Cálculo Diferencial en $\mathbb{R}^n$

Francisco Javier Pérez González

Departamento de Análisis Matemático

Universidad de Granada

Septiembre de 2016

Índice general

Capítulo 1

Estructura euclídea y topología de $\mathbb{R}^n$

1.1. Producto escalar y norma euclídeos

Como sabes, $\mathbb{R}^n$ es un espacio vectorial en el que suele destacarse la llamada base canónica formada por los vectores $\{\mathbf{e_1}, \mathbf{e_2}, \dots, \mathbf{e_n}\}$ donde $\mathbf{e_k}$ es el vector cuyas componentes son todas nulas excepto la que ocupa el lugar k que es igual a 1.

1.1 Definición. Dados dos vectores $\mathbf{x} = (x_1, x_2, \dots, x_n)$, $\mathbf{y} = (y_1, y_2, \dots, y_n)$ se define su producto escalar por:

$$\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Este producto escalar se llama *producto escalar euclídeo*.

Observa que el producto escalar de dos vectores no es un vector sino un número real. La notación **x.y** es frecuentemente usada en los libros de Física para representar el producto escalar de los vectores **x** e **y**.

Propiedades del producto escalar.

1. $\langle \mathbf{x} | \mathbf{x} \rangle \geqslant 0$, y $\langle \mathbf{x} | \mathbf{x} \rangle = 0 \Longleftrightarrow \mathbf{x} = \mathbf{0}$.

2. *Simetría.* $\langle \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{y} | \mathbf{x} \rangle$ para todos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

3. *Linealidad.* $\langle \alpha \mathbf{x} + \beta \mathbf{y} | \mathbf{z} \rangle = \alpha \langle \mathbf{x} | \mathbf{z} \rangle + \beta \langle \mathbf{y} | \mathbf{z} \rangle$ para todos $\alpha, \beta \in \mathbb{R}$ y para todos $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$.

Dichas propiedades del producto escalar se deducen fácilmente de su definición.

1.2 Definición. La *norma euclídea* de un vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ se define por

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

Desigualdad de Cauchy-Schwarz.

Para todos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *se verifica que*

$$\left|\langle \mathbf{x}|\mathbf{y}\rangle\right| \leqslant \|\mathbf{x}\|_2 \|\mathbf{y}\|_2$$

Además, supuesto que $\mathbf{x}$ *e* $\mathbf{y}$ *no son nulos, la igualdad* $\left|\langle \mathbf{x}|\mathbf{y}\rangle\right| = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2$ *equivale a que exista un número* $\lambda \in \mathbb{R}$ *tal que* $\mathbf{x} = \lambda \mathbf{y}$ *(es decir, los vectores* $\mathbf{x}$ *e* $\mathbf{y}$ *están en una misma recta que pasa por el origen).*

Propiedades de la norma euclídea.

1. $\|\mathbf{x}\|_2 \geqslant 0$, y $\|\mathbf{x}\|_2 = 0 \Longleftrightarrow \mathbf{x} = \mathbf{0}$.

2. *Homogeneidad*. $\|\lambda \mathbf{x}\|_2 = |\lambda| \, \|\mathbf{x}\|_2$ para todo $\mathbf{x} \in \mathbb{R}^n$ y todo $\lambda \in \mathbb{R}$.

3. *Desigualdad triangular.* Para todos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ se verifica que

$$\|\mathbf{x} + \mathbf{y}\|_2 \leqslant \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2$$

Además, supuesto que $\mathbf{x}$ e $\mathbf{y}$ no son nulos, la igualdad $\|\mathbf{x} + \mathbf{y}\|_2 = \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2$ equivale a que hay un número $\lambda > 0$ tal que $\mathbf{x} = \lambda \mathbf{y}$ (es decir, los vectores $\mathbf{x}$ e $\mathbf{y}$ están en una misma semirrecta que pasa por el origen).

1.3 Definición. Se dice que los vectores $\mathbf{x}$ e $\mathbf{y}$ son **ortogonales**, y escribimos $\mathbf{x} \perp \mathbf{y}$, cuando su producto escalar es cero. Se dice que un vector $\mathbf{x}$ es ortogonal a un conjunto de vectores $E \subset \mathbb{R}^n$ cuando $\mathbf{x}$ es ortogonal a todo vector en E. Un conjunto de vectores no nulos que son mutuamente ortogonales se dice que es un **conjunto ortogonal** de vectores; si, además, los vectores tienen todos norma 1 se dice que es un **conjunto ortonormal** de vectores. Una base vectorial que también es un conjunto ortogonal (ortonormal) se llama una **base ortogonal** (ortonormal).

Si $\mathbf{x}$ e $\mathbf{y}$ son vectores no nulos, el vector

$$\prod_{\mathbf{y}}(\mathbf{x}) = \frac{\langle \mathbf{x}|\mathbf{y}\rangle}{\langle \mathbf{y}|\mathbf{y}\rangle}\mathbf{y}$$

se llama **proyección ortogonal de x sobre y**.

El vector $\mathbf{x} - \prod_{\mathbf{y}}(\mathbf{x})$ es ortogonal a $\mathbf{y}$. En particular, si $\mathbf{y}$ es un **vector unitario** (de norma 1) entonces el vector $\mathbf{x} - \langle \mathbf{x}|\mathbf{y}\rangle \mathbf{y}$ es ortogonal a $\mathbf{y}$.

Ejercicios propuestos

---

1. Particulariza las definiciones y propiedades anteriores para el caso $n = 1$, es decir, para $\mathbb{R}$.

2. Prueba la desigualdad de Cauchy-Schwarz.

Sugerencia. Comprueba que la ecuación $\langle \mathbf{x} - \lambda\mathbf{y} | \mathbf{x} - \lambda\mathbf{y} \rangle = 0$, en la que λ es un número real arbitrario y $\mathbf{x}$ e $\mathbf{y}$ son vectores que se suponen fijos, es un trinomio de segundo grado en la variable λ. Ten en cuenta que dicho trinomio toma siempre valores mayores o iguales que cero lo que proporciona información sobre su discriminante.

3. Prueba la desigualdad triangular.

Sugerencia. Una estrategia para probar desigualdades entre normas euclídeas es elevar al cuadrado. La desigualdad $\|\mathbf{x}+\mathbf{y}\|_2^2 \leqslant \left(\|\mathbf{x}\|_2 + \|\mathbf{y}\|_2\right)^2$ es equivalente a la desigualdad triangular pero es fácil de probar desarrollando $\|\mathbf{x}+\mathbf{y}\|_2^2 = \langle \mathbf{x}+\mathbf{y} | \mathbf{x}+\mathbf{y} \rangle$ y usando la desigualdad de Cauchy-Schwarz.

4. **Teorema de Pitágoras**. Prueba que los vectores $\mathbf{x}$ e $\mathbf{y}$ son ortogonales si, y solo si

$$\|\mathbf{x}+\mathbf{y}\|_2^2 = \|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2$$

5. Prueba que el vector $\mathbf{x} - \prod_{\mathbf{y}}(\mathbf{x})$ es ortogonal a $\mathbf{y}$.

1.4 Definición. La *distancia euclídea* en $\mathbb{R}^n$ es la aplicación $\mathrm{d}_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_o^+$ definida por:

$$\mathrm{d}_2(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_2 \qquad (\mathbf{x},\mathbf{y} \in \mathbb{R}^n)$$

La distancia euclídea entre los vectores $\mathbf{x}$ e $\mathbf{y}$ es el número $\mathrm{d}_2(\mathbf{x},\mathbf{y})$.

Las siguientes propiedades de la distancia euclídea se deducen fácilmente de las de la norma euclídea.

Propiedades de la distancia euclídea.

1. $\mathrm{d}_2(\mathbf{x},\mathbf{y}) \geqslant 0$, y $\mathrm{d}_2(\mathbf{x},\mathbf{y}) = 0 \Longleftrightarrow \mathbf{x} = \mathbf{y}$.

2. *Simetría.* $\mathrm{d}_2(\mathbf{x},\mathbf{y}) = \mathrm{d}_2(\mathbf{y},\mathbf{x})$ para todos $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$.

3. *Homogeneidad.* $\mathrm{d}_2(\lambda\mathbf{x},\lambda\mathbf{y}) = |\lambda|\,\mathrm{d}_2(\mathbf{x},\mathbf{y})$ para todos $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ y todo $\lambda \in \mathbb{R}$.

4. *Desigualdad triangular.* $\mathrm{d}_2(\mathbf{x},\mathbf{y}) \leqslant \mathrm{d}_2(\mathbf{x},\mathbf{z}) + \mathrm{d}_2(\mathbf{z},\mathbf{y})$ para todos $\mathbf{x},\mathbf{y},\mathbf{z} \in \mathbb{R}^n$.

1.2. Espacios normados y espacios métricos

Las propiedades de la norma y de la distancia euclídeas en $\mathbb{R}^n$ se pueden abstraer dando lugar a los conceptos de espacio normado y espacio métrico.

1.5 Definición. Sea X un espacio vectorial real. Una norma sobre X es una aplicación $\|\ \| : X \to \mathbb{R}_o^+$ que verifica las siguientes propiedades:

1. $\|\mathbf{x}\| = 0 \Longleftrightarrow \mathbf{x} = \mathbf{0}$.
2. *Homogeneidad*. $\|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\|$ para todo $\lambda\in\mathbb{R}$ y todo $\mathbf{x}\in X$.
3. *Desigualdad triangular*. $\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$ para todos $\mathbf{x},\mathbf{y}\in X$.

El par ordenado $(X,\|\ \|)$ se llama un *espacio normado*.

Naturalmente, sobre un mismo espacio vectorial pueden considerarse distintas normas, cada una de ellas da lugar a un espacio normado diferente. Para tener en cuenta este hecho se dice que un espacio normado es un par ordenado $(X,\|\ \|)$ formado por un espacio vectorial real X y una norma. No obstante, con frecuencia se dice simplemente “sea X un espacio normado” y se sobreentiende que X es un espacio vectorial real en el que está definida una norma concreta.

1.6 Ejemplos. En $\mathbb{R}^n$ suelen considerarse, además de la norma euclídea, la *norma de la suma*, $\|\ \|_1$, y la *norma del máximo*, $\|\ \|_\infty$, definidas para todo $\mathbf{x}\in\mathbb{R}^n$ por:

$$\|\mathbf{x}\|_1 = \sum_{k=1}^{n} |x_k|$$
$$\|\mathbf{x}\|_\infty = \text{máx}\,\{|x_k| : 1 \leqslant k \leqslant n\}$$

En el espacio vectorial, $\mathcal{B}(A)$, de todas las funciones reales acotadas definidas en un conjunto no vacío $A\subset\mathbb{R}$, se define la *norma uniforme* dada para toda $f\in\mathcal{B}(A)$ por:

$$\|f\|_\infty = \sup\{|f(t)| : t\in A\}\,.$$

En el espacio vectorial, $\mathcal{C}([a,b])$, de todas las funciones reales continuas definidas en un intervalo $[a,b]$, se define la *norma integral de orden 1* dada para todo $f\in\mathcal{C}([a,b])$ por:

$$\|f\|_1 = \int_a^b |f(t)|\,\mathrm{d}t$$

1.7 Definición. Sea E un conjunto cualquiera no vacío. Una *distancia* en E es una aplicación $\mathrm{d} : E\times E \to \mathbb{R}_o^+$ que verifica las siguientes propiedades:

1. $\mathrm{d}(x,y) = 0 \Longleftrightarrow x = y$.
2. *Simetría*. $\mathrm{d}(x,y) = \mathrm{d}(y,x)$ para todos $x,y\in E$.
3. *Desigualdad triangular*. $\mathrm{d}(x,y) \leqslant \mathrm{d}(x,z) + \mathrm{d}(z,y)$ para todos $x,y,z\in E$.

El par ordenado (E,d) se llama un *espacio métrico*. Los elementos de un espacio métrico suelen llamarse *puntos* de dicho espacio métrico.

Naturalmente, sobre un mismo conjunto pueden considerarse distintas distancias, cada una de ellas da lugar a un espacio métrico diferente. Para tener en cuenta este hecho se dice que un espacio métrico es un par ordenado (E,d) formado por un conjunto no vacío y una distancia. No obstante, con frecuencia se dice simplemente "sea E un espacio métrico" y se sobreentiende que en E está definida una distancia concreta.

Dado un espacio normado, $(X,\|\ \|)$, la aplicación $\mathrm{d}: X\times X\to\mathbb{R}_o^+$ dada por:

$$\mathrm{d}(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| \qquad (\mathbf{x},\mathbf{y}\in X)$$

es una distancia en X que se llama *distancia asociada a la norma*.

Todo espacio normado se considera siempre como espacio métrico con la distancia asociada a su norma.

Ejercicios propuestos

6. Prueba que $\|\ \|_1$ y $\|\ \|_\infty$ son normas en $\mathbb{R}^n$ y que para todo $\mathbf{x}\in\mathbb{R}^n$ se verifican las desigualdades:

$$\|\mathbf{x}\|_\infty \leqslant \|\mathbf{x}\|_2 \leqslant \|\mathbf{x}\|_1 \leqslant n\|\mathbf{x}\|_\infty$$

7. Prueba que $\|\ \|_\infty$ es una norma en $\mathcal{B}(A)$ y que $\|\ \|_1$ es una norma en $\mathcal{C}([a,b])$..

8. Sea $n\in\mathbb{N}$ y $S(n)=\{1,2,\dots,n\}$. Prueba que los espacios vectoriales $\mathcal{B}(S(n))$ y $\mathbb{R}^n$ son isomorfos. ¿Qué relación hay entre la norma uniforme y la del máximo en dichos espacios?

9. Sea (E,d) un espacio métrico. Prueba la desigualdade:

$$|\mathrm{d}(x,y)-\mathrm{d}(y,z)| \leqslant \mathrm{d}(x,z) \qquad (x,y,z\in E) \tag{1.1}$$

Deduce que en todo espacio normado se verifica la desigualdad:

$$\big|\|\mathbf{x}\|-\|\mathbf{y}\|\big| \leqslant \|\mathbf{x}-\mathbf{y}\| \qquad (\mathbf{x},\mathbf{y}\in E) \tag{1.2}$$

10. Sea (E,d) un espacio métrico y sean $x_1,x_2,\dots,x_n$ puntos de E. Prueba que:

$$\mathrm{d}(x_1,x_n) \leqslant \sum_{j=1}^{n-1}\mathrm{d}(x_j,x_{j+1})$$

11. Prueba que la distancia asociada a una norma verifica las propiedades de la definición 1.7. Comprueba que dicha distancia es invariante por traslaciones y es homogénea.

12. Prueba que en todo espacio vectorial real se pueden definir normas y en todo conjunto se pueden definir distancias.

1.3. Topología de un espacio métrico

Sea (E,d) un espacio métrico. Dados un punto $a \in E$ y un número $r > 0$ definimos el conjunto (al que por ahora no pondremos nombre):

$$B(a,r) = \{x \in E : \mathrm{d}(x,a) < r\}$$

Observa que el conjunto así definido depende claramente de la distancia d por lo que una notación más apropiada sería $B_{\mathrm{d}}(a,r)$, pero dicha notación es incómoda y solamente se usa cuando se consideran varias distancias diferentes en un mismo contexto.

1.8 Definición. Se dice que un conjunto $A \subset E$ es *abierto* en el espacio métrico (E,d) si para cada punto $x \in A$ hay un número $r_x > 0$ tal que $B(x,r_x) \subset A$. Por convenio, el conjunto vacío, Ø, se considera abierto.

1.9 Proposición. *Sean* (E,d) *un espacio métrico,* $a \in E$ *y* $r > 0$*. Se verifica que el conjunto* $B(a,r)$ *es abierto. Dicho conjunto se llama* bola abierta *de centro* a *y radio* r*.*

1.10 Proposición (Propiedades de los conjuntos abiertos de un espacio métrico)**.** *En todo espacio métrico* (E,d) *se verifica que:*

1. *Los conjuntos E y Ø son abiertos.*
2. *La unión de cualquier colección de conjuntos abiertos es un conjunto abierto.*
3. *La intersección de una colección finita de conjuntos abiertos es un conjunto abierto.*

Las propiedades anteriores se expresan diciendo que si $\mathcal{T}$ *es la clase de todos los conjuntos abiertos de un espacio métrico* (E,d)*, entonces* $\mathcal{T}$ *es una* topología *en* E*. Se dice que dicha topología está asociada a la distancia* d*.*

1.11 Definición. Se dice que un conjunto $F \subset E$ es *cerrado* en el espacio métrico (E,d) si su complementario $E \setminus F$ es abierto en dicho espacio métrico.

1.12 Proposición (Propiedades de los conjuntos cerrados de un espacio métrico)**.** *En todo espacio métrico* (E,d) *se verifica que:*

1. *Los conjuntos E y Ø son cerrados.*
2. *La intersección de cualquier colección de conjuntos cerrados es un conjunto cerrado.*
3. *La unión de una colección finita de conjuntos cerrados es un conjunto cerrado.*

1.13 Proposición. *Sean* (E,d) *un espacio métrico,* $a \in E$ *y* $r \geqslant 0$*. Se verifica que el conjunto*

$$\overline{B}(a,r) = \{x \in E : \mathrm{d}(x,a) \leqslant r\}$$

es cerrado. Dicho conjunto se llama bola cerrada *de centro a radio* r*.*

Recuerda que todo espacio normado se considera automáticamente como espacio métrico con la distancia asociada a la norma, por lo que las definiciones y resultados anteriores tienen sentido también para espacios normados. En particular, *la topología de un espacio normado es la topología asociada a la distancia asociada a la norma del mismo*. En $\mathbb{R}$ se considera siempre la topología y la distancia asociadas al valor absoluto, es decir, $\mathbb{R}$ se considera siempre como el espacio normado $(\mathbb{R}, |\ |)$.

Ejercicios propuestos

13. Prueba que si d es una distancia en E la función $\rho : E \times E \to \mathbb{R}$ dada por:

$$\rho(x,y) = \frac{\mathrm{d}(x,y)}{1 + \mathrm{d}(x,y)} \qquad (x,y \in E)$$

también es una distancia en E.

Sugerencia. Prueba que la aplicación $t \mapsto \dfrac{t}{1+t}$ es creciente en $\mathbb{R}_o^+$.

