Índice general

Capítulo 1

El Espacio Normado $\mathbb{R}^n$

En este curso supondremos conocida la estructura de $\mathbb{R}$ y su topología, así como las propiedades de las funciones continuas o derivables de una variable. Todo este bagaje inicial se usará, sin necesidad de hacer las demostraciones que correspondan. Nuestro objetivo fundamental serán las funciones de varias variables, más exactamente el cálculo diferencial para funciones definidas en subconjuntos de $\mathbb{R}^n$. Muchos de los resultados serán extensiones a las varias variables de otros ya estudiados en una variable, otros en cambio serán nuevos por responder a problemas que no tienen sentido en 1-variable.

A pesar de que sólo estamos interesados en las funciones de un número finito de variables, muchos conceptos y demostraciones se establecerán para funciones definidas entre espacios vectoriales reales de cualquier dimensión, sentando así las bases para un Cálculo Diferencial en espacios funcionales y, evitando además el uso de coordenadas, cuando éstas no sean necesarias. Consideramos pues funciones $f : A \subset E \to F$, siendo E y F espacios vectoriales sobre el cuerpo de los números reales de dimensión ≥ 1 . Cuando E sea de dimensión n diremos que f es una función de las n-variables reales $(x_1, x_2, \ldots, x_n)$. En tal caso, puesto que E es isomorfo a $\mathbb{R}^n$, estamos identificando E y $\mathbb{R}^n$. También es habitual decir que una función que toma sus valores en $\mathbb{R}$ (i.e. $F = \mathbb{R}$) es una función escalar, mientras que la función se dice vectorial cuando $\dim(F) > 1$. En el caso en que F sea el producto de un número finito de espacios vectoriales,

$$f : A \subset E \to F_1 \times F_2 \times \cdots F_p,$$

si se denota por $f_i(x)$ la i-ésima coordenada de $f(x)$, escribiremos $f = (f_1, \ldots, f_p)$ y diremos que las funciones $f_i : A \subset E \to F_i;\ i = 1, 2, \ldots, p$, son las funciones coordenadas de f.

Sea E un espacio vectorial sobre un cuerpo $\mathbb{K} = \mathbb{R}$ ó $\mathbb{C}$ indistintamente. Una *norma* sobre E es una aplicación de E en $\mathbb{R}$ que satisface las tres propiedades siguientes:

NOR1. $\|x\| = 0$ si y sólo si $x = 0$

NOR2. $\|\lambda x\| = |\lambda|\,\|x\|, \quad \forall \lambda \in \mathbb{K},\ \forall x \in E$

NOR3. $\|x + y\| \leq \|x\| + \|y\|,\ \forall x, y \in E$

Al número real $\|x\|$ se le denomina norma del vector x y se dice que el par $(E, \|\ \|)$ es un espacio normado. En lo sucesivo, todos los espacios vectoriales serán reales, es decir $\mathbb{K} = \mathbb{R}$.

Ejemplos 1.1 (1) Las únicas normas sobre $\mathbb{R}$ son el valor absoluto y sus múltiplos positivos. En efecto, sea $\|\ \|$ una norma cualquiera sobre $\mathbb{R}$ y sea $k = \|1\|$. Entonces

$$\|x\| = \|x{\cdot}1\| = |x|\|1\| = k|x|.$$

Para probar que k debe ser un número real mayor estrictamente que 0, sólo hay que tener en cuenta que toda norma sobre un espacio vectorial E satisface las propiedades:

NOR4. Toda norma es simétrica, es decir $\| - x\| = \|x\|$, para todo $x \in E$,

NOR5. La norma de todo vector de E es un número real positivo,

La propiedad NOR4 se obtiene trivialmente de la segunda condición de norma. NOR5 se demuestra así:

$$0 = \|x - x\| \leq \|x\| + \| - x\| = 2\|x\| \quad \Rightarrow \quad \|x\| \geq 0.$$

(2) En $\mathbb{R}^n$ las normas más utilizadas son

$$\|(x_1, \dots, x_n)\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p},\ p \geq 1$$

$$\|(x_1, \dots, x_n)\|_\infty = \text{máx}\{|x_1|, \dots, |x_n|\}.$$

La comprobación, en las del tipo p, de la tercera propiedad de norma se basa en la desigualdad de Hölder (Ver ejercico 1A , aunque para el caso $p = 2$ cabe una demostración alternativa, basada en la desigualdad de Cauchy-Schwartz, que veremos a continuación. La $\|\ \|_2$ es la norma de la geometría euclídea, ella forma parte del importante grupo de normas que se derivan de un producto escalar las Normas Euclídeas. La norma $\| \|_\infty$ es conocida como la norma producto.

Como es bien conocido mediante la igualdad

$$(x, y) \cdot (u, v) = xu + yv,$$

se define una aplicación bilineal de $\mathbb{R}^2 \times \mathbb{R}^2$ en $\mathbb{R}$, el producto escalar euclídeo (ejercicio). Obviamente, se tiene que el producto escalar de un vector por sí mismo es justamente el cuadrado de su norma euclídea: $(x, y) \cdot (x, y) = x^2 + y^2 = \|(x, y)\|^2$.

Vamos a ver que se satisface la desigualdad,

$$\big|(x, y) \cdot (u, v)\big| \leq \|(x, y)\|\|(u, v)\|,$$

conocida como *Desigualdad de Cauchy-Schwartz.*
Es obvio que esta desigualdad es cierta sii

$$\begin{aligned}
&(xu + yv)^2 \leq (x^2 + y^2)(u^2 + v^2) \quad \text{sii} \\
&2xyuv \leq x^2v2 + y^2u^2 \quad \text{sii} \\
&x^2v2 + y^2u^2 - 2xvyu \geq 0,
\end{aligned}$$

y es claro que esto último es cierto pues $x^2v2 + y^2u^2 - 2xvyu = (xv - yu)^2$.

Deducimos entonces que

$$\begin{aligned}
\|(x, y) + (u, v)\|^2 &= \Big((x, y) + (u, v)\Big) \cdot \Big((x, y) + (u, v)\Big) \\
&= (x, y) \cdot (x, y) + 2(x, y) \cdot (u, v) + (u, v) \cdot (u, v) \\
&\leq \|(x, y\|^2 + 2\|(x, y)\|\|(u, v)\| + \|(u, v)\|^2 = \Big(\|(x, y)\| + \|(u, v)\|\Big)^2.
\end{aligned}$$

Las normas anteriores pueden definirse sobre productos finitos de espacios normados, o sea si E en en vez de $\mathbb{R}^n$ es un producto de los espacios normados E_i, $i = 1, \ldots, n$, entonces sobre el espacio vectorial $E = E_1 \times \cdots \times E_n$ podemos considerar de forma análoga la norma producto y las normas p.

Ejercicio. Una norma $\| \|$ sobre $\mathbb{R}^n$ la llamaremos monótona si satisface la condición:

$$|x_i| \leq |y_i|, \quad i = 1, \ldots, n \quad \Rightarrow \quad \|(x_1, \ldots, x_n)\| \leq \|(y_1, \ldots, y_n)\|.$$

1. Probar que las normas $\|\ \|_p$ y la norma producto son normas monótonas de $\mathbb{R}^n$.

2. Probar que la norma $\|(x,y)\| = |x|+|x-y|$ no es una norma monótona de $\mathbb{R}^2$.

3. Si $E_i,\ i=1,\ldots,n$ son espacios normados y $\|\ \|^*$ es una norma monótona de $\mathbb{R}^n$, entonces mediante la fórmula

$$\|(x_1,\ldots,x_n)\| = \big\|\big(\|x_1\|,\ldots,\|x_n\|\big)\big\|^*$$

   se define una norma en $E = E_1 \times E_2 \times \cdots \times E_n$.

4. Si $\|\ \|_1,\ldots,\|\ \|_n$ son normas sobre el mismo espacio vectorial E y $\|\ \|^*$ es una norma monótona de $\mathbb{R}^n$, entonces mediante la fórmula

$$\|x\| = \big\|\big(\|x\|_1,\ldots,\|x\|_n\big)\big\|^*$$

   se define una norma en E.

5. Utilizar uno de los dos apartados anteriores para justificar que las siguientes expresiones son normas

   en $\mathbb{R}^2$:

$$\|(x,y)\| = |x|+|y|+\sqrt{x^2+y^2}$$
$$\|(x,y)\| = \sqrt{(\text{máx}\{|x|,|y|\})^2+x^2+y^2+(|x|+|y|)^2}$$

   en $\mathbb{R}^3$:

$$\|(x,y,z)\| = 2|x|+3|y|+|z|$$
$$\|(x,y,z)\| = \text{máx}\{|x|,\sqrt[3]{|y|^3+|z|^3}\}$$
$$\|(x,y,z)\| = \sqrt{x^2+y^2+z^2+(|x|+|y|+|z|)^2}$$

6. Procediendo de la misma forma, construir más normas.

1.2 Toda norma lleva asociada de forma natural una distancia d definida por $d(x, y) = \|x - y\|$. En particular, en $\mathbb{R}$ la distancia asociada a la norma "valor absoluto" es la usual $d(x, y) = |x - y|$. Para $E = \mathbb{R}^2$ la distancia asociada a $\| \ \|_2$ es la distancia euclídea. Para $\| \ \|_\infty$ es fácil ver que si P, Q son dos puntos del plano situados en un rectángulo de lados paralelos a los ejes de medidas s, t, entonces $d_\infty(P, Q) \leq \max\{t, s\}$.

Definición 1.3 En un espacio normado $(E, \| \ \|)$,

i) al conjunto de puntos de E que distan de un punto a menos que $r > 0$ lo llamaremos *bola abierta* de centro $a \in E$ y radio $r > 0$ y se denotará por $B(a, r)$. Luego $B(a, r) = \{x \in E : \|x - a\| < r\}$. Análogamente la *bola cerrada* será, $B[a, r] = \{x \in E : \|x - a\| \leq r\}$ y la *esfera*, $S[a, r] = \{x \in E : \|x - a\| = r\}$;

ii) un subconjunto $A \subset E$ se dirá acotado si existe alguna constante $\alpha \geq 0$ tal que $\|x\| \leq \alpha$ para todo $x \in A$. Equivalentemente si A está contenido en alguna bola.

1.4 Algunas de las propiedades elementales de las bolas en un espacio normado son las siguientes:

1. *"La geometría" de las bolas no depende del centro ni del radio. Esto es debido a la igualdad* $B(a, r) = a + rB(0, 1)$ (ejercicio), que nos dice que toda bola se obtiene mediante la traslación de una homotética de la bola unidad (de centro 0 y radio 1).

2. *Cada bola tiene infinitos puntos.* En efecto, consideremos la bola $B(a, r)$. Entonces para cada de los infinitos puntos $b \in E$, $b \neq 0; b \neq a$ es inmediato ver que los infinitos puntos $a + tb$, con $t \in (-r/\|b\|, r/\|b\|)$ pertenecen a $B(a, r)$.

3. *Para cada par de puntos distintos $a, b \in E$ existes dos bolas centradas en a y en b respectivamente que son disjuntas:* Si $a \neq b$ entonces $r = 1/2\|a - b\|$ es mayor estrictamente que 0 y las bolas $B(a, r)$ y $B(b, r)$ son disjuntas, pues si $x \in B(a, r)$ entonces $\|a - b\| \leq \|a - x\| + \|x - b\| < r + \|x - b\|$, lo que implica que $\|x - b\| > r$ y por tanto $x \notin B(b, r)$.

Ejercicio. Dibujar en $\mathbb{R}^2$ una bola respecto a la normas $\| \ \|_\infty$, $\| \ \|_1$, $\| \ \|_2$ y también respecto a la norma $\|(x, y)\| = |x| + |x - y|$.

1. Normas equivalentes

Análogamente a cómo se hacía en $\mathbb{R}$, a partir de la distancia asociada a una norma, podemos dar la definición de sucesión convergente en espacios normados . Concretamente:

Definición 1.5 La sucesión de puntos del espacio normado $(E, \| \|)$ se dice que converge (o que $\| \|$-converge, si pudiera haber confusión) al punto x, $\{x_p\} \to x$, si para cada $\varepsilon > 0$ existe un índice ν tal que si $p \geq \nu$ entonces $\|x_p - x\| < \varepsilon$. En otras palabras, si a partir de ν todos los términos de la sucesión están en la bola $B(x, \varepsilon)$.

Por tanto, es claro que una sucesión no puede converger a dos puntos distintos. Si la sucesión $\{x_p\}$ convergiese a los dos puntos distintos $x_1 \neq x_2$, tomando ε tal que $B(x_1, \varepsilon) \cap B(x_2, \varepsilon) = \emptyset$, todos los términos de la sucesión a partir de uno en adelante deberían estar en ambas bolas, lo cual es absurdo. Cuando la sucesión $\{x_p\}$ converge a x, también se dice que x es el punto límite de la misma y se escribirá $\lim\limits_{p \to \infty} x_p = x$.

Es inmediato comprobar que si $\{x_p\}$ es una sucesión convergente entonces también es de Cauchy, i.e para cada $\varepsilon > 0$ existe un índice ν tal que si $p, q \geq \nu$ entonces $\|x_p - x_q\| \leq \varepsilon$.

Ejemplo 1.6 Es fácil ver que en $\mathbb{R}^2$ respecto a la norma producto, una sucesión $\{(x_p, y_p)\}$ converge al punto (x, y) si y sólo si las sucesiones coordenadas $\{x_p\}$ y $\{y_p\}$ convergen respectivamente a x e y. Igualmente en $\mathbb{R}^n$.

Puesto que el concepto de sucesión convergente lo hemos expresado en términos de la norma, parece posible que una misma sucesión de puntos del espacio vectorial E pueda ser convergente o no dependiendo de cuál sea la norma con la que trabajemos. De hecho, como veremos a continuación, dos normas sobre E que proporcionan las mismas sucesiones convergentes son justamente las que denominaremos equivalentes.

Definición 1.7 Dos normas $\| \|, \| \|^*$ sobre el mismo espacio vectorial E se dicen equivalentes si existen dos números reales α, β mayores estrictamente que 0 tales que $\|x\| \leq \alpha \|x\|^*$; $\|x\|^* \leq \beta \|x\|, \forall x \in E$.

Proposición 1.8 *Sean $\| \|, \| \|^*$ dos normas sobre el mismo espacio vectorial E. Los enunciados siguientes son equivalentes:*

1) *Las normas $\| \|, \| \|^*$ son equivalentes,*

2) *Cada sucesión que converge respecto a una de ellas converge también respecto a la otra y al mismo límite,*

3) *Cada sucesión que converge respecto a una de ellas converge también respecto a la otra.*

Demostración. 1) implica 2) Supongamos que las normas son equivalentes i.e existen α, β mayores estrictamente que 0 tales que $\|x\| \leq \alpha\|x\|^*$; $\|x\|^* \leq \beta\|x\|, \forall x \in E$, y sea $\{x_p\}$ una sucesión $\| \|$-convergente a x. Luego dado $\varepsilon > 0$, a partir de un cierto índice, los términos de la sucesión $\| \|$-distan de x menos que ε. Se deduce entonces que, a partir de ese índice, $\|x_p - x\|^* \leq \beta\|x_p - x\| < \beta\varepsilon$. Es decir la sucesión $\{x_p\}$ también $\| \|^*$-tiende a x.

2) implica 1) Recíprocamente, supongamos que las normas no son equivalentes, por ejemplo que no se satisface $\| \|^* \leq \beta\| \|$ para ningún $\beta > 0$. En particular, que para cada $p \in \mathbb{N}$ existe $x_p \in E$ tal que $\|x_p\|^* > p\|x_p\|$. O lo que es lo mismo, que

$$\frac{\|x_p\|}{\|x_p\|^*} < 1/p,$$

lo que significa que la sucesión, $\{\frac{x_p}{\|x_p\|^*}\}$, $\| \|$-converge a 0. Pero, en cambio, esta sucesión no $\| \|^*$-converge a 0, pues $\left\|\frac{x_p}{\|x_p\|^*}\right\|^* = 1$.

Trivialmente 2) implica 3).

3) implica 2) Teniendo en cuenta que una sucesión $\{x_p\}$ converge a a si y sólo si la sucesión $\{x_p - a\}$ converge a 0, bastará probar que si $\{x_p\}$ es una sucesión que $\| \|$-converge a 0 entonces esta sucesión $\| \|^*$-converge a 0. Por 3) $\{x_p\}$ debe ser $\| \|^*$-convergente. Lo que se trata de probar es que además debe hacerlo a 0. En efecto, supongamos que $\{x_p\}$ $\| \|^*$-converge a b, entonces es obvio que la sucesión $\{2x_p\}$ $\| \|$-converge a 0 y $\| \|^*$-converge a $2b$. Sea $\{y_p\}$ la sucesión definida por $y_p = x_p$ si p es impar e $y_p = 2x_p$ si p es par. Obviamente $\{y_p\}$ $\| \|$-converge a 0, por lo que según 3), debe ser $\| \|^*$-convergente a algún c. Pero la subsucesión de los términos impares $\| \|^*$-converge a b mientras que la de los términos pares converge a $2b$. Por tanto $c = b = 2b$ que implica $b = 0$ como queríamos probar. ∎

Teorema 1.9 *En $\mathbb{R}^n$ todas las normas son equivalentes.*

Demostración. Vamos a probar que cada norma $\| \|$ sobre $\mathbb{R}^n$ es equivalente a la norma $\| \|_\infty$.

1) Denotemos por $e_i = (0, \ldots, 1, \ldots, 0)$ entonces

$$\|x\| = \|(x_1, \ldots, x_n)\| == \|\sum_{i=1}^{n} x_i e_i\|$$

$$\leq \sum_{i=1}^{n} |x_i|\|e_i\| \leq \sum(\|x\|_\infty \|e_i\|) = (\sum \|e_i\|)\|x\|_\infty.$$

Entonces, si $\beta = \sum \|e_i\|$, se ha probado que para todo $x \in \mathbb{R}^n$, $\|x\| \leq \beta\|x\|_\infty$.

2) Para probar la otra desigualdad razonaremos por inducción sobre n. Ya sabemos que es cierta para $n = 1$, pues las únicas normas sobre $\mathbb{R}$ son los múltiplos positivos del valor absoluto. Supongamos cierto en $\mathbb{R}^{n-1}$ la existencia de una desigualdad en este sentido entre cada norma y la norma producto. Para probar esto mismo sucede en $\mathbb{R}^n$, observemos que si $v = (x_1, x_2, \ldots, x_n)$ es un vector no nulo, entonces para cada $x_j \neq 0$ podemos escribir $v = x_j(x_1/x_j, \ldots, 1, \ldots, x_n/x_j)$, de lo que se deduce que $\|v\| = \|v\|_\infty \|u\|$, donde u es un vector que tiene un 1 en alguna de sus coordenadas. La desigualdad que buscamos se obtendría si fuese cierto que existe alguna constante $\alpha > 0$ tal que para cada vector u de este tipo $\|u\|$ fuese mayor que α. Denotemos por $L_1, L_2, \ldots$ los conjuntos formados por los puntos que tienen un 1 en la 1ª coordenada, un 1 en la 2ª coordenada,... y veamos que en cada uno de ellos es cierto lo anterior. Supongamos que no y que por ejemplo existe una sucesión de vectores $u_p = (1, y_2^p, \ldots, y_n^p) \in L_1, p = 1, 2, \ldots$ tal que $\|u_p\| \leq 1/p$. Teniendo en cuenta que

$$\|u_p\| = \|(1, y_2^p, \ldots, y_n^p)\| = \|(1, 0, \ldots, 0) + (0, y_2^p, \ldots, y_n^p)\|,$$

eso significaría que la sucesión $\{(0, y_2^p, \ldots, y_n^p)\}$ $\| \ \|$-converge a $(-1, 0, \ldots, 0)$ (luego también es $\| \ \|$-Cauchy).

Denotemos por $\| \ \|^*$ a la norma sobre $\mathbb{R}^{n-1}$ definida así: $\|(y_2, \ldots, y_n)\|^* = \|(0, y_2, \ldots, y_n)\|$ (comprobar que es norma). Por hipótesis de inducción debe existir una constante $s > 0$ tal que $\|(y_2, \ldots, y_n)\|_\infty \leq s\|(y_2, \ldots, y_n)\|^* = s\|(0, y_2, \ldots, y_n)\|$. De esta desigualdad se deduce entonces que el carácter $\| \ \|$-Cauchy de la sucesión $\{(0, y_2^p, \ldots, y_n^p)\}$ implicaría que la sucesión $\{(y_2^p, \ldots, y_n^p)\}$ es $\| \ \|_\infty$-Cauchy, equivalentemente que las sucesiones coordenadas son sucesiones de Cauchy de números reales, luego convergentes. Si $c_2, c_3, \ldots, c_n$ fuesen sus límites entonces la sucesión $\{(y_2^p, \ldots, y_n^p)\}$ convergería respecto a la norma $\| \ \|_\infty$ al punto $c = (c_2, \ldots, c_n)$, pero entonces, teniendo en cuenta que (según vimos en 1) $\| \ \| \leq \beta\| \ \|_\infty$, la sucesión $\{(0, y_2^p, \ldots, y_n^p)\}$ $\| \ \|$-convergería a $\|(0, c_2, \ldots, c_n)\|$. Habríamos obtenido dos límite diferentes para la misma sucesión. ∎

Corolario 1.10 *$\mathbb{R}^n$ es un espacio normado completo (un espacio de Banach) respecto a cualquier norma i.e., cada sucesión de Cauchy en $(\mathbb{R}^n, \| \|)$ es convergente cualquiera que sea la norma $\| \|$.*

Demostración. Puesto que en $\mathbb{R}^n$ todas las normas son equivalente, una sucesión es $\| \|$-Cauchy si y sólo es $\| \|_\infty$-Cauchy y, por tanto, si y sólo si cada sucesión coordenada es una sucesión de números reales de Cauchy (luego convergente, pues suponemos conocido que $\mathbb{R}$ es completo). Se tiene pues que dicha sucesión, según veíamos antes, es $\| \|_\infty$-convergente lo que, según la proposición 1.8, equivale a ser $\| \|$-convergente. ∎

2. Continuidad y limites de funciones

Comenzamos aquí con el estudio de las funciones de varias variables, y más generalmente con el de las funciones del tipo $f : A \subset E \to F$ con E y F espacios normados. Para ello necesitaremos definir previamente algunos conceptos de topología:

Definición 1.11 Un subconjunto A de un espacio normado $(E, \| \|)$ se dirá entorno del punto a (o equivalentemente que a es un punto interior de A: $a \in \overset{\circ}{A}$) si existe un número real $r > 0$ tal que $B(a, r) \subset A$. Un subconjunto U se dirá abierto cuando sea entorno de todos sus puntos o dicho de otra forma si $\overset{\circ}{U} = U$. A los complementarios de abiertos se les llama cerrados. Nos referiremos a la familia τ de los abiertos como a la topología del espacio normado.

Es obvio que todo conjunto que contenga un entorno de un punto a es también entorno de a y que la intersección de dos entornos del punto a es también entorno de a. Asimismo es claro que la unión arbitraria de abiertos y la intersección finita de abiertos es un abierto.

Proposición 1.12 *Un conjunto U es abierto del espacio normado E si y sólo si U es unión de bolas abiertas.*

Demostración. Si U es abierto entonces $U = \bigcup_{x \in U} B(x, r_x)$. Recíprocamente, cada $B(a, r)$ es un conjunto abierto. En efecto si $x \in B(a, r)$ y $0 < r_x < r - \|x - a\|$ entonces cada $y \in B(x, r_x)$ verifica que $\|y - a\| \leq \|y - x\| + \|x - a\| < r_x + \|x - a\| < r$. ∎

Proposición 1.13 *Dos normas sobre el espacio vectorial E son equivalentes si y sólo si inducen la misma topología sobre E. También, si y sólo si los entornos de cada punto son los mismos para las dos normas.*

Demostración. De acuerdo a la definición de abierto, bastará probar la segunda de las afirmaciones. Sean $\| \;\|$ y $\| \;\|^*$ dos normas sobre E y para $a \in E$, denotemos por $\mathcal{V}(a)$ y $\mathcal{V}_*(a)$ a la familia de entornos de a respecto a esas normas. Supongamos que son equivalentes y que $U \in \mathcal{V}(a)$, es decir que existe $B(a, r) \subset U$. Al ser equivalentes ambas normas existe $\alpha > 0$ tal que $\|x - a\| \leq \alpha \|x - a\|^*$ para todo x. Por tanto si $\|x - a\|^* < r/\alpha$ entonces $\|x - a\| < r$ o sea que $B_*(a, r/\alpha) \subset B(a, r) \subset U$, lo que implica que $U \in \mathcal{V}_*(a)$.

Recíprocamente, supongamos que los entornos de cada punto son los mismos para ambas normas y veamos que entonces las sucesiones convergentes a cada punto son también las mismas para ambas normas. Supongamos que $\{x_p\}$ es una sucesión que $\| \;\|$-convergente a b, y sea $\varepsilon > 0$, como la bola $B_*(b, \varepsilon)$ es un entorno de b (respecto a ambas normas) debe existir $B(b, \varepsilon') \subset B_*(b, \varepsilon)$. Se deduce pues que todos los términos de la sucesión a partir de uno en adelante están en $B(b, \varepsilon')$ y por tanto en $B_*(b, \varepsilon)$, lo que significa que $\{x_p\}$ es una sucesión que $\| \;\|^*$-converge a b. ∎

Ejercicio. Probar que todo subespacio vectorial de $\mathbb{R}^n$ es cerrado.

Definiciones 1.14 La función $f : A \subset E \to F$ se dice:

1) *continua en el punto* $a \in A$, si para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que si $\|x - a\| < \delta$ $(x \in A)$ entonces $\|f(x) - f(a)\| < \varepsilon$.

Es fácil ver que un toda función continua en un punto es *secuencialmente continua* en ese punto, i.e. si la función f es continua en el punto a y $\{x_p\}$ es una sucesión de puntos que converge al punto a, entonces la sucesión de sus imágenes $\{f(x_p)\}$ converge a $f(a)$.

2) *continua en* A, cuando sea continua en todos los puntos de A.

3) *uniformemente continua en* A, si para cada $\varepsilon > 0$ existe un $\delta > 0$ tal que si $\|x - y\| < \delta$ $(x, y \in A)$ entonces $\|f(x) - f(y)\| < \varepsilon$.

De la definición se deduce que si f es uniformemente continua en A entonces es continua en A, es decir continua en cada punto $x \in A$, pero con la particularidad de que el δ no depende más que de ε, es decir δ_x puede tomarse el mismo para todo x.

4) *lipschitziana*, si existe una constante $k > 0$ tal que $\|f(x) - f(y)\| \leq k\|x - y\|$.

Obviamente cada función lipschitziana es uniformemente continua.

Ejemplos 1.15 1. Cada función constante es trivialmente continua.

2. En todo espacio normado $(E, \| \;\|)$ la aplicación norma, $\| \;\| : E \to \mathbb{R}$, es continua.

En efecto, de hecho esta aplicación es lipschitziana, pues como es fácil de comprobar se tiene la siguiente desigualdad (ejercicio):

$$\Big| \|x\| - \|y\| \Big| \leq \|x - y\|$$

3. Sean $(E_i, \| \|)$, $i = 1, 2, \ldots, p$ espacios normados. Consideremos sobre $E = E_1 \times E_2 \times \cdots \times E_p$ la norma producto

$$\|(x_1, \ldots, x_p)\|_\infty = \max_{1 \leq i \leq p} \|x_i\|$$

Entonces, cualquiera que sea j, la aplicación proyección $\pi_j : E \to E_j$ $\pi_j(x_1, \ldots, x_p) = x_j$ es continua (lipschitziana), pues

$$\|\pi_j(x_1, \ldots, x_p) - \pi_j(y_1, \ldots, y_p)\| = \|x_j - y_j\|$$
$$\leq \|(x_1 - y_1, \ldots, x_p - y_p)\|_\infty = \|(x_1, \ldots, x_p) - (y_1, \ldots, y_p)\|_\infty.$$

Proposición 1.16 *La función $f\colon A \subset E \to F$ es continua en A si y sólo si para cada conjunto abierto V de F la antimagen $f^{-1}(V)$ es un abierto de A i.e. existe algún abierto U de E tal que $f^{-1}(V) = U \cap A$.*

Demostración. Supongamos f continua en A y sea V un abierto de F. Veamos $f^{-1}(V)$ abierto de A: si $x \in f^{-1}(V)$ se tiene $f(x) \in V$ luego existe $B(f(x), \varepsilon_x) \subset V$ y por la continuidad de f en x debe existir δ_x tal que $f(B(x, \delta_x) \cap A) \subset B(f(x), \varepsilon_x) \subset V \;\Rightarrow\; B(x, \delta_x) \cap A \subset f^{-1}(V)$. Entonces:

$$f^{-1}(V) \subset \bigcup_{x \in f^{-1}(V)} B(x, \delta_x) \cap A \subset f^{-1}(V).$$

Luego tomando el abierto $U = \displaystyle\bigcup_{x \in f^{-1}(V)} B(x, \delta_x)$ se tiene que $f^{-1}(V) = U \cap A$.

(Probar como ejercicio la implicación en el otro sentido). ∎

Completaremos la serie de definiciones anteriores con la definición de límite. Para ello necesitaremos referirnos previamente a la noción de punto de acumulación:

Definición 1.17 Sea A un subconjunto de un espacio normado E. Un punto a se dice de acumulación del conjunto A (lo que se expresará como que $a \in A'$) si cada entorno de a tiene algún punto de A distinto de a. De hecho en espacios normado el punto a será de acumulación de A si y sólo si cada entorno de A contiene infinitos puntos de A (ejercicio). Observar que a puede ser de acumulación de A sin necesidad de pertenecer a A. Por otra parte también es inmediato probar que si un punto a es interior a A o más generalmente si existe una $B(a,r) \subset A \cup \{a\}$ entonces $a \in A'$.

Definición 1.18 Sea $f\colon A \subset E \to F$ y a un punto de acumulación de A. Diremos que el punto $l \in F$ es límite de la función en el punto a, lo que denotaremos como

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

si para $\varepsilon > 0$, existe $\delta > 0$ tal que si $x \in A,\ x \neq a$ y $\|x-a\| < \delta$ entonces $\|f(x) - l\| < \varepsilon$.

Consecuencias directas de la definición son:

a) Una función no puede tener dos límites diferentes en el mismo punto. Como antes con las sucesiones, la demostración de la unicidad del límite se basa en que puntos distintos admiten entornos disjuntos.

b) Sea $f\colon A \subset E \to F$ y $a \in A$. Se tiene:

- Si a no es de acumulación de A (en ese caso se dice que a es un punto aislado de A) entonces f es continua en a.
- Si $a \in A'$ entonces f es continua en a si y sólo si $\lim_{x\to a} f(x) = f(a)$.

Si a es aislado de A entonces existe δ tal que el único x de A con la propiedad $\|x-a\| < \delta$ es justamente el punto a. Por lo tanto es obvio que f es continua en a.

Si $a \in A' \cap A$, entonces:

(f continua en a) significa:

$$\forall\varepsilon > 0,\ \exists\delta\ \|x-a\| < \delta,\ x \in A\ \Rightarrow\ \|f(x) - f(a)\| < \varepsilon$$

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

$$\forall\varepsilon > 0,\ \exists\delta\ \|x-a\| < \delta,\ x \in A, x \neq a\ \Rightarrow\ \|f(x) - f(a)\| < \varepsilon$$

Obviamente los dos enunciados coinciden pues para $x = a$, $\|f(a) - f(a)\| = 0 < \varepsilon$.

c) Si $f\colon A \subset E \to F$ y $a \in A'$ (respectivamente $a \in A$), entonces $\lim_{x\to a} f(x) = l$ (respect. f continua en a) si y sólo si para cada $V \in \mathcal{V}(l)$ (respect. $\mathcal{V}(f(a))$ exista algún $U \in \mathcal{V}(a)$ tal que $f(U^* \cap A) \subset V$ (respect. $f(U \cap A) \subset V$), donde $U^* = U \setminus \{a\}$.

