty)$ es convexa, y que $g^{-1}(0) = C$.

6. Demostrar también que $Dg(x) \neq 0$ si $x \notin C$ (ver el problema 9.22).

Cuando tengamos el teorema de la función implícita a nuestra disposición podremos deducir de aquí varios resultados interesantes sobre aproximación regular de conjuntos y funciones convexos; ver los problemas 15.40 y 15.41.

Problema 11.14. Demostrar que, si definimos $(e_{i_1}^* \otimes ... \otimes e_{i_k}^*) : \mathbb{R}^n \times ... \times \mathbb{R}^n \to \mathbb{R}$ (k veces) por
$$(e_{i_1}^* \otimes ... \otimes e_{i_k}^*)(h^1, ..., h^k) = h_{i_1}^1 ... h_{i_k}^k,$$
entonces $\{(e_{i_1}^* \otimes ... \otimes e_{i_k}^*) : 1 \leq i_1, ..., i_k \leq n\}$ es una base de $\mathcal{L}^k(\mathbb{R}^n, \mathbb{R})$, que tiene así dimensión n^k.

Capítulo 12

Teorema de Taylor. Aproximación.

Como ya es conocido en el caso de funciones reales de una variable real, el polinomio de Taylor de una función varias veces diferenciable nos permite aproximar localmente la función por un polinomio, lo que permite, entre muchas otras cosas, decidir sobre el carácter de los puntos críticos de una función sin más que examinar sus derivadas de orden superior en dichos puntos. En este capítulo veremos cómo la fórmula de Taylor puede extenderse al caso de funciones de varias variables. Definiremos también las funciones real analíticas de varias variables, y concluiremos estudiando otros métodos de aproximación (global) de funciones por funciones más regulares (de clase C^∞).

Recordemos que un monomio de orden k en $\mathbb{R}^n$ es una función de la forma

$$x = (x_1, ..., x_n) \mapsto x_{i_1}...x_{i_k},$$

con $i_1, ..., i_k \in \{1, ..., n\}$. Un polinomio k-homogéneo en $\mathbb{R}^n$ se define entonces como una función $P : \mathbb{R}^n \to \mathbb{R}$ de la forma

$$P(x_1, ..., x_n) = \sum_{i_1,...,i_k=1}^{n} a_{i_1,...,i_k} x_{i_1}...x_{i_k}, \qquad (*)$$

donde $a_{i_1,...,i_k}$ son constantes reales. Toda forma k-lineal simétrica $A \in \mathcal{L}^k(\mathbb{R}^n, \mathbb{R})$ define un polinomio k-homogéneo P_A mediante su restricción a la diagonal, es decir,

$$P_A(x) = A(x, ..., x).$$

Recíprocamente, es posible demostrar que para todo polinomio k-homogéneo P existe una única forma k-lineal *simétrica* A tal que $P = P_A$ (de hecho $A = D^k P$), aunque esto no lo usaremos en absoluto.

Nótese que si $P : \mathbb{R}^n \to \mathbb{R}$ es un polinomio k-homogéneo entonces $P(tv) = t^k P(v)$ para todo $t \in \mathbb{R}$ y todo $v \in \mathbb{R}^n$.

Se define un polinomio de grado m en $\mathbb{R}^n$ como toda suma (finita) de polinomios k-homogéneos de grado $k \leq m$ y constantes (que pueden considerarse monomios de grado cero).

En particular, si $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función de clase C^k, la diferencial $D^k f(a)$ define el polinomio k-homogéneo

$$P_{D^k f(a)}(h) = D^k f(a)(h, ..., h) = \sum_{i_1,...,i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1}...\partial x_{i_k}}(a) h_{i_1}...h_{i_k}.$$

Por abuso de notación, identificaremos la diferencial $D^k f(a)$ con este polinomio, es decir con su restricción a la diagonal.

Es claro que al escribir un polinomio k-homogéneo en la forma $(*)$ estamos repitiendo muchos monomios que podrían agruparse, sumando las constantes $a_{i_1...i_k}$ correspondientes y sacando factor común. Por ejemplo, un polinomio

$$\begin{aligned} P(x,y) = \quad & a_{111}x^3 + a_{112}x^2y + a_{121}xyx + a_{122}xy^2 + \\ & + a_{211}x^2y + a_{212}xy^2 + a_{221}y^2x + a_{222}y^3 \end{aligned}$$

se escribe de manera más conveniente como

$$P(x,y) = b_{(3,0)}x^3y^0 + b_{(2,1)}x^2y^1 + b_{(1,2)}x^1y^2 + b_{(0,3)}x^0y^3,$$

donde $b_{(3,0)} = a_{111}$, $b_{(2,1)} = a_{112} + a_{121} + a_{211}$, $b_{(1,2)} = a_{122} + a_{212} + a_{221}$, y $b_{(0,3)} = a_{222}$.

Más en general, cualquier monomio $x_{i_1}...x_{i_k}$ de grado k puede escribirse como

$$x_1^{\alpha_1}...x_n^{\alpha_n},$$

donde $\alpha_i \in \mathbb{N} \cup \{0\}$, y $\alpha_1 + ... + \alpha_n = k$. Por tanto todo polinomio k-homogéneo P en $\mathbb{R}^n$ podrá escribirse en la forma

$$P(x_1,...,x_n) = \sum_{\alpha_j \geq 0, \alpha_1+...+\alpha_n=k} b_{\alpha_1...\alpha_n} x_1^{\alpha_1}...x_n^{\alpha_n}.$$

Con esta notación, puede comprobarse que si f es de clase C^k entonces $D^kf(a)$ se expresa como

$$D^kf(a)(h) = \sum_{\alpha_j \geq 0, \alpha_1+...+\alpha_n=k} \frac{k!}{\alpha_1!...\alpha_n!} \frac{\partial^k f}{\partial x_1^{\alpha_1}...\partial x_n^{\alpha_n}} h_1^{\alpha_1}...h_n^{\alpha_n}.$$

Hay aún otra manera de escribir esto, usando la notación de *multi-índices*, a saber, poniendo $\alpha = (\alpha_1,...,\alpha_n)$, y

$$\partial^\alpha f(a) := \partial^{(\alpha_1,...,\alpha_n)} f(a) := \frac{\partial^k f(a)}{\partial x_1^{\alpha_1}...\partial x_n^{\alpha_n}}.$$

Si definimos $|\alpha| = \alpha_1 + ... + \alpha_n$, entonces $D^kf(x)$ puede escribirse

$$D^kf(x)(h) = \sum_{|\alpha|=k} \frac{k!}{\alpha_1!...\alpha_n!} \partial^\alpha f(x) h_1^{\alpha_1}...h_n^{\alpha_n}.$$

De esta expresión se sigue que

$$|D^kf(x)(h)| \leq \left(n^k \sup_{|\alpha|=k} |\partial^\alpha f(x)| \right) \|h\|^k.$$

Si f tiene derivadas parciales de orden k continuas en $a \in U$, entonces dichas derivadas estarán acotadas en un entorno de a, y por consiguiente se deduce que existen $M, \delta > 0$ tal que $\|D^kf(x)(h)\| \leq M\|h\|^k$ para todos $x \in B(a,\delta)$, $h \in \mathbb{R}^n$. Esta propiedad será usada en la demostración del teorema de Taylor.

Definición 12.1. Sean $U \subseteq \mathbb{R}^n$ abierto, $f : U \to \mathbb{R}$ una función de clase C^p. Se define el polinomio de Taylor de orden m de f en $a \in U$ como

$$P_{f,a}^k(h) = f(a) + Df(a)(h) + \frac{1}{2!}D^2f(a)(h) + \frac{1}{3!}D^3f(a)(h) + ... + \frac{1}{m!}D^mf(a)(h).$$

Ejemplo 12.1. Calcular el polinomio de Taylor de orden 3 de la función $f(x,y) = x\cos y$ en el punto $a = (1,0)$.

A continuación vemos cómo la versión general del teorema de Taylor puede deducirse fácilmente de la versión ya conocida para funciones de variable real.

Teorema 12.1 (fórmula de Taylor). *Sean U abierto de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ una función $m+1$ veces diferenciable en U, con $m \geq 1$, $a \in U$, $h \in \mathbb{R}^n$ con $[a, a+h] \subset U$. Entonces existe $\xi = \xi_h \in [a, a+h]$ tal que*

$$\begin{aligned} f(a+h) = \quad & f(a) + Df(a)(h) + \tfrac{1}{2!}D^2f(a)(h) + ... + \tfrac{1}{m!}D^m f(a)(h) + \\ & + \tfrac{1}{(m+1)!}D^{m+1}f(\xi_h)(h) = \\ & = P^m_{f,a}(h) + \tfrac{1}{(m+1)!}D^{m+1}f(\xi_h)(h). \end{aligned}$$

Por otra parte, si $k \geq 2$, suponiendo únicamente que f es $k-1$ veces diferenciable en el abierto U, y que $D^{k-1}f$ es diferenciable en un punto $x \in U$, se tiene que

$$\lim_{h\to 0} \frac{f(x+h) - P^k_{f,x}(h)}{\|h\|^k} = 0.$$

Demostración. Definamos

$$g(t) = f(a+th),$$

para $t \in [0,1]$. Por el teorema de Taylor para funciones de una variable real, existe $\zeta \in [0,1]$ tal que

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

Nótese que $g(1) = f(a+h)$, y $g(0) = f(a)$. Además, es fácil comprobar (ver el ejercicio 12.4) que

$$g^{(k)}(t) = D^k f(a+th)(h).$$

Por tanto, sustituyendo,

$$f(a+h) = f(a) + \sum_{k=1}^{m} \frac{1}{k!} D^k f(a)(h) + \frac{1}{(m+1)!} D^{m+1} f(a+\zeta h)(h),$$

lo que prueba la primera parte del enunciado si ponemos $\xi_h = a + \zeta h$.

Para ver la segunda parte del enunciado, para cada dirección $v \in \mathbb{R}^n$, $\|v\| = 1$, y cada $t \in (0,r)$, donde r es tal que $B(x,r) \subseteq U$, aplicamos el teorema del valor medio de Cauchy para encontrar un número $\alpha_{t,v} \in (0,t)$ tal que

$$\begin{aligned} & \frac{f(x+tv) - f(x) - tDf(x)(v) - ... - \frac{1}{k!}t^k D^k f(x)(v)}{t^k} = \\ & \frac{Df(x+\alpha_{t,v}v)(v) - Df(x)(v) - \alpha_{t,v}D^2 f(x)(v) - ... - \frac{1}{(k-1)!}\alpha_{t,v}^{k-1} D^k f(x)(v)}{k\alpha_{t,v}^{k-1}}. \end{aligned}$$

Volviendo a aplicar el teorema del valor medio de Cauchy obtenemos un número $\beta_{t,v} \in (0, \alpha_{t,v})$ tal que

$$\begin{aligned} & \frac{Df(x+\alpha_{t,v}v)(v) - Df(x)(v) - \alpha_{t,v}D^2 f(x)(v) - ... - \frac{1}{(k-1)!}\alpha_{t,v}^{k-1} D^k f(x)(v)}{k\alpha_{t,v}^{k-1}} = \\ & \frac{D^2 f(x+\beta_{t,v}v)(v) - D^2 f(x)(v) - \beta_{t,v}D^3 f(x)(v) - ... - \frac{1}{(k-2)!}\beta_{t,v}^{k-2} D^k f(x)(v)}{k(k-1)\beta_{t,v}^{k-2}}. \end{aligned}$$

Aplicando reiteradamente el teorema del valor medio de Cauchy hasta $k-3$ veces más, obtenemos un número $s_{t,v}\in(0,t)$ tal que

$$\frac{f(x+tv)-f(x)-tDf(x)(v)-...-\frac{1}{k!}t^kD^kf(x)(v)}{t^k}= \tag{12.1}$$
$$\frac{D^{k-1}f(x+s_{t,v}v)(v)-D^{k-1}f(x)(v)-s_{t,v}D^kf(x)(v)}{k!s_{t,v}}. \tag{12.2}$$

Ahora bien, puesto que $D^{k-1}f$ es diferenciable en x, se tiene que

$$\lim_{s\to 0^+}\frac{D^{k-1}f(x+sv)(v)-D^{k-1}f(x)(v)-sD^kf(x)(v)}{k!s}=0,$$

uniformemente en $v\in\mathbb{S}^{n-1}$ (la esfera unidad de $\mathbb{R}^n$). Es decir, dado $\varepsilon>0$ existe $\delta>0$ tal que si $0<s<\delta$ entonces

$$\left|\frac{D^{k-1}f(x+sv)(v)-D^{k-1}f(x)(v)-sD^kf(x)(v)}{k!s}\right|\leq\varepsilon$$

para todo $v\in\mathbb{S}^n$. Luego, si $0<t<\delta$, como $s_{t,v}\in(0,t)\subset(0,\delta)$ para todo $v\in\mathbb{S}^n$, deducimos de (12.1) que si $0<t<\delta$ entonces

$$\left|\frac{f(x+tv)-f(x)-tDf(x)(v)-...-\frac{1}{k!}t^kD^kf(x)(v)}{t^k}\right|\leq\varepsilon$$

para todo $v\in\mathbb{S}^n$. Esto significa que

$$\lim_{t\to 0^+}\frac{f(x+tv)-f(x)-tDf(x)(v)-...-\frac{1}{k!}t^kD^kf(x)(v)}{t^k}=0$$

uniformemente en $v\in\mathbb{S}^{n-1}$, lo que equivale a decir que

$$\lim_{h\to 0}\frac{f(x+h)-f(x)-Df(x)(h)-...-\frac{1}{k!}D^kf(x)(h)}{|h|^k}=0,$$

como queríamos comprobar. □

De manera análoga, usando la versión del teorema de Taylor con resto en forma integral para funciones de una variable, se demuestra el siguiente

Teorema 12.2 (fórmula de Taylor con resto integral)**.** *Sean U abierto de $\mathbb{R}^n$, $f:U\to\mathbb{R}$ una función de clase m en U, con $m\geq 1$, $a\in U$, $h\in\mathbb{R}^n$ con $[a,a+h]\subset U$. Entonces*

$$f(a+h)-\left(f(a)+Df(a)(h)+\frac{1}{2!}D^2f(a)(h)+...+\frac{1}{m!}D^mf(a)(h)\right)=$$
$$\int_0^1\frac{(1-t)^{m-1}}{(m-1)!}\left(D^mf(a+th)(h)-D^mf(a)(h)\right)dt.$$

Observación 12.3. A la cantidad

$$R^m_{f,a}(h) := f(a+h) - P^m_{f,a}(h)$$

se le suele llamar resto de Taylor. Así el teorema de Taylor nos dice que

$$\lim_{h\to 0} \frac{R^m_{f,a}(h)}{\|h\|^m} = 0.$$

Veamos ahora que el polinomio de Taylor $P^m_{f,a}$ es el único polinomio de grado m que tiene esta propiedad.

Para ello usaremos el siguiente lema.

Lema 12.4. *Sea $Q : \mathbb{R}^n \to \mathbb{R}$ un polinomio de grado menor o igual que m con la propiedad de que*

$$\lim_{h\to 0} \frac{Q(h)}{\|h\|^m} = 0.$$

Entonces $Q = 0$.

Demostración. Escribamos $Q = Q_0 + Q_1 + ... + Q_k$, donde Q_0 es una constante, Q_j es un polinomio homogéneo de orden j, y $k \leq m$. Tomando h de la forma $h = tv$, donde $\|v\| = 1$, y haciendo $t \to 0$, tenemos que

$$0 = \lim_{h\to 0} \frac{Q(h)}{\|h\|^m} = \lim_{t\to 0} \frac{Q(tv)}{\|tv\|^m} = \lim_{t\to 0} \frac{Q_0(tv) + Q_1(tv) + ... + Q_k(tv)}{t^m\|v\|} =$$
$$\lim_{t\to 0} \frac{Q_0(tv) + tQ_1(v) + ... + t^k Q_k(v)}{t^m} = \lim_{t\to 0} \frac{a_0 + a_1 t + ... + a_k t^k}{t^m},$$

donde $a_j = Q_j(v)$, $j = 0, 1, ..., k$. Ahora bien, el límite

$$\ell = \lim_{t\to 0} \frac{a_0 + a_1 t + ... + a_k t^k}{t^m}$$

nunca es igual a cero cuando $k \leq m$, salvo en el caso de que $a_0 = a_1 = ... = a_k = 0$. Esto muestra que, para todo $v \in \mathbb{R}^n$ con $\|v\| = 1$ se tiene $Q_j(v) = 0$ para todo $j = 0, 1, ..., k$, y como Q_j es j-homogéneo, también $Q_j(x) = 0$ para todo $x \in \mathbb{R}^n$, $j = 0, 1, ..., k$. Así $Q = 0$. □

Proposición 12.5. *Sean U un abierto de $\mathbb{R}^n$, f una función m veces diferenciable en un punto $a \in U$, y $P : \mathbb{R}^n \to \mathbb{R}$ un polinomio de grado menor o igual que m. Entonces, P es el polinomio de Taylor de orden m de f en a si y sólo si*

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

Demostración. ($\Rightarrow$): es consecuencia del Teorema 12.1.

($\Leftarrow$): Como tanto P como $P^m_{f,a}$ tienen la propiedad del enunciado, es inmediato que su diferencia, $Q = P - P^m_{f,a}$, satisface que

$$\lim_{h\to 0} \frac{P(h) - P^m_{f,a}(h)}{\|h\|^m} = \lim_{x\to a} \frac{f(a+h) - P^m(h)}{\|h\|^m} - \lim_{x\to a} \frac{f(a+h) - P(h)}{\|h\|^m} = 0,$$

luego, aplicando el lema precedente, concluimos que $Q = 0$. □

Supongamos que $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función de clase C^∞. En cada punto $a \in U$ podemos formar una serie infinita con las derivadas de f,

$$\sum_{k=0}^{\infty} \frac{1}{k!} D^k f(a)(x-a)$$

(donde $D^0 = f$). A esta serie se le llama la *serie de Taylor* de f en el punto a. Supongamos también que esta serie es convergente para x en un entorno de a. Es natural preguntarse si se tendrá

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} D^k f(a)(x-a)$$

al menos para todo x en un entorno suficientemente pequeño de a. Desgraciadamente esto no es siempre así, como prueba el ejemplo siguiente.

Ejemplo 12.2. Sea $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(t) = \begin{cases} e^{-1/t^2} & \text{si } t \neq 0, \\ 0 & \text{si } t = 0. \end{cases}.$$

No es difícil comprobar que $f \in C^\infty(\mathbb{R}, \mathbb{R})$ y que $D^k f(0) = 0$ para todo $k \in \mathbb{N}$. Por tanto, la serie de Taylor de f en 0 es constantemente cero, y sin embargo para todo $\delta > 0$, se tiene $f(\delta) \neq 0$, luego la serie de Taylor de f no converge a la función f en nigún entorno de 0.

También puede suceder, por supuesto, que la serie de Taylor ni siquiera sea convergente en ningéun entorno del punto

Definición 12.2. Se dice que $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ es real-analítica en U si para cada $a \in U$ existe $r > 0$ tal que

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} D^k f(a)(x-a)$$

para todo $x \in B(a, r)$. Es decir, f es real analítica si el resto de Taylor $R^k_{f,a}$ converge a cero cuando $k \to \infty$ en un entorno de cada $a \in U$.

Ejercicio 12.1. Probar que las funciones $f, g : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \cos(x + y)$, $g(x, y) = e^{-x^2-y^2}$ son real-analíticas.

En este curso no estudiaremos en profundidad las funciones real-analíticas; lo natural es hacerlo en el marco de la teoría de funciones de variable compleja. Baste con decir que los polinomios, las funciones trigonométricas, y las funciones exponencial y logaritmo son real-analíticas, y que puede demostrarse que la suma, producto, cociente (donde esté bien definico) y composición de funciones real-analíticas son real-analíticas.

El teorema de Taylor nos dice cómo aproximar localmente una función de clase C^p por un polinomio, y cuál es el polinomio con el que mejor es esta aproximación local. Sin embargo, cuando se pretende pasar de lo local a lo global y obtener resultados de aproximación uniforme de funciones por polinomios, ejemplos como el 12.2 muestran que este enfoque tiene sus limitaciones. A continuación estudiamos algunos resultados que nos garantizan que es posible aproximar uniformemente funciones continuas por funciones más regulares (como mínimo de clase C^∞).

En realidad uno de estos resultados ya lo conocemos: el teorema de Stone-Weierstrass visto en el Capítulo 7 nos dice que, dados cualquier conjunto compacto K de $\mathbb{R}^n$, una función continua $f : K \to \mathbb{R}$, y $\varepsilon > 0$, existe un polinomio $p : \mathbb{R}^n \to \mathbb{R}$ tal que

$$|f(x) - p(x)| \leq \varepsilon$$

para todo $x \in K$. Como todo polinomio es una función de clase C^∞ (de hecho también real-analítica), se sigue que el espacio de las funciones que tienen una extensión de clase C^∞ a un abierto que contiene a K es denso en el espacio de las continuas sobre cualquier compacto K de $\mathbb{R}^n$.

En el resto de este capítulo analizaremos otro método de aproximación que nos permite alcanzar la misma conclusión en situaciones más generales en que el dominio de la función no es compacto, y que además es muy empleado en multitud de situaciones diversas del análisis, la topología y la geometría diferencial: las llamadas particiones diferenciables de la unidad.

Si $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ es una función, se define su soporte abierto como

$$\text{sop}(f) = \{x \in U : f(x) \neq 0\},$$

y su soporte cerrado, o soporte, como

$$\overline{\text{sop}}(f) = \overline{\{x \in U : f(x) \neq 0\}}.$$

Por ejemplo, el problema 11.12 nos asegura que todo abierto de $\mathbb{R}^n$ es el soporte abierto de alguna función de $C^\infty(\mathbb{R}^n, \mathbb{R})$.

Si $(C_i)_{i\in I}$ es una familia de subconjuntos de $\mathbb{R}^n$, se dice que $(C_i)_{i\in I}$ es *localmente finita* si para cada $x \in \mathbb{R}^n$ existe $r = r_x > 0$ tal que $B(x, r)$ corta a lo sumo a una cantidad finita de subconjuntos de $(C_i)_{i\in I}$, es decir, el conjunto $\{i \in I : B(x, r) \cap C_i \neq \emptyset\}$ es finito.

Definición 12.3. Sea $\mathcal{U} = \{U_i\}_{i\in I}$ un recubrimiento por abiertos de un abierto U de $\mathbb{R}^n$. Diremos que $(\varphi_i)_{i\in I}$ es una partición de la unidad de clase C^∞ subordinada a $\mathcal{U}$ si

1. $\overline{\text{sop}}\varphi_i \subset U_i$ para todo $i \in I$;

2. $\{\overline{\text{sop}}\varphi_i\}_{i\in I}$ es localmente finita;

3. $\sum_{i\in I} \varphi_i(x) = 1$ para todo $x \in U$.

Si $\mathcal{V} = \{V_\alpha\}_{\alpha\in A}$ y $\mathcal{U} = \{U_i\}_{i\in I}$ son dos familias de subconjuntos de $\mathbb{R}^n$, diremos que $\mathcal{V}$ es un refinamiento de $\mathcal{U}$ si para todo $\alpha \in A$ existe $i \in I$ tal que $V_\alpha \subseteq U_i$.

Observación 12.6. Es útil observar que si $\mathcal{V} = \{V_\alpha\}_{\alpha\in A}$ es un recubrimiento por abiertos de U que refina otro recubrimiento por abiertos $\mathcal{U} = \{U_i\}_{i\in I}$ y $\mathcal{V}$ tiene subordinada una partición de la unidad $\{\psi_\alpha\}_{\alpha\in A}$ de clase C^∞, entonces también existe otra partición C^∞ de la unidad $\{\varphi_i\}_{i\in I}$ subordinada a $\mathcal{U}$. En efecto, si para cada $\alpha \in A$ escogemos $i(\alpha) \in I$ tal que $V_\alpha \subseteq U_{i(\alpha)}$, y consideramos la función $f : A \to I$, $f(\alpha) = i(\alpha)$, podemos poner

$$\varphi_i(x) = \sum_{\alpha\in f^{-1}(i)} \psi_\alpha(x).$$

Estas funciones están bien definidas y son de clase C^∞, puesto que la suma que define cada φ_i es localmente finita: dado $x \in U$, como $\{\overline{\text{sop}}\psi_\alpha\}_{\alpha\in A}$ es localmente finito, existe $r > 0$ tal que $\{\alpha \in A : B(x, r) \cap \overline{\text{sop}}\psi_\alpha \neq \emptyset\}$ es finito, y por tanto todas las ψ_α salvo una cantidad finita de

ellas se anulan en la bola $B(x,r)$. Además, si $\{\alpha \in A : B(x,r) \cap \overline{\text{sop}}\psi_\alpha \neq \emptyset\} = \{\alpha_1, ..., \alpha_k\}$, entonces
$$\{i \in I : B(x,r) \cap \overline{\text{sop}}\varphi_i \neq \emptyset\} \subseteq \{f(\alpha_1), ..., f(\alpha_k)\}$$
y en particular se ve que $\{\overline{\text{sop}}\varphi_i\}_{i\in I}$ es localmente finita. Por último, como los conjuntos $\{f^{-1}(i)\}_{i\in I}$ forman una partición de A, es obvio que
$$\sum_{i\in I} \varphi_i = \sum_{i\in I} \sum_{\alpha \in f^{-1}(i)} \psi_\alpha = \sum_{\alpha\in A} \psi_\alpha = 1,$$
y también que $\overline{\text{sop}}\varphi_i \subset U_i$ para cada $i \in I$.

Proposición 12.7. *Sea U un abierto de $\mathbb{R}^n$, y $\mathcal{U} = \{U_i\}_{i\in I}$ un recubrimiento por abiertos de U. Entonces existe $\mathcal{V} = \{V_k\}_{k\in\mathbb{N}}$ un recubrimiento por abiertos de U tal que*

1. $\widetilde{\mathcal{V}} := \{\overline{V_k}\}_{k\in\mathbb{N}}$ *refina a $\mathcal{U}$ y es localmente finito;*
2. $\overline{V_k}$ *es compacto para todo $k \in \mathbb{N}$.*

Demostración. Como cada U_i es abierto, para cada $x \in U_i$ existe $r_x^i > 0$ tal que $\overline{B}(x, r_x^i) \subset U_i$. Podemos escribir por tanto
$$U = \bigcup_{i\in I} U_i = \bigcup_{i\in I} \bigcup_{x\in U_i} B(x, r_x^i),$$
y como U es separable, existen sucesiones $\{x_k\}_{k\in\mathbb{N}} \subset U$ e $\{i(k)\}_{k\in\mathbb{N}} \subset I$ tales que $x_k \in U_{i(k)}$ y
$$U = \bigcup_{k=1}^{\infty} B(x_k, r_k^{i(k)}),$$
donde se denota $r_k = r_{x_k}^{i(k)}$. Así se tiene que $\{B(x_k, r_k)\}_{k\in\mathbb{N}}$ es un refinamiento de $\mathcal{U}$ que recubre U.

Ahora pongamos $V_1 = B(x_1, r_1)$, y para $k \geq 2$ definamos
$$V_k = B(x_k, r_k) \setminus \bigcup_{j=1}^{k-1} \overline{B}(x_j, \lambda_k r_j),$$
donde (λ_k) es una sucesión de números positivos tales que $\lambda_k < \lambda_{k+1}$ para todo k y $\lim_{k\to\infty} \lambda_k = 1$. Es claro que los V_k son abiertos, y que
$$\bigcup_{k=1}^{\infty} V_k = \bigcup_{k=1}^{\infty} B(x_k, r_k) = U,$$
y también es obvio que $\overline{V_k}$ es compacto para todo $k \in \mathbb{N}$. Además, puesto que para cada k existe $i = i(k)$ tal que $\overline{V_k} \subseteq \overline{B}(x_k, r_k) \subset U_i$, se tiene que $\{\overline{V_k} : k \in \mathbb{N}\}$ refina a $\mathcal{U}$. Veamos por último que la familia $\{\overline{V_k} : k \in \mathbb{N}\}$ es localmente finito. Sea $x \in U$, y escojamos $m = m_x \in \mathbb{N}$ tal que $x \in B(x_m, r_m)$. Sea $s = r_m - d(x, x_m) > 0$. Entonces, como $\lambda_k \nearrow 1$, existe $k_0 \in \mathbb{N}$ tal que $k_0 > m$ y $\lambda_k r_m > r_m - s/2$ para todo $k \geq k_0$. Luego, si $k \geq k_0$ se tiene que $B(x, s/2) \subset B(x_m, \lambda_k r_m)$, y por tanto
$$B(x, s/2) \cap \overline{V_k} = B(x, s/2) \cap \left(\overline{B}(x_k, r_k) \setminus \bigcup_{j=1}^{k-1} B(x_j, \lambda_k r_j) \right) = \emptyset,$$
lo que prueba que $B(x, s/2)$ corta a una cantidad finita de miembros de $\{\overline{V_k} : k \in \mathbb{N}\}$. $\square$

Ya podemos demostrar el teorema de existencia de particiones de la unidad de clase C^∞ subordinadas a cualquier recubrimiento por abiertos de $\mathbb{R}^n$.