14. Prueba que un conjunto $A \subset E$ es abierto en el espacio métrico (E,d) si, y sólo si, A es unión de bolas abiertas.

15. Describe las bolas abiertas y cerradas en los espacios normados $(\mathbb{R}, |\ |)$, $(\mathbb{R}^2, \|\ \|_2)$, $(\mathbb{R}^2, \|\ \|_1)$, $(\mathbb{R}^2, \|\ \|_\infty)$ y $(\mathcal{B}([0,1]), \|\ \|_\infty)$.

16. Da un ejemplo de una colección de conjuntos abiertos (resp. cerrados) en $(\mathbb{R}, |\ |)$ cuya intersección (resp. unión) no sea un conjunto abierto (resp. cerrado).

17. Prueba que todo intervalo abierto (resp. cerrado) en $\mathbb{R}$ es un conjunto abierto (resp. cerrado) en $(\mathbb{R}, |\ |)$.

18. Sean (E_i, d_i) $i = 1,2,\ldots,n$ espacios métricos. En el conjunto producto cartesiano de los E_i, $E = E_1 \times E_2 \times \cdots \times E_n$, definimos:

$$\rho((x_1,x_2,\ldots,x_n),(y_1,y_2,\ldots,y_n)) = \text{máx}\{\mathrm{d}_i(x_i,y_i) : 1 \leqslant i \leqslant n\}$$

a) Prueba que (E,ρ) es un espacio métrico. Se dice que (E,ρ) es el **espacio métrico producto** de los espacios métricos (E_i, d_i) $i = 1,2,\ldots,n$.

b) Describe las bolas abiertas en (E,ρ).

1.14 Definición. Sea (E,d) un espacio métrico y $A \subset E$.

Decimos que un punto $x \in E$ es **adherente** al conjunto A si toda bola abierta centrada en x tiene puntos de A. El conjunto de todos los puntos adherentes a A se llama la **adherencia** de A y se representa por $\overline{A}$.

Decimos que un punto $x \in E$ es un punto de **acumulación** del conjunto A si toda bola abierta centrada en x tiene puntos de A *distintos* de x. El conjunto de todos los puntos de acumulación de A se llama la **acumulación** de A y se representa por A'.

El conjunto de todos los puntos adherentes a A y a E/A se llama la **frontera** de A y se representa por $\mathrm{Fr}(A)$.

Decimos que un punto $x \in A$ es un punto **interior** al conjunto A si hay alguna bola abierta centrada en x contenida en A. El conjunto de todos los puntos interiores de A se llama el **interior** de A y se representa por $\overset{\circ}{A}$.

Decimos que un punto $x \in A$ es un punto **aislado** en A si hay alguna bola abierta $B(x,r)$ tal que $B(x,r) \cap A = \{x\}$.

Ejercicios propuestos

19. Sea (E,d) un espacio métrico y $A \subset E$. Prueba que:

1. La adherencia de A es el más pequeño conjunto cerrado que contiene a A. Por tanto A es cerrado si, y sólo si $A = \overline{A}$.
2. El interior de A es el más grande conjunto abierto contenido en A.
3. La frontera de A son los puntos adherentes a A que no son interiores de A.
4. La adherencia de A es la unión de A con la frontera de A.

20. Prueba que los puntos adherentes de A que no son de acumulación son puntos aislados de A.

21. Da ejemplos de conjuntos en $(\mathbb{R},|\,|)$ y en $(\mathbb{R}^2,\|\,\|_2)$ que no sean abiertos ni cerrados.

22. Prueba que el supremo y el ínfimo de un conjunto no vacío y acotado de números reales son adherentes a dicho conjunto.

23. Calcula la adherencia en $\mathbb{R}$ de los conjuntos $\mathbb{Q}$ y $\mathbb{R}/\mathbb{Q}$.

Conceptos topológicos

En todo espacio métrico conviven dos estructuras, la propia de espacio métrico, (E,d), y la de espacio topológico, $(E,\mathcal{T})$, donde $\mathcal{T}$ es la topología asociada a la distancia d. Sucede que en los espacios métricos es cómodo definir algunos conceptos usando la distancia, pero a veces dichos conceptos pueden formularse también únicamente en términos de abiertos, sin referencia ninguna a la distancia original. Los conceptos que pueden formularse únicamente en términos de abiertos se llaman *topológicos*. Es importante saber cuando un concepto es topológico porque puede ocurrir que distancias distintas sobre un mismo conjunto E den lugar a la misma topología y, por tanto, las propiedades que solamente dependen de la topología son las mismas para ambas distancias.

Todos los conceptos introducidos en la definición 1.14 dependen aparentemente de la distancia porque en todos ellos se utilizan bolas abiertas. Pero es fácil comprobar que de hecho son conceptos topológicos. Para ello basta con que compruebes que en dichas definiciones puede sustituirse la expresión *"bola abierta centrada en x"* por *"conjunto abierto que contiene a x"* sin que ello afecte para nada a los conceptos allí definidos.

1.15 Definición. Se dice que dos distancias d y ρ sobre un mismo conjunto E son *equivalentes* si ambas definen la misma topología en E.

Se dice que dos normas sobre un mismo espacio vectorial son equivalentes si sus distancias asociadas son equivalentes.

1.16 Proposición. *Sean $\|\,\|$ y $|||\,|||$ dos normas sobre un espacio vectorial X. Equivalen las afirmaciones:*

a) $\|\,\|$ y $|||\,|||$ son normas equivalentes.

b) Existen números $m > 0$, $M > 0$ verificándose que:

$$m\|x\| \leqslant |||x||| \leqslant M\|x\| \qquad (x\in X)$$

Convergencia en espacios métricos

1.17 Definición. Se dice que una sucesión $\{x_n\}$ de puntos de un espacio métrico (E,d) es **convergente** si hay un elemento $x \in E$ tal que $\{\mathrm{d}(x_n,x)\} \to 0$, es decir, para todo $\varepsilon > 0$ existe un $m_\varepsilon \in \mathbb{N}$ tal que para todo $n \geqslant m_\varepsilon$ se verifica que $\mathrm{d}(x_n,x) < \varepsilon$.

Es fácil comprobar que si hay algún elemento $x \in E$ que verifica la condición de convergencia anterior dicho elemento x es único. Tal elemento se llama *límite* de la sucesión $\{x_n\}$ y escribimos $\lim\limits_{n\to\infty}\{x_n\} = x$ o, simplemente, $\{x_n\} \to x$.

Observa que $\{x_n\} \to x$ quiere decir que para toda bola $B(x,\varepsilon)$ existe un $m_\varepsilon \in \mathbb{N}$ tal que para todo $n \geqslant m_\varepsilon$ se verifica que $x_n \in B(x,\varepsilon)$.

Ejercicios propuestos

24. Prueba que el concepto de sucesión convergente es topológico. Por tanto, distancias equivalentes tienen las mismas sucesiones convergentes.

25. Sea (E,d) un espacio métrico. Prueba la desigualdad:

$$|\mathrm{d}(x,y) - \mathrm{d}(z,u)| \leqslant \mathrm{d}(x,z) + \mathrm{d}(y,u) \qquad (x,y,z,u \in E)$$

Deduce que si $\{x_n\} \to x$ e $\{y_n\} \to y$ entonces $\{\mathrm{d}(x_n,y_n)\} \to \mathrm{d}(x,y)$.

26. Para cada $n \in \mathbb{N}$ sea $f_n : [0,1] \to \mathbb{R}$ la función dada por $f(x) = x^n$. Prueba que la sucesión $\{f_n\}$ converge a la función nula $f = 0$ en el espacio $(\mathcal{C}([0,1]), \|\ \|_1)$ pero no es convergente en $(\mathcal{C}([0,1]), \|\ \|_\infty)$.

27. Sea $\{\mathbf{x}_m\}$ una sucesión de puntos de $(\mathbb{R}^n, \|\ \|_2)$. Representaremos por $\mathbf{x}_m(k)$ la coordenada $k-sima$ del vector $\mathbf{x}_m$. Prueba que $\{\mathbf{x}_m\} \to \mathbf{x}$ si, y sólo si, $\{\mathbf{x}_m(k)\} \to \mathbf{x}(k)$ para $1 \leqslant k \leqslant n$. Esto es, **la convergencia en $(\mathbb{R}^n, \|\ \|_2)$ equivale a la convergencia por coordenadas**.

28. Sea $\{\mathbf{x}_m\}$ una sucesión de puntos de $(\mathbb{R}^n, \|\ \|_2)$. Prueba que $\{\mathbf{x}_m\} \to \mathbf{x}$ si, y sólo si, se verifican las siguientes dos condiciones:

a) $\{\langle \mathbf{x}_m | \mathbf{y} \rangle\} \to \langle \mathbf{x} | \mathbf{y} \rangle \quad \forall \mathbf{y} \in \mathbb{R}^n$.

b) $\{\|\mathbf{x}_m\|_2\} \to \|\mathbf{x}\|_2$.

1.18 Proposición. *Sean (E,d) un espacio métrico, $A \subset E$ y $x \in E$. Equivalen las siguientes afirmaciones:*

a) x es un punto adherente a A.

b) x es límite de una sucesión de puntos de A.

La condición de Cauchy para una sucesión de números reales puede formularse para sucesiones de puntos de un espacio métrico.

1.19 Definición (Sucesión de Cauchy)**.** Se dice que una sucesión $\{x_n\}$ de puntos de un espacio métrico (E,d) cumple la condición de Cauchy o que es una sucesión de Cauchy, si para todo $\varepsilon > 0$ existe un $m_\varepsilon \in \mathbb{N}$ tal que para todos $p \geqslant m_\varepsilon$, $q \geqslant m_\varepsilon$ se verifica que $\mathrm{d}(x_p, x_q) < \varepsilon$.

Se dice que un espacio métrico (E,d) es *completo* si toda sucesión de puntos de E que verifica la condición de Cauchy es convergente en dicho espacio.

Un espacio normado que es completo como espacio métrico se llama un *espacio de Banach*.

El concepto de *sucesión de Cauchy* no es topológico, es decir, dos distancias equivalentes pueden no tener las mismas sucesiones de Cauchy.

1.20 Definición. Sea (E,d) un espacio métrico y A un subconjunto no vacío de E. Si el conjunto de números reales:

$$C = \{\mathrm{d}(x,y) : x\in A, y\in A\}$$

está mayorado, se dice que A está *acotado*, en cuyo caso el número $\sup(C)$ se llama *diámetro* de A y se representa por $\mathrm{diam}(A)$. Si A no está acotado se define $\mathrm{diam}(A) = +\infty$.

Se dice que una sucesión $\{x_n\}$ está acotada si el conjunto $\{x_n : n\in\mathbb{N}\}$ está acotado.

Series en un espacio normado. Sea $(X, \|.\|)$ un espacio vectorial normado. Dada una sucesión $\{a_n\}$ de elementos de X podemos formar otra sucesión $\{\sum_{k=1}^n a_k\}_{n\in\mathbb{N}}$ que se obtiene sumando *consecutivamente* los términos de $\{a_n\}$. Dicha sucesión se representa por $\sum_{n\geqslant 1} a_n$ y se llama *serie de término general* a_n. Concretamente, $\sum_{n\geqslant 1} a_n$, es la aplicación de $\mathbb{N}$ en X dada por $n \mapsto \sum_{k=1}^n a_k$ para todo $n\in\mathbb{N}$. Las series son sucesiones por lo que es innecesario especificar lo que significa que una serie es convergente. El límite de una serie convergente $\sum_{n\geqslant 1} a_n$ se llama *suma* de la serie y se representa por $\sum_{n=1}^{\infty} a_n$.

Ejercicios propuestos

29. Prueba que toda sucesión convergente es de Cauchy.

30. Prueba que toda sucesión de Cauchy está acotada.

31. Prueba que si una sucesión de Cauchy tiene una sucesión parcial convergente entonces es convergente.

32. Sea $(X, \| \|)$ un espacio normado. Prueba que un conjunto $A \subset X$ está acotado si, y sólo si, existe un número $M > 0$ tal que $\|x\| \leqslant M$ para todo $x\in A$.

33. Sea A un subconjunto no vacío de un espacio métrico (E,d). Prueba que $\mathrm{diam}(A) = \mathrm{diam}(\overline{A})$.

34. Prueba que en un espacio normado se verifica que $\overline{B(a,r)} = \overline{B}(a,r)$ y $\mathrm{diam}(B(a,r) = 2r$. ¿Son ciertas estas igualdades para un espacio métrico cualquiera?

35. Prueba que $(\mathbb{R}^n, \| \|_2)$ es un espacio de Banach.

36. Prueba que el espacio $(\mathcal{B}(A), \| \|_\infty)$ es completo.

37. Para cada $n \in \mathbb{N}$ sea $f_n : [0,2] \to \mathbb{R}$ la función dada por:

$$f_n(x) = \begin{cases} x^n, & 0 \leqslant x \leqslant 1; \\ 1, & 1 \leqslant x \leqslant 2. \end{cases}$$

Prueba que la sucesión $\{f_n\}$ es de Cauchy en el espacio $(\mathcal{C}([0,2]), \| \|_1)$ pero no es convergente.

38. Sea $A \subset \mathbb{R}$ un conjunto no vacío de números reales. Prueba que el conjunto $\mathcal{C}(A)$ de las funciones reales continuas y acotadas en A es un conjunto cerrado en el espacio $(\mathcal{B}(A), \| \|_\infty)$.

39. Sea ℓ_∞ el espacio normado de las sucesiones acotadas de números reales con la norma uniforme, es decir, $\ell_\infty = (\mathcal{B}(\mathbb{N}), \| \|_\infty)$. Para cada $n \in \mathbb{N}$ sea $\delta_n : \mathbb{N} \to \mathbb{R}$ la sucesión definida por

$$\delta_n(k) = \begin{cases} 1, & \text{si } k = n; \\ 0, & \text{si } k \neq n. \end{cases}$$

Prueba que la sucesión $\{\delta_n\}$ no tiene ninguna sucesión parcial convergente.

40. Sea $(X, \|.\|)$ un espacio de Banach y $\sum\limits_{n\geqslant 1} a_n$ una serie de elementos de X. Prueba que si la serie $\sum\limits_{n\geqslant 1} \|a_n\|$ es convergente entonces también es convergente la serie $\sum\limits_{n\geqslant 1} a_n$.

41. Sea

$$\ell_1 = \left\{ \varphi \in \mathbb{R}^{\mathbb{N}} : \sum_{n=1}^{\infty} |\varphi(n)| < \infty \right\}.$$

Para $\varphi \in \ell_1$ se define

$$\|\varphi\|_1 = \sum_{n=1}^{\infty} |\varphi(n)|$$

a) Prueba que $(\ell_1, \|.\|_1)$ es un espacio de Banach.

b) Sea δ_q la sucesión dada por $\delta_q(q) = 1$ y $\delta_q(n) = 0$ para $n \neq q$. Dada $\varphi \in \ell_1$ prueba que

$$\varphi = \sum_{n=1}^{\infty} \varphi(n)\delta_n$$

42. Prueba que un espacio métrico (E, d) es completo si, y sólo si, para toda sucesión $\{F_n\}$ de subconjuntos cerrados no vacíos de E tal que $F_{n+1} \subset F_n$ para todo $n \in \mathbb{N}$, y $\lim\{\operatorname{diam}(F_n)\} = 0$, se verifica que $\bigcap\limits_{n\in\mathbb{N}} F_n \neq \emptyset$.

Da un ejemplo de una sucesión $\{F_n\}$ de conjuntos cerrados no vacíos de $\mathbb{R}$ verificando que $F_{n+1} \subset F_n$ para todo $n \in \mathbb{N}$ y cuya intersección es vacía.

1.21 Teorema (Bolzano – Weierstrass en $(\mathbb{R}^n, \| \|_2)$). *Toda sucesión acotada de puntos de $(\mathbb{R}^n, \| \|_2)$ tiene alguna sucesión parcial convergente en $(\mathbb{R}^n, \| \|_2)$.*

1.22 Teorema (Hausdorff). *Todas las normas en un espacio vectorial de dimensión finita son equivalentes.*

Como consecuencia de este teorema en $\mathbb{R}^n$ hay una única topología que procede de una norma, dicha topología se llama *la* topología de la norma y es la que consideraremos siempre en $\mathbb{R}^n$. En dicha topología los abiertos son uniones de bolas abiertas para alguna norma.

Puesto que el concepto de sucesión convergente es topológico, deducimos, como consecuencia del teorema de Hausdorff, que las sucesiones convergentes en $\mathbb{R}^n$ son las mismas para todas las normas.

1.23 Corolario. *En todo espacio normado de dimensión finita la convergencia de una sucesión de puntos equivale a la convergencia por coordenadas.*

Otra consecuencia del teorema anterior y de la proposición 1.16 es que las sucesiones de Cauchy y las sucesiones acotadas en un espacio normado de dimensión finita son las mismas para todas las normas.

Hemos visto en el ejercicio 35 que $(\mathbb{R}^n, \| \|_2)$ es un espacio de Banach, es decir en dicho espacio las sucesiones convergentes coinciden con las de Cauchy. Deducimos que en $\mathbb{R}^n$ con cualquier norma las sucesiones de Cauchy coinciden con las sucesiones convergentes. A partir de aquí es fácil probar los siguientes resultados.

1.24 Teorema (Complitud). *Todo espacio normado de dimensión finita es un espacio de Banach.*

1.25 Teorema (Bolzano–Weierstrass). *Toda sucesión acotada de puntos de un espacio normado de dimensión finita tiene alguna sucesión parcial convergente.*

1.26 Definición. Un subconjunto K de un espacio métrico se dice que es **compacto** si toda sucesión de puntos de K tiene alguna sucesión parcial que converge a un punto de K.

Puesto que el concepto de sucesión convergente es topológico, la compacidad es una propiedad topológica.

1.27 Proposición. *Todo conjunto compacto de un espacio métrico es cerrado y acotado.*

Como consecuencia de lo visto en el ejercicio 39, se tiene que la bola cerrada $\overline{B}(\mathbf{0}, 1)$ en el espacio ℓ_∞ no es un conjunto compacto aunque, evidentemente, dicho conjunto es cerrado y acotado.