La demostración consiste en una sencilla adaptación de la definición $\varepsilon-\delta$ en términos de entornos: supongamos que $\lim_{x\to a} f(x) = l$ y sea $V \in \mathcal{V}(l)$.

Entonces existe $\varepsilon > 0$ tal que $B(l,\varepsilon) \subset V$ y para este ε existe δ tal que si $x \in A,\ x \neq a$ y $\|x-a\| < \delta$ entonces $\|f(x)-l\| < \varepsilon$. Dicho de otra forma si $x \in B^*(a,\delta) \cap A$ entonces $f(x) \in B(l,\varepsilon) \subset V$ i.e., si consideramos el entorno $U = B(a,\delta)$ se tiene que $f(U^* \cap A) \subset V$. La implicación en el otro sentido es análoga (ejercicio).

De la definición de límite por entornos dada en b) se deduce:

d) la existencia y el valor del límite de una función en un punto o la continuidad se mantienen, si se cambian las normas de E y F por otro par de normas equivalentes. En particular, si los espacios E y F son de dimensión finita, para estudiar la existencia de límite de una función en un punto o la continuidad, podemos utilizar las normas que queramos.

2.1. Reglas de cálculo para límites

1.19 *Sean $E, F_1, \dots, F_p$ espacios normado $f : A \subset E \to F$ donde $F = F_1 \times F_2 \times \cdots F_p$ con la norma producto. Sea $a \in A'$ y denotemos por $f_i,\ i = 1, 2, \dots, p$ a las funciones coordenadas de f, entonces*

$$\lim_{x\to a} f(x) = l = (l_1, \dots, l_p) \quad \Leftrightarrow \quad \lim_{x\to a} f_i(x) = l_i.$$

Demostración. Teniendo en cuenta que

$$\|f(x)-l\|_\infty = \|(f_1(x)-l_1, \dots, f_p(x)-l_p)\|_\infty = \max_{1\le i\le p} \|f_i(x)-l_i\|,$$

es obvio que $\|f(x)-l\|_\infty \to 0$ sii $\|f_i(x)-l_i\| \to 0,\ \forall i.$ ∎

1.20 *Sea $f : A \subset E \to F$ tal que en el punto $a \in A'$, $\lim_{x\to a} f(x) = b$ y supongamos que g es una función definida en un subconjunto B de F que contiene a $f(A) \cup \{b\}$ y que es continua en b. Entonces,*

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

Demostración. Trivial.

1.21 *Sea $f : A \subset E \to F$, $a \in A'$. Si $\lim_{x\to a} f(x) = b$ y $b \neq 0$, entonces existe un entorno U de a y $\alpha > 0$ tal que $\|f(x)\| \geq \alpha$ para cada $x \in U^* \cap A$. En particular, $f(x) \neq 0$ para cada $x \in U^* \cap A$.*

Demostración. Como consecuencia de lo anterior y de la aplicación "norma" es continua en todo punto se tiene que $\lim_{x\to a}\|f(x)\| = \|\lim_{x\to a} f(x)\| = \|b\|$. Por tanto, tomando $\varepsilon = \|b\|/2$, existirá $\delta > 0$ tal que si $\|x-a\| < \delta, x \in A, x \neq a$ entonces

$$\Big|\|f(x)\| - \|b\|\Big| \leq \|b\|/2 \quad \Leftrightarrow \quad \|b\|/2 \leq \|f(x)\| \leq \frac{3}{2}\|b\|.$$

Luego tomando $U = B(a,\delta)$ se tiene que $\|f(x)\| \geq \alpha = \|b\|/2$ para todo $x \in U^* \cap A$. ∎

1.22 *Si $f : A \subset E \to F$ y $\lim_{x\to a} f(x) = l$, entonces si $B \subset A$ y $a \in B'$, $\lim_{x\to a} f_{|B}(x) = l$.*

Demostración. Trivial.

Es claro que, contrariamente, la existencia de $\lim_{x\to a} f_{|B}(x)$ no implica la del $\lim_{x\to a} f(x)$. Sin embargo en el caso particular en el que $B = B(a,r)\cap A$, entonces

$$a \in B' \;\Leftrightarrow\; a \in A' \text{ y } \lim_{x\to a} f(x) = l \;\Leftrightarrow\; \lim_{x\to a} f_{|B}(x) = l$$

1.23 Para el cálculo de límites son aplicables las siguientes fórmulas:

1) $\lim_{x\to a}(f+g)(x) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x); \quad \lim_{x\to a}(\lambda f)(x) = \lambda \lim_{x\to a} f(x)$

2) Si f, g son funciones escalares, entonces $\lim_{x\to a}(fg)(x) = \lim_{x\to a} f(x) \lim_{x\to a} g(x)$

3) Si f, g son escalares y $\lim_{x\to a} g(x) \neq 0$, entonces $\lim_{x\to a}(1/g)(x) = \dfrac{1}{\lim_{x\to a} g(x).}$

Por tanto $\lim_{x\to a}(f/g)(x) = \dfrac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x).}$

Demostración. Las demostraciones son totalmente análogas a las correspondientes para funciones de 1-variable:

1) es inmediata.

2) Si $\lim_{x\to a} f(x) = l$; $\lim_{x\to a} g(x) = r$, entonces:

$$\begin{aligned}|(fg)(x) - lr| &= |f(x)(g(x)-r) + r(f(x)-l)| \\ &\leq |f(x)||g(x)-r| + |r||f(x)-l| \\ &\leq |f(x)-l||g(x)-r| + |l||g(x)-r| + |r||f(x)-l|.\end{aligned}$$

Puesto que cada uno de los tres sumandos anteriores tienden a 0 cuando $x \to a$, se deduce que $\lim_{x\to a}(fg)(x) = lr$.

3) Puesto que $\lim_{x\to a} g(x) = r \neq 0$, de 1.21 se deduce que existe $\alpha > 0$ y $\delta_1 > 0$ tal que si $\|x - a\| < \delta_1, x \in A, x \neq a$ entonces $|g(x)| > \alpha$. Por lo tanto,

$$\Big|\frac{1}{g(x)} - \frac{1}{r}\Big| = \frac{1}{|r||g(x)|}|g(x) - r| \leq \frac{1}{|r|\alpha}|g(x) - r| \to 0.$$

Los resultados anteriores sobre límites se trasladan de forma automática a resultados análogos sobre continuidad:

Corolario 1.24

- *La función $f = (f_1, \ldots, f_p)$ es continua en un punto a si y sólo si las funciones coordenadas f_i son todas continuas en a.*
- *Si f continua en a, g continua en $b = f(a)$, entonces $g \circ f$ es continua en a.*
- *Si f, g continuas en a , entonces $f + g, \lambda f$, (y si f, g son escalares) fg y f/g(cuando $g(a) \neq 0$), son continuas en a.*

Otras consecuencias

- En espacios normados la suma y la multiplicación por escalares son aplicaciones continuas.
- Cada función polinómica definida en $\mathbb{R}^n$ es continua en todo punto.
- Si f, g son funciones continuas sobre un subconjunto A de un espacio normado y toman sus valores en $\mathbb{R}$, entonces $\{x \in A : f(x) < g(x)\}$ es abierto de A y los conjuntos $\{x \in A : f(x) \leq g(x)\}$ y $\{x \in A : f(x) = g(x)\}$ son cerrados de A.

2.2. Límites iterados y límites direccionales

A continuación establecemos algunas condiciones necesarias para la existencia de límite de funciones (escalares, si se quiere) de varias variables reales.

Definición 1.25 (Límites iterados) Sea $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ y $(x_0, y_0) \in A'$. A cada uno de los límites

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

se les denomina límites iterados.

Proposición 1.26 *Con las notaciones anteriores, si existe*

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = l,$$

y en algún entorno U de (x_0, y_0) existe el límite $\lim_{y\to y_0} f_{|U^*\cap A}(x,y)$ *entonces también existen y es igual a "l" el límite iterado* $\lim_{x\to x_0}(\lim_{y\to y_0} f(x,y))$*. Mismo enunciado para el otro límite iterado.*

Demostración. Resulta directamente de aplicar la definición de límite.

Definición 1.27 (Límites direccionales) Dada una recta $\boldsymbol{r}$ que pasa por el punto (x_0, y_0) Llamaremos límite de la función f en el punto (x_0, y_0) siguiendo esa recta al $\lim_{(x,y)\to(x_0,y_0)} f_{|B}(x,y)$, donde $B = \boldsymbol{r} \cap A$ (supuesto que $(x_0, y_0) \in B'$). Por tanto, si $\boldsymbol{r}$ es la recta de pendiente m, $y - y_0 = m(x - x_0)$, entonces el límite siguiendo esa recta será:

$$\lim_{x\to x_0} f(x, y_0 + m(x - x_0)).$$

Los límites siguiendo rectas serán los límites direccionales.(Análoga definición para límite siguiendo curvas que pasan por el punto).

Nota. La definición 1.25 se generalizan de manera natural al caso de funciones de 3 o más variables. Para generalizar también la noción de límites direccionales de una función en un punto, deberemos escribir en forma paramétrica las ecuaciones de las rectas que pasan por el punto. Así si $a = (a_1, \ldots, a_n)$, entonces

$$x_1 = a_1 + th_1,\ x_2 = a_2 + th_2,\ \ldots\ , x_n = a_n + th_n$$

es la ecuación de la recta que tiene como vector director $h = (h_1, \ldots, h_n)$ y que pasa por a. El límite siguiendo esta recta será entonces

$$\lim_{t\to 0} f(a_1 + th_1, \ldots, a_n + th_n).$$

Para $n = 2$ el límite anterior coincide con el límite direccional en el sentido de la definición 1.27, siguiendo la recta de pendiente $m = h_2/h_1$.

Como en el caso de los límites iterados, es evidente que la existencia de límite implica la de los límites direccionales y siguiendo curvas. Se deduce, pues, que la existencia de los límites iterados, direccionales y siguiendo curvas son *condiciones necesarias* para la existencia del límite. Por lo tanto:

NO existe límite cuando

1. *No existe alguno de los límites iterados o existen pero son distintos.*

Ejemplo 1.28 Consideremos la función

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

Esta función no es continua en $(0,0)$ ya que uno de los límites iterados no existe:

$$\lim_{y\to 0}(\lim_{x\to 0} \frac{x^2+y}{\sqrt{x^2+y^2}}) = \lim_{y\to 0} \frac{y}{|y|} = \pm 1.$$

Ejemplo 1.29 Este es un ejemplo de una función para la que los límites iterados existen pero son diferentes (luego el límite no existe)

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

Se tiene que

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

2. *No existe alguno de los límites direccionales o existen, pero no son iguales.*

Ejemplo 1.30 Sea

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

Los límites direccionales de esta función no son todos iguales. En efecto:

$$\lim_{x\to 0} f(x,mx) = \lim_{x\to 0} \frac{mx^2}{(m^2+1)x^2} = \frac{m}{m^2+1}$$

que depende de m. Se deduce pues que el límite no existe.

Ejemplo 1.31 Sea

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

Es inmediato comprobar que $\lim_{x\to 0} f(x, mx) = 0$ para $m \neq 1$. En cambio para $m = 1$ el límite anterior no existe, es decir la función no tiene límite en (0,0) siguiendo la recta $y = x$ y por lo tanto no admite límite en ese punto (no es continua en (0,0)).

3. *No existe el límite siguiendo alguna curva que pasa por el punto o el límite varía dependiendo de la curva que se tome.*

Ejemplo 1.32 Consideremos la función

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

Tanto los límite iterados como los límites direccionales en el punto (0,0) existen y valen 0, sin embargo esta función no tiene límite en ese punto, ya que si tomamos las curvas $y = m\sqrt{x}$, se tiene:

$$\lim_{x\to 0} f(x, m\sqrt{x}) = \lim_{x\to 0} \frac{m^2 x^2}{x^2 + m^4 x^2} = \frac{m^2}{m^4 + 1}.$$

Es decir los límites siguiendo esa familia de curvas existen todos pero son diferentes entre sí, luego el límite no existe.

Ejemplo 1.33 Sea

$$f(x,y) = \frac{x^2 \ln y}{x^4 + (x^2 + \ln y)^2}, \qquad \text{si } (x,y) \neq (0,1); \qquad f(0,1) = 0.$$

Es inmediato comprobar que también en este caso los límites iterados en (0,1) valen 0. En cuanto a los límites direccionales

$$\lim_{x\to 0} \frac{x^2 \ln(1+mx)}{x^4 + (x + \ln(1+mx))^2} = \lim_{x\to 0} \frac{\ln(1+mx)}{x^2 + (x + 1/x \ln(1+mx))^2} = \frac{\ln 1}{m^2} = 0.$$

Sin embargo tampoco existe el límite ya que si consideramos la curva $y = e^{-x^2}$, que obviamente pasa por (0,1), la función admite límite siguiendo esta curva, pero es diferente de 0.

Ejercicios. 1. Sea f la función definida en $A = \{(x, y) \in \mathbb{R}^2 : x \neq 0;\ y \neq 0;\ xy \neq 1\}$ por

$$f(x, y) = \frac{x^3 - y^3}{x^2y^2 - xy}.$$

a) Probar que f es continua es cada punto de A.

b) Estudiar la existencia de límite de f en los puntos $(a, 0),\ (0, b), (c, 1/c)$ siendo $a \neq 0$, $b \neq 0$ y $|c| \neq 1$.

c) Obtener los límites iterados de f en los puntos $(1, 1)$ y $(-1, -1)$ para estudiar si existe límite de f en esos puntos.

d) Comprobar que los límites iterados de de f en $(0, 0)$ no están definidos; que los límites direccionales en ese punto son todos iguales, pero que la función f no tiene límite en $(0, 0)$.

e) Considerar la función g definida en A por

$$g(x, y) = (x + y) \operatorname{sen} \frac{x^3 - y^3}{x^2y^2 - xy},$$

y demostrar que tiene límite en $(0, 0)$ a pesar de que los límites iterados no están definidos en $(0, 0)$.

2. Considerar las funciones definidas en $A = \mathbb{R}^2 \setminus \{(0, 0)\}$:

$$f_1(x, y) = \frac{(x + y)^2(e^y - 1)}{x^2 + y^2};\ f_2(x, y) = \frac{(x + y)^2(e^y - 1)}{x^4 + y^2}$$
$$f_3(x, y) = \frac{xy(e^y - 1)}{x^4 + y^2}.$$

a) Probar que f_1 tiene límite en $(0, 0)$.

b) Probar que los límites iterados y direccionales de f_2 en $(0, 0)$ existen, pero f_2 no tiene límite en $(0, 0)$.

c) Probar que la función f_3 tiene límite en $(0, 0)$.

3. *a*) Probar que la función de 1-variable

$$g(t) = \frac{\operatorname{sen} t}{t} \text{ si } t \neq 0;\ g(0) = 1,$$

es continua en todo punto $t \in \mathbb{R}$.

b) Utilizar el apartado anterior para probar que la función

$$f(x,y) = \frac{\operatorname{sen} xy}{y} \text{ si } y \neq 0;\ f(x,0) = 1,$$

es continua en todo punto de $\mathbb{R}^2$.

c) Utilizar de nuevo el apartado a) para obtener $\lim_{(x,y)\to(a,a)} f(x,y)$, $a \in \mathbb{R}$, siendo f la función definida en el conjunto $A = \{(x,y) \in \mathbb{R}^2 : |x| \neq |y|\}$ por

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

3. Funciones continuas sobre compactos

De nuevo nos referiremos a la topología de un espacio normado para enunciar, sin demostración, algunos resultados de las funciones continuas, por otra parte ya conocidos para las funciones de una variable. Las demostraciones de estos resultados se verán en la asignatura de Topología del segundo semestre de este curso o bien pueden encontrarse en los apuntes de esta misma asignatura del curso 2011/12 (Compacidad).

Definición 1.34 Un conjunto K de un espacio normado se dice compacto si de cada recubrimiento de K por conjuntos abiertos se puede extraer un subrecubrimiento finito. Abreviadamente si $K \subset \bigcup_{i\in I} U_i$ entonces existe un subconjunto finito $J \subset I$ tal que $K \subset \bigcup_{i\in J} U_i$.

- Puesto que la definición de compacto se hace exclusivamente en términos de los abiertos, es claro que que *si dos normas sobre el mismo espacio vectorial E son equivalente, un conjunto $K \subset E$ es compacto respecto a una de la normas si y sólo si lo es respecto a la otra*. En particular en $\mathbb{R}^n$ la compacidad de un conjunto no depende de la norma que se tome.
- Es cierto, aunque no lo demostraremos, que *todo compacto de un espacio normado es un conjunto cerrado y acotado. Ya sabemos que en $\mathbb{R}$ el recíproco también se satisface: los conjuntos compactos son justamente los que son cerrados y acotados . Esto también es cierto en* $\mathbb{R}^n$, es decir en los espacios normados de dimensión finita y, según un teorema de F. Riesz (Ver Manual, Teorema 2.9), sólo en ellos.

1.35 (Funciones continuas sobre compactos). *Sean E, F espacios normados y $f\colon A \subset E \to F$ una aplicación continua, entonces:*

i) *La imagen por f de cada compacto $K \subset A$, $f(K)$, es un conjunto compacto de F y por tanto cerrado y acotado.*

ii) *f es uniformemente continua sobre cada compacto $K \subset A$.*

iii) *Si $F = \mathbb{R}$ entonces f alcanza un valor máximo y un valor mínimo sobre cada compacto $K \subset A$.*

Ejercicios

1A Sean p, q números reales positivos tales $1/p + 1/q = 1$ (observar que en estas condiciones p y q deben ser mayores que 1).

(a) Demostrar la desigualdad:

$$xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q, \quad x, y \geq 0.$$

Indicación. Escribir $\quad xy = e^{\frac{1}{p}\ln x^p + \frac{1}{q}\ln y^q} \quad$ y tener en cuenta que la función e^x es convexa.

(b) (*Desigualdad de Hölder*) Utilizar el apartado anterior para demostrar que

$$\sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q}.$$

En otros términos, $\langle x, y\rangle \leq \|x\|_p \|y\|_q, \quad x = (x_1, \ldots, x_n);\ y = (y_1, \ldots, y_n)$.
Indicación. Suponer en una primera etapa que $\|x\|_p = 1, \|y\|_q = 1$ y demostrar que entonces $\langle x, y\rangle \leq 1$.

(c) Demostrar que $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$ es una norma sobre $\mathbb{R}^n$.

1B Sea p un número real mayor o igual que 1 y denotemos por l_p al conjunto de sucesiones de números reales (x_n) tales que $\sum_{n=1}^{\infty} |x_n|^p < \infty$. Definamos también l_∞ como el conjunto de las sucesiones acotadas de números reales.

(a) Probar que l_p y l_∞ son espacios vectoriales y que la expresiones

$$\|\mathbf{x}\|_p = \Big(\sum_{n=1}^{\infty} |x_n|^p\Big)^{1/p}; \quad \|\mathbf{x}\|_\infty = \sup_{n\in\mathbb{N}} |x_n|$$

definen sendas normas sobre l_p y l_∞

(b) Demostrar que la adherencia en l_∞ del conjunto de sucesiones que tienen todos sus términos nulos, salvo un número finito de ellos, es c_0: el espacio vectorial de sucesiones reales que convergen a 0.

1C Demostrar que si $\|\cdot\|$ es una norma sobre $\mathbb{R}^n$ tal que

$$(*) \qquad \|(u_1,\ldots,u_n)\| \leq 1 \quad \Rightarrow \quad |u_i| \leq 1,$$

entonces $|x_i| \leq \|x\|$, para cada $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$. Dar ejemplos de normas que no satisfagan la condición (*) para ningún i.

1D Estudiar si las expresiones siguientes definen una norma sobre $\mathbb{R}^2$:

1. $\|(x,y)\| = \sqrt{4x^2 + y^2}$.
2. $\|(x,y)\| = \sqrt{|x| + |y|}$.
3. $\|(x,y)\| = |x| + |\sqrt[3]{x^3 + y^3}|$.
4. $\|(x,y)\| = \sqrt{(x-y)^2 + y^2}$.

1E Sean $(E_i,\ i = 1,2,\ldots,n)$ una familia finita de espacios normados y empleemos la notación común $\|\cdot\|$ para designar a las normas de E_i.

(a) Demostrar que

$$\|(x_1,\ldots,x_n)\| = \sum_{i=1}^{n} \alpha_i \|x_i\|, \quad \alpha_i \geq 0,$$
$$\|(x_1,\ldots,x_n)\| = \sqrt{\|x_1\|^2 + \ldots + \|x_n\|^2}$$

son normas sobre $E = E_1 \times \ldots \times E_n$.

(b) Utilizar lo anterior para demostrar que

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

es una norma sobre $\mathbb{R}^3$.

1F Demostrar que la expresión

$$\|(x,y,z)\| = \sqrt{x^2 + (y-x)^2 + (z-y)^2}$$

define una norma sobre $\mathbb{R}^3$. Compararla con la norma euclídea.

1G Sea $(E, \|\cdot\|)$ un espacio normado. Estudiar si la aplicación de E en sí mismo, $f(x) = x\|x\|$, es continua, uniformemente continua o lipschitziana.

1H Encontrar una norma sobre $\mathbb{R}^2$ para la que la esfera unidad sea la elipse de ecuación $x^2 + 4y^2 = 4$.

1I Sea $\{x_n\}$ con $x_n \neq 0$ para todo n, una sucesión de Cauchy en un espacio normado.

(a) Probar que la sucesión de números reales $\{\|x_n\|\}$ es convergente. Sea α su límite.

(b) Probar que si $\alpha > 0$ entonces la sucesión $\{\frac{x_n}{\|x_n\|}\}$ es de Cauchy.

(c) Demostrar con un ejemplo que si $\alpha = 0$, la sucesión $\{\frac{x_n}{\|x_n\|}\}$ no es necesariamente de Cauchy.

1J Sea $\{x_n\}$ una sucesión convergente a 0 en un espacio normado. Probar que también converge a 0 la sucesión:

$$y_n = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

1K Sea E el espacio vectorial de las funciones polinómicas sobre el intervalo $[0,1]$. Consideremos sobre él las normas:

$$\|a_0 + a_1 x + \ldots + a_n x^n\|_\infty = \text{máx}(|a_0|, \ldots, |a_n|)$$
$$\|a_0 + a_1 x + \ldots + a_n x^n\|_1 = |a_0| + \cdots + |a_n|$$
$$\|a_0 + a_1 x + \ldots + a_n x^n\| = \max_{x \in [0,1]} |a_0 + a_1 x + \ldots + a_n x^n|.$$

Establecer las comparaciones posibles entre ellas, probando, en particular, que la primera y la tercera no son comparables.

Capítulo 2

La Diferencial de Fréchet

Iniciamos, con este capítulo, el cálculo diferencial para funciones de varias variables reales. Aunque el marco de trabajo será, con frecuencia, el de los espacios normados, nuestro interés se centra en la generalización del concepto de derivada, y el estudio de sus propiedades, a las funciones de varias variables reales.

1. Funciones diferenciables

Si f es una función real de una variable real, sabemos que f es derivable en el punto a si existe

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}. \tag{2.1}$$

Es obvio que el concepto anterior de función derivable puede extenderse, sin modificación alguna, a las funciones de una sola variable, pero que toman valores en un espacio normado cualquiera F. En particular si $f\colon A \subset \mathbb{R} \to \mathbb{R}^p$, es fácil ver que $f'(a) = (f_1'(a), f_2'(a), \ldots, f_p'(a))$. Esta fórmula es igualmente válida si f es una función de 1 variable que toma sus valores en un producto finito de espacios normados. Sin embargo, cuando f es una función varias variables (o de "variable vectorial"), no podemos definir $f'(a)$ como en (2.1) pues el "h" por el que habría que dividir no sería, en ese caso, elemento de un cuerpo. Aún sería esto posible para las funciones de variable compleja, pero éstas no son objeto de estudio en este curso. En lo sucesivo, por tanto, el término variable habrá que entenderlo como variable real, y del mismo modo un espacio normado será, siempre, un espacio normado real. No obstante, hemos de señalar que no existen diferencias esenciales entre un cálculo diferencial real y un cálculo diferencial complejo.

Antes de proceder a la extensión definitiva del concepto de derivada a las funciones de varias variables, vamos a dedicar un primer una sección la introducción de dos conceptos básicos, el de derivada parcial y el de derivada direccional. Aunque pueda parecer exagerado, se podría afirmar que el Cálculo Diferencial en dimensión finita consiste en el cálculo con derivadas parciales.

1.1. Derivadas parciales

Para funciones de una variable disponemos de unas reglas de cálculo de derivadas, las llamadas reglas de derivación. Por ejemplo sabemos que la función si $f(x) = x^2 + x \operatorname{sen} x$ satisface las hipótesis necesarias para poder aplicar estas reglas para calcular su derivada y así obtener que $f'(x) = 2x + \operatorname{sen} x + x \cos x$. En particular podemos obtener el valor de la derivada en un punto concreto a sin más que sustituir a en la igualdad anterior, de este modo $f'(\pi) = 2\pi + 0 - \pi = \pi$. Cuando estas reglas no son aplicables en algún punto a, aún la función puede admitir derivada en a. Por ejemplo si $f(x) = \sqrt{x^2 + x \operatorname{sen} x}$, $x \in [-\pi/2, \pi/2]$ y queremos calcular $f'(x)$, se puede aplicar la regla de la cadena para hacer este cálculo en todo punto x distinto de 0. De este modo para $x \neq 0$ se tiene $f'(x) = 1/2(2x + \operatorname{sen} x + x \cos x)(2x + \operatorname{sen} x + x \cos x)^{-1/2}$. En cambio no podemos aplicar la regla de la cadena para calcular $f'(0)$ pues para ello sería necesario que la raíz cuadrada (como función de una variable) fuese derivable en 0. Para estudiar la existencia de $f'(0)$ recurrimos a la definición de derivada en un punto:

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

Entonces para la función del ejemplo tendríamos

$$f'(0) = \lim_{x \to 0} \frac{\sqrt{x^2 + x \operatorname{sen} x}}{x} = \begin{cases} \sqrt{2} & \text{si } x \to 0^+ \\ -\sqrt{2} & \text{si } x \to 0^- \end{cases}$$

Una situación idéntica se plantea cuando se tiene una función f de varias variables, supongamos para concretar de 2-variables x, y, y se quiere calcular la derivada (parcial) de f respecto a una de las variables. Por ejemplo si $f(x, y) = x^2 + y \operatorname{sen} y$ la derivada parcial de f respecto de x en un punto (x, y) no es otra cosa que la derivada de la función de 1-variable que se obtiene al considerar y en la derivación como una constante. Luego $\frac{\partial f}{\partial x}(x, y) = 2x$ y análogamente $\frac{\partial f}{\partial y}(x, y) = \operatorname{sen} y + y \cos y$. En particular podemos obtener el valor de la derivadas parciales en un punto concreto

(a, b) sin más que sustituir (a, b) en las igualdades anteriores. De igual modo si $f(x, y) = \sqrt{x^2 + y \operatorname{sen} y}$, $y \in [-\pi/2, \pi/2]$, se pueden aplicar en cada $(x, y) \neq (0, 0)$ las reglas de derivación de las funciones de una variable para deducir que

$$\frac{\partial f}{\partial x}(x, y) = x(x^2 + y \operatorname{sen} y)^{-1/2}; \frac{\partial f}{\partial y}(x, y) = \frac{1}{2}(\operatorname{sen} y + y \cos y)(x^2 + y \operatorname{sen} y)^{\frac{1}{2}}.$$

Como antes, no podemos aplicar la regla de la cadena para estudiar la existencia de la derivada parcial respecto a x (o respecto a y) en el punto $(0, 0)$, sino que habremos de recurrir a la definición de derivada parcial respecto a x (o respecto a y) en un punto (a, b), que de acuerdo a lo anterior no es más que derivada en a de la función de la variable x (o y) que se obtiene al sustituir en la definición de f la variable y (o x) por b (o a) i.e.,

$$\frac{\partial f}{\partial x}(a, b) = \lim_{x \to a} \frac{f(x, b) - f(a, b)}{x - a}; \quad \frac{\partial f}{\partial y}(a, b) = \lim_{y \to b} \frac{f(a, y) - f(a, b)}{y - b}$$

Más generalmente,

Definición 2.1 Sea $f\colon A \subset \mathbb{R}^n \to F$ y sea $a \in \mathring{A}$. Llamaremos derivada parcial j-ésima de f en a a la derivada en a_j de la aplicación de 1-variable

$$F_j\colon x_j \to f(a_1, .., a_{j-1}, x_j, a_{j+1}, .., a_n)$$

La denotaremos por $(\partial f/\partial x_j)(a)$ o, también, $D_j f(a)$. Por tanto:

$$\frac{\partial f}{\partial x_j}(a) = F_j'(a_j) = \lim_{x_j \to a_j} \frac{f(a_1, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_n) - f(a)}{x_j - a_j}.$$

Teniendo en cuenta que las derivadas parciales no son más que derivadas de funciones de una variable, podemos aplicar las reglas de derivación ya conocidas para estas últimas para deducir las siguientes reglas de cálculo con derivadas parciales:

1) Si $f = (f_1, f_2, \ldots, f_p)$, $\quad \dfrac{\partial f}{\partial x_j}(a) = \left(\dfrac{\partial f_1}{\partial x_j}(a), \dfrac{\partial f_2}{\partial x_j}(a), \ldots, \dfrac{\partial f_p}{\partial x_j}(a)\right)$.

2) $\dfrac{\partial (f+g)}{\partial x_j}(a) = \dfrac{\partial f}{\partial x_j}(a) + \dfrac{\partial g}{\partial x_j}(a); \quad \dfrac{\partial (\lambda f)}{\partial x_j}(a) = \lambda \dfrac{\partial f}{\partial x_j}(a)$.

3) Si f, g son funciones escalares, $\quad \dfrac{\partial (fg)}{\partial x_j}(a) = f(a)\dfrac{\partial f}{\partial x_j}(a) + \dfrac{\partial g}{\partial x_j}(a)$;

y si además $g(a) \neq 0$, $\quad \dfrac{\partial (f/g)}{\partial x_j}(a) = \dfrac{g(a)(\partial f/\partial x_j)(a) - f(a)(\partial g/\partial x_j)(a)}{g(a)^2}$

Ejemplo (*Una función no continua en un punto, que admite en ese punto derivadas parciales respecto a cualquier índice*).

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

Es fácil ver que las dos derivadas parciales de esta función son nulas en (0,0). Por otras parte la función no es continua en (0,0) pues los límites direccionales de f en (0,0) existen pero no son todos iguales:

$$\lim_{x\to 0} \frac{mx^2}{x^2+m^2x^2} = \frac{m}{1+m^2}$$

1.2. Derivadas direccionales

Definición 2.2 Sea $f\colon A \subset E \to F, a \in \overset{\circ}{A}$ y $h \neq 0$ un vector de E. Se dirá que f es derivable en el punto a, siguiendo el vector h, si existe

$$D_h f(a) = \lim_{t\to 0} \frac{f(a+th)-f(a)}{t}. \tag{2.2}$$

Al elemento de F, $D_h f(a)$, se le denominará *derivada de f en a, siguiendo el vector h*. Cuando f admite derivada siguiendo cualquier vector no nulo, se dirá también que f admite derivadas en todos las direcciones.