Teorema 12.8. *Sea U un abierto de $\mathbb{R}^n$, y sea $\mathcal{U} = \{U_i\}_{i\in I}$ un recubrimiento por abiertos de U. Entonces existe una partición de la unidad de clase C^∞ subordinada a $\mathcal{U}$.*

Demostración. Por la Proposición anterior, existe $\mathcal{V} = \{V_\alpha\}_{\alpha\in A}$ recubrimiento por abiertos de U tal que $\{\overline{V_\alpha} : \alpha \in A\}$ refina a $\mathcal{U}$. Por la observación 12.6, basta demostrar que existe una partición de la unidad de clase C^∞ subordinada a $\mathcal{V}$. Por el problema 11.12, para cada $\alpha \in A$ existe una función $f_\alpha \in C^\infty(\mathbb{R}^n, \mathbb{R})$ tal que $f_\alpha^{-1}(0) = \mathbb{R}^n \setminus V_\alpha$, y en particular $\mathrm{sop}(f_\alpha) = V_\alpha$. Luego

$$\overline{\mathrm{sop}}(f_\alpha) = \overline{V_\alpha}.$$

Para cada $\alpha \in A$ definamos ahora $\psi_\alpha : U \to \mathbb{R}$ por

$$\psi_\alpha(x) = \frac{f_\alpha(x)}{\sum_{\beta\in A} f_\beta(x)}.$$

Es inmediato ver que el denominador no se anula en U y la suma es localmente finita, por lo que ψ_α está bien definida y es de clase C^∞ en U. Como $\overline{\mathrm{sop}}(\psi_\alpha) = \overline{\mathrm{sop}}(f_\alpha) = \overline{V_\alpha}$, es claro también que $\{\overline{\mathrm{sop}}(\psi_\alpha) : \alpha \in A\}$ es localmente finita. Por último, es evidente que

$$\sum_{\alpha\in A} \psi_\alpha(x) = 1$$

para todo $x \in U$. Por tanto $\{\psi_\alpha : \alpha \in A\}$ es una partición de la unidad de clase $C^\infty(U)$ subordinada a $\mathcal{V}$. $\square$

Las particiones de la unidad suelen utilizarse para pegar funciones diferenciables definidas localmente y obtener así funciones diferenciables definidas globalmente y que conservan ciertas propiedades interesantes, por ejemplo de aproximación. Para concluir, veamos un ejemplo típico de uso de las particiones de la unidad.

Teorema 12.9. *Sean U un abierto de $\mathbb{R}^n$, $f : U \to \mathbb{R}^m$ una función continua, y $\varepsilon : U \to (0, +\infty)$ una función continua estrictamente positiva. Entonces existe $g \in C^\infty(U, \mathbb{R}^m)$ tal que*

$$\|g(x) - f(x)\| \leq \varepsilon(x)$$

para todo $x \in U$.

Demostración. Como f y ε son funciones continuas, para cada $x \in U$ existe $\delta_x > 0$ tal que

$$\|f(y) - f(x)\| \leq \frac{\varepsilon(x)}{2} \leq \varepsilon(y) \ \text{ si } y \in B(x, \delta_x).$$

Sea $\{\varphi_x\}_{x\in U}$ una partición de la unidad de clase C^∞ subordinanda al recubrimiento $\mathcal{U} := \{B(x, \delta_x)\}_{x\in U}$. Definamos entonces $g : U \to \mathbb{R}^m$ por

$$g(y) = \sum_{x\in U} f(x)\varphi_x(y).$$

Como la suma es localmente finita, g está bien definida y es de clase $C^\infty(U)$. Además se tiene que, si fijamos $y \in U$ cualquiera,

$$\|g(y) - f(y)\| = \|\sum_{x\in U} f(x)\varphi_x(y) - f(y)\| =$$
$$\|\sum_{x\in U} f(x)\varphi_x(y) - \sum_{x\in U} f(y)\varphi_x(y)\| = \|\sum_{x\in U}(f(x) - f(y))\varphi_x(y)\|.$$

En la suma dentro del último término de estas igualdades, todos los sumandos son nulos salvo una cantidad finita de ellos, precisamente los correspondientes a los x que satisfacen $y \in \text{sop}\varphi_x$, en cuyo caso, puesto que $\text{sop}\varphi_x \subseteq B(x, \delta_x)$, se cumple que

$$\|f(x) - f(y)\| \leq \varepsilon(y).$$

Por tanto deducimos que

$$\|g(y) - f(y)\| \leq \sum_{x\in U,\, z\in \text{sop}(\varphi_x)} \|(f(x) - f(y))\|\varphi_x(y) \leq$$
$$\sum_{x\in U,\, z\in \text{sop}(\varphi_x)} \varepsilon(y)\varphi_x(y) \leq \sum_{x\in U} \varepsilon(y)\varphi_x(y) = \varepsilon(y)\sum_{x\in U}\varphi_x(y) = \varepsilon(y).$$

□

12.1. Problemas

Problema 12.1. Calcúlese el Polinomio de Taylor de grado menor o igual que 3 para las funciones:

$$f(x,y) = e^{(x-1)^2}\cos y \text{ en } (1,0), \quad g(x,y) = x\sin y + y\sin x \text{ en } (0,0),$$
$$h(x,y) = xy^2 \text{ en } (0,0), \qquad \varphi(x,y) = \log(x+y) \text{ en } (1,1),$$
$$\psi(x,y) = \frac{1}{x^2+y^2+1} \text{ en } (0,0), \qquad \Phi(x,y) = \tfrac{y^2}{x^3} \text{ en } (1,-1).$$

Problema 12.2. Escríbase el polinomio $x^3 + y^2(1+x)$ en potencias de $(x-1)$ e $(y-2)$.

Problema 12.3. Demuéstrese que $f(x,y) = y^x$ es diferenciable para $y > 0$. Encuéntrese un polinomio $P(x,y)$ en dos variables, de grado menor o igual que 3, tal que

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

Problema 12.4. Si $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ es una aplicación afín (es decir, suma de una lineal más una constante), y $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^p$ es k veces diferenciable, probar que

$$D^k(f\circ\varphi)(a)(h_1, ..., h_k) = D^k f(\varphi(a))(D\varphi(a)(h_1), ..., D\varphi(a)(h_k)).$$

En el caso particular en que $\varphi : \mathbb{R} \to \mathbb{R}^m$, $\varphi(t) = a + th$, concluir que, si $g(t) = f(a+th)$, entonces
$$g^{(k)}(t) = D^k f(a+th)(h).$$

Problema 12.5. Usar el teorema de Taylor para probar el teorema binomial:

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

Indicación: Considerar $f(x) = (a+x)^n$.

Problema 12.6. Para cada $A \in \mathcal{L}^k(\mathbb{R}^n, \mathbb{R})$ defínase

$$\|A\| = \text{ínf}\{M \geq 0 : |A(h_1, ..., h_k)| \leq M\|h_1\|...\|h_k\|\}.$$

Probar que esta expresión define una norma en el espacio $\mathcal{L}^k(\mathbb{R}^n, \mathbb{R})$.

Problema 12.7. Sean U un abierto de $\mathbb{R}^n$ y $f : U \to \mathbb{R}$ una función de clase C^∞ tal que existe $M > 0$ de modo que para cada $x \in U$ se tiene

$$\|D^k f(x)\| \leq M^k$$

(donde $\|\cdot\|$ es la norma en $\mathcal{L}^k(\mathbb{R}^n, \mathbb{R})$ definida en el problema anterior). Demostrar que f es analítica en U.

Capítulo 13

Extremos locales y puntos críticos

Como ya es conocido para el alumno, para encontrar los máximos y mínimos locales de una función real definida sobre un intervalo de $\mathbb{R}$, la técnica más socorrida es hallar los puntos donde su derivada se anula (los puntos críticos), y ver después cuál es el signo de la derivada segunda sobre esos puntos críticos. Si se quiere después hallar los máximos y mínimos absolutos de la función en el intervalo, se comparan los valores de f en los puntos críticos con los que toma en los extremos del intervalo. Como veremos en este capítulo, el mismo método, esencialmente, sirve para funciones reales de varias variables, con algunas diferencias importantes. Primera, que la segunda derivada ya no es un número, sino una forma bilineal simétrica. En este caso el *signo* de la derivada será sustituido por ser definida positiva o definida negativa. Segundo, que la frontera del conjunto abierto donde esté definida la función ya no va a ser un conjunto tan simple como un par de puntos: en general habrá infinitos puntos en dicha frontera, y la tarea de comparar uno por uno los valores de f en estos puntos es obviamente impracticable. Será por tanto necesario desarrollar nuevas técnicas que permitan hallar los máximos y mínimos de una función diferenciable restringida a la frontera de un abierto. Esto, dicho así en toda generalidad, es un objetivo excesivamente ambicioso, pero si la frontera del abierto es un conjunto suficientemente regular (técnicamente, unión finita de variedades diferenciables) entonces el problema tendrá solución satisfactoria, aunque habrá que esperar al Capítulo 14 para ver completado este esquema.

Definición 13.1. Se dice que $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ tiene un máximo local (o relativo) en un punto $a \in A$ si existe $r > 0$ tal que

$$f(x) \leq f(a) \text{ para todo } x \in A \cap B(x, r).$$

Se dice que f tiene un mínimo local en a si existe $r > 0$ tal que

$$f(a) \leq f(x) \text{ para todo } x \in A \cap B(x, r).$$

En cualquiera de los casos se dirá que f tiene un extremo local en a, y que es estricto si la desigualdad es estricta cuando $x \neq a$.

Definición 13.2. Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable en $a \in U$. Se dice que a es un *punto crítico* de f si $Df(a) = 0$.

Como ocurre con las funciones de una variable, todos los extremos locales de las funciones diferenciables son puntos críticos.

Proposición 13.1. *Sean U abierto de $\mathbb{R}^n$, y $f : U \to \mathbb{R}$ diferenciable en U. Supongamos que f tiene un extremo local en $a \in U$. Entonces*

$$Df(a) = 0,$$

es decir a es un punto crítico de f.

Demostración. Sea $v \in \mathbb{R}^n$ cualquiera con $\|v\| = 1$. Definamos $g_v : (-r, r) \to \mathbb{R}$ por

$$g_v(t) = f(a + tv),$$

donde $r > 0$ es tal que $B(a, r) \subseteq U$. Como f tiene un extremo local en a, es evidente que g_v también tiene un extremo local (del mismo tipo) en $t = 0$. Entonces, por el resultado ya conocido para funciones de variable real, se tiene que

$$0 = g'(0) = Df(a)(v).$$

Esto prueba que $Df(a)(v) = 0$ para todo v con $\|v\| = 1$, y por tanto, como $Df(a)$ es lineal, $Df(a) = 0$. □

El recíproco no es cierto, es decir, no todo punto crítico de una función diferenciable es un exremo local. Por ejemplo, la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$f(x, y) = xy$$

tiene un punto crítico en $(0, 0)$, que no es ni máximo ni mínimo (téngase en cuenta que $f(t, t) = t^2$, y $f(t, -t) = -t^2$). La gráfica de f es la típica silla de montar, y de $(0, 0)$ se dice que es un *punto de silla* de f.

Para determinar si un punto crítico es un máximo o un mínimo local, o ninguna de las dos cosas, es suficiente, en muchos casos un análisis de la matriz de la segunda derivada.

Definición 13.3. Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función de clase C^2. Se dice llama hessiana de f en a a la segunda derivada de f en a, y matriz hessiana a la matriz de esta forma cuadrática, es decir,

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(a) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(a) \\ \vdots & \dots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(a) \end{pmatrix}.$$

Recordemos que, si Q es una forma cuadrática en $\mathbb{R}^n$, se dice que:

1. Q es definida positiva si $Q(h) > 0$ para todo $h \neq 0$;
2. Q es semidefinida positiva si $Q(h) \geq 0$ para todo $h \in \mathbb{R}^n$;
3. Q es definida negativa si $Q(h) < 0$ para todo $h \neq 0$;
4. Q es semidefinida negativa si $Q(h) \leq 0$ para todo $h \in \mathbb{R}^n$;
5. Q es indefinida si no es ni semidefinida positiva ni semidefinida negativa, es decir, si existen $h_1, h_2 \in \mathbb{R}^n$ tales que $Q(h_1) < 0 < Q(h_2)$.

El teorema de Taylor nos permite dar condiciones suficientes sobre la segunda derivada de una función en un punto crítico para asegurar que se trata de un mínimo o un máximo local, o decidir que no es ni lo uno ni lo otro.

Teorema 13.2. *Sean U un abierto de $\mathbb{R}^n$, y $f \in C^2(U, \mathbb{R})$. Supongamos que $Df(a) = 0$, es decir, $a \in U$ es un punto crítico de f. Entonces:*

1. *Si $D^2 f(a)$ es definida positiva, f tiene un mínimo local estricto en a.*
2. *Si $D^2 f(a)$ es definida negativa, f tiene un máximo local estricto en a.*
3. *Si $D^2 f(a)$ es indefinida, f tiene un punto de silla en a (es decir, no tiene mi máximo ni mínimo local en este punto crítico).*

   Además, se tiene un recíproco parcial (no total; ver los ejemplos a continuación) de lo anterior:
4. *Si f tiene un mínimo local en a entonces $D^2 f(a)$ es semidefinida positiva.*
5. *Si f tiene un máximo local en a entonces $D^2 f(a)$ es semidefinida negativa.*

Demostración. (1) Por el teorema de Taylor, y puesto que $Df(a) = 0$, sabemos que

$$f(a+h) = f(a) + \frac{1}{2} D^2 f(a)(h) + R(h),$$

donde $\lim_{h\to 0} R(h)/\|h\|^2 = 0$. Sea

$$m := \inf\{D^2 f(a)(v) : \|v\| = 1\}.$$

Como $D^2 f(a)(\cdot)$ es continua y la esfera unidad $S = \{h : \|h\| = 1\}$ es compacta, este ínfimo es en realidad un mínimo, y puesto que $D^2 f(a)$ es definida positiva, debe ser $m > 0$. Así, para todo $h \in \mathbb{R}^n$, $h \neq 0$, tenemos

$$D^2 f(a)(\frac{h}{\|h\|}) \geq m > 0,$$

luego, puesto que $D^2 f(a)$ es 2-homogénea,

$$D^2 f(a)(h) \geq m\|h\|^2$$

para todo $h \in \mathbb{R}^n$. Sea $\varepsilon \in (0, m)$. Como $\lim_{h\to 0} R(h)/\|h\|^2 = 0$, existe $\delta > 0$ tal que si $\|h\| \leq \delta$ entonces $|R(h)| \leq \varepsilon\|h\|^2$, y por tanto

$$|f(a+h) - f(a) - D^2 f(a)(h)| \leq \varepsilon\|h\|^2,$$

luego

$$f(a+h) - f(a) \geq D^2 f(a)(h) - \varepsilon\|h\|^2 \geq (m - \varepsilon)\|h\|^2$$

para todo h con $\|h\| \leq \delta$. Esto significa que

$$f(x) \geq f(a) \;\text{ si }\; x \in B(a, \delta),$$

y que la desigualdad es estricta cuando $x \neq a$, es decir, f tiene un mínimo local estricto en a.
(2): es análogo (o bien aplicar (1) a la función $-f$).
(4): Sea $v \in \mathbb{R}^n$ con $\|v\| = 1$. Por el teorema de Taylor, y teniendo en cuenta que f tiene un mínimo local en a, podemos escribir

$$0 \leq f(a+tv) - f(a) = \frac{1}{2} D^2 f(a)(tv) + R(tv)$$

si t es suficientemente pequeño, donde $\lim_{t\to 0} R(tv)/\|tv\|^2 = 0$. Entonces, tomando límites,

$$0 \le \lim_{t\to 0} \frac{f(a+tv)-f(a)}{t^2} = \lim_{t\to 0} \frac{1}{t^2}\left(\frac{1}{2}D^2 f(a)(tv) + R(tv)\right) =$$
$$\lim_{t\to 0}\left(\frac{1}{2}D^2 f(a)(v) + \frac{R(tv)}{\|tv\|^2}\right) = \frac{1}{2}D^2 f(a)(v) + 0,$$

es decir, $D^2 f(a)(v) \ge 0$ para todo v con $\|v\| = 1$, lo que significa que f es semidefinida positiva.
(5): análogo (o considerar $-f$).
(3): es consecuencia inmediata de (4) y (5). □

A continuación recordamos algunos criterios para decidir si una forma cuadrática simétrica (por ejemplo $D^2 f(a)$) es definida positiva, definida negativa, o indefinida.

Proposición 13.3. *Sea Q una forma cuadrática simétrica en $\mathbb{R}^n$. Entonces*

1. *Q es definida positiva si y sólo si todos sus autovalores son estrictamente positivos;*
2. *Q es definida negativa si y sólo si todos sus autovalores son estrictamente negativos;*
3. *Q es indefinida si y sólo si tiene autovalores positivos y negativos (estrictamente);*
4. *Q es semidefinida positiva si y sólo si todos sus autovalores son mayores o iguales que cero;*
5. *Q es semidefinida negativa si y sólo si todos sus autovalores son menores o iguales que cero.*

Recordemos que los autovalores de Q puedan calcularse hallando las raíces de su polinomio característico,

$$P(\lambda) = \det(A - \lambda I),$$

donde A es la matriz de Q, y que todas las matrices correspondientes a formas cuadráticas simétricas tienen autovalores reales y son diagonalizables. Por tanto, para decidir si $D^2 f(a)$ es definida positiva o definida negativa, o indefinida, basta hallar sus autovalores y mirar el signo de éstos.

Si uno no desea calcular autovalores, el siguiente criterio puede ser igualmente útil, por lo menos para dimensiones bajas.

Proposición 13.4 (criterio de Silvester). *Sea Q una forma cuadrática simétrica en $\mathbb{R}^n$, con matriz $A = (a_{ij})$. Consideremos los menores angulares de A,*

$$A_j = \det \begin{pmatrix} a_{11} & ... & a_{1j} \\ \vdots & ... & \vdots \\ a_{j1} & ... & a_{jj} \end{pmatrix}.$$

Entonces:

1. *Q es definida positiva si y sólo si $A_j > 0$ para todo $j = 1, ..., n$;*
2. *Q es definida negativa si y sólo si $(-1)^j A_j > 0$ para todo $j = 1, ..., n$;*

3. *Si Q es semidefinida positiva entonces $A_j \geq 0$ para todo $j = 1, ..., n$;*

4. *Si Q es semidefinida negativa entonces $(-1)^j A_j \geq 0$ para todo $j = 1, ..., n$.*

Si algún A_j con j par es negativo, entonces Q es indefinida.

Ejemplo 13.1. Estudiar los puntos críticos de las siguientes funciones, tratando de determinar si corresponden a máximos locales, mínimos locales, o puntos de silla:

1. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = e^{xy}$;

2. $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x^2 + y^2 + z^2 + 4yz + y^2x - z^3$;

3. $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x^3 + y^3 + z^3 - 3(xy + xz + yz)$;

4. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^3 + y^3 - 3ax$, según los valores de $a \in \mathbb{R}$;

5. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2$;

6. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 - y^4$.

Cuando la función f a estudiar es convexa el problema de determinar sus extremos es mucho más fácil: todo punto crítico de una función convexa diferenciable es un mínimo absoluto.

Proposición 13.5. *Sean U un abierto convexo de $\mathbb{R}^n$, y $f : U \to \mathbb{R}$ convexa. Supongamos que f es diferenciable en $a \in U$, y que $Df(a) = 0$. Entonces f tiene un mínimo absoluto en a.*

Demostración. Supongamos lo contrario. Entonces existe $b \in U$ tal que $f(b) < f(a)$. Para todo $t \in [0, 1]$ tenemos

$$f(tb + (1-t)a) \leq tf(b) + (1-t)f(a) = f(a) + t(f(b) - f(a)).$$

Si definimos $g(t) = f(tb + (1-t)a)$, $t \in [0, 1]$, esto quiere decir que

$$g(t) \leq g(0) + t(f(b) - f(a))$$

para todo $t \in [0, 1]$. Por tanto

$$Df(a)(b-a) = g'(0) = \lim_{t \to 0^+} \frac{g(t) - g(0)}{t} \leq f(b) - f(a) < 0,$$

en contradicción con la hipótesis de que a es un punto crítico de f. □

Ejemplo 13.2. Sea $f : \mathbb{R}^3 \to \mathbb{R}$ definida por $f(x, y, z) = x^4 + y^6 + y^8$. Esta función es convexa (es suma de convexas), y $(0, 0, 0)$ es un punto crítico de f; por tanto f alcanza un mínimo absoluto en $(0, 0, 0)$ (nótese que $D^2 f(0, 0, 0) = 0$, y por tanto el criterio de la derivada segunda no nos da información suficiente en este caso).

Por otro lado, como sucede en el caso de funciones reales de variable real, una función $f \in C^2(U, \mathbb{R})$ es convexa si y sólo si $D^2 f(x)$ es semidefinida positiva para todo $x \in U$.

Proposición 13.6. *Sea U un abierto convexo de $\mathbb{R}^n$, y sea $f : U \to \mathbb{R}$ de clase C^2. Entonces, f es convexa si y sólo si $D^2 f(x)$ es semidefinida positiva para todo $x \in U$.*

Demostración. La función f es convexa si y sólo si para todo $x, y \in U$ la función $\varphi : [0,1] \to \mathbb{R}$ definida por

$$\varphi_{x,y}(t) = \varphi(t) = f(ty + (1-t)x)$$

es convexa. Por tanto, basta ver que $\varphi_{x,y}$ es convexa para todo $x, y \in U$ si y sólo si $D^2 f(a)$ es semidefinida positiva para todo $a \in U$. Como $\varphi_{x,y}$ es una función de variable real de clase C^2, esto equivale a decir que $\varphi''_{x,y}(t) \geq 0$ para todos $t \in [0,1]$, $x, y \in U$ si y sólo si $D^2 f(a)$ es definida positiva para todo $a \in U$. Pero

$$\varphi''_{x,y}(t) = D^2 f(ty + (1-t)x)(y-x)^2.$$

Por tanto, si $D^2 f(a)$ es semidefinida positiva para todo a, está claro que $\varphi''_{x,y}(t) \geq 0$ para todos $t \in [0,1]$, $x, y \in U$.

Recíprocamente, supongamos que todas estas funciones $\varphi_{x,y}$ tienen segunda derivada positiva. Si $a \in U$, como U es abierto podemos tomar $r > 0$ tal que $\overline{B}(a,r) \subset U$. Sea $v \in \mathbb{R}^n$ cualquiera con $\|v\| = 1$, y pongamos $x = a$, $y = a + rv$. Entonces

$$\varphi_{x,y}(t) = \varphi(t) = f(a + trv),$$

luego

$$0 \leq \varphi''(t) = D^2 f(a)(rv) = r^2 D^2 f(a)(v).$$

Esto prueba que $D^2 f(a)(v) \geq 0$ para todo $v \in \mathbb{R}^n$ con $\|v\| = 1$, y por tanto que $D^2 f(a)$ es semidefinida positiva. $\square$

Pasamos ahora a considerar brevemente la cuestión de cómo identificar los máximos y mínimos globales de una función en un conjunto no necesariamente abierto.

En general es necesario asegurarse primero de que existen los extremos globales que se buscan. En general no tienen por qué existir si el conjunto no es compacto o la función no es continua, como ya sabemos. Incluso si la función es continua y acotada, el ínfimo y el supremo de su rango podrían no alcanzarse, como en el caso de la función $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \arctan(x)$.

Por tanto, supondremos que nuestra función f está definida sobre un compacto K de $\mathbb{R}^n$ que es la adherencia de un conjunto abierto acotado U sobre el cual la función f es diferenciable, y que además f es continua en K. En estas condiciones sabemos que f alcanza un máximo y un mínimo absolutos sobre K. La cuestión que debemos intentar responder ahora es cómo calcularlos y localizar los puntos donde se alcanzan. Cada caso concreto puede ser diferente, pero podemos dar algunas reglas generales. Lo primero que podemos hacer es intentar hallar los puntos críticos de f en U, tarea que en general puede ser muy difícil en sí misma, ya que puede ser imposible resolver explícitamente el sistema de ecuaciones $Df(x) = 0$, aún suponiendo que tenga solución. Esto puede ocurrir perfectamente aún en el caso de las funciones más simples como los polinomios (es bien conocido que no hay fórmulas algebraicas para resolver ecuaciones polinómicas de grado igual o superior a cinco).

Seamos optimistas y supongamos que hemos sido capaces de resolver este sistema de ecuaciones $Df(x) = 0$, y supongamos para empezar que tiene un número finito de soluciones, $a_1, ..., a_k \in U$. Entre estos puntos críticos puede que se encuentre el máximo o el mínimo de f en K. Para decidir si esto es así, habrá que comparar los valores de f en los puntos a_i con los valores que toma f en la frontera $\partial K = \partial U$. El problema es que en dimensiones $n \geq 2$, ∂U consta siempre de una cantidad infinita de puntos, de modo que esto no puede hacerse uno por uno como en el caso de funciones de una variable definidas en intervalos. Pero puede que hayamos tenido suerte y de alguna manera podamos saber lo que valen

$$\alpha = \inf\{f(x) : x \in \partial U\} \text{ y } \beta = \sup\{f(x) : x \in \partial U\}.$$

En este caso, el mayor de los valores $\{\alpha, \beta, f(a_1), ..., f(a_k)\}$ corresponderá al máximo de f en K, y el menor de ellos al mínimo. El siguiente ejemplo ilustra esta situación.

Ejemplo 13.3. Hallar los máximos y mínimos de $f(x,y) = x^2 + x + y^3 - 3y$ en el conjunto $K = \{(x,y) \in \mathbb{R}^2 : x^2 + x + y^3 - 3y \leq 7\}$.

En este ejemplo $\alpha = \beta = 7$, y basta hallar los extremos relativos de f en el interior de K y comparar el valor de f en estos puntos con 7.

Un caso un poco más complicado que puede presentarse es que el conjunto de puntos críticos de f no sea finito. Parecería entonces que nuestra tarea está abocada al fracaso, pero en muchas situaciones sucede que puede identificarse una sucesión (finita o infinita) de conjuntos (infinitos) sobre los cuales f es constante. Es decir, pueden encontrarse $C_k \subset K$, $k \in \mathbb{N}$, tales que $Df(x) = 0$ si y sólo si $x \in \bigcup_{k=1}^{\infty} C_k$, y f es constante en C_k para cada k, pongamos $f(C_k) = \{c_k\}$. Entonces el mayor de los valores $\{\alpha, \beta\} \cup \{c_k : k \in \mathbb{N}\}$ corresponderá al máximo de f en K, y el menor de ellos al mínimo (¿por qué este conjunto, en general infinito, tiene un máximo y un mínimo?).

Ejemplo 13.4. Sea $F : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$f(x,y) = \begin{cases} e^{-1/(x^2+y^2)} \sin \frac{1}{x^2+y^2} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0). \end{cases},$$

y sea $K = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 5\}$. La función f es de clase C^∞ (por ser composición de la función $h(t) = e^{-1/t} \sin(1/t)$, $h(0) = 0$, que es de clase C^∞, con el polinomio $(x,y) \to x^2 + y^2$). Es inmediato ver que

$$Df(x,y) = (0,0) \iff x^2 + y^2 = \frac{1}{\frac{\pi}{4} + k\pi}, k \in \mathbb{N} \cup \{0\}, \text{ o } (x,y) = (0,0).$$

Así pues los conjuntos $C_k = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1/(\pi/4 + k\pi)\}$ están formados todos ellos por puntos críticos de f, y f es constante sobre cada C_k (donde vale $(-1)^k e^{-(\pi/4+k\pi)}\sqrt{2}/2$, $k = 0, 1, 2, ...$). El origen es también un punto crítico de f, y no hay más. Es fácil ver que f tiene máximos relativos sobre cada punto de C_k con k par, y mínimos relativos sobre cada punto de C_k con k impar, mientras que el origen es un punto de silla. Por otro lado, f es constante en ∂K (donde vale $e^{-1/5}\sin(1/5)$). De todo esto se concluye, comparando los valores

$$0, \ e^{-1/5}\sin(1/5), \ (-1)^k e^{-(\pi/4+k\pi)}\sqrt{2}/2 \ (k \in \mathbb{N} \cup \{0\}),$$

que f alcanza su máximo absoluto sobre K en cada punto de C_0, y su mínimo en cada punto de C_1. De hecho, puesto que

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

estos mínimos y máximos de f en K lo son también de f en todo $\mathbb{R}^2$.

Los ejemplos anteriores eran muy simples en cuanto al comportamiento de f en la frontera de K, porque f era constante en ∂K, y así su mínimo y máximo en la frontera eran iguales y se alcanzaban en cualquier punto de ésta. Supongamos ahora que no hemos tenido tanta suerte y que f no es constante en ∂K. ¿Qué podemos hacer para determinar los números

$$\alpha = \inf\{f(x) : x \in \partial U\} \ \text{ y } \ \beta = \sup\{f(x) : x \in \partial U\} \ ?$$

Una primera observación que puede hacerse es que a veces no es necesario calcular α y β, sino que puede bastar con acotarlos, si conseguimos una cota que nos permita concluir, cotejándola con otros valores de f en el interior de K, que el máximo y el mínimo de f en K no pueden alcanzarse en la frontera ∂K y deben por tanto hallarse en $U = \text{int}(K)$.