1.28 Teorema (Caracterización de la compacidad en dimensión finita). *Los conjuntos compactos en un espacio normado de dimensión finita son los conjuntos cerrados y acotados.*

Ejercicios propuestos

43. Sea K un subconjunto compacto no vacío de $\mathbb{R}^n$. Sea $r > 0$ un número fijo y definamos $A = \bigcup_{x \in K} \overline{B}(x, r)$ donde $\overline{B}(x, r)$ indica la bola cerrada de centro x y radio r para una norma que se supone fijada en $\mathbb{R}^n$. Prueba que A es compacto.

44. Sea N un número natural y sean $\alpha_0, \alpha_1, \dots, \alpha_N$ números reales distintos. En el espacio vectorial X de las funciones polinómicas de grado menor o igual que N se define:

$$\|\| = \sum_{k=0}^{N} |(\alpha_k)|, \qquad (\in X).$$

Prueba que:

1. $\|.\|$ es una norma en X.
2. La topología que genera esta norma no depende de la elección de los reales α_k.
3. Una sucesión $\{_n\}$ en X converge uniformemente (es decir, con la norma $\|.\|_\infty$) en un intervalo $[a,b]$ si, y sólo si, existen $N+1$ números reales distintos del intervalo $[a,b]$, $\beta_0, \beta_1, \dots, \beta_N$, tales que las $N+1$ sucesiones $\{_n(\beta_k)\}_{n\in\mathbb{N}}$ $(k = 0, 1, \dots, N)$ convergen.

45. Sea $\{K_n\}$ una sucesión de subconjuntos compactos no vacíos de un espacio métrico y supongamos que para todo $n\in\mathbb{N}$ es $K_{n+1} \subset K_n$. Prueba que $\bigcap\limits_{n\in\mathbb{N}} K_n \neq \emptyset$.

1.4. Continuidad

1.29 Definición. Sean (E,d), (F,ρ) espacios métricos, $A \subset E$ y $f : A \to F$. Se dice que f es continua en un punto $\alpha\in A$ si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que para todo $x\in A$ con $\mathrm{d}(x,\alpha) < \delta$ se verifica que $\rho(f(x), f(\alpha)) < \varepsilon$.

Si f es continua en todos los puntos de un conjunto $B \subset A$ se dice que f es continua en B. Si f es continua en A se dice simplemente que f es continua.

1.30 Proposición (Caracterización secuencial de la continuidad). *Sean (E,d), (F,ρ) espacios métricos, $A \subset E$, $f : A \to F$ y $\alpha\in A$. Equivalen las siguientes afirmaciones:*

a) f es continua en α.

b) Para toda sucesión $\{x_n\} \to \alpha$ con $x_n \in A$ se verifica que $\{f(x_n)\} \to f(\alpha)$.

Como el concepto de sucesión convergente es topológico, del resultado anterior se sigue que la continuidad es una propiedad topológica.

Ejercicios propuestos

46. Enuncia apropiadamente y prueba los siguientes resultados:

a) La composición de funciones continuas es continua.

b) Cualquier restricción de una función continua es continua.

47. Prueba que toda función es continua en los puntos aislados de su conjunto de definición.

48. Prueba que en todo espacio normado la norma es una aplicación continua.

49. Prueba que todo espacio normado $(X, \|.\|)$ es *homeomorfo* a su bola abierta unidad $B(0,1)$, es decir, existe una biyección continua y con inversa continua de X sobre $B(0,1)$.

50. Prueba que la imagen de un compacto por una función continua es un compacto.

51. Prueba que todo compacto en $\mathbb{R}$ tiene máximo y mínimo.

52. **Propiedad de compacidad.** Sea $f : K \to \mathbb{R}$ una función continua definida en un conjunto compacto K de un espacio métrico. Entonces f alcanza en K un máximo y un mínimo absolutos.

53. Sean (E, d) un espacio métrico, $f_i : E \to E$, $i = 1,2,\ldots,n$ funciones continuas. Consideremos el espacio métrico producto (E^n, ρ) y definamos $F : E^n \to E^n$ por

$$F(x_1, x_2, \ldots, x_n) = (f_1(x_1), f_2(x_2), \ldots, f_n(x_n))$$

Prueba que F es continua.

54. Sea $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^n$ una función continua tal que para todos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ se verifica que $\|\mathbf{x} - \mathbf{y}\| \leqslant \|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})\|$. Prueba que la imagen de $\mathbf{F}$ es un conjunto cerrado en $\mathbb{R}^n$.

1.31 Proposición (Carácter local de la continuidad). *Sean (E, d), (F, ρ) espacios métricos, $A \subset E$, $f : A \to F$ y $\alpha \in A$. Supongamos que hay alguna bola abierta $B(\alpha, r)$ tal que la restricción $f_{|B}$ de f al conjunto $B = B(\alpha, r) \cap A$ es continua en α. Entonces f es continua en α.*

En particular, si $U \subset A$ es un conjunto abierto y $f_{|U}$ es continua, entonces f es continua en U.

1.32 Proposición (Caracterización topológica de la continuidad). *Sean (E, d) y (F, ρ) espacios métricos y $f : E \to F$. Equivalen las siguientes afirmaciones:*

a) f es continua en E.

b) La imagen inversa por f de todo abierto de F es un abierto de E.

c) La imagen inversa por f de todo cerrado de F es un cerrado de E.

1.33 Corolario. *Sea (E, d) un espacio métrico y $f : E \to \mathbb{R}$ una aplicación continua. Entonces para todo $t \in \mathbb{R}$ se verifica que:*

a) Los conjuntos $\{x \in E : f(x) < t\}$ y $\{x \in E : f(x) > t\}$ son abiertos en E.

b) Los conjuntos $\{x \in E : f(x) \leqslant t\}$, $\{x \in E : f(x) \geqslant t\}$ y $\{x \in E : f(x) = t\}$ son cerrados en E.

1.34 Corolario. *Sean d y ρ dos distancias en un conjunto no vacío E. Equivalen las afirmaciones siguientes:*

a) d y ρ son distancias equivalentes.

b) Las aplicaciones identidad entre los espacios métricos (E, d) y (E, ρ) son continuas.

c) Los espacios métricos (E, d) y (E, ρ) tienen las mismas sucesiones convergentes.

Sea (E,d) un espacio métrico, A un subconjunto no vacío de E y notemos $\mathrm{d}_{|A}$ la restricción de d a $A\times A$; es decir, $\mathrm{d}_{|A} : A\times A \to \mathbb{R}_o^+$ es la aplicación definida por:

$$\mathrm{d}_{|A}(x,y) = \mathrm{d}(x,y) \qquad \forall (x,y)\in A\times A$$

Es inmediato comprobar que $\mathrm{d}_{|A}$ es una distancia en A que se llama *distancia inducida* en A por la distancia d. Por tanto $(A,\mathrm{d}_{|A})$ es un espacio métrico. Se dice que $(A,\mathrm{d}_{|A})$ es un *subespacio métrico* de (E,d).

No suele distinguirse entre la restricción $\mathrm{d}_{|A}$ de la distancia d al conjunto A y la distancia d, y ambas se representan con la misma letra d, aunque no hay que olvidar que (A,d) y (E,d) son espacios métricos diferentes.

Naturalmente, los conceptos que se han definido antes para un espacio métrico (E,d) conservan su significado cuando nos referimos a un subespacio métrico. Por ejemplo, la expresión "A es un subespacio completo de E" quiere decir que A, considerado como subespacio métrico de (E,d), es un espacio métrico completo: toda sucesión de Cauchy de puntos de A converge a un punto de A. Es frecuente, sin embargo, no usar la palabra "subespacio" sino "subconjunto" con el convenio de que ***todo subconjunto no vacío de un espacio métrico*** (E,d) ***se considera automáticamente como subespacio métrico***.

Ejercicios propuestos

55. Da ejemplos de subconjuntos de $\mathbb{R}$ que no sean completos.

56. Prueba que todo conjunto compacto de un espacio métrico es completo.

57. Prueba que todo conjunto completo de un espacio métrico es cerrado.

58. Prueba que todo conjunto cerrado de un espacio métrico completo es completo.

59. Sea $f:\mathbb{R}\to\mathbb{R}$ una función continua e inyectiva. Definamos $\rho:\mathbb{R}\times\mathbb{R}\to\mathbb{R}_o^+$ por

$$\rho(x,y) = |f(x)-f(y)| \qquad \forall (x,y)\in\mathbb{R}\times\mathbb{R}$$

Prueba que:

a) ρ es una distancia en $\mathbb{R}$ equivalente a la usual.

b) Si la función f está mayorada o minorada entonces el espacio métrico $(\mathbb{R},\rho)$ no es completo.

60. En el intervalo $E =]0,1]$ se define la distancia:

$$\rho(x,y) = \left|\frac{1}{x}-\frac{1}{y}\right| \qquad \forall (x,y)\in E\times E$$

Prueba que (E,ρ) es un espacio métrico completo y que la distancia ρ es equivalente a la distancia usual en E.

61. Un *homeomorfismo* es una aplicación biyectiva y continua con inversa continua. Prueba que una biyección continua entre dos espacios métricos compactos es un homeomorfismo.

Como sabemos, todo espacio métrico se considera como espacio topológico con la topología asociada a la distancia, por tanto, **todo subconjunto no vacío A de un espacio métrico (E,d) también se considera como espacio topológico con la topología que tiene como subespacio métrico (A,d)**. Los abiertos en esta topología se llaman *abiertos relativos* y los cerrados se llaman *cerrados relativos* de A. Teniendo en cuenta que las bolas abiertas, $B_A(a,r)$, de un punto $\alpha\in A$ en (A,d) son de la forma:

$$B_A(a,r)=\{x\in A : d(x,a)<r\}=\{x\in E : d(x,a)<r\}\cap A=B(a,r)\cap A$$

es fácil describir los abiertos y los cerrados relativos de A.

1.35 Proposición. *Sea A un subconjunto no vacío de un espacio métrico E. Se verifica que:*

a) Los abiertos relativos de A son las intersecciones de los abiertos de E con A.

b) Los cerrados relativos de A son las intersecciones de los cerrados de E con A.

Consideremos ahora una función f definida en un subconjunto no vacío A de un espacio métrico (E,d) con valores en otro espacio métrico (F,ρ). La continuidad de f en un punto $\alpha\in A$ quiere decir que para todo $\varepsilon>0$ existe un $\delta>0$ tal que para todo $x\in A$ con $\mathrm{d}(x,\alpha)<\delta$ se verifica que $\rho(f(x),f(\alpha))<\varepsilon$. Observa que aquí solamente intervienen el conjunto A, la distancia d *trabajando en A* y el espacio (F,ρ). Es decir, para definir la continuidad de f en un punto $\alpha\in A$ podemos olvidarnos del espacio (E,d) y quedarnos con el subespacio (A,d). En particular, la continuidad de f en A quiere decir que la aplicación f del espacio métrico (A,d) en el espacio métrico (F,ρ) es continua. Como consecuencia de las proposiciones 1.32 y 1.35, deducimos el siguiente resultado.

1.36 Proposición. *Sean E y F espacios métricos, $\varnothing\neq A\subset E$ y $f:A\to F$. Equivalen las siguientes afirmaciones:*

a) f es continua en A.

b) La imagen inversa por f de todo abierto de F es un abierto relativo de A.

c) La imagen inversa por f de todo cerrado de F es un cerrado relativo de A.

Continuidad uniforme

Piensa un par de minutos antes de responder a la siguiente pregunta. Supongamos que f es una función real continua en un intervalo I. ¿Es cierto que si tomamos valores $x,y\in I$ muy próximos entre sí los correspondientes valores de la función $f(x),f(y)$ también están muy próximos entre sí?

Si tu respuesta ha sido afirmativa, como suele ser, te equivocas. Considera la función continua $f:]0,1]\to\mathbb{R}$ dada por $f(x)=1/x$. Los puntos 10^{-10} y 10^{-20} están muy próximos entre sí: $10^{-10}-10^{-20}<10^{-10}$, pero $f(10^{-10})=10^{10}$ y $f(10^{-20})=10^{20}$ están muy distantes entre sí.

No hay nada extraño en este comportamiento. A cualquier función continua cuya gráfica tenga una asíntota vertical le pasa lo mismo: hay puntos muy próximos entre sí en los que la función toma valores muy distantes entre sí.

Pero también hay funciones continuas y acotadas que se comportan de forma parecida. Considera la función continua $g :]0,1] \to \mathbb{R}$ dada por $g(x) = \operatorname{sen}(1/x)$. Es una función acotada: el mayor valor que toma es 1 y el menor valor que toma es -1, de hecho se tiene que $g(]0,1]) = [-1,1]$. Sea n un número natural. Los puntos $x_n = \frac{1}{2n\pi + \pi/2}$ e $y_n = \frac{1}{2n\pi - \pi/2}$ están, para n suficientemente grande, muy próximos entre sí; de hecho $\{x_n - y_n\} \to 0$. Los valores que toma en ellos la función $g(x_n) = 1$ y $g(y_n) = -1$ distan entre sí 2 unidades (que es la máxima distancia que puede haber entre valores tomados por esta función).

Si lo piensas un poco, te darás cuenta de que en ambos ejemplos este comportamiento se debe a que las funciones f y g "oscilan mucho" en intervalos arbitrariamente pequeños. Conviene precisar la idea de "oscilación en un intervalo".

Se define la *oscilación* de una función real f en un intervalo J contenido en el dominio de definición de f como:

$$\omega(f,J) = \begin{cases} \sup f(J) - \inf f(J), & \text{si } f(J) \text{ está acotado;} \\ +\infty, & \text{si } f(J) \text{ no está acotado.} \end{cases}$$

En otros términos: la oscilación de f en J es la longitud del intervalo más pequeño que contiene a $f(J)$.

Para la función $f(x) = 1/x$ se tiene que $\omega\big(f, [1/2n, 1/n]\big) = n$ y $\omega\big(f,]0, 1/n]\big) = +\infty$. Para la función $g(x) = \operatorname{sen}(1/x)$ tenemos que $\omega\big(g, [1/(2n\pi + \pi/2), 1/(2n\pi - \pi/2)]\big) = 2$. Estas funciones tienen una oscilación "grande" en intervalos arbitrariamente pequeños. En algunas circunstancias interesa poder controlar el tamaño de la oscilación de una función de manera que dicha oscilación sea menor que una cierta cantidad fijada, $\varepsilon > 0$, en *cualquier* intervalo de longitud menor que un cierto número $\delta > 0$. Las funciones para las que esto puede hacerse cualquiera sea $\varepsilon > 0$, se llaman *uniformemente continuas*.

Se dice que una función $f : I \to \mathbb{R}$ es *uniformemente continua* en un intervalo I si para todo $\varepsilon > 0$ es posible encontrar un $\delta > 0$ de manera que siempre que J sea un intervalo contenido en I de longitud menor que δ, se verifica que la oscilación de f en J es menor o igual que ε.

Teniendo en cuenta que $\omega(f,J) \leqslant \varepsilon \Longleftrightarrow |f(x) - f(y)| \leqslant \varepsilon$ para todos $x,y \in J$, la definición dada puede expresarse de forma equivalente como sigue.

Una función f es uniformemente continua en un intervalo I si para todo $\varepsilon > 0$ es posible encontrar un $\delta > 0$ de manera que siempre que x, y sean puntos de I con $|x-y| \leqslant \delta$, se verifica que $|f(x) - f(y| \leqslant \varepsilon$. Simbólicamente:

$$\forall \varepsilon \in \mathbb{R}^+ \ \exists \delta \in \mathbb{R}^+ : \left.\begin{array}{c} |x-y| \leqslant \delta \\ x,y \in I \end{array}\right\} \Longrightarrow \quad |f(x) - f(y)| \leqslant \varepsilon$$

Esta definición puede generalizarse fácilmente para funciones entre espacios métricos. Llegamos así a la siguiente definición.

1.37 Definición. Sean (E,d), (F,ρ) espacios métricos, A un subconjunto no vacío de E y $f : A \to F$ una función definida en A. Se dice que f es *uniformemente continua* en A si para todo $\varepsilon > 0$ es posible encontrar un $\delta > 0$ de manera que siempre que x, y sean puntos de A con $\mathrm{d}(x,y) \leqslant \delta$, se verifica que $\rho(f(x), f(y)) \leqslant \varepsilon$. Simbólicamente:

$$\forall \varepsilon \in \mathbb{R}^+ \;\; \exists \delta \in \mathbb{R}^+ \; : \left.\begin{array}{c} \mathrm{d}(x,y) \leqslant \delta \\ x,y \in A \end{array}\right\} \Longrightarrow \quad \rho(f(x), f(y)) \leqslant \varepsilon \tag{1.3}$$

1.38 Observaciones.

- El concepto de "continuidad uniforme" es un concepto global: depende del comportamiento de la función en todo un conjunto. No tiene sentido decir que una función es uniformemente continua en un punto: la continuidad uniforme no es un concepto local.

- Es muy interesante comparar las definiciones de continuidad puntual 1.29 y de continuidad uniforme 1.37. Resulta evidente que la continuidad uniforme en A implica la continuidad en todo punto de A: *toda función uniformemente continua en un conjunto es continua en dicho conjunto*.

En general, no es cierto que una función continua en un conjunto A sea uniformemente continua en A como lo prueban los ejemplos dados al principio. Pero hay una situación particular en la que dicha afirmación sí es cierta. Este es el contenido del siguiente teorema. Se trata de un resultado importante en el que pueden destacarse aportaciones de varios matemáticos. Dirichlet ya lo incluyó en sus lecciones de 1862 y en 1872 Heine dio una primera demostración del mismo. Posteriormente Weierstrass, Borel y Lebesgue generalizaron el resultado inicial.

1.39 Teorema (Heine). *Toda función continua en un conjunto compacto K de un espacio métrico con valores en un espacio métrico es uniformemente continua en K.*

Por los resultados vistos hasta ahora, queda claro que la compacidad es una propiedad que tiene importantes consecuencias de muy variado tipo. El siguiente resultado nos dice que dicha propiedad debe usarse con cuidado en espacios normados de dimensión infinita.

Se dice que un espacio normado es *localmente compacto* si las bolas cerradas son compactos. Teniendo en cuenta que en un espacio normado dos bolas cerradas cualesquiera de radios positivos son homeomorfas, resulta que un espacio normado es localmente compacto si, y sólo si, hay alguna bola cerrada de radio positivo que sea compacta.