Consideremos la recta de ecuación $x = a+th$, $t \in \mathbb{R}$ (recta que pasa por a y tiene a h como vector director). Entonces $f(a+th)$ son los valores que toma f sobre esta recta, y por tanto, por analogía con los límites direccionales, podría pensarse en denominar al límite 2.2, como la derivada de la función f en a siguiendo la recta $x = a+th$. Esto sería correcto, de no ser porque para cada vector director de esa recta puede resultar un valor distinto para $D_h f(a)$. Concretamente, es fácil ver que

$$D_{\lambda h} f(a) = \lambda D_h f(a).$$

2.3 Observemos que las derivadas parciales en un punto no son más que las derivadas en ese punto siguiendo ciertos vectores particulares, los vectores $e_1 = (1,0,\ldots,0)$, $e_2 = (0,1,0,\ldots,0), \ldots$. En efecto,

$$\frac{\partial f}{\partial x_j}(a) = \lim_{x_j\to a_j} \frac{f(a_1,\ldots,a_{j-1},x_j,a_{j+1},\ldots,a_n)-f(a_1,\ldots,a_n)}{x_j-a_j}$$
$$= \lim_{x_j\to a_j} \frac{f(a+(x_j-a_j)e_j)-f(a)}{x_j-a_j} = \lim_{t\to 0} \frac{f(a+te_j)-f(a)}{t} = D_{e_j} f(a).$$

2.4 La existencia de derivadas en todas las direcciones será una condición necesaria para que una función sea derivable en un punto. Pero de nuevo esta condición es muy débil. Como pasaba con las derivadas parciales, es posible que una función sea derivable en un punto en cualquier dirección y no ser continua en ese punto.

Ejemplo (*Una función no continua en un punto, que admite en ese punto derivadas en todas las direcciones*).

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

Si tomamos $v = (h,k)$ y aplicamos la definición para calcular la derivada en el punto (0,0) siguiendo el vector v, resulta

$$\text{Si } k \neq 0, \quad D_v f(0,0) = \lim_{t\to 0} \frac{f(th,tk)}{t} = \lim_{t\to 0} \frac{t^3h^2k}{(t^4h^4 + t^2k^2)t} = \frac{h^2}{k}$$

$$\text{Si } k = 0, \quad D_v f(0,0) = 0.$$

Sin embargo, esta función no es continua en (0,0), pues aunque los límites iterados y direccionales existen todos y valen 0, los límites siguiendo las curvas $y = mx^2$ son todos diferentes.

1.3. Diferenciabilidad en un punto

Hemos visto que una función de varias variables puede admitir derivadas parciales e incluso direccionales en un punto a sin ser continua en a. Esto indica que si se quiere que las funciones de varias variables que sean derivables en un punto sean continuas en ese punto, no va a bastar que se les exija sólo la existencia de derivadas parciales o direccionales.

Antes de proceder a extender a varias variables la noción de función derivable en un punto, conviene establecer un resultado elemental sobre aplicaciones lineales, que habremos de tener en cuenta en lo sucesivo.

Lema 2.5 i) *Una aplicación L de $\mathbb{R}^n$ en F es lineal si y sólo existen $c_j \in F$, $j = 1,\dots,n$ tales que $L(x_1,x_2,\dots,x_n) = c_1x_1 + c_2x_2 + \cdots + c_nx_n$. Se dice que L está determinada por la matriz $(c_1,c_2,\dots,c_n)$.*

ii) *Toda aplicación lineal L de $\mathbb{R}^n$ en F es continua, de hecho lipschitziana. (Esto no es verdad para $L : E \to F$ y E de dimensión infinita):*

Demostración. i) Obviamente cualquier aplicación de ese tipo es lineal. Recíprocamente, si L es lineal entonces

$$L(x_1, x_2, \dots, x_n) = L(\sum x_j e_j) = \sum x_j L(e_j) = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n,$$

siendo $c_j = L(e_j)$.

ii) También es inmediato comprobar que si L es lineal entonces es lipschitziana, pues $\|L(x) - L(y)\| = \|\sum c_j(x_j - y_j)\| \leq (\sum \|c_j\|)\|x_j - y_j\|_\infty$. ∎

Ejercicio. Probar que toda aplicación lineal y continua entre espacios normados es lipschitziana.

Definición 2.6 Sea $f : A \subset \mathbb{R}^n \to F,\ a \in \overset{\circ}{A}$. Se dice que f es diferenciable o derivable en a (en el sentido de Fréchet), si admite derivadas parciales respecto a cualquier índice en a y además se satisface la siguiente condición:

$$\lim_{x \to a} \frac{f(x) - f(a) - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a)(x_j - a_j)}{\|x - a\|} = 0, \tag{2.3}$$

(O equivalentemente, si

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) h_j}{\|h\|} = 0, \tag{D}$$

Notas. 1. Observemos en primer lugar que la definición anterior es independiente de la norma. Es decir, que si f es diferenciable en a respecto a las normas $\|\ \|$ de $\mathbb{R}^n$ y F y sustituimos estas normas por las normas equivalentes $\|\ \|^*$, f es también diferenciable en a respecto a las nuevas normas. sabemos que existen dos constantes mayores que 0, α, β, tales que $\|x\| \leq \alpha\|x\|^*$ para cada $x \in \mathbb{R}^n$ y $\|y\|^* \leq \beta\|y\|$ para cada $y \in F$. Entonces si

$$\frac{\|f(a+h) - f(a) - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) h_j\|}{\|h\|} \leq \varepsilon \quad \text{cuando } \|h\| \leq \delta,$$

se tiene también que

$$\frac{\|f(a+h) - f(a) - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) h_j\|^*}{\|h\|^*}$$

$$\leq \alpha\beta \frac{\|f(a+h) - f(a) - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) h_j\|}{\|h\|} \leq \alpha\beta\varepsilon \quad \text{cuando } \|h\|^* \leq 1/\alpha\delta,$$

2. Para funciones de una variable, la definición anterior coincide con la de función derivable en un punto. En efecto, supongamos que f es una función de una variable (con valores en un espacio normado F). Si f es diferenciable en a entonces por la definición debe existir $(\partial f/\partial x)(a) = f'(a)$. Recíprocamente si f es derivable en a, entonces

$$\begin{aligned}
&\lim_{x\to a} \frac{f(x)-f(a)-f'(a)(x-a)}{|x-a|} = 0 \\
&\Leftrightarrow \lim_{x\to a} \frac{f(x)-f(a)-f'(a)(x-a)}{x-a} = 0 \\
&\Leftrightarrow \lim_{x\to a} \left[\frac{f(x)-f(a)}{x-a} - f'(a)\right] = 0 \quad \Leftrightarrow \quad \lim_{x\to a} \frac{f(x)-f(a)}{x-a} = f'(a).
\end{aligned}$$

Ejemplo 2.7 La función

$$f(x,y) = \frac{x^3-y^3}{x^2+y^2}; \quad f(0,0) = 0$$

es continua y admite derivadas parciales respecto a ambas coordenadas en (0,0) pero no es diferenciable en (0,0).

Proposición 2.8 *Sea* $f : A \subset \mathbb{R}^n \to F$ *y* $a \in \overset{o}{A}$. *Son equivalentes:*
i) f *es diferenciable en* a.
ii) *Existe una aplicación lineal* $J : \mathbb{R}^n \to F$ *tal que*

$$\lim_{h\to 0} \frac{f(a+h)-f(a)-J(h)}{\|h\|} = 0. \tag{2.4}$$

Demostración. i) inplica ii). De la definición de función diferenciable se deduce que la condición f diferenciable en a implica que la aplicación lineal J cuya matriz asociada es la matriz jacobiana $(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a))$, satisface la condición ii).

ii) implica i) De hecho de la condición ii) se deduce:

- f es derivable en todas las direcciones.
- La aplicación J es única, $J(h) = D_h f(a) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) h_j$.
- f es diferenciable en a

En efecto, supongamos que la aplicación lineal $J : \mathbb{R}^n \to F$ satisface

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

entonces para cada $h \neq 0$ (fijado) se tiene:

$$\lim_{t\to 0} \frac{f(a+th)-f(a)-J(th)}{\|th\|} = 0 = \lim_{t\to 0} \frac{f(a+th)-f(a)-tJ(h)}{|t|\|h\|},$$

que claramente implica que

$$\lim_{t\to 0} \frac{f(a+th)-f(a)-tJ(h)}{t} = 0$$

y que, por tanto,

$$J(h) = \lim_{t\to 0} \frac{f(a+th)-f(a)}{t} = D_h f(a).$$

Se ha probado pues que J no puede ser otra aplicación que la definida por $J(h) = D_h f(a)$, determinada (en dimensión finita) por la matriz

$$(J(e_1), J(e_2), \ldots, J(e_n)) = (D_{e_1} f(a), \ldots, D_{e_n} f(a)) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right),$$

es decir $J(h) = D_h f(a) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) h_j$, que nos dice que la condición ii) no es otra cosa que la condición (D) de diferenciabilidad en a. ∎

Definición 2.9 La función $f : E \to F$ se dice diferenciable (o derivable) en el punto $a \in \mathring{A}$ si existe una aplicación lineal y continua $J : E \to F$ tal que

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

Teniendo en cuenta la proposición anterior si f es diferenciable en el punto a existe única aplicación lineal J con esta propiedad, la aplicación $J(h) = D_h f(a)$, que en lo sucesivo se denotará por $Df(a)$ y se llamará la diferencial de f en a. En dimensión finita, es decir cuando $E = \mathbb{R}^n$, ya sabemos que $Df(a)$ está determinada por la matriz jacobiana, es decir $Df(a)(h) \equiv Df(a)h = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a) h_j$.

Proposición 2.10 *Una función diferenciable en un punto a es continua en ese punto. De hecho, existe alguna bola centrada en $B(a,r)$ y alguna constante $M \geq 0$ tal que si $x \in B(a,r)$, entonces $\|f(x)-f(a)\| \leq M\|x-a\|$.*

Demostración. Por la diferenciabilidad de f en a existe $r > 0$ tal que si $\|x-a\| \leq r$ entonces

$$\frac{\|f(x)-f(a)-Df(a)(x-a)\|}{\|x-a\|} \leq 1$$
$$\Rightarrow \|f(x)-f(a)-Df(a)(x-a)\| \leq \|x-a\|.$$

Sea $V = B(a,r)$ y $x \in V$, entonces teniendo en cuenta que $Df(a)$ es lipschitziana, se tiene

$$\begin{aligned}\|f(x)-f(a)\| &\leq \|f(x)-f(a)-Df(a)(x-a)\| + \|Df(a)(x-a)\| \\ &\leq (1+K)\|x-a\|.\end{aligned}$$

∎

Ejercicio. Sean las funciones

$$\begin{aligned}
f_1(x,y,z) &= \frac{xyz}{x^2+y^4+z^4}; \quad f_1(0,0,0) = 0 \\
f_2(x,y,z) &= \frac{xyz}{x^2+y^2+z^2}; \quad f_2(0,0,0) = 0 \\
f_3(x,y,z) &= \frac{xyz^2}{x^2+y^4+z^4}; \quad f_3(0,0,0) = 0.
\end{aligned}$$

1. Probar que todas las funciones anteriores admiten derivadas en $(0,0,0)$ en todas las direcciones.
2. La única función que no es continua en $(0,0,0)$ es f_1.
3. Comprobar que la aplicaciones $T_2(u) = D_u f_2(0,0,0)$ no es lineal y por tanto f_2 no es diferenciable en $(0,0,0)$.
4. Estudiar cuales de las funciones f_i satisfacen la igualdad $D_u f_i(0,0,0) = \frac{\partial f_i}{\partial x}(0,0,0)h + \frac{\partial f_i}{\partial y}(0,0,0)k + \frac{\partial f_i}{\partial z}(0,0,0)p$, siendo $u = (h,k,p)$.
5. Comprobar que $T_3(u) = D_u f_3(0,0,0)$ es lineal, pero f no es diferenciable en $(0,0,0)$ es decir $\lim_{u\to 0} \frac{f_3(a+u)-f_3(a)-D_u f_3(a)}{\|u\|} \neq 0$, $a = (0,0,0)$; $u = (h,k,p)$.
6. Comprobar que $\lim_{u\to 0} \frac{f_2(a+u)-f_2(a)-D_u f_2(a)}{\|u\|} = 0$, $a = (0,0,0)$; $u = (h,k,p)$, aunque f_2 no es diferenciable en $(0,0,0)$ (Observar que $T_2(u) = D_u f_2(0,0,0)$ no es lineal).

1.4. Operaciones con funciones diferenciables

Todas las operaciones permitidas para las funciones derivables de 1-variable serán válidas para las funciones diferenciables definidas entre espacios normados:

2.11 *Si f, g son funciones diferenciables en a entonces $f+g$ es diferenciable en a*, ya que la aplicación lineal $Df(a) + Dg(a)$ satisface la condición de diferenciabilidad (D) para $f+g$. *Del mismo modo sucede con λf* (ejercicio).

2.12 *Sea $f\colon A \subset \mathbb{R}^n \to \mathbb{R}^p$, diferenciable en $a \in \overset{o}{A}$. Si $B \supset f(A)$ y $g\colon B \subset \mathbb{R}^p \to F$ es diferenciable en $b = f(a) \in \overset{o}{B}$ entonces $g \circ f$ es diferenciable en a y se tiene que*

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a) \quad (\textit{Regla de la Cadena}). \tag{2.5}$$

Demostración. De acuerdo a la proposición 2.8 sólo tenemos que probar que $g \circ f$ satisface en a la condición de diferenciabilidad respecto a la aplicación lineal y continua $J(h) = (Dg(f(a)) \circ Df(a))h$, es decir:

$$\lim_{x \to a} \frac{(g \circ f)(x) - (g \circ f)(a) - J(x-a)}{\|x-a\|} = 0.$$

En efecto, teniendo en cuenta que

$$\begin{aligned}
&\frac{(g \circ f)(x) - (g \circ f)(a) - (Dg(f(a) \circ Df(a))(x-a)}{\|x-a\|} \\
&= \frac{g(f(x)) - g(f(a)) - Dg(f(a)(Df(a)(x-a))}{\|x-a\|} \\
&= \frac{g(f(x)) - g(f(a)) - Dg(f(a)(f(x) - f(a))}{\|x-a\|} \\
&\qquad + \frac{Dg(f(a))(f(x) - f(a) - Df(a)(x-a))}{\|x-a\|},
\end{aligned}$$

bastará ver que cada uno de los sumandos en la expresión última tiende a 0 cuando x tiende a a. Consideremos la función

$$\varphi(y) = \frac{g(y) - g(b) - Dg(b)(y-b)}{\|y-b\|} \text{ si } y \neq b; \quad \varphi(b) = 0.$$

La diferenciabilidad de g en b implica que φ es continua en 0. Por otra parte, teniendo en cuenta que en algún entorno U de a existe un $M \geq 0$ tal que

$\|f(x) - f(a)\| \leq M\|x - a\|$ se sigue que si $x \in U$

$$\left\| \frac{g(f(x)) - g(f(a)) - Dg(f(a))(f(x) - f(a))}{\|x - a\|} \right\|$$
$$= \left\| \frac{\|f(x) - f(a)\|}{\|x - a\|} \varphi(f(x)) \right\| \leq M\|\varphi(f(x))\|.$$

Y puesto que $\lim_{x\to a} \varphi(f(x)) = \varphi(\lim_{x\to a} f(x)) = \varphi(f(a)) = 0$, se deduce que el primero de los sumandos tiende a 0 cuando x tienda a a.

Y en cuanto al segundo sumando:

$$\lim_{x\to a} \frac{Dg(f(a))(f(x) - f(a) - Df(a)(x - a))}{\|x - a\|}$$
$$= \lim_{x\to a} Dg(f(a))\left(\frac{f(x) - f(a) - Df(a)(x - a)}{\|x - a\|} \right)$$
$$= Dg(f(a))\left(\lim_{x\to a} \frac{f(x) - f(a) - Df(a)(x - a)}{\|x - a\|} \right) = Dg(f(a))(0) = 0.$$

∎

En particular, la igualdad $D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$ nos dice que que la matriz jacobiana de $g \circ f$ (la matriz de la aplicación lineal $D(g \circ f)(a)$) es el producto de las matrices jacobianas de g en $f(a)$ por la jacobiana de f en a.

$$\left(\frac{\partial(g \circ f)}{\partial x_1}(a), \ldots, \frac{\partial(g \circ f)}{\partial x_n}(a)\right) =$$
$$\left(\frac{\partial g}{\partial y_1}(f(a)), \quad \cdots \quad , \frac{\partial g}{\partial y_p}(f(a)) \right) \times \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\ \cdots & \cdots & \cdots \\ \frac{\partial f_p}{\partial x_1}(a) & \cdots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

Esta igualdad matricial nos permite escribir entonces lo que podemos llamar la *regla de la cadena para derivadas parciales*:

$$\frac{\partial(g \circ f)}{\partial x_j}(a) = \sum_{s=1}^{p} \frac{\partial g}{\partial y_s}(f(a)) \frac{\partial f_s}{\partial x_j}(a).$$

Nota. Observar, no obstante, que la demostración dada para 2.12 es válida aunque los espacios normados no sean de dimensión finita.

2.13 *Una función $f = (f_1, \ldots, f_p)$ es diferenciable en un punto a si y sólo si las funciones coordenadas $f_1, \ldots, f_p$ son diferenciables en a.*

Demostración. Es inmediato comprobar que si las funciones f_i son diferenciables, entonces la aplicación lineal y continua

$$J(h) = (Df_1(a)h, Df_2(a)h, \ldots, Df_p(a)h)$$

satisface la condición de diferenciabilidad en a para la función $f = (f_1, \ldots, f_p)$. Recíprocamente si f es diferenciable en a, entonces cada f_i es diferenciable en a pues $f_i = \pi_i \circ f$, donde π_i es la proyección i-ésima que es lineal y continua y por tanto diferenciable en todo punto, según el lema que damos más abajo.

2.14 *Si f, g son funciones escalares diferenciables en el punto a entonces el producto fg y el cociente f/g (si $g(a) \neq 0$) son diferenciables en a.*

Demostración. La aplicación fg se obtiene por la composición $x \to (f(x), g(x)) \to f(x)g(x)$. La segunda función no es más que la aplicación de $\mathbb{R}^2$ en $\mathbb{R}$ producto de dos números reales, que veremos en el lema de abajo que es diferenciable en todo punto. La primera es diferenciable en a pues se trata de una función con dos funciones coordenadas f, g, que por hipótesis son diferenciables en a. Para el cociente Basta ver que si $g(a) \neq 0$, entonces la aplicación $1/g$ es diferenciable en a. Escribiendo $1/g$ como la composición de las aplicaciones

$$x \to g(x) \to \frac{1}{g(x)},$$

eso se sigue de que la aplicación de $\mathbb{R}$ en $\mathbb{R}$, $t \to 1/t$, es diferenciable en $\mathbb{R} \setminus \{0\}$. ∎

Lema 2.15 i) *Si T es una aplicación lineal y continua entre los espacios normados E y F entonces T es diferenciable en cada punto de $x \in E$ y $DT(x) = T$.*

ii) *La aplicación producto de números reales es una aplicación de $\mathbb{R}^2$ en $\mathbb{R}$ diferenciable en todo punto.*

Demostración. i) Puesto que

$$\frac{T(x+h) - T(x) - T(h)}{\|h\|} = 0,$$

Es obvio que se satisface la condición de diferenciabilidad en x con $T = DT(x)$.

ii) Sea $p : \mathbb{R}^2 \to \mathbb{R};\ p(x,y) = xy$. Entonces $\frac{\partial p}{\partial x}(a,b) = b$ y $\frac{\partial p}{\partial y}(a,b) = a$, por lo que

$$\frac{p((a,b)+(h,k)) - p(a,b) - (\frac{\partial p}{\partial x}(a,b)h + \frac{\partial p}{\partial y}(a,b)k)}{\|h\|}$$
$$= \frac{(a+h)(b+k) - ab - (bh+ak)}{\|h\|} = \frac{hk}{\sqrt{h^2+k^2}}.$$

Puesto que $\lim_{(h,k)\to(0,0)} \frac{hk}{\sqrt{h^2+k^2}} = 0$, se tiene que p satisface la condición de diferenciabilidad en (a,b). ∎

Nota. No consideraremos aquí la diferenciabilidad de la inversa de una función, ya que el estudio de funciones inversas será el objeto de otro capítulo posterior.

2.16 Los resultados anteriores nos permitirán construir una extensa familia de funciones diferenciables. Empecemos, por ejemplo, con los *polinomios*: Como se sabe un polinomio de varias variables es una aplicación de la forma

$$f(x_1, x_2, \ldots, x_n) = \sum_{\text{finita}} a_{i_1\ldots i_n} x_1^{i_1} \ldots x_n^{i_n}\ , \quad i_k = 0, 1, \ldots$$

Puesto que la suma y el producto de aplicaciones diferenciables es una aplicación diferenciable, es evidente que los polinomios son diferenciables ya que las aplicaciones

$$(x_1, x_2, \ldots, x_n) \to a_{i_1\ldots i_n}$$
$$(x_1, x_2, \ldots, x_n) \to x_{i_k}$$

lo son, como fácilmente puede verificarse.

Aplicando 2.14, *las funciones racionales* (cocientes de dos polinomios) son diferenciables sobre el conjunto de puntos donde el polinomio del denominador no se anula.

Por último, mediante la composición de las funciones anteriores con funciones de una variable y diferenciables, se obtienen nuevas funciones de varias variables y diferenciables

$$\text{sen}\,(x^2+y^2), \quad \log(1+x^2+y^2), \quad \frac{1}{1+\cos^2(xyz)}, \quad \ldots$$

1.5. Interpretación geométrica de la diferenciabilidad

Es bien conocido que una función de una variable f es derivable en un punto a si y sólo si su gráfica admite una recta tangente (no vertical) en el punto $(a, f(a))$. De igual forma el hecho de que una función (escalar) $z = f(x_1, \ldots, x_n)$, de n-variables, sea diferenciable en un punto a significa en términos geométricos que existe un hiperplano de $\mathbb{R}^{n+1}$ de ecuación $z = f(a_1, \ldots, a_n) + J(x_1 - a_1, \ldots, x_n - a_n)$ (J lineal), tangente a la gráfica de f en el punto $c = (a, f(a))$. Previamente debemos formalizar la noción de tangencia, a fin de que la condición de diferenciabilidad: existe J tal que

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

pueda interpretarse en el sentido geométrico anterior.

Definición 2.17 Llamaremos curva de $\mathbb{R}^k$ ($k > 1$) a toda aplicación continua $\gamma : I \subset \mathbb{R} \to \mathbb{R}^k$, donde I es un intervalo de $\mathbb{R}$. Podemos pensar en $\gamma(t) = (x_1(t), \ldots, x_k(t))$ como en la posición de un móvil en el instante de tiempo t. Al subconjunto de $\mathbb{R}^k$ imagen de γ

$$T(\gamma) = \{(x_1(t), \ldots, x_k(t)) : t \in I\}$$

se le denomina la traza de la curva.

(Idéntica definición para curva en un espacio normado).

Si γ es derivable en un punto $t_0 \in \overset{\circ}{I}$ diremos que el vector $\gamma'(t_0) = (x_1'(t_0), \ldots, x_k'(t_0))$ es tangente a la curva en $c = \gamma(t_0)$. Es fácil ver que el vector $\gamma'(t_0)$ es, conforme a la idea intuitiva que se tiene de tangencia, un vector tangente a la traza.

Ejemplo 2.18 Si la función de 1-variable $f : I \subset \mathbb{R} \to \mathbb{R}$ es continua en el intervalo I entonces su gráfica es la traza de una curva de $\mathbb{R}^2$.

Obviamente la gráfica de f, $Gr(f) = \{(x, f(x)) : x \in I\}$, es la traza de la curva $\gamma : I \subset \mathbb{R} \to \mathbb{R}^2$, $\gamma(x) = (x, f(x))$.

Definición 2.19 Un vector w de $\mathbb{R}^k$ se dirá tangente al conjunto $M \subset \mathbb{R}^k$ en el punto $c \in M$, si existe alguna curva $\gamma : I \subset \mathbb{R} \to \mathbb{R}^k$ contenida en M que pasa por c y tiene a w como vector tangente en c i.e.,

i) $\gamma(t) \in M$ para cada t,

ii) existe $t_0 \in I$ tal que $\gamma(t_0) = c$,

iii) γ es derivable en t_0 y $\gamma'(t_0) = w$.

Geométricamente, un vector tangente a M en c es pues un vector tangente a alguna curva trazada sobre M pasando por c (ver ejercicio 2M para una definición más general de vector tangente.) Al conjunto de vectores tangentes a M en el punto $c \in M$ se le denotará por $T_c(M)$.

Proposición 2.20 *Sea $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ una aplicación diferenciable en el punto $a = (x_0, y_0) \in \overset{o}{A}$, M la gráfica de f y $c = (a, f(a)) = (x_0, y_0, f(x_0, y_0))$. Entonces el conjunto de vectores tangentes a M en c, $T_c(M)$, es el plano vectorial director del plano afín $A_c(M)$ que pasa por c y tiene de ecuación*

$$z = f(a) + Df(a)(x - x_0, y - y_0) = f(a) + \frac{\partial f}{\partial x}(a)(x - x_0) + \frac{\partial f}{\partial y}(a)(y - y_0).$$

Demostración. Es claro que el plano vectorial director de $A_c(M)$ es $L = \{(x, y, z) : z = Df(a)(x, y)$. Veamos primero que $L \subset T_c(M)$. En efecto, sea $w = (h, k, p) \in L$ es decir tal que $p = Df(a)(h, k)$ y consideremos la curva

$$\gamma(t) = (x_0 + th, y_0 + tk, f(x_0 + th, y_0 + tk)).$$

Obviamente γ es una curva contenida en M que pasa por $c = (x_0, y_0, f(x_0, y_0))$. Geométricamente podemos obtenerla mediante la intersección de M con el plano perpendicular al XY y contiene a la recta $x = x_0 + th$; $y = y_0 + tk$; $z = 0$. Observemos que γ es derivable en $t = 0$,

$$\begin{aligned}\gamma'(0) &= (h, k, \lim_{t \to 0} \frac{f(x_0 + th, y_0 + tk) - f(x_0, y_0)}{t}) \\ &= (h, k, D_{(h,k)} f(a)) = (h, k, Df(a)(h, k)) \in T_c(M).\end{aligned}$$

Veamos ahora que $T_c(M) \subset L$. Sea $w = (h, k, p) \in T_c(M)$ y, por tanto, sea $\gamma \colon I \subset \mathbb{R} \to \mathbb{R}^3 : \ \gamma(t) = (x(t), y(t), z(t))$ una curva cuya traza esté contenida en M es decir $z(t) = f(x(t), y(t))$ y tal que en algún punto $t_0 \in \overset{o}{I}$ $\gamma(t_0) = (x(t_0), y(t_0), z(t_0)) = c = (x_0, y_0, f(x_0, y_0))$ y $\gamma'(t_0) = (x'(t_0), y'(t_0), z'(t_0)) = (h, k, p)$.

Se trata pues de ver que de lo anterior se deduce que $p = Df(a)(h, k)$. En efecto, puesto que $z(t) = f(x(t), y(t))$, aplicando la regla de la cadena se tiene que $p = z'(t_0) = Df(x(t_0), y(t_0))(x'(t_0), y'(t_0)) = Df(x_0, y_0)(h, k)$. ∎

Observar que el plano vectorial $T_c(M)$ de la función f no contiene al vector $(0, 0, 1)$. Por eso, la interpretación geométrica que acabamos de dar del hecho de que f sea diferenciable en a suele expresarse diciendo que la gráfica de f admite un plano tangente no vertical en el punto $c = (a, f(a))$.

La proposición anterior se extiende a más variables y a funciones entre espacios normados con idéntica demostración. El enunciado preciso es el siguiente:

Proposición 2.21 *Si $f : A \subset E \to F$ es una aplicación diferenciable en el punto $a \in \overset{\circ}{A}$, $M \subset E \times F$ es la gráfica de f y $c = (a, f(a))$, entonces el conjunto de vectores tangentes a M en c, $T_c(M)$, es un subespacio vectorial de $E \times F$ isomorfo a E, que no contiene vectores de la forma $(0, v)$ con $v \neq 0$. $T_c(M)$ es asimismo el espacio vectorial director de la variedad afín de $E \times F$ de ecuación $z = f(a) + Df(a)(x - a)$.*

Ejercicios

2A Estudiar continuidad y existencia de derivadas parciales para las funciones

1. $f(x,y) = \begin{cases} \ln(1 + (x-y)^2) & \text{si } x - y > 1 \\ x - y + \ln 2 & \text{si } x - y \leq 1 \end{cases}$
2. $f(x,y) = \sqrt{x^4 + \operatorname{sen}^2 xy}$
3. $f(x,y) = \begin{cases} \dfrac{x^3}{x^2 + y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$
4. $f(x,y) = \begin{cases} \frac{\operatorname{sen} x - \operatorname{sen} y}{x - y} & \text{si } x \neq y \\ \cos x & \text{si } x = y \end{cases}$

2B Supongamos que f es una función escalar de dos variables, continua en el punto (0,0), y sea $g(x,y) = xf(x,y)$. Probar que g es diferenciable en (0,0).

2C Sean E, F espacios normados, $f\colon A \subset E \to F$ y $a \in \overset{\circ}{A}$.

(a) Si f es diferenciable en a, probar que

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{\|h\|} = \begin{cases} 0 & \text{si } Df(a) = 0 \\ \text{No existe} & \text{si } Df(a) \neq 0 \end{cases}$$

(b) Probar que si f admite derivadas en a siguiendo cualquier vector y

$$(*) \qquad \lim_{h \to 0} \frac{\|f(a+h) - f(a)\|}{\|h\|} = \alpha \neq 0$$

entonces $\|D_h f(a)\| = \alpha \|h\|$, para todo $h \in E$.

(c) Deducir de (b) que si $\dim E > \dim F$ y f satisface $(*)$ entonces f no puede ser diferenciable en a.
INDICACIÓN. De ser diferenciable en a, existiría algún vector no nulo en el núcleo de $Df(a)$.

(d) Si f es diferenciable en a, entonces son equivalentes:

i) Existe $\lim_{h\to 0} \dfrac{\|f(a+h)-f(a)\|}{\|h\|}$

ii) $\|Df(a)h\| = \|Df(a)\|\|h\|$ para todo $h \in E$.

(e) Considerar las funciones

$$f_1(x,y) = \sqrt{x^2+y^2}\,; \quad f_2(x,y) = \begin{cases} \sqrt{x^2+y^2} & \text{si } x \geq 0 \\ -\sqrt{x^2+y^2} & \text{si } x < 0 \end{cases}$$

Probar que ambas funciones satisfacen la condición $(*)$ en (0,0), cuando se considera la norma euclídea en $\mathbb{R}^2$. Ninguna de las dos funciones son diferenciables en (0,0), pero mientras que la función f_1 no admite derivadas direccionales, la función f_2 sí.

2D Probar que la función $f(x) = x^{3/2} \operatorname{sen} 1/x$, $\quad f(0) = 0$ es derivable en 0, pero no es lipschitziana en ningún entorno de 0.

2E Estudiar continuidad, existencia de derivadas parciales, existencia de derivadas direccionales y diferenciabilidad en el origen de coordenadas, para cada una de las funciones siguientes. (Supondremos que todas estas funciones toman el valor 0 en el origen)

1. $f(x,y) = \dfrac{x^4+x^2y^2+y^5}{x^2+y^4}$
2. $f(x,y) = \dfrac{x^3-y^3}{x^2+y^2}$
3. $f(x,y) = \dfrac{x^4+y^4}{x^2+y^2}$
4. $f(x,y) = \dfrac{\operatorname{sen}(x^3+xyz)}{x^2+y^2+z^2}$
5. $f(x,y) = \dfrac{\ln(1+xy)}{\sqrt{x^2+y^2}}$
6. $f(x,y) = \dfrac{\ln(1+xy)}{\sqrt[3]{x^2+y^2}}$
7. $f(x,y) = \dfrac{(x^2+y^3)(x^2+y^2)}{x^2+y^4}$
8. $f(x,y) = \dfrac{xy}{x^2+\sqrt{x^2+y^2}}$

2F

1. Estudiar la diferenciabilidad de la función $f(x,y) = \operatorname{sen}|x^2-y^2|$, en los puntos (0,0) y (1,1).
2. Estudiar la diferenciabilidad en (0,0) de la función $f(x,y) = |xy|^\alpha$, según los valores de $\alpha \geq 0$.