Ejemplo 13.5. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ la función del ejemplo anterior, y $K = \{(x,y) \in \mathbb{R}^2 : x^2 + 3y^2 \leq 25\}$. Ahora f ya no es constante en $\partial K = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 25\}$, pero, teniendo en cuenta que la función $t \mapsto e^{-1/t} \sin(1/t)$ es creciente para $t > 4/\pi$, podemos acotar α y β como sigue. Para todo $(x,y) \in \partial K$ tenemos que $x^2 + y^2 \leq x^2 + 3y^2 = 25$, y por tanto

$$e^{-1/(x^2+y^2)} \sin \frac{1}{x^2+y^2} \leq e^{-1/(x^2+3y^2)} \sin \frac{1}{x^2+3y^2} = e^{-1/25} \sin \frac{1}{25},$$

de donde se deduce que

$$\beta := \sup\{f(x,y) : (x,y) \in \partial K\} \leq e^{-1/25} \sin \frac{1}{25}.$$

Análogamente, teniendo en cuenta que si $(x,y) \in \partial K$ entonces $x^2 + y^2 \geq \frac{1}{3}(x^2 + 3y^2) = 25/3 > 8$, se deduce que

$$\alpha := \inf\{f(x,y) : (x,y) \in \partial K\} \geq e^{-1/8} \sin \frac{1}{8}.$$

Por tanto, puesto que

$$\beta \leq e^{-1/25} \sin \frac{1}{25} < f(C_0) = e^{-\pi/4} \frac{\sqrt{2}}{2},$$

y que

$$f(C_1) = -e^{-\pi/4 - \pi} \sin \frac{\sqrt{2}}{2} < 0 < e^{-1/8} \sin \frac{1}{8} \leq \alpha,$$

es obvio que f alcanza su máximo absoluto sobre K en cada uno de los puntos de C_0, y su mínimo absoluto sobre K en los puntos de C_1.

Supongamos ahora que no somos capaces de acotar α y β satisfactoriamente para ver que el mínimo y el máximo deben alcanzarse en el interior de K, o incluso que tenemos la sospecha de que el mínimo o el máximo de hecho deben alcanzarse en la frontera ∂K. ¿Cómo podemos intentar calcular α y β, y hallar los puntos donde se alcanzan? Este problema, en general, es bastante difícil, y de hecho no tendremos todas las herramientas para abordarlo rigurosamente hasta el Capítulo 15 (en donde probaremos el teorema de los multiplicadores de Lagrange), pero podemos avanzar, mediante algunos ejemplos, algunas de las ideas que suelen emplearse en este tipo de cuestiones.

Ejemplo 13.6. Sean $f(x,y) = x + y + 1$, $K = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1$. Calcular el máximo y el mínimo de f sobre K.

La función f no tiene ningún punto crítico en $\mathbb{R}^2$. Por tanto su máximo y mínimo sobre K no pueden alcanzarse en el interior de K, sino que deben estar forzosamente en la frontera. Puesto que la frontera de K es una curva diferenciable, la circunferencia $\partial K = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, podemos parametrizar esta curva y estudiar los mínimos y máximos de la composición de dicha parametrización con f, que deberán corresponder con los de f en ∂K. En efecto, sea $\gamma : [0, 2\pi] \to \partial K$ definida por

$$\gamma(t) = (\cos t, \sin t).$$

La curva γ es diferenciable y sobreyectiva. Por tanto

$$\max_{t \in [0,2\pi]} f(\gamma(t)) = \max_{(x,y) \in \partial K} f(x,y),$$

luego, analizando los puntos críticos de $g(t) = f(\gamma(t))$, podemos determinar el máximo de g en $[0, 2\pi]$ y por tanto también el de f en ∂K, y lo mismo puede decirse de los mínimos. Tenemos que

$$g'(t) = -\sin t + \cos t = 0 \iff t \in \{\pi/4, 3\pi/4\},$$

y como $g(\pi/4) = f(\sqrt{2}/2, \sqrt{2}/2) = 1 + \sqrt{2}$ y $g(3\pi/4) = f(-\sqrt{2}/2, -\sqrt{2}/2) = 1 - \sqrt{2}$, es claro que el valor máximo de f en ∂K (y por tanto en K) es $1 + \sqrt{2}$ y se alcanza en $(\sqrt{2}/2, \sqrt{2}/2)$, y el mínimo es $1 - \sqrt{2}$ y se alcanza en $(\sqrt{2}/2, \sqrt{2}/2)$.

Ejemplo 13.7. Hallar los extremos absolutos de $f(x, y) = x - x^2 - y^2$, en $K = [-1, 1] \times [-1, 1]$.

En este caso f sólo tiene un punto crítico en $(1/2, 0)$, que corresponde a un máximo local. Por tanto, al menos el mínimo de f en K debe alcanzarse en la frontera. Para hallar los extremos de f en ∂K podemos parametrizar los cuatro segmentos que componen ∂K y componer f con estas curvas, obteniendo así cuatro funciones,

$$h_1(t) = f(1, t); \ h_2(t) = f(-1, t); \ h_3(t) = f(t, 1) \ , h_4(t) = f(t, -1),$$

$t \in [-1, 1]$, de las cuales podemos calcular sus extremos en el intervalo $[-1, 1]$. El problema se reduce entonces a comparar los valores de $f \circ h_i$ en los puntos críticos de las funciones h_i con los valores que toma f en el punto $(1/2, 0)$ y en los vértices del cuadrado K. El resultado es que el máximo de f en K es $1/4$ y se alcanza en el punto $(1/2, 0)$, mientras que el mínimo es -3 y se alcanza en los vértices $(-1, 1)$ y $(-1, -1)$.

Cuando se use esta técnica de parametrizar las curvas que componen el borde de un compacto K, es importante asegurar que las parametrizaciones γ sean regulares, es decir, que $\gamma'(t) \neq 0$ para todo t. De lo contrario, entre los puntos críticos de la función $g(t) = f(\gamma(t))$ estarían todos los $\gamma(t)$ para los cuales $\gamma'(t) = 0$ (ya que $g'(t) = Df(\gamma(t))(\gamma'(t))$, por la regla de la cadena), que no tienen por qué ser extremos de la restricción de f a γ.

Ejemplo 13.8. Hallar los extremos de $f(x, y, z) = x + y + z$ en $K = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1\}$.

Como f no tiene ningún punto crítico en $\mathbb{R}^3$, los extremos de f en el compacto K deben estar necesariamente en la frontera de K, la esfera unidad $\partial K = \{(x, y, z) \in R^3 : x^2 + y^2 + z^2 = 1\}$. La técnica análoga a la que hemos recurrido en el ejemplo 13.6 sería parametrizar la superficie de ∂K mediante una función diferenciable y sobreyectiva de dos variables, y calcular los puntos críticos de esta. Sea $\Phi : \mathbb{R}^2 \to \partial K$ definida por

$$\Phi(u, v) = (\sin u \cos v, \sin u \sin v, \cos u).$$

La aplicación Φ es diferenciable y sobreyectiva. Es además una parametrización *regular* de la esfera, en el sentido de que

$$\text{rango} D\Phi(u, v) = 2$$

para todo (u, v). Si definimos

$$\varphi(u, v) = f(\Phi(u, v)) = \sin u \cos v + \sin u \sin v + \cos u,$$

se tiene que

$$\max_{(u,v) \in [-\pi,\pi] \times [-\pi,\pi]} \varphi(u, v) = \max_{(x,y,z) \in \partial K} f(x, y, z),$$

y otro tanto ocurre con el mínimo. Los puntos críticos de φ vienen dados por el sistema de ecuaciones

$$\begin{cases} \frac{\partial\varphi}{\partial u} = \cos u(\cos v + \sin v) - \sin u = 0 \\ \frac{\partial\varphi}{\partial v} = \cos u(\cos v - \sin v) = \sin u \end{cases}.$$

Es fácil ver que este sistema tiene un número finito de soluciones para $u, v \in [-\pi, \pi]$, que corresponden a un número finito de puntos $(x, y, z) \in \partial K$ en algunos de los cuales f alcanzará su máximo y su mínimo en ∂K. Comparando los valores de f en estos puntos se ve inmediatamente que f alcanza su máximo sobre ∂K para $x = y = z = \sqrt{3}/3$, y su mínimo para $x = y = z = -\sqrt{3}/3$.

Este método de buscar puntos críticos de la composición de f con parametrizaciones regulares de la frontera de K es en la práctica poco conveniente, y resulta mucho más adecuado utilizar el siguiente resultado, cuya versión más general se conoce como teorema de los multiplicadores de Lagrange y será demostrado en el Capítulo 15.

Teorema 13.7. *Sean $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ funciones de clase C^1. Sean $x_0 \in U$, $r = g(x_0)$, $S_r = g^{-1}(r)$ el conjunto de nivel r de g. Supongamos que $\nabla g(x_0) \neq 0$, y que la restricción de f a S_r, denotada $f_{|S_r}$, tiene un máximo o un mínimo en x_0. Entonces existe $\lambda \in \mathbb{R}$ tal que*

$$\nabla f(x_0) = \lambda \nabla g(x_0).$$

Es decir, los extremos de $f_{|S_r}$ se alcanzan en puntos en los que el gradiente de f es perpendicular a la superficie de nivel S_r de g.

Aunque aún no podemos dar una demostración rigurosa de este teorema, sí que podemos dar una justificación intuitiva del mismo.

Por un lado, recuérdese que el espacio tangente a S_r en x_0 se define como el espacio ortogonal a $\nabla g(x_0)$ (ver la Definición 9.3), y la justificación de esto fue precisamente el hecho de que $\nabla g(x_0)$ es perpendicular al vector tangente $\gamma'(0)$ a cualquier camino diferenciable $\gamma : (-1, 1) \to S_r$ tal que $\gamma(0) = x_0$ (ver el Teorema 9.11). Por otro lado, si γ es una de estas curvas, la función $h(t) = f(\gamma(t))$ debe tener un extremo en $t = 0$, y por tanto

$$0 = h'(0) = \langle \nabla f(x_0), \gamma'(0) \rangle,$$

es decir, $\nabla f(x_0)$ es perpendicular a todos los vectores $\gamma'(0)$. Así $\nabla f(x_0)$ es también ortogonal al espacio tangente a S_r en x_0. Por tanto $\nabla f(x_0)$ y $\nabla g(x_0)$ deben ser paralelos (al ser ambos perpendiculares al mismo hiperplano), es decir, existe $\lambda \in R$ tal que $\nabla f(x_0) = \lambda \nabla g(x_0)$.

Para que este argumento sea una verdadera demostración sólo queda comprobar un hecho, y es que para todo vector v del hiperplano tangente a S_r en x_0 (es decir, para todo vector v perpendicular a $\nabla g(x_0)$) existe una curva diferenciable $\gamma : (-1, 1) \to S_r$ tal que $\gamma'(0) = v$. Esto es lo que aún no estamos en condiciones de probar rigurosamente; habremos de esperar a tener a nuestra disposición el teorema de la función implícita.

Para ilustrar la mayor efectividad de este método en comparación con el de parametrizar la superficie de nivel S_r, repitamos el ejemplo 13.8 para una función lineal f genérica.

Ejemplo 13.9. Hallar los extremos de $f(x, y, z) = ax + by + cz$ en $K = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1\}$.

Salvo en el caso trivial en que $a = b = c = 0$, f no tiene ningún punto crítico en el interior de K (ni en $\mathbb{R}^3$), luego los extremos de f en K deben alcanzarse en ∂K. Definamos $g(x, y, z) = x^2 + y^2 + z^2$. Nótese que $\partial K = g^{-1}(1)$ y que $\nabla g(x, y, z) \neq 0$ para todo $(x, y, z) \in \partial K$. Si

(x, y, z) es uno de los extremos, deberá satisfacer, según el teorema anterior, que, para algún $\lambda \in \mathbb{R}$,

$$(a, b, c) = \nabla f(x, y, z) = \lambda \nabla g(x, y, z) = \lambda(2x, 2y, 2z),$$

es decir (x, y, z) es proporcional a (a, b, c). Como además es $x^2 + y^2 + z^2 = 1$, se deduce que

$$(x, y, z) \in \left\{ \frac{1}{\sqrt{a^2 + b^2 + c^2}}(a, b, c),\ -\frac{1}{\sqrt{a^2 + b^2 + c^2}}(a, b, c) \right\}.$$

Es obvio entonces que el máximo de f en ∂K es

$$\frac{a + b + c}{\sqrt{a^2 + b^2 + c^2}},$$

y se alcanza en el punto $\frac{1}{\sqrt{a^2+b^2+c^2}}(a, b, c)$, mientras que el mínimo es

$$-\frac{a + b + c}{\sqrt{a^2 + b^2 + c^2}},$$

y se alcanza en $-\frac{1}{\sqrt{a^2+b^2+c^2}}(a, b, c)$.

13.1. Problemas

Problema 13.1. Estúdiense los puntos críticos, máximos y mínimos relativos de las siguientes funciones:

$$\begin{array}{rl} f(x, y) = x - x^2 - y^2 & g(x, y) = x^2 - y^2 + xy \\ h(x, y) = x^2 + y^2 + 3xy & F(x, y) = y^2 - x^3 \\ H(x, y) = x^3 + y^3 - 3axy & \psi(x, y) = \sin(x^2 + y^2) \\ \varphi(x, y) = xye^{x+2y} & \Phi(x, y) = e^{1+x^2-y^2} \\ \phi(x, y, z) = x^2 + y^2 - z^2 & P(x, y, z) = x^2 + y^2 + z^2 - 2x + yz \end{array}$$

Problema 13.2. Calcúlense:

1. Los extremos absolutos de las funciones siguientes, en los conjuntos que se indican:
   a) $f(x, y) = x - x^2 - y^2$, en $K = \{(x, y) \in \mathbb{R}^2, \quad \text{máx}\{|\, x \,|, |\, y \,|\} \leq 1\}$.
   b) $f(x, y) = \sin(xy)$, en $K = \{(x, y) \in \mathbb{R}^2, \quad \text{máx}\{|\, x \,|, |\, y \,|\} \leq 1\}$.
   c) $f(x, y) = \sin(x + y) + \cos(x - y)$, en $K = [0, 2\pi] \times [0, 2\pi]$.
2. $\sup\{4x^2 + y^2 - 4x - 3y;\ y \geq 0,\ 4x^2 + y^2 \leq 4\}$.
3. $\sup\{(x + y)e^{-(x^2+y^2)};\ (x, y) \in \mathbb{R}^2\}$.

Problema 13.3. Sea $f(x, y) = (y - 3x^2)(y - x^2)$

1. Pruébese que (0,0) es un punto crítico de f.
2. Pruébese que f tiene un mínimo relativo en (0,0) sobre cada recta que pasa por (0,0).
3. Pruébese que (0,0) no es un mínimo relativo de f, sino un punto de silla.

Problema 13.4. Encuéntrense los puntos $(x, y) \in \mathbb{R}^2$ y las direcciones para las que la derivada direccional de $f(x, y) = 3x^2 + y$ alcanza el máximo valor, suponiendo que (x, y) verifica $x^2 + y^2 \leq 1$.

Problema 13.5. Hállese la ecuación del plano que pasa por $(1, 1, 2)$ y determina con los ejes de coordenadas un tetraedro de volumen mínimo (no nulo).

Problema 13.6. Sean $f(x, y) = e^{ax+y^2} + b\sin(x^2 + y^2)$ $(a, b \in \mathbb{R})$

$$g(x,y) = \begin{cases} \frac{xy}{x+y} & \text{si } x + y \neq 0 \\ 0 & \text{si } x + y = 0 \end{cases}$$

1. Demuéstrese que g posee derivada según cualquier dirección en el punto (0,0), pero que no es diferenciable en (0,0).
2. Determínense los valores de los parámetros a y b sabiendo que
   a) f tiene un extremo en (0,0).
   b) El polinomio de Taylor de grado 2 de f en el origen toma el valor 6 en el punto (1,2).
3. Con los resultados obtenidos en (b), ¿qué clase de extremo es el punto (0,0) para f?

Problema 13.7. Sea $f(x, y) = a[2xy + y^2 + yx^2 + \cos(x + y)] + x^2(a^2 - y)$. Discútase la existencia de extremos en el origen, según los valores de a.

Problema 13.8. Encuéntrense los extremos relativos y los absolutos, si existen, de las siguientes funciones:

1. $f(x, y) = (ax^2 + by^2)e^{-(x^2+y^2)}, \quad a, b > 0.$
2. $f(x, y, z) = (x + z^2)e^{x(y^2+z^2+1)}$.

Problema 13.9. Hállense los extremos de $f(x, y, z) = xy^2z^3(a - x - y - z)$.

Problema 13.10. Sea D un abierto acotado de $\mathbb{R}^n$ y sea $f : \overline{D} \longrightarrow \mathbb{R}$ una función continua, diferenciable en D. Si f se anula en la frontera de D, demuéstrese que existe al menos un punto $a \in D$ tal que $Df(a) = 0$. Este resultado puede considerarse como la generalización del teorema de Rolle a funciones de varias variables.

Problema 13.11. Hallar los puntos de la superficie $z^2 = xy + 1$ de $\mathbb{R}^3$ que están más próximos al origen $(0, 0, 0)$.

Problema 13.12. Probar que el máximo de $x^2y^2z^2$ sobre la esfera $x^2 + y^2 + z^2 = r^2$ es $(r^2/3)^3$.

Problema 13.13. Probar que el mínimo de $x^2 + y^2 + z^2$ sobre la superficie $x^2y^2z^2 = (r^2/3)^3$ es r^2.

Problema 13.14. Encontrar el máximo de $x_1^2x_2^2...x_n^2$ sobre la esfera $x_1^2 + x_2^2 + ... + x_n^2 = 1$. Deducir que para cualesquiera n números positivos a_i se tiene que

$$(a_1a_2...a_2)^{1/n} \leq \frac{a_1 + a_2 + ... + a_n}{n},$$

es decir, la media geométrica es siempre menor o igual que la media aritmética.

Problema 13.15. Sean $a_1, ..., a_n$ números positivos. Para $p > 1$, sea q tal que

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Hallar el máximo de $f(x) = \sum_{i=1}^n a_i x_i$ sobre la superficie S definida por $\sum_{i=1}^n |x_i|^q = 1$. Deducir que

$$\sum_{i=1}^n a_i x_i \leq \left(\sum_{i=1}^n a_i^p\right)^{1/p} \left(\sum_{i=1}^n |x_i|^q\right)^{1/q}$$

(la *desigualdad de Hölder* para sumas). Probar también que

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

(desigualdad de Minkowski para sumas).

Problema 13.16. Sea $f : \mathbb{R} \to \mathbb{R}$ diferenciable en $\mathbb{R}$ tal que f tiene un mínimo local en un punto x_0, y que no hay más puntos críticos. Probar que x_0 es entonces un mínimo global.

Problema 13.17. Probar que el problema anterior no es válido para funciones $f : \mathbb{R}^2 \to \mathbb{R}$. Considérese para ello la función $f(x, y) = x^2 + y^2(1 - x)^3$, mostrando que $(0, 0)$ es el único punto crítico, que es un mínimo local estricto, y que f no tiene ningún mínimo global.

Capítulo 14

Teoremas de la función inversa e implícita

En este capítulo estudiaremos los dos teoremas más importantes del curso. El primero de ellos nos proporciona condiciones suficientes bajo las cuales una función diferenciable es invertible con inversa también diferenciable. Es decir, si tenemos $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$, deseamos saber si dado $y \in \mathbb{R}^n$ existe un único $x \in U$ tal que $f(x) = y$, y en caso afirmativo si la aplicación $y \mapsto x := f^{-1}(y)$ es diferenciable.

Si f es una función de clase C^1 entonces f se podrá aproximar bien por su derivada en un entorno de cada punto $a \in U$, y nos cabe esperar que las propiedades de invertibilidad de la aplicación lineal $Df(a)$ puedan traducirse en propiedades de invertibilidad de f en un entorno de a. Esto sugiere que el primer paso para abordar el problema que nos planteamos es estudiar lo que ocurre cuando f es lineal.

Si $L : \mathbb{R}^n \to \mathbb{R}^n$ es una aplicación lineal, el problema de encontrar una inversa sabemos que equivale a resolver el sistema de ecuaciones lineales $L(x) = y$, donde x es la incógnita, es decir, si $A = (a_{ij})$ es la matriz de L, queremos encontrar un único $x = (x_1, ..., x_n)$ tal que

$$\begin{array}{ccccc} a_{11}x_1 & + & ... & + & a_{1n}x_n = y_1 \\ \vdots & \vdots & ... & \vdots & \vdots \\ a_{n1}x_1 & + & ... & + & a_{nn}x_n = y_n. \end{array}$$

Por la teoría de álgebra lineal sabemos que esto es posible si y sólo si $\det A \neq 0$, en cuyo caso $x = L^{-1}(y)$, donde L^{-1} es la aplicación lineal cuya matriz viene dada por

$$A^{-1} = \frac{1}{\det A}\text{adj}(A)^t,$$

donde adj(A) es la matriz $B = (b_{ij})$ tal que $b_{ij} = (-1)^{i+j} \det A(j|i)$, donde $\det A(j|i)$ denota el determinante de la matriz que se obtiene de A al suprimir la j-ésima fila y la i-ésima columna, y donde B^t denota transposición, es decir, cambiar en B las filas por las columnas.

Es verosímil entonces que, si $\det Df(a) \neq 0$ entonces $y = f(a) + Df(a)(x)$ tiene solución en x, y como f puede ser bien aproximada por $f(a) + Df(a)$ en un entorno de a, la ecuación $y = f(x)$ debería tener solución en x para todo y en un entorno de $f(a)$, es decir, debería existir $x = f^{-1}(y)$ para y en un entorno de $f(a)$. Esto es lo que viene a decirnos el teorema de la función inversa.

Más en general, casi todos los teoremas que veremos en este capítulo vienen a decir: un sistema de ecuaciones que involucre funciones diferenciables se comporta localmente (es decir en un entorno de cada punto) de manera análoga a como se comportaría el sistema de ecuaciones lineales correspondiente a la aplicación lineal formada por las diferenciales de dichas funciones en el punto en cuestión, siempre y cuando dicha aplicación lineal tenga rango máximo.

Teorema 14.1 (de la función inversa)**.** *Sean U un subconjunto abierto de $\mathbb{R}^n$, y $f : U \to \mathbb{R}^n$ una aplicación de clase C^1 en U. Supongamos que* $\det Df(a) \neq 0$ *(es decir $Df(a)$ es invertible). Entonces f es localmente invertible en un entorno de a, con inversa de clase C^1. Es decir, existen V entorno abierto de a, W entorno abierto de $b = f(a)$ tales que $f_{|V} : V \to W$ es biyectiva, y $f^{-1} : W \to V$ es diferenciable en b. Además, si $f \in C^p(U, \mathbb{R}^n)$ entonces $f^{-1} \in C^p(W, \mathbb{R}^n)$. Por último, la derivada de f^{-1} en $y = f(x)$ es la inversa de la aplicación lineal $Df(x)$, es decir,*

$$Df^{-1}(y) = [Df(x)]^{-1}.$$

Antes de probar este teorema haremos algunos comentarios sobre el mismo. Una primera observación es que la hipótesis de que $Df(a)$ sea invertible es claramente necesaria para obtener una inversa diferenciable: puesto que $f \circ f^{-1} = I$ y $f^{-1} \circ f = I$, aplicando la regla de la cadena obtenemos que

$$Df(a) \circ Df^{-1}(b) = I, \text{ y } Df^{-1}(b) \circ Df(a) = I,$$

con lo que queda claro que $Df(a)$ debe ser invertible (y de paso se ve también por qué las dimensiones del espacio dominio y del espacio imagen deben ser iguales: una aplicación lineal entre espacios de dimensiones diferentes nunca puede ser un isomorfismo; recuérdese también el teorema de invariancia del dominio).

Esta hipótesis no es necesaria, sin embargo, si lo único que deseamos es obtener la inyectividad local de f, como prueba el ejemplo $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$: aquí se tiene $f'(0) = 0$ y f es inyectiva.

Una segunda observación es que, a diferencia de lo que ocurre en dimensión uno (donde una función diferenciable sobre un intervalo con derivada no nula es estrictamente creciente o estrictamente decreciente y por tanto inyectiva globalmente; ver el problema 14.1), el resultado no es, en principio, globalizable: no basta, en general, que la diferencial de una función $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ de clase C^1 sea invertible en cada punto para que la función f sea inyectiva en U. El teorema de la función inversa nos dice que f es inyectiva en un entorno de cada punto, pero no necesariamente va a ser globalmente inyectiva, como prueba el siguiente ejemplo.

Ejemplo 14.1. Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

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

Es obvio que f es de clase $C^\infty(\mathbb{R}^2, \mathbb{R}^2)$, y la matriz jacobiana de f es

$$\begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}$$

por tanto

$$\det Df(x, y) = e^{2x}(\cos^2 y + \sin^2 y) = e^{2x} \neq 0$$

para todo $(x, y) \in \mathbb{R}^2$, y según el teorema de la función inversa f es localmente inyectiva con inversa local de clase C^∞. Sin embargo, puesto que $f(0, 2\pi) = f(0, 0)$, la función f no es inyectiva en $\mathbb{R}^2$.

Otra observación útil en la práctica es que decir que f es localmente invertible en un entorno de $b = f(a)$ equivale a asegurar que la ecuación $y = f(x)$ tiene solución única en x para y cerca de b, x cerca de a.

Ejercicio 14.1. Estudiar si el sistema de ecuaciones

$$u = x + xyz$$
$$v = y + xy$$
$$w = z + 2x + 3z^2$$

tiene solución para x, y, z en términos de u, v, w cerca del punto $(0, 0, 0)$.

La prueba del teorema de la función inversa no es sencilla, pero se trata de un resultado esencial del análisis matemático y debe por tanto ser dominada; además en su demostración aparecen técnicas, como la aplicación del teorema del punto fijo para aplicaciones contractivas, que son frecuentemente usadas en otros problemas. Para tratar de hacerla más inteligible, dividiremos la demostración de este resultado en varios pasos, todos ellos de interés en sí mismos.

Comenzamos con un lema acerca del conjunto de aplicaciones lineales invertibles y la operación de inversión de una de estas aplicaciones. Si $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ es el espacio de las aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^n$, denotaremos por $GL(\mathbb{R}^n)$ el conjunto de todas las aplicaciones lineales invertibles de $\mathbb{R}^n$ en $\mathbb{R}^n$; a este conjunto se le llama el grupo lineal general. Así, $GL(\mathbb{R}^n) = \{A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) : \det A \neq 0\}$.

Proposición 14.2. *El conjunto $GL(\mathbb{R}^n)$ es abierto en el espacio $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, y la aplicación $\mathcal{J} : GL(\mathbb{R}^n) \to GL(\mathbb{R}^n)$ definida por*

$$A \to \mathcal{J}(A) = A^{-1}$$

es de clase C^∞.

Demostración. Identificando $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$ con $\mathbb{R}^{n^2}$, es claro que la aplicación $\det : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \to \mathbb{R}$ es un polinomio en n^2 variables, luego es una aplicación de clase C^∞ y en particular continua. Entonces $GL(\mathbb{R}^n) = \det^{-1}(\mathbb{R} \setminus \{0\})$ es abierto en $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$.

Por otra parte, $\mathcal{J}(A) = A^{-1} = \frac{1}{\det A}\text{adj}(A)^t$, luego, como las aplicaciones determinante (por ser un polinomio) y la transposición (por ser lineal) son de clase C^∞ basta ver que la aplicación $A \mapsto \text{adj}(A)$ es también de clase C^∞. Sabemos que

$$(\text{adj}(A))_{ij} = (-1)^{i+j} \det A(j|i),$$

donde $A(j|i)$ es la matriz obtenida de A al suprimir la fila j-ésima y la columna i-ésima. Las funciones componentes de $A \mapsto \text{adj}(A)$ son por tanto de la forma $A \mapsto (-1)^{i+j} \det A(j|i)$ y, nuevamente, como $\det$ es de clase C^∞, es suficiente ver que cada aplicación $A \mapsto A(j|i)$ es de clase C^∞. Pero, vista como una aplicación de $\mathbb{R}^{n^2}$ en $\mathbb{R}^{(n-1)^2}$, la función $A \mapsto A(j|i)$ no es más que una proyección lineal, y es por tanto obvio que es de clase C^∞. $\square$

Como consecuencia del teorema del punto fijo para aplicaciones contractivas, tenemos el siguiente resultado, que, aparte de ser interesante en sí mismo, es una de las piezas fundamentales de la demostración del teorema de la función inversa. Nos dice que toda perturbación de la identidad por una suma de una función contractiva es homeomorfismo.

Teorema 14.3. *Sean U un abierto de un espacio de Banach E, y $f : U \to E$ una aplicación contractiva (es decir, $\|f(x) - f(y)\| \leq K\|x - y\|$ para todo $x, y \in U$, donde K es una constante tal que $0 < K < 1$). Entonces la aplicación $g : x \mapsto x + f(x)$, es un homeomorfismo de U sobre un abierto V de E, y la aplicación $h : y \mapsto y - g^{-1}(y)$ es $K/(1 - K)$-Lipschitz de V en E.*

Demostración. Veamos primero que g es inyectiva y g^{-1} es $1/(1-K)$-Lipschitz. En efecto, para todos $x, x' \in U$ se tiene

$$\begin{aligned}\|g(x) - g(x')\| &= \|(x - x') + (f(x) - f(x'))\| \geq \|x - x'\| - \|f(x) - f(x')\| \\ &\geq \|x - x'\| - K\|x - x'\| = (1 - K)\|x - x'\|,\end{aligned}$$

lo que muestra que g es inyectiva y $g^{-1} : g(U) \to U$ es $1/(1-K)$-Lipschitz.