1.40 Teorema (Riesz, 1918). *Un espacio normado es de dimensión finita si, y sólo si, es localmente compacto.*

Ejercicios propuestos

62. Justifica con un ejemplo que la imagen de una sucesión de Cauchy por una función continua puede no ser una sucesión de Cauchy.

63. Prueba que la imagen de una sucesión de Cauchy por una función uniformemente continua es una sucesión de Cauchy.

64. Sea A un conjunto no cerrado de un espacio métrico (E,d). Prueba que existe una aplicación continua de A en $\mathbb{R}$ que no es uniformemente continua.

65. Sea A un conjunto no vacío de un espacio métrico (E,d) y $f : A \to F$ una aplicación de A en un espacio métrico completo (F,ρ). Prueba que si f es uniformemente continua en A entonces existe una extensión única de f a $\overline{A}$ que es uniformemente continua.

66. Prueba que un espacio normado es localmente compacto si, y sólo si, hay alguna bola cerrada de radio positivo que sea compacta.

67. Sea ℓ_∞ el espacio normado de las sucesiones acotadas de números reales con la norma uniforme, es decir, $\ell_\infty = (\mathcal{B}(\mathbb{N}), \| \ \|_\infty)$. Supongamos que $A \subset \overline{B}(\mathbf{0},1)$ es tal que $\overline{B}(\mathbf{0},1) \subset \bigcup\limits_{x\in A} B(x,1/2)$. Prueba que A es un conjunto infinito.

Algunas clases de funciones continuas

1.41 Definición. Sean (E,d), (F,ρ) espacios métricos, A un subconjunto no vacío de E y $f : A \to F$. Se dice que f es *lipchiciana* si existe $M > 0$ tal que

$$\rho(f(x), f(y)) \leqslant M\mathrm{d}(x,y) \qquad \forall x,y\in A.$$

Si $M < 1$ se dice que f es una *aplicación contractiva*. Cualquier $M > 0$ que satisfaga la desigualdad anterior se llama una *constante de Lipschitz* para f.

Es claro que las funciones lipchicianas son uniformemente continuas.

1.42 Proposición. *Sean $(E,\| \ \|)$, $(F,|||\ |||)$ espacios normados y $T : E \to F$ una aplicación lineal. Equivalen las siguientes afirmaciones:*

a) T *es continua.*

b) T *es continua en* $\mathbf{0}$.

c) Existe $M > 0$ *tal que* $|||T(\mathbf{x})||| \leqslant M\|\mathbf{x}\|$ *para todo* $\mathbf{x}\in E$.

d) T *es lipchiciana.*

Ejercicios propuestos

---

68. Prueba que la aplicación $x \mapsto \|x\|$ es lipchiciana con constante de Lipschitz 1.

69. Sea (E,d) un espacio métrico y $\varnothing \neq A \subset E$. Sea $\mathrm{dist}(\,.\,,A) : E \to \mathbb{R}$ la aplicación definida por:

$$\mathrm{dist}(x,A) = \operatorname{ínf}\{\mathrm{d}(x,a) : a\in A\} \qquad \forall x\in E$$

1. Prueba que dicha aplicación es lipchiciana con constante de Lipschitz 1.

2. Prueba que $\overset{\circ}{A} = \{x \in E : \operatorname{dist}(x, E \backslash A) > 0\}$ y $\overline{A} = \{x \in E : \operatorname{dist}(x, A) = 0\}$.
3. Prueba que en un espacio métrico todo conjunto abierto se puede expresar como unión numerable de conjuntos cerrados y cada conjunto cerrado puede expresarse como intersección numerable de abiertos.
4. Sean F y H subconjuntos de E no vacíos cerrados y disjuntos. Prueba que existe una aplicación continua $f : E \to \mathbb{R}$ verificando que:

$$0 \leqslant f(x) \leqslant 1 \ \ \forall x \in E, \ f(F) = \{0\}, \ f(H) = \{1\}$$

Deduce que existen abiertos U, V de E tales que $F \subset U$, $H \subset V$ y $U \cap V = \emptyset$.

70. Sea $A \subset \mathbb{R}^n$ un conjunto abierto no vacío y distinto de $\mathbb{R}^n$. Para cada $p \in \mathbb{N}$ se define:

$$K_p = \left\{ \mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \leqslant p, \operatorname{dist}(\mathbf{x}, \mathbb{R}^n \backslash A) \geqslant \frac{1}{p} \right\}$$

Prueba que:

a) K_p es compacto y $K_p \subset \overset{\circ}{K}_{p+1}$.

b) $\displaystyle\bigcup_{p=1}^{\infty} K_p = A$.

c) Si $K \subset A$ es compacto, existe un $p \in \mathbb{N}$ tal que $K \subset K_p$.

71. Sean $F \subset \mathbb{R}^n$ un conjunto cerrado y $K \subset \mathbb{R}^n$ un conjunto compacto no vacíos y disjuntos $F \cap K = \emptyset$. Sea $\| \ \|$ una norma en $\mathbb{R}^n$. Prueba que hay puntos $\mathbf{a} \in F$ y $\mathbf{b} \in K$ tales que

$$\|\mathbf{a} - \mathbf{b}\| = \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{x} \in F, \ \mathbf{y} \in K\}$$

72. Sean F un subespacio finito dimensional propio de un espacio normado $(E, \| \ \|)$ y $a \in E$. Prueba que existe $b \in F$ tal que:

$$\|a - b\| = \operatorname{dist}(a, F)$$

Deduce que hay un vector $\mathbf{x}$ de norma 1 que verifica que $\operatorname{dist}(\mathbf{x}, F) = 1$.

Prueba que dados $n \in \mathbb{N}$ y una función $f \in \mathcal{C}([0,1])$, existe una función polinómica h de grado menor o igual que n tal que:

$$\int_0^1 |f(x) - h(x)| \, dx \leqslant \int_0^1 |f(x) - p(x)| \, dx$$

cualquiera sea la función polinómica p de grado menor o igual que n.

73. Sea ℓ_∞ el espacio normado de las sucesiones acotadas de números reales con la norma uniforme, es decir, $\ell_\infty = (\mathcal{B}(\mathbb{N}), \| \ \|_\infty)$.

a) Prueba que el subespacio de las sucesiones convergentes es cerrado en ℓ_∞.

b) Prueba que la aplicación $T : \to \mathbb{R}$ dada por $T(\varphi) = \lim\{\varphi(n)\}$ es continua.

c) Prueba que el subespacio $_0$ de las sucesiones convergentes a cero es cerrado en ℓ_∞.

1.43 Proposición. *Toda aplicación lineal de un espacio normado de dimensión finita en otro espacio normado es continua.*

Reciben el nombre de *campos escalares* las funciones definidas en subconjuntos de $\mathbb{R}^n$ que toman valores en $\mathbb{R}$. Un campo escalar es, por tanto, una función real que depende de n variables.

Un campo escalar de una variables es, simplemente, una función real de variable real; un campo escalar de dos variables es una función definida en un subconjunto del plano que toma valores reales; un campo escalar de tres variables es una función definida en un subconjunto del espacio que toma valores reales.

Los campos escalares de una o dos variables se pueden visualizar por medio de sus representaciones gráficas que son, respectivamente, curvas en el plano o superficies en el espacio. No es posible visualizar campos escalares de tres o más variables porque sus gráficas están en espacios de dimensión mayor o igual que 4.

Naturalmente, los campos escalares se pueden sumar, multiplicar y dividir (donde no se anule el denominador) al igual que lo hacemos con las funciones reales. Es fácil probar, usando la caracterización secuencial de la continuidad, que dichas operaciones conservan la continuidad.

Ejemplos de campos escalares continuos lo proporcionan las *proyecciones* sobre los ejes coordenados. Representaremos por π_k la aplicación $\pi_k : \mathbb{R}^n \to \mathbb{R}$ que a cada vector $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ hace corresponder su coordenada k-ésima en la base canónica:

$$\pi_k(\mathbf{x}) = \pi_k(x_1, x_2, \dots, x_n) = x_k$$

Dicha aplicación se llama *proyección coordenada k-ésima*. Es una aplicación lineal.

Los campos escalares más sencillos son las *funciones polinómicas de varias variables*. Dichas funciones se obtienen como sumas de productos de proyecciones sobre los ejes coordenados y son, por tanto, continuas.

Para $n = 3$ las proyecciones sobre los ejes coordenados son

$$\pi_1(x,y,z) = x, \quad \pi_2(x,y,z) = y, \quad \pi_3(x,y,z) = z$$

Un producto de estas funciones es una función de la forma $f(x,y,z) = x^m y^p z^q$ donde m, p, q son números naturales o nulos. Las funciones polinómicas en tres variables son combinaciones lineales de este tipo de funciones.

Las *funciones racionales* de n variables son las funciones de la forma

$$R(x_1, x_2, \dots, x_n) = \frac{P(x_1, x_2, \dots, x_n)}{Q(x_1, x_2, \dots, x_n)}$$

Donde $P(x_1, x_2, \dots, x_n)$ y $Q(x_1, x_2, \dots, x_n)$ son funciones polinómicas de n variables. El *dominio natural* de definición de una función racional es el conjunto de puntos donde no se anula el denominador $\Omega = \{\mathbf{x} \in \mathbb{R}^n : Q(\mathbf{x}) \neq 0\}$. *Las funciones racionales son continuas en su dominio natural de definición.*

Componiendo funciones continuas reales de una variable con funciones polinómicas y racionales en varias variables obtenemos muchísimos ejemplos de campos escalares continuos.

Aquí tienes unos pocos.

$$f(x,y) = \operatorname{sen}(xy), \quad f(x,y) = \log(1+x^2+y^2), \quad f(x,y,z) = \frac{1+xy^2+xz^2}{2+\operatorname{arc\,tg}(xyz)}$$

1.5. Límite funcional

El concepto de límite para funciones reales de una variable real, que ya conoces de primer curso, se generaliza con facilidad para funciones entre espacios métricos.

1.44 Definición. Sean (E,d), (F,ρ) espacios métricos, $A \subset E$, $f : A \to F$ y $\alpha \in A'$. Se dice que f tiene límite en α si hay un elemento $\ell \in F$ con la propiedad de que para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que para todo $x \in A$ con $0 < \mathrm{d}(x,\alpha) < \delta$ se verifica que $\rho(f(x),\ell) < \varepsilon$.

Es fácil comprobar que solamente puede haber un elemento $\ell \in F$ que cumpla la condición anterior. Se dice que ℓ es el límite de f en el punto α y escribimos $\lim\limits_{x\to\alpha} f(x) = \ell$.

Con el abuso usual de notación, podemos escribir:

$$\lim_{x\to\alpha} f(x) = \ell \Longleftrightarrow \left[\forall \varepsilon \in \mathbb{R}^+ \ \exists \delta \in \mathbb{R}^+ : \left. \begin{array}{c} 0 < \mathrm{d}(x,\alpha) < \delta \\ x \in A \end{array} \right\} \Longrightarrow \quad \rho(f(x),\ell) < \varepsilon \right] \tag{1.4}$$

Una notación mejor para el límite sería $\lim\limits_{\substack{x\to\alpha \\ x\in A}} f(x) = \ell$ para indicar que solamente se consideran valores de la variable $x \in A$. Esta precisión es importante porque la existencia y el valor de un límite pueden depender del conjunto A donde se considera definida la función. No obstante, no usaremos esta notación porque es algo incómoda pero no debes olvidar que la variable x siempre toma valores en el conjunto donde se supone que está definida la función.

Observa que en la definición se supone que $\alpha \in A'$ es un punto de acumulación de A. Esto asegura que para todo $\delta > 0$, hay puntos $x \in A$ que cumplen la condición $0 < \mathrm{d}(x,\alpha) < \delta$. Por otra parte, puede ocurrir que $\alpha \notin A$, en cuyo caso f puede no estar definida en α.

La condición $0 < \mathrm{d}(x,\alpha) < \delta$ es una forma de escribir $x \neq \alpha$ y $\mathrm{d}(x,\alpha) < \delta$. Es decir, en la condición de límite (1.4) no interviene para nada, en el caso de que f esté definida en α, el valor que f pueda tener en α.

Podemos escribir (1.4) usando bolas abiertas:

$$\lim_{x\to\alpha} f(x) = \ell \Longleftrightarrow \left[\forall B_\rho(\ell,\varepsilon) \ \exists B_\mathrm{d}(\alpha,\delta) : \left. \begin{array}{c} x \in B_\mathrm{d}(\alpha,\delta) \cap A \\ x \neq \alpha \end{array} \right\} \Longrightarrow \quad f(x) \in B_\rho(\ell,\varepsilon) \right]$$

Puedes comprobar que si en lo anterior sustituyes “bolas abiertas centradas en α o en ℓ” por “abiertos que contengan a α o a ℓ” la condición que se obtiene es equivalente. Es decir, el concepto de límite de una función en un punto es un concepto topológico.

La siguiente caracterización del límite funcional mediante sucesiones es una generalización directa de la que se estudia en primer curso.

1.45 Proposición (Caracterización secuencial del límite funcional)**.** *Sean* (E,d), (F,ρ) *espacios métricos,* $A \subset E$, $f : A \to F$ *y* $\alpha \in A'$. *Equivalen las siguientes afirmaciones:*

a) $\lim\limits_{x\to\alpha} f(x) = \ell$.

b) Para toda sucesión $\{x_n\}$ *de puntos de* A *distintos de* α, $x_n \in A \setminus \{\alpha\}$, *que converge a* α, $\{x_n\} \xrightarrow{\mathrm{d}} \alpha$, *se verifica que* $\{f(x_n)\} \xrightarrow{\rho} \ell$.

La relación entre límite funcional y continuidad es la que ya conoces de primer curso.

1.46 Proposición. *Sean* (E,d), (F,ρ) *espacios métricos,* $A \subset E$, $f : A \to F$ *y* $\alpha \in A \cap A'$. *Equivalen las siguientes afirmaciones:*

a) f *es continua en* α.

b) $\lim\limits_{x\to\alpha} f(x) = f(\alpha)$.

1.47 Proposición (Carácter local del límite funcional). *Sean* (E,d), (F,ρ) *espacios métricos,* $A \subset E$, $f : A \to F$ *y* $\alpha \in A'$. *Supongamos que hay alguna bola abierta* $B(\alpha,r)$ *tal que la restricción* $f_{|B}$ *de* f *al conjunto* $B = B(\alpha,r) \cap A$ *tiene límite en* α *igual a* ℓ. *Entonces* f *también tiene límite en* α *igual a* ℓ.

1.6. Continuidad y límites de campos escalares y vectoriales

Las funciones que más nos interesan en este curso son los campos escalares y las funciones vectoriales. Un *campo escalar* de n variables es una función definida en un conjunto no vacío $A \subset \mathbb{R}^n$ que toma valores reales. Una *función vectorial* de n variables es una función definida en un conjunto no vacío $A \subset \mathbb{R}^n$ que toma valores vectoriales en un espacio $\mathbb{R}^m$ con $m \geqslant 2$. Por comodidad en la expresión, llamaremos a veces *campos vectoriales* a las funciones vectoriales aunque dicho nombre suele reservarse para aquellas funciones vectoriales definidas en un subconjunto de un espacio vectorial y que toman valores en dicho espacio vectorial.

1.48 Definición (Componentes de una función vectorial)**.** Sea $A \subset \mathbb{R}^n$ y $\mathbf{F} : A \to \mathbb{R}^m$ una función vectorial. Para cada $\mathbf{x} \in A$ se tiene que $\mathbf{F}(\mathbf{x})$ es un vector de $\mathbb{R}^m$. Para $1 \leqslant k \leqslant m$ definamos $f_k = \pi_k \circ \mathbf{F} : A \to \mathbb{R}$ donde π_k es la proyección coordenada k-ésima. Las funciones f_k así definidas son campos escalares que se llaman *componentes* de $\mathbf{F}$. Para cada $\mathbf{x} \in A$ se tiene que

$$\mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_m(\mathbf{x})) = \sum_{k=1}^{m} f_k(\mathbf{x})\mathbf{u}_k$$

donde $\{\mathbf{u}_k : 1 \leqslant k \leqslant m\}$ es la base canónica de $\mathbb{R}^m$.

Escribiremos $\mathbf{F} = (f_1, f_2, \dots, f_m)$ para indicar que $\mathbf{F}$ es un campo vectorial cuyos campos escalares componentes son los f_k $(1 \leqslant k \leqslant m)$.

Ejercicios propuestos

74. Sea $\mathbf{F} = (f_1, f_2, \dots, f_m) : A \to \mathbb{R}^m$ un campo vectorial de n variables. Prueba que:

1. $\mathbf{F}$ es continuo en un punto $\alpha \in A$ si, y sólo si, todos los campos escalares f_k son continuos en α.

2. $\lim\limits_{\mathbf{x}\to\alpha} \mathbf{F}(\mathbf{x}) = \beta \Longleftrightarrow \lim\limits_{\mathbf{x}\to\alpha} f_k(\mathbf{x}) = \beta_k \quad (1 \leqslant k \leqslant m)$.

Teniendo en cuenta este ejercicio, para estudiar límites o continuidad de campos vectoriales es suficiente hacerlo para campos escalares.

Campos escalares definidos en un mismo conjunto se pueden sumar, multiplicar y dividir. Así mismo, campos vectoriales con igual número de funciones componentes pueden sumarse, y también podemos multiplicar un campo vectorial por un campo escalar. Todo ello en el supuesto de que los campos considerados estén definidos en un mismo conjunto. También pueden componerse campos vectoriales entre sí o con campos escalares. Todas estas operaciones se comportan respecto a la continuidad y al límite funcional en la forma que cabe esperar.

Tiene interés definir el concepto de campo escalar divergente en un punto que generaliza lo que ya conoces de primer curso.

1.49 Definición. Sea $f : A \to \mathbb{R}$ donde $A \subset \mathbb{R}^n$ y $\alpha \in A'$. Se dice que f es positivamente divergente en α si para todo $M > 0$ existe $\delta > 0$ tal que para todo $\mathbf{x} \in A$ con $0 < \|\mathbf{x} - \alpha\| < \delta$ se verifica que $f(\mathbf{x}) > M$. Simbólicamente escribimos $\lim\limits_{\mathbf{x}\to\alpha} f(\mathbf{x}) = +\infty$.