3. Estudiar continuidad, existencia de derivadas parciales, existencia de derivadas direccionales y diferenciabilidad en los puntos (0,0) y (0,1), para la función

$$f(x,y) = \begin{cases} \sqrt{x^4 + \operatorname{sen}^2(xy)} & \text{si } x \geq 0 \\ -\sqrt{x^4 + \operatorname{sen}^2(xy)} & \text{si } x < 0 \end{cases}$$

2G Demostrar que la función

$$f(x,y) = \frac{x^5 - y^3}{\sqrt{x^6 + y^4}}\,; \quad f(0,0) = 0$$

es diferenciable en (0,0).

Indicación: Probar que

$$0 \leq \sqrt{x^6 + y^4} - y^2 \leq |x|^3, \quad \forall x, y.$$

2H En este ejercicio g y φ serán en todos los casos una función escalar diferenciable en todo punto (aunque no siempre del mismo número de variables). Supuesto esto, se trata de probar que la función h construida a partir de g es también diferenciable y de calcular sus derivadas parciales en términos de las funciones g y φ y sus derivadas parciales:

1. $h(x,y,z) = g(x^2 - z, \operatorname{sen} xyz)$
2. $h(x,y,z) = g(x + y - z^2)$
3. $h(x) = g(x^3, \operatorname{sen} x, x - 1)$
4. $h(x) = g(x^2, g(x, \operatorname{sen} x))$
5. $h(x,y) = g(x^2, g(x, \operatorname{sen} y))$
6. $h(x,y,z) = xg(xy) + yg(xz) + zg(yz)$
7. $h(x) = xg(x + g(x))$
8. $h(x,y,z) = g(x,y) + g(x,z) + g(y,z)$
9. $h(x,y) = g(x + \varphi(x,y))$
10. $h(x,y) = g(x, y\varphi(x,y))$

2I Como en el ejercicio anterior, g y φ denotarán funciones escalares diferenciables en todo punto. Sean:

(a) $h(x,y,z) = g(xy, \varphi(yz))$, con

$$\varphi(0) = 0\,; \quad \varphi'(0) = 1\,; \quad Dg(0,0) \equiv (2,3)$$

Calcular las derivadas parciales de h en $(0,1,0)$.

(b) $h(x,y) = x \cdot g(x,y,y)$, con

$$g(1,0,0) = 1\,; \quad Dg(1,0,0) \equiv (1,2,-2).$$

Calcular las derivadas parciales de h en $(1,0)$.

(c) $h(x,y,z) = g(xz, g(y,z))$, con

$$g(0,1) = 0\,; \quad Dg(0,0) \equiv (1,2)\,; \quad Dg(0,1) \equiv (-3,4).$$

Calcular las derivadas parciales de h en $(0,0,1)$.

2J Probar que la función

$$h(x,y) = \int_0^{x+y} e^{-t^2+x} dt$$

es diferenciable en cada punto y calcular su diferencial.
INDICACIÓN: Tener en cuenta el teorema fundamental del cálculo integral: Si f es una función continua en el intervalo $[a,b]$, entonces la función $F(x) = \int_a^x f(t)dt$ es derivable y su derivada es $F'(x) = f(x)$.

2K Considerar la función $f(x,y) = x^2 - 2y$.

(a) Obtener la ecuación del plano tangente a la gráfica de f en el punto (0,1,-2).

(b) Calcular la derivada de f en (0,1) siguiendo el vector $v = (2,3)$.

(c) Encontrar alguna curva sobre la gráfica de la función f, que pase por el punto (0,1,-2) y tenga como vector tangente en ese punto a $(2, 3, D_{(2,3)}f(0,1))$.

2L Estudiar la existencia de derivadas direccionales en (0,0) para las funciones

$$f_1(x,y) = \sqrt[3]{x^3 + y^3}; \quad f_2(x,y) = \sqrt[3]{x^3 + y^4}$$

¿Son diferenciables en (0,0)?
INDICACIÓN: Para estudiar la diferenciabilidad de f_2 en (0,0), puede resultar útil saber que, si r es un número real > 0 y denotamos por g a la función $g(u) = \sqrt[3]{u+r} - \sqrt[3]{u}$, entonces g es no negativa y alcanza un máximo absoluto en el punto $u = -r/2$.

2M (**Una definición más general de vector tangente**) Sea M un subconjunto de un espacio normado G, $c \in M$ y $v \in G$. Se dirá que v es tangente a M en c, si existe una sucesión $\{z_n\}$ de puntos de M y una sucesión de escalares $\{\lambda_n > 0\}$ tal que:

$$z_n \to c, \quad \lambda_n(z_n - c) \to v.$$

Se denotará por $T_c(M)$ al conjunto de vectores tangentes a M en c.

(a) Probar que un vector no nulo v es tangente a M en c si y sólo si existe una sucesión $\{z_n\} \subset M,\ z_n \neq c$ tal que

$$z_n \to c, \quad \frac{z_n - c}{\|z_n - c\|} \to v.$$

(b) Sea $\gamma\colon (a,b) \subset \mathbb{R} \to G$ una curva contenida en M que pasa por el punto $c \in M$. Para concretar, sea $c = \gamma(t_0)$. Probar que si γ es derivable en t_0 entonces el vector $v = \gamma'(t_0)$ es un vector tangente a M en c.

(c) Sean E, F espacios normados y $f\colon A \subset E \to F$ una aplicación continua y derivable siguiendo un vector h en el punto $a \in \overset{\circ}{A}$. Probar que el vector $(h, D_h f(a))$ es un vector tangente en el punto $c = (a, f(a))$ a la gráfica de la función f.

(e) Sea $f\colon A \subset E \to F$ una función diferenciable en un punto a y M su gráfica:

1. Demostrar que el conjunto de los vectores tangentes a M en el punto $c = (a, f(a))$ es el subespacio vectorial isomorfo a E,

$$T_c(M) = \{(h, Df(a)h)\colon h \in E\}.$$

2. Demostrar que el vector v es tangente a M en $c = (a, f(a))$ si y sólo si v es tangente en c a alguna curva contenida en M que pasa por c.

2. El teorema del valor medio

Comenzaremos esta sección recordando dos versiones del teorema del valor medido para funciones de 1-variable y por tanto ya conocidas:

2.22 *Sea* $f\colon [a,b] \subset \mathbb{R} \to \mathbb{R}$ *una función continua en* $[a,b]$ *y derivable en* (a,b)*. Entonces:*

(i) *Existe algún punto* $c \in (a,b)$ *tal que* $f(b) - f(a) = f'(c)(b-a)$*.*

(ii) *Si además existe alguna constante* M *tal que* $|f'(x)| \leq M$ *para todo* $x \in (a,b)$*, entonces* $|f(x) - f(y)| \leq M|x-y|, \quad \forall x,y \in [a,b]$ (*Fórmula de los incrementos finitos*)

En este tema extenderemos a las varias variables la fórmula de los incrementos finitos. Esta extensión, a la que nos referiremos como teorema del valor medio estará presente en mayor o en menor medida, como sucedía para 1-variable, en gran parte de los resultados del Calculo Diferencial en varias variables. En el enunciado del teorema necesitaremos referirnos a la noción de segmento:

Definición 2.23 En un espacio vectorial E se denomina *segmento* de extremos a, b, al conjunto

$$[a,b] = \{a + t(b-a)\colon t \in [0,1]\} = \{(1-t)a + tb\colon t \in [0,1]\}.$$

Llamaremos segmento abierto de extremos a y b al conjunto $(a,b) = [a,b] \setminus \{a,b\}$. El conjunto A se dirá *convexo* si para cada par de puntos de A, el segmento que los une está totalmente contenido en A.

Proposición 2.24 *Toda bola es un conjunto convexo.*

Demostración. Sean x, y dos puntos de la bola $B(a,r)$ y sea $z = (1-t)x + ty$ un punto del segmento $[x,y]$. Entonces

$$\begin{aligned} \|z-a\| &= \|(1-t)x + ty - ((1-t)a + ta\| \\ &\leq (1-t)\|x-a\| + t\|y-a\| < (1-t)r + tr = r. \end{aligned}$$

∎

El primero de los resultados (2.22(i)) se extiende tal cual a funciones ESCALARES de varias variables. Pero no es cierto, en general, para funciones vectoriales.

Teorema 2.25 *Sea* $[a,b]$ *un segmento de* $\mathbb{R}^n$ *y* $f\colon A\subset\mathbb{R}^n\to\mathbb{R}$ *una función continua en* $[a,b]\subset \overset{o}{A}$ *y derivable en* (a,b)*. Entonces existe algún punto* $c\in(a,b)$ *tal que* $f(b)-f(a)=Df(c)(b-a)$*.*

Demostración. Consideremos la aplicación $\lambda\colon[0,1]\subset\mathbb{R}\to[a,b]\subset\mathbb{R}^n$ definida por $\lambda(t)=a+t(b-a)=(a_1+t(b_1-a_1),\dots,a_n+t(b_n-a_n))$. Esta aplicación es claramente derivable en [0,1], siendo $\lambda'(t)=b-a$. Sea $g=f\circ\lambda$, g es una función escalar de una variable que es continua en $[0,1]$ y derivable en $(0,1)$, pues f es continua en cada punto $\lambda(t)$ con $t\in[0,1]$ y derivable en $\lambda(t)$ con $t\in(0,1)$. Aplicando la regla de la cadena se tiene que

$$g'(t)=\sum\frac{\partial f}{\partial x_j}(\lambda(t))(b_j-a_j).$$

Aplicando ahora el teorema de valor medio para funciones escalares de una variable a la función g en $[0,1]$, se tiene que existe algún $\theta\in(0,1)$

$$f(b)-f(a)=g(1)-g(0)=g'(\theta)=\sum\frac{\partial f}{\partial x_j}(\lambda(\theta))(b_j-a_j)=Df(\lambda(\theta))(b-a).$$

∎

En cuanto al segundo de los resultados (2.22(ii)) se extiende, sin ninguna restricción, a funciones con valores en un espacio normado F, aunque aquí sólo lo demostraremos para $F=(\mathbb{R}^p,\|\ \|_\infty)$:

Teorema 2.26 *Sea* $[a,b]$ *un segmento de* $\mathbb{R}^n$ *y* $f\colon A\subset\mathbb{R}^n\to\mathbb{R}^p$ *una función continua en* $[a,b]\subset \overset{o}{A}$ *y derivable en* (a,b)*. Si para para algún* $M\geq 0$ *se tiene*

$$\left|\frac{\partial f_i}{\partial x_j}(x)\right|\leq M,\qquad \forall x\in(a,b);\quad i=1,\dots,p;\ j=1,\dots,n.$$

Entonces

$$\|f(b)-f(a)\|_\infty\leq M\,\|b-a\|_1. \tag{2.6}$$

Demostración. Es una consecuencia directa del teorema anterior (Teorema 2.25) aplicado a cada función coordenada de la función f. En efecto si $f=(f_1,f_2,...,f_p)$, entonces para cada $i=1,\dots,p$, existe algún punto $c_i\in(a,b)$ tal que $f_i(b)-f_i(a)=\sum\frac{\partial f_i}{\partial x_j}(c_i)(b_j-a_j)$, luego

$$\|f(b)-f(a)\|_\infty=\max_{1\leq i\leq p}|f_i(b)-f_i(a)|\leq\sum M|b_j-a_j|=M\|b-a\|_1.$$

∎

El siguiente teorema no es una extensión a varias variables de alguno ya conocido para 1-variable:

Teorema 2.27 *Sea $[a,b]$ un segmento de $\mathbb{R}^n$ y $f\colon A \subset \mathbb{R}^n \to \mathbb{R}^p$ una función continua en $[a,b] \subset \overset{o}{A}$. Supongamos que existe algún abierto U tal que $[a,b] \subset U \subset A$ en el que las derivadas parciales de f existen y están acotadas, es decir existe $M \geq 0$ tal que*

$$\left|\frac{\partial f_i}{\partial x_j}(x)\right| \leq M, \qquad \forall x \in U; \quad i = 1, \dots, p\,; \; j = 1, \dots, n.$$

Entonces $\|f(b) - f(a)\|_\infty \leq M\,\|b-a\|_1$.

Demostración. Puede suponerse que f es una función escalar, pues si el teorema fuese cierto para funciones escalares, entonces para una función vectorial, $f = (f_1, f_2, ..., f_p)$, se tendría también:

$$\|f(b) - f(a)\|_\infty = \max_{1\leq i\leq p} |f_i(b) - f_i(a)| \leq M\,\|b-a\|_1.$$

(i) Etapa 1:

Supongamos en primer lugar que el segmento $[a,b]$ está contenido en algún n-rectángulo cerrado R contenido a su vez en U. Por definición, un *n-rectángulo* (cerrado) es un producto cartesiano de n intervalos (cerrados) de $\mathbb{R}$, es decir un conjunto de la forma

$$R = [c_1, d_1] \times \ldots \times [c_n, d_n].$$

Por tanto un punto $z = (z_i)$ está en R si y sólo si $c_i \leq z_i \leq d_i$, para todo i. Entonces

$$\begin{aligned} f(a) - f(b) = & f(a_1, a_2, \dots, a_n) - f(b_1, a_2, \dots, a_n) \\ & + f(b_1, a_2, \dots, a_n) - f(b_1, b_2, a_3, \dots, a_n) \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & + f(b_1, b_2, \dots, b_{n-1}, a_n) - f(b_1, \dots, b_n). \end{aligned} \tag{2.7}$$

Evidentemente cada uno de los nuevos puntos que utilizamos en esta descomposición pertenecen a R, y en cada paso los dos puntos que aparecen sólo se diferencian en una de las coordenadas. Entonces, la existencia de derivadas parciales en cada punto de R, nos permite aplicar en cada uno de los pasos anteriores el teorema de valor medio para funciones de una variable, en efecto vayamos al primer paso y consideremos la función de la variable

x_1, $F_1 : x_1 \to f(x_1, a_2, \ldots, a_n)$. Obviamente esta función y su derivada está bien definida en $[c_1, d_1]$ ya que $F_1'(x_1) = \dfrac{\partial f}{\partial x_1}(x_1, a_2, \ldots, a_n)$ y las derivadas parciales existen en cada punto de $R \subset U$. Puesto que $a_1, b_1 \in [c_1, d_1]$ y $|F_1'(x_1)| \leq M$ si $x_1 \in [c_1, d_1]$, aplicando a F_1 el teorema del valor medio (2.22)(ii) se tiene que

$$|f(a_1, a_2, \ldots, a_n) - f(b_1, a_2, \ldots, a_n)| = |F_1(a_1) - F_1(b_1)| \leq M|a_1 - b_1|$$

Procediendo de igual modo con cada línea de (2.7) obtendríamos

$$\begin{aligned} &|f(b_1, \ldots, b_{j-1}, a_j, \ldots, a_n) - f(b_1, \ldots, b_j, a_{j+1}, \ldots, a_n)| \\ &\qquad \leq \sum M|a_j - b_j| = M\|a - b\|_1. \end{aligned}$$

Etapa 2:

Veamos ya que la desigualdad anterior se verifica en el caso general. Para ello vamos a utilizar el siguiente

Lema 2.28 *Si el segmento $[a, b]$ está contenido en el abierto U, entonces existe algún $\lambda > 0$ tal que para todo $x \in [a, b]$, el n-cubo (cuadrado en $\mathbb{R}^2$) $B_\infty[x, \lambda]$ está contenido en U.*

En consecuencia es fácil ver que si p es un número natural tal que $\frac{1}{p}\|b - a\|_\infty \leq \lambda$, los puntos del segmento $[a, b]$:

$$a = c_0,\ c_1 = c_0 + \frac{1}{p}(b - a),\ c_2 = c_1 + \frac{1}{p}(b - a), \ldots, c_p = c_{p-1} + \frac{1}{p}(b - a) = b$$

tienen la propiedad siguiente: Cada dos consecutivos pertenecen a un mismo n-cubo cerrado contenido en U. Entonces, conforme a lo probado en la Etapa 1 , se tiene:

$$\begin{aligned} \|f(a) - f(b)\|_\infty &\leq \|f(c_0) - f(c_1)\|_\infty + \|f(c_1) - f(c_2)\|_\infty + \cdots \\ &\leq M\|c_1 - c_0\|_1 + M\|c_2 - c_1\|_1 + \cdots \\ &= M(\frac{1}{p}\|a - b\|_1 + \frac{1}{p}\|a - b\|_1 + \cdots) \\ &= M\|a - b\|_1. \end{aligned}$$

∎

Demostración del Lema. Supongamos que para cada p existe un $d_p \in [a, b]$ y un $y_p \in B_\infty[d_p, \frac{1}{p}]$ tal que $y_p \in U^c$. Sea $d_p = a + t_p(b - a)$ con $t_p \in [0, 1]$. Puesto que $[0, 1]$ es compacto la sucesión $\{t_p\}$ tiene una subsucesión $\{t_{p_k}\}$

que converge a un punto s de $[0,1]$ y por lo tanto $\{d_{p_k}\}$ converge al punto $d = a + s(b-a)$ de $[a,b]$ (comprobarlo como ejercicio). También

$$\|y_{p_k} - d\|_\infty \leq \|y_{p_k} - d_{p_k}\|_\infty + \|d_{p_k} - d\|_\infty \leq \frac{1}{p_k} + \|d_{p_k} - d\|_\infty \to 0,$$

lo que nos dice que la sucesión $\{y_{p_k}\}$ de puntos de U^c converge a d. Pero esto es absurdo, pues como $d \in U$ y U es abierto cualquier sucesión que converja a d debe tener sus términos a partir de uno en adelante contenidos en U. ∎

Nota. Los teoremas 2.26 y 2.27 pueden extenderse a funciones con valores en un espacio normado de cualquier dimensión $(F, \| \|)$. En tal caso la hipótesis sobre las derivadas parciales habría habría de ser

$$\|\frac{\partial f}{\partial x_j}(x)\| \leq M,$$

para cada x bien en (a,b) o en algún abierto conteniendo a $[a,b]$, según estemos en las condiciones del Teorema 2.26 o en las del Teorema 2.27. La conclusión, la misma:

$$\|f(b) - f(a)\| \leq M\|b-a\|_1.$$

La demostración de este hecho se deduce claramente del siguiente

Lema 2.29 *Sea F un espacio normado y $f\colon [a,b] \subset \mathbb{R} \to F$ una función continua en $[a,b]$ y derivable en (a,b). Supongamos que existe alguna constante M tal que $\|f'(t)\| \leq M$ para todo $t \in (a,b)$, entonces $\|f(b) - f(a)\| \leq M(b-a)$*

Hay varias demostraciones para este lema. Una de ellas, tomada básicamente del libro de T.M. Flett [12], puede verse en Manual (Lema 8.10(2) y Teorema 8.12). Otra puede verse en el libro de H. Cartan [5]. La demostración que damos a continuación, quizá la más corta, es la que se basa en el potente resultado del Análisis Funcional conocido como teorema de Hahn-Banach:

2.30 [Hahn-Banach] *Sea $(E, \| \|)$ un espacio normado, L un subespacio vectorial de E y $\varphi : L \to \mathbb{R}$ una forma lineal continua sobre L tal que $|\varphi(u)| \leq \|u\|$ para cada $u \in L$. Entonces φ puede extenderse a una forma lineal y continua en todo E y tal que $|\varphi(x)| \leq \|x\|$ para cada $x \in E$.*

Demostración Lema 2.29 Sea $h = f(b) - f(a) \in F$, $L =< h >$ el subespacio vectorial de F generado por h y $\varphi : L \to \mathbb{R}$ definida por $\varphi(th) = t\|h\|$. Es obvio que φ es lineal y continua sobre L y para cada $u \in L$, $|\varphi(u)| = \|u\|$. Sea $\varphi : F \to \mathbb{R}$ la extensión a F que nos proporciona el teorema de Hahn-Banach y, por tanto, satisfaciendo la desigualdad $|\varphi(x)| \le \|x\|$ para todo $x \in E$. Puesto que toda aplicación lineal y continua es diferenciable en cada punto, la aplicación $g = \varphi \circ f$ es una función continua en $[a, b]$ y derivable en (a, b) y que toma sus valores en $\mathbb{R}$. Además,

$$|g'(t)| = |D\varphi(f(t))(f'(t))| = |\varphi(f'(t))| \le \|f'(t)\| \le M.$$

Por tanto, de 2.22(ii) se deduce que

$$\|f(b) - f(a)\| = |\varphi(f(b) - f(a))| = |g(b) - g(a)| \le M(b - a).$$

∎

2.1. Consecuencias

2.31 *Sea U un conjunto abierto* CONVEXO *de $\mathbb{R}^n$ y $f : U \to \mathbb{R}^p$ una función que admite derivadas parciales acotadas en U. Entonces f es lipschitziana en U.*

Demostración. Basta aplicar el apartado (i) del teorema anterior en el segmento $[x, y]$ para cada $x, y \in U$.

2.32 *Si todas las derivadas parciales de una función $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ son nulas en el abierto* CONEXO *U, entonces f es constante sobre U.*

Demostración. Por definición el abierto U se dice conexo si NO se puede descomponer como unión de otros dos abiertos no vacíos y disjuntos. Sea $a \in U$ y $B = \{x \in U : f(x) = f(a)$. Se trata de probar que $B = U$. Para ello, veamos en primer lugar que B es abierto, es decir que si $x \in B$ entonces B es entorno de x. Puesto que U es abierto, U es entorno de x y por tanto existe alguna bola centrada en x, $B(x, r(x))$ contenida en U. Para que B sea entorno de x bastará ver que esta bola está contenida en B. Puesto que las derivadas parciales son nulas (acotadas, por tanto, por $M = 0$) la función f debe ser, según 2.25, lipschitziana de constante de Lipschitz $M = 0$ en cada abierto y convexo contenido en U, en particular en la bola $B(x, r(x))$. Es decir que si $y \in B(x, r(x))$ entonces $\|f(y) - f(x)\| \le 0\|x - y\|$, luego $f(y) = f(x)$ y, como $f(x) = f(a)$ ya que $x \in B$, se tiene que $y \in B$.

Con el mismo argumento se prueba que también $U \setminus B$ es abierto, por lo que U se escribe como unión de los abiertos disjuntos B y $U \setminus B$. Como U es conexo y $B \ne \emptyset (a \in B)$, se tiene que $U \setminus B = \emptyset$ o sea que $B = U$. ∎

Proposición 2.33 *Un conjunto abierto U de $\mathbb{R}^n$ es conexo si y sólo si es conexo por arcos i.e., si para cada $x, y \in U$ existe una curva $r : [\alpha, \beta] \to \mathbb{R}^n$ de extremos $x = r(\alpha);\ y = r(\beta)$ cuya traza está contenida en U.*

Demostración. Ver Manual (Proposición 1.10).

2.34 [Condición suficiente de diferenciabilidad] *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^p$, y $a \in \overset{o}{A}$. Si f admite derivadas parciales, respecto a cualquier índice, en un entorno del punto a y éstas son aplicaciones continuas en a, entonces f es lipschitziana en alguna bola centrada en a y diferenciable en a.*

Demostración. Puesto que suponemos que las aplicaciones

$$\frac{\partial f_i}{\partial x_j} : x \to \frac{\partial f}{\partial x_j}(x)$$

son continuas en a debe existir alguna $B(a, r_{ij})$ tal que si $x \in B(a, r_{ij})$ entonces

$$|\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)| \leq 1 \ \Rightarrow\ |\frac{\partial f_i}{\partial x_j}(x)| \leq 1 + |\frac{\partial f_i}{\partial x_j}(a)| \leq M = (1 + \max_{i,j} |\frac{\partial f_i}{\partial x_i}(a)|).$$

Tomando $r = \min r_{ij}$, de 2.31 se deduce entonces que f es lispchitziana en $B(a, r)$.

Para que f sea diferenciable en a hemos de ver que

$$\lim_{x \to a} \frac{f(x) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j)}{\|x - a\|} = 0. \tag{2.8}$$

Para ello vamos a aplicar de nuevo 2.31 a la función

$$g(x) = f(x) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j).$$

Es claro que

$$\frac{\partial g_i}{\partial x_j}(x) = \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a).$$

Por tanto, aplicando el hecho de que las derivadas parciales de f son continuas en a, se tiene que para todo $\varepsilon > 0$ existe algún $\delta > 0$ tal que si $x \in V = B(a, \delta)$ entonces

$$|\frac{\partial g_i}{\partial x_j}(x)| = |\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(a)| \leq \varepsilon,\ \forall i, j.$$

Se deduce, pues, que la función g cumple en V las hipótesis de 2.31, luego es lipschitziana en V. En particular, si $x \in V$

$$\|g(x) - g(a)\|_\infty = \Big\|f(x) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j)\Big\|_\infty \leq \varepsilon\|x - a\|_1,$$

que, obviamente, significa que f satisface la condición 2.8. ∎

Nota. Obsérvese que del hecho de que la función g de la demostración anterior sea lipschitziana en V, se deduce que

$$\lim_{(x,y)\to(a,a)} \frac{f(x) - f(y) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - y_j)}{\|x - y\|} = 0.$$

Una función que satisface la condición anterior se dice que es *estrictamente diferenciable en a*. Por lo tanto, se ha demostrado que si f es una función cuyas derivadas parciales son continuas en a, entonces f es algo más de diferenciable en a, es estrictamente diferenciable en a.

Consecuencia inmediata del teorema anterior es el siguiente

Corolario 2.35 *Si todas las derivadas parciales de una función f son continuas en un abierto U de $\mathbb{R}^n$ entonces f es locamente lipschitziana y diferenciable en U.*

Definición 2.36 (Función de clase C^1) Sea A un subconjunto de $\mathbb{R}^n$. Una función f se dice de clase C^1 sobre A, lo cual lo expresaremos con la notación $f \in C^1(A)$, si f admite derivadas parciales en algún abierto que contiene a A y éstas son continuas en cada punto de A.

El siguiente corolario es consecuencia directa del teorema anterior:

Corolario 2.37 *Si f es una función de clase C^1 sobre A entonces f es diferenciable en cada punto de A.*

Ejercicios

2N (a) Probar que si $\|\cdot\|$ es una norma cualquiera sobre $\mathbb{R}^n$, entonces la aplicación $x \to \|x\|$ es una aplicación lipschitziana que no admite derivadas direccionales en 0.

(b) Sea $U = \{(x,y) \in \mathbb{R}^2 \colon x^2 + y^2 < 1\} \setminus \{0\} \times [0,1]$, y consideremos la función f definida sobre U por

$$f(x,y) = \begin{cases} y^2 & \text{si } \ x > 0 \text{ e } y \geq 0 \\ 0 & \text{en otro caso} \end{cases}$$

Probar que U es un abierto conexo (no convexo) sobre el que f es continua, admite derivadas parciales acotadas, pero no es lipschitziana.

2Ñ (a) Probar que si f es una función lipschitziana sobre un abierto U de $\mathbb{R}^n$ y admite derivadas parciales, respecto a cualquier índice, en todo punto de U, entonces sus derivadas parciales están acotadas en U.

(b) Estudiar si la función $f(x,y,z) = \operatorname{sen}(x^2 - y^2 + z^2)$ es lipschitziana o localmente lipschitziana en $\mathbb{R}^3$.

2O (a) Probar que toda aplicación lipschitziana $f \colon A \subset E \to F$, donde E y F son espacios de Banach, se extiende a una aplicación lipschitziana sobre $\overline{A}$.

(b) Sean A, B dos conjuntos no vacíos de un espacio normado, con $B \subset A$, y supongamos que cada uno de los conjuntos $B, A \setminus B$ y A es convexo. Probar entonces que una aplicación f es lipschitziana sobre A si y sólo si es lipschitziana sobre B y sobre $A \setminus B$.

(c) Estudiar si las aplicaciones

$$f(x,y) = \operatorname{sen}|x-y|, \quad g(x,y,z) = \operatorname{sen}|x^2 + y^2 - z^2|$$

son lipschizianas o localmente lipschitzianas.

2P Consideremos la función

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

Probar que f es una función continua en todo punto, que admite derivadas parciales acotadas en $\mathbb{R}^2 \setminus (0,0)$ ¿Es lipschitziana?

2Q (a) Sea U un abierto convexo de $\mathbb{R}^n$ y supongamos que $f, g \colon U \to \mathbb{R}^p$ son dos funciones tales que, en cada punto $x \in U$, $\partial f_i / \partial x_j(x) = \partial g_i / \partial x_j(x)$, cualesquiera que sean los índices i, j. Probar entonces que las funciones f y g se diferencian en una constante.

(b) Determinar las funciones $f\colon \mathbb{R}^2 \to \mathbb{R}$ que satisfacen las ecuaciones

$$\frac{\partial f}{\partial x}(x, y) = 1\,; \quad \frac{\partial f}{\partial y}(x, y) = y, \quad \forall (x, y).$$

2R Sea I un intervalo abierto de $\mathbb{R}$, U un abierto de $\mathbb{R}^n$ y $f : (t, x) \in I \times U \to f(t, x)$ una función escalar. Demostrar que si

$$\frac{\partial f}{\partial t}(t, x) = 0, \quad \forall (t, x) \in I \times U$$

entonces f no depende de t, es decir $f(t_1, x) = f(t_2, x)$ cualesquiera que sean t_1, t_2, x.

3. Derivadas parciales de orden superior

Conviene tener presente que, por convenio, una función f derivable (aunque sea parcialmente) en un punto a debe estar definida en un entorno de a. Recordemos también que la derivada parcial $\partial f/\partial x_j(a)$, coincide con la derivada en el punto a_j de la aplicación:

$$x_j \to f(a_1, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_n).$$

Sea $f : A \subset \mathbb{R}^n \to F, \quad a \in \overset{\circ}{A}$. Llamaremos derivada parcial segunda de f respecto a x_i y x_j en el punto a, a la derivada respecto x_i de la función $\partial f/\partial x_j$ en el punto a. Abreviadamente

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

Se deduce, pues, que la función f es derivable respecto a las variables x_i y x_j en el punto a, si y sólo si la aplicación

$$\frac{\partial f}{\partial x_j} : x \to \frac{\partial f}{\partial x_j}(x)$$

está definida en algún entorno de a y admite derivada parcial respecto a x_i en el punto a.

Más generalmente, si $j_1, j_2, \ldots, j_r$ son números naturales (independientes entre sí) comprendido entre 1 y n, definiremos inductivamente

$$\frac{\partial^r f}{\partial x_{j_1} \ldots \partial x_{j_r}}(a) = \frac{\partial}{\partial x_{j_1}}(\frac{\partial^{r-1} f}{\partial x_{j_2} \ldots \partial x_{j_r}})(a).$$

Cuando el resultado final de una derivación parcial sólo dependa del número de veces que se deriva respecto a cada variable, y no del orden en que se realiza tal proceso (esto no sucederá siempre), cabe utilizar una notación abreviada para designar a las derivadas parciales de orden superior. Así, en ese caso, utilizaremos la expresión

$$\frac{\partial^r f}{\partial x_1^{i_1} \partial x_2^{i_2} \ldots \partial x_n^{i_n}}(a),$$

para denotar al resultado de efectuar r derivaciones parciales: i_n respecto a x_n , i_{n-1} respecto a x_{n-1}, etc. y por último i_1 derivaciones respecto a x_1. Por tanto $i_1 + i_2 + \cdots + i_n = r$. Algunos de los i_k pueden ser iguales a 0, lo que expresará que no se realiza derivación alguna respecto a la variable x_k (en cuyo caso omitiremos en la expresión anterior el término $\partial x_k^{i_k}$).