Veamos ahora que $V = g(U)$ es abierto. Sea $b = g(a) \in V = g(U)$. Como U es abierto existe $r > 0$ tal que $\overline{B}(a, r) \subset U$. Pongamos $s := (1 - K)r$, y veamos que $B(b, s) \subset V$, lo que permitirá concluir que V es abierto. Si $y \in B(b, (1-K)r)$ entonces, para todo $x \in \overline{B}(a, r)$ tenemos que

$$\begin{aligned}&\|y - f(x) - a\| \leq \|y - b\| + \|b - a - f(x)\| = \\ &\|y - b\| + \|a + f(a) - a - f(x)\| \\ &\leq \|y - b\| + K\|a - x\| \leq (1 - K)r + Kr = r,\end{aligned}$$

es decir, $y - f(x) \in \overline{B}(a, r)$. Por tanto la aplicación $\varphi : x \mapsto y - f(x)$ lleva la bola cerrada $\overline{B}(a, r)$ dentro de sí misma, y puesto que

$$\|\varphi(x) - \varphi(x')\| = \|f(x) - f(x')\| \leq K\|x - x'\|,$$

resulta que $\varphi : \overline{B}(a, r) \to \overline{B}(a, r)$ es una aplicación contractiva del espacio métrico completo $\overline{B}(a, r)$ en sí mismo. Entonces, por el Teorema del punto fijo 4.18, φ tiene un único punto fijo x_y, que satisface por tanto $x_y = y - f(x_y)$, es decir, $y = g(x_y)$, lo que muestra que $y \in g(\overline{B}(a, r)) \subset g(U) = V$.

Nótese que este argumento también prueba que $g^{-1}(y) = x_y$, donde x_y es el único punto fijo de la aplicación contractiva $\varphi = \varphi_y : x \mapsto y - f(x)$.

Veamos por último que h es $K/(1 - K)$-Lipschitz. Sea $y = g(x)$. Entonces $x = y - f(x)$, luego $h(y) = y - g^{-1}(y) = y - x = f(x)$. Es decir, $h(y) = f(g^{-1}(y))$. Como f es K-Lipschitz y g^{-1} es $1/(1 - K)$-Lipschitz se deduce inmediatamente que la composición $h = f \circ g^{-1}$ es $K/(1 - K)$-Lipschitz (ver el problema 5.25). □

Corolario 14.4. *Sea $L : \mathbb{R}^n \to \mathbb{R}^n$ una aplicación lineal con $\|L\| < 1$. Entonces $I - L$ es invertible, y*

$$\|(I - L)^{-1}\| \leq \frac{1}{1 - \|L\|}.$$

Demostración. Puede aplicarse el teorema anterior ya que $\|L(x) - L(y)\| \leq \|L\|\|x - y\|$ para todo $x, y \in \mathbb{R}^n$. □

Ya estamos en condiciones de demostrar el teorema de la función inversa.

Demostración del Teorema 14.1.

Sea L la inversa de $Df(a)$. La aplicación $g : U \to \mathbb{R}^n$ definida por

$$g(x) = L \circ f(x) - x,$$

es de clase C^1 y su derivada es

$$Dg(x) = L \circ Df(x) - I;$$

en particular se tiene $Dg(a) = 0$, y como Dg es continua existe $r > 0$ tal que $B(a,r) \subset U$ y $\|Dg(x)\| \leq 1/2$ para todo $x \in B(a,r)$, luego, por la desigualdad del valor medio,

$$\|g(x) - g(x')\| \leq \frac{1}{2}\|x - x'\|$$

para todos $x, x' \in B(a,r)$. Entonces, en la bola $B(a,r)$, la aplicación $L \circ f$ es la suma de la identidad y la aplicación $1/2$-Lipschitz g. Por el Teorema 14.3 resulta que $L \circ f$ es un homeomorfismo de $B(a,r)$ sobre un abierto A de $\mathbb{R}^n$, luego $f = L^{-1} \circ (L \circ f)$ es un homeomorfismo de $V := B(a,r)$ sobre el abierto $W := L^{-1}(A) = Df(a)(A)$ de $\mathbb{R}^n$.

Veamos ahora que f^{-1} es diferenciable en $b = f(a)$, con $Df^{-1}(b) = Df(a)^{-1}$. Sea $\varepsilon > 0$. Como Dg es continua, existe $\rho \in (0,r)$ tal que $\|Dg(x)\| \leq \varepsilon$ para todo $x \in B(a,\rho)$. En particular g es ε-Lipschitz en $B(a,\rho)$, y otra vez por el Teorema 14.3, sabemos que $(L \circ f)^{-1} = f^{-1} \circ L^{-1}$ es, sobre el abierto $W_\rho = L \circ f(B(a,\rho))$ (que es entorno de $L(b)$) la suma de la identidad y de una función $\varepsilon/(1-\varepsilon)$-Lipschitz. Así, para todo $y \in W_\rho$ tendremos que

$$\|(L \circ f)^{-1}(y) - y - ((L \circ f)^{-1}(L(b)) - L(b))\| \leq \frac{\varepsilon}{1-\varepsilon}\|y - L(b)\|,$$

es decir

$$\|(L \circ f)^{-1}(y) - (L \circ f)^{-1}(L(b)) - (y - L(b))\| \leq \frac{\varepsilon}{1-\varepsilon}\|y - L(b)\|$$

para todo y en el entorno W_ρ de $L(b)$. Como $\lim_{\varepsilon \to 0} \frac{\varepsilon}{1-\varepsilon} = 0$, esto implica que $(L \circ f)^{-1}$ es diferenciable en $L(b)$, y su diferencial en este punto es la identidad. Entonces, por la regla de la cadena, $f^{-1} = (L \circ f)^{-1} \circ L$ es diferenciable en b, con diferencial $Df^{-1}(b) = I \circ L = L = Df(a)^{-1}$.

Finalmente, veamos que si $f \in C^p(V)$ entonces $f^{-1} \in C^p(W)$. Para $p = 1$, ya sabemos que f es un homeomorfismo de V en W y en particular f^{-1} es continua en W. Entonces, como

$$Df^{-1}(y) = (Df(f^{-1}(y)))^{-1} = \mathcal{J} \circ (Df) \circ f^{-1}(y)$$

para todo $y \in W$, y puesto que la aplicación $\mathcal{J} : GL(\mathbb{R}^n) \to GL(\mathbb{R}^n)$ es de clase C^∞ y Df es continua, se concluye que $Df^{-1} : W \to V$ es también continua. En el caso general se razona por inducción: suponiendo que ya sabemos que f^{-1} es de clase C^k (con $k \leq p-1$), como Df es de clase $p - 1 \geq k$, la fórmula

$$Df^{-1} = \mathcal{J} \circ Df \circ f^{-1}$$

nos permite concluir que Df^{-1} es de clase C^k (por ser composición de aplicaciones C^k), luego f^{-1} es de clase C^{k+1}. □

Veamos ahora algunas consecuencias importantes del teorema de la función inversa.

Definición 14.1. Sean V y W abiertos en dos espacios vectoriales de dimensión finita. Se dice que $f : V \to W$ es un difeomorfismo de clase C^p si f es una biyección de V en W y tanto $f : V \to W$ como $f^{-1} : W \to V$ son diferenciables de clase C^p. En este caso se dice también que U y V son difeomorfos.

En particular todo difeomorfismo es también un homeomorfismo, y por tanto, teniendo en cuenta el teorema de invariancia del dominio, dos abiertos sólo pueden ser difeomorfos si están en espacios de la misma dimensión. Este hecho puede probarse también directamente sin recurrir al teorema de invariancia del dominio: si $U \subseteq E$, $V \subseteq F$ y $f : U \to V$ es un difeomorfismo de clase C^1 entonces, por la regla de la cadena, $I_F = Df(x) \circ Df^{-1}(f(x))$,

$I_E = Df^{-1}(f(x)) \circ Df(x)$, luego $Df(x) : E \to F$ es un isomorfismo lineal con inversa $Df^{-1}(f(x))$, y en particular E y F tienen la misma dimensión.

Lo que sigue es una reformulación del teorema de la función inversa, y algunos corolarios útiles del mismo.

Teorema 14.5 (de la función inversa)**.** *Sean U abierto en $\mathbb{R}^n$, $f : U \to \mathbb{R}^n$ de clase C^p, y supongamos que* $\det Df(a) \neq 0$. *Entonces existen abiertos $V \subseteq U$ y W de $\mathbb{R}^n$ tales que $a \in V$ y $f : V \to W$ es un difeomorfismo de clase C^p.*

Se dice que una aplicación entre dos espacios métricos es abierta si la imagen por ella de cualquier abierto es un abierto.

Corolario 14.6. *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ de clase C^1, y supongamos que* $\det Df(x) \neq 0$ *para cada $x \in U$. Entonces f es una aplicación abierta.*

Si la aplicación f es además inyectiva globalmente se concluye que f es un difeomorfismo global.

Corolario 14.7. *Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ de clase C^1 e inyectiva. Entonces, f es un difeomorfismo sobre su imagen $f(U)$ si y sólo si* $\det Df(x) \neq 0$ *para todo $x \in U$.*

Pasamos ahora a estudiar el teorema de la función implícita. Así como el teorema de la función inversa nos permite asegurar la existencia local de soluciones en x de ecuaciones del tipo $y = f(x)$, donde f es una función diferenciable con derivada invertible, el teorema de la función implícita nos proporcionará condiciones suficientes para la existencia de soluciones de ecuaciones del tipo $F(x, y) = 0$, llamadas implícitas, donde ahora se pretende expresar y como función de x, o bien x en función de y. Este problema parece más general a primera vista (y en efecto, ya que siempre podemos considerar $F(x, y) = f(x) - y$, la primera ecuación es un tipo especial de la segunda), pero en realidad es equivalente al anterior, ya que como veremos el teorema de la función implícita se deduce del teorema de la función inversa.

Se dirá que $F(x, y) = 0$ define implícitamente y como función de x si existe una función f tal que la ecuación $F(x, y) = 0$ equivale a $y = f(x)$. Por supuesto, no debe esperarse una fórmula explícita para f, lo que quiere decir esto es que para cada x puede determinarse unívocamente un $y = f(x)$ tal que $F(x, f(x)) = 0$, y de tal manera que f sea diferenciable si F lo es. Si $F : \mathbb{R}^2 \to \mathbb{R}$, el conjunto de los (x, y) que satisfacen $F(x, y) = 0$ suele ser una curva; lo que estamos pidiendo entonces es que esta curva sea la gráfica de una función $y = f(x)$. El siguiente ejemplo nos advierte que, incluso en los casos más sencillos, esto no es siempre posible.

Ejemplo 14.2. Sea $F : \mathbb{R}^2 \to \mathbb{R}$ definida por $F(x, y) = x^2 + y^2 - 1$. La ecuación $F(x, y) = 0$ define, como bien sabemos, la circunferencia unidad de $\mathbb{R}^2$. Esta curva no es la gráfica de ninguna función f. En efecto, para $-1 < x < 1$ hay siempre dos valores diferentes de y que satisfacen la ecuación, a saber $y = f_1(x) := \sqrt{1 - x^2}$ e $y = f_2(x) := -\sqrt{1 - x^2}$.

Sin embargo, si tomamos un punto (x_0, y_0) de la circunferencia y nos restringimos a un entorno suficientemente pequeño de este punto, sí es verdad que la porción de la circunferencia contenida en este entorno es o bien parte de la gráfica de f_1 o bien parte de la gráfica de f_2, pero nunca de ambas, es decir, sí hay unicidad *local*, salvo cuando $(x_0, y_0) = (-1, 0)$ o $(x_0, y_0) = (1, 0)$ (precisamente donde la circunferencia se dobla y tiene tangente vertical). Los puntos $(-1, 0)$ y $(1, 0)$ son *excepcionales* por varios motivos: en primer lugar, ni f_1 ni f_2 son diferenciables en estos puntos; por otra parte, en cualquier entorno de estos puntos la solución para y en función de x podría ser tanto la raíz positiva, f_1, como la negativa, f_2, y por tanto no hay una f diferenciable, ni siquiera continua, univocamente determinada. Nótese además que en los puntos $(-1,0)$ y $(1,0)$ se tiene que $\partial F/\partial y = 0$ (lo que significa que la tangente a la circunferencia es vertical).

Estas consideraciones muestran que cualquier teorema general que permita asegurar la existencia de funciones $y = f(x)$ diferenciables definidas implícitamente por $F(x,y) = 0$, donde F es diferenciable, debe, por una parte, ser local, y además, por otra, imponer condiciones del tipo $\partial F/\partial y \neq 0$. Esta última condición surge también de modo natural si suponemos que $y = \varphi(x)$ es una función diferenciable que satisface $F(x,\varphi(x)) = 0$ y aplicamos la regla de la cadena:

$$\frac{\partial}{\partial x}F(x,\varphi(x)) = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial \varphi}{\partial x} = 0;$$

si ahora pretendemos despejar $\partial\varphi/\partial x$, deberemos exigir que $\partial F/\partial y \neq 0$.

En el caso más general consideraremos, pues, una función diferenciable $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ y estudiaremos la ecuación $F(x,y) = 0$, que corresponde al sistema de ecuaciones

$$\begin{gathered} F_1(x_1,...,x_n,y_1,...,y_m) = 0 \\ \vdots \\ F_m(x_1,...,x_n,y_1,...,y_m) = 0. \end{gathered}$$

El objetivo es expresar $y_1,...,y_m$ como funciones diferenciables de $(x_1,...,x_n)$ al menos en un entorno de cada punto. ¿Cuál es la condición análoga a la de $\partial F/\partial y \neq 0$ en el ejemplo del círculo que debemos exigir en este caso para que nuestro problema tenga solución? Para ver más claro el problema, consideremos el caso ya conocido por los cursos de álgebra lineal en que F es lineal, o un poco más general, en que F es de la forma $F(x) = h(x) + L(y)$, donde L es lineal con matriz (a_{ij}) respecto de las bases canónicas. Entonces la ecuación $F(x,y) = 0$ corresponde al sistema de ecuaciones lineales en y siguiente:

$$\begin{array}{cccc} a_{11}y_1+ & ... & +a_{1m}y_m = & -h_1(x) \\ \vdots & ... & \vdots & \vdots \\ a_{m1}y_1+ & ... & +a_{mm}y_m = & -h_n(x). \end{array}$$

En este caso sabemos por la teoría de álgebra lineal que el sistema tiene solución única si y sólo si en rango del sistema es máximo, es decir, si y sólo si el determinante de la matriz (a_{ij}) (o sea el de la aplicación lineal L) es no nulo. En este caso, además, $y = L^{-1}(-h(x))$. Puesto que L es la diferencial parcial de F con respecto a $y = (y_1,...,y_m)$, parece lógico pensar que la condición que necesitamos en el caso general es que la diferencial de $y \to F(x,y)$ sea invertible.

Notación 14.8. Sea $F : U \times V \subseteq \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$. Si la función parcial

$$F_x : V \subseteq \mathbb{R}^m \to \mathbb{R}^k,\ y \mapsto F(x,y)$$

es diferenciable, denotaremos su diferencial por

$$\frac{\partial F}{\partial y} = \frac{\partial[F_1,...,F_k]}{\partial[y_1,...,y_m]}.$$

Ya podemos enunciar y demostrar el teorema de la función implícita.

Teorema 14.9 (de la función implícita)**.** *Sea A un subconjunto abierto de $\mathbb{R}^n \times \mathbb{R}^m$, y sea $F : A \to \mathbb{R}^m$ una función de clase C^p. Supongamos que $F(a,b) = 0$ y que*

$$\det\left(\frac{\partial F}{\partial y}(a,b)\right) \neq 0.$$

Entonces existen U entorno abierto de a en $\mathbb{R}^n$, V entorno abierto de b en $\mathbb{R}^m$, y una única función $\varphi : U \to V$ tal que

$$F(x, \varphi(x)) = 0$$

para todo $x \in U$. Además, φ es de clase C^p, y satisface que

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\frac{\partial \varphi}{\partial x} = 0, \ \textit{es decir} \ \frac{\partial F_i}{\partial x_j} + \frac{\partial F_i}{\partial y_k}\frac{\partial \varphi_k}{\partial x_j} = 0,$$

o, en términos de matrices,

$$\begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1} & \cdots & \frac{\partial \varphi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_m}{\partial x_1} & \cdots & \frac{\partial \varphi_m}{\partial x_n} \end{pmatrix} = -\begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}.$$

Demostración. Vamos a aplicar el teorema de la función inversa a la función $f : A \subseteq \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^m$ definida por

$$f(x,y) = (x, F(x,y)).$$

La ecuación $F(x,y) = 0$ equivale pues a la ecuación $f(x,y) = (x,0)$. Se tiene que $f(a,b) = (a,0)$, y la diferencial de f en (a,b) viene dada por

$$\begin{pmatrix} I & O \\ \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y}, \end{pmatrix}$$

donde I es la matriz identidad. Es decir,

$$Df(a,b)(h,k) = (h, \frac{\partial F}{\partial x}(a,b)(h) + \frac{\partial F}{\partial y}(a,b)(k)).$$

para todo $(h,k) \in \mathbb{R}^n \times \mathbb{R}^m$. Si $Df(a,b)(h,k) = 0$ entonces $h = 0$ y $\frac{\partial F}{\partial y}(a,b)(k) = 0$, luego (al ser $\frac{\partial F}{\partial x}(a,b)$ invertible) también $k = 0$. Esto muestra que $Df(a,b)$ es invertible. Podemos aplicar entonces el teorema de la función inversa a f: existe Ω entorno abierto de (a,b) en A tal que $f : \Omega \to f(\Omega)$ es un difeomorfismo de clase C^p. Sean U_0 entorno abierto de a y V_0 entorno abierto de b tales que $U_0 \times V_0 \subset \Omega$. Entonces $f(U_0 \times V_0)$ es abierto en $\mathbb{R}^n \times \mathbb{R}^m$, y para todo $(x,y) \in U_0 \times V_0$ se tiene que

$$F(x,y) = 0 \iff f(x,y) = (x,0) \iff (x,y) = f^{-1}(x,0).$$

Nótese que, como f es de la forma $f(x,y) = (x, F(x,y))$ (es decir, su primera función coordenada es la proyección $(x,y) \to x$), entonces f^{-1} es también de la forma $f^{-1}(u,v) = (u, \psi(u,v))$, para alguna función $\psi : f(U_0 \times V_0) \to \mathbb{R}^m$ de clase C^p.

Definamos $U = \{x \in \mathbb{R}^n : (x,0) \in f(U_0 \times V_0)\}$. Es inmediato que U es abierto (se tiene $U = j^{-1}(f(U_0 \times V_0))$, donde la inyección $j : x \to (x,0)$ es continua), y es no vacío puesto que $(a,0) \in f(U_0 \times V_0)$, es decir $a \in U$. Sea $\varphi : U \to \mathbb{R}^m$ la aplicación definida por

$$\varphi(x) = \psi(x,0),$$

que es claramente de clase C^p sobre el abierto U. Para cada $x \in U$ se tiene $(x, \varphi(x)) = f^{-1}(x, 0) \in U_0 \times V_0$, y en particular $x \in U_0$, $\varphi(x) \in V_0$. Luego φ lleva U en $V := V_0$.

En conclusión, para todo $(x, y) \in U \times V$ se tiene que

$$F(x, y) = 0 \iff (x, y) = f^{-1}(x, 0) = (x, \varphi(x)) \iff y = \varphi(x).$$

Finalmente, diferenciando la aplicación $x \mapsto F(x, \varphi(x))$, que es constante en U, obtenemos que

$$\begin{aligned} 0 = \frac{\partial F}{\partial x}(x, \varphi(x)) \circ I + \frac{\partial F}{\partial y}(x, \varphi(x)) \circ \varphi'(x) = \\ \frac{\partial F}{\partial x}(x, \varphi(x)) + \frac{\partial F}{\partial y}(x, \varphi(x)) \circ \varphi'(x) = \\ \frac{\partial F}{\partial x}(x, y) + \frac{\partial F}{\partial y}(x, y) \circ \varphi'(x), \end{aligned}$$

lo que termina la demostración del teorema. □

Observación 14.10. En la práctica, en lugar de calcular la matriz inversa de $\partial F/\partial y$ para hallar $\varphi'(x)$, suele ser mucho más conveniente derivar ambos miembros de las expresiones $F_j(x, \varphi(x)) = 0$, $j = 1, ..., m$, respecto de las diferentes variables $x_1, ..., x_n$, para después despejar del sistema así obtenido las derivadas parciales $\partial \varphi_j / \partial x_i$.

Por otra parte, si bien hemos supuesto por simplicidad que las variables están ordenadas de modo que

$$\det\left(\frac{\partial[F_1, ..., F_m]}{\partial[y_1, ..., y_m]}\right) \neq 0,$$

es obvio que puede darse un enunciado análogo para cualquier otra disposición de las variables, siempre que haya un determinante no nulo de orden m en la matriz de $Df(a)$.

Usando el teorema del punto fijo para aplicaciones contractivas, es fácil dar una demostración directa (es decir, sin pasar por el teorema de la función inversa) de la existencia y continuida de φ (ver el problema 14.11), aunque establecer la diferenciabilidad de φ por este procedimiento requiere más trabajo (equivalente al de la demostración del teorema de la función inversa).

Ejemplo 14.3. Estudiar si la ecuación $z^3 + x(z - y) = 1$ define implícitamente una función diferenciable $z = f(x, y)$ en un entorno del punto $(0, 1, 1)$. Derivar implícitamente la ecuación para calcular las derivadas parciales de f, y estudiar sus puntos críticos.

Ejemplo 14.4. Consideremos el siguiente sistema de ecuaciones

$$\begin{cases} xu + yv^2 = 0 \\ xv^3 + y^2u^6 = 0. \end{cases}$$

Estudiar si existe solución única para u, v en términos de x, y cerca de los puntos $x = 1, y = -1, u = 1, v = -1$, por un lado, y $x = 0, y = 1, u = 0, v = 0$ por otro. Calcular $\partial u/\partial x$ en $x = 1, y = -1$, y en $x = 0, y = 1$, si existe.

En lo que resta de capítulo estudiaremos otras consecuencias del teorema de la función inversa que son muy útiles, por ejemplo, en la teoría de variedades diferenciables, como veremos más adelante.

El siguiente resultado nos dice que una aplicación diferenciable de $\mathbb{R}^n$ en $\mathbb{R}^m$ cuya diferencial en un punto tenga rango m (máximo) puede verse en esencia (salvo composición con un difeomorfismo de $\mathbb{R}^n$) una proyección lineal sobre las m últimas coordenadas.

Teorema 14.11. *Sean A un subconjunto abierto de $\mathbb{R}^n$, $m \leq n$, $a \in A$, y $f : A \to \mathbb{R}^m$ una aplicación de clase C^p, y supongamos que*

$$\textit{rango}\,(Df(a)) = m.$$

Entonces existen abiertos U y V de $\mathbb{R}^n$, con $a \in V$, y un difeomorfismo $\varphi : U \to V$ tal que

$$f \circ \varphi = P,$$

en U, donde $P : \mathbb{R}^n \to \mathbb{R}^m$ está definida por

$$P(x_1, ..., x_n) = (x_{n-m+1}, ..., x_n).$$

Cuando $m = 1$ existe otra interpretación geométrica interesante de este resultado: componiendo con el difeomorfismo φ podemos *estirar* los conjuntos de nivel de f, consiguiendo así que los conjuntos de nivel de $f \circ \varphi : U \to \mathbb{R}$ sean hiperplanos paralelos a $x_n = 0$.

Demostración. Por hipótesis, la matriz de $Df(a)$tiene m columnas linealmente independientes. Salvo una permutación de las variables (operación que es un isomorfismo lineal y en particular un difeomorfismo de clase C^∞ de $\mathbb{R}^n$), puede suponerse que se trata de las m últimas columnas de $Df(a)$, es decir, los vectores

$$\frac{\partial f}{\partial x_{n-m+1}}(a), ..., \frac{\partial f}{\partial x_n}(a)$$

son linealmente independientes.

Definamos $\Phi : A \to \mathbb{R}^n$ por

$$\Phi(x_1, ..., x_n) = (x_1, ..., x_{n-m}, f_1(x), ..., f_m(x)).$$

Es claro que

$$\det D\Phi(a) = \det\left(\frac{\partial f}{\partial x_{n-m+1}}(a), ..., \frac{\partial f}{\partial x_n}(a)\right) \neq 0.$$

Por tanto, por el teorema de la función inversa, Φ es localmente invertible: existen U, V abiertos de $\mathbb{R}^n$ con $a \in V$ tales que $\Phi : U \to V$ es un difeomorfismo de clase C^p. Sea $\varphi = \Phi^{-1}$. Entonces

$$\Phi \circ \varphi = I,$$

y por tanto, para cada $x = (x_1, ..., x_n) \in U$,

$$f_1 \circ \varphi(x) = x_{n-m+1}, ...f_m \circ \varphi(x) = x_n,$$

es decir $f \circ \varphi = P$ en U. □

El siguiente corolario nos dice que la propiedad de una diferencial de tener rango máximo se conserva localmente.

Corolario 14.12. *Si $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es de clase C^1 y*

$$\textit{rango}\,(Df(a)) = m$$

entonces existe V entorno abierto de a tal que

$$\textit{rango}\,(Df(x)) = m$$

para todo $x \in V$.

Demostración. Por el Teorema anterior, $f = P \circ \varphi^{-1}$ en V. Como φ^{-1} es un difeomorfismo su diferencial es un isomorfismo lineal en todo punto, y como P es una proyección lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$, su diferencial es ella misma y es sobreyectiva en todo punto. Por tanto,

$$Df(x) = P \circ D\varphi^{-1}(x)$$

es una aplicación lineal sobreyectiva de $\mathbb{R}^n$ en $\mathbb{R}^m$, es decir $Df(x)$ tiene rango m en todo punto $x \in V$. □

Corolario 14.13. *Si $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ es de clase C^1 y*

$$\textit{rango}\,(Df(a)) = m$$

entonces existe U entorno abierto de a tal que $f(W)$ es abierto en $\mathbb{R}^m$ para todo V abierto de contenido en U.

En particular, si $Df(x)$ tiene rango m para todo $x \in A$, entonces f es una aplicación abierta.

Demostración. Como $f \circ \varphi = P$ es una aplicación abierta (ver el problema 5.13), su composición con el difeomorfismo φ^{-1}, es decir, $f = P\varphi^{-1}$, también es abierta de U en $\mathbb{R}^m$.

Supongamos ahora que $Df(x)$ tiene rango máximo para cada $x \in A$, y veamos que $f : A \to \mathbb{R}^m$ es abierta. En efecto, si $G \subseteq A$ es abierto, entonces, para cada $y = f(x) \in f(G)$, por lo que acabamos de ver x tiene un entorno en el que f es abierta, luego existe $r > 0$ tal que $B(x, r) \subseteq G$ y $f(B(x, r))$ es abierto. Luego $f(B(x, r))$ es un entorno abierto de y que está contenido en $f(G)$. □

Corolario 14.14. *Sean A un abierto de $\mathbb{R}^{n+k}$, $a \in A$, $F : A \to \mathbb{R}^k$ de clase C^p, con $F(a) = 0$. Supongamos que*

$$\textit{rango}\,(Df(a)) = k.$$

Entonces existe un entorno abierto U de a en A, un abierto V de $\mathbb{R}^n$ y una aplicación $\psi : V \to U$ de clase C^p tal que rango $(D\psi(v)) = n$ para todo $v \in V$, y

$$\psi(V) = \{x \in U : F(x) = 0\}.$$

Esto significa que en un entorno de a el conjunto $F^{-1}(0)$ posee una parametrización regular de clase C^p por un abierto de $\mathbb{R}^n$. Este resultado será de gran utilidad en el estudio de las variedades diferenciables que haremos en el capítulo siguiente.

Demostración. Por el Teorema 14.11, existe un difeomorfismo $\varphi : W \to U$ entre dos abiertos de $\mathbb{R}^{n+k}$, con $a \in U$, tal que

$$F \circ \varphi(x_1, ..., x_{n+k}) = P(x_1, ..., x_{n+k}) = (x_{n+1}, ..., x_{n+k})$$

en W. Por tanto $F = P \circ \varphi^{-1}$, y en particular $F(x) = 0$ si y sólo si las últimas k coordenadas de $\varphi^{-1}(x)$ son todas cero.

Sea $V = \{(x_1, ..., x_n) \in \mathbb{R}^n : (x_1, ..., x_n, 0, ..., 0) \in W\}$, que es abierto (por ser $V = j^{-1}(W)$, donde j es la inyección canónica de $\mathbb{R}^n$ en $\mathbb{R}^{n+k}$, que es continua), y que contiene al punto cuyas coordenadas son las n primeras coordenadas de $\varphi^{-1}(a)$. Definamos $\psi : V \to U$ por

$$\psi(x_1, ..., x_n) = \varphi(x_1, ..., x_n, 0, ..., 0).$$

Como $D\varphi(w)$ tiene rango $n+k$ para todo $w \in W$ (por ser difeomorfismo de W en U), es inmediato que $D\psi(x_1,...,x_n)$ tiene rango n para todo $v \in V$ (las n columnas de su matriz son las primeras n columnas de la matriz de $D\varphi(v,0)$, que son linealmente independientes al tener ésta rango máximo).

Usando que $F = P \circ \varphi^{-1}$ se comprueba inmediatamente que $\psi(V) = \{x \in U : F(x) = 0\}$. $\square$

Por su parte, el resultado siguiente nos indica que todo conjunto de $\mathbb{R}^{n+k}$ que pueda parametrizarse de manera regular, localmente, por abiertos de $\mathbb{R}^n$ (tal y como el conjunto $F^{-1}(0)$ del corolario anterior) puede verse también, salvo permutación de las coordenadas de $\mathbb{R}^{n+k}$, como la gráfica de una función de un abierto de $\mathbb{R}^n$ en $\mathbb{R}^k$.