Se dice que f es negativamente divergente en α si $-f$ es positivamente divergente en α, en cuyo caso escribimos $\lim\limits_{\mathbf{x}\to\alpha} f(\mathbf{x}) = -\infty$.

Se dice que f es divergente en α si $|f|$ es positivamente divergente en α, en cuyo caso escribimos $\lim\limits_{\mathbf{x}\to\alpha} f(\mathbf{x}) = \infty$.

En la definición anterior $\| \ \|$ es cualquier norma en $\mathbb{R}^n$, pues sabemos que todas ellas son topológicamente equivalentes por lo que podemos usar la que queramos para estudiar propiedades topológicas como son límites o continuidad.

En lo que sigue vamos a considerar el caso más sencillo de un campo escalar de dos variables: $f : A \to \mathbb{R}$ donde $\emptyset \neq A \subset \mathbb{R}^2$. Supongamos que $\alpha \in A'$ y queremos estudiar el límite $\lim\limits_{\mathbf{x}\to\alpha} f(\mathbf{x})$. En $\mathbb{R}^2$ tenemos que $\mathbf{x} = (x,y)$, $\alpha = (a,b)$, por lo que el límite anterior se escribe usualmente en la forma $\lim\limits_{(x,y)\to(a,b)} f(x,y)$ (y no debes olvidar el paréntesis que indica vector: (x,y) y (a,b), porque no tiene sentido escribir $\lim\limits_{x,y\to a,b} f(x,y)$). Siempre que te encuentres en esta situación el resultado de intentar evaluar la función f en (a,b) va a ser una indeterminación, es decir, una expresión del tipo $0/0, \infty/\infty$ o 1^∞ y otras que ya conoces de primer curso. Veamos un ejemplo.

1.50 Ejemplo. Sea $f(x,y) = \dfrac{\operatorname{sen}(x^2+y^2)}{x^2+y^2}$. Esta función es un campo escalar definido (y continuo ¿por qué?) en $\mathbb{R}^2 \setminus \{(0,0)\}$. Si tratas de evaluar $f(x,y)$ en $(0,0)$ obtienes la indeterminación $0/0$. En este caso, como en otros parecidos, hay una forma sencilla de proceder. Basta darse cuenta de que nuestra función es de la forma $\dfrac{\operatorname{sen}(t)}{t}$ donde $t = x^2+y^2$. Dicho de otra forma, el campo escalar f es la composición de la función $g(t) = \dfrac{\operatorname{sen}(t)}{t}$ con la función polinómica $h(x,y) = x^2+y^2$. Llegado aquí ya debes intuir lo que vale el límite $\lim\limits_{(x,y)\to(0,0)} f(x,y)$: cuando

(x,y) está muy próximo a $(0,0)$ el número $t = x^2 + y^2$ está muy próximo también a 0 por lo que $\dfrac{\operatorname{sen}(t)}{t}$ está muy próximo a 1.

Justifiquemos que $\displaystyle\lim_{(x,y)\to(0,0)} f(x,y) = 1$. Para ello definamos $g(0) = 1$, con lo que la función g resulta continua en 0 (porque ya conoces el famoso límite $\displaystyle\lim_{t\to 0} \frac{\operatorname{sen} t}{t} = 1$). Ahora, está claro que definiendo $f(0,0) = 1$ la función f así extendida es continua en $(0,0)$ porque es composición de funciones continuas: la función $h(x,y) = x^2 + y^2$ que es continua y $h(0,0) = 0$ y la función g que es continua y $g(0) = 1$. Luego $\displaystyle\lim_{(x,y)\to(0,0)} f(x,y) = 1$.

Otra forma de proceder, más cómoda en este caso, es trabajar con sucesiones. Para ello sea $\{(x_n, y_n)\} \to (0,0)$ con $(x_n, y_n) \neq (0,0)$. Pongamos $z_n = x_n^2 + y_n^2$. Claramente, $\{z_n\} \to 0$ y $z_n \neq 0$ por lo que:

$$h(x_n, y_n) = \frac{\operatorname{sen}(x_n^2 + y_n^2)}{x_n^2 + y_n^2} = \frac{\operatorname{sen}(z_n)}{z_n} \to 1.$$

Como esto es cierto para toda sucesión en las condiciones anteriores, deducimos, por la proposición 1.45, que $\displaystyle\lim_{(x,y)\to(0,0)} f(x,y) = 1$. ♦

Ejercicios propuestos

75. Sean f y g funciones de una variable real tales que $\displaystyle\lim_{x\to a} f(x) = \alpha$ y $\displaystyle\lim_{x\to b} g(x) = \beta$. Prueba que:

$$\lim_{(x,y)\to(a,b)} f(x)g(y) = \alpha\beta, \qquad \lim_{(x,y)\to(a,b)} \big(f(x) + g(y)\big) = \alpha + \beta.$$

76. Estudia la continuidad de los siguientes campos escalares y el límite en el punto α indicado en cada caso:

a) $f(x,y) = (1 + xy)^{\frac{1}{xy}}$ definido en $\mathbb{R}^+ \times \mathbb{R}^+$, $\alpha = (0,0)$.

b) $f(x,y) = \dfrac{\log(1 + x^2 + y^2)}{x^2 + y^2}$ definido en $\mathbb{R}^2 \setminus \{(0,0)\}$, $\alpha = (0,0)$.

c) $f(x,y) = \dfrac{e^{x^2}\, e^{y^2} - 1}{x^2 + y^2}$ definido en $\mathbb{R}^2 \setminus \{(0,0)\}$, $\alpha = (0,0)$.

d) $f(x,y) = \dfrac{\cos(x^2 + y^2) - 1}{(x^2 + y^2)^2}$ definido en $\mathbb{R}^2 \setminus \{(0,0)\}$, $\alpha = (0,0)$.

e) $f(x,y) = \dfrac{\operatorname{sen} x \operatorname{arc\,tg} y}{xy}$ definido en $A = \big\{(x,y) \in \mathbb{R}^2 : xy \neq 0\big\}$, $\alpha = (0,0)$.

f) $f(x,y) = \dfrac{x^2 + y^2}{\sqrt{x^2 + y^2 + 1} - 1}$ definido en $\mathbb{R}^2 \setminus \{(0,0)\}$, $\alpha = (0,0)$.

g) $f(x,y) = \dfrac{1 - \cos(x + y)}{x^2 + y^2 + 2xy}$ definido en $\mathbb{R}^+ \times \mathbb{R}^+$, $\alpha = (0,0)$.

77. Sea $f : \mathbb{R}^* \times \mathbb{R}^* \to \mathbb{R}$ la función dada por:

$$f(x,y) = (x+y)\operatorname{sen}\frac{\pi}{x}\operatorname{sen}\frac{\pi}{y}$$

Estudia si dicha función tiene límite en los puntos $(0,0),(0,1)$ y $(0,\pi)$. ¿Es f uniformemente continua?

78. Sea $f : I \to \mathbb{R}$ una función de clase $\mathcal{C}^1$ en un intervalo I. Se define la función $h : I \times I \to \mathbb{R}$ por:

$$h(x,y) = \begin{cases} \dfrac{f(x)-f(y)}{x-y}, & x \neq y; \\ f'(x), & x = y. \end{cases}$$

Estudia la continuidad de h.

79. Sea $f : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ la función dada por:

$$f(x,y) = \begin{cases} \dfrac{e^x - e^y}{\log x - \log y}, & x \neq y; \\ x e^x, & x = y. \end{cases}$$

Estudia la continuidad de f y la existencia de límite de f en $(0,0)$.

Para no tener que indicar en cada caso el conjunto donde está definido un campo escalar, te recuerdo que si un campo escalar viene dado por medio de funciones elementales su *dominio natural* de definición es el conjunto más grande en el que las operaciones que definen a dicho campo tienen sentido para valores reales de las variables.

Límites direccionales y límite de un campo escalar a lo largo de una curva

En la recta real solamente podemos aproximarnos a un punto $\alpha \in \mathbb{R}$ por dos direcciones: por valores mayores y por valores menores que α. Como sabes esto da lugar a los límites laterales en α. Muy diferente situación tenemos si $\alpha \in \mathbb{R}^2$ pues podemos acercarnos a α por infinitas direcciones, a saber, todas las de la forma $\alpha + t\mathbf{u}$ donde $\mathbf{u}$ es un vector unitario en $\mathbb{R}^2$ y $t \in \mathbb{R}$. Naturalmente, si existe el límite $\displaystyle\lim_{\substack{\mathbf{x}\to\alpha \\ \mathbf{x}\in A}} f(\mathbf{x}) = \ell$ entonces para todo conjunto $B \subset A$ con $\alpha \in B'$ también se verifica que $\displaystyle\lim_{\substack{\mathbf{x}\to\alpha \\ \mathbf{x}\in B}} f(\mathbf{x}) = \ell$. En particular, si $B = \{\alpha + t\mathbf{u} : t \in \mathbb{R}\} \cap A$, entonces se tiene que:

$$\lim_{\substack{\mathbf{x}\to\alpha \\ \mathbf{x}\in B}} f(\mathbf{x}) = \lim_{t\to 0} f(\alpha + t\mathbf{u})$$

Este límite se llama *límite de* f en α *según la dirección dada por el vector* $\mathbf{u}$. Según acabamos de decir, si existe $\displaystyle\lim_{\substack{\mathbf{x}\to\alpha \\ \mathbf{x}\in A}} f(\mathbf{x}) = \ell$ entonces también debe ser $\displaystyle\lim_{t\to 0} f(\alpha + t\mathbf{u}) = \ell$ para todo vector unitario $\mathbf{u}$.

Podemos generalizar considerando en vez de una recta que pasa por α otro tipo de curvas que pasan por α, por ejemplo, parábolas $\gamma(t) = (t, \lambda t^2)$ o $\gamma(t) = (\lambda t^2, t)$. En general, si $\gamma(t) =$

$(x(t),y(t))$ es una curva que pasa por α, $\gamma(0)=\alpha$, el límite $\lim\limits_{t\to 0} f(\gamma(t))$ se llama *límite de* f en α *a lo largo de la curva* γ. Si existe el límite de f en α entonces existe el límite de f a lo largo de toda curva que pasa por α y todos ellos coinciden con el valor del límite.

Estos resultados pueden ser útiles para probar que un cierto límite no existe: para ello basta encontrar que no existe el límite a lo largo de una determinada curva o dirección, o bien probar que los límites a lo largo de dos curvas o direcciones distintas son diferentes.

Puesto que un vector unitario en $\mathbb{R}^2$ con la norma euclídea es de la forma $\mathbf{u}=(\cos\vartheta,\operatorname{sen}\vartheta)$, y $\alpha=(a,b)$, suelen escribirse los límites direccionales en la forma:

$$\lim_{t\to 0} f(\alpha+t\mathbf{u})=\lim_{t\to 0} f(a+t\cos\vartheta,b+t\operatorname{sen}\vartheta)$$

Es frecuente escribir una recta que pasa por (a,b) en la forma $\gamma(t)=(a+t,b+\lambda t)$ donde $\lambda\in\mathbb{R}$ es un parámetro fijo – la pendiente de la recta – en cuyo caso los límites direccionales en (a,b) vienen dados por:

$$\lim_{t\to 0} f(a+t,b+\lambda t)$$

Cuando $(a,b)=(0,0)$ las rectas por el origen suelen representarse en la forma $\gamma(t)=(t,\lambda t)$ o $\gamma(t)=(\lambda t,t)$ donde λ es un parámetro fijo.

Con frecuencia es más cómodo estudiar los límites en $(0,0)$. Esto puede hacerse siempre mediante una traslación pues:

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

Hay que observar que la existencia de todos los límites direccionales en un punto siendo, además, todos ellos iguales, no garantiza la existencia del límite en dicho punto.

1.51 Ejemplo. Sea $f(x,y)=\dfrac{y}{x}\operatorname{sen}(x^2+y^2)$ definido en $A=\left\{(x,y)\in\mathbb{R}^2, x\neq 0\right\}$. Estudiemos el límite de f en $\alpha=(0,0)$.

Consideremos rectas por el origen $\gamma(t)=(t,\lambda t)$. Tenemos $f(\gamma(t))=\lambda\operatorname{sen}((1+\lambda^2)t^2)$, por lo que $\lim\limits_{t\to 0} f(t,\lambda t)=0$. Por tanto, los límites según todas las direcciones que pasan por el origen (excepto la dirección del eje de ordenadas $\gamma(t)=(0,t)$ en la que no está definida f) existen y son todos iguales a 0.

Observa que si $\gamma(t)=(x(t),y(t))$ es una curva con $\gamma(0,0)=0$, se tiene que:

$$f(\gamma(t))=\frac{y(t)}{x(t)}\operatorname{sen}(x(t)^2+y(t)^2)$$

Como las curvas son funciones continuas $\lim\limits_{t\to 0}\operatorname{sen}(x(t)^2+y(t)^2)=\operatorname{sen}(0)=0$. Por lo que para toda curva en estas condiciones que verifique que el cociente $\dfrac{y(t)}{x(t)}$ esté acotado se tendrá que $\lim\limits_{t\to 0} f(\gamma(t))=0$.

A la vista de lo anterior, podemos buscar una curva $\gamma(t)=(x(t),y(t))$ con $\gamma(0,0)=0$ tal que $\lim\limits_{t\to 0}\dfrac{y(t)}{x(t)}=\infty$ y que $\operatorname{sen}(x(t)^2+y(t)^2)$ tienda a 0 más lentamente. Como sabemos que $\operatorname{sen}(t)\sim t$

para $t \to 0$, la idea es tomar una curva $\gamma(t) = (x(t), y(t))$ que verifique:

$$\left.\begin{array}{r} \dfrac{y(t)}{x(t)} = \dfrac{1}{t^2} \\ \\ x(t)^2 + y(t)^2 = |t| \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} x(t) = t^2\sqrt{\dfrac{|t|}{1+t^4}} \\ y(t) = \sqrt{\dfrac{|t|}{1+t^4}} \end{array}\right.$$

Para dicha curva (que pasa por $(0,0)$) tenemos que:

$$f(\gamma(t)) = f(x(t), y(t)) = \frac{1}{t^2}\operatorname{sen}(|t|) = \frac{1}{|t|}\frac{\operatorname{sen}(|t|)}{|t|} \Longrightarrow \lim_{t\to 0} f(\gamma(t)) = +\infty$$

El campo escalar f diverge positivamente a lo largo de dicha curva, por lo que dicho campo no tiene límite en $(0,0)$. ♦

Límites iterados

Para estudiar la existencia del límite de un campo escalar de dos variables suelen usarse los *límites iterados*. Consideremos un campo escalar $f : A \to \mathbb{R}$ donde A es de la forma $A = I \times J \setminus \{(0,0)\}$ siendo I, J intervalos abiertos en $\mathbb{R}$ que contienen al 0. Esta situación no es restrictiva dado el carácter local del límite. Supongamos que para todo $x \in I \setminus \{0\}$ existe el límite $\lim\limits_{y\to 0} f(x,y)$. Tal límite dependerá en general del valor fijado de $x \in I \setminus \{0\}$. Definimos:

$$G : I \setminus \{0\} \to \mathbb{R} \qquad G(x) = \lim_{y\to 0} f(x,y) \tag{1.5}$$

Si la función G tiene límite en 0 dicho límite se llama un *límite iterado* de f en $(0,0)$ y suele representarse en la forma:

$$\lim_{x\to 0} G(x) = \lim_{x\to 0}\left(\lim_{y\to 0} f(x,y)\right) \tag{1.6}$$

Análogamente, si para todo $y \in J \setminus \{0\}$ existe el límite $\lim\limits_{x\to 0} f(x,y)$, podemos definir la función:

$$H : J \setminus \{0\} \to \mathbb{R} \qquad H(y) = \lim_{x\to 0} f(x,y) \tag{1.7}$$

Si la función H tiene límite en 0 dicho límite se llama un *límite iterado* de f en $(0,0)$ y suele representarse en la forma:

$$\lim_{y\to 0} H(y) = \lim_{y\to 0}\left(\lim_{x\to 0} f(x,y)\right) \tag{1.8}$$

Los límites iterados pueden no existir incluso aunque exista el límite de f en $(0,0)$.

1.52 Proposición. *En la situación que estamos considerando, supongamos que se verifica alguna de las siguientes afirmaciones:*

- *Los dos límites iterados* (1.6) *y* (1.8) *existen y son distintos.*
- *La función G definida en* (1.5) *existe pero no existe el límite de G en 0.*

- *La función H definida en (1.7) existe pero no existe el límite de H en 0.*

Entonces no existe el límite de f en $(0,0)$.

Si existe solamente uno de los límites iterados o bien ambos y coinciden, entonces el límite de f en $(0,0)$ puede no existir, pero si existe es igual al valor de dicho límite iterado.

Observa que los límites iterados son límites de funciones de una variable por lo que suelen ser fáciles de calcular.

Ejercicios propuestos

80. Estudia la existencia del límite en $(0,0)$ para cada uno de los siguientes campos escalares definidos en $\mathbb{R}^2 \setminus \{(0,0)\}$.

a) $f(x,y) = \dfrac{y}{\sqrt{x^2+y^2}}$ b) $f(x,y) = \dfrac{x^2-y^2}{x^2+y^2}$ c) $f(x,y) = \dfrac{x^2y}{x^4+y^2}$

d) $f(x,y) = \dfrac{xy}{x^2+y^2}$ e) $f(x,y) = \dfrac{(x+y)^2}{x^2+y^2}$ f) $f(x,y) = \dfrac{x\operatorname{sen} y + y\operatorname{sen} x}{x^2+y^2}$

g) $f(x,y) = \dfrac{y^3}{(x^2+y^2)^{3/2}}$ h) $f(x,y) = \begin{cases} y\operatorname{sen}(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases}$

i) $f(x,y) = \dfrac{x^2y^2}{(x^2+y^2)^{3/2}}$ j) $f(x,y) = \dfrac{x^4+3x^2y^2+2xy^3}{(x^2+y^2)^2}$

k) $f(x,y) = \dfrac{2x^5+2y^3(2x^2-y^2)}{(x^2+y^2)^2}$

l) $f(x,y) = \dfrac{y^2(x^3+y^2)+x^4}{x^4+y^4}$

m) $f(x,y) = \dfrac{x^2y+x\operatorname{sen} y}{\sqrt{x^2+y^2-xy}}$

n) $f(x,y) = \dfrac{1+x-\cos(x^2+y^2)-\operatorname{arc\,tg} x}{x^2+y^2}$

En los ejemplos de campos escalares de dos variables que hemos considerado aparece con frecuencia la expresión x^2+y^2. En los campos escalares de tres variables aparece con frecuencia la expresión $x^2+y^2+z^2$. Esto no es un capricho matemático sino que procede de la Física: los potenciales newtonianos y los potenciales eléctricos varían en razón inversa a la distancia euclídea y las fuerzas de atracción gravitatoria y eléctrica varían en razón inversa al cuadrado de la distancia euclídea. La expresión x^2+y^2 se simplifica usando coordenadas polares, es decir, escribiendo $(x,y) = (\rho\cos\vartheta, \rho\operatorname{sen}\vartheta)$ con lo cual $x^2+y^2 = \rho^2$.