El teorema más clásico en relación al problema de la permutabilidad de las derivadas es el conocido como teorema de Schwarz o de las derivadas parciales segundas cruzadas.

Teorema 2.38 (Schwarz) *Sea f una función escalar de dos variables. Supongamos que para cada (x, y) de alguna bola centrada en el punto (a, b), existen*

$$\frac{\partial f}{\partial x}(x,y), \quad \frac{\partial f}{\partial y}(x,y), \quad \frac{\partial^2 f}{\partial x \partial y}(x,y),$$

y que la aplicación $\dfrac{\partial^2 f}{\partial x \partial y}$ es continua en (a, b). Entonces también existe la otra derivada cruzada en (a, b), y se verifica que

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

Una útil consecuencia del teorema de Schwarz es el siguiente resultado para funciones de n-variables:

Corolario 2.39 *Si todas las derivadas parciales de orden r de la función escalar f de n-variables existen en algún entorno de un punto a y son funciones continuas en a, entonces cada derivada parcial de orden r de f en a es independiente del orden en que se efectúen las derivaciones.*

(Ver Apuntes 2010: Teoremas de Schwarz).

Observemos que la condición del corolario anterior para $r = 1$ es justamente la que denominábamos çondición suficiente de diferenciabilidad"(2.34). En la proposición siguiente vemos qué consecuencias se derivan ahora de esta condición sobre las derivadas de orden r.

Proposición 2.40 *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^p$, y $a \in \overset{o}{A}$. Si todas las derivadas parciales de orden r de la función f existen en algún entorno del punto a y éstas son aplicaciones continuas en a, entonces:*

1. *Cada derivada de orden $r-1$ de f, $\partial^{r-1} f$, es una función lipschitziana en algún entorno de a (luego continua en cada punto de ese entorno) y diferenciable en a.*

2. *f y cada derivada parcial de orden $s < r-1$ es diferenciable en algún entorno de a.*

Demostración. Consideremos una derivada de orden $r-1$, $\partial^{r-1}f$. Llamemos, por ejemplo, φ a esta aplicación. Entonces la función φ satisface la condición suficiente de diferenciabilidad, ya que $\frac{\partial \varphi}{\partial x_j}$ no es otra cosa que una derivada parcial de orden r de f (que por hipótesis es continua en a):

$$\frac{\partial \varphi}{\partial x_j} = \frac{\partial}{\partial x_j}(\partial^{r-1}f) \equiv \partial^r f.$$

Por tanto, según (2.34) $\varphi(= \partial^{r-1}f)$ es lipschitziana en alguna bola $B(a,\delta)$ y diferenciable en a.

Con el mismo argumento, aplicado ahora a cada $\partial^{r-1}f$, que es continua en cada punto $x \in B(a,\delta)$, se tiene que cada $\partial^{r-2}f$ es diferenciable en x. Y así sucesivamente, cada $\partial^{r-3}f, \ldots, \frac{\partial f}{\partial x_j}$ y, finalmente, también f, son diferenciables en cada $x \in B(a,\delta)$. ∎

Definición 2.41 (Función de clase C^r) Sea A un subconjunto de $\mathbb{R}^n$. Una función f se dice de clase C^r ($r \in \mathbb{N} \setminus \{0\}$) sobre A, lo cual lo expresaremos con la notación $f \in C^r(A)$, si f admite derivadas parciales de orden r en algún abierto que contiene a A y éstas son continuas en cada punto de A.

De acuerdo con el resultado tipo teorema de Schwarz del corolario anterior y la Proposición 2.40 es obvio que

Corolario 2.42 *Si f es una función de clase C^r sobre un conjunto A, entonces en el cálculo de la las derivadas parciales de orden r de la función f en cada punto de A o en el las derivadas de orden $s < r$ en cada punto x de algún abierto que contiene a A, no importa el orden en que se efectúen las derivaciones.*

Ejercicios

2S Comprobar si en las funciones siguientes se da la igualdad entre las derivadas parciales cruzadas en $(0,0)$. Estudiar en cada caso si se satisfacen las condiciones del teorema de Schwartz.

1. $f(x,y) = xy\dfrac{x^2-y^2}{x^2+y^2}\,; \quad f(0,0)=0.$
2. $f(x,y) = x^2y^2\cos 1/x\,; \quad f(0,y)=0.$
3. $f(x,y) = x^2y^2 \operatorname{sen}\dfrac{1}{xy^2}\,; \quad f(x,0)=f(0,y)=0.$

2T Consideremos los operadores diferenciales

$$\nabla f = \left(\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_n}\right); \qquad \Delta f = \frac{\partial^2 f}{\partial x_1^2}+\cdots+\frac{\partial^2 f}{\partial x_n^2}$$

$$\mathrm{H}\, f = x_1\frac{\partial f}{\partial x_1}+\cdots+x_n\frac{\partial f}{\partial x_n}\,; \qquad \operatorname{div}(F_1,\ldots,F_n) = \frac{\partial F_1}{\partial x_1}+\cdots+\frac{\partial F_n}{\partial x_n}$$

$$\operatorname{rot}(F_1,F_2,F_3) = \nabla\times F = \left(\frac{\partial F_3}{\partial y}-\frac{\partial F_2}{\partial z}\,,\,\frac{\partial F_1}{\partial z}-\frac{\partial F_3}{\partial x}\,,\,\frac{\partial F_2}{\partial x}-\frac{\partial F_1}{\partial y}\right)$$

A los operadores anteriores se les conoce, en el orden en que han sido definido, como operador: Gradiente, Laplaciano, Hamiltoniano, Divergencia y Rotacional.

Supuesto que se pueden permutar las derivaciones, demostrar que

1. $\Delta\, f = \operatorname{div}\nabla f$
2. $\operatorname{div}(\operatorname{rot} F) = 0.$
3. $\mathrm{H}\,\Delta - \Delta\,\mathrm{H} = -2\,\Delta$
4. $\Delta\, f = 0 \;\Rightarrow\; \Delta\,\Delta((x_1^2+\cdots+x_n^2)f) = 0.$

Capítulo 3

Teoremas de Taylor

Una vez más nos disponemos a extender a las funciones de varias variables resultados ya conocidos para funciones de una variable, los teoremas de aproximación de Taylor. Por buena lógica deberíamos empezar con la extensión a las varias variables del concepto de función r-veces diferenciable en un punto. Sin embargo, para evitar complicaciones formales (y de paso ganar tiempo) no haremos esta extensión. Sólo nos referiremos al polinomio de Taylor de orden r de una función f en un punto a, cuando f sea de clase C^r en el punto a, es decir cuando f admite derivadas parciales de orden r en algún entorno de a y éstas son continuas en a.

Esto es pedirle a f algo más que ser r-veces diferenciable en el punto a. Ya lo sabíamos para $r = 1$, esa condición implica que f es diferenciable en a y lipschitziana en algún entorno de a. En consecuencia, de la Proposición 2.40 y el Corolario 2.39, se deduce que para $r > 1$, si una función es de clase C^r en a entonces:

1. Cada derivada de orden $r-1$ de f, $\partial^{r-1} f$, es una función lipschitziana en algún entorno de a (luego continua en cada punto de ese entorno i.e. de clase C^{r-1} en ese entorno) y diferenciable en a.

2. f y cada derivada parcial de orden $s < r-1$ es diferenciable en algún entorno de a.

3. En el cálculo de las derivadas parciales de orden r de f en a y de las derivadas parciales de orden $\leq r-1$ de f en cada x de algún entorno de a, se pueden permutar las derivaciones.

Ejercicio. Probar por inducción:

1. Si $f, g : A \subset \mathbb{R}^n \to \mathbb{R}$ son de clase C^r en un punto $a \in \overset{\circ}{A}$, entonces su producto $f.g$ también es de clase C^r en a.
2. Si $f : A \subset \mathbb{R}^n \to \mathbb{R}^p$ y $g : B \subset \mathbb{R}^p \to \mathbb{R}$ son funciones de clase C^r en algún entorno los puntos $a \in \overset{\circ}{A}$ y $f(a) \in \overset{\circ}{B}$, entonces $g \circ f$ es de clase C^r en algún entorno de a.

Definición 3.1 Sea $f : A \subset \mathbb{R}^n \to F$ una función que admite derivadas parciales de orden r en algún entorno de un punto $a \in \overset{\circ}{A}$. Llamaremos Polinomio de Taylor de la función f en el punto a al polinomio cuyo término independiente es $f(a)$ y los términos de grado $1 \leq k \leq r$ son los siguientes:

$$\frac{1}{k!} \sum_{1 \leq i_1, \ldots, i_k \leq n} \frac{\partial^k f(a)}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}} x_{i_1} x_{i_2} \ldots x_{i_k}$$

Es habitual emplear la notación

$$D^k f(a) x^k = \sum_{1 \leq i_1, \ldots, i_k \leq n} \frac{\partial^k f(a)}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}} x_{i_1} x_{i_2} \ldots x_{i_k}$$

y referirse a $D^k f(a)$ como la diferencial de orden k de f en a. De hecho, la igualdad anterior, caso de ser f k-veces diferenciable, establece la forma de calcular el valor que toma la aplicación $D^k f(a)$ sobre el vector $(x, x, \ldots, x)$. Puesto que en este curso no cabe hablar de diferenciales de orden superior, la igualdad anterior nos permite dar una expresión del polinomio de Taylor del mismo tipo que para las funciones de una variable:

$$P_r f(a) x = f(a) + D f(a) x + \frac{1}{2!} D^2 f(a) x^2 + \cdots + \frac{1}{2!} D^r f(a) x^r.$$

Proposición 3.2 *Si la aplicación f es de clase C^r en a, entonces*

$$\frac{1}{k!} D^k f(a) x^k = \sum_{k_1 + \ldots + k_n = k} \frac{1}{k_1! k_2! \ldots k_n!} \frac{\partial^k f(a)}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$$

Demostración. Consideremos uno de los sumandos de $D^k f(a) x^k$:

$$\frac{\partial^k f(a)}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}} x_{i_1} x_{i_2} \ldots x_{i_k}.$$

Nos preguntamos si puede haber otros sumandos iguales a éste. Es obvio que, todos los sumandos en que intervengan las mismas coordenadas y el mismo número de veces, son iguales, pues sólo diferenciarán en el orden en que se efectúan las derivaciones y esto, de acuerdo a la hipótesis con que trabajamos, no importa. Si $k_1, k_2, k_3, \ldots, k_n$ es el número de derivaciones respecto $x_1, x_2, x_3, \ldots, x_n$ (donde algunos de los naturales k_i puede ser igual a 0, ya que $k = k_1 + k_2 + \cdots + k_n$), es claro entonces que el número de sumandos iguales al fijado es exactamente igual al de las permutaciones de k elementos en los que se repiten $k_1, k_2, \ldots, k_n$, es decir igual a $\frac{k!}{k_1! \cdots k_n!}$, y el valor común de ellos:

$$\frac{\partial^k f(a)}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n},$$

Observar que cuando algún k_i es 0 (para concretar por ejemplo si $k_1 = 0$) se tiene que

$$\frac{\partial^k f(a)}{\partial x_1^0 \partial x_2^{k_2} \ldots \partial x_n^{k_n}} x_1^0 x_2^{k_2} \ldots x_n^{k_n} = \frac{\partial^k f(a)}{\partial x_2^{k_2} \ldots \partial x_n^{k_n}} x_2^{k_2} \ldots x_n^{k_n}.$$

Por tanto,

$$\begin{aligned} \frac{1}{k!} D^k f(a) x^k &= \frac{1}{k!} \sum_{k_1 + \ldots + k_n = k} \frac{k!}{k_1! k_2! \ldots k_n!} \frac{\partial^k f(a)}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n} \\ &= \sum_{k_1 + k_2 + \ldots + k_n = k} \frac{1}{k_1! k_2! \ldots k_n!} \frac{\partial^k f(a)}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}. \end{aligned}$$

∎

1. Teoremas de Taylor

Extenderemos a continuación a las varias variables el teorema (global) de Taylor con resto de Lagrange y el conocido como teorema local de Taylor. Como ya dijimos las hipótesis que usaremos serán algo más fuerte que las que realmente se necesitan.

Recordando:

Teorema 3.3 *i)(Global) Si $f : [a, b] \subset \mathbb{R} \to \mathbb{R}$ es una función $(r + 1)$-veces derivable en el intervalo $[a, b]$, entonces existe algún punto $\xi \in (a, b)$ tal que*

$$f(b) - P_r f(a)(b - a) = \frac{1}{(r + 1)!} f^{(r+1)}(\xi)(b - a)^{r+1}.$$

ii) Si $f : [a-\delta, a+\delta] \subset \mathbb{R} \to \mathbb{R}$ es r-veces derivable en a entonces

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

Veremos en primer lugar la extensión a varias variables del teorema global.

1.1. Teorema global de Taylor

Teorema 3.4 *Sea $f\colon A \subset \mathbb{R}^n \to \mathbb{R}$, una función de clase C^{r+1} en el segmento $[a, a+h] \subset \overset{\circ}{A}$, entonces existe un punto $\xi \in (a, a+h)$ tal que*

$$f(a+h) - P_r f(a)h = \frac{1}{(r+1)!} D^{r+1} f(\xi) h^{r+1}$$

Demostración. Puesto que $f \in C^{r+1}[a, a+h]$, sabemos que f y cada $\partial^s f$ con $s \leq r$ es diferenciable en cada punto del segmento $[a, a+h]$. Vamos a ver que la función de 1-variable $g(t) = f(a+th)$ es $(r+1)$-veces derivable en $[0,1]$ y, por tanto, le podemos aplicar el teorema global de Taylor en $[0,1]$. En efecto, g es la composición de las aplicaciones $t \to a+th \to f(a+th)$, ambas diferenciables luego

$$g'(t) = \sum_j \frac{\partial f}{\partial x_j}(a+th)h_j.$$

De acuerdo con esto, también podremos derivar g' ya que las derivadas parciales $\frac{\partial f}{\partial x_j}$ también son derivables en cada punto $a+th$. Luego,

$$g''(t) = \sum_j \Big(\sum_i \frac{\partial^2 f(a+th)}{\partial x_i \partial x_j} h_i\Big) h_j = \sum_{i,j} \frac{\partial^2 f(a+th)}{\partial x_i \partial x_j} h_i h_j.$$

Procediendo de esta forma, es claro que podemos derivar g hasta el orden $r+1$ obteniendo,

$$g^{(r+1)}(t) = \sum_{1 \leq i_1, \ldots, i_{r+1} \leq n} \frac{\partial^{r+1} f(a+th)}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_{r+1}}} h_{i_1} h_{i_2} \ldots h_{i_{r+1}}.$$

Aplicando entonces el teorema global de Taylor a g en $[0,1]$, y teniendo en cuenta que $g(1) = f(a+h); g(0) = f(a)$ se tiene ya el resultado que buscábamos. ∎

Nota. El teorema global de Taylor anterior (sólo válido en esta forma para funciones escalares), es para el caso $r = 1$ el teorema del valor medio para funciones escalares: *Si $f : A \subset \mathbb{R}^n \to \mathbb{R}$ es una función continua en el segmento $[a, b] \subset A$ y derivable en (a, b), entonces existe un punto $\xi \in (a, b)$ tal que*

$$f(b) - f(a) = Df(\xi)(b - a) = \sum \frac{\partial f}{\partial x_j}(\xi)(b_j - a_j).$$

1.2. Teorema local de Taylor

En la extensión del teorema local de Taylor a varias variables usaremos el siguiente lema:

Lema 3.5 *Si f es una función de clase C^r en el punto a entonces*

$$\frac{\partial}{\partial x_j}\Big(P_r f(a)(x-a)\Big)(x) = P_{r-1}\Big(\frac{\partial f}{\partial x_j}\Big)(a)(x-a)$$

Demostración. Sea φ_k la parte homogénea de grado k del polinomio de Taylor, es decir

$$\varphi_k(x) = \sum_{i_1+i_2+\ldots+i_n=k} \frac{1}{i_1!i_2!\ldots i_n!} \frac{\partial^k f(a)}{\partial x_1^{i_1}\partial x_2^{i_2}\ldots\partial x_n^{i_n}}(x_1-a_1)^{i_1}(x_2-a_2)^{i_2}\ldots(x_n-a_n)^{i_n}.$$

Entonces

$$\frac{\partial \varphi_k}{\partial x_j}(x) = \frac{\partial}{\partial x_j}\Bigg(\sum_{\substack{i_1+i_2+\ldots+i_n=k\\ i_j\geq 1}} \frac{1}{i_1!i_2!\ldots i_n!} \frac{\partial^k f(a)}{\partial x_1^{i_1}\partial x_2^{i_2}\ldots\partial x_n^{i_n}}(x_1-a_1)^{i_1}(x_2-a_2)^{i_2}\ldots(x_n-a_n)^{i_n}\Bigg)$$

$$= \sum_{\substack{i_1+i_2+\ldots+i_n=k\\ i_j\geq 1}} \frac{i_j}{i_1!i_2!\ldots i_n!} \frac{\partial^k f(a)}{\partial x_1^{i_1}\partial x_2^{i_2}\ldots\partial x_n^{i_n}}(x_1-a_1)^{i_1}\ldots(x_j-a_j)^{i_j-1}\ldots(x_n-a_n)^{i_n}$$

$$= \sum_{\substack{i_1+\ldots+(i_j-1)\\ +\ldots+i_n=k-1\\ i_j\geq 1}} \frac{1}{i_1!\ldots(i_j-1)!\ldots i_n!} \frac{\partial^k f(a)}{\partial x_1^{i_1}\partial x_2^{i_2}\ldots\partial x_n^{i_n}}(x_1-a_1)^{i_1}..(x_j-a_j)^{i_j-1}..(x_n-a_n)^{i_n}$$

$$= \sum_{i_1+i_2+\ldots+i_n=k-1} \frac{1}{i_1!i_2!\ldots i_n!} \frac{\partial^{k-1}(\partial f/\partial x_j)(a)}{\partial x_1^{i_1}\partial x_2^{i_2}\ldots\partial x_n^{i_n}}(x_1-a_1)^{i_1}\ldots(x_j-a_j)^{i_j}\ldots(x_n-a_n)^{i_n}.$$

(observar que la validez de la última de las igualdades anteriores se basa en que se pueden permutar las derivaciones). Es evidente que de lo anterior se sigue ya lo que queríamos, es decir que

$$\frac{\partial}{\partial x_j}\Big(P_r f(a)(x-a)\Big)(x) = P_{r-1}\Big(\frac{\partial f}{\partial x_j}\Big)(a)(x-a).$$

Teorema 3.6 (Local de Taylor) *Si la aplicación $f\colon A\subset \mathbb{R}^n \to \mathbb{R}^p$ es de clase C^r en un punto $a \in \overset{o}{A}$, entonces*

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

Demostración. Obviamente se puede suponer que f es una función escalar, es decir $p = 1$. Razonaremos por inducción sobre r. Para $r = 1$ la hipótesis implica que f es diferenciable en a, luego

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

Como $P_1 f(a)(x-a) = f(a) + Df(a)(x-a)$, de lo anterior se deduce que el teorema es cierto para $r = 1$.

Supongamos pues, que para $r > 1$, f es de clase C^r en a (por tanto f es diferenciable en algún entorno de a) y tomemos como hipótesis de inducción que el teorema es cierto para $r-1$. En particular, estamos suponiendo que el teorema es cierto para las funciones $\frac{\partial f}{\partial x_j}$, ya que cada $\partial^{r-1}(\frac{\partial f}{\partial x_j})$ es una derivada parcial de orden r y por tanto continua en a. Aplicando el lema anterior, se deduce que:

$$\begin{aligned} 0 &= \lim_{x\to a} \frac{\frac{\partial f}{\partial x_j}(x) - P_{r-1}(\partial f/\partial x_j)(a)(x-a)}{\|x-a\|_1^{r-1}} \\ &= \lim_{x\to a} \frac{\frac{\partial}{\partial x_j}(f(x) - P_r f(a)(x-a))}{\|x-a\|_1^{r-1}}. \end{aligned}$$

Si $g(x) = f(x) - P_r f(a)(x-a)$ que, igual que f, es diferenciable an algún entorno de a, lo anterior nos dice que para cada $\varepsilon > 0$ existe $B(a,\delta)$ tal que si $x \in B(a,\delta)$ entonces g diferenciable en x y

$$\left|\frac{\partial g}{\partial x_j}(x)\right| \le \varepsilon \|x-a\|_1^{r-1}, \ \forall j.$$

Fijado x con $\|x-a\|_1 < \delta$ se puede aplicar el teorema del valor medio (Teorema 2.26) a la función g en el segmento $[a,x]$. En efecto si $z \in [a,x]$ entonces $|\frac{\partial g}{\partial x_j}(z)| \leq \varepsilon \|z-a\|_1^{r-1} \leq \varepsilon \|x-a\|_1^{r-1}$. Es decir g tiene sus derivadas parciales acotadas en $[a,x]$ y además es diferenciable en $[a,x]$, luego

$$|f(x) - P_r f(a)(x-a)| = |g(x) - g(a)| \leq \varepsilon \|x-a\|_1^{r-1} \|x-a\|_1 = \varepsilon \|x-a\|_1^r,$$

es decir

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

∎

A $P_r f(a)(x-a)$, es decir al polinomio de Taylor de orden r en a expresado en potencias de $(x_i - a_i)$ se le suele denominar “el desarrollo de Taylor de orden de f en a”. La conclusión del teorema anterior suele expresarse diciendo que el desarrollo de Taylor de f en a aproxima a f hasta el orden r, o también que f y su desarrollo de Taylor son tangentes de orden r o tienen un punto de tangencia de orden r en a.

1.3. Unicidad del polinomio de Taylor

Lema 3.7 (Unicidad) *Sea $f\colon A \subset \mathbb{R}^n \to F$ y $a \in \overset{o}{A}$, entonces para cada número natural $r \geq 1$ existe a lo sumo un polinomio Q de grado $\leq r$ tal que*

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

Demostración. Supongamos que existen dos polinomios de grado $\leq r$ tales

$$\lim_{x \to 0} \frac{f(h) - Q_1(h)}{\|h\|^r} = 0$$
$$\lim_{h \to 0} \frac{f(h) - Q_2(h)}{\|h\|^r} = 0$$

y probemos que ambos polinomios coinciden.

Si llamamos $S(h) = Q_1(h) - Q_2(h)$, nuestra hipótesis implica que

$$(*) \qquad \lim_{h \to 0} \frac{S(h)}{\|h\|^r} = 0.$$

Se trata de ver, por tanto, que el único polinomio de grado $\leq r$ que tiene la propiedad anterior es el idénticamente nulo, o sea que $S(h) = 0$ para todo $h \in \mathbb{R}^n$.

Escribamos $S = S_0 + S_1 + \cdots + S_r$, donde S_k es la parte homogénea del polinomio S formada por los términos de grado k, es decir

$$S_k(h_1, \ldots, h_n) = \sum_{i_1+i_2+\ldots+i_n=k} c_{i_1 i_2 .. i_n} h_1^{i_1} h_2^{i_2} \ldots h_n^{i_n}.$$

Fijemos $h = (h_1, \ldots, h_n) \neq (0, \ldots, 0)$ y veamos que $S(h) = 0$. De la expresión dada anteriormente para S_k, se deduce que si $t \in \mathbb{R}$ entonces

$$S_k(th) = S_k(th_1, \ldots, th_n) = t^k S_k(h_1, \ldots, h_n) = t^k S_k(h).$$

De la condición $(*)$ se sigue entonces que

$$0 = \lim_{t \to 0^+} \frac{S(th)}{t^r \|h\|^r} = \lim_{t \to 0^+} \frac{S_0 + tS_1(h) + \cdots + t^r S_r(h)}{t^r \|h\|^r} = \frac{S_0}{0} \;\Rightarrow\; S_0 = 0.$$

En consecuencia

$$0 = \lim_{t \to 0^+} \frac{tS_1(h) + \cdots + t^r S_r(h)}{t^r \|h\|^r} = \lim_{t \to 0^+} \frac{S_1(h) + \cdots + t^{r-1} S_r(h)}{t^{r-1} \|h\|^r}$$
$$= \frac{S_1(h)}{0} \;\Rightarrow\; S_1(h) = 0.$$

De este modo se deduce pues $S_0 = 0;\; S_1(h) = 0; \ldots S_{r-1}(h) = 0$ y que

$$0 = \lim_{t \to 0^+} \frac{tS_r(h)}{t\|h\|^r} = \frac{S_r(h)}{\|h\|^r}$$

que obviamente implica que también $S_r(h) = 0$. ▮

Como consecuencia del teorema local de Taylor y el lema de unicidad se obtiene el siguiente corolario:

Corolario 3.8 *Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}^p$ una función de clase C^r en un punto $a \in \overset{\circ}{A}$ y supongamos que Q es un polinomio tal que*

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

Entonces $P_r f(a)(x) = [Q(x)]_{\leq r}$, donde $[Q(x)]_{\leq r}$ es el polinomio formado con los términos de grado $\leq r$ del polinomio Q.

Demostración. Si $R(h) = Q(h) - [Q(h)]_{\leq r}$, entonces es obvio que todos los términos de R son de grado $> r$. Puesto que podemos elegir la norma que queramos, podemos suponer que para cada $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ se tiene que $|h_i| \leq \|h\|$. Entonces si denotamos por M al máximo de las normas de los coeficientes de R y por s al número de términos de R, es claro que en algún entorno de 0, $\|R(h)\| \leq sM\|h\|^{r+1}$. Luego

$$\left\| \frac{R(h)}{\|h\|^r} \right\| \leq sM\|h\| \to 0 \quad \text{cuando } h \to 0.$$

Se deduce pues que

$$\lim_{h\to 0} \frac{f(a+h) - [Q(h)]_{\leq r}}{\|h\|^r} = \lim_{h\to 0} \frac{(f(a+h) - Q(h)) + R(h)}{\|h\|^r} = 0.$$

Puesto que único polinomio de grado $\leq r$ que aproxima a f hasta el orden r en un entorno de a es $P_r f(a)x$, de lo anterior se deduce que $P_r f(a)x = [Q(x)]_{\leq r}$.

Ejemplo 3.9 (Una función que se aproxima hasta el orden r en un entorno de a por un polinomio y no es de clase C^r en a)

Sea la función

$$f(x) = x^3 \operatorname{sen} \frac{1}{x}; \quad f(0) = 0.$$

El polinomio 0 aproxima hasta el orden 2 a la función f en un entorno de 0, pues

$$\lim_{x\to 0} \frac{f(x)}{x^2} = \lim_{x\to 0} x \operatorname{sen} \frac{1}{x} = 0.$$

En cambio, f no es de clase C^2 en 0, ya que ni siquiera es 2-veces derivable en 0.

Por contra para $r = 1$ se tiene:

Ejercicio. Probar que una función f continua en un punto a es diferenciable en a si y sólo si existe un polinomio de grado ≤ 1 tal que

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

.

2. ANEXO: El álgebra de los desarrollos de Taylor

Sea $f : A \subset \mathbb{R}^n \to F$ una función continua en $a \in \overset{\circ}{A}$ y Q un polinomio. Emplearemos la notación $f \overset{r}{\sim} Q$ en a para expresar que

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

3.10 *Sean $f, g : A \subset \mathbb{R}^n \to F$ continuas en $a \in \overset{\circ}{A}$, tales que $f \overset{r}{\sim} Q$ y $g \overset{r}{\sim} P$ en a, entonces:*

1. $\lambda f + \mu g \overset{r}{\sim} \lambda Q + \mu P$ *en a, $\forall \lambda, \mu \in \mathbb{R}$.*
2. *(Para $F = \mathbb{R}$), $fg \overset{r}{\sim} QP$ en a.*

Demostración. 1. La demostración se deduce trivialmente de la definición.

2. Escribiendo $f(x)g(x) - Q(x-a)P(x-a) = (f(x) - Q(x-a))g(x) + (g(x) - P(x-a))Q(x-a)$, de la hipótesis se deduce que

$$\lim_{x \to a} \frac{f(x)g(x) - Q(x-a)P(x-a)}{\|x-a\|^r} = 0 \cdot g(a) + 0 \cdot f(a) = 0,$$

∎

3.11 *Sean $f : A \subset \mathbb{R}^n \to \mathbb{R}^p$ y $g : B \subset \mathbb{R}^p \to \mathbb{R}$ funciones continuas en $a \in \overset{\circ}{A}$ y $f(a) \in \overset{\circ}{B}$, respectivamente. Si Q y P son polinomios tales que $f \overset{r}{\sim} Q$ en a y $g \overset{r}{\sim} P$ en $f(a)$, entonces $g \circ f \overset{r}{\sim} P \circ (Q - f(a))$ en a.*

Demostración. Sea $\varepsilon > 0$. Como f es diferenciable en a, existe $\delta_1 > 0$ y $\alpha > 0$ tal que $\|f(x) - f(a)\| \leq \alpha \|x - a\|$. Por hipótesis existe $\eta > 0$ tal que si $\|y - f(a)\| < \eta$ entonces

$$\|g(y) - P(y - f(a))\| \leq \frac{\varepsilon}{\alpha^r} \|y - f(a)\|^r.$$

Entonces tomando $\|x-a\| \leq \delta_2 = \text{mín}(\delta_1, \frac{\eta}{\alpha})$ (lo que implica $\|f(x)-f(a)\| \leq \alpha\|x-a\| \leq \eta$), se tiene que

$$\|g(f(x)) - P(f(x) - f(a))\| \leq \frac{\varepsilon}{\alpha^r} \|f(x) - f(a)\|^r \leq \frac{\varepsilon}{\alpha^r} \alpha^r \|x-a\|^r = \varepsilon \|x-a\|^r.$$

Se deduce pues que cuando $\|x - a\| < \delta_2$,

$$\begin{aligned} \|g(f(x)) - P(Q(x-a) - f(a))\| &\leq \|g(f(x)) - P(f(x)) - f(a))\| \\ &+ \|P(f(x) - f(a)) - P(Q(x-a) - f(a))\| \\ &\leq \varepsilon \|x-a\|^r + \|P(f(x) - f(a)) - P(Q(x-a) - f(a))\|. \end{aligned}$$

Por último, sea $\beta > 0$ mayor que $\|f(x) - f(a)\|$ y $\|Q(x-a) - f(a)\|$ cuando $\|x-a\| \leq \delta_2$ (observemos que cuando $\|x-a\| \leq \delta_2$, entonces $\|f(x)-f(a)\| \leq \alpha\delta_2$ y $\|Q(x-a) - f(a)\|$ está acotado, pues Q es continua en el compacto $B[0, \delta_2]$). Puesto que P es de clase C^1, sus derivadas parciales están acotadas en $B[0, \beta]$, luego P es lipschitziana en $B(0, \beta)$, y por tanto existe $M > 0$ tal si $\|x - a\| \leq \delta_2$ entonces $\|P(f(x) - f(a)) - P(Q(x-a) - f(a))\| \leq M\|f(x) - Q(x-a)\|$. Como $f \overset{r}{\sim} Q$ podemos encontrar $\delta_3 > 0$ tal que si $\|x - a\| < \delta_3$ entonces $\|f(x) - Q(x-a)\| \leq \varepsilon/M\|x-a\|^r$, y en consecuencia

$$\|P(f(x) - f(a)) - P(Q(x-a) - f(a))\| \leq \varepsilon\|x-a\|^r.$$

∎

Teorema 3.12 *(i) Sean $f, g : A \subset \mathbb{R}^n \to \mathbb{R}^p$ de clase C^r en un punto $a \in \overset{o}{A}$. Entonces*

1. *$\lambda f + \mu g$ es de clase C^r en a y*

$$P_r(\lambda f + \mu g)(a)(x-a) = \lambda P_r f(a)(x-a) + \mu P_r g(a)(x-a)$$

2. *(Para p=1), la función $f.g$ es de clase C^r en a y*

$$P_r(fg)(a)(x-a) = [P_r f(a)(x-a) \cdot P_r g(a)(x-a)]_{\leq r}$$

(ii) Si $f : A \subset \mathbb{R}^n \to \mathbb{R}^p$ y $g : B \subset \mathbb{R}^p \to \mathbb{R}$ son funciones de clase C^r en los puntos $a \in \overset{o}{A}$ y $f(a) \in \overset{o}{B}$, entonces $g \circ f$ es de clase C^r en a y se tiene que

$$P_r(g \circ f)(a)(x-a) = [P_r g(f(a))(P_r f(a)(x-a) - f(a))]_r$$

Demostración. Que la suma, producto y composición son funciones de clase C^r era el objetivo del ejercicio propuesto al principio del capítulo. Por otra parte ya sabemos que si una función φ es de clase C^r en un punto c, entonces $P_r\varphi(c)$ es el único polinomio de grado $\leq r$ con la propiedad $\varphi \overset{r}{\sim} P_r\varphi$ en c (ver Corolario 3.8). En (i) el polinomio de grado $\leq r$ $\lambda P_r f(a)(x-a) + \mu P_r g(a)(x-a)$ tiene esta propiedad, luego según lo anterior este polinomio debe ser el desarrollo de Taylor de orden r de $\lambda f + \mu g$. En cuanto a fg también sabemos que $fg \overset{r}{\sim} P_r f \cdot P_r g$ en a lo que implica, según el corolario 3.8, que $fg \overset{r}{\sim} [P_r f \cdot P_r g]_r$. De nuevo, se tiene que el polinomio de grado $\leq r$, $[P_r f \cdot P_r g]_r$ debe ser precisamente $P_r(fg)(a)$.