Corolario 14.15. *Sean A abierto de $\mathbb{R}^n$, $f : A \to \mathbb{R}^{n+k}$ de clase C^p, $a \in A$, $b = f(a) = (b_1,...,b_{n+k})$. Supongamos que*

$$\textit{rango}\,(Df(a)) = n.$$

Entonces, salvo permutación de las coordenadas de $\mathbb{R}^{n+k}$, existen U entorno de a en A, V entorno abierto de $(b_1,...,b_n)$ en $\mathbb{R}^n$ y una función $\varphi : V \to \mathbb{R}^k$ de clase C^p tal que $f(U)$ coincide con la gráfica de φ, es decir,

$$(x_1,...,x_{n+k}) \in f(U) \iff (x_1,...,x_n) \in V,\ (x_{n+1},...,x_{n+k}) = \varphi(x_1,...,x_n).$$

Demostración. Por hipótesis la matriz de $Df(a)$ tiene n filas linealmente independientes que, salvo permutación de las coordenadas, podemos suponer que son las n primeras. Es decir,

$$\det\left(\frac{\partial(f_1,...,f_n)}{\partial(x_1,...,x_n)}(a)\right) \neq 0.$$

Entonces, por el teorema de la función inversa, la aplicación

$$(x_1,...,x_n) \mapsto \psi(x) := (f_1(x),...,f_n(x))$$

es un difeomorfismo de clase C^p de un entorno abierto U de $(a_1,...,a_n)$ en un entorno abierto V de $(b_1,...,b_n)$. Así, para cada punto $(x_1,...,x_n) \in V$, el punto $f \circ \psi^{-1}(x_1,...,x_n)$ tiene $x_1,...,x_n$ como n primeras coordenadas. Basta definir entonces

$$\varphi(x) = \varphi(x_1,...,x_n) = (f_{n+1} \circ \psi^{-1}(x),...,f_{n+k} \circ \psi^{-1}(x))$$

para expresar $f(U)$ como la gráfica de φ. $\square$

El siguiente resultado es la contrapartida del teorema 14.11: si la diferencial de una función $f : \mathbb{R}^k \to \mathbb{R}^n$ tiene rango $k \leq n$ en un punto a, entonces f puede verse como una inyección lineal (salvo composición con un difeomorfismo) en un entorno de a.

Teorema 14.16. *Sean A un abierto de $\mathbb{R}^k$, $k \leq n$, $a \in A$, $f : A \to \mathbb{R}^n$ una aplicación de clase C^p, y supongamos que*

$$\textit{rango}\,(Df(a)) = k.$$

Entonces existen abiertos U y V de $\mathbb{R}^n$ con $f(a) \in U$, $(a,0) \in V$ y un difeomorfismo $\psi : U \to V$ tal que

$$\psi \circ f = J$$

en $W := \{x \in \mathbb{R}^k : (x,0) \in V\}$, donde $J : \mathbb{R}^k \to \mathbb{R}^n$ es la inyección lineal definida por

$$J(x_1,...,x_k) = (x_1,...,x_k,0,...,0).$$

Demostración. Por hipótesis, existen k filas en la matriz de $Df(a)$ que son linealmente independientes. Salvo permutación de las variables de $\mathbb{R}^n$ podemos suponer que se trata de las k primeras filas, es decir, que los vectores

$$\nabla f_1(a), ..., \nabla f_k(a)$$

son linealmente independientes. Definamos $\Phi : A \times \mathbb{R}^{n-p} \to \mathbb{R}^n$ por

$$\Phi(x_1, ..., x_n) = f(x_1, ..., x_k) + (0, ..., 0, x_{k+1}, ..., x_n).$$

Es inmediato que

$$\det D\Phi(a_1, ..., a_k, 0, ..., 0) = \det \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_k}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1}(a) & \dots & \frac{\partial f_k}{\partial x_k}(a) \end{pmatrix} \neq 0.$$

Por el teorema de la función inversa existen U entorno abierto de $f(a)$ en $\mathbb{R}^n$ y V entorno abierto de $(a, 0) = (a_1, ..., a_k, 0..., 0)$ en $\mathbb{R}^n$ tales que $\Phi : V \to U$ es un difeomorfismo C^p. Sea $\psi = \Phi^{-1}$. Puesto que $\psi \circ \Phi = I$ en U, se deduce que

$$\begin{aligned} J(x) = (x_1, ..., x_k, 0, ..., 0) = \psi \circ \Phi(x_1, ..., x_k, 0, ..., 0) = \\ = \psi(f(x_1, ..., x_k) + (0, ..., 0)) = \psi(f(x_1, ..., x_k)) \end{aligned}$$

para todo $(x_1, ..., x_k)$ con $(x_1, ..., x_k, 0, ..., 0) \in V$. □

Corolario 14.17. *Si $f : A \subseteq \mathbb{R}^k \to \mathbb{R}^n$ es de clase C^p, $k \leq n$, y*

$$\text{rango}\,(Df(a)) = k$$

entonces f es inyectiva en un entorno abierto W de a, y su inversa f^{-1} es continua de $f(W) \subset \mathbb{R}^n$ en $W \subset \mathbb{R}^k$.

En particular, si $Df(x)$ tiene rango k para todo $x \in A$, f es localmente inyectiva.

Demostración. Basta aplicar el teorema anterior: $f = \psi^{-1} \circ J$ es inyectiva por ser composición de dos inyecciones. Además $f^{-1} : f(W) \to W$ viene dada por

$$f^{-1} = (g_1, ..., g_k),$$

donde $g_j = \psi_j$, $j = 1, ..., k$, luego f^{-1} es continua por serlo sus funciones componentes. □

14.1. Problemas

Problema 14.1. Probar que si $f : I \subseteq \mathbb{R} \to \mathbb{R}$ es una función derivable en todo punto del intervalo abierto I y $f'(t) \neq 0$ para todo $t \in I$, entonces f es estrictamente creciente o estrictamente decreciente en todo I, y en particular es inyectiva. *Indicación:* usar el teorema del valor medio.

Problema 14.2. a) Estúdiese si la función $f(x, y) = \left(\frac{x^2-y^2}{x^2+y^2}, \frac{xy}{x^2+y^2}\right)$ tiene inversa en un entorno de (0,1).

b) Pruébese que la función $f(x, y, z) = (z\cos(xy),\ z\sin(xy),\ x + z)$ admite una inversa local g alrededor del punto $(x, y, z) = (1, 0, 1)$; calcúlese $Dg(f(1, 0, 1))$.

c) Pruébese que la función $f(x, y) = (e^x \cos y,\ e^x \sin y)$ es localmente invertible en un entorno de cada punto, pero no lo es globalmente.

Problema 14.3. Sea $\varphi(r,\theta) = (r\cos\theta,\, r\sin\theta)$. Pruébese que φ es localmente inversible en cada punto (r,θ) con $r > 0$, de clase C^∞. Calcúlese una inversa de φ en el conjunto $A = \{(x,y) \in \mathbb{R}^2,\, x > 0,\, y > 0\}$. Calcúlense las derivadas parciales $\frac{\partial r}{\partial x}$, $\frac{\partial r}{\partial y}$, $\frac{\partial \theta}{\partial x}$, $\frac{\partial \theta}{\partial y}$.

Problema 14.4. Resolver las ecuaciones

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta \end{aligned}$$

para r, θ, ϕ en términos de x, y, z. ¿Dónde es diferenciable la función inversa? Comparar con las conclusiones derivadas de aplicar el teorema de la función inversa. A este cambio de coordenadas se le llama *coordenadas esféricas*.

Problema 14.5. Sea $g : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definida por $g(x,y,z) = (e^{2y} + e^{2z}, e^{2x} - e^{2z}, x - y)$

1. Pruébese que g es diferenciable en todo punto de $\mathbb{R}^3$ y que admite inversa diferenciable en un entorno de cada punto.
2. ¿Admite g inversa global?

Problema 14.6. Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por $f(x,y) = y\cos x$. Pruébese que f define una función implícita diferenciable $y = h(x)$ en un entorno del punto (0,0). Hállese $Dh(0)$. ¿Tiene h inversa diferenciable en un entorno de 0?

Problema 14.7. Considérese el sistema de ecuaciones:

$$\begin{aligned} 3x + y - z + u^2 &= 0 \\ x - y + 2z + u &= 0 \\ 2x + 2y - 3z + 2u &= 0 \end{aligned}$$

Determínese si el sistema puede resolverse en las siguientes condiciones:

1. para (x,y,z) en función de u;
2. para (x,y,u) en función de z;
3. para (x,z,u) en función de y.

Problema 14.8. Sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $f(x,y,z,u,v) = (u + v + x^2 - y^2 + z^2, u^2 + v^2 + u - 2xyz)$. Pruébese que f define, en un entorno de $(0,0,0,-\frac{1}{2},\frac{1}{2})$, una función implícita diferenciable

$$(u,v) = (h_1(x,y,z), h_2(x,y,z)),$$

y calcúlese $Dh(0,0,0)$.

Problema 14.9. Supóngase que $F(x,y,z) = 0$ define implícitamente cada variable como función de las otras dos. Probar que

$$\frac{\partial y}{\partial x}\frac{\partial z}{\partial y}\frac{\partial x}{\partial z} = -1.$$

El término $\partial y/\partial x$ significa $\partial g/\partial x$, donde $y = g(x,z)$, y análogamente para los demás.

Problema 14.10. Sea $f : \mathbb{R} \to \mathbb{R}$ definida por

$$f(x) = \begin{cases} \frac{x}{2} + x^2 \sin(\frac{1}{x}) & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases}$$

Probar que f es diferenciable en todo punto, con $f'(0) \neq 0$, pero f no es inyectiva en ningún entorno de 0. ¿Contradice esto el teorema de la función inversa?

Problema 14.11. Supongamos que $f : \mathbb{R}^n \to \mathbb{R}^n$ es diferenciable e invertible en un entorno de un punto a, y que $\det Df(a) = 0$. Probar que f^{-1} no es diferenciable en $f(a)$.

Problema 14.12. Supóngase que F satisface las hipótesis del teorema de la función implícita, salvo que F no necesariamente es de clase C^1 en las dos variables x e y, sino solamente en y, es decir, que todas las derivadas parciales $\partial F_i/\partial y_j$ son continuas. Demostrar directamente (sin usar el teorema de la función inversa) que existen entornos U de a y V de b y una única función continua $\varphi : U \to V$ tal que $F(x, \varphi(x)) = 0$. *Indicación:* seguir estos pasos:

1. Puede suponerse $(a, b) = (0, 0)$.

2. Sea $L = \partial F/\partial y(0, 0)$. Probar que existen $r, s > 0$ tal que para todo $x \in B(x, s)$ la aplicación $G_x(y) = y - L^{-1} \circ F(x, y)$ tiene un único punto fijo, que denotaremos y_x, en $B(0, r)$.

3. Escribir $\varphi(x) = y_x$, y probar que $F(x, \varphi(x)) = 0$, y que φ es continua.

Problema 14.13. Sea $f : R^2 \to \mathbb{R}^2$ una aplicación de clase C^1, y supongamos que

$$\frac{\partial f_1}{\partial x} = \frac{\partial f_2}{\partial y}, \ \frac{\partial f_1}{\partial y} = -\frac{\partial f_2}{\partial x}$$

(a estas igualdades se les llama *ecuaciones de Cauchy-Riemann*, y surgen naturalmente en la teoría de variable compleja. Probar que $\det Df(x, y) \neq 0$ si y sólo si $Df(x, y) = 0$, y por tanto que f es localmente invertible en un entorno de (x, y) si y sólo si $Df(x, y) \neq 0$. Probar también que en este caso la función inversa f^{-1} también satisface las ecuaciones de Cauchy-Riemann.

Problema 14.14. Consideremos la aplicación $\mathcal{J} : GL(\mathbb{R}^n) \to GL(\mathbb{R}^n)$, $A \mapsto \mathcal{J}(A) = A^{-1}$. La Proposición 14.2 nos dice que $\mathcal{J}$ es de clase C^∞, pero no cuál es la derivada de $\mathcal{J}$. Demostrar que

$$D\mathcal{J}^{-1}(A)(B) = -A^{-1} \circ B \circ A^{-1}.$$

Problema 14.15. Sea $h : \mathbb{R}^2 \to \mathbb{R}$ la función definida por $h(x, y) = x^2 + y^3 + xy + x^3 + ay$, donde a es un parámetro real.

1. ¿Para qué valores de a la ecuación $h(x, y) = 0$ define y como función implícita de clase infinito de x, en un entorno de $(0, 0)$? ¿Define la anterior ecuación a x como función implícita diferenciable de y en un entorno de $(0, 0)$, para algún valor de a?

2. Sea $y = f(x)$ la función implícita determinada por $h(x, y) = 0$, definida en un cierto abierto U con $0 \in U$. Calcúlese el valor de a para que el polinomio de Taylor de segundo grado de f en el origen valga 1 en el punto $x = 1$. ¿Para qué valores de a tiene f un extremo en $x = 0$?

3. Sea $F : U \times \mathbb{R} \to \mathbb{R}^2$ la función $F(x, y) = (e^{x+y} + x^2 - 1, f(x) + y \cos x)$. Demuéstrese que F admite función inversa diferenciable en un entorno de $(0, 0)$ y que $G = F \circ F + F^{-1}$ es diferenciable en $(0, 0)$ y calcúlese $DG(0, 0)$.

Problema 14.16. Calcúlese el desarrollo de Taylor hasta el orden 2 de la función $z = f(x, y)$, definida implícitamente por $2x^2 + 2y^2 + z^2 - 8xz - z + 8 = 0$ en un entorno de $(2, 0, 1)$.

Problema 14.17. Calcúlense las constantes a y b sabiendo que

1. El sistema
$$\begin{cases} F(x, y, z) = x + a \sin y - z^3 = 0 \\ G(x, y, z) = x - a^2y^2 + e^{bz} = 1 \end{cases}$$
define implícitamente las funciones de clase infinito $y = f(x)$, $z = g(x)$, en un entorno de $(0, 0, 0)$.

2. El polinomio de Taylor de segundo grado de g en 0 vale 0 para $x = b^2$.

3. La función $h = f - g$ tiene un máximo relativo en $x = 0$.

Problema 14.18. Sea $a > 0$, definimos $f : B(0, a) \longrightarrow \mathbb{R}^n$ como $f(x) = \frac{ax}{\sqrt{a^2 - x_1^2 - \cdots - x_n^2}}$. Pruébese que es biyectiva, diferenciable y con inversa diferenciable (difeomorfismo).

Problema 14.19. Demuéstrese que las ecuaciones:
$$\begin{cases} x^2 - y^2 + uv &= 0 \\ xy + u^2 - v^2 &= 0 \end{cases}$$
definen funciones implícitas $u(x, y)$, $v(x, y)$ en un entorno del punto $(x, y, u, v) = (0, 1, 1, 1)$. Considérese ahora la función: $\varphi(x, y) = (u(x, y), v(x, y))$, definida en un entorno de $(0, 1)$.

1. Calcúlese $D\varphi(0, 1)$. ¿Admite φ inversa local alrededor de $(0, 1)$.

2. Calcúlese la derivada direccional de φ^{-1} en $(1, 1)$, según la dirección $(1/\sqrt{2}, -1/\sqrt{2})$.

3. Si $\gamma(t) = (t, t^2)$, ¿cuánto vale $(\varphi^{-1} \circ \gamma)'(1)$?

Problema 14.20. Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la función $f(x, y) = (x^2 + y^2, 2xy)$.

1. Hállese la imagen mediante la función f de las circunferencias $x^2 + y^2 = R^2$ y las hipérbolas $xy = c$.

   En cada uno de los casos anteriores hágase un dibujo indicando cómo se mueven los puntos de la imagen cuando los puntos del original describen las figuras indicadas.

2. Hállese la imagen mediante f del disco $\{(x, y) \in \mathbb{R}^2 : (x - 1)^2 + (y - 1)^2 \leq 1\}$.

3. Hállese la imagen de f.

4. Encuéntrese un subconjunto conexo A de $\mathbb{R}^2$ tal que $f(A) = f(\mathbb{R}^2)$ y $f|_A$ sea inyectiva.

5. Hállese la inversa de la función $f|_A$ y estúdiese si es diferenciable.

6. Calcúlese la matriz jacobiana de f en un punto arbitrario (x_0, y_0).

7. Si $(x_0, y_0) \in \mathbb{R}^2$ es tal que el jacobiano de f en ese punto es no nulo, hállese la matriz jacobiana de una inversa local de f en $(x_0^2 + y_0^2, 2x_0y_0)$.

8. Demuéstrese que el conjunto $\{(x,y) \in \mathbb{R}^2 : f$ tiene jacobiano no nulo en $(x,y)\}$ se puede poner como la unión de cuatro conjuntos, Δ_1, Δ_2, Δ_3, Δ_4, abiertos conexos.

9. Muéstrese que para $i = 1,2,3,4$ la función $f|_{\overline{\Delta_i}}$ es inyectiva.

10. Calcúlense las inversas de las funciones del apartado anterior.

11. Sea $i_0 \in \{1,2,3,4\}$ tal que $\Delta_{i_0} \cap A \neq \emptyset$. Hállese $f(\Delta_{i_0})$ y compruébese que es abierto.

12. Demuéstrese que la inversa de $f|_{\overline{\Delta_{i_0}}}$ es de clase C^∞ en $f(\Delta_{i_0})$ y hállese su matriz jacobiana.

13. Compárese el apartado anterior con los apartados d), e) y g).

14. ¿Es posible definir una inversa local diferenciable de f en un entorno del punto (2,2)?

Problema 14.21. Sea $f(x,y) = x^3 + x + y$.

1. Represéntense las curvas de nivel de la función f para valores próximos a 0.

2. Sea $G : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ la función $G(x,y) = (x, f(x,y))$. Demuéstrese que G es inyectiva.

3. Represéntense las imagenes mediante G de las curvas del apartado a).

4. Calcúlese el jacobiano de G.

5. Demuéstrese que $D = G(\mathbb{R}^2)$ es abierto y que la inversa de G, que denotaremos H, es de clase C^∞ en D.

6. Compruébese que $f(H(x,y)) = y, \forall (x,y) \in D$.

7. Hágase un estudio análogo para la función $f(x,y) = x^2 + y^2$ considerándola definida en $\mathbb{R}^2 \setminus (0,0)$.

8. ¿Qué ocurriría en g) si añadiésemos (0,0) al dominio de f? Explíquese la respuesta.

9. Generalícense los resultados obtenidos.

Problema 14.22. Sea $F : \mathbb{R}^3 \longrightarrow \mathbb{R}$ la función

$$F(x,y,z) = x^2 + y^2 + z^2 - 2xz - 4$$

y sea $V = \{(x,y,z) \in \mathbb{R}^3 : \ F(x,y,z) = 0\}$.

1. Represéntese V gráficamente.

2. Demuéstrese que $\{y \in \mathbb{R} : \ F(x,y,z) = 0\} = [-2,2]$.

3. Demuéstrese que para cada punto $(x,y) \in \mathbb{R} \times [-2,2]$ existe un $z \in \mathbb{R}$ tal que $F(x,y,z) = 0$.

4. ¿En qué casos el z obtenido en el apartado anterior es único?

5. Encuéntrese una función $f : \mathbb{R} \times [-2, 2] \longrightarrow \mathbb{R}$, de clase C^∞ en $\mathbb{R} \times (-2, 2)$, tal que para cada $(x, y) \in \mathbb{R} \times [-2, 2]$ se verifique que $F(x, y, f(x, y)) = 0$.

6. ¿Es única la función del apartado anterior?

7. Hállese la matriz jacobiana de F.

8. Sea $(x_0, y_0, z_0) \in V$ tal que $\frac{\partial F}{\partial z}(x_0, y_0, z_0) \neq 0$. Demuéstrese que existe una única función φ de clase C^∞ en un entorno U de (x_0, y_0), tal que $\varphi(x_0, y_0) = z_0$ y $F(x, y, \varphi(x, y)) = 0$ $\forall (x, y) \in U$. Compárese con el apartado f).

9. ¿Es posible encontrar una función φ de clase C^∞ en un entorno U de (0,2) tal que $\varphi(0, 2) = 0$ y $F(x, y, \varphi(x, y)) = 0$ $\forall (x, y) \in U$?

10. Resuélvanse las cuestiones análogas cambiando x por z e y por z en el apartado anterior.

Problema 14.23. Se considera el sistema de ecuaciones

$$\begin{cases} x^2 - yu &= 0 \\ xy + uv &= 0 \end{cases}$$

1. Usando el teorema de la función implícita, indíquese en qué condidiones se puede resolver el sistema en u y v.

2. Resuélvase el sistema directamente y compárese con el apartado a).

Problema 14.24. Demuéstrese que las ecuaciones

$$\begin{cases} x^2 - y^2 - u^3 + v^2 + 4 &= 0 \\ 2xy + y^2 - 2u^2 + 3v^4 + 8 &= 0 \end{cases}$$

determinan funciones de clase C^∞, $u(x, y)$ y $v(x, y)$, en un entorno de (2,-1), tales que $u(2, -1) = 2$ y $v(2, -1) = 1$. Calcular $\frac{\partial u}{\partial x}$ en (2,-1).

Problema 14.25. El punto (1,-1,1) pertenece a las superficies $x^3(y^3 + z^3) = 0$ y $(x - y)^3 - z^2 = 7$. Demuéstrese que, en un entorno de este punto, la intersección de las dos superficies determina una curva de clase C^∞ de ecuaciones $y = f(x)$, $z = g(x)$. Calcúlese la tangente a dicha curva en $x = 1$.

Capítulo 15

Variedades diferenciables en $\mathbb{R}^n$. Extremos condicionados.

En los capítulos precedentes han aparecido con frecuencia curvas y superficies en $\mathbb{R}^n$ que se han descrito de diversas maneras. Por ejemplo, muchas curvas en el plano $\mathbb{R}^2$ pueden describirse mediante una ecuación del tipo $F(x, y) = 0$. Un caso particular de esta situación es la gráfica de una función, donde se tiene $y = f(x)$ (es decir, $F(x, y) = 0$ para $F(x, y) = f(x) - y$). También pueden venir definidas en forma paramétrica, es decir, $\gamma(t) = (x(t), y(t))$, $t \in I$.

Una curva en $\mathbb{R}^3$ puede venir definida también en forma paramétrica, $\gamma(t) = (x(t), y(t), z(t))$, $t \in I$. Sin embargo, las ecuaciones implícitas del tipo $F(x, y, z) = 0$ no definen en general curvas, sino superficies, en $\mathbb{R}^3$. La intersección de dos de estas superficies es en general una curva, que puede así ser descrita en forma implícita como un par de ecuaciones

$$\begin{cases} F_1(x, y, z) = 0 \\ F_2(x, y, z) = 0. \end{cases}$$

En general, y por analogía con lo que ocurre con los sistemas de ecuaciones lineales, uno estaría tentado de definir una superficie de dimensión k en $\mathbb{R}^n$, bien como conjunto de ceros de una función $F : \mathbb{R}^n \to \mathbb{R}^{n-k}$, bien como imagen de una función $\varphi : U \subseteq \mathbb{R}^k \to \mathbb{R}^n$. Además, estamos interesados en obtener superficies suaves, en el sentido de que posean planos tangentes en cualquiera de sus puntos, lo que (teniendo presente el caso en que la superficie es la gráfica de una función $z = f(x, y)$) aconseja que estas funciones F o φ sean diferenciables.

Sin embargo, una definición así no sería en absoluto satisfactoria, puesto que, incluso exigiendo que F sea diferenciable de clase C^∞, permitiría que cualquier conjunto cerrado de $\mathbb{R}^n$ fuera una hipersuperficie de dimensión $n - 1$, lo que contradice cualquier idea intuitiva que pueda tenerse de lo que es o debe ser una superficie. En efecto, para todo cerrado C de $\mathbb{R}^n$ existe una función $F : \mathbb{R}^n \to \mathbb{R}$ de clase C^∞ tal que $C = F^{-1}(0)$ (ver el problema 11.12). Por otra parte, puesto que la imagen de una función $\varphi : \mathbb{R}^k \to \mathbb{R}^n$ puede bien ser un solo punto, tampoco $\varphi(\mathbb{R}^k)$ corresponderá, en general, a nuestra idea de superficie de dimensión k.

Hace falta, pues, pedir algo más a F o a φ para que la definición apuntada anteriormente se corresponda con el concepto intuitivo que tenemos de superficie. Nuevamente, un análisis más cuidadoso de lo que sucede con los sistemas de ecuaciones lineales y las aplicaciones lineales, más los teoremas estudiados en el capítulo anterior, nos sugieren que la definición será buena si se añade que F o φ tengan rango máximo.

Definición 15.1. Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ una función de clase C^1. Se dice que $a \in U$ es un *punto regular* de la función f si $Df(a)$ tiene rango máximo (es decir, rango n cuando $n \leq m$, y rango m cuando $n \geq m$). Se dice que b es un *valor regular* de f si x es un punto regular de f para todo x tal que $f(x) = b$. Los puntos que no son regulares se llaman *puntos críticos*, y la imagen por f de cualquier punto crítico se llama *valor crítico* de f.

Se dice que una aplicación diferenciable f es *regular* si todos los puntos de su dominio son puntos regulares, es decir, si Df tiene rango máximo en todos los puntos.

Definición 15.2. Sea M un subconjunto de $\mathbb{R}^n$. Se dice que M es una variedad diferenciable (respectivamente, de clase C^p) de dimensión k de $\mathbb{R}^n$ si para cada punto $a \in M$ existen un entorno abierto W de a, un abierto V de $\mathbb{R}^k$, y una aplicación regular $\varphi : V \to W$ de clase C^1 (resp. de clase C^p) tal que $M \cap W = \varphi(V)$.

Se dice en este caso que cada $\varphi : V \to W$ es una parametrización regular de $M \cap W$.

Observación 15.1. Por el Corolario 14.17 sabemos que una aplicación φ de estas características es localmente inyectiva, por lo que siempre puede suponerse (reduciendo los abiertos V y W si es necesario) que $\varphi : V \to W$ es inyectiva.

Ejemplo 15.1. Sea $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función de clase C^p. Entonces su gráfica $G_f = \{(x, z) \in \mathbb{R}^{n+1} : z = f(x), x \in U\}$ es una variedad de clase C^p y dimensión n en $\mathbb{R}^{n+1}$.

Más en general, si $\Phi : U \subseteq \mathbb{R}^k \to \mathbb{R}^{n-k}$ es diferenciable de clase C^p entonces su gráfica G_Φ es una variedad de dimensión k en $\mathbb{R}^n$. Basta considerar la parametrización definida por $\varphi(x) = (x, \Phi(x))$, $x \in U$.

Teorema 15.2. *Sea M un subconjunto de $\mathbb{R}^n$. Entonces M es una variedad de clase C^p y dimensión k si y sólo si para cada punto $a \in M$ existen un entorno abierto U de a y una función $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^{n-k}$ regular de clase C^p tales que*

$$M \cap U = \{x \in U : F(x) = 0\}.$$

Demostración. ($\Rightarrow$): Sean $a \in M$, $\varphi : V \subseteq \mathbb{R}^k \to W \subseteq \mathbb{R}^n$ parametrización regular como en la definición de variedad. Por el Corolario 14.15 sabemos que, reduciendo si es preciso los abiertos V y W, $\varphi(V)$ es, salvo permutación de coordenadas, una gráfica de una función $g : U \subseteq \mathbb{R}^k \to \mathbb{R}^{n-k}$ de clase C^p, es decir

$$x \in M \cap W = \varphi(V) \iff x = (u, g(u)), u \in U.$$

Definamos entonces $F : U \times \mathbb{R}^{n-k} \subseteq \mathbb{R}^k \times \mathbb{R}^{n-k} = \mathbb{R}^n \to \mathbb{R}^{n-k}$ por

$$F(u, z) = z - g(u).$$

Es obvio que $M \cap W = \{x : F(x) = 0\}$, que F es de clase C^p, y que DF tiene rango $n - k$, máximo, en todos los puntos de $U \times \mathbb{R}^{n-k}$.

($\Leftarrow$): es el Corolario 14.14. $\square$

Corolario 15.3. *Sea $F : \mathbb{R}^n \to \mathbb{R}^{n-k}$ una función diferenciable de clase C^p, y supongamos que $b \in \mathbb{R}^{n-k}$ es un valor regular de F. Entonces $M = F^{-1}(b)$ es una variedad de clase C^p y dimensión k en $\mathbb{R}^n$.*

Demostración. Sea $a \in M = F^{-1}(b)$. Como b es un valor regular de F, $DF(a)$ tiene rango máximo. Por el Corolario 14.12 sabemos que $DF(x)$ tiene rango máximo en un entorno U de a en $\mathbb{R}^n$. Es decir, $F = F_{|_U} : U \subset \mathbb{R}^n \to \mathbb{R}^{n-k}$ es una función regular de clase C^p, luego también lo es $G : U \to \mathbb{R}^{n-k}$, $G(x) = F(x) - b$, y es obvio que $M \cap U = \{x \in U : G(x) = 0\}$. Por el Teorema anterior resulta que M es una variedad de dimensión k y clase C^p de $\mathbb{R}^n$. $\square$

Ejemplo 15.2. La esfera $x^2+y^2+z^2=1$ de $\mathbb{R}^3$ es una variedad diferenciable de dimensión 2.

Ejercicio 15.1. Probar directamente que la esfera es una variedad diferenciable, expresándola como unión de imágenes de parametrizaciones regulares por abiertos de $\mathbb{R}^2$.