1.53 Proposición. *Sea $f : A \to \mathbb{R}$ donde $A = B_2((0,0),r) \setminus \{(0,0)\}$ y $B_2((0,0),r)$ es la bola euclídea abierta en $\mathbb{R}^2$ centrada en el origen de radio r. Definamos la función:*

$$g :]0,r[\times [0,2\pi] \to \mathbb{R} \quad \textit{dada por} \quad g(\rho,\vartheta) = f(\rho\cos\vartheta, \rho\operatorname{sen}\vartheta)$$

Entonces equivalen las siguientes afirmaciones:

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

b) Para todo $\varepsilon>0$ *existe* $\delta>0$ *tal que para todo* $\rho\in]0,\delta[$ *se verifica que*

$$\sup\{|f(\rho\cos\vartheta,\rho\operatorname{sen}\vartheta)-\ell| : 0\leqslant\vartheta\leqslant 2\pi\}\leqslant\varepsilon.$$

En el estudio de límites conviene tener presente la desigualdad $|xy|\leqslant\frac{1}{2}(x^2+y^2)$ que se deduce de $0\leqslant(|x|-|y|)^2$.

Ejercicios propuestos

81. Estudia la existencia del límite en $(0,0)$, y calcula su valor cuando exista, para cada uno de los siguientes campos escalares definidos en $\mathbb{R}^2\setminus\{(0,0)\}$.

a) $f(x,y)=\dfrac{x^2y}{x^2+y^2}$ b) $f(x,y)=\dfrac{x^3-y^3}{x^2+y^2}$ c) $f(x,y)=\dfrac{\log(1+\sqrt{x^2+y^2})}{|x|+|y|}$

d) $f(x,y)=\dfrac{\operatorname{sen}y\log(1+x^2)}{x^2+y^2}$ e) $f(x,y)=\dfrac{x\operatorname{sen}y-y\operatorname{sen}x}{(x^2+y^2)^{\frac{3}{2}}}$

82. Estudia la existencia del límite en $(0,0)$, y calcula su valor cuando exista, para cada uno de los siguientes campos escalares definidos en $\mathbb{R}^+\times\mathbb{R}^+$.

a) $f(x,y)=\dfrac{\operatorname{sen}x+\operatorname{sen}y}{x+y}$ b) $f(x,y)=\dfrac{x\,\mathrm{e}^y+y\,\mathrm{e}^x}{x+y}$ c) $f(x,y)=\left(\dfrac{1}{x^2}+\dfrac{1}{y^2}\right)\operatorname{sen}(xy)$

83. Estudia para qué números positivos α el campo escalar $f:\mathbb{R}^2\setminus\{(0,0)\}\to\mathbb{R}$ dado por:

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

tiene límite en $(0,0)$.

Conjuntos conexos

Queremos definir una clase de conjuntos en los espacios métricos que desempeñe un papel análogo al de los intervalos en la recta real. Los intervalos pueden caracterizarse fácilmente con el teorema de Bolzano.

1.54 Proposición. *Sea* $\emptyset\neq C\subset\mathbb{R}$*. Equivalen las afirmaciones:*

a) C *es un intervalo.*

b) Para toda función continua $f:C\to\mathbb{R}$ *se verifica que* $f(C)$ *es un intervalo.*

La implicación $a)\Rightarrow b)$ es el famoso teorema de Bolzano. La otra implicación es un sencillo ejercicio que debes hacer.

1.55 Definición. Sea E un espacio métrico y $\emptyset \neq C \subset E$ se dice que C es un conjunto *conexo* si para *toda* función continua $f : C \to \mathbb{R}$ se verifica que $f(C)$ es un intervalo.

Los conexos de $\mathbb{R}$ son, en virtud de la proposición anterior, los intervalos. Como la continuidad es una propiedad topológica, el concepto de conjunto conexo es topológico. Seguidamente lo vamos a caracterizar de una forma más cómoda, pues la definición que hemos dado hace intervenir a *todas* las funciones continuas $f : C \to \mathbb{R}$ y eso la hace incómoda de usar en casos concretos.

Es muy fácil caracterizar los conjuntos que *no* son conexos. Un conjunto C no es conexo si, y sólo si, *existe* alguna función continua $f : C \to \mathbb{R}$ tal que $f(C)$ no es un intervalo. Que $f(C)$ no sea un intervalo equivale a que existe un $z \in \mathbb{R}$ con $z \notin f(C)$ tal que $]-\infty, z[\cap f(C) \neq \emptyset$ y $]z, +\infty[\cap f(C) \neq \emptyset$. Pero entonces se tiene que $C = f^{-1}(]-\infty, z[) \cup f^{-1}(]z, +\infty[)$. Puesto que f es continua, sabemos, por la proposición 1.36 que los conjuntos $A = f^{-1}(]-\infty, z[)$ y $B = f^{-1}(]z, +\infty[)$ son abiertos relativos de C. Observa que $A \neq \emptyset$ y $B \neq \emptyset$, $A \cap B = \emptyset$ y $C = A \cup B$.

Hemos probado así que si C *no* es conexo entonces existen *abiertos relativos* de C *disjuntos y no vacíos* A y B tales que $C = A \cup B$. Se dice que el par $\{A, B\}$ es una partición *no trivial* de C por abiertos relativos.

Recíprocamente, sea C un conjunto no vacío de un espacio métrico y supongamos que hay una partición no trivial de C por abiertos relativos $\{A, B\}$. Entones es fácil comprobar usando la proposición 1.36 que la función $f : C \to \mathbb{R}$ dada por:

$$f(x) = \begin{cases} -1, & \text{si } x \in A; \\ 1, & \text{si } x \in B. \end{cases}$$

es continua. Y como su imagen $f(C) = \{-1, 1\}$ no es un intervalo concluimos que C no es conexo.

Hemos probado que C no es conexo equivale a que exista una partición no trivial de C por abiertos relativos. Dicho de otra forma:

1.56 Proposición. *Un conjunto $\emptyset \neq C \subset E$ de un espacio métrico E es conexo si, y sólo si, la única partición de C por abiertos relativos es la trivial. Es decir, si A y B son abiertos relativos de C con $A \cap B = \emptyset$ y $C = A \cup B$ entonces el par $\{A, B\}$ ha de ser la partición trivial $\{A, B\} = \{\emptyset, C\}$.*

Al igual que vimos con la compacidad, la conexión se conserva por funciones continuas.

1.57 Proposición. *Sean E y F espacios métrico, C un subconjunto conexo de E y $f : C \to F$ una función continua. Entonces se verifica que $f(C)$ es conexo.*

Ejercicios propuestos

84. Prueba la proposición 1.57.

85. Prueba que la unión de conexos con intersección no vacía es conexo.

Los conjuntos abiertos y conexos se llaman *dominios*. Los dominios de $\mathbb{R}^n$ tienen una caracterización muy útil.

1.58 Proposición. *Sea $\Omega \subset \mathbb{R}^n$ un conjunto abierto. Equivalen las siguientes afirmaciones:*

a) Ω es conexo.

b) Todo par de puntos de Ω puede unirse por medio de una curva sin salirse de Ω. Es decir, para cada par de puntos $\mathbf{a}, \mathbf{b} \in \Omega$ existe una función continua $\gamma : [0,1] \to \mathbb{R}^n$, tal que $\gamma(0) = \mathbf{a}$, $\gamma(1) = \mathbf{b}$ y para todo $t \in [0,1]$ se tiene que $\gamma(t) \in \Omega$.

Intuitivamente, un dominio es un conjunto abierto *de un solo trozo*.

Capítulo 2

Derivadas parciales y extremos relativos de campos escalares

Curvas en $\mathbb{R}^n$

Como ya debes saber, una curva en $\mathbb{R}^n$ es una aplicación continua $\gamma : [a,b] \to \mathbb{R}^n$. El punto $\gamma(a)$ se llama *origen* y el punto $\gamma(b)$ *extremo* de la curva. Naturalmente, como $\gamma(t) \in \mathbb{R}^n$ podremos expresarlo por medio de sus componentes en la base canónica que serán funciones de t.

$$\gamma(t) = (\gamma_1(t), \gamma_2(t), \dots, \gamma_n(t))$$

Las funciones $\gamma_k(t)$ se llaman funciones componentes de γ. Se dice que γ es derivable en un punto t cuando todas sus funciones componentes son derivables en dicho punto, en cuyo caso la derivada de γ en t es, por definición, el vector

$$\gamma'(t) = (\gamma_1{}'(t), \gamma_2{}'(t), \dots, \gamma_n{}'(t))$$

Dado un punto $\mathbf{a} = \gamma(t_0)$ tal que $\gamma'(t_0) \neq \mathbf{0}$, se define la **recta tangente** a γ en el punto $\mathbf{a}$ (aunque es más apropiado decir *en el punto* t_0) como la recta de ecuación paramétrica $\mathbf{a} + t\gamma'(t_0)$, es decir, la recta que pasa por $\mathbf{a}$ con vector de dirección $\gamma'(t_0)$.

Cuando se interpreta $\gamma(t)$ como la función de trayectoria de un móvil, entonces su **velocidad** en un instante t es el vector $\gamma'(t)$ y su **rapidez** es $\|\gamma'(t)\|_2$. La distancia que recorre dicho móvil entre dos instantes $t = a$ y $t = b$ viene dada por

$$\int_a^b \|\gamma'(t)\|_2 \,\mathrm{d}t\,.$$

También consideraremos en lo que sigue funciones $\gamma : I \to \mathbb{R}^n$ definidas y continuas en un intervalo $I \subset \mathbb{R}$ no necesariamente cerrado y acotado. Las seguiremos llamando, por abuso de lenguaje, *curvas*.

2.1. Derivadas parciales. Vector gradiente

Hemos visto en el capítulo 1 que los conceptos de continuidad y límite para funciones reales de una variable se generalizan fácilmente para campos escalares de varias variables. No ocurre lo mismo con el concepto de derivabilidad el cual no puede generalizarse de forma inmediata. La razón es que el concepto de derivabilidad hace intervenir la división de números reales, pues una derivada es un límite de cocientes incrementales, y en $\mathbb{R}^n$ no podemos dividir por vectores, es decir, la estructura algebraica de $\mathbb{R}^n$ no permite generalizar algo parecido a un "cociente incremental". Si f es un campo escalar de dos o más variables, la expresión

$$\frac{f(\mathbf{x}) - f(\mathbf{a})}{\mathbf{x} - \mathbf{a}}$$

no tiene ningún sentido.

Otra diferencia importante es que en la recta real, $\mathbb{R}$, solamente podemos acercarnos a un punto de ella a través de la propia recta, mientras que en $\mathbb{R}^n$ para $n \geqslant 2$ hay muchísimas más posibilidades de acercarse a un punto dado; por ejemplo, podemos acercarnos a través de cualquier curva que pase por dicho punto. Surge así una primera idea que consiste en acercarse a un punto dado a través de una recta dada. Parece que esta situación es más parecida a lo que conocemos para funciones reales de una variable.

Las cuestiones que vamos a estudiar son de naturaleza algebraica y topológica, no dependen para nada de la norma que se use; por eso, en todo lo que sigue, representaremos con $\| \ \|$ una norma en $\mathbb{R}^n$. Si para ti es más cómodo, puedes considerar que siempre que aparezca una norma en $\mathbb{R}^n$ se trata de la norma euclídea. Cuando en algún resultado concreto sea preciso considerar la norma euclídea la representaremos en la forma usual por $\| \ \|_2$.

Una **dirección** o un **vector unitario** en $(\mathbb{R}^n, \|.\|)$ es un vector $\mathbf{u}$ con $\|\mathbf{u}\| = 1$.

Dados un punto $\mathbf{a} \in \mathbb{R}^n$ y una dirección $\mathbf{u}$, la recta que pasa por $\mathbf{a}$ con dirección $\mathbf{u}$ es la imagen de la aplicación $\gamma : \mathbb{R}^n \to \mathbb{R}$ dada por $\gamma(t) = \mathbf{a} + t\mathbf{u}$, es decir, es el conjunto de puntos $\{\mathbf{a} + t\mathbf{u} : t \in \mathbb{R}\}$.

2.1 Definición. Sea f un campo escalar definido en un conjunto abierto $\Omega \subset \mathbb{R}^n$, sea $\mathbf{a} \in \Omega$ y $\mathbf{u}$ una dirección. Se define la **derivada de f en a en la dirección u** como el límite

$$D_{\mathbf{u}} f(\mathbf{a}) = \lim_{t \to 0} \frac{f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a})}{t} \tag{2.1}$$

supuesto, claro está, que dicho límite exista.

La derivada direccional de un campo escalar f en un punto $\mathbf{a}$ en la dirección del vector $\mathbf{e_k}$ de la base canónica, se llama **derivada parcial** (de primer orden) de f en $\mathbf{a}$ respecto a la variable k-ésima. Está definida por:

$$\begin{aligned} D_{\mathbf{e_k}} f(\mathbf{a}) &= \lim_{t \to 0} \frac{f(\mathbf{a} + t\mathbf{e_k}) - f(\mathbf{a})}{t} = \lim_{t \to 0} \frac{f(a_1, \dots, a_k + t, \dots, a_n) - f(a_1, \dots, a_k, \dots, a_n)}{t} \\ &= \lim_{x_k \to a_k} \frac{f(a_1, \dots, x_k, \dots, a_n) - f(a_1, \dots, a_k, \dots, a_n)}{x_k - a_k} \end{aligned} \tag{2.2}$$

y se representa con los símbolos $D_k f(\mathbf{a})$ y $\dfrac{\partial f}{\partial x_k}(\mathbf{a})$.

Observa que las derivadas que acabamos de definir son derivadas de funciones reales de una variable real pues para calcular la derivada de un campo escalar f en un punto $\mathbf{a}$ en la dirección $\mathbf{u}$ lo que se hace es derivar en $t=0$ la función $t \mapsto f(\mathbf{a}+t\mathbf{u})$ que es una función real de una variable real.

Observa que la segunda igualdad de (2.2) nos dice que, *para calcular la derivada parcial $D_k f(\mathbf{a})$, lo que se hace es derivar f respecto a la variable k-ésima considerando fijas las demás variables*. Por eso se llaman derivadas *parciales*.

2.1.1. Interpretación geométrica de las derivadas parciales

Es importante que entiendas el significado de las derivadas parciales de una función en un punto. Para poder visualizarlo vamos a considerar un campo escalar f de dos variables definido en un abierto $\Omega \subset \mathbb{R}^2$. Fijemos un punto $(a,b)\in\Omega$. Las derivadas parciales de f en (a,b) son, por definición

$$\begin{aligned} D_1 f(a,b) &= \lim_{t\to 0}\frac{f(a+t,b)-f(a,b)}{t} = \lim_{x\to a}\frac{f(x,b)-f(a,b)}{x-a} \\ D_2 f(a,b) &= \lim_{t\to 0}\frac{f(a,b+t)-f(a,b)}{t} = \lim_{y\to b}\frac{f(a,y)-f(a,b)}{y-b} \end{aligned}$$

Es decir, lo que hacemos es derivar las funciones parciales $x \mapsto f(x,b)$ y $y \mapsto f(a,y)$ en los puntos $x=a$ e $y=b$ respectivamente.

La gráfica de f, es decir, el conjunto $S=\{(x,y,f(x,y)) : (x,y)\in\Omega\}$ es una superficie en $\mathbb{R}^3$. Las funciones

$$\gamma_1(x)=(x,b,f(x,b)), \quad \gamma_2(y)=(a,y,f(a,y))$$

son curvas contenidas en dicha superficie que pasan por el punto $(a,b,f(a,b))$. Dichas curvas se obtienen cortando la superficie S por los planos $y=b$ y $x=a$ respectivamente. Los vectores tangentes a dichas curvas en los puntos $\gamma_1(a)$ y $\gamma_2(b)$ son, respectivamente

$$\gamma_1{}'(a)=(1,0,D_1 f(a,b)), \quad \gamma_2{}'(b)=(0,1,D_2 f(a,b))$$

En la figura (2.1) se ha representado la gráfica de f y las curvas obtenidas cortándola por los planos $x=a$ e $y=b$ junto a sus vectores tangentes en el punto $(a,b,f(a,b))$

Cuando un campo escalar f tiene derivadas parciales en todos los puntos de un conjunto abierto $\Omega \subset \mathbb{R}^n$, podemos definir las *funciones derivadas parciales* (de primer orden) de f, $D_k f : \Omega \to \mathbb{R}$ que a cada punto $\mathbf{x}\in E$ hace corresponder el número $D_k f(\mathbf{x})$. Dichas funciones son también campos escalares.

2.2 Definición. Sea f un campo escalar. Se define el **vector gradiente** de f en un punto $\mathbf{a}$ como el vector

$$\nabla f(\mathbf{a}) = \big(D_1 f(\mathbf{a}), D_2 f(\mathbf{a}), \ldots, D_n f(\mathbf{a})\big)$$

supuesto, claro está, que dichas derivadas parciales existan.