La demostración para la composición (apartado (ii)) es idéntica a la del producto (ejercicio). ∎

Ejercicio. (a) Sea $g(t) = (1-t)^{-1}$. Probar que $P_r g(0)t = 1 + t + \cdots + t^r$.

(b) Sea
$$h(x,y) = \frac{1}{\cos x - \operatorname{sen} y}.$$
Utilizar el teorema 3.12(ii) y el apartado anterior para calcular el desarrollo de Taylor de orden 4 de h en el punto $(0,\pi)$.

Sugerencia: Si se escribe
$$h(x,y) = \frac{1}{1-(1-(\cos x - \operatorname{sen} y))}$$
y llamamos $f(x,y) = 1 - (\cos x - \operatorname{sen} y)$ entonces $h = g \circ f$.

Ejercicios

3A Obtener el coeficiente del término en x^4yz^2 del desarrollo de Taylor en el origen de una función de las variables x, y, z.

3B Si en el desarrollo de Taylor en el origen de una función de las variables x, y, z el único término de grado 7 es $3x^4yz^2$, ¿cuáles son las derivadas parciales de orden 7 de esta función en $(0,0,0)$?

3C Supuesta conocida la función f y sus derivadas, obtener el polinomio de Taylor de orden 2 de la función $g(x,y,z) = f(xy, xz)$ en un entorno del punto $(0,0,0)$.

3D Obtener el polinomio de Taylor de orden 3 de la funciones
$$f(x,y) = \frac{\cos x}{\cos y}\,; \quad g(x,y) = \cos xy$$
en un entorno de (0,0)

3E Sea $f(x,y) = 3x^2y^2 + x^4 + y^4 + \operatorname{sen}^3 xy$. Demostrar que el polinomio de Taylor de orden 5 para la función f en (0,0) es igual a $3x^2y^2 + x^4 + y^4$.

3F Supongamos que f es una función de clase C^6, tal que
$$\lim_{(x,y,z)\to(0,0,0)} \frac{f(x,y,z) - (2x - yz^2 + x^2yz + z^3x^2)}{(x^2+y^2+z^2)^3} = 0.$$

(a) Obtener $P_6 f(0,0,0)(x,y,z), P_5 f(0,0,0)(x,y,z)$ y $P_3 f(0,0,0)(x,y,z)$

(b) Calcular las derivadas parciales de f en $(0,0,0)$ hasta el orden 6.

3G Demostrar, sin tener que hacer el cálculo, que todas las derivadas parciales de orden ≤ 8 de la función $f(x,y) = \operatorname{sen}(x^9 + y^9)$ son nulas en (0,0). Probar asimismo que igual sucede con las de orden ≤ 5 para la función $g(x,y,z) = \cos xyz$.

3H Estudiar la existencia de los límites siguientes haciendo el desarrollo de Taylor que convenga y aplicando después el teorema de Taylor que proceda.

1. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy - \operatorname{sen} x \operatorname{sen} y}{x^2+y^2}$
2. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy - \operatorname{sen} x \operatorname{sen} y}{(x^2+y^2)^{3/2}}$
3. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy - \operatorname{sen} x \operatorname{sen} y}{\sqrt{x^6+y^6}}$
4. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\operatorname{sen} x + \cos y - \cos x - x + y^2}{x^2+y^2}$
5. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{2xy - 1 + \sqrt{\cos 2(x+y)}}{x^2+y^2}$
6. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\operatorname{sen}^2 x + \operatorname{sen}^2 y}{x^2+y^2}$
7. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xe^y - ye^x - x + y}{x^2+y^2}$
8. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xe^y - ye^x - x + y + x^3}{x^2+y^2}$
9. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xe^y - ye^x}{x^2+y^2}$
10. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xe^y - ye^x - x + y}{x^2+y^2}$

3I En cada uno de los ejemplos estudiar la existencia del límite en (0,0).

1. $h(x,y) = \dfrac{(x-y)^3}{(x-y)^2 + x^4}$.
2. $h(x,y) = \dfrac{x^5 + y^3x - x^2y^2 + y^4x}{x^4 + y^4 - x^2y^2 + y^6}$
3. $h(x,y) = \dfrac{x^5 + y^3x^2 - x^3y^2 + y^4x}{x^4 + y^4 - x^2y^2 + y^6}$
4. $h(x,y) = \dfrac{x^5 + y^3x^2 - x^3y^2 + y^4x}{x^4 + x^2y^2 + y^6}$
5. $h(x,y) = \dfrac{x^4 + y^2x^2 + y^4x + y^6}{x^4 + x^2y^2 + y^6}$
6. $h(x,y) = \dfrac{x^5 + y^4x + y^7}{x^4 + x^2y^2 + y^6}$
7. $h(x,y) = \dfrac{x^4 + y^2x^2 + (x+y)^6}{x^4 + x^2y^2 + y^6}$
8. $h(x,y) = \dfrac{x \operatorname{sen} y - y \operatorname{sen} x}{\operatorname{sen}^2 x + \operatorname{sen}^2 y}$
9. $h(x,y) = \dfrac{x \operatorname{sen} y - y \operatorname{sen} x}{(\operatorname{sen}^2 x + \operatorname{sen}^2 y)^2}$
10. $h(x,y) = \dfrac{xy - \operatorname{sen} x \operatorname{sen} y}{x^4 + y^4}$

3J Sea P un polinomio homogéneo de n variables y grado $k \geq 1$ y $f\colon \mathbb{R} \to \mathbb{R}$ una función derivable hasta el orden que necesitemos. Consideremos la función $g = f \circ P$.

(a) Probar que para cada $0 \leq j < k$ se verifica que

$$P_{kr+j}g(0)x = P_r f(0)(P(x))$$

(b) Aplicar (a) para calcular el polinomio de Taylor de orden 7 en (0,0) de la función $f(x,y) = \cos(x^2 - xy)$.

3K (a) Sea $f\colon \mathbb{R} \to \mathbb{R}$ una función n-veces diferenciable en 0. Probar que si $1 \leq k < n$, la función

$$g(t) = \frac{f(t) - P_n f(0)t}{t^{n-k}}\,; \quad g(0) = 0$$

es k-veces diferenciable en 0.

INDICACIÓN: Razonar por inducción sobre k.

(b) Deducir de (a) que si f es n-veces diferenciable en 0 ($n \geq 2$), entonces la función

$$h(t) = \frac{f(t) - f(0)}{t}\,; \quad h(0) = f'(0)$$

es $(n-1)$-veces diferenciable en 0, siendo

$$P_{n-1}g(0)t = f'(0) + 1/2!f''(0)t + \ldots + 1/n!f^{(n)}(0)t^{n-1}$$

(c) Utilizar lo anterior para demostrar que las funciones

$$g_1(x,y) = \begin{cases} \dfrac{\operatorname{sen} x - \operatorname{sen} y}{x-y} & \text{si } x \neq y \\ \cos x & \text{si } x = y \end{cases}$$

$$g_2(x,y) = \begin{cases} \dfrac{1-\sqrt{1+xy}}{xy} & \text{si } xy \neq 0 \\ 0 & \text{si } xy = 0 \end{cases}$$

son de clase C^∞ (g_2 en algún entorno de (0,0)). Obtener el polinomio de Taylor de orden 4 en (0,0) de ambas funciones.

Capítulo 4

Funciones Implícitas

El teorema de las funciones implícitas constituye, junto con el de la función inversa, la herramienta teórica básica para el estudio de las variedades diferenciables (curvas, superficies,...). Nosotros lo obtendremos como corolario del teorema de Banach de la aplicación contractiva.

1. Teorema de punto fijo

Lema 4.1 *Sea $B[b,r]$ una bola cerrada de $\mathbb{R}^p$ y k una aplicación de la bola $B[b,r]$ en sí misma. Si k es una aplicación contractiva i.e., lipschitziana con constante de Lipschitz $0 < c < 1$, entonces existe un único punto $v \in B[b,r]$ tal $k(v) = v$ (un punto fijo).*

Demostración. (*Existencia*) Supongamos $k(b) \neq b$ y consideremos la sucesión

$$b, k(b), k(k(b)), \ldots$$

Observemos que para que k tenga un punto fijo basta que esta sucesión sea de Cauchy. En efecto, en ese caso la sucesión sería convergente a un punto v de la bola $B[b,r]$ pues cada término de la sucesión está en $B[b,r]$ y los conjuntos cerrados contienen a los límites de sus sucesiones convergentes. Pero al ser k continua, la imagen por k de esta sucesión convergería a $k(v)$ i.e.,

$$b, k(b), k(k(b)), \ldots \to v \;\Rightarrow\; k(b), k(k(b)), \ldots \to k(v).$$

Pero es obvio que esta última sucesión (la sucesión imagen) es una subsucesión de la anterior y por tanto también ha de converger a v. Por la unicidad del límite resultaría entonces que $k(v) = v$.

Denotemos por $v_1 = b, v_2 = k(b), \ldots, v_{j+1} = k(v_j), \ldots$ y veamos que la sucesión $\{v_j\}$ es de Cauchy: Si $q > m$ entonces

$$\begin{aligned}
\|v_q - v_m\| &\leq \|v_q - v_{q+1}\| + \|v_{q+1} - v_{q+2}\| + \cdots + \|v_{m-1} - v_m\| \\
&= \|k(v_{q-1}) - k(v_q)\| + \|k(v_q - k(v_{q+1})\| + \cdots + \|k(v_{m-2}) - k(v_{m-1})\| \\
&\leq c(\|v_{q-1} - v_q\| + \|v_q - v_{q+1}\| + \cdots + \|v_{m-2} - v_{m-1}\|) \\
&- \\
&\leq c^q(1 + c + \cdots + c^{m-q-1})\|b - k(b)\| \\
&\leq c^q \frac{1}{1-c}\|b - k(b)\| \to 0 \quad \text{cuando } q \to \infty.
\end{aligned}$$

(*Unicidad.*) Si v_1 y v_2 fueran puntos fijos de k, entonces

$$\|v_1 - v_2\| = \|k(v_1) - k(v_2)\| \leq c\|v_1 - v_2\| \;\Rightarrow\; v_1 = v_2.$$

∎

2. El problema de las funciones implícitas

Antes de enunciar el teorema de existencia de Funciones Implícitas, haremos algunas consideraciones sobre el tipo de problemas que vamos a tratar:

Si X, Y son espacios normados y M un subconjunto no vacío de $X \times Y$, un punto $(a, b) \in M$ se dice que es punto regular de M si existe algún entorno W de (a, b) tal que el trozo de M contenido en W es exactamente bien

- la gráfica de una función de x, es decir satisface la condición

$$(x, y_1) \in M \cap W,\ (x, y_2) \in M \cap W \;\Rightarrow\; y_1 = y_2.$$

  equivalentemente existe $A \subset X$ y una función $h : A \subset X \to Y$ tal que $M \cap W = \{(x, h(x)) : x \in A\}$,

o bien,

- la gráfica de una función de y, es decir satisface la condición

$$(x_1, y) \in M \cap W,\ (x_2, y) \in M \cap W \;\Rightarrow\; x_1 = x_2.$$

  equivalentemente existe $B \subset Y$ y una función $g : B \subset Y \to X$ tal que $M \cap W = \{(y, g(y)) : y \in B\}$.

Ejercicio. Sean X, Y espacios normados, $M \subset X \times Y$ y $(a, b) \in M$. Probar que las condiciones siguientes son equivalentes:

1. Existe algún entorno W de (a, b) tal $M \cap W$ es la gráfica de una función $y = h(x)$ definida en algún entorno U de a y continua en a.
2. Existen $B(a, s), B(b, r)$ tales que $M \cap (B(a, s) \times B(b, r))$ es la gráfica de un función $y = h(x)$ definida en $B(a, s)$ y continua en a.
3. Existen $B(a, s), B(b, r)$ tales que para cada $x \in B(a, s)$ existe un único punto $y = h(x) \in B(b, r)$ tal que $(x, h(x)) \in M$ y h continua en a.
4. Existe $B(a, s)$ y una única función $h : B(a, s) \to Y$ continua en a y $h(a) = b$ tal que $(x, h(x)) \in M$.

El teorema de las funciones implícitas constituye la principal herramienta para detectar en los casos interesantes si un punto es regular. El punto de partida en este este teorema será el subconjunto M de los puntos $(x, y) \in \mathbb{R}^n \times \mathbb{R}^p$ que satisfacen alguna ecuación del tipo $f(x, y) = 0$, donde $f\colon A \subset \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^p$. Por tanto M es el lugar geométrico de los puntos de $\mathbb{R}^n \times \mathbb{R}^p$ que satisfacen el sistema de p ecuaciones:

$$
\begin{array}{l}
f_1(x_1, \ldots, x_n, y_1, \ldots, y_p) = 0 \\
f_2(x_1, \ldots, x_n, y_1, \ldots, y_p) = 0 \\
-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,- \\
f_p(x_1, \ldots, x_n, y_1, \ldots, y_p) = 0
\end{array}
$$

El objetivo del teorema no es pues resolver la ecuación $f(x, y) = 0$, que puede ser un problema bien difícil e incluso insoluble de forma exacta, sino el de establecer condiciones (suficientes) sobre f para, dado un punto $(a, b) \in M$, exista algún entorno W de (a, b) tal que $M \cap W$ sea la gráfica de una función de x y, por lo tanto, que (a, b) sea regular.

Observemos en primer lugar que muchas curvas y superficies conocidas se expresan como un lugar geométrico de este tipo: la circunferencia de centro (0,0) y radio r es el lugar geométrico de los puntos del plano que son solución de la ecuación $x^2 + y^2 = r^2$, análogamente el conjunto de puntos de $\mathbb{R}^3$ que satisfacen la ecuación $x^2 + y^2 - z^2 = r^2$ es una superficie esférica y los que satisfacen el sistema $x^2 + y^2 - z^2 = r^2;\ x + y + z = 0$ es una curva, concretamente una circunferencia. Pero también una ecuación del tipo $f(x, y) = 0$ puede originar lugares geométricos más complejos.

Ejercicio. Dibujar los conjuntos M_1 y M_2 de puntos de $\mathbb{R}^2$ que son, respectivamente, solución de la ecuación $\operatorname{sen}(x^2+y^2)=0$ y de la ecuación $\operatorname{sen} xy = 0$.

Sólo nos limitaremos a obtener herramientas para saber si al despejar y vamos a tener una sola posibilidad de hacerlo, es decir si sólo vamos a encontrar una función $y = h(x)$ tal que $f(x, h(x)) = 0$ o varias posibilidades, es decir varias funciones $y = h_i(x)$ tales que $f(x, h_i(x)) = 0$.

Ejemplo 4.2 Sea M el conjunto de puntos de $\mathbb{R}^2$ que satisfacen la ecuación $(y-1)(x-y^2) = 0$. Vamos a analizar M en un entorno de cada uno de estos tres puntos de M: $(4,-2), (1,1), (0,0)$. Consideremos el punto $(4,-2)$ y el entorno de este punto $(3,5)\times(-3,-1)$. Es obvio que la intersección de este entorno y M es justamente la gráfica de una función de x, la función h definida en $(3,5)$ (entorno del punto 4) por $h(x) = -\sqrt{x}$. Luego el punto $(4,-2)$ es un punto regular de M. Sin embargo el punto $(1,1)$ no es regular, pues la intersección de M con un entorno del punto (1,1) no es la gráfica de una única función de x, en cualquier entorno de (1,1) la ecuación permite despejar y de más de una forma pues para cada x los puntos $(x,1)$ y $(x,+\sqrt{x})$ son puntos solución. Tampoco M es la gráfica de una función de x en ningún entorno del punto $(0,0)$, sin embargo es claro que en este caso M sí que es la gráfica de una función de una función de y definida en un entorno de 0. Luego $(0,0)$ es un punto regular.

2.1. Existencia de funciones implícitas

Teorema 4.3 *Sea $f : A \subset \mathbb{R}^n \times \mathbb{R}^p \longrightarrow \mathbb{R}^p$ y (a,b) un punto de $\overset{\circ}{A}$ tal que $f(a,b) = 0$. Supongamos que*

(i) *f es continua en (a,b).*

(ii) *Las derivadas parciales $\partial f_i/\partial y_j$ existen en algún entorno del punto (a,b) y son continuas en (a,b).*

(iii) $det\left(\dfrac{\partial f_i}{\partial y_j}(a,b)\right) \neq 0.$

En estas condiciones, existen dos bolas abiertas $U = B(a,s)$ y $V = B(b,r)$ tales que si M es el lugar geométrico de los puntos que son solución de la ecuación $f(x,y) = 0$, entonces $M\cap(U\times V)$ es la gráfica de una única función h definida en U i.e., para cada $x \in U$ existe un único punto $y = h(x) \in V$ verificando la ecuación $f(x,y) = 0$. Además la función h es continua en a.

Demostración. [1] [**?**] Para concretar utilizaremos la norma $\| \|_\infty$ es decir, a lo largo de la demostración, $\| \| \equiv \| \|_\infty$. Es un ejercicio fácil comprobar que la conclusión es independiente de la norma. Vamos a ponernos en situación de aplicar el lema 4.1. Definamos para ello

$$k(x,y) = y - D_2f(a,b)^{-1}(f(x,y)),$$

donde con $D_2f(a,b)$ denotamos a la aplicación lineal de $\mathbb{R}^p$ en $\mathbb{R}^p$ cuya matriz asociada es la matriz de las derivadas parciales de $f = (f_1, f_2, \ldots, f_p)$ respecto a las coordenadas y_j en el punto (a,b), o sea

$$D_2f(a,b) \equiv \left(\frac{\partial f_i}{\partial y_j}(a,b)\right)_{i,j}$$

Observemos que k está bien definida, ya que la hipótesis (iii) garantiza la existencia de la inversa de la aplicación lineal $D_2f(a,b)$. Además es inmediato comprobar que

$$f(x,y) = 0 \iff k(x,y) = y.$$

Luego, en particular, $k(a,b) = b$.

Nos va a interesar escribir

$$k(x,y) = D_2f(a,b)^{-1}(D_2f(a,b)(y) - f(x,y))$$

Lo que significa que si llamamos

$$g(x,y) = D_2f(a,b)(y) - f(x,y) = \sum \frac{\partial f}{\partial y_j}(a,b)y_j - f(x,y),$$

entonces $k = D_2f(a,b)^{-1} \circ g$. Es claro que, de la continuidad de f en (a,b) y del hecho de que (en dimensión finita) toda aplicación lineal es lipschitziana (ver demostración de Proposición 2.8), se deduce que k también es continua en (a,b). Por otra parte, si se tiene en cuenta que

$$\frac{\partial g}{\partial y_j}(x,y) = \frac{\partial f}{\partial y_j}(a,b) - \frac{\partial f}{\partial y_j}(x,y),$$

de la continuidad en el punto (a,b) de $\frac{\partial f_i}{\partial y_j}$ se deduce que dado $\varepsilon > 0$ existe $r > 0$ tal que si $\|x - a\| \leq r$ y $\|y - b\| \leq r$ entonces $\|(\partial g_i)/\partial y_j)(x,y)\| \leq \varepsilon$.

[1] La demostración está sugerida por el método de Newton de obtención numérica de raíces en una ecuación

Resulta pues que para cada x fijo con $\|x-a\| \leq r$, la aplicación $g_x : y \to g(x,y)$ tiene sus derivadas parciales acotadas en la bola $B[b,r]$. Aplicando el teorema del valor medio a g_x, se tiene entonces que si $\|x-a\| \leq r$ entonces para todos u,v en $B[b,r]$

$$\|g(x,u)-g(x,v)\| \leq \varepsilon\|u-v\|_1 \leq p\varepsilon\|u-v\|_\infty.$$

Si β es una constante de Lipschitz para la aplicación lineal $D_2f(a,b)^{-1}$, de lo anterior se deriva que para cada x tal que $\|x-a\| \leq r$, se tiene

$$\|k(x,u)-k(x,v)\| \leq \beta\|g(x,u)-g(x,v)\| \leq p\beta\varepsilon\|u-v\|_\infty.$$

Tomando ε tal que $c = p\beta\varepsilon < 1$, la desigualdad anterior nos dice que para cada $x \in B[a,r]$ la aplicación k_x es contractiva en $B[b,r]$. Para poder aplicar el teorema del punto fijo a k_x sólo necesitaríamos que k_x aplique la $B[b,r]$ en sí misma. Esto será posible, pero habrá que tomar x en una bola centrada en a y radio (quizás) más pequeño que r. En efecto, de la continuidad de k en (a,b) se deriva que existe un número real $0 < s < r$ tal que

$$\|x-a\| \leq s \implies \|k(x,b)-k(a,b)\| < (\text{estricto})r(1-c).$$

Entonces, si $\|x-a\| \leq s$ e $y \in B[b,r]$

$$\begin{aligned}\|k(x,y)-b\| &= \|k(x,y)-k(a,b)\| \leq \|k(x,y)-k(x,b)\| + \|k(x,b)-k(a,b)\| \\ &< c\|y-b\| + r(1-c) \leq r.\end{aligned}$$

Obsérvese que lo anterior prueba que k_x aplica la bola cerrada $B[b,r]$ en la bola abierta $B(b,r)$.

En consecuencia, del teorema del punto fijo se deduce que para cada $x \in B(a,s)$ la aplicación $k_x : B[b,r] \to B[b,r]$ tiene un único punto fijo $y = h(x)$, es decir para cada $x \in B(a,s)$ existe un único punto $y = h(x) \in B(b,r)$ tal que $k(x,y) = y \iff f(x,y) = 0$.

Sólo queda probar la continuidad de la aplicación h en a: Utilizando de nuevo la continuidad de k en (a,b), dado $\varepsilon > 0$ existe $\delta > 0$ tal que si $\|x-a\| < \delta$ entonces,

$$\begin{gathered}\|h(x)-h(a)\| = \|k(x,h(x))-k(a,h(a))\| \\ \leq \|k(x,h(x))-k(x,h(a))\| + \|k(x,h(a))-k(a,h(a))\| \\ \leq c\|h(x)-h(a)\| + \varepsilon(1-c) \\ \Downarrow \\ (1-c)\|h(x)-h(a)\| \leq \varepsilon(1-c) \Leftrightarrow \|h(x)-h(a)\| \leq \varepsilon\end{gathered}$$

∎

Demostración para $n = 1,\ p = 1.$

De acuerdo con las hipótesis, suponemos pues que $\frac{\partial f}{\partial y}$ es una aplicación continua en (a, b) y que $\frac{\partial f}{\partial y}(a, b) \neq 0$. La demostración está sugerida por el método de Newton de obtención numérica de raíces en una ecuación. Vamos a ponernos en situación de aplicar el Teorema de Punto Fijo. Definamos para ello

$$k(x, y) = y - \frac{1}{\partial f/\partial y(a, b)} f(x, y).$$

Es inmediato comprobar que

$$f(x, y) = 0 \iff k(x, y) = y.$$

Luego, en particular, $k(a, b) = b$. Por tanto lo que se tiene que probar es que existen entornos U y V de a y b respectivamente tales que para cada $x \in U$ existe un único $y = h(x) \in V$ tal que $k(x, h(x)) = h(x)$. O lo que es lo mismo que si $x \in U$ la aplicación de la variable y, $k_x(y) = k(x, y)$, tiene un único punto fijo $y = h(x)$ en V.

Pues vamos a ello, intentando aplicar a k_x el teorema del punto fijo:

1. ¿k_x es contractiva? Según el teorema del valor medio así será si $|k'_x(y)| \leq c$, con $0 < c < 1$.

$$k'_x(y) = \frac{\partial k}{\partial y}(x, y) = 1 - \frac{1}{\partial f/\partial y(a, b)} \frac{\partial f}{\partial y}(x, y).$$

En particular $\frac{\partial k}{\partial y}(a, b) = 0$ y la continuidad en (a, b) de $\frac{\partial f}{\partial y}$ implica también la continuidad en (a, b) de $\frac{\partial k}{\partial y}$. Por tanto, tomando $0 < c < 1$ existe $r > 0$ tal que cuando $|x - a| \leq r,\ |y - b| \leq r$ se tiene que

$$|k'_x(y)| = |\frac{\partial k}{\partial y}(x, y) - \frac{\partial k}{\partial y}(a, b)| \leq c,$$

Luego si $x \in [a - r, a + r]$, la aplicación k_x es lipschitziana de constante c (contractiva) en el intervalo cerrado $[b - r, b + r]$, i.e.

$$|k_x(y_1) - k_x(y_2)| = |k(x, y_1) - k(x, y_2)| \leq c|y_1 - y_2|,\ \forall y_1, y_2 \in [b - r, b + r].$$

2. Para poder aplicar el teorema del punto fijo a k_x sólo necesitaríamos que k_x aplique el intervalo $[b - r, b + r]$ en sí mismo. Esto será posible, pero habrá que tomar x en un intervalo centrado en a y radio (quizás) más

pequeño que r. En efecto, de la continuidad de k en (a, b) se deriva que existe un número real $0 < s < r$ tal que

$$|x - a| \leq s \quad \Longrightarrow \quad |k(x, b) - k(a, b)| < r(1 - c).$$

Entonces, si $|x - a| \leq s$ e $|y - b| \leq r$

$$\begin{aligned} |k_x(y) - b| &= |k(x, y) - k(a, b)| \leq |k(x, y) - k(x, b)| + |k(x, b) - k(a, b)| \\ &< c|y - b| + r(1 - c) \leq r. \end{aligned}$$

Obsérvese que lo anterior prueba que k_x aplica el intervalo cerrada $[b-r, b+r]$ en en intervalo abierto $(b - r, b + r)$.

En consecuencia, del teorema del punto fijo se deduce que para cada $x \in (a - s, a + s)$ la aplicación $k_x : [b - r, b + r] \to [b - r, b + r]$ tiene un único punto fijo $y = h(x)$, es decir para cada $x \in (a - s, a + s)(= U)$ existe un único punto $y = h(x) \in (b - r, b + r)(= V)$ tal que $k(x, h(x)) = h(x) \Leftrightarrow f(x, h(x)) = 0$.

Sólo queda probar la continuidad de la aplicación h en a: Utilizando de nuevo la continuidad de k en (a, b), dado $\varepsilon > 0$ existe $\delta > 0$ tal que si $|x - a| < \delta$ entonces,

$$\begin{gathered} |h(x) - h(a)| = |k(x, h(x)) - k(a, h(a))| \\ \leq |k(x, h(x)) - k(x, h(a))| + |k(x, h(a)) - k(a, h(a))| \\ \leq c|h(x) - h(a)| + \varepsilon(1 - c) \\ \Downarrow \\ (1 - c)|h(x) - h(a)| \leq \varepsilon(1 - c) \Leftrightarrow |h(x) - h(a)| \leq \varepsilon \end{gathered}$$

La conclusión del teorema anterior suele expresarse también de alguna de las formas equivalentes siguientes:

- En $B(a, s) \times B(b, r)$ la ecuación $f(x, y) = 0$ permite despejar a y como una (única) función $y = h(x)$ definida en $B(a, s)$ y continua en a.
- En $B(a, s) \times B(b, r)$ la ecuación $f(x, y) = 0$ define implícitamente una única función $y = h(x)$ continua en a.

2.2. Funciones implícitas: derivación

El teorema de existencia de funciones implícitas y sobre todo su corolario, muestra cómo las condiciones de continuidad de la aplicación f las hereda íntegramente la función h definida implícitamente a partir de la ecuación $f(x, y) = 0$. Vamos a ver ahora que igual sucede con la diferenciabilidad.

Lema 4.4 (Lema fundamental) *Sea $f : A \subset \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^p$ y (a,b) un punto de $\overset{\circ}{A}$ tal que $f(a,b) = 0$. Supongamos que h es una función continua en a tal que $h(a) = b$ y que verifica $f(x, h(x)) = 0$ para cada x de alguna bola centrada en a. Entonces, si*

(i) *f es diferenciable en (a,b), y*

(ii) *det $\left(\dfrac{\partial f_i}{\partial y_j}(a,b)\right) \neq 0$,*

la aplicación h es también diferenciable en a y su diferencial en a se calcula mediante la fórmula:

$$Dh(a) = -D_2 f(a,b)^{-1} \circ D_1 f(a,b).$$

Demostración. Vamos a probar en primer lugar que, supuesta h diferenciable en a, su diferencial en a es la de la fórmula.