Observación 15.4. *Sea M una variedad diferenciable de dimensión k en $\mathbb{R}^n$. Para cada punto $a \in M$ sabemos por la definición que existen abiertos V de $\mathbb{R}^k$ y W de $\mathbb{R}^n$ y una aplicación inyectiva y regular $\varphi : V \to W$ tal que $\varphi(V) = W \cap M$. Consideremos su inversa, $\psi = \varphi^{-1} : M \cap W \to V \subset \mathbb{R}^k$. Entonces ψ es un homeomorfismo de $M \cap W$ en V.*

En efecto, $\psi = \varphi^{-1} : M \cap M = \varphi(V) \to V$ es continua por el Corolario 14.17. Y como $\psi^{-1} = \varphi$ es también continua (porque de hecho es diferenciable), resulta que ψ es un homeomorfismo.

Consideremos una variedad M de dimensión k y clase C^p en $\mathbb{R}^n$. Para cada punto $a \in M$ sabemos por la definición y la observación anterior que existen abiertos V de $\mathbb{R}^k$ y W de $\mathbb{R}^n$ y un homeomorfismo $\psi : M \cap W \to V$ tal que $\varphi = \psi^{-1} : V \to \mathbb{R}^n$ es una aplicación regular de clase C^p. Evidentemente ψ no es única (por ejemplo, componiéndola con cualquier difeomorfismo de V en sí mismo obtenemos otra aplicación con las mismas propiedades). ¿Qué relación habrá entre dos aplicaciones ψ_1 y ψ_2 con estas propiedades?

Más en general, sean W_1 y W_2 abiertos de $\mathbb{R}^n$; V_1 y V_2 abiertos de $\mathbb{R}^k$, $\psi_j : M \cap W_j \to V_j$ homeomorfismos tales que $\varphi_j := \psi_j^{-1} : V_j \to M \subset \mathbb{R}^n$ es una aplicación regular de clase C^p ($i = 1, 2$), y supongamos que $W_1 \cap W_2 \cap M \neq \emptyset$. Sea $a \in W_1 \cap W_2 \cap M$. Por el Teorema 14.16 existen abiertos U_j, A_j de $\mathbb{R}^n$, con $a \in U_j$, y difeomorfismos $\widetilde{\psi}_j : U_j \to A_j$ ($i = 1, 2$) tales que

$$\widetilde{\psi}_j \circ \varphi_j = J,$$

donde $J(x_1, ..., x_k) = (x_1, ..., x_k, 0, ..., 0)$ es la inyección canónica de $\mathbb{R}^k$ en $\mathbb{R}^n$. Reduciendo los abiertos si fuera necesario, puede suponerse $U_j = U$, $A_j = A$, $i = 1, 2$. Si denotamos $P : \mathbb{R}^n \to \mathbb{R}^k$ la proyección $P(x_1, ..., x_n) = (x_1, ..., x_k)$, se tiene que $P \circ J = I$, la identidad, y por tanto

$$P \circ \widetilde{\psi}_j \circ \varphi_j = I,$$

es decir,

$$\psi_j = \varphi_j^{-1} = P \circ \widetilde{\psi}_j$$

en $M \cap U$. Entonces, por la regla de la cadena,

$$\psi_2 \circ \psi_1^{-1} = P \circ \widetilde{\psi}_2 \circ \varphi_1$$

es diferenciable de $\psi_1(M \cap U)$ en $\psi_2(M \cap U)$, con inversa

$$(\psi_2 \circ \psi_1^{-1})^{-1} = \psi_1 \circ \psi_2^{-1} = P \circ \widetilde{\psi}_1 \circ \varphi_2$$

también diferenciable de $\psi_2(M \cap U)$ en $\psi_1(M \cap U)$. Como este argumento es válido para cualquier punto $a \in M \cap W_1 \cap W_2$, esto prueba que $\psi_2 \circ \psi_1^{-1}$ es un difeomorfismo de clase C^p de $\psi_1(M \cap W_1 \cap W_2)$ en $\psi_2(M \cap W_1 \cap W_2)$.

Observación 15.5. En muchos textos avanzados de geometría diferencial se usa la siguiente definición más general de variedad diferenciable:

Sea M un espacio métrico. Se dice que M es una variedad diferenciable de dimensión k y clase C^p, y que $\mathcal{A} = \{(U_i, \psi_i) : i \in I\}$ es un atlas (o estructura de variedad diferenciable) para

M, si para cada $a \in M$ existe $i \in I$ tal que U_i es un entorno abierto de a y $\psi_i : U_i \to \psi_i(U_i)$ es un homeomorfismo sobre un subconjunto abierto de $\mathbb{R}^k$ y si, además, para todos $i, j \in I$ con $U_i \cap U_j \neq \emptyset$ la aplicación $\psi_i \circ \psi_j^{-1}$ es un difeomorfismo de clase C^p del abierto $\psi_j(U_i \cap U_j)$ en el abierto $\psi_i(U_i \cap U_j)$ de $\mathbb{R}^k$. De las aplicaciones $\psi_j : U_j \to \psi_j(U_j)$ se dice que son cartas *de la variedad M.*

La discusión precedente muestra que toda variedad diferenciable de dimensión k en $\mathbb{R}^n$ según la Definición 15.2 es también una variedad diferenciable según la definición anterior. Sin embargo, el recíproco no es cierto, es decir, existen espacios métricos (¡incluso subconjuntos de $\mathbb{R}^n$!) que satisfacen esta definición pero no la Definición 15.2. Ver el problema 15.8. La solución a esta discrepancia de definiciones viene dada por el concepto de subvariedad de una variedad diferenciable. Aunque no entraremos en detalles técnicos que escapan de los objetivos de este curso, diremos que $\mathbb{R}^n$ tiene una estructura natural de variedad diferenciable según la definición anterior (donde todas las cartas serían la identidad), y que puede definirse con rigor el concepto de subvariedad de una variedad diferenciable. De acuerdo con esta definición de subvariedad (que no daremos aquí) resulta que un subconjunto M de $\mathbb{R}^n$ que sea variedad diferenciable de dimensión k según la definición general es una variedad según la Definición 15.2 si, y sólo si, M es subvariedad diferenciable de $\mathbb{R}^n$ (con su estructura natural).

Por otro lado, es conveniente señalar que no se pierde apenas generalidad en la teoría de variedades diferenciables si se decide utilizar exclusivamente la Definición 15.2 (como hemos hecho nosotros en este texto), puesto que, gracias a un importante teorema debido a Whitney, toda variedad diferenciable de clase C^p y dimensión k (según la definición general) es C^p difeomorfa a una subvariedad diferenciable de clase C^∞ y de dimensión k de $\mathbb{R}^m$ (que satisface por tanto 15.2), siempre y cuando m sea *suficientemente* grande (por ejemplo $m \geq 2k + 1$).

Veamos ahora cómo puede definirse el espacio tangente a una variedad en un punto de ésta. Probablemente la definición más intuitiva que pueda hacerse del espacio vectorial tangente a una variedad M en un punto a sea la del conjunto de vectores tangentes en a a todos los caminos diferenciables que pasan por a y están contenidos en M, pero hay otras que son también muy prácticas, por lo que en lugar de decidirnos por una definición concreta, veremos primero que tres de todas las posibles son equivalentes, y definiremos el espacio tangente como el conjunto de todos los vectores que satisfacen una (y por tanto todas) de estas tres propiedades equivalentes.

Teorema 15.6. *Sea M una variedad diferenciable de dimensión k en $\mathbb{R}^n$. Sea $a \in M$, y sean U entorno abierto de a en $\mathbb{R}^n$, V abierto en $\mathbb{R}^k$, $z \in V$, $a = \varphi(z)$, y $F : U \to \mathbb{R}^{n-k}$, $\varphi : V \to U$ funciones regulares tales que*

$$M \cap U = \varphi(V) = \{x \in U : F(x) = 0\}.$$

Para cualquier vector $v \in \mathbb{R}^n$, las siguientes afirmaciones son equivalentes:

1. *$v \in D\varphi(z)(\mathbb{R}^k)$.*

2. *$v \in KerDF(a)$ (es decir, $DF(a)(v) = 0$).*

3. *Existe una curva diferenciable $\gamma : (-\varepsilon, \varepsilon) \to M$ tal que $\gamma(0) = a$ y $\gamma'(0) = v$.*

Definición 15.3. Si v satisface una (y por tanto cualquiera) de estas tres propiedades, se dirá que v es un vector tangente a M en el punto a. Se denotará el espacio vectorial de todos estos v como TM_a, y se dirá que TM_a es el espacio vectorial tangente a M en a. Es decir,

$$TM_a = D\varphi(z)(\mathbb{R}^k) = \mathrm{Ker}DF(a).$$

Se define el espacio afín tangente a M en a como el espacio afín $a + TM_a$.

Observación 15.7.

1. Puesto que $D\varphi(z) : \mathbb{R}^k \to \mathbb{R}^n$ tiene rango máximo, y $k \leq n$, $TM_a = D\varphi(z)(\mathbb{R}^k)$ tiene dimensión k, la misma que M.

2. Puesto que para una F fija pueden considerarse todas las φ posibles y se mantiene la igualdad $D\varphi(\mathbb{R}^k) = \text{Ker}DF(a)$, TM_a no depende de la parametrización local φ escogida.

3. El mismo argumento, manteniendo fija φ y considerando todas las F posibles, prueba que TM_a tampoco depende de la representación implícita local $F = 0$ que hayamos escogido.

4. Puesto que $D\varphi(z) : \mathbb{R}^k \to TM_a$ es un isomorfismo lineal, si $\{e_1, ..., e_k\}$ es la base canónica de $\mathbb{R}^k$ entonces los vectores $D\varphi(z)(e_j)$, $j = 1, ..., k$, forman una base de TM_a. Es decir
$$\{\frac{\partial\varphi}{\partial x_1}(z),\ ...,\ \frac{\partial\varphi}{\partial x_k}(z)\}$$
es una base de TM_a.

Demostración del Teorema 15.6:
(1) $\iff$ (2): Como $M \cap U = \{x \in U : F(x) = 0\} = \varphi(V)$, la función $g : V \to \mathbb{R}^{n-k}$ definida por
$$g(x) = F(\varphi(x))$$
es idénticamente nula en V. Por tanto
$$0 = Dg(a) = DF(\varphi(z)) \circ D\varphi(z),$$
luego $D\varphi(z)(\mathbb{R}^k) \subseteq \text{Ker}DF(a)$. Ahora bien, como F y φ son ambas regulares, se tiene que
$$\dim \text{Ker}DF(a) = n - \dim \text{Im}DF(a) = n - (n-k) = k = \dim D\varphi(z)(\mathbb{R}^k),$$
y puesto que dos espacios vectoriales de la misma dimensión entre los cuales haya una relación de inclusión forzosamente son el mismo, se concluye que $D\varphi(z)(\mathbb{R}^k) = \text{Ker}DF(a)$.
(1) $\implies$ (3): Si $v \in D\varphi(z)(\mathbb{R}^k)$, existe $h \in \mathbb{R}^k$ tal que $v = D\varphi(z)(h)$. Definamos $\gamma : (-\varepsilon, \varepsilon) \to M$ por
$$\gamma(t) = \varphi(z + th).$$
Es obvio que γ es un camino diferenciable regular (es decir con derivada no nula en todos sus puntos) contenido en M. Se tiene además que $\gamma(0) = \varphi(z) = a$, y
$$\gamma'(0) = D\varphi(z)(h) = v.$$
(3) $\implies$ (2): Si $v = \gamma'(0)$ para un camino diferenciable $\gamma : (-\varepsilon, \varepsilon) \to M$ con $\gamma(0) = a$, tomando si es necesario un ε más pequeño, puede suponerse que $\gamma(t) \in M \cap U = \{x \in U : F(x) = 0\}$ para todo t, con lo que la función $f : (-\varepsilon, \varepsilon) \to \mathbb{R}$ definida por
$$f(t) = F(\gamma(t))$$
es idénticamente nula, y por tanto
$$0 = f'(0) = DF(\gamma(0))(\gamma'(0)) = DF(a)(v),$$
es decir, $v \in \text{Ker}DF(a)$. $\square$

Existen otras posibles definiciones del espacio tangente (ver por ejemplo el problema 15.10), aunque las tres anteriores son las más usadas.

Ejercicio 15.2. Si M y N son variedades diferenciables de $\mathbb{R}^n$ tales que $M \subset N$ en un entorno de $a \in M$, entonces
$$TM_a \subseteq TN_a.$$

Definición 15.4. Se define el espacio normal a una variedad M de $\mathbb{R}^n$ en un punto $a \in M$ como el complemento ortogonal del espacio tangente TM_a en $\mathbb{R}^n$. Si un vector v está en el espacio normal a M en a, se dirá que v es un vector normal a M en a.

Teorema 15.8. *Sea M una variedad diferenciable de dimensión k en $\mathbb{R}^n$. Sea $a \in M$, y sean U entorno abierto de a en $\mathbb{R}^n$, V abierto en $\mathbb{R}^k$, $z \in V$, $a = \varphi(z)$, y $F : U \to \mathbb{R}^{n-k}$, $\varphi : V \to U$ funciones regulares tales que*
$$M \cap U = \varphi(V) = \{x \in U : F(x) = 0\}.$$
Para cualquier vector $v \in \mathbb{R}^n$, las siguientes afirmaciones son equivalentes:

1. *v es un vector normal a M en a.*
2. *$v \in [\nabla F_1(a), ..., \nabla F_{n-k}(a)]$, es decir, existen $\lambda_j \in \mathbb{R}$, $j = 1, ..., n-k$, tales que $v = \sum_{j=1}^{n-k} \lambda_j \nabla F_j(a)$.*
3. *Para todo $j = 1, ..., k$ se tiene que $\sum_{i=1}^{n} \frac{\partial \varphi_i}{\partial x_j}(z) v_i = 0$, es decir,*
$$\langle \frac{\partial \varphi}{\partial x_j}(z), v\rangle = 0.$$

Demostración. (1) $\iff$ (2): Si $w \in TM_a$ se tiene $DF(a)(w) = 0$, o bien $\langle \nabla F_j(a), w\rangle = 0$ para $j = 1, ..., n-k$. Así w es ortogonal a cada $\nabla F_j(a)$ y por tanto w es también ortogonal al subespacio engendrado por los vectores $\nabla F_j(a)$, es decir,
$$w \perp [\nabla F_1(a), ..., \nabla F_{n-k}(a)] := E.$$
Como F es regular, los vectores $\nabla F_j(a)$ son linealmente independientes, y E tiene dimensión $n-k$. Así pues, E es ortogonal a cada vector de TM_a, y como TM_a tiene dimensión k y E tiene dimensión $n-k$, se concluye que E debe ser el complemento ortogonal de TM_a en $\mathbb{R}^n$.
(1) $\iff$ (3): Puesto que
$$\{\frac{\partial \varphi}{\partial x_1}(z),\ ...,\ \frac{\partial \varphi}{\partial x_k}(z)\}$$
es una base de TM_a, un vector $v \in \mathbb{R}^n$ será ortogonal a TM_a si y sólo si lo es a cada uno de los vectores de esta base. □

Observación 15.9. Sea M una variedad diferenciable tal que $M \subseteq \mathbb{R}^n \subseteq \mathbb{R}^m$. Mientras que el espacio tangente TM_a es el mismo tanto si contemplamos M como subvariedad de $\mathbb{R}^n$ como si lo hacemos respecto de $\mathbb{R}^m$, el espacio normal, por el contrario, es muy diferente: por un lado se toma el complemento ortogonal a TM_a en $\mathbb{R}^n$, y por otro en $\mathbb{R}^m$, siendo el último mayor siempre que $m > n$. No obstante, el espacio normal a M en a relativo a $\mathbb{R}^n$ es la intersección de $\mathbb{R}^n$ con el espacio normal a M en a relativo a $\mathbb{R}^m$.

Pasemos ahora a estudiar la diferenciabilidad de funciones entre variedades. Recordemos que a la inversa de una parametrización local inyectiva la llamamos también carta; ver la Observación 15.5.

Definición 15.5. Sean M y N variedades diferenciables de dimensión m y n respectivamente, $a \in M$, y $f : M \to N$. Se dice que f es diferenciable en a si existen un entorno abierto U de a en M, un entorno abierto V de $f(a)$ en N, y cartas $\psi : U \to \mathbb{R}^m$, $\Phi : V \to \mathbb{R}^m$, tales que la aplicación

$$\Phi \circ f \circ \psi^{-1} : \psi(U) \subseteq \mathbb{R}^m \to \Phi(V) \subseteq \mathbb{R}^n$$

es diferenciable en a.

Se dice que f es de clase $C^p(M, N)$ si para todo punto $a \in M$ existen tales cartas ψ, Φ de modo que la aplicación $\Phi \circ f \circ \psi^{-1}$ es de clase C^p en $\psi(U)$.

Observación 15.10.

1. Es inmediato ver que si f satisface esta condición para dos cartas particulares ψ y Φ entonces también la cumple para *todas* las cartas ψ en un entorno de a y Φ en un entorno de $f(a)$.

2. También es inmediato que si $f : M \to N$ es diferenciable en a y $g : N \to L$ entonces $g \circ f : M \to L$ es diferenciable en a.

3. Si $M = \mathbb{R}^m$, $N = \mathbb{R}^n$, entonces esta definición equivale a la ya conocida.

A pesar de su generalidad esta definición no nos dice gran cosa. Por ejemplo, podemos saber que una determinada función $f : M \to N$ es diferenciable, incluso de clase C^∞, pero ¿qué es su diferencial y cómo se puede hallar? Para responder a esta pregunta, examinemos qué es y cómo se comporta la diferencial de una función entre dos abiertos de $\mathbb{R}^m$ y $\mathbb{R}^n$. Si $f : U \subseteq \mathbb{R}^m \to V \subseteq \mathbb{R}^n$ es diferenciable en a, su diferencial $Df(a)$ es una aplicación lineal que lleva vectores de $\mathbb{R}^m$ en vectores de $\mathbb{R}^n$. Si v es un vector de $\mathbb{R}^m$ y $\gamma : (-\varepsilon, \varepsilon) \to U$ es una curva diferenciable tal que $\gamma(0) = a$ y $\gamma'(0) = v$ entonces $w = Df(a)(v)$ es precisamente el vector tangente en 0 a la curva $\sigma : (-\varepsilon, \varepsilon) \to N$ definida por $\sigma(t) = f(\gamma(t))$. En efecto, por la regla de la cadena

$$\sigma'(0) = Df(\gamma(0))(\gamma'(0)) = Df(a)(v).$$

La diferencial de una aplicación f entre dos variedades M y N en un punto $a \in M$ podrá entonces definirse por analogía como la aplicación $df(a) : TM_a \to TN_{f(a)}$ que lleva un vector v tangente en a a una curva γ contenida en M en el vector w tangente a la curva $f \circ \gamma$ en el punto $f(a)$. Por supuesto, puesto que hay infinitas curvas cuya tangente en a es el mismo vector v, para que esto sea una buena definición hay que comprobar que no depende de la curva γ escogida. Más tarde veremos que esto es, en efecto, así.

Otra posibilidad es (suponiendo M variedad de $\mathbb{R}^k$ y N variedad de $\mathbb{R}^{k'}$) la de extender la función $f : M \to N$ a una función $\widetilde{f} : \mathbb{R}^k \to \mathbb{R}^{k'}$ que sea también diferenciable en a (¡pero ahora como función de $\mathbb{R}^k$ en $\mathbb{R}^{k'}$!) y definir entonces $df(a)$ como la restricción de $D\widetilde{f}(a)$ a TM_a. Con este enfoque surgen dos dificultades: no sólo hay que comprobar que la definición de $df(a) : TM_a \to TM_{f(a)}$ no depende de la extensión $\widetilde{f}$ utilizada, sino que tales extensiones realmente existen. El siguiente teorema establece la existencia de dichas $\widetilde{f}$.

Teorema 15.11. *Sea $f : M \to N$ una función diferenciable en $a \in M$, donde M y N son variedades de $\mathbb{R}^k$ y $\mathbb{R}^{k'}$ respectivamente. Entonces existe una función $\widetilde{f} : \mathbb{R}^k \to \mathbb{R}^{k'}$ tal que $\widetilde{f}_{|M} = f$ y $\widetilde{f}$ es diferenciable en a (como función de $\mathbb{R}^k$ en $\mathbb{R}^{k'}$).*

Además, si $f : M \to N$ es una función de clase $C^p(M, N)$, entonces existe un abierto U de $\mathbb{R}^k$ tal que $M \subset U$ y una función $\widetilde{f} \in C^p(U, \mathbb{R}^{k'})$ tal que $\widetilde{f}_{|M} = f$.

Demostración. Sean ψ una carta de M definida en un entorno de a, y Φ una carta de N definida en un entorno de $f(a)$. Puesto que $\varphi = \psi^{-1}$ es una parametrización local de M, por el Teorema 14.16 existen abiertos U, A de $\mathbb{R}^n$, con $a \in U$, y un difeomorfismo $\widetilde{\psi} : U \to A$ tal que

$$\widetilde{\psi} \circ \varphi = J,$$

donde J es la inyección canónica de $\mathbb{R}^m$ en $\mathbb{R}^k$. Si denotamos $P : \mathbb{R}^k \to \mathbb{R}^m$ la proyección $P(x_1, ..., x_k) = (x_1, ..., x_m)$, se tiene que $P \circ J = I$, la identidad, y por tanto

$$P \circ \widetilde{\psi} \circ \varphi = I,$$

la identidad, en un entorno de $\psi(a)$ que puede suponerse que es $\psi(M \cap U) := V$. Definamos $\widetilde{f} : W \subseteq \mathbb{R}^k \to \mathbb{R}^{k'}$ por

$$\widetilde{f} = f \circ \psi^{-1} \circ (P \circ \widetilde{\psi}).$$

Puesto que $\widetilde{f} = \Phi^{-1} \circ (\Phi \circ f \circ \psi^{-1}) \circ (P \circ \widetilde{\psi})$ y las aplicaciones Φ^{-1} y $(P \circ \widetilde{\psi})$ son diferenciables, y $(\Phi \circ f \circ \psi^{-1})$ es diferenciable en $\psi(a)$ por hipótesis, resulta que $\widetilde{f}$ es diferenciable en a.

Además, si $x = \psi^{-1}(y) \in \psi^{-1}(V) = M \cap U$, se tiene que

$$\widetilde{f}(x) = f \circ \psi^{-1} \circ (P \circ \widetilde{\psi}) \circ \psi^{-1}(y) = f \circ \psi^{-1} \circ I(y) = f \circ \psi^{-1}(y) = f(x),$$

es decir, $\widetilde{f}_{|M \cap U} = f$. Luego $\widetilde{f}$ es una extensión de f al entorno U de a. Ahora puede extenderse $\widetilde{f}$ de manera arbitraria a todo $\mathbb{R}^m$ (por ejemplo poniendo $\widetilde{f}(x) = 0$ si $x \in \mathbb{R}^k \setminus (U \cup M)$, y $\widetilde{f}(x) = f(x)$ si $x \in M$), de modo que $\widetilde{f}$ es una extensión de f a $\mathbb{R}^k$ que es diferenciable en a.

Nótese que, si f es de clase $C^p(M, N)$, entonces esta construcción proporciona una función $\widetilde{f}$ que es de clase $C^p(U, \mathbb{R}^{k'})$.

Finalmente, usando particiones diferenciables de la unidad, vamos a ver cómo podemos pegar todas estas extensiones locales de f para obtener una extensión de f que es de clase C^p en todo un abierto que contiene a M. Supongamos pues que $f \in C^p(M, N)$. Por lo anterior, para cada punto $a \in M$ existe U_a entorno abierto de a en $\mathbb{R}^k$ y $\widetilde{f}_a : U_a \to \mathbb{R}^{k'}$ de clase C^p en U_a tal que $\widetilde{f}_a(x) = f(x)$ para todo $x \in M \cap U_a$.

Consideremos el recubrimiento por abiertos $\mathcal{U} = \{U_a : a \in M\}$ de $U = \bigcup_{a \in M} U_a$ (que es un abierto de $\mathbb{R}^k$). Por el Teorema 12.8 existe una partición de la unidad $\{\mu_a : a \in M\}$ de clase C^∞ subordinada al recubrimiento $\mathcal{U}$. Definamos $\widetilde{f} : U \to \mathbb{R}^{k'}$ por

$$\widetilde{f}(x) = \sum_{a \in M} \mu_a(x) \widetilde{f}_a(x).$$

Como la suma es localmente finita, $\widetilde{f}$ está bien definida y es de clase C^p en U. Además, como $\sum_{a \in M} \mu_a(x) = 1$ para todo $x \in U$ y $\widetilde{f}_a(x) = f(x)$ si $x \in M$, se tiene que

$$\widetilde{f}(x) = \sum_{a \in M} \mu_a(x) \widetilde{f}_a(x) = \sum_{a \in M} \mu_a(x) f(x) = f(x) \sum_{a \in M} \mu_a(x) = f(x)$$

para todo $x \in M$. □

Observación 15.12. En general no puede obtenerse $\widetilde{f} \in C^p(\mathbb{R}^k, \mathbb{R}^{k'})$ tal que $\widetilde{f}_{|M} = f$; ver el problema 15.11.

Teorema 15.13. *Sean M y N variedades diferenciables de dimensiones m y n de $\mathbb{R}^k$ y $\mathbb{R}^{k'}$ respectivamente, $a \in M$, y $f : M \to N$ una aplicación diferenciable en a. Consideremos dos vectores $v \in TM_a$, $w \in TN_{f(a)}$. Sea $\gamma : (-\varepsilon, \varepsilon) \to M$ una curva diferenciable con $\gamma(0) = a$ y $\gamma'(0) = v$. Sea $\widetilde{f} : \mathbb{R}^k \to \mathbb{R}^{k'}$ una extensión de f diferenciable en a (como función de $\mathbb{R}^k$ en $\mathbb{R}^{k'}$). Y sean $\psi : U \to \mathbb{R}^m$ carta de M en a, $\Phi : V \to \mathbb{R}^n$ carta de N en $f(a)$, y $v' \in \mathbb{R}^m$, $w' \in \mathbb{R}^n$ tales que $v = D\psi^{-1}(y)(v')$, $w = D\Phi^{-1}(z)(w')$, donde $y = \psi(a)$, $z = \Phi(f(a))$.*

Las siguientes afirmaciones son equivalentes:

1. $w = (f \circ \gamma)'(0)$.
2. $w = D\widetilde{f}(a)(v)$.
3. $w' = D(\Phi \circ f \circ \psi^{-1})(y)(v')$.

Demostración. $(1) \implies (2)$: Al ser $\widetilde{f} : \mathbb{R}^k \to \mathbb{R}^{k'}$ diferenciable en a, y $\widetilde{f}_{|M} = f$, tenemos que, por la regla de la cadena,

$$w = (f \circ \gamma)'(0) = (\widetilde{f} \circ \gamma)'(0) = D\widetilde{f}(\gamma(0))(\gamma'(0)) = D\widetilde{f}(a)(v).$$

$(2) \implies (3)$: Como la parametrización local $\psi^{-1} : \psi(U) \to \mathbb{R}^k$ es diferenciable, la composición $\widetilde{f} \circ \psi^{-1} : \psi(U) \to \mathbb{R}^{k'}$ es diferenciable en a. Por otro lado, como $\varphi := \Phi^{-1} : \Phi(V) \to \mathbb{R}^{k'}$ es una parametrización local de N, por el Teorema 14.16 existe un difeomorfismo $\widetilde{\Phi}$ entre dos abiertos de $\mathbb{R}^{k'}$ de modo que

$$\widetilde{\Phi} \circ \varphi = J,$$

donde J representa la inyección canónica de $\mathbb{R}^n$ en $\mathbb{R}^{k'}$. Si $P : \mathbb{R}^{k'} \to \mathbb{R}^n$ es la proyección tal que $P \circ J = I$ (identidad), podemos añadir que $P \circ \widetilde{\Phi}$ es una extensión diferenciable de Φ a un entorno abierto de $f(a)$ en $\mathbb{R}^{k'}$, y $P \circ \widetilde{\Phi} \circ \varphi = I$, de donde se deduce que $D(P \circ \widetilde{\Phi}(f(a)) \circ D\varphi(z)(w') = w'$, es decir, $D(P \circ \widetilde{\Phi}(f(a))(w) = w'$. Por tanto,

$$\begin{aligned} &D(\Phi \circ f \circ \psi^{-1})(y)(v') = D(P \circ \widetilde{\Phi} \circ \widetilde{f} \circ \psi^{-1})(y)(v') = \\ &D(P \circ \widetilde{\Phi})(f(a)) \circ D\widetilde{f}(a) \circ D\psi^{-1}(y)(v') = \\ &D(P \circ \widetilde{\Phi})(f(a)) \left(D\widetilde{f}(a)(v)\right) = D(P \circ \widetilde{\Phi})(f(a))(w) = w'. \end{aligned}$$

$(3) \implies (1)$: Siguiendo con la notación anterior, podemos tomar un difeomorfismo $\widetilde{\psi}$ entre dos abiertos de $\mathbb{R}^m$ de modo que $P \circ \widetilde{\psi}$ es una extensión diferenciable de ψ a un abierto de $\mathbb{R}^m$. Definamos la curva $\sigma : (-\varepsilon, \varepsilon) \to \mathbb{R}^m$ por

$$\sigma(t) = P \circ \widetilde{\psi}(\gamma(t)).$$

Es inmediato que σ es una curva diferenciable tal que $\sigma(0) = y$, $\sigma'(0) = v'$. Entonces

$$\begin{aligned} &(\Phi \circ f \circ \gamma)'(0) = (\Phi \circ f \circ \psi^{-1} \circ \sigma)'(0) = \\ &D(\Phi \circ f \circ \psi^{-1})(\sigma(0))((\sigma)'(0)) = D(\Phi \circ f \circ \psi^{-1})(y)(v') = w' \end{aligned}$$

y por tanto

$$\begin{aligned} &(f \circ \gamma)'(0) = (\Phi^{-1} \circ \Phi \circ f \circ \gamma)'(0) = \\ &D\Phi^{-1}(f(a))((\Phi \circ f \circ \gamma)'(0)) = D\Phi^{-1}(f(a))(w') = w. \end{aligned}$$

Esto prueba la equivalencia entre las tres condiciones. $\square$

Corolario 15.14. *Sean M y N variedades diferenciables de dimensiones m y n de $\mathbb{R}^k$ y $\mathbb{R}^{k'}$, respectivamente. Sean $a \in M$, y $f : M \to N$ una aplicación diferenciable en a. Entonces existe una única aplicación lineal $df(a) : TM_a \to TN_{f(a)}$ con las siguientes propiedades:*

1. *Para todo $v \in TM_a$ y toda curva $\gamma : (-\varepsilon, \varepsilon) \to M$ con $\gamma(0) = a$, $\gamma'(0) = v$, el vector $w := df(a)(v)$ satisface que*
$$w = (f \circ \gamma)'(0).$$
2. *Para toda función $\widetilde{f} : \mathbb{R}^k \to \mathbb{R}^{k'}$ diferenciable en a y tal que $\widetilde{f}_{|M} = f$ se tiene que*
$$df(a)(v) = D\widetilde{f}(a)(v);$$
es decir, $df(a)$ es la restricción de $D\widetilde{f}(a) : \mathbb{R}^k \to \mathbb{R}^{k'}$ al subespacio TM_a de $\mathbb{R}^k$.
3. *Para todas $\psi : U \to \mathbb{R}^m$ carta de M en a, $\Phi : V \to \mathbb{R}^n$ carta de N en $f(a)$, y $v' \in \mathbb{R}^m$, $w' \in \mathbb{R}^n$ tales que $v = D\psi^{-1}(y)(v')$, $df(a)(v) = D\Phi^{-1}(z)(w')$, donde $y = \psi(a)$, $z = \Phi(f(a))$, se cumple que*
$$w' = D(\Phi \circ f \circ \psi^{-1})(y)(v').$$

Definición 15.6. A la aplicación $df(a) : TM_a \to TN_{f(a)}$ así definida se la llamará diferencial de f en a.