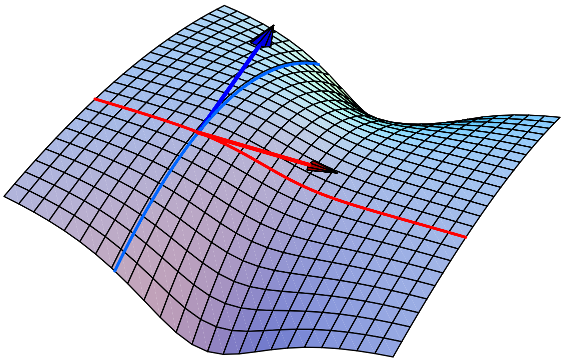


Figura 2.1. Derivadas parciales

Supongamos que f es una función real de una variable real. La derivabilidad de f en un punto $a \in \mathbb{R}$ se expresa por

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = f'(a) \Longleftrightarrow \lim_{x \to a} \frac{f(x) - f(a) - f'(a)(x - a)}{x - a} = 0$$

Recuerda que la recta de ecuación cartesiana $y = f(a) + f'(a)(x - a)$ es la recta tangente a la gráfica de f en el punto $(a, f(a))$.

Si ahora f es un campo escalar definido en un abierto $\Omega \subset \mathbb{R}^n$, cuyo vector gradiente $\nabla f(\mathbf{a})$ está definido en un punto $a \in \Omega$, podemos considerar el hiperplano en $\mathbb{R}^{n+1}$ de ecuación cartesiana $x_{n+1} = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}) | \mathbf{x} - \mathbf{a} \rangle$. Este hiperplano pasa por el punto $(\mathbf{a}, f(\mathbf{a})) \in \mathbb{R}^{n+1}$ y es la generalización natural de la recta tangente a la gráfica de una función. Observa el parecido formal entre las expresiones

$$y = f(a) + f'(a)(x - a), \qquad x_{n+1} = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}) | \mathbf{x} - \mathbf{a} \rangle$$

Ambas representan hiperplanos (un hiperplano en $\mathbb{R}^2$ es una recta) y la segunda se deduce de la primera sustituyendo la derivada por el vector gradiente y el producto usual de números reales por el producto escalar de vectores. Esto nos lleva a la siguiente definición.

2.1.2. Campos escalares diferenciables

2.3 Definición. Sea f un campo escalar definido en un abierto $\Omega \subset \mathbb{R}^n$ y sea $\mathbf{a} \in \Omega$. Supongamos que está definido el vector gradiente $\nabla f(\mathbf{a})$. Se dice que f es **diferenciable** o **derivable** en $\mathbf{a}$ si se verifica que

$$\lim_{\mathbf{x} \to \mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - \langle \nabla f(\mathbf{a}) | \mathbf{x} - \mathbf{a} \rangle}{\|\mathbf{x} - \mathbf{a}\|} = 0 \tag{2.3}$$

Definamos

$$R(\mathbf{x}, \mathbf{a}) = \frac{f(\mathbf{x}) - f(\mathbf{a}) - \langle \nabla f(\mathbf{a}) | \mathbf{x} - \mathbf{a} \rangle}{\|\mathbf{x} - \mathbf{a}\|}$$

La igualdad (2.3) dice que $\lim\limits_{\mathbf{x}\to\mathbf{a}} R(\mathbf{x},\mathbf{a}) = 0$. Con lo que, otra forma equivalente de escribir la igualdad (2.3) es la siguiente.

$$f(\mathbf{x}) = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}) | \mathbf{x}-\mathbf{a} \rangle + R(\mathbf{x},\mathbf{a})\|\mathbf{x}-\mathbf{a}\| \quad \text{donde} \quad \lim_{\mathbf{x}\to\mathbf{a}} R(\mathbf{x},\mathbf{a}) = 0 \tag{2.4}$$

La siguiente proposición expresa lo dicho en la definición anterior de una forma más abstracta. Recuerda que si $\alpha : \mathbb{R}^n \to \mathbb{R}$ es una aplicación lineal entonces para todo vector $\mathbf{x} = (x_1, x_2, \dots, x_n) = \sum\limits_{j=1}^{n} x_j \mathbf{e}_j$ en $\mathbb{R}^n$ se tiene que:

$$\alpha(\mathbf{x}) = \sum_{j=1}^{n} x_j \alpha(\mathbf{e}_j) = \langle \hat{\alpha} | \mathbf{x} \rangle$$

donde $\hat{\alpha} = \big(\alpha(\mathbf{e}_1), \alpha(\mathbf{e}_2), \dots, \alpha(\mathbf{e}_n)\big)$.

2.4 Proposición. *Sea f un campo escalar definido en un abierto $\Omega \subset \mathbb{R}^n$ y sea $\mathbf{a} \in \Omega$. Equivalen las siguientes afirmaciones:*

a) f es diferenciable en $\mathbf{a}$.

b) Existe una aplicación lineal $\alpha : \mathbb{R}^n \to \mathbb{R}$ tal que

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - \alpha(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} = 0 \tag{2.5}$$

Además, caso de que se cumpla b) se tiene que α es la aplicación lineal dada por $\alpha(\mathbf{x}) = \langle \nabla f(\mathbf{a}) | \mathbf{x} \rangle$.

2.5 Definición. Sea f un campo escalar diferenciable en un punto $\mathbf{a}$. El hiperplano en $\mathbb{R}^{n+1}$ de ecuación cartesiana

$$x_{n+1} = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}) | \mathbf{x}-\mathbf{a} \rangle$$

se llama hiperplano tangente a f en $\mathbf{a}$ o **hiperplano tangente** a la gráfica de f en el punto $(\mathbf{a}, f(\mathbf{a}))$.

2.6 Proposición. *Sea f un campo escalar diferenciable en un punto $\mathbf{a}$ y sea $\mathbf{u}$ una dirección en $\mathbb{R}^n$. Entonces se verifica que*

$$D_{\mathbf{u}} f(\mathbf{a}) = \langle \nabla f(\mathbf{a}) | \mathbf{u} \rangle$$

Demostración. En la igualdad (2.4) pongamos $\mathbf{x} = \mathbf{a} + t\mathbf{u}$ con lo que obtenemos

$$\begin{aligned} f(\mathbf{a}+t\mathbf{u}) &= f(\mathbf{a}) + \langle \nabla f(\mathbf{a}) | t\mathbf{u} \rangle + R(\mathbf{a}+t\mathbf{u},\mathbf{a})\|t\mathbf{u}\| \\ &= f(\mathbf{a}) + t\langle \nabla f(\mathbf{a}) | \mathbf{u} \rangle + R(\mathbf{a}+t\mathbf{u},\mathbf{a})|t| \end{aligned}$$

Deducimos que

$$\lim_{t\to 0} \frac{f(\mathbf{a}+t\mathbf{u}) - f(\mathbf{a})}{t} = \langle \nabla f(\mathbf{a}) | \mathbf{u} \rangle + \lim_{t\to 0} R(\mathbf{a}+t\mathbf{u},\mathbf{a}) \frac{|t|}{t} = \langle \nabla f(\mathbf{a}) | \mathbf{u} \rangle .$$

□

2.7 Corolario. *Sea f un campo escalar diferenciable en un punto $\mathbf{a}$ con vector gradiente no nulo en $\mathbf{a}$.*

a) La dirección en la que la derivada direccional de f en $\mathbf{a}$ es máxima es la dirección dada por el gradiente de f en $\mathbf{a}$, es decir, la dirección $\mathbf{u} = \dfrac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|_2}$.

b) La dirección en la que la derivada direccional de f en $\mathbf{a}$ es mínima es la dirección opuesta a la dada por el gradiente de f en $\mathbf{a}$, es decir, la dirección $\mathbf{v} = -\dfrac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|_2}$.

Demostración. Las afirmaciones hechas son consecuencia de la proposición anterior y de la desigualdad de Cauchy–Schwarz, pues para toda dirección $\mathbf{w}$ se tiene que

$$|\langle \nabla f(\mathbf{a}) | \mathbf{w} \rangle| \leqslant \|\nabla f(\mathbf{a})\|_2 \|\mathbf{w}\|_2 = \|\nabla f(\mathbf{a})\|_2$$

Y la igualdad se da si, y solo si, hay un número $\lambda \in \mathbb{R}$ tal que $\mathbf{w} = \lambda \nabla f(\mathbf{a})$. Tomando normas en esta igualdad se deduce que $|\lambda| = 1/\|\nabla f(\mathbf{a})\|_2$, es decir las direcciones $\mathbf{w}$ que hacen máximo $|D_{\mathbf{w}} f(\mathbf{a})| = |\langle \nabla f(\mathbf{a}) | \mathbf{w} \rangle|$ son $\mathbf{u} = \dfrac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|_2}$ y $\mathbf{v} = -\dfrac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|_2}$.

Para la primera se tiene que

$$D_{\mathbf{u}} f(\mathbf{a}) = \left\langle \nabla f(\mathbf{a}) \middle| \frac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|_2} \right\rangle = \frac{1}{\|\nabla f(\mathbf{a})\|_2} \langle \nabla f(\mathbf{a}) | \nabla f(\mathbf{a}) \rangle = \|\nabla f(\mathbf{a})\|_2$$

que es el valor máximo que puede tener una derivada direccional.

Análogamente, para la segunda se tiene que

$$D_{\mathbf{v}} f(\mathbf{a}) = -\|\nabla f(\mathbf{a})\|_2$$

que es el valor mínimo que puede tener una derivada direccional. □

El resultado anterior nos dice que el *vector gradiente en un punto señala la dirección en la que el campo tiene máximo crecimiento en dicho punto*. Mientras que en la dirección opuesta a la del vector gradiente en un punto el campo tiene máximo decrecimiento.

2.8 Proposición. *Sean f un campo escalar definido en un abierto $\Omega \subset \mathbb{R}^n$ y $\gamma : I \to \mathbb{R}^n$ una curva en $\mathbb{R}^n$ que toma valores en Ω. Supongamos que γ es derivable en un punto t_0 y que f es diferenciable en el punto $\mathbf{a} = \gamma(t_0) \in \Omega$. Entonces se verifica que la función $h(t) = f(\gamma(t))$ es derivable en t_0 y su derivada viene dada por:*

$$h'(t_0) = \langle \nabla f(\mathbf{a}) | \gamma'(t_0) \rangle = \sum_{k=1}^{n} D_k f(\mathbf{a}) \gamma_k{}'(t_0) \tag{2.6}$$

Si suponemos que f es diferenciable en Ω y que γ es derivable en I entonces para todo $t \in I$ se tiene que:

$$h'(t) = \langle \nabla f(\gamma(t)) | \gamma'(t) \rangle = \sum_{k=1}^{n} D_k f(\gamma(t)) \gamma_k{}'(t) \tag{2.7}$$

Demostración. Se tiene que

$$h(t)-h(t_0)=f(\gamma(t))-f(\gamma(t_0))=\left\langle \nabla f(\mathbf{a}) \middle| \gamma(t)-\gamma(t_0)\right\rangle + R(\gamma(t),\gamma(t_0))\|\gamma(t)-\gamma(t_0)\|$$

Dividiendo por $t-t_0$ tenemos

$$\frac{h(t)-h(t_0}{t-t_0}=\frac{f(\gamma(t))-f(\gamma(t_0))}{t-t_0}=\left\langle \nabla f(\mathbf{a}) \middle| \frac{\gamma(t)-\gamma(t_0)}{t-t_0}\right\rangle + R(\gamma(t),\gamma(t_0))\frac{\|\gamma(t)-\gamma(t_0)\|}{t-t_0}$$

Teniendo en cuenta que $\lim\limits_{t\to t_0}\dfrac{\gamma(t)-\gamma(t_0)}{t-t_0}=\gamma'(t_0)$ se deduce que

$$\lim_{t\to t_0}\frac{h(t)-h(t_0)}{t-t_0}=\left\langle \nabla f(\mathbf{a}) \middle| \gamma'(t_0)\right\rangle$$

como queríamos demostrar. □

2.9 Proposición. *Sean f un campo escalar diferenciable en un abierto $\Omega\subset\mathbb{R}^n$ y $g: I\to\mathbb{R}$ una función derivable en un intervalo I. Supongamos que $f(\Omega)\subset I$ y sea $h=g\circ f:\Omega\to\mathbb{R}$. Entonces el campo escalar h es diferenciable en Ω y su gradiente viene dado por:*

$$\nabla h(\mathbf{x})=g'(f(\mathbf{x}))\nabla f(\mathbf{x}) \qquad \forall \mathbf{x}\in\Omega \tag{2.8}$$

Demostración. Sean $\mathbf{a}\in\Omega$ y $b=f(\mathbf{a})\in I$, fijos en lo que sigue. Como g es derivable en b la función:

$$\varphi(t)=\frac{g(t)-g(b)}{t-b}-g'(b) \quad t\in I\setminus\{b\},\ \varphi(b)=0$$

es continua. Sustituyendo en la igualdad:

$$g(t)-g(b)=(g'(b)+\varphi(t))(t-b)$$

$t=f(\mathbf{x})$, $b=f(\mathbf{a})$ y usando la igualdad (2.4), tenemos:

$$g(f(\mathbf{x}))-g(f(\mathbf{a}))=\big(g'(b)+\varphi(f(\mathbf{x}))\big)(f(\mathbf{x})-f(\mathbf{a}))=$$

$$=g'(b)\left\langle \nabla f(\mathbf{a}) \middle| \mathbf{x}-\mathbf{a}\right\rangle+\big(g'(b)+\varphi(f(\mathbf{x}))\big)\|\mathbf{x}-\mathbf{a}\|R(\mathbf{x},\mathbf{a})+\varphi(f(\mathbf{x}))\left\langle \nabla f(\mathbf{a}) \middle| \mathbf{x}-\mathbf{a}\right\rangle$$

Por tanto:

$$\frac{h(\mathbf{x})-h(\mathbf{a})-g'(b)\left\langle \nabla f(\mathbf{a}) \middle| \mathbf{x}-\mathbf{a}\right\rangle}{\|\mathbf{x}-\mathbf{a}\|}=\big(g'(b)+\varphi(f(\mathbf{x}))\big)R(\mathbf{x},\mathbf{a})+\varphi(f(\mathbf{x}))\frac{\left\langle \nabla f(\mathbf{a}) \middle| \mathbf{x}-\mathbf{a}\right\rangle}{\|\mathbf{x}-\mathbf{a}\|}$$

Podemos suponer que la norma, $\|\ \|$, que hemos fijado en $\mathbb{R}^n$ es la euclídea para aplicar la desigualdad de Cauchy-Schwarz:

$$\left|\varphi(f(\mathbf{x}))\frac{\left\langle \nabla f(\mathbf{a}) \middle| \mathbf{x}-\mathbf{a}\right\rangle}{\|\mathbf{x}-\mathbf{a}\|}\right|\leqslant|\varphi(f(\mathbf{x}))|\,\|\nabla f(\mathbf{a})\|$$

Y teniendo en cuenta que $\lim\limits_{\mathbf{x}\to\mathbf{a}} R(\mathbf{x},\mathbf{a}) = 0$ y $\lim\limits_{\mathbf{x}\to\mathbf{a}} \varphi(f(\mathbf{x})) = 0$, deducimos que

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{h(\mathbf{x}) - h(\mathbf{a}) - g'(b)\left\langle \nabla f(\mathbf{a}) \middle| \mathbf{x}-\mathbf{a}\right\rangle}{\|\mathbf{x}-\mathbf{a}\|} = 0$$

Lo que prueba que h es diferenciable en $\mathbf{a}$ y que $\nabla h(\mathbf{a}) = g'(f(\mathbf{a}))\nabla f(\mathbf{a})$. □

Que un campo escalar tenga derivadas parciales en un punto es una propiedad muy débil. Por ejemplo, el campo escalar $f(x,y) = \dfrac{xy}{x^2+y^2}$, $f(0,0) = 0$ tiene derivadas parciales nulas en $(0,0)$ pero no es continuo en dicho punto.

La propiedad de ser diferenciable es mucho más fuerte que tener derivadas parciales.

2.10 Proposición. *Un campo escalar diferenciable en un punto es continuo en dicho punto.*

Para campos escalares diferenciables hay un análogo al teorema del valor medio para funciones derivables reales.

El segmento que une dos puntos $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ es el conjunto $[\mathbf{u},\mathbf{v}] = \{\mathbf{u}+t(\mathbf{v}-\mathbf{u}) : 0 \leqslant t \leqslant 1\}$.

2.11 Teorema (Del valor medio para campos escalares)**.** *Sea f un campo escalar diferenciable en un abierto $\Omega \subset \mathbb{R}^n$. Sean $\mathbf{a},\mathbf{b}\in\Omega$ y supongamos que el segmento $[\mathbf{a},\mathbf{b}] \subset \Omega$. Entonces existe un punto $\mathbf{c} \in [\mathbf{a},\mathbf{b}]$ tal que se verifica la igualdad:*

$$f(\mathbf{b}) - f(\mathbf{a}) = \left\langle \nabla f(\mathbf{c}) \middle| \mathbf{b}-\mathbf{a}\right\rangle$$

De donde se deduce:

$$|f(\mathbf{b}) - f(\mathbf{a})| \leqslant \|\mathbf{b}-\mathbf{a}\|_2 \sup\{\|\nabla f(\mathbf{x})\|_2 : \mathbf{x} \in [\mathbf{a},\mathbf{b}]\}$$

El siguiente resultado proporciona una condición suficiente de diferenciabilidad muy útil.

2.12 Teorema (Condición suficiente de diferenciabilidad). *Un campo escalar que tiene derivadas parciales continuas en un conjunto abierto es diferenciable en todo punto de dicho conjunto.*

La proposición 2.9 puede usarse para estudiar la diferenciabilidad de campos escalares. Con frecuencia un campo escalar f puede expresarse en la forma $f(\mathbf{x}) = g(p(\mathbf{x}))$ donde p es una función polinómica en n variables. Como las funciones polinómicas en n variables tienen derivadas parciales continuas, son diferenciables y, por la proposición antes citada, deducimos que el campo escalar $f(\mathbf{x}) = g(p(\mathbf{x}))$ es diferenciable en todo punto $\mathbf{x}$ tal que g sea derivable en $p(\mathbf{x})$.

2.13 Ejemplo. La función $g(t) = \dfrac{\operatorname{sen} t}{t}$, para $t \neq 0$ y $g(0) = 1$ es derivable en $\mathbb{R}$. Por tanto, el campo escalar $f(x,y) = \dfrac{\operatorname{sen}(x^2+y^2)}{x^2+y^2}$ y $f(0,0) = 1$ es diferenciable en $\mathbb{R}^2$. ♦

En la práctica suele suponerse que los campos escalares tienen derivadas parciales continuas. Esta hipótesis garantiza que son diferenciables y es suficiente para justificar la mayoría de los resultados que siguen.

Es sabido que una función derivable en un intervalo con derivada nula es constante. Para campos escalares hay un resultado análogo. Observa la hipótesis de que el campo esté definido en un *dominio*.

2.14 Proposición. *Un campo escalar definido en un dominio con derivadas parciales nulas en todo punto del mismo es constante.*

Ejercicios propuestos

86. Sean $f,g:\Omega\to\mathbb{R}$ campos escalares diferenciables en un abierto Ω. Prueba que los campos escalares $f+g$, fg son diferenciables en Ω y calcula sus vectores gradientes. Si, además, $g(\mathbf{x})\neq 0$ para todo $\mathbf{x}\in\Omega$ entonces prueba que el campo escalar $\dfrac{f}{g}$ es diferenciable en Ω y calcula su vector gradiente.

87. Sean f_i, $1\leqslant i\leqslant n$ funciones reales derivables en un intervalo abierto I. Definamos $f: I^n\to\mathbb{R}$ por:

$$f(x_1,x_2,\dots,x_n)=\sum_{j=1}^{n} f_j(x_j) \qquad \forall\mathbf{x}=(x_1,x_2,\dots,x_n)\in I^n$$

Prueba que el campo escalar f es diferenciable en I^n.

88. Sean $f,g:\Omega\to\mathbb{R}$ campos escalares definidos en un abierto $\Omega\subset\mathbb{R}^n$. Sea $\mathbf{a}\in\Omega$ y supongamos que f es diferenciable en $\mathbf{a}$, $f(\mathbf{a})=0$ y g es continua en $\mathbf{a}$. Prueba que fg es diferenciable en $\mathbf{a}$.

2.2. Rectas tangentes y planos tangentes

En esta sección vamos calcular rectas y planos tangentes a curvas y superficies considerando las distintas formas en que éstas pueden venir dadas. Mi propósito es esencialmente práctico, a saber, que entiendas la forma de proceder en cada caso; por lo que no me preocupo de justificar con detalle todo lo que digo.

Curvas en el plano

Una curva Γ en el plano puede venir dada de tres formas:

a) Como la *gráfica de una función* $y=f(x)$ donde $x\in I$ siendo I un intervalo de $\mathbb{R}$.

$$\Gamma=\{(x,f(x)):x\in I\}$$

b) Por medio de *ecuaciones paramétricas* $\gamma(t) = (x(t), y(t))$.

$$\Gamma = \gamma(I) = \{(x(t), y(t)) : t \in I\}$$

c) De *forma implícita* como el conjunto de puntos $g(x,y) = 0$ donde se anula una función diferenciable de dos variables.

$$\Gamma = \left\{(x,y) \in \mathbb{R}^2 : g(x,y) = 0\right\}$$

Suele usarse la siguiente terminología. Si $h(x,y)$ es un campo escalar diferenciable, las curvas de ecuación implícita $h(x,y) = c$ o, lo que es igual $h(x,y) - c = 0$, donde c es una constante, se llaman *curvas de nivel*. Dichas curvas se obtienen cortando la gráfica de h con planos de la forma $z = c$. Estas curvas son las que ves representadas en los mapas topográficos.

Observa que **a)** es un caso particular de **c)** (basta considerar $g(x,y) = f(x) - y$) y también es un caso particular de **b)** (basta considerar $\gamma(x) = (x, f(x))$).

La tangente en un punto de Γ viene dada en cada caso como sigue.

a′) La tangente en un punto $(a,b) = (a, f(a)) \in \Gamma$ es la recta de ecuación cartesiana

$$y - b = f'(a)(x - a)$$

El vector $(1, f'(a))$ es tangente a Γ en el punto (a,b) y el vector $(f'(a), -1)$ es ortogonal a Γ en el punto (a,b).

b′) La tangente en un punto $\gamma(t_0) = (a,b) \in \Gamma$ es la recta de ecuaciones paramétricas

$$(x,y) = \gamma(t_0) + t\gamma'(t_0) = (a,b) + t(x'(t_0), y'(t_0))$$

El vector $\gamma'(t_0) = (x'(t_0), y'(t_0))$ es tangente a Γ en (a,b).

c′) La tangente en un punto $(a,b) \in \Gamma$ es la recta de ecuación implícita

$$\left\langle \nabla g(a,b) \,\middle|\, (x-a, y-b) \right\rangle = 0$$

Se supone que $\nabla g(a,b) \neq 0$ pues en otro caso, la tangente en (a,b) no está definida. El vector gradiente $\nabla g(a,b)$ es ortogonal a Γ en el punto (a,b).

Estas últimas afirmaciones requieren alguna justificación. Para ello, supongamos que conocemos una *representación paramétrica local* de Γ en torno al punto (a,b). Es decir, hay una curva de la forma $\alpha(t) = (\alpha_1(t), \alpha_2(t)) \in \Gamma$ que pasa por el punto (a,b) y que es derivable[1]. Pongamos $\alpha(t_0) = (a,b)$. Por lo visto en **b′)**, sabemos que la tangente a Γ en (a,b) es la recta que pasa por el punto (a,b) con vector de dirección $\alpha'(t_0)$. Pongamos $h(t) = g(\alpha(t))$. En virtud de la igualdad (2.7), tenemos que $h'(t) = \left\langle \nabla g(\alpha(t)) \,\middle|\, \alpha'(t) \right\rangle$. Pero $h(t) = 0$, por lo que $h'(t) = \left\langle \nabla g(\alpha(t)) \,\middle|\, \alpha'(t) \right\rangle = 0$. Resulta así que el vector $\nabla g(\alpha(t))$ es ortogonal al vector tangente $\alpha'(t)$. En particular, el vector $\nabla g(a,b)$ es ortogonal al vector $\alpha'(t_0)$ tangente a Γ en (a,b). Concluimos que la recta que pasa por (a,b) y tiene como

[1] El teorema de la función implícita, que se verá más adelante, garantiza la existencia de dicha curva siempre que el vector gradiente $\nabla g(a,b) \neq \mathbf{0}$.

vector ortogonal $\nabla g(a,b)$ es la recta tangente a Γ en (a,b), pero dicha recta es justamente la recta de ecuación cartesiana $\left\langle \nabla g(a,b) \middle| (x-a, y-b) \right\rangle = 0$.

De lo antes visto, merece la pena destacar la siguiente propiedad.

El vector gradiente $\nabla g(x,y)$ de un campo escalar es ortogonal en todo punto (x,y) (en el que $\nabla g(x,y) \neq \mathbf{0}$) a la curva de nivel que pasa por dicho punto.

Superficies en $\mathbb{R}^3$

Una superficie S en el espacio $\mathbb{R}^3$ puede venir dada de tres formas:

a) Como la gráfica de una función $z = f(x,y)$ donde $(x,y) \in A$ siendo A un conjunto de $\mathbb{R}^2$.

$$S = \{(x,y,f(x,y)) : (x,y) \in A\}$$

b) Por ecuaciones paramétricas $\gamma(s,t) = (x(s,t), y(s,t), z(s,t))$ donde $(s,t) \in A \subset \mathbb{R}^2$.

$$S = \gamma(A) = \{(x(s,t), y(s,t), z(s,t)) : (s,t) \in A\}$$

c) De forma implícita como el conjunto de puntos $g(x,y,z) = 0$ donde se anula una función diferenciable de tres variables.

$$S = \left\{(x,y,z) \in \mathbb{R}^3 : g(x,y,z) = 0\right\}$$

Observa que **a)** es un caso particular de **c)** (basta considerar $g(x,y,z) = f(x,y) - z$) y también es un caso particular de **b)** (basta considerar $\gamma(s,t) = (s,t,f(s,t))$). El plano tangente en un punto de S viene dado en cada caso como sigue.

a′) El plano tangente en un punto $(a,b,c) = (a,b,f(a,b)) \in S$ es el plano de ecuación cartesiana

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

Los vectores $\left(1, 0, \frac{\partial f}{\partial x}(a,b)\right)$ y $\left(0, 1, \frac{\partial f}{\partial y}(a,b)\right)$ son tangentes a S en (a,b,c) y el vector

$$\left(\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b), -1\right)$$

es ortogonal a S en el punto (a,b,c).

b′) El plano tangente en un punto $\gamma(s_0,t_0) = (a,b,c) \in S$ es el plano de ecuaciones paramétricas

$$(x,y,z) = \gamma(s_0,t_0) + s\frac{\partial \gamma}{\partial s}(s_0,t_0) + t\frac{\partial \gamma}{\partial t}(s_0,t_0)$$

Donde

$$\frac{\partial \gamma}{\partial s}(s_0,t_0) = \left(\frac{\partial x}{\partial s}(s_0,t_0), \frac{\partial y}{\partial s}(s_0,t_0), \frac{\partial z}{\partial s}(s_0,t_0)\right)$$

y

$$\frac{\partial \gamma}{\partial t}(s_0,t_0) = \left(\frac{\partial x}{\partial t}(s_0,t_0), \frac{\partial y}{\partial t}(s_0,t_0), \frac{\partial z}{\partial t}(s_0,t_0)\right)$$

Dichos vectores son tangentes a S en (a,b,c).

c′) El plano tangente en un punto $(a,b,c) \in S$ es el plano de ecuación implícita

$$\left\langle \nabla g(a,b,c) \middle| (x-a, y-b, z-c) \right\rangle = 0$$

Se supone que $\nabla g(a,b,c) \neq 0$ pues en otro caso, el plano tangente a S en (a,b,c) no está definido. El vector gradiente $\nabla g(a,b,c)$ es ortogonal a S en el punto (a,b,c).

Si $g(x,y,z)$ es un campo escalar, las superficies de ecuación implícita $g(x,y,z) = c$ o, lo que es igual $g(x,y,z) - c = 0$, donde c es una constante, se llaman *superficies de nivel* (cuando el campo se interpreta como un potencial se llaman *superficies equipotenciales*). De lo dicho en **c′)**, se sigue que *el vector gradiente* $\nabla g(x,y,z)$ *es ortogonal en todo punto* (x,y,z) *(en el que* $\nabla g(x,y,z) \neq \mathbf{0}$*) a la superficie de nivel que pasa por dicho punto.*

Curvas en $\mathbb{R}^3$

Una curva Γ en el espacio puede venir dada de dos formas.

a) Como intersección de dos superficies S_1 y S_2.

b) Por medio de ecuaciones paramétricas $\gamma(t) = (x(t), y(t), z(t))$ donde $t \in I \subset \mathbb{R}$ e I es un intervalo.

$$\Gamma = \gamma(I) = \{(x(t), y(t), z(t)) : t \in I\}$$

La tangente en un punto de Γ viene dada en cada caso como sigue.

a′) La tangente en un punto $(a,b,c) \in \Gamma$ es la recta intersección de los planos tangentes a S_1 y a S_2 en (a,b,c). Por ejemplo, si las superficies vienen dadas por sus ecuaciones implícitas.

$$\begin{array}{l} S_1 = \left\{(x,y,z) \in \mathbb{R}^3 : f(x,y,z) = 0\right\} \\ S_2 = \left\{(x,y,z) \in \mathbb{R}^3 : g(x,y,z) = 0\right\} \end{array} \qquad \Gamma = \left\{(x,y,z) \in \mathbb{R}^3 : g(x,y,z) = f(x,y,z) = 0\right\}$$

Entonces, las ecuaciones implícitas de la recta tangente son

$$\left\{ \begin{array}{l} \left\langle \nabla f(a,b,c) \middle| (x-a, y-b, z-c) \right\rangle = 0 \\[2ex] \left\langle \nabla g(a,b,c) \middle| (x-a, y-b, z-c) \right\rangle = 0 \end{array} \right.$$

Donde se supone que los vectores gradiente $\nabla f(a,b,c)$, $\nabla g(a,b,c)$ son linealmente independientes pues, en otro caso, la recta tangente a la curva Γ en (a,b,c) no está definida.

b′) La tangente en un punto $\gamma(t_0) = (a,b,c) \in \Gamma$ es la recta de ecuaciones paramétricas

$$(x,y,z) = \gamma(t_0) + t\gamma'(t_0) = (a,b,c) + t(x'(t_0), y'(t_0), z'(t_0))$$

El vector $\gamma'(t_0) = (x'(t_0), y'(t_0), z'(t_0))$ es tangente a Γ en (a,b,c).

2.3. Derivadas parciales de orden superior

Supongamos un campo escalar f con derivadas parciales $D_k f$ en un abierto $\Omega \subset \mathbb{R}^n$. Las funciones $D_k f$ son también campos escalares de n variables que podemos, cuando se dejen, volver a derivar parcialmente en puntos de Ω. Obtenemos de esta forma las *derivadas parciales de segundo orden* de f, es decir las funciones $D_{jk} f = D_j(D_k f)$, que se representan también de las formas:

$$D_{jk} f(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \partial x_k}(\mathbf{x}), \quad D_{jj} f(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j^2}(\mathbf{x})$$

De manera análoga se definen las derivadas parciales de tercer orden de f como las derivadas parciales de las derivadas parciales de segundo orden de f, $D_{ijk} f = D_i(D_j(D_k f))$, y se representan también de las formas:

$$D_{ijk} f(\mathbf{x}) = \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathbf{x});\ D_{jjj} f(\mathbf{x}) = \frac{\partial^3 f}{\partial x_j^3}(\mathbf{x});\ D_{jjk} f(\mathbf{x}) = \frac{\partial^3 f}{\partial x_j^2 \partial x_k}(\mathbf{x})$$

Las derivadas parciales de cuarto orden son las derivadas parciales de las derivadas parciales de tercer orden y así sucesivamente. Es natural preguntarse si el orden en que se realizan las derivadas debe ser o no tenido en cuenta. Afortunadamente, en la mayoría de los casos podemos olvidarlo porque se verifica el siguiente utilísimo resultado.

2.15 Teorema (Schwarz). *Las derivadas parciales de orden menor o igual que k de un campo escalar con derivadas parciales de orden k continuas solamente dependen del número de veces que se deriva parcialmente respecto de cada variable, pero el orden en que se realicen dichas derivaciones no afecta para nada al resultado final.*

2.16 Definición. Se dice que un campo escalar f es de clase $\mathcal{C}^k$ en un abierto Ω si f tiene derivadas parciales de orden k continuas en Ω, en cuyo caso escribimos $f \in \mathcal{C}^k(\Omega)$. Se dice que f es de clase $\mathcal{C}^\infty$ en Ω si tiene derivadas parciales continuas de todos órdenes en Ω.

Ejercicios propuestos

Como para calcular derivadas parciales de una función de varias variables se consideran fijas todas las variables menos aquella respecto a la que se deriva, calcular derivadas parciales es lo mismo que derivar funciones de una variable. Solamente debes tener cuidado para darte cuenta qué tipo de función es la que tienes que derivar porque ello puede depender de la variable respecto de la que derivas. Por ejemplo, la función $f(x,y) = x^y$ cuando fijas y (para derivar respecto a x) es una función potencia (la variable está en la base y el exponente está fijo) y cuando fijas x (para derivar respecto a y) es una función exponencial (la variable está en el exponente y la base está fija).

Te recuerdo que es muy frecuente, sobre todo en libros de Física e ingenierías diversas, representar las funciones por letras. Así, lo que los matemáticos solemos escribir $f(x,y) = \cos(xy) + xy^2$, para indicar que f es una función de dos variables x e y cuyo valor en el punto (x,y) viene dado por $\cos(xy) + xy^2$, suele expresarse de forma menos precisa en la forma $z = \cos(xy) + xy^2$, cuyo significado es exactamente el mismo que el anterior cambiando f por z. Naturalmente, en vez de z puede usarse cualquier otro símbolo que sea distinto de x e y. Tienes que acostumbrarte a esta notación y entender cuando una letra representa una variable y cuando representa una función.

90. Calcula las derivadas parciales de primer orden de los campos escalares:

(a) $f(x,y)=x^2y+z^2x+y\operatorname{sen}(xz)$ (b) $z=(x^2+y^3)\,\mathrm{e}^{-xy}$ (c) $w=x\,\mathrm{e}^z+z\,\mathrm{e}^y+xyz$.

91. Calcula las derivadas parciales de primer y segundo orden del campo escalar:

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

92. Calcula las derivadas parciales de primer y segundo orden de los campos escalares:

(a) $z=\operatorname{sen}\left(\cos\left(\mathrm{e}^{xy}\right)\right)$ (b) $w=\log\left(4+\operatorname{arc\,tg}(x/y)\right)$

(c) $u=\operatorname{tg}\left((xy)^z\right)$ (d) $v=\operatorname{arc\,tg}\left(z^{xy}\right)$

93. Estudia la existencia de las derivadas en el origen según las distintas direcciones del campo escalar $f:\mathbb{R}^2\to\mathbb{R}$ definido por:

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

¿Es f continuo en $(0,0)$?

Te recuerdo que una dirección viene dada por un vector de norma euclídea 1. Si $\mathbf{a}$ y $\mathbf{b}$ son puntos de $\mathbb{R}^n$ la dirección del punto $\mathbf{a}$ hacia el punto $\mathbf{b}$ viene dada por el vector $\dfrac{\mathbf{b}-\mathbf{a}}{\|\mathbf{b}-\mathbf{a}\|_2}$.

94. Calcula la derivada direccional de $f(x,y)=\log(1+\sqrt{x^2+y^2})$ en el punto $(1,2)$ en la dirección hacia el origen.

95. Calcula la derivada direccional de $z(x,y)=\operatorname{arc\,tg}\left(\dfrac{xy}{x^2+y^2}\right)$ en el punto $(1,1)$ en la dirección hacia el punto $(2,1)$.

96. Calcula valores de a, b y c para que la derivada direccional de la función

$$f(x,y,z)=axy^2+byz+cz^2x^3$$

en el punto $(1,2,-1)$ tenga un valor máximo igual a 64 en la dirección