Por hipótesis existe una bola $U = B(a,s)$ tal que $f(x, h(x)) = 0$ para cada $x \in U$, es decir

$$f_i(x_1, \dots, x_n, h_1(x_1, \dots, x_n), \dots, h_p(x_1, \dots, x_n)) = 0, \; i = 1, 2, \dots, p$$

con otras palabras, para cada i, la composición de las aplicaciones

$$(x_1, \dots, x_n) \xrightarrow{\varphi} (x_1, \dots, x_n, h_1(x_1, \dots, x_n), \dots, h_p(x_1, \dots, x_n)) \xrightarrow{f_i} 0$$

es la aplicación idénticamente nula sobre U. Aplicando entonces la regla de la cadena, se tiene que

$$0 = \frac{\partial (f_i \circ \varphi)}{\partial x_j}(a,b) = \frac{\partial f_i}{\partial x_j}(a,b) + \sum_{k=1}^{p} \frac{\partial f_i}{\partial y_k}(a,b) \frac{\partial h_k}{\partial x_j}(a),$$

que matricialmente se puede expresar mediante la igualdad:

$$D_1 f(a,b) + D_2 f(a,b) \circ Dh(a) = 0 \;\Leftrightarrow\; D_2 f(a,b)^{-1} \circ D_1 f(a,b) + Dh(a) = 0.$$

Probemos que h es diferenciable en a:

Hemos de ver que

$$\lim_{u \to 0} \frac{h(a+u) - h(a) + (D_2 f(a,b)^{-1} \circ D_1 f(a,b))u}{\|u\|} = 0.$$

Como hipótesis tenemos que f es diferenciable en el punto (a,b), luego

$$\lim_{(u,v)\to(0,0)} \frac{f(a+u,b+v)-f(a,b)-D_1f(a,b)u-D_2f(a,b)v}{\|u\|+\|v\|}=0.$$

Consideremos en la expresión anterior $v=h(a+u)-h(a)$. Teniendo en cuenta que h es continua en a, es decir $(h(a+u)-h(a))\to 0$ cuando $u\to 0$) y que $f(x,h(x))=0$, se deduce que

$$\lim_{u\to 0} \frac{D_2f(a,b)(h(a+u)-h(a))+D_1f(a,b)u}{\|u\|+\|h(a+u)-h(a)\|}=0,$$

o equivalentemente (componiendo con $D_2f(a,b)^{-1}$) que

$$\lim_{u\to 0} \frac{h(a+u)-h(a)+(D_2f(a,b)^{-1}\circ D_1f(a,b))u}{\|u\|+\|v\|}=0.$$

Observamos entonces que la diferencia entre las expresiones que dan la diferenciabilidad de f en (a,b) y la diferenciabilidad de h en a está únicamente en el denominador de las mismas. Si escribimos

$$\frac{h(a+u)-h(a)+(D_2f(a,b)^{-1}\circ D_1f(a,b))u}{\|u\|}=$$

$$\frac{h(a+u)-h(a)+(D_2f(a,b)^{-1}\circ D_1f(a,b))u}{\|u\|+\|v\|}\cdot\frac{\|u\|+\|v\|}{\|u\|},$$

donde $v=h(a+u)-h(a)$, bastará probar para terminar que cuando $u\to 0$, la expresión

$$\frac{\|u\|+\|v\|}{\|u\|}$$

está acotada. Para ello tendremos en cuenta que $D_2f(a,b)^{-1}\circ D_1f(a,b)$ es lipschitziana, por ser lineal.

$$\begin{aligned}\|v\|=\|h(a+u)-h(a)\| \leq & \|h(a+u)-h(a)+(D_2f(a,b)^{-1}\circ D_1f(a,b))u\| \\ & +\|(D_2f(a,b)^{-1}\circ D_1f(a,b))u\| \\ \leq & \varepsilon(\|u\|+\|v\|)+\beta\|\|u\|,\end{aligned}$$

donde ε es arbitrario y las desigualdades se verifican para u suficientemente pequeño, es decir siempre que $\|u\|<\delta$ (que depende de ε).

Se deduce que

$$(1-\varepsilon)\|v\|\leq(\varepsilon+\beta)\|u\|.$$

Por lo que tomando por ejemplo $\varepsilon=1/2$, se tiene ya que

$$\frac{\|u\|+\|v\|}{\|u\|}=1+\frac{\|v\|}{\|u\|}\leq 1+2\beta.$$

∎

Teorema 4.5 (De las funciones implícitas) *Si en el teorema de existencia de funciones implícitas (Teorema 4.3) se sustituye la hipótesis "f continua en (a, b)"por una de las siguientes:*

- *f es continua en algún entorno de (a, b)*
- *f es diferenciable en (a, b),*
- *f es de clase C^r en algún entorno de (a, b),*

entonces, si h es la aplicación definida implícitamente por la ecuación $f(x, y) = 0$, se tiene que:

- *h es continua en algún entorno de a*
- *h es diferenciable en a,*
- *h es de clase C^r en algún entorno de a.*

Demostración. Supongamos f continua en algún entorno de (a, b). Es obvio que en la demostración del teorema de existencia podemos tomar las bolas abiertas U y V de radios s y r suficientemente pequeños para que f (y por tanto k) sea continua en cada punto de $U \times V$. La demostración de que entonces h es continua en cada punto de U es idéntica a la que la se dio para probar la continuidad en a.

La diferenciabilidad de h en a cuando f es diferenciable en (a, b) es consecuencia directa del lema de derivación (Lema 4.4).

Supongamos f de clase C^r en un abierto W que contenga al punto (a, b). Podemos suponer también que en todo punto $(x, y) \in W$ se tiene que $\det(D_2 f(x, y)) \neq 0$. En efecto, escribamos la aplicación $(x, y) \to \det(D_2 f(x, y))$ como composición de las funciones del diagrama

$$(x, y) \to \left(\frac{\partial f_i}{\partial y_j}(x, y)\right)_{i,j} \to \det \left(\frac{\partial f_i}{\partial y_j}(x, y)\right)_{i,j}.$$

Puesto que, por hipótesis, las derivadas parciales de f respecto a las coordenadas y son continuas en (a, b) y la aplicación *determinante* es también continua, la función composición es continua en (a, b). De la hipótesis $\det(D_2 f(a, b)) \neq 0$ resulta entonces que para todo (x, y) en algún entorno de (a, b), $\det(D_2 f(x, y)) \neq 0$.

Tomando ahora las bolas U y V del teorema de existencia de funciones implícitas satisfaciendo que $U \times V \subset W$, la función h está en las condiciones del lema 4.4 en cada punto $x \in U$, pues h es continua en x según veíamos antes, f diferenciable en $(x, h(x))$ y $\det(D_2 f(x, h(x)) \neq 0$. Se deduce pues que h es diferenciable en cada x de U, siendo

$$Dh(x) = -D_2 f(x, h(x))^{-1} \circ D_1 f(x, h(x)).$$

De lo anterior se deduce que en cada x de U

$$\frac{\partial h_i}{\partial x_j}(x) = \frac{-1}{\det D_2 f(x, h(x))} \sum A_{ik}(x, h(x)) \frac{\partial f_k}{\partial x_j}(x, h(x)),$$

donde las funciones $A_{ik}(x, h(x))$ son los elementos de la fila i de la matriz adjunta de $D_2 f(x, h(x))$ y por lo tanto son sumas de productos de derivadas parciales de primer orden de la función f. Lo mismo cabe decir para la función $\det D_2 f(x, h(x))$, de la que además sabemos que es distinta de cero en U. De todo ello se deduce fácilmente que las aplicaciones $\partial h_i / \partial x_j$ son aplicaciones continuas. Razonando por inducción y teniendo en cuenta el ejercicio siguiente se obtendría finalmente que h es de clase C^r (hacerlo!). ∎

Ejercicios

4A Dar ejemplos de funciones de clase C^∞, f, para las que el conjunto

$$M = \{(x, y) \colon f(x, y) = 0\}$$

sea respectivamente $\emptyset$, finito e infinito numerable.

4B Probar que el conjunto $M = \{(x, y) \in \mathbb{R}^2 \colon y^2 + \operatorname{sen}^2 1/x = 0\}$ es, en un entorno de $(0, 0)$, la gráfica de una función h, pero esta función h no está definida en ningún entorno de 0.

4C Consideremos las funciones $f_1(x, y) = -1 + \cos(x - y)$; $f_2(x, y) = (x^2 + y^2)(x^2 + (y-1)^2 - 2)$; $f_3(x, y) = (x^2 + y^2)(xy - 1)$ y sea $M_i = \{(x, y) \colon f_i(x, y) = 0\}$ $(i = 1, 2, 3)$.

(a) Probar que en $(0, 0)$ las funciones f_i no satisfacen alguna de las hipótesis del teorema de existencia de funciones implícitas.

(b) Probar que existen intervalos abiertos U, V centrados en 0 tales que $M_1 \cap (U \times V)$ es la gráfica de una función de x definida en U y continua en 0.

(c) Probar que existen intervalos abiertos U, V centrados en 0 tales que $M_2 \cap (U \times V)$ es la gráfica de una función de x definida en U, pero no continua en 0.

(d) Probar que no existen dos intervalos abiertos U, V centrados en 0 tales que $M_3 \cap (U \times V)$ sea la gráfica de una función de x definida en U.

4D Estudiar si el sistema

$$\begin{aligned} x^2 - y^2 + 2t &= 2 \\ z^2 - t^2 - xy &= -1 \end{aligned}$$

permite despejar dos de las variables como función de las otras dos en algún entorno del punto $(x_0, y_0, z_0, t_0) = (\sqrt{6}, 0, \sqrt{3}, -2)$ y en algún entorno del punto $(0, 0, 0, 1)$.

4E Sean x, y, z, t cuatro variables ligadas entre sí por las ecuaciones

$$\begin{aligned} x^3 + y^3 + z^3 + t^3 &= 0 \\ x^2 + y^2 + z^2 + t^2 &= 1 \end{aligned}$$

Comprobar que la expresión $\partial x/\partial y$ puede tener más de un significado, y calcular todos sus posibles valores para $(x_0, y_0, z_0, t_0) = (-\sqrt{3/8}, -\sqrt{1/8}, \sqrt{3/8}, \sqrt{1/8})$. ¿Qué ocurre para $(x_0, y_0, z_0, t_0) = (1/2, -1/2, 1/2, -1/2)$?

4F Considerar las funciones

$$f_1(x, y) = y - x^2 \sqrt[3]{y}; \quad f_2(x, y) = y - x \sqrt[3]{y^3 + x}.$$

Probar que las dos son diferenciables en $(0, 0)$, sus derivadas respecto a y en (0,0) son distintas de 0, pero no son continuas en (0,0). Comprobar que $f_1(x, h(x)) = 0$ para más de una función h diferenciable en 0. En cambio sólo existe una función h verificando $f_2(x, h(x)) = 0$.

Capítulo 5

Subvariedades Diferenciables de $\mathbb{R}^k$

Vamos a completar lo visto en los capítulos anteriores sobre el teorema de las Funciones Implícitas y Funciones Inversas con un tema de iniciación al estudio de las Variedades Diferenciables.

1. Variedades

Definición 5.1 (Definición explícita) Sea M un subconjunto no vacío de $\mathbb{R}^k$. Se dirá que M es una variedad diferenciable de dimensión $1 \leq n < k$ y clase C^r si para cada $c \in M$ existe algún entorno W de c tal que $M \cap W$ es la gráfica de alguna función h definida en algún abierto U de $\mathbb{R}^n$ y de clase C^r sobre U.

La primera observación que es necesario hacer es que la función h, debido al carácter local de la misma, puede cambiar de un punto a otro, y asimismo las n coordenadas de las que depende. Por tanto si M es variedad, de acuerdo con la definición, para el punto $c \in M$ existe un $W \in \mathscr{V}(c)$ y una distribución de las k coordenadas de los puntos z de W en dos bloques x, y de n y $p = k-n$ coordenadas, de forma que $M \cap W = \{(x, h(x)) : x \in U\}$, con U abierto de $\mathbb{R}^n$ y $h : U \subset \mathbb{R}^n \to \mathbb{R}^p$ de clase C^r. (Obviamente, el bloque de coordenadas x no es necesariamente el de las n primeras coordenadas de los puntos $z \in \mathbb{R}^k$, aunque, por comodidad, escribamos $z = (x, y)$).

La segunda observación es que W puede tomarse abierto. En efecto, si con las notaciones anteriores $M \cap W$ es la gráfica de la función $h \in C^r(U)$, entonces la función $x \to (x, h(x))$ es continua sobre el abierto U y por tanto el conjunto $V = \{x \in U : (x, h(x)) \in \overset{\circ}{W}\}$ es un abierto. Se tiene pues que $M \cap \overset{\circ}{W}$ es la gráfica de la restricción de h a V ¿Pueden ser distintos $M \cap W$ y $M \cap \overset{\circ}{W}$?

Teorema 5.2 (Definición implícita) *El conjunto no vacío $M \subset \mathbb{R}^k$ es una variedad diferenciable de dimensión $1 \leq n < k$ y clase C^r si y sólo si para cada $c \in M$ existe un entorno abierto W de c y una función $f\colon W \subset \mathbb{R}^k \to \mathbb{R}^{k-n}$ de clase C^r tal que*

1. Rango$(Df(c)) = k - n$.
2. $M \cap W = \{z \in W\colon f(z) = 0\}$.

Demostración. Supongamos que M es un variedad diferenciable de dimensión n y sea c un punto de M. Sea W entorno abierto de c tal que $M \cap W = \{(x, h(x))\colon x \in U\}$ para alguna función h de las n-variables x y de clase C^r en U. Podemos escribir entonces que

$$M \cap W = \{(x, y) \in U \times \mathbb{R}^p : y - h(x) = 0,$$

lo cual es obvio que implica:

$$\begin{aligned} M \cap W = M \cap W \cap (U \times \mathbb{R}^p) &= \{(x, y) \in U \times \mathbb{R}^p : y - h(x) = 0\} \\ &= \{(x, y) \in U \times \mathbb{R}^p : y - h(x) = 0\} \cap W \\ &= \{(x, y) \in (U \times \mathbb{R}^p) \cap W : y - h(x) = 0\}. \end{aligned}$$

Si f es la función definida en $U \times \mathbb{R}^p$ como $f(x, y) = y - h(x)$ y $W_1 = W \cap (U \times \mathbb{R}^p)$, entonces lo anterior nos permite escribir que

$$M \cap W_1 = \{(x, y) \in W_1 : f(x, y) = 0\}.$$

Claramente W_1 es un entorno abierto de c y $f \in C^r(W_1)$. Además, la matriz jacobiana de $Df(c)$ es la siguiente

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(c) & \cdots & \dfrac{\partial f_1}{\partial x_n}(c) & 1 & 0 & \cdots & 0 \\ \dfrac{\partial f_2}{\partial x_1}(c) & \cdots & \dfrac{\partial f_2}{\partial x_n}(c) & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial f_p}{\partial x_1}(c) & \cdots & \dfrac{\partial f_p}{\partial x_n}(c) & 0 & 0 & \cdots & 1 \end{pmatrix},$$

que tiene un menor de orden $p = k - n$ no nulo, luego $\mathrm{rg}(Df(c)) = p$.

Recíprocamente, si M satisface la condición del teorema en $c \in M$, denotemos por $y = (y_1, \ldots, y_p)$ a uno de los grupos de coordenadas tal que el menor de $Df(c)$ correspondiente a las derivaciones respecto a y_j es diferente de 0, y por $x = (x_1, \ldots, x_n)$ al grupo formado con el resto de las coordenadas. Sin pérdida de generalidad, podemos suponer que el bloque x es el de las n primeras coordenadas. Según esto, los puntos de M que están en el entorno W son los que verifican la ecuación $f(x, y) = 0$. Además, de la hipótesis se deduce que f satisface las condiciones del teorema de existencia de Funciones Implícitas en $c = (a, b)$. Luego, existen entornos abiertos U y V de a y b, respectivamente, con $U \times V \subset W$ y una función h de clase C^r de U en V, cuya gráfica es $M \cap (U \times V)$. ∎

2. Variedad tangente

En el capítulo 2 se definió el concepto de vector tangente en un punto c de un conjunto $M \subset \mathbb{R}^k$. Geométricamente, un vector tangente a M en c no era más que un vector tangente a alguna curva trazada sobre M pasando por c (ver ejercicio 2M para una definición más general de vector tangente). Se vio entonces que, en el caso particular de que M sea la gráfica de una función h de n variables reales y diferenciable en un punto a, el conjunto $T_c(M)$ de vectores tangentes a M en el punto $c = (a, h(a))$ es un espacio vectorial de dimensión n, cuya expresión es $T_c(M) = \{(u, Dh(a)u) \colon u \in \mathbb{R}^n\}$.

Proposición 5.3 *Si M es una variedad diferenciable de $\mathbb{R}^k$ de dimensión $1 \leq n < k$ y clase C^r y $c \in M$, entonces el conjunto $T_c(M)$ es un subespacio vectorial de $\mathbb{R}^k$ de dimensión n*

Demostración. Puesto que, localmente, M es la gráfica de una función de n variables y clase C^r, basta tener en cuenta lo comentado en el párrafo anterior. ∎

A $T_c(M)$ se le llamará el espacio vectorial tangente a la variedad M en c y al trasladado a c de este subespacio, i.e., $c + T_c(M)$ se le llamará pues la variedad tangente a M en c. Un punto z de $\mathbb{R}^k$ pertenecerá a la variedad tangente a M en c si $(z - c) \in T_c(M)$.

Vamos a caracterizar ahora el espacio vectorial tangente a un variedad en un punto a partir de la definición implícita. Suponemos, pues, que M es una variedad de $\mathbb{R}^k$ determinada, en un entorno W del punto $c \in M$, como el lugar geométrico de los puntos de W que satisfacen la ecuación

$f(z_1, z_2, \ldots, z_k) = 0$, donde $f = (f_1, f_2, \ldots, f_p)$ es una función de clase C^r tal que $\mathrm{rg}(Df(c)) \neq 0$. (Abreviadamente, nos referiremos a lo anterior diciendo que M está determinada en un entorno de c por la función $f = (f_1, \ldots, f_p)$). Se tiene entonces

Proposición 5.4 *Si M es la variedad determinada en un entorno del punto $c \in M$ por la función $f = (f_1, \ldots, f_p)$, entonces*

$$T_c(M) = \ker Df(c) = \cap \ker Df_i(c).$$

Demostración. Sea $v \in T_c(M)$. Entonces v es el vector velocidad en c de alguna curva γ contenida en M y que pasa por c. Es decir $\gamma(t) \in M$ y para algún t_0, $\gamma(t_0) = c$ y $\gamma'(t_0) = v$. Puesto que M está determinada en un entorno de c por f y γ esta contenida en M, se deduce que, para t suficientemente próximo a t_0, $f(\gamma(t)) = 0$. Luego derivando en t_0, se obtiene que $Df(\gamma(t_0))\gamma'(t_0) = 0$, o sea $Df(c)v = 0$, que nos dice que $v \in \ker Df(c) = \bigcap_{i=1}^p \ker Df_i(c)$. Para demostrar que $T_c(M) \supset \ker Df(c)$ basta tener en cuenta que ambos subespacios tienen la misma dimensión. En efecto, como por hipótesis, el rango de la aplicación lineal $Df(c)$ es igual a p, su núcleo debe ser un subespacio de dimensión igual a $n = k - p$. ∎

Nota. Para otras definiciones equivalentes de una variedad diferenciable ver Manual (Anexo: distintas presentaciones de una variedad).

Ejercicios

5A Estudiar si el conjunto M de los puntos que satisfacen las ecuaciones siguientes son variedades diferenciables y en tal caso de qué dimensión.

1. $x^2 + y^2 + z^2 + xy + yz - xz = 0$
2. $x^2 + y^2 + z^2 + xy + yz - xz = 1$

5B Sea M el conjunto de puntos de $\mathbb{R}^3$ que satisfacen el sistema

$$\begin{cases} x^2 + y^2 - z = 1 \\ x - y + z^2 = 1 \end{cases}$$

(a) Probar que M es una variedad diferenciable.

(b) Hallar la ecuación de la recta tangente a M en $(1, 0, 0)$.

5C Sea $M = \{(x, y, z) \neq (0, 0, 0) \colon x^3 + y^3 + z^3 = 2(x + y + z)^2\}$.

(a) Probar que M es una variedad diferenciable.

(b) Determinar los posibles puntos de M en los que el plano $x + y = 0$ sea tangente a M.

(c) Determinar la recta tangente a la curva intersección de M con el plano $x+y = 0$ en el punto $(1, -1, 2)$.

Capítulo 6

Funciones Inversas

En este capítulo estudiaremos condiciones para la derivación de la inversa de una función de varias variables y, en particular, extenderemos a estas funciones la fórmula $(g^{-1})'(g(a)) = 1/g'(a)$.

Sin embargo, el principal resultado de este capítulo es el teorema conocido como *Teorema de la Función Inversa*. Lo obtendremos a partir del teorema de las funciones implícitas y, junto a éste, constituirá otra importante herramienta de la Geometría Diferencial. Formalmente este teorema consiste en una condición suficiente para que una función de varias variables, g, admita localmente una función inversa con las mismas propiedades de diferenciabilidad que g.

1. Derivada de funciones inversas

En lo que sigue la frase "la función $g : A \subset X \to Y$ admite inversa" expresará cualquiera de los enunciados (evidentemente) equivalentes:

- cada elemento $y \in g(A)$ tiene una única anti-imagen en A;
- para cada elemento $y \in g(A)$, $A \cap g^{-1}(y)$ es un conjunto con un sólo elemento;
- para cada elemento $y \in g(A)$, existe un único $x \in A$ tal que $g(x) = y$;
- g es inyectiva en A.
- $g^{-1} : g(A) \to X$ es aplicación.

Nuestro punto de partida en esta sección será una función $g : A \subset \mathbb{R}^n \to \mathbb{R}^n$ inyectiva y diferenciable en un punto $a \in \overset{\circ}{A}$. Es bien conocido que, en general, su inversa no es diferenciable. Por ejemplo, la función de 1-variable $g(x) = x^3$ es una biyección de $\mathbb{R}$ en $\mathbb{R}$ diferenciable en 0 cuya inversa $g^{-1}(y) = \sqrt[3]{y}$ no es diferenciable en $0 = g(0)$. El motivo de esto es que para funciones de 1-variable una condición necesaria para la derivabilidad de g^{-1} en $g(a)$ es que $g'(a) \neq 0$. Esta condición se mantiene para funciones de varias variables:

Proposición 6.1 *Sea $g : A \subset \mathbb{R}^n \to \mathbb{R}^n$ inyectiva y diferenciable en un punto $a \in \overset{\circ}{A}$. Si g^{-1} es diferenciable en $g(a)$ entonces*

$$Dg^{-1}(g(a)) = Dg(a)^{-1},$$

y por tanto

$$\det Dg(a) = \det\left(\frac{\partial g_i}{\partial x_j}(a)\right) \neq 0.$$

Demostración. Puesto que $g^{-1} \circ g = \mathrm{id}_{\mathbb{R}^n}$(la identidad), aplicando la regla de la cadena resulta que

$$Dg^{-1}(g(a)) \circ Dg(a) = \mathrm{id}_{\mathbb{R}^n},$$

de lo que se deduce que $Dg(a)$ es una aplicación lineal inversible, siendo $Dg^{-1}(g(a)) = (Dg(a))^{-1}$, y por tanto $\det Dg(a) \neq 0$.

Ejercicio. Comprobar que la aplicación $g : (x, y) \to (u, v)$ cuyas funciones coordenadas son $u(x, y) = x^3 + xy^2$; $v(x, y) = y$ es una biyección de $\mathbb{R}^2$ en $\mathbb{R}^2$ cuya inversa no es diferenciable en $g(0, 0) = 0$.

La función g^{-1} del ejercicio anterior sí será diferenciable en $(u, v) = g(x, y)$ para cada (x, y) en el que $\det Dg(x, y) \neq 0$. El motivo de esto es que esta condición del determinante es también suficiente, como veremos, cuando g es de clase C^1, como es el caso del ejemplo.

Ejercicio. Siendo g la función del ejercicio anterior y $g^{-1} : (u, v) \to (x, y)$, observar que no resulta fácil obtener explícitamente g^{-1}, es decir x e y como funciones de (u, v). Obtener, sin necesidad de conocer g^{-1}, $Dg^{-1}(-3, 2)$ es decir

$$\frac{\partial x}{\partial u}(-3, 2);\ \frac{\partial x}{\partial v}(-3, 2);\ \frac{\partial x}{\partial u}(-3, 2);\ \frac{\partial y}{\partial v}(-3, 2).$$

Damos a continuación el principal resultado de esta sección. En él se establecen las condiciones justas para que la inversa de una función biyectiva y diferenciable sea también diferenciable. Si no se ha incluido en el capítulo dedicado a las reglas de derivación, es porque todo el estudio sobre funciones inversas está estrechamente ligado al de las funciones implícitas.

Teorema 6.2 *Sea $g : A \subset \mathbb{R}^n \to \mathbb{R}^n$ una aplicación inyectiva y diferenciable en un punto $a \in \overset{o}{A}$. Entonces la aplicación g^{-1} es diferenciable en el punto $b = g(a)$ si y sólo si*

(a) $b \in g(\overset{o}{A})$ *y* g^{-1} *es continua en* b.

(b) $\det\left(\dfrac{\partial g_i}{\partial x_j}(a)\right) \neq 0$.

Demostración. Si g^{-1} es diferenciable en $b = g(a)$ entonces, como hemos convenido siempre, el punto b debe ser interior al conjunto donde está definida la función g^{-1}, es decir a $g(A)$. Por otra parte g^{-1} debe ser continua en b, pues toda función diferenciable en un punto es continua en ese punto. Por último ya vimos antes que la condición $\det Dg(a) \neq 0$ era necesaria para la diferenciabilidad de g^{-1}.

Que las condiciones anteriores también son suficientes resulta del lema 4.4, aplicado a las funciones $f(x, y) = g(x) - y$; $h = g^{-1}$, cambiando en él los papeles de las coordenadas y y las coordenadas x (verificar como ejercicio todos los detalles). ∎

En la práctica los problemas que el teorema anterior plantea son dos:

1. Cómo reconocer si g es inyectiva (al menos en algún entorno del punto a).

2. Supuesta g inyectiva, cómo saber si g^{-1} está definida en algún entorno de $g(a)$ y es continua en $g(a)$. Tengamos presente que no siempre es fácil obtener explícitamente g^{-1}.

Por fortuna disponemos de herramientas de fácil aplicación (condiciones suficientes) para resolver ambos problemas. Concretamente, en el segundo de ellos, bastará que g, además de ser inyectiva, sea continua en algún entorno de a para garantizar que g^{-1} sea continua en $g(a)$. De hecho con esta hipótesis adicional g^{-1} resultaría continua en algún entorno de $g(a)$. Esto es consecuencia de un importante teorema de topología, que no tiene cabida aquí. Lo que sí veremos a continuación es otro importante teorema, que establece condiciones suficientes fácilmente verificables en la práctica, que garantizan a la vez el carácter inyectivo de g en algún entorno de a y la continuidad de g^{-1} en algún entorno de $g(a)$.

2. Inversión local

Teorema 6.3 (Lema fundamental) *Sea* $g : A \subset \mathbb{R}^n \to \mathbb{R}^n$, $a \in \overset{o}{A}$ *y supongamos que*

(a) *g admite derivadas parciales en algún entorno del punto a, continuas en a.*

(b) $\det\left(\dfrac{\partial g_i}{\partial x_j}(a)\right) \neq 0.$

Entonces existen U y V, entornos abiertos de a y $b = g(a)$ respectivamente, tales que la restricción de g a U es una biyección de U sobre V, cuya inversa es diferenciable en b.

Demostración. Observemos en primer lugar que de (a) se deriva que g es una aplicación diferenciable en a y continua en algún entorno de a (ver Teorema 2.34). Teniendo en cuenta esto, es fácil ver que la función $f(x, y) = g(x) - y$ satisface la condiciones para poder aplicar el teorema de las Funciones Implícitas en el punto (a, b) respecto de las coordenadas x, es decir siendo x el bloque de coordenadas a despejar. Concretamente, $f(a, b) = 0$, f es una función derivable en (a, b) y continua en algún entorno de (a, b), admite derivadas parciales continuas en (a, b) y

$$\det(\partial f_i/\partial x_j)(a, b) = \frac{\partial g_i}{\partial x_j}(a) \neq 0.$$

Existen, por tanto, dos entornos U_1 y V, de a y b respectivamente (si se quiere, bolas abiertas), tales que

$$\forall y \in V,\ \exists(\text{único})\, x = h(y) \in U_1 \text{ tal que } y = g(x),$$

y la aplicación h hereda las propiedades de f, luego es continua en V y diferenciable en b. De acuerdo con esto, resulta pues que cada punto y de V tiene una única antimagen x en U_1. Esto significa que si denotamos

$$\begin{aligned} U &= \{\text{los puntos de } U_1 \text{ que son antimágenes por } g \text{ de los puntos de } V\} \\ &= g^{-1}(V) \cap U_1, \end{aligned}$$

entonces U es abierto (para probarlo tened en cuenta que g puede suponerse continua sobre U_1 y aplicar la Proposición 1.16) y la restricción $g_{|U}$ es una biyección entre U y V, cuya inversa (que no es otra que la aplicación h es continua en V y diferenciable en b. ∎

Nota. Observemos que si la hipótesis (a) del lema anterior la cambiamos por la más fuerte, "g de clase C^r en algún entorno de a", entonces la función f hereda esa misma propiedad y en consecuencia la función $(g_{|U})^{-1}$ definida implícitamente por la ecuación $f(x,y)=0$ resultará entonces de clase C^r en el entorno de $g(a)$, $g(U)$.

Como consecuencia directa del lema de inversión local se tiene:

Teorema 6.4 (Teorema de la función inversa) *Sea U un abierto de $\mathbb{R}^n$ y $g: U \subset \mathbb{R}^n \to \mathbb{R}^n$ una aplicación de clase C^r sobre U. Se tiene entonces,*

1. *Las condiciones siguientes son equivalentes:*

   a) Para cada $x \in U$

   $$\det Dg(x) = \det\left(\frac{\partial g_i}{\partial x_j}(x)\right) \neq 0.$$

   b) Para cada x de U existe algún entorno abierto $U_x \subset U$ tal que $g_{|U_x}$ es inyectiva, $g(U_x)$ abierto y $g_{|U_x}^{-1}$ de clase C^r (abreviadamente una tal g se dice que es un "difeomorfismo local" de clase C^r sobre U).

2. *En particular, cuando además g sea inyectiva en U, la equivalencia anterior establece que la aplicación g^{-1} es de clase C^r (en cuyo caso se dice que g es un "difeomorfismo" entre los abiertos U y $g(U)$) si y sólo si* $\det Dg(x) \neq 0$.

Demostración. 1. a) implica b) Sólo hay que aplicar el lema de inversión y la nota posterior en cada $x \in U$. b) implica a) Como ya sabemos (Teorema 6.2), una condición necesaria para que la inversa $g_{|U_x}$ sea diferenciable en $g(x)$ es que $\det Dg(x) \neq 0$.

2. Puesto que g es inyectiva en U, la equivalencia anterior nos dice que $\det Dg(x) \neq 0$ para cada $x \in U$ si y sólo si g^{-1} es de clase C^r en el entorno abierto de $g(x)$, $g(U_x) \subset g(U)$, luego $g(U)$ abierto y las derivadas parciales de orden r de g^{-1} son continuas en cada $x \in U$.

Ejercicios

6A (T. aplicación inyectiva) Sea $g\colon U \subset \mathbb{R}^n \to \mathbb{R}^p$ una aplicación de clase C^1 sobre el abierto U y supongamos que para todo punto $x \in U$, $Dg(x)$ es inyectiva. Probar entonces que g es una aplicación localmente inyectiva.

6B (T. aplicación suprayectiva) Sea $g\colon U \subset \mathbb{R}^n \to \mathbb{R}^p$ una aplicación de clase C^1 sobre el abierto U y supongamos que para todo punto $x \in U$, $Dg(x)$ es suprayectiva. Probar entonces que g es una aplicación abierta.

6C Probar que no puede existir una aplicación de clase C^1 e inyectiva de un abierto de $\mathbb{R}^2$ en $\mathbb{R}$.

6D Sea $g\colon \mathbb{R} \to \mathbb{R}^2$ una aplicación de clase C^1. Probar que si $x \in \mathbb{R}$ existe algún entorno de x cuya imagen por g no es entorno de $g(x)$.

6E Sea g la transformación de coordenadas dada por las ecuaciones

$$u = x^2 - y; \quad v = xy$$

y consideremos el abierto $U = \{(x,y)\colon \det Dg(x,y) \neq 0\}$.

(a) Estudiar si g es un difeomorfismo local o global sobre U.

(b) ¿Es g localmente inyectiva sobre $\mathbb{R}^2$?

(c) Probar que la transformación anterior define un cambio de coordenadas (difeomorfismo global) sobre el abierto $V = \{(x,y)\colon x^2 - y < 0\}$ y utilizase en la ecuación funcional

$$\frac{\partial}{\partial y}\left[f(x,y) + x\frac{\partial f}{\partial x}(x,y)\right] = xy.$$

6F Estudiar si la aplicación

$$g(x,y,z) = (x^2 - y - z, 2x + y + z, x + y - z)$$

es un difeomorfismo del abierto $V = \{(x,y,z)\colon x > -1\}$ sobre $g(V)$.