Observación 15.15.

1. Cuando $M = \mathbb{R}^m$, $N = \mathbb{R}^n$, esta definición de diferencial coincide con la ya conocida (ya que puede tomarse $k = m$, $k' = n$, y $\widetilde{f} = f : \mathbb{R}^m \to \mathbb{R}^n$).
2. Cuando $f : M \to \mathbb{R}$, $df(a)$ es una aplicación lineal de TM_a en $\mathbb{R}$, y por tanto puede identificarse a un vector de TM_a (si E es un subespacio de $\mathbb{R}^k$, la aplicación $E \ni x \to T_x \in \mathcal{L}(E, \mathbb{R})$ definida por $T_x(v) = \langle x, v\rangle$ es una isometría lineal); a este vector $df(a) \in TM_a$ lo llamaremos vector gradiente de f en a, y lo denotaremos a veces también por $\nabla f(a)$.

Demostración del Corolario 15.14: Si $v \in TM_a$, por definición de espacio tangente sabemos que existe una curva $\gamma : (-\varepsilon, \varepsilon) \to M$ tal que $\gamma(0) = a$ y $\gamma'(0) = v$. Entonces la curva $f \circ \gamma$ está contenida en N, y así $w_v = w := (f \circ \gamma)'(0) \in TN_{f(\gamma(0))} = TN_{f(a)}$. Además este w_v es único y no depende de la curva γ: si $w_1 = (f \circ \gamma_1)'(0)$ y $w_2 = (f \circ \gamma_2)'(0)$ para dos curvas γ_1, γ_2 (posiblemente diferentes) con $\gamma_1(0) = a = \gamma_2(0)$, $\gamma_1'(0) = v = \gamma_2'(0)$, entonces, fijando una misma extensión $\widetilde{f} : \mathbb{R}^k \to \mathbb{R}^{k'}$ de f diferenciable en a, y aplicando dos veces la equivalencia $(1) \iff (2)$ del Teorema 15.13, deducimos que
$$w_1 = (f \circ \gamma_1)'(0) = D\widetilde{f}(a)(v) = (f \circ \gamma_2)'(0) = w_2.$$
Por tanto la aplicación $v \mapsto w_v$ está bien definida de TM_a en $TN_{f(a)}$. Denotaremos esta aplicación por $v \mapsto df(a)(v) = w_v$. Si en este razonamiento dejamos fija la aplicación $\widetilde{f}$ y consideramos todos los $v \in TM_a$ y γ posibles con estas propiedades, el argumento prueba que
$$df(a)(v) = D\widetilde{f}(a)(v)$$
para todo $v \in TM_a$, es decir, $df(a)$ es la restricción de la aplicación lineal $D\widetilde{f}(a)$ al subespacio TM_a de $\mathbb{R}^m$, y en particular $df(a)$ es una aplicación lineal. Como $\widetilde{f}$ es arbitraria esto prueba también que $df(a)$ satisface (2). Es evidente por la misma definición que $df(a)$ es la única aplicación de TM_a en $TN_{f(a)}$ que satisface (1). Finalmente, la equivalencia $(1) \iff (3)$ del Teorema 15.13 nos dice que $df(a)$ también satisface (3). □

Definición 15.7. Por analogía con la terminología usada para funciones entre abiertos de espacios vectoriales de dimensión finita, se dirá que $a \in M$ es un *punto regular* de una función $f : M \to N$ si su diferencial $df(a) : TM_a \to TN_{f(a)}$ tiene rango máximo (es decir, es sobreyectiva si $n = \dim(N) \leq \dim(M) = m$, o es inyectiva si $n \geq m$; cuando las dimensiones n y m son iguales esto ocurre si y sólo si $df(a)$ es un isomorfismo lineal).

Se dirá que $b \in N$ es un *valor regular* de f si x es un punto regular de f para todo $x \in f^{-1}(b)$.

Si a no es un punto regular de f, diremos que a es un *punto crítico* de f, y a $f(a)$ se le llamará *valor crítico.*

Un *difeomorfismo* entre dos variedades diferenciables es una biyección diferenciable con inversa también diferenciable. Se dirá que dos variedades son *difeomorfas* si existe un difeomorfismo entre ellas.

El siguiente teorema, cuya demostración puede dejarse como ejercicio, es una reformulación del teorema de la función inversa en el contexto de variedades.

Teorema 15.16. *Sean M y N dos variedades diferenciables de la misma dimensión, y supongamos que $f : M \to N$ es regular en $a \in M$. Entonces f es un difeomorfismo local en un entorno de a, es decir, existen U entorno abierto de a en M, V entorno abierto de $f(a)$ en N tal que $f_{|U} : U \to V$ es un difeomorfismo.*

Para profundizar un poco más sobre las consecuencias locales y globales de la existencia de puntos y valores regulares, pueden consultarse los problemas 15.37 y 15.38. Ver también la demostración del teorema fundamental del álgebra que se proporciona al final de este capítulo.

Estudiaremos ahora cómo pueden localizarse los extremos de una función diferenciable sobre una variedad, deduciendo el teorema de los multiplicadores de Lagrange cuya versión más simple ya fue enunciada al final del capítulo 13.

Definición 15.8. Se dice que una función $f : M \to \mathbb{R}$ definida sobre una variedad M tiene un máximo local en $a \in M$ si existe un entorno U de a en M tal que $f(x) \leq f(a)$ para todo $x \in U$. Análogamente se define mínimo local. A los máximos y mínimos locales se le llaman extremos locales de f.

Teorema 15.17. *Sean M una variedad diferenciable, $f : M \to \mathbb{R}$ una función diferenciable en $a \in M$, y supongamos que f tiene un extremo local en a. Entonces $df(a) = 0$.*

Demostración. Sea $v \in TM_a$, y consideremos una curva diferenciable $\gamma : (-\varepsilon, \varepsilon) \to M$ tal que $\gamma(0) = a$, $\gamma'(0) = v$. Entonces la función $g_v = g : (-\varepsilon, \varepsilon) \to \mathbb{R}$ definida por

$$g(t) = f(\gamma(t))$$

tiene un extremo local en 0 (del mismo tipo que f lo tenga en a), y por tanto $g'(0) = 0$, es decir,

$$0 = g'(0) = df(\gamma(0))(\gamma'(0)) = df(a)(v).$$

Esto prueba que $df(a)(v) = 0$ para todo $v \in TM_a$, y por consiguiente que $df(a) = 0$. $\square$

Combinando este resultado con la caracterización del espacio normal proporcionada por el Teorema 15.8 podemos deducir la versión general del Teorema 13.7.

Teorema 15.18 (de los multiplicadores de Lagrange)**.** *Sea M una variedad diferenciable de dimensión m en $\mathbb{R}^n$, descrita en un entorno U de un punto $a \in M$ por la ecuación $F = 0$, donde $F = (F_1, ..., F_{n-m}) : U \subseteq \mathbb{R}^n \to \mathbb{R}^{n-m}$. Sea $\widetilde{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}$ una función diferenciable en a, y supongamos que $f := \widetilde{f}_{|M} : M \to \mathbb{R}$ tiene un extremo local en a. Entonces existen $\lambda_1, ..., \lambda_{n-m} \in \mathbb{R}$ tales que*

$$\nabla\widetilde{f}(a) = \sum_{j=1}^{n-m} \lambda_j \nabla F_j(a).$$

A los números $\lambda_1, ..., \lambda_{n-m}$ se los llama multiplicadores de Lagrange.

Demostración. Sea $v \in TM_a \subset \mathbb{R}^n$, y tomemos $\gamma : (-\varepsilon, \varepsilon) \to M$ tal que $\gamma(0) = a$ y $\gamma'(0) = v$. Entonces, puesto que $df(a) = D\widetilde{f}(a)_{|TM_a}$, y además $df(a) = 0$ por el teorema anterior, se tiene que

$$\langle \nabla\widetilde{f}(a), v\rangle = D\widetilde{f}(a)(v) = df(a)(v) = 0.$$

Esto prueba que $\nabla\widetilde{f}(a)$ es perpendicular a todo vector $v \in TM_a$, es decir, $\nabla\widetilde{f}(a) \in TM_a^{\perp}$, el espaco normal a M en a.

Por el Teorema 15.8, esto quiere decir que el vector $\nabla\widetilde{f}(a)$ está en $[\nabla F_1(a), ..., \nabla F_{n-m}(a)]$, es decir, existen números reales $\lambda_1, ..., \lambda_{n-m}$ tales que $\nabla\widetilde{f}(a) = \sum_{j=1}^{n-m} \lambda_j \nabla F_j(a)$. □

Observación 15.19. Como se ve en la demostración, los puntos críticos de $f = \widetilde{f}_{|M}$ son los que satisfacen la condición de los multiplicadores de Lagrange del Teorema anterior.

Corolario 15.20. *Sean $F = (F_1, ..., F_k) : U \subseteq \mathbb{R}^n \to \mathbb{R}^k$ una función de clase C^1, y supongamos que 0 es un valor regular de F (es decir, rango$(DF(x)) = k$ para todo $x \in \mathbb{R}^n$ con $F(x) = 0$. Sea $f : U \to \mathbb{R}$ diferenciable en U. Si a es un punto de máximo o mínimo de $f(x)$ sujeta a las restricciones $F_1(x) = ... = F_k(x) = 0$, entonces*

$$\nabla f(a) = \lambda_1 \nabla F_1(a) + ... + \lambda_k \nabla F_k(a)$$

para ciertos números $\lambda_1, ..., \lambda_k \in \mathbb{R}$.

Podemos ahora retomar la discusión iniciada en el capítulo 13 acerca de cómo identificar los máximos y mínimos de una función diferenciable en un compacto de $\mathbb{R}^n$. Como ya observamos, el problema es demasiado difícil en su forma más general, pero cuando $K = \overline{U}$, donde U es un abierto acotado de $\mathbb{R}^n$, y la frontera ∂K es unión finita de variedades diferenciables (un caso muy frecuente en la práctica), entonces sí que podemos dar una estrategia para encontrar los puntos donde una función f (que se supone diferenciable en un abierto que contiene a K) pueda alcanzar sus extremos absolutos (que siempre existen al ser f continua en K). Primero deben hallarse los puntos críticos de f en U. Después, suponiendo

$$\partial K = M_1 \cup ... \cup M_k,$$

donde M_j son variedades diferenciables disjuntas de dimensiones diversas, deben hallarse, usando el Teorema de los multiplicadores de Lagrange, los puntos críticos de las restricciones $f_{|M_j}$, para $j = 1, ..., k$. Finalmente, deben comparase los valores de f en todos los puntos críticos de f en U y de sus restricciones a las variedades M_j para identificar así el máximo y el mínimo de f en K.

Ejemplo 15.3. Calcular los extremos de la función $f(x, y, z) = x - y + z - 5$ en el compacto $K = \{(x, y, z) \in \mathbb{R}^3 : z \geq x^2 + y^2,\ x^2 + 3y^2 + 5z^2 \leq 27\}$.

En este caso, $K = \overline{U}$, donde $U = \{(x, y, z) \in \mathbb{R}^3 : z > x^2 + y^2,\ x^2 + 3y^2 + 5z^2 < 27\}$ es abierto, y

$$\partial K = M_1 \cup M_2 \cup M_3,$$

donde

$$M_1 = \{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2, x^2 + 3y^2 + 5z^2 < 27\},$$
$$M_2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + 3y^2 + 5z^2 = 27, z > x^2 + y^2\},$$
$$M_3 = \{(x, y, z) \in \mathbb{R}^3 : z = x^2 + y^2, x^2 + 3y^2 + 5z^2 = 27\}.$$

Es inmediato comprobar que M_1 y M_2 son variedades de dimensión 2 (es decir superficies) de $\mathbb{R}^3$, y M_3 es una variedad unidimensional (curva) en $\mathbb{R}^3$. Por tanto puede aplicarse la estrategia anterior para encontrar $\inf_K f$ y $\sup_K f$ y los puntos donde éstos se alcanzan.

Concluiremos este capítulo mostrando cómo puede aplicarse la teoría de variedades diferenciables hasta aquí desarrollada y la noción de valor regular para obtener una bella demostración del teorema fundamental del álgebra.

Para ello necesitamos primero hacer una sencilla observación sobre los puntos regulares de una función diferenciable entre dos variedades de igual dimensión.

Lema 15.21. *Sean M y N variedades diferenciables de la misma dimensión. Supongamos M compacta, y sea $f : M \to N$. Entonces:*

1. *Para cualquier valor regular $y \in N$ de f, el conjunto $f^{-1}(y)$ es finito.*
2. *La aplicación $y \mapsto \sharp f^{-1}(y)$ (donde $\sharp f^{-1}(y)$ denota el número de puntos de $f^{-1}(y)$), definida sobre el conjunto de valores regulares de f, es localmente constante.*

Demostración. Sea $y \in N$ un valor regular de f. Por el teorema de la función inversa para variedades enunciado más arriba, sabemos que para cada $x \in f^{-1}(y)$ existe un entorno abierto U_x de x en M tal que $f : U_x \to f(U_x)$ es un difeomorfismo. Esto implica que $f^{-1}(y)$ consta exclusivamente de puntos aislados y, como es compacto (por ser cerrado en la variedad compacta M), debe ser finito.

Por otra parte, si fijamos un valor regular $y \in N$ y $x_1, ..., x_k$ son los puntos de $f^{-1}(y)$, el argumento anterior muestra que existen entornos $U_1, ..., U_k$ de $x_1, ..., x_k$ tales que $f : U_j \to f(U_j) := V_j$ es un difeomorfismo para cada $j = 1, ..., k$, y obviamente puede suponerse que los U_j son disjuntos dos a dos. Pongamos

$$V = (V_1 \cap ... \cap V_k) \setminus (f (M \setminus (U_1 \cup ... \cup U_k))) .$$

Entonces V es un entorno abierto de y en N que consta sólo de valores regulares de f y tal que para todo $z \in V$ se tiene $\sharp f^{-1}(z) = k = \sharp f^{-1}(y)$. □

Teorema 15.22. *Sea $P : \mathbb{C} \to \mathbb{C}$ un polinomio de grado $n \geq 1$. Entonces P tiene exactamente n raíces complejas.*

Demostración. Puesto que todo monomio de la forma $z - z_0$, donde z_0 es una raíz de P, divide a P, y en cada división el grado del polinomio cociente disminuye una unidad, es obvio que un P tiene a lo sumo n raíces, y que basta probar que *todo polinomio no constante tiene al menos una raíz* (ya que entonces, aplicando este resultado a cada uno de dichos cocientes sucesivos llegamos a que P tiene exactamente n raíces).

Para demostrar esto identificaremos, mediante la proyeccción estereográfica, el plano complejo como un subconjunto de la esfera unidad de $\mathbb{R}^3$. Consideremos la esfera unidad S de $\mathbb{R}^3$, y la proyección estereográfica desde el *polo norte* de S,

$$h_+ : S \setminus \{(0,0,1)\} \to \mathbb{R}^2 \times \{0\} \equiv \mathbb{R}^2 \equiv \mathbb{C},$$

y la proyección estereográfica desde el *polo sur* de S,

$$h_- : S \setminus \{(0,0,-1)\} \to \mathbb{R}^2 \times \{0\} \equiv \mathbb{R}^2 \equiv \mathbb{C}.$$

Las aplicaciones h_+ y h_- son cartas de S, es decir sus inversas son parametrizaciones regulares e inyectivas cuyas imágenes recubren S (ver el problema 15.16).

Podemos identificar el polinomio $P : \mathbb{C} \to \mathbb{C}$ con la aplicación $f : S \to S$ definida por

$$f(x) = \begin{cases} h_+^{-1} \circ P \circ h_+(x) & \text{si } x \neq (0,0,1), \\ (0,0,1) & \text{si } x = (0,0,1). \end{cases}$$

Vamos a ver que la aplicación $f : S \to S$ es de clase C^∞. Puesto que $P : \mathbb{R}^2 \to \mathbb{R}^2$ es de clase C^∞ (problema 11.6), $f = h_+^{-1} \circ P \circ h_+(x)$ es de clase C^∞ en $S \setminus \{(0,0,1)\}$, por definición. Veamos que f también es de clase C^∞ en un entorno de $(0,0,1)$. Esto quedará garantizado, por definición, si probamos que la aplicación $Q : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

$$Q(z) = h_- \circ f \circ h_-^{-1}(z)$$

es de clase C^∞ en un entorno de $h_-(0,0,1) = (0,0) \in \mathbb{R}^2$. Es fácil ver (problema 11.6) que

$$(h_+ \circ h_-^{-1})(z) = \frac{z}{|z|^2} = \frac{1}{\overline{z}},$$

y usando este hecho un cálculo directo muestra que, si $P(z) = a_0 z^n + a_1 z^{n-1} + ... + a_{n-1} z + a_n$, con $a_0 \neq 0$, entonces

$$Q(z) = \frac{z^n}{\overline{a}_0 + \overline{a}_1 z + ... + \overline{a}_n z^n},$$

y por tanto Q es de clase C^∞ en un entorno de 0 (ver el problema 11.7).

Obsérvese que $f : S \to S$ tiene una cantidad finita de puntos críticos, ya que P deja de ser un difeomorfismo local únicamente en las raíces de su derivada P', que como es un polinomio de grado $n-1$, tiene a lo sumo $n-1$ ceros (digamos $c_1, ..., c_m$). Por tanto, el conjunto $\mathcal{R}$ de valores regulares de f es la esfera S menos una cantidad finita de puntos (a saber, los puntos $f(c_1), ..., f(c_m)$), lo que en cualquier caso resulta un conjunto conexo. Por tanto, la función $y \mapsto \sharp f^{-1}(y)$, al ser localmente constante (por el lema anterior) en el conjunto conexo $\mathcal{R}$, es de hecho constante en este conjunto. Puesto que $\sharp f^{-1}(y)$ no puede ser cero para todo $y \in \mathcal{R}$ (esto significaría que $f(S) = S \setminus \mathcal{R}$ sería finito, lo cual es imposible porque es conexo y f no es constante), se deduce que $\sharp f^{-1}(y) = k > 0$ para todo $y \in \mathcal{R}$, y en particular $\mathcal{R} \subseteq f(S)$. Por tanto,

$$\begin{aligned} S = (S \setminus \{f(c_1), ..., f(c_m)\}) \cup \{f(c_1), ..., f(c_m)\} = \\ \mathcal{R} \cup \{f(c_1), ..., f(c_m)\} \subseteq f(S), \end{aligned}$$

es decir, $f : S \to S$ es sobreyectiva. Existe, por consiguiente, $x_0 \in S$ tal que $f(x) = (0,0,-1)$, lo que equivale a $P(z_0) = 0$ para $z_0 = h_+(x_0)$. $\square$

15.1. Problemas

Problema 15.1. Demuéstrese que las ecuaciones

$$\begin{cases} x^2 - y^2 + xyz^2 - 11 = 0 \\ x^3 + y^3 + z^3 - xyz - 30 = 0 \end{cases}$$

definen una variedad de clase C^∞ y dimensión 1 en $\mathbb{R}^3$, en un entorno del punto $(3, 2, 1)$. Hállense las ecuaciones de la tangente y del plano normal a la variedad en dicho punto.

Problema 15.2. Demuéstrese que las ecuaciones

$$\begin{cases} 2x + y + 2z + u - v - 1 = 0 \\ xy + z - u + 2v - 1 = 0 \\ yz + xz + u^2 + v = 0 \end{cases}$$

definen una variedad en un entorno de (1,1,1,1,-1). Determínese su dimensión y calcúlese el espacio tangente en (1,1,1,1,-1).

Problema 15.3. Calcúlense las tangentes a las curvas siguientes en los puntos indicados:

$$\varphi(t) = (1 + t, 1 + 2t), \quad \varphi(t) = (1 + t^3, 1 + 2t^3), \text{ en } t = 0, t = 1$$

$$\begin{cases} xyz = 1 \\ y = z \end{cases} \text{ en } (1, 1, 1) \quad \begin{cases} xy + z^2 = 1 \\ x + y = 0 \end{cases} \text{ en } (0, 0, 1)$$

Problema 15.4. Probar que la dimensión de una variedad diferenciable es única, es decir, si $M \subseteq \mathbb{R}^n$ es una variedad diferenciable de dimensión k y también de dimensión m, entonces $k = m$.

Problema 15.5. Sea $M = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 - z^2 + 2xy - 16 = 0\}$.

1. Demuéstrese que M es una variedad.
2. Sea $f(x, y, z) = x^2 + y^2 + z^2$. Calcúlense los posibles extremos de f en M.
3. ¿Es (-2,-2,0) un punto extremo de f sobre M?
4. ¿Tiene f extremos absolutos sobre M?

Problema 15.6. Indíquese en qué puntos las ecuaciones dadas definen variedades diferenciables de dimensión 2 en $\mathbb{R}^3$ y calcúlense el plano tangente y la normal en los puntos que se indican:

(1) $z^2 = x^2 + y^2$ en $(1, 0, 1)$ (2) $z^3 = x^2 + y^2$ en $(0, -1, 1)$

(3) $z = x^2 + y^2$ en $(1, -2, 5)$ (4) $\dfrac{x^2}{16} + \dfrac{y^2}{9} - \dfrac{z^2}{8} = 0$ en $(4, 3, 4)$

Problema 15.7. Hállense en la superficie $x^2 + y^2 - z^2 - 2x = 0$ los puntos en que los planos tangentes a ella sean paralelos a los planos coordenados.

Problema 15.8. Consideremos el subespacio métrico de $\mathbb{R}^2$ definido por $M = \{(x, y) \in \mathbb{R}^2 : y = |x|\}$.

1. Probar que M es homeomorfo a $\mathbb{R}$. Utilizar este hecho para dotar al espacio métrico M de una estructura de variedad diferenciable de dimensión uno y clase C^∞ según la definición general de la Observación 15.5.
2. Probar que no existe ninguna curva diferenciable $\gamma : (-\varepsilon, \varepsilon) \subset \mathbb{R} \to M$ tal que $\gamma(0) = (0,0)$ y $\gamma'(0) \neq 0$.
3. Concluir que no existe ninguna parametrización regular de ningún entorno de $(0,0)$ en M por un abierto de $\mathbb{R}$. Por tanto M no es una variedad diferenciable según la Definición 15.2.
4. Poner otros ejemplos de subconjuntos de $\mathbb{R}^2$ y $\mathbb{R}^3$ que sean variedades diferenciables según la definición de la Observación 15.5, pero que no sean subvariedades de $\mathbb{R}^3$, es decir, que no satisfagan la Definición 15.2.

Problema 15.9. Supongamos $M \subset \mathbb{R}^n \subset \mathbb{R}^m$: Demostrar que si M es una variedad k-dimensional de $\mathbb{R}^n$ entonces M también es una variedad k-dimensional de $\mathbb{R}^m$.

Problema 15.10. Sea a un punto de una variedad diferenciable M de $\mathbb{R}^n$. Sea $v \in \mathbb{R}^n$ un vector con $\|v\| = 1$. Probar que $v \in TM_a$ si y sólo si existe una sucesión (x_j) en M tal que

$$\lim_{j\to\infty} x_j = a \ \text{ y } \ \lim_{j\to\infty} \frac{x_j - a}{\|x_j - a\|} = v.$$

Por tanto, el espacio tangente a M en a puede también caracterizarse como el conjunto de límites de direcciones de cuerdas en $\mathbb{R}^n$ que unen puntos de la variedad al punto a.

Problema 15.11. Sea $M = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 < 1, z = 0\}$, y consideremos $f : M \to \mathbb{R}$ definida por

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

Demostrar que M es una variedad diferenciable de $\mathbb{R}^3$ y que $f : M \to \mathbb{R}$ es diferenciable de clase C^∞, pero que no existe ninguna función $\widetilde{f} : \mathbb{R}^3 \to \mathbb{R}$ diferenciable en $\mathbb{R}^3$ tal que $\widetilde{f}_{|M} = f$.

Problema 15.12. Encontrar la distancia del punto $(-1,1)$ a la curva $xy = 1$, $x > 0$.

Problema 15.13. Sea $b \in \mathbb{R}^n$, sea M una variedad de $\mathbb{R}^n$, y sea a un punto de M de distancia mínima a b. Probar que $b - a$ es normal a M en a. ¿Existe siempre tal punto $a \in M$?

Problema 15.14. Probar que el *toro* obtenido al rotar la circunferencia $x^2 + (y-2)^2 = 1$ alrededor del eje x cumple la ecuación

$$F(x,y,z) := (x^2 + y^2 + z^2 + 3)^2 - 16(y^2 + z^2) = 0.$$

Probar que 0 es un valor regular de F, y por tanto el toro es una variedad bidimensional de $\mathbb{R}^3$.

Problema 15.15. Hallar el máximo y el mínimo de la función $f(x,y,z) = z$ sobre el toro del problema anterior.

Problema 15.16. Sea S la esfera unidad de $\mathbb{R}^3$, y consideremos el polo norte $p = (0,0,1)$ y el polo sur $s = (0,0,-1)$. Para cada $x \in S \setminus \{p\}$, denotemos por $h_+(x)$ la intersección de la recta que pasa por p y x con el plano $\mathbb{R}^2 \times \{0\}$. Análogamente, para cada $x \in S \setminus \{s\}$, sea $h_-(x)$ la intersección de este plano con la recta que pasa por x y s. A la aplicación $h_+ : S \setminus \{p\} \to \mathbb{R}^2$ se la llama *proyección estereográfica* de S sobre el plano $\mathbb{R}^2$ desde el polo norte, y a $h_- : S \setminus \{s\} \to \mathbb{R}^2$ se la llama proyección estereográfica de S desde el polo sur.

1. Probar que h_+ y h_- son biyecciones.
2. Hallar las ecuaciones de h_+ y h_- y sus inversas.
3. Probar que h_+^{-1} y h_-^{-1} son parametrizaciones regulares de S.
4. Comprobar que
$$(h_+ \circ h_-^{-1})(x) = \frac{x}{\|x\|^2}.$$
5. Generalizar todo esto para la esfera unidad de $\mathbb{R}^n$, $n \geq 2$.

Problema 15.17. Probar que toda variedad es localmente conexa, es decir, cada punto tiene un entorno abierto en la variedad que es conexo.

Problema 15.18. Probar que toda variedad compacta tiene un número finito de componentes conexas.

Problema 15.19. Probar que un subconjuto M de $\mathbb{R}^n$ es una variedad de dimensión n en $\mathbb{R}^n$ si y sólo si M es abierto en $\mathbb{R}^n$.

Problema 15.20. Probar que un subconjunto M de $\mathbb{R}^n$ es una variedad de dimensión cero si y sólo si M consta únicamente de puntos aislados. Concluir que M es a lo sumo numerable.

Problema 15.21. Sea M una variedad conexa de dimensión uno en $\mathbb{R}^n$. Demostrar que M es homeomorfa (de hecho difeomorfa) o bien a una circunferencia, o bien a un intervalo abierto de $\mathbb{R}$.