6G Sea $g\colon (a,b) \subset \mathbb{R} \to \mathbb{R}$ una aplicación derivable en (a,b) con $g'(x) \neq 0$ para todo $x \in (a,b)$. Probar que g admite función inversa diferenciable en cada punto del abierto $g(a,b)$.

6H Si f es una función de 2 variables y clase C^2, transformar la expresión

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

mediante el cambio $x = e^u$; $y = e^v$.

6I (a) Transformar la ecuación diferencial $y' = f(x,y)$ mediante el cambio de variables $x = x(u,v)$; $y = y(u,v)$.

(b) Utilizar coordenadas polares para plantear y resolver el siguiente problema: Obtener el perfil que debe tener unas tijeras para que corten en ángulo recto.

Capítulo 7

Extremos de funciones de varias variables

En este capítulo vamos a considerar la teoría clásica de extremos para funciones diferenciables de varias variables, cuyos dos tópicos habituales son los Extremos Relativos y los Extremos Condicionados.

1. Extremos relativos

Una aplicación clásica del Teorema Local de Taylor es el estudio de los extremos relativos de una función escalar. Aunque la analogía con el caso de una variable es total, hay algunas diferencias que surgen de manera natural por el paso a una dimensión superior.

Trabajaremos con funciones escalares, definidas sobre un conjunto $A \subset \mathbb{R}^k$. Se dirá que la función $f : A \subset \mathbb{R}^k \to \mathbb{R}$ presenta un *mínimo (máximo) absoluto* en el punto $a \in A$ si $f(x) \geq f(a)$ $(f(x) \leq f(a))$ para todo $x \in A$. Y se dirá que f presenta extremo relativo en a, si existe un entorno V de a contenido en A, tal que la diferencia $f(x) - f(a)$ no cambia de signo cuando $x \in V$:

Máximo Si $f(x) - f(a) \leq 0$.

Mínimo Si $f(x) - f(a) \geq 0$.

Luego sólo cuando $a \in \overset{\circ}{A}$, es decir cuando f esté definida en alguna bola centrada en a, podremos considerar la cuestión de si f presenta un extremo relativo en a.

1.1. Condiciones necesarias de extremo

Recordemos que si queremos que una función derivable de 1-variable presente un máximo o un mínimo relativo en un punto a, la recta tangente a su gráfica en el punto $(a, f(a))$, $y = f(a)+f'(a)(x-a)$ debe ser paralela al eje X, es decir $f'(a) = 0$. Aunque sabemos que esta condición no es suficiente, pues a puede ser un punto de inflexión. Del mismo modo si f es de 2-variables es intuitivamente claro que una condición necesaria para que f presente un extremo relativo en el punto $a = (x_0, y_0)$ es que (supuesta f diferenciable en a) el plano tangente a su gráfica en en el punto $(x_0, y_0, f(x_0, y_0))$, $z = f(x_0, y_0) + Df(a)(x - x_0, y - y_0)$ sea paralelo al plano XY, es decir

$$Df(a) \equiv (\frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a)) = (0, 0).$$

Más generalmente cuando f es una función diferenciable se obtiene la siguiente condición necesaria de extremo, totalmente análoga a la de funciones de una variable.

Proposición 7.1 *Si f es diferenciable en a y presenta un extremo relativo en ese punto, entonces $Df(a) = 0$ y por tanto*

$$\frac{\partial f}{\partial x_i}(a) = 0,\ i = 1, 2, \ldots, k.$$

Demostración. Supongamos, para concretar, que f presenta un mínimo en a. Sea entonces h un vector cualquiera. Entonces:

$$Df(a)h = D_h f(a) = \lim_{t\to 0} \frac{f(a+th) - f(a)}{t} = \begin{cases} \text{lím}_{t\to 0^+} \frac{f(a+th)-f(a)}{t} & \geq 0 \\ \text{lím}_{t\to 0^-} \frac{f(a+th)-f(a)}{t} & \leq 0 \end{cases}. \blacksquare$$

7.2 Por tanto, el proceso para encontrar los posibles puntos de extremo relativo para una función diferenciable comienza con el planteamiento del sistema

$$Df(x) = 0 \ \Leftrightarrow\ \frac{\partial f}{\partial x_i}(x_1, \cdots, x_n) = 0,\ i = 1, 2, \cdots, n.$$

Los puntos solución de este sistema de n ecuaciones con n incógnitas se denominan *puntos críticos*. Después de la proposición 7.1, una condición necesaria para que la función f presente un extremo relativo en un punto x es que x sea un punto crítico. Pero esto no es, en general suficiente. ¿Cómo continuar? Recordemos que cuando f es de una variable, para saber si esta función presenta o no un extremo en un punto crítico a, se obtenían las derivadas sucesivas $f''(a), f'''(a)$.... Si todas ellas son nulas, no disponemos de ningún criterio general para saber si f presenta un extremo en a. En otro caso, del teorema de Taylor se deduce que una condición necesaria y suficiente para que f presente un extremo relativo en a es que la primera derivada que no se anule en a sea de orden par. Pero esto **no se extiende tal cual a las varias variables.** Dicha condición, aunque necesaria, no será suficiente para garantizar la existencia de extremo.

El parecido y la diferencia entre una y varias variables quedan patentes en la proposición que enunciaremos a continuación:

Consideremos pues una función $f : A \subset \mathbb{R}^k \longrightarrow \mathbb{R}$ cuyas derivadas parciales sucesivas no son todas nulas en el punto $a \in \overset{\circ}{A}$ y supongamos que la primera de estas derivadas que no se anulan en a es la de orden r. Para poder aplicar el teorema de Taylor, supondremos también que f es de clase C^r en a. En este caso, puesto que todas las derivadas parciales de orden menor estrictamente que r se anulan en a, es claro que

$$\begin{aligned} P_r f(a)h &= f(a) + \frac{1}{r!} D^r f(a) h^r \\ &= f(a) + \sum_{r_1+\ldots+r_k=r} \frac{1}{r_1!\ldots r_n!} \frac{\partial^r f(a)}{\partial x_1^{r_1} \partial x_2^{r_2} \ldots \partial x_k^{r_k}} h_1^{r_1} h_2^{r_2} \ldots h_k^{r_k}. \end{aligned}$$

En lo sucesivo, siempre que no haya lugar a confusión, denotaremos por Q_p a la suma de los términos de grado p del polinomio de Taylor de f en a. Es decir:

$$P_r f(a) = Q_0 + Q_1 + \ldots + Q_r; \; Q_p(h) = \frac{1}{p!} D^p f(a) h^p.$$

Proposición 7.3 *En estas condiciones, es decir si la primera de las derivadas parciales sucesivas de f que no se anulan en a es la de orden r y f es de clase C^r en a, entonces:*

1) *Una condición necesaria para que f presente un mínimo en a es que r sea un número par y $Q_r(h) \geq 0$ para todo $h \in \mathbb{R}^k$ (≤ 0 para todo h para máximo).*
2) *Una condición suficiente para que f presente un mínimo en a es que r sea par y $Q_r(h) > 0$ para todo $h \in \mathbb{R}^k$, $h \neq 0$ (< 0 para todo $h \neq 0$ para máximo).*

Demostración. Para probar 1), supongamos para concretar que f presenta un mínimo relativo en a. De acuerdo con las hipótesis y teniendo en cuenta que para cada polinomio homogéneo de grado r se satisface la igualdad $Q(th) = t^r Q(h)$, es fácil ver que el teorema local de Taylor nos dice que si h es un vector de $\mathbb{R}^k$ entonces:

$$\lim_{t\to 0} \frac{f(a+th)-f(a)}{t^r} = \frac{1}{r!} D^r f(a)h^r = Q_r(h). \tag{7.1}$$

Como estamos suponiendo que f presenta un mínimo en a, $f(a+th) - f(a)$ es ≥ 0 para t suficientemente pequeño. Por tanto r debe ser par, pues si r fuese impar el límite en (7.1) debería ser igual a 0 , ya que este límite por la derecha sería ≥ 0 mientras que por la izquierda sería ≤ 0. Como el valor de este límite es $\frac{1}{r!}D^r f(a)h^r$, entonces de ser r impar se tendría que $D^r f(a)h^r = 0$ para todo h, en contra de nuestra suposición de que las derivadas parciales de orden r no son todas nulas. Luego la igualdad (7.1) prueba que r es par y que $Q_r(h) \geq 0$ para todo h.

2) Supongamos ahora que (el polinomio homogéneo de grado r) $Q_r(h) = 1/(r!)D^r f(a)h^r$ es estrictamente mayor que 0 para todo $h \neq 0$ de $\mathbb{R}^k$, y sea λ el valor mínimo que alcanza este polinomio sobre la esfera $\{u : \|u\| = 1\}$ (que es un compacto). De acuerdo a nuestra hipótesis es claro que $\lambda > 0$. Teniendo en cuenta que, al ser Q_r un polinomio homogéneo de grado r,

$$\frac{Q_r(h)}{\|h\|^r} = Q_r\left(\frac{h}{\|h\|}\right),$$

del teorema local de Taylor, se deduce que

$$0 = \lim_{h\to 0} \frac{f(a+h) - f(a) - \frac{1}{r!}D^r f(a)h^r}{\|h\|^r} = \lim_{h\to 0} \frac{f(a+h)-f(a)-Q_r(h)}{\|h\|^r}.$$

Y, por lo tanto, para $\varepsilon = \lambda/2$ existe un $\delta > 0$ tal que si $0 < \|h\| < \delta$, entonces

$$\begin{aligned} -\frac{\lambda}{2} &\leq \frac{f(a+h)-f(a)-Q_r(h)}{\|h\|^r} \\ &= \frac{f(a+h)-f(a)}{\|h\|^r} - Q_r(\frac{h}{\|h\|})^r \\ &\leq \frac{f(a+h)-f(a)}{\|h\|^r} - \lambda. \end{aligned}$$

Lo que implica que

$$\frac{f(a+h)-f(a)}{\|h\|^r} \geq \frac{\lambda}{2}$$

si $0 < \|h\| < \delta$ y, por lo tanto que $f(a+h) - f(a) > 0$. Así pues f presenta un mínimo (estricto) en a. ∎

7.4 En resumen, el procedimiento general para la obtención de los puntos de extremos relativos es el siguiente:

1. Obtención de los Puntos Críticos

$$\frac{\partial f}{\partial x_i}(x_1, \cdots, x_n) = 0,\ \ i = 1, 2, \cdots, n$$

   Supongamos que a es un punto crítico y la primera derivada de la función f que no se anula en a es la de orden r, entonces:

2. r **impar** La función f no tiene extremos en a

3. r **par** y $Q_r(h) > 0$ para todo $h \neq 0$. La función presenta un **Mínimo** en a.

4. r **par** y $Q_r(h) < 0$ para todo $h \neq 0$. La función presenta un **Máximo** en a.

5. r **par** pero $Q_r(h)$ no tienen signo constante. La función no tiene extremos en a.

6. r **par**, $Q_r(h)$ tiene signo constante, pero existe algún $h \neq 0$ tal que $Q_r(h) = 0$. Caso **Dudoso**.

Aplicamos a continuación el estudio anterior a algunos ejemplos:

Ejemplo 7.5 Sea

$$f(x, y) = 6x^2y - 2y^3 - 3x^4.$$

1. Puntos críticos:

$$\frac{\partial f}{\partial x}(x, y) = 12xy - 12x^3 = 0$$
$$\frac{\partial f}{\partial y}(x, y) = 6x^2 - 6y^2 = 0.$$

Es claro que los únicos puntos que satisfacen el sistema anterior (los puntos críticos), son: $(0, 0), (1, 1), (-1, 1)$

Consideremos en primer lugar el (0,0):

Puesto que f es un polinomio de grado 4, sabemos que $P_4f(x_0, y_0)(x, y) = f(x + x_0, y + y_0)$. En particular $P_4f(0, 0)(x, y) = f(x, y) = 6x^2y - 2y^3 - 3x^4$.

Se tiene pues que todas las derivadas parciales de 1º y 2º orden de f en (0,0) son nulas, luego

2. $r = 3$ (impar) y por tanto f no presenta un extremo en $(0, 0)$.

Estudiamos ahora f en el punto crítico (1,1):

$$\begin{aligned} P_4f(1,1)(x,y) = f(x+1,y+1) &= 6(x+1)^2(y+1) - 2(y+1)^3 - 3(x+1)^4 \\ &= 1 - 12x^2 - 6y^2 + 12xy + 6x^2y - 12x^3 - 2y^3 - 3x^4. \end{aligned}$$

Por lo tanto las derivadas parciales de primer orden de f en (1,1) son nulas mientras que las de orden 2 no, luego

2. $r = 2$ (par) y por tanto no podemos descartar aún la posibilidad de que f presente un extremo en este punto.

3. $Q_2(x, y) = -12x^2 - 6y^2 + 12xy$, y es fácil ver que podemos escribir $Q_2(x, y) = -6[x^2 + (x-y)^2]$, de lo que se deduce que $Q_2(x, y) < 0$ para todo $(x, y) \neq (0, 0)$ y que , por lo tanto, f presenta un máximo relativo estricto en $(1, 1)$.

Por último analizamos qué sucede en (-1,1):

$$\begin{aligned} P_4f(-1,1)(x,y) = f(x-1,y+1) &= 6(x-1)^2(y+1) - 2(y+1)^3 - 3(x-1)^4 \\ &= 1 - 12x^2 - 6y^2 - 12xy + 6x^2y + 12x^3 - 2y^3 - 3x^4. \end{aligned}$$

Por lo tanto las derivadas parciales de primer orden de f en (-1,1) son nulas mientras que las de orden 2 no, luego de nuevo

2. $r = 2$ (par).

3. $Q_2(x, y) = -12x^2 - 6y^2 - 12xy = -6[x^2 + (x+y)^2]$. Por tanto también ahora $Q_2(x, y) < 0$ para todo $(x, y) \neq (0, 0)$ y por lo tanto f presenta un máximo relativo estricto en $(-1, 1)$.

En los casos en que la primera derivada que no se anule en un punto sea la segunda, es decir cuando $r = 2$, existen ciertos atajos para estudiar la existencia de extremo en ese punto, que vemos a continuación:

1.2. Caso particular

Según lo visto antes, en el caso particular de que la primera de las derivada parciales sucesivas de la función f que no se anule en el punto a sea la de orden 2, la existencia de extremo de f en a está estrechamente relacionada con el signo de $Q_2(h)$ (el polinomio formado por los términos de grado 2 del polinomio de Taylor en a). Para estudiar el posible signo de este polinomio (en este caso y sólo en este caso) podemos utilizar un resultado que se deriva de la teoría de las formas cuadráticas (ver GANTMACHER, F.R., "Théorie des Matrices" Tome 1, Dunod, Paris, 1966.):

Denotemos por $H_f(a)$ a la matriz de las derivadas parciales segundas de f en a (que se denomina *Hessiano* de f en a).

$$H_f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_k}(a) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_k \partial x_1}(a) & \frac{\partial^2 f}{\partial x_k \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_k^2}(a) \end{pmatrix}$$

Se tiene entonces :

(i) $Q_2(h) \geq 0,\ \forall h$ si y sólo si todos los menores principales de $H_f(a)$ son no negativos (condición necesaria de mínimo).

(ii) $Q_2(h) \leq 0,\ \forall h$ si y sólo si los menores principales de orden par de $H_f(a)$ son no negativos y los menores principales de orden impar no positivos (condición necesaria de máximo).

(iii) $Q_2(h) > 0,\ \forall h \neq 0$ si y sólo si los menores principales de $H_f(a)$ formado por las i primeras filas y las i primeras columnas $i = 1, 2, \ldots, k$

$$\Delta_i = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_i}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_i}(a) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_i \partial x_1}(a) & \frac{\partial^2 f}{\partial x_i \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_i^2}(a) \end{vmatrix},$$

son mayores (estrictamente) que 0 (condición suficiente de mínimo).

(iv) $Q_2(h) < 0,\ \forall h \neq 0$ si y sólo si $\Delta_1 < 0,\ \Delta_2 > 0, \ldots, (-1)^i \Delta_i > 0, \ldots,$ (condición suficiente de máximo).

Aplicando este criterio al ejemplo anterior en el punto (-1,1), para el que $Q_2(x,y) = -12x^2 - 6y^2 - 12xy$, se tendría que

$$H_f(-1,1) = \begin{pmatrix} -24 & -12 \\ -12 & -12 \end{pmatrix}$$

Por lo que $\Delta_1 = -24 < 0,\ \Delta_2 = 144 > 0$. Según el criterio anterior, esto significa que $Q_2(x,y) < 0$ para todo $(x,y) \neq (0,0)$ y por lo tanto que f presenta un máximo en $(-1,1)$.

Ejemplo 7.6 Sea

$$f(x,y,z) = \frac{\cos xz}{1 - xyz}$$

definida en el abierto $U = \{(x,y,z) \in \mathbb{R}^3 : xyz < 1\}$.

1. Puntos críticos.

Es fácil ver (ejercicio) que los únicos puntos críticos de f en U son los puntos $(c,0,0), (0,c,0), (0,0,c)$ para cualquier $c \in \mathbb{R}$. A continuación vamos a analizar si f presenta extremo en los puntos $(0,0,0), (1,0,0), (0,1,0)$.

2. En general, parece razonable tratar de aplicar, si fuera posible, el atajo anterior. Para lo cual, dado un punto crítico, el Hessiano de f en ese punto debería ser distinto de la matriz idénticamente nula. En caso contrario, es decir si las primeras derivadas no nulas son las de orden r y $r > 2$, deberemos recurrir (de acuerdo al estudio general) a analizar Q_r, el polinomio formado por los términos de grado r del polinomio de Taylor en el punto. En este ejemplo, $r = 2$ en el punto (1,0,0) y en el (0,1,0), pero es mayor que 2 en el punto (0,0,0), como puede comprobarse haciendo el cálculo directo de las derivadas parciales segundas (que resulta un poco pesado) en esos puntos. Podemos evitar este cálculo directo de derivadas obteniendo los polinomios de Taylor y deducirlas de los coeficientes de estos polinomios. Consideremos pues, en primer lugar, el punto crítico $(0,0,0)$. Teniendo en cuenta que los desarrollos de Taylor de las funciones de 1-variable $\cos t$, $\frac{1}{1-t}$ en 0 son:

$$\cos t \equiv 1 - \frac{1}{2}t^2 + \cdots ; \quad \frac{1}{1-t} \equiv 1 + t + t^2 + \cdots ,$$

es fácil ver (ejercicio) que los sucesivos polinomios de Taylor de f en (0,0,0) se obtienen así:

$$Pf(0,0,0)(x,y,z) \equiv (1 - \frac{1}{2}x^2z^2 + \cdots)(1 + xyz + \cdots) \equiv 1 + xyz + \cdots$$

Lo que nos indica que la primera derivadas no nula de f en (0,0,0) es

$$\frac{\partial f}{\partial x \partial y \partial z}(0,0,0) = 1,$$

es decir que $r = 3$ (impar) y por tanto f no presenta extremos en $(0, 0, 0)$.

Análogamente los sucesivos polinomios de Taylor de f en (1,0,0) y (0,1,1) se obtienen así:

$$\begin{aligned} Pf(1,0,0)(x,y,z) &\equiv (1 - \frac{1}{2}(x+1)^2 z^2 + \cdots)(1 + (x+1)yz + \cdots) \\ &\equiv 1 - \frac{1}{2}z^2 + yz + \cdots \\ Pf(0,1,0)(x,y,z) &\equiv (1 - \frac{1}{2}x^2 z^2 + \cdots)(1 + x(y+1)z + \cdots) \\ &\equiv 1 + xz + \cdots . \end{aligned}$$

Se deduce pues que

$$H_f(1,0,0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}; \quad H_f(0,1,0) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

lo que de acuerdo al criterio del hessiano nos dice que f no presenta tampoco extremos en ninguno de estos puntos, ya que en ambos hessianos existe un menor principal de orden 2 que es menor que cero.

2. Extremos condicionados

Vamos a desarrollar en esta sección una técnica clásica, basada en el uso de Multiplicadores, para la obtención de condiciones de extremo sobre una variedad . Dichas condiciones no diferirán formalmente de las ya vimos para extremos relativos.

Definición 7.7 Sea $\varphi \colon A \subset \mathbb{R}^k \to \mathbb{R}$ una función escalar, $c \in \mathring{A}$ y M una variedad diferenciable que contiene a c. Se dirá que φ presenta un extremo sobre M (o condicionado) en el punto c, si existe un entorno W de c tal que $\varphi(z) - \varphi(c)$ no cambia de signo cuando $z \in W \cap M$.

Máximo Si $\varphi(z) - \varphi(a) \leq 0$.

Mínimo Si $\varphi(z) - \varphi(a) \geq 0$.

Proposición 7.8 *En la situación anterior y supuesta φ diferenciable en c, una condición necesaria para que φ presente un extremo en c sobre la variedad M es que $D\varphi(c)w = 0$, para todo $w \in T_c(M)$. Se dice en ese caso que c es un punto crítico de φ sobre la variedad M.*

Demostración. Sea $w \in T_c(M)$ y γ una curva contenida en M que pase por c y tenga a w por vector tangente en c. Es decir, $\gamma(t_0) = c$, y $\gamma'(t_0) = w$ en algún punto t_0. Entonces la función de la variable t, $\varphi \circ \gamma$, presenta un extremo relativo en el punto t_0, luego $(\varphi \circ \gamma)'(t_0) = 0$. Aplicando la regla de la cadena se tiene pues

$$0 = (\varphi \circ \gamma)'(t_0) = D\varphi(\gamma(t_0))\gamma'(t_0) = D\varphi(c)w.$$

∎

Corolario 7.9 (Multiplicadores de Lagrange) *Sean M y φ como en la proposición anterior y supongamos que M está determinada, en un entorno del punto $c \in M$, por las funciones $f_1, f_2, \ldots, f_p$. Entonces, una condición necesaria para que la aplicación φ presente un extremo sobre la variedad M en el punto c, es que existan p números reales $\lambda_1, \ldots, \lambda_p$ tales que c sea un punto crítico de la función $F = \varphi + \lambda_1 f_1 + \ldots + \lambda_p f_p$.*

Demostración. Por la proposición anterior, si φ presenta un extremo condicionado en c entonces $D\varphi(c)w = 0$ para todo $w \in T_c(M) = \cap \ker Df_i(c)$. Luego

$$\ker D\varphi(c) \supset \cap \ker Df_i(c),$$

y esto implica ya lo que queríamos, en virtud del lema algebraico siguiente

Lema 7.10 *Si $g_1, g_2, \ldots, g_p$ son formas lineales independientes de $\mathbb{R}^k$ y g es otra forma lineal tal que $\ker g \supset \cap_{i=1}^p \ker g_i$, entonces existen p números reales $\lambda_1, \ldots, \lambda_p$ tales que*

$$g + \sum \lambda_i g_i = 0.$$

Demostración. Completemos la familia de formas lineales $g_1, g_2, \ldots, g_p$ hasta obtener una base $\mathcal{B} = \{g_1, \ldots, g_p, g_{p+1}, \ldots, g_k\}$. Entonces

$$g = \sum_{i=1}^{k} \mu_i g_i.$$

Sea $\{e_1, \ldots, e_k\}$ la base dual de $\mathcal{B}$, es decir $g_i(e_j) = \delta_{ij}$. Es evidente entonces que si $i \leq p < j$,

$$e_j \in \bigcap_{i=1}^{p} \ker g_i \subset \ker g,$$

luego

$$0 = g(e_j) = \sum_{i=1}^{k} \mu_i g_i(e_j) = \mu_j g_j(e_j) = \mu_j,$$

lo que implica que $g = \sum_{i=1}^{p} \mu_i g_i$.
Tomando $\lambda_i = -\mu_i$ resulta lo que queríamos. ∎

Nota. Los ejercicios 7E,7F,7H, 7I y 7J constituyen ejemplos prácticos de aplicación del resultado anterior sobre multiplicadores de Lagrange. En el ejercicio 7E se da una indicación que quizás pueda servir como modelo en los demás ejercicios. En todos ellos se parte de un función diferenciable φ, que se sabe que alcanza un máximo o mínimo absoluto sobre una variedad M, y se trata de obtener este valor máximo o mínimo y algún punto $c \in M$ en donde se alcance. Es inmediato comprobar que tales puntos c deben ser puntos de extremo (condicionado) sobre M y por lo tanto puntos críticos de φ sobre M.

Ejercicios

7A Estudiar la existencia de extremos relativos para las funciones

1. $f(x, y) = x^4 + y^4 - 2(x - y)^2$
2. $f(x, y) = x^3 + y^2 - xy^2$
3. $f(x, y) = x^3 + y^3 - 3x^2y^2 + 1$
4. $f(x, y, z) = x^4 + y^4 - 3x^2y^2xz$
5. $f(x, y, z) = 2x^4 + 2y^4 + z^4 - 4x^2y^2z$
6. $f(x, y) = 2x^4 + 3y^4 - 4x^2y^3$
7. $f(x, y) = x^2y + x^2 + 2xy + xy^2 + y^2$
8. $f(x, y) = x^4 + y^4 - xy^3 + 1$
9. $f(x, y) = x^4 + y^4 - xy^4 + 1$
10. $f(x, y) = y^2 - 3x^2y + 2x^4$
11. $f(x, y) = x^3 + 3x^2y - xy^2 + 1$
12. $f(x, y) = x^2y^2 + xy^4$

7B Demostrar que los siguientes enunciados son equivalentes

(a) $x^2 + y^2 + z^2 \geq \alpha(xy + xz + yz)$, para todos x, y, z.

(b) La función $f(x, y, z) = x^2 + y^2 + z^2 - \alpha(xy + xz + yz)$, presenta un mínimo relativo en el punto $(0, 0, 0)$.

(c) $\alpha \in [-2, 1]$.

7C Estudiar la existencia de extremos relativos para las funciones

$f(x,y) = \operatorname{sen}|xy|, \quad 0 < x \leq \pi/2, \quad 0 < y \leq \pi/2.$

$f(x,y) = e^{ax+y^2} + b\operatorname{sen}(x^2+by^2) - y^2.$

$f(x_1,\ldots,x_n) = x_1x_2\cdots x_n + a^{n+1}\left(\frac{1}{x_1}+\ldots+\frac{1}{x_n}\right), \ (a>0)\ , \ x \in \Omega = (0,\infty)^n$

$f(x_1,\ldots,x_n) = (\sum a_ix_i)e^{-(x_1^2+\ldots+x_n^2)}$

7D Demostrar que entre los polígonos convexos de n lados inscritos en una circunferencia, el polígono regular es el de mayor perímetro y también el de mayor área.

7E Hallar la distancia

(a) entre la recta $x+y=1$ y la hipérbola $xy=2$.

INDICACIÓN: Considerar $M = \{(x,y,z,t) \in \mathbb{R}^4 : x+y=1; zt=2\}$ y φ la aplicación $\varphi(x,y,z,t) = (x-z)^2+(y-t)^2$. M es una variedad diferenciable de $\mathbb{R}^4$ y el cuadrado de la distancia entre las dos curvas, el mínimo absoluto de φ sobre M.

(b) entre las curvas

$$\frac{x}{1} = \frac{y}{2} = \frac{z}{4}\ ; \qquad \begin{cases} x^2-xy+y^2-z^2 & =1 \\ x^2+y^2 & =1 \end{cases}$$

7F Hallar la distancia al origen de la curva intersección de las superficies

$$xyz = a; \quad y = bx, \quad a>0,\ b>0.$$

7G Sea M el conjunto de puntos de $\mathbb{R}^3$ que satisfacen el sistema

$$\begin{cases} x^2+y^2-z=1 \\ x-y+z^2=1 \end{cases}$$

(a) Probar que M es una variedad diferenciable.

(b) Hallar la ecuación de la recta tangente a M en $(1,0,0)$.

(c) Estudiar si $(1,0,0)$ es un punto crítico de la función $g(x,y,z) = xyz-x-y+z$ sobre M.

7H (a) Obtener el paralelepípedo recto de menor área entre los que tienen igual volumen.

(b) Obtener el paralelepípedo de mayor volumen entre los que tienen igual área.

7I Sea $K = \{(x, y, z) \in \mathbb{R}^3 \colon z = x^2 + y^2\,; z = x + 1\}$. Probar que K es un compacto y hallar el máximo y el mínimo absoluto de la función $\varphi(x, y, z) = y^2 - 2x + z$ sobre K.

7J Calcular el máximo y el mínimo absoluto de la función $\varphi(x, y, z) = x + y - z$ sobre el compacto $K = \{(x, y, z) \colon x^2 + y^2 \leq 1,\ x^2 + y^2 + z^2 \leq 4\}$.

Bibliografía

[1] Apostol, T., "Análisis Matemático", Reverté, Barcelona, 1979.

[2] Avez, A., "Calcul Différentiel", Masson, Paris, 1983.

[3] Benedetto, J.J., "Real Variable and Integration ", B.G. Teubner, Stuttgart, 1976.

[4] Berstein, S., Quelques remarques sur l'interpolation, *Math. Ann.*, **79** (1918), 1 – 12.

[5] Cartan, H., "Cálculo Diferencial", Ediciones Omega, Barcelona, 1972.

[6] Cheney, E.W., "Introduction to Approximation Theory", McGraw-Hill, New York, 1966.

[7] Choquet, "Topología", Toray-Masson, Barcelona, 1971.

[8] Cohen, P.J., A minimal model for set theory, *BAMS*, **69** (1963), 537 – 540.

[9] Cohn, D.L., "Measure Theory", Birkhäuser, Boston, 1980.

[10] Dieudonné, J., "Fundamentos de Análisis Moderno", Reverté, Barcelona, 1966.

[11] Fleming, W.H., "Funciones de Varias Variables", C.E.C.S.A., 1976.

[12] Flett, T.M., "Differential Analysis", Cambridge University Press, 1980.

[13] Gantmacher, F.R., "Théorie des Matrices", Tome 1 Dunod, Paris, 1966.

[14] Garnir, H.G., "Teoría de Funciones I", Marcombo, Barcelona, 1966.

[15] Garnir, H.G., "Functions de Variables Reelles", tome II, Gauthier-Villars, París, 1965.

[16] Hewitt, E., Stromberg, K., "Real and Abbstract Analysis", Springer-Verlag, New York, 1965.

[17] Hewitt, E., Ross, K.A., Extensions of Haar measure and of harmonic analysis for locally compact abelian groups, *Math. Annalen* , **160** (1965), 171 – 194.

[18] Isaacson, E. and Keller, H.B., “Analysis of Numerical Methods”, Dover, New Yory, 1994.

[19] Kakutani, S., Oxtoby, J.C., , *Ann. of Math.,* **(2) 52** (1950), 580 – 590.

[20] Kolmogorov, A.N., Fomin, S.V., “Elementos de la Teoría de Funciones y del Análisis Funcional ”, Mir, Moscú, 1975.

[21] Moore, G.H., “Zermelo’s Axiom of Choice ”, Springer-Verlag, New York, 1982.

[22] Nachbin, L., “Topology on Spaces of Holomorphic Mappings”, Springer, Berlin, 1969.

[23] Natanson, I.P., “Theory of Functions of a Real Variable ”, Ungar, New York, 1967.

[24] Von Newman, J., “Functional Operators Vol.I: Measures and Integrals”, Princeton Univ. Press, , 1950.

[25] Rudin, W., “Análisis Real y Complejo”, McGraw Hill, Madrid, 1973.

[26] Schwartz, L., “Cours d’Analyse”, Hermann, Paris, 1967.

[27] Solovay, R.M., A model of set-theory in which every set of reals is Lebesgue medible, *Ann. of Math.,* **92** (1970), 1 – 56.

[28] Williamsom, J.H., “Integración Lebesgue”, Tecnos, Madrid, 1973.

[29] Yosida, K., “Funtional Analysis”, Sspringer, Berlin, 1974.

[30] Zaanen, A.C., “Integration”, North-Holland, Amsterdam, 1967.