Problema 15.22. Calcúlese el plano tangente a las superficies siguientes en los puntos indicados:

$(1)\ z = \max\{|\,x\,|, |\,y\,|\}$ en $(1, 2, 2)$ $\qquad$ $(2)\ e^x + e^y + e^z = 3$ en $(0, 0, 0)$

$$(3) \begin{cases} x = \cos u \sin v \\ y = \sin u \sin v \\ z = \cos v \end{cases} \text{en } (\pi/4, \pi/4) \qquad (4) \begin{cases} x = s\cos t \\ y = s \sin t \\ z = t \end{cases} (0 < s < 2) \text{ en } (1, 2\pi)$$

Problema 15.23. Hállense los planos tangentes a la superficie $x^2 + 2y^2 + 3z^2 = 21$, paralelos al plano $x + 4y + 6z = 0$.

Problema 15.24. Parametrícense las siguientes curvas y superficies alrededor de los puntos indicados, y calcúlese la recta o plano tangente (respectivamente) en dichos puntos.

$(1)\ x^2 - y^2 = 9$ en $(5, 4)$ $\qquad$ $(2)\ x^2 + y^2 = z^2 + 1$ en $(0, \sqrt{2}, 1)$

$(3)\ xyz = 1$ en $(1, 1, 1)$ $\qquad$ $(4)\ 4x^2 + y^2 + 9z^2 = 1$ en $(0, 0, \frac{1}{3})$

Problema 15.25. Estúdiense los extremos de las siguientes funciones con las condiciones que se indican:

1. $f(x, y, z) = x - y + z;\ x^2 + y^2 + z^2 = 2$
2. $f(x, y) = x;\ x^2 + 2y^2 = 3$
3. $f(x, y, z) = x^2 + y^2 + z^2;\ x^2 - xy + y^2 - z^2 = 1$

4. $f(x,y) = 3x + 2y;\ 2x^2 + 3y^2 = 3$

5. $f(x,y,z) = x + y + z;\ x^2 - y^2 = 1; 2x + z = 1$

6. $f(x,y) = x - y;\ x^2 - y^2 = 2.$

Problema 15.26. Calcúlese la distancia de los siguientes conjuntos al origen de coordenadas: $A = \{(x,y,z);\ x^2+y^2 = 1;\ x+y+z = 1\}$; $B = \{(x,y,z);\ x^2+y^2-z^2 = 1;\ x+y+z = 1\}$.

Problema 15.27. Calcúlense los extremos de la función $f(x,y) = (x-y)^n$, con la condición $x^2 + y^2 = 1$ $(n \geq 1)$.

Problema 15.28. Sea $f(x,y,z) = x^2 + y^2 + z^2 + x + y + z$.

1. Pruébese que el conjunto $M = \{(x,y,z) \in \mathbb{R}^3;\ x^2 + y^2 + z^2 \leq 4;\ z \leq 1\}$ es compacto.

2. Calcúlense los puntos de máximo y mínimo absoluto de f sobre M.

Problema 15.29. a) Estúdiense los extremos de la función $f(x,y,z) = \log x + \log y + 3\log z$ sobre la porción de esfera $x^2 + y^2 + z^2 = 5r^2$ en la que $x > 0,\ y > 0,\ z > 0$.

b) Utilícese el resultado obtenido en a) para demostrar que, si $a,\ b$ y c son números reales positivos, se cumple:

$$(abc)^3 \leq 27\left(\frac{a+b+c}{5}\right)^5.$$

Problema 15.30. Se considera la función $f(x,y,z) = x + y + z$.

1. Calcúlense los extremos relativos de f sobre el hiperboloide $x^2 + y^2 - 2z^2 = 6$

2. Calcúlense los extremos absolutos de f sobre el compacto

$$K = \{(x,y,z) :\ x^2 + y^2 - 2z^2 \leq 6,\ 0 \leq z \leq 2\}$$

Problema 15.31. Hállense los ejes de la elipse $5x^2 + 8xy + 5y^2 = 9$.

Problema 15.32. Hállese el punto más alto y más bajo del elipsoide $2x^2 + 3y^2 + z^2 + 12xy + 4xz = 25$.

Problema 15.33. Sea $f : (0,\infty)\times(0,\infty) \longrightarrow \mathbb{R}$ la función definida por $f(x,y) = \frac{x^p}{p} + \frac{y^q}{q} - xy$, donde p, $q > 1$.

1. Discútanse los puntos críticos de f en función de p y q.

2. Si $\frac{1}{p} + \frac{1}{q} = 1$, demuéstrese la *desigualdad de Young*: $xy \leq \frac{x^p}{p} + \frac{y^q}{q} \quad \forall x > 0, \quad \forall y > 0.$

Problema 15.34. Se considera el conjunto $A = \{(x,y,z,t) \in \mathbb{R}^4 :\ x^2+y^2+z^2 \leq t^2,\ x^2+y^2 \leq 1\}$.

1. Exprésese A como unión de variedades conexas.

2. Calcúlense el supremo y el ínfimo de la función $f(x,y,z,t) = (x^2 + z^2)e^{-t}$ en A.

Problema 15.35. Se consideran el conjunto $M = \{(x,y,z) \in \mathbb{R}^3 :\ z(x^2 + y^2) \leq 1,\ z \geq 1\}$ y la función $f(x,y,z) = xy + z(x^2 + y^2)$. Calcúlense el supremo y el ínfimo de f en M.

Problema 15.36. Demuéstrese que el volumen máximo de un paralelepípedo rectangular inscrito en el elipsoide $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$, es $\frac{8abc}{3\sqrt{3}}$

Problema 15.37. Sea $f : M \to N$ una función de clase C^1 entre dos variedades diferenciables de dimensiones $m \geq n$. Sea $b \in N$ un valor regular de f. Probar que $f^{-1}(b)$ es una variedad diferenciable contenida en M.

Problema 15.38. Sea $f : M \to N$ una función de clase C^1 entre dos variedades diferenciables de dimensiones $m \leq n$, y supongamos que f es regular en $a \in M$. Probar que existe U entorno abierto de a tal que $f(U)$ es una variedad diferenciable contenida en N, y $f_U : U \to f(U)$ es un difeomorfismo.

Problema 15.39. Sea $f : M \to N$ una función de clase C^1 entre dos variedades diferenciables de dimensiones m y n respectivamente. Probar que el conjunto de puntos regulares de f es abierto en M. Además, si $m \geq n$, probar que el conjunto de valores regulares de f es abierto en N.

Problema 15.40. Se dice que $C \subset \mathbb{R}^n$ es un cuerpo convexo de clase C^p en $\mathbb{R}^n$ si C es cerrado, convexo, con interior no vacío, y su frontera ∂C es una variedad de dimensión $n-1$ y de clase C^p en $\mathbb{R}^n$. Demostrar que todo subconjunto convexo y compacto C de $\mathbb{R}^n$ puede aproximarse por cuerpos convexos de clase C^∞, en el sentido siguiente: para todo abierto U de $\mathbb{R}^n$ con $C \subset U$, existe D cuerpo convexo de clase C^∞ en $\mathbb{R}^n$ tal que $C \subset D \subset U$.

Indicación: Seguir estos pasos:

1. Por el problema 11.13, existe una función convexa $f : \mathbb{R}^n \to [0, \infty)$ de clase C^∞ tal que $C = f^{-1}(0)$, y $Df(x) \neq 0$ si $x \in \mathbb{R}^n \setminus C$. Defínase $C_k = f^{-1}\left((-\infty, 1/k]\right)$ para cada $k \in \mathbb{N}$.

2. Probar que C_k es convexo y compacto, y que $C \subset C_{k+1} \subset C_k$ para todo $k \in \mathbb{N}$. Para ver la acotación de C_k puede ayudar el problema 6.33.

3. Probar que todo número $r > 0$ es un valor regular de f, y concluir que $\partial C_k = \{x \in \mathbb{R}^n : f(x) = 1/k\}$ es una variedad de dimensión $n-1$ y de clase C^∞ de $\mathbb{R}^n$. Es decir, C_k es un cuerpo convexo de clase C^∞ para todo $k \in \mathbb{N}$.

4. Demostrar que
$$\lim_{k\to\infty} d(C, \partial C_k) = 0,$$
donde $d(C, \partial C_k)$ es la distancia de C a la frontera de C_k (razonar por reducción al absurdo, usando la compacidad de C_1 y la continuidad de la función f).

5. Concluir que si U es un abierto que contiene a C entonces existe $k_U \in \mathbb{N}$ tal que $C \subset C_k \subset U$ para todo $k \geq k_U$.

Problema 15.41. Sea D un convexo compacto de $\mathbb{R}^n$, y sea $f : D \to \mathbb{R}$ una función convexa y continua. Demostrar que existe una sucesión de funciones $(f_k) : D \to \mathbb{R}$ convexas y de clase C^∞ tal que $\lim_{k\to\infty} f_k = f$ uniformemente en D. (Decir que f_k es de clase C^∞ en el compacto D es, por definición, decir que f_k tiene una extensión de clase C^∞ a un abierto U_k que contiene a D.) *Indicación:* Seguir estos pasos:

1. Sea N una cota superior de f en D. Defínase $C = \{(x, z) \in \mathbb{R}^n \times \mathbb{R} : x \in D, f(x) \leq z \leq N\}$. Probar que C es un convexo compacto de $\mathbb{R}^{n+1}$.

2. Aplicar el problema anterior para encontrar una sucesión de cuerpos convexos compactos de clase C^∞ que aproxima a C. Para cada $k \in \mathbb{N}$, póngase $U_k = P(\text{int}(C_k))$, donde $P : \mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1} \to \mathbb{R}^n$ es la proyección $P(x, z) = x$, y defínase $f_k : U_k \to \mathbb{R}$ por

$$f_k(x) = \text{ínf}\{t \in \mathbb{R} : (x, t) \in C_k\}.$$

   Pruébese que f_k es convexa y que, geométricamente, la gráfica de f corresponde a la *parte de abajo* de la intersección de ∂C_k con el abierto $P^{-1}(U_k)$.

3. Demostrar que la gráfica de f_k es una variedad diferenciable de dimensión $n - 1$ en $\mathbb{R}^n$.

4. Concluir que f_k es una función convexa de clase C^∞en U_k, y que

$$\lim_{k\to\infty} f_k = f$$

   uniformemente en D.

Apéndice A

Si bien el cálculo diferencial es una herramienta imprescindible en muchas ramas de las matemáticas, lo cierto es que no llega a alcanzar su máxima potencia y profundidad hasta que no se combina con el cálculo integral y la teoría de la medida. En este sentido, conviene quizás anunciar algunos resultados que pertenecen esencialmente al ámbito del cálculo diferencial aunque en su enunciado y demostración intervengan la integral y la teoría de la medida, lo que quizás es la causa de que, con demasiada frecuencia, sean en buena parte ignorados en los programas de las asignaturas de análisis (en las de cálculo diferencial porque no pueden demostrarse sin ayuda de esas otras herramientas, y en las asignaturas más avanzadas bien por falta de tiempo, bien porque se salen de la corriente principal de resultados). Sin embargo se trata de teoremas fundamentales, que en nuestra opinión todo aquél a quien le guste el análisis (y seguramente también todo licenciado en matemáticas) debería conocer. Un buen alumno de segundo, después de haber cursado la asignatura de cálculo integral que suele impartirse a continuación de ésta, no debería tener dificultades para apreciar los enunciados e incluso estudiar (el verano puede ser un momento propicio) las demostraciones y aplicaciones de los teoremas que a continuación enunciamos.

A propósito de puntos y valores regulares de las funciones diferenciables (estudiados en los capítulos 14 y 15) es de obligada mención el teorema de Sard: si una función $f : \mathbb{R}^m \to \mathbb{R}^n$ es de clase C^p con p suficientemente grande, entonces el conjunto de valores regulares de f tiene medida cero en $\mathbb{R}^n$. Es decir, ¡casi todos los valores que toma la función son valores regulares! En particular, $f^{-1}(b)$ es una subvariedad de $\mathbb{R}^m$ para casi todo $b \in f(\mathbb{R}^n)$. El resultado es también válido para funciones entre variedades diferenciables.

Teorema 15.23 (de Sard)**.** *Si M es una variedad diferenciable de dimensión m y $f : M \to \mathbb{R}^n$ es una función de clase C^p, con $p > \max\{m-n, 0\}$, entonces el conjunto de valores regulares de f tiene medida cero en $\mathbb{R}^n$.*

La hipótesis sobre el orden de diferenciabilidad de f es necesaria, como demuestra el clásico ejemplo de Whitney [15] que ya citamos en el capítulo 10: ver la observación 10.7.

Las aplicaciones del teorema de Sard son numerosas, y puede llegar a decirse que existen áreas de la matemática, como la topología diferencial, que dependen totalmente de este resultado. En el libro de Milnor [9] puede encontrarse una demostración relativamente simple del teorema de Sard en el caso en que f es de clase C^∞ (lo que suele ser suficiente en la inmensa mayoría de las aplicaciones). Para una demostración en el caso general puede consultarse el artículo de Bates [3], en donde incluso se prueba que la hipótesis de diferenciabilidad sobre f puede relajarse: basta pedir que f sea de clase C^{p-1} con derivada $(p-1)$-ésima Lipschitz.

Al hilo de funciones Lispchitz y desigualdades del valor medio, no debe olvidarse el teorema de Rademacher: las funciones Lipschitz es derivable en casi todo punto.

Teorema 15.24 (de Rademacher)**.** *Sea $f : \mathbb{R}^n \to \mathbb{R}^m$ una función Lipschitz. Entonces f es diferenciable en casi todo punto (es decir, el conjunto de puntos donde f no es diferenciable tiene medida cero).*

Las aplicaciones de este teorema son también numerosas, y nuevamente puede decirse que sin el mismo, algunas áreas de la matemática, como la teoría geométrica de la medida, serían muy diferentes. El libro de Smith [11] contiene una demostración de este resultado.

Finalmente, sobre los métodos y teoremas de aproximación de funciones continuas o diferenciables por funciones diferenciables de clase C^∞, es preciso mencionar el método de aproximación por convolución integral con sucesiones *molificantes*, es decir, con núcleos integrales que son diferenciables de clase C^∞, tienen integral unidad, y que *convergen* a la función delta de Dirac (la que vale infinito en el origen y cero en el resto, y que *por definición* tiene integral uno). Este método ya aparece en el artículoss clásico de Whitney sobre aproximación y extensión de funciones e inmersión de variedades diferenciables; consultar por ejemplo [14, 6, 12].

Apéndice B

En este segundo Apéndice vamos a presentar una demostración elemental del teorema del punto fijo de Brouwer que sólo requiere cálculo diferencial e integral de funciones de varias variables. Esta demostración se debe a J. Milnor [10]. El teorema se deducirá de otro resultado también clásico de la Topología diferencial, conocido como el *teorema de la bola peluda*. Para cada $n \in \mathbb{N}$, denotemos por $\mathbb{S}^{n-1}$ el conjunto $\{(x \in \mathbb{R}^n : x_1^2 + ... + x_n^2 = 1\}$, es decir la esfera unidad de $\mathbb{R}^n$. Un campo de vectores tangentes a $\mathbb{S}^{n-1}$ es una aplicación $f : \mathbb{S}^{n-1} \to \mathbb{R}^n$ tal que $\langle x, f(x)\rangle = 0$ para cada $x \in \mathbb{S}^{n-1}$.

Teorema 15.25. *Si n es par entonces no existe ningún campo* continuo *de vectores tangentes a $\mathbb{S}^n$ que no se anule en algún punto.*

Es decir, si n es par, no existe ninguna aplicación continua $f : \mathbb{S}^{n-1} \to \mathbb{R}^n$ tal que $f(x) \neq 0$ y $\langle x, f(x)\rangle = 0$ para cada $x \in \mathbb{S}^{n-1}$ (cabría decir, pues, que una cabeza esférica totalmente recubierta de pelo no puede peinarse sin incurrir en discontinuidades). La hipótesis de que n sea par es por supuesto esencial: si n es impar entonces la fórmula $f(x_1, ..., x_{n+1}) = (x_2, -x_1, ..., x_{n+1}, -x_n)$ define un campo continuo de vectores unitarios tangentes a $\mathbb{S}^n$.

La demostración del teorema de la bola peluda se basará en los dos lemas siguientes.

Lema 15.26. *Sea A un compacto de $\mathbb{R}^n$, y $A \ni x \mapsto v(x) \in \mathbb{R}^n$ un campo vectorial de clase C^1 definido en un abierto U que contiene a A. Para cada $t \in \mathbb{R}$ definamos $f_t : A \to \mathbb{R}^n$ por*

$$f_t(x) = x + tv(x).$$

Entonces existe $\varepsilon > 0$ tal que para cada $t \in (-\varepsilon, \varepsilon)$ la aplicación $f_t : A \to \mathbb{R}^n$ es inyectiva y transforma A en una región $f_t(A)$ cuyo volumen viene expresado por un polinomio en t.

Demostración. Puesto que A es compacto y v es de clase C^1, esta aplicación es Lipschitz en A, luego existe $c > 0$ tal que

$$|v(x) - v(y)| \leq c|x - y|$$

para todo $x, y \in A$. Definamos $\varepsilon = 1/c$. Entonces para cada $t \in (-\varepsilon, \varepsilon)$ la aplicación $f_t : A \to \mathbb{R}^n$ es inyectiva (en efecto, si $f_t(x) = f_t(y)$ entonces $x - y = t\,(v(y) - v(x))$, luego $|x - y| \leq |t|c|x - y|$, lo que implica que $x = y$ porque $|t|c < 1$). La matriz de $Df_t(x)$ es de la forma $I + t[\partial v_i/\partial x_j]$, y un cálculo directo muestra que su determinante es una función polinomial de t, de la forma

$$\det Df_t(x) = 1 + tg_1(x) + ... + t^n g_n(x),$$

donde las g_j son funciones continuas. Este determinante es positivo para $|t|$ suficientemente pequeño. Podemos suponer pues que c es suficientemente grande para que el determinante sea positivo para todo $t \in (-\varepsilon, \varepsilon)$. Aplicando el teorema del cambio de variable, podemos calcular el volumen de $f_t(A)$ integrando este determinante sobre A, es decir:

$$\text{vol}\,(f_t(A)) = \int_A (1 + tg_1(x) + ... + t^n g_n(x))\,dx = a_0 + a_1 t + ... + a_n t^n,$$

donde $a_0 = \mathrm{vol}(A)$ y $a_j = \int_A g_j(x)dx$ para $j = 1, ..., n$. □

Ahora supongamos que existe $v : \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$ de clase C^1 y tal que $\langle u, v(u)\rangle = 0$ para todo $u \in \mathbb{S}^{n-1}$ (es decir v es un campo de clase C^1 de vectores unitarios tangentes a $\mathbb{S}^{n-1}$), y definamos

$$f_t(u) = u + tv(u).$$

Nótese que $|u + tv(u)| = \sqrt{1 + t^2}$.

Lema 15.27. *Con la notación anterior, si t es suficientemente pequeño entonces f_t es una biyección entre S^{n-1} y la esfera de centro 0 y radio $\sqrt{1 + t^2}$.*

Demostración. Sea $A = \{x \in \mathbb{R}^n : \frac{1}{2} \leq |x| \leq \frac{3}{2}\}$. Extendamos las aplicaciones v y f_t a A mediante la fórmulas

$$v(ru) = rv(u), \quad f_t(ru) = ru + trv(u)$$

para cada $u \in \mathbb{S}^{n-1}$, $r \in [\frac{1}{2}, \frac{3}{2}]$. Nótese que

$$f_t(ru) = rf_t(u).$$

Se sigue que f_t lleva la esfera de centro 0 y radio r dentro de la esfera de centro 0 y radio $r\sqrt{1 + t^2}$.

Sea ahora t tal que

$$|t| < \text{mín}\{\frac{1}{3}, \frac{1}{c}\},$$

donde c es como en la demostración del lema anterior. Fijemos $u_0 \in \mathbb{S}^{n-1}$, y consideremos la aplicación $\varphi : A \to A$ definida por

$$\varphi(y) = u_0 - tv(y).$$

Es inmediato comprobar que φ es contractiva (tiene constante de Lipschitz menor que 1). Por el Teorema del punto fijo para aplicaciones contractivas existe un único $y_0 \in A$ tal que $\varphi(y_0) = y_0$, es decir

$$u_0 - tv(y_0) = y_0,$$

lo que equivale a decir que la ecuación $f_t(y) = u_0$ tiene como única solución $y = y_0$. Multiplicando por $\sqrt{1 + t^2}$ obtenemos

$$\sqrt{1 + t^2}u_0 = \sqrt{1 + t^2}f_t(y_0) = f_t\left(\sqrt{1 + t^2}y_0\right),$$

y recordando que f_t lleva la esfera de radio r dentro de la esfera de radio $r\sqrt{1 + t^2}$, se sigue que $x_0 := \sqrt{1 + t^2}y_0$ está en la esfera de radio 1. Esto prueba que t_t aplica $\mathbb{S}^{n-1}$ en la esfera de centro radio $\sqrt{1 + t^2}$ de manera sobreyectiva. Por otro lado, la inyectividad de esta aplicación queda garantizada por el lema anterior. □

Con esto ya podemos dar una demostración del teorema de la bola peluda en el siguiente caso particular.

Teorema 15.28. *Si n es par entonces no existe ningún campo de clase C^1 de vectores unitarios y tangentes a $\mathbb{S}^{n-1}$.*

Demostración. Supongamos que existiera un tal campo v. Sean A, t, y f_t como en la demostración del lema anterior. Entonces se tiene que f_t aplica la esfera de centro 0 y radio r de manera sobreyectiva en la esfera de centro 0 y radio $r\sqrt{1+t^2}$. Por tanto f_t aplica la región A de forma sobreyectiva en el conjunto $\{x \in \mathbb{R}^n \ : \ \frac{1}{2}\sqrt{1+t^2} \leq |x| \leq \frac{3}{2}\sqrt{1+t^2}\}$. Consiguientemente,

$$\mathrm{vol}\,(f_t(A)) = \left(\sqrt{1+t^2}\right)^n \mathrm{vol}(A).$$

Ahora bien, si n es par este volumen no es una función polinomial de t, en contradicción con el Lema 15.26. $\square$

Demostración del Teorema 15.25.

Supongamos que exista $v : \mathbb{S}^{n-1} \to \mathbb{R}^n$ tal que $v(x) \neq 0$ y $\langle x, v(x)\rangle = 0$ para todo $x \in \mathbb{S}^{n-1}$. Sea

$$m =: \min_{u\in\mathbb{S}^{n-1}} |v(u)| > 0.$$

Por el teorema de aproximación de Weierstrass existe una aplicación polinomial $p : \mathbb{S}^{n-1} \to \mathbb{R}^n$ tal que

$$|p(u) - v(u)| < \frac{m}{2}$$

para todo $u \in \mathbb{S}^{n-1}$. Definamos $w : \mathbb{S}^{n-1} \to \mathbb{R}^n$ por la fórmula

$$w(u) = p(u) - \langle p(u), u\rangle u.$$

Es inmediato comprobar que $\langle u, w(u)\rangle = 0$, es decir que w es tangente a $\mathbb{S}^{n-1}$, y también (usando que $\langle u, v(u)\rangle = 0$) que

$$|w(u) - p(u)| = |\langle p(u), u\rangle| < \frac{m}{2},$$

lo que junto con $|p(u) - v(u)| < \frac{m}{2}$ y la desigualdad triangular prueba que $|w(u)| > 0$ para todo $u \in \mathbb{S}^{n-1}$. Entonces el cociente $w(u)/|w(u)|$ define un campo C^∞ de vectores unitarios tangentes a $\mathbb{S}^{n-1}$. Si $n-1$ es par, esto es imposible, por el teorema anterior. $\square$

A continuación deduciremos el teorema del punto fijo de Brouwer del teorema de la bola peluda.

Teorema 15.29 (del punto fijo de Brouwer)**.** *Si B es homeomorfo a una bola cerrada de $\mathbb{R}^n$, entonces para toda aplicación continua $f : B \to B$ existe al menos un $x_0 \in K$ tal que $f(x_0) = x_0$.*

Demostración. Podemos suponer que $B = B_n$ es la bola unidad de $\mathbb{R}^n$. Si f no tiene al menos un punto fijo entonces podemos obtener un campo $w : B \to \mathbb{R}^n$ tal que $w(u) = u$ para cada $u \in \mathbb{S}^{n-1}$ (es decir, w coincide con la normal exterior a ∂B en cada punto de ∂B). Por ejemplo, podemos definir

$$w(x) = x - \frac{1 - |x|^2}{1 - \langle x, f(x)\rangle}\, f(x).$$

Obviamente $w(x) = x$ si $|x| = 1$, y $w : B \to \mathbb{R}^n$ es continua. Cuando $\{x, f(x)\}$ son linealmente independientes, es claro que $w(x) \neq 0$. Y cuando estos vectores son linelmente dependientes se tiene que $\langle x, x\rangle f(x) = \langle x, f(x)\rangle x$, lo que implica que $w(x) = (x - f(x))/(1 - \langle x, f(x)\rangle) \neq 0$.

Ahora, mediante una proyección estereográfica, vamos a trasplantar este campo $w(x)$ al hemisferio sur de la esfera S^n de $\mathbb{R}^{n+1}$. Identificando $\mathbb{R}^n$ con el hiperplano $x_{n+1} = 0$ de $\mathbb{R}^{n+1}$, que pasa por el ecuador de $\mathbb{S}^n$, definimos la proyección estereográfica desde el polo norte $(0, ..., 0, 1)$ de $\mathbb{S}^n$ por

$$s(x) = \frac{1}{1 + |x|^2} \left(2x_1, ..., 2x_n, |x|^2 - 1\right).$$

La aplicación s lleva cada punto x de $B_n = B_{n+1} \cap \{x_{n+1} = 0\}$ a un punto $u = s(x)$ de $\mathbb{S}^n \cap \{z \in \mathbb{R}^{n+1} : z_{n+1} \leq 0\}$. Aplicando la derivada de s en x al vector $w(x)$ obtenemos un vector $W(u)$ tangente a $\mathbb{S}^n$ en el punto $u = s(x)$. El vector $W(u)$ puede describirse como el vector velocidad $ds(x + tw(x))/dt$ de la curva esférica $t \mapsto s(x + tw(x))$, evaluado en $t = 0$. De este modo obtenemos un campo vectorial continuo tangente al hemisferio sur de $\mathbb{S}^n$, que no se anula en ningún punto de este. En cada punto $u = s(u)$ del ecuador de esta esfera, es fácil ver (usando que $w(u) = u$) que $W(u) = (0, ..., 0, 1)$ apunta directamente hacia el norte.

De manera análoga, usando una proyección estereográfica desde el polo sur, podemos trasplantar el campo vectorial $-w(x)$ para obtener un campo vectorial continuo definido en el hemisferio norte de $\mathbb{S}^n$, que no se anula en ningún punto de este, y que también apunta directamente hacia el norte en el ecuador. Juntando estos dos campos vectoriales, obtenemos pues un campo vectorial continuo W, tangente a $\mathbb{S}^n$, que no se anula en ningún punto de esta esfera. Si n es par, esto es imposible por el Teorema 15.25.

Esto prueba el teorema del punto fijo de Brouwer en el caso en que n es par. Ahora supongamos que $n = 2k - 1$ es impar y que existiera una aplicación continua $f : B_{2k-1} \to B_{2k-1}$ sin ningún punto fijo. Entonces la fórmula $F(x_1, ..., x_{2k}) = (f(x_1, ..., x_{2k-1}), 0)$ definiría una aplicación continua $F : B_{2k} \to B_{2k}$ sin ningún punto fijo. Lo que según acabamos de probar es imposible. Por tanto el enunciado del teorema se cumple también en el caso en que n es impar. $\square$

Bibliografía

[1] T. M. Apostol, *Análisis Matemático*, Reverté, 1979.

[2] T. M. Apostol, *Calculus*, Reverté, 1979.

[3] S. M. Bates, *Toward a precise smoothness hypothesis in Sard's theorem*, Proc. Amer. Math. Soc. 117 (1993), no. 1, 279-283.

[4] F. Bombal, L. Rodríguez Marín, G. Vera, *Problemas de Análisis Matemático*, Editorial AC, 1987.

[5] C. Fernández, F. Vázquez, J. M. Vegas, *Cálculo diferencial de varias variables*, Thomson, 2002.

[6] M. Hirsch, *Differential Topology*, Undergraduate Texts in Mathematics, Springer-Verlag, New York.

[7] J. E. Marsden, M. J. Hoffman, *Elementary Classical Analysis*, W. H. Freeman and Company, New York, 1993.

[8] J. E. Marsden, A. J. Tromba, *Cálculo vectorial*, Addison-Wesley Iberoamericana, 1991.

[9] J. W. Milnor, *Topology from the Differentiable Viewpoint* (1965) Princeton Landmarks in Mathematics, Princeton University Press, reedición de 1997.

[10] J. W. Milnor, *Analytic proofs of the "hairy ball theorem" and the Brouwer fixed-point theorem*. Amer. Math. Monthly 85 (1978), no. 7, 521–524

[11] K. T. Smith, *Primer of Modern Analysis*, Springer, New York, 1983.

[12] E. Stein, *Singular integrals and differentiability properties of functions*. Princeton, University Press, 1970.

[13] J. R. L. Webb, *Functions of Several Real Variables*, Ellis Horwood, London 1991.

[14] H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc. 36 (1934), 63–89.

[15] H. Whitney, *A function not constant on a connected set of critical points*, Duke Math. J. 1 (1935), 514-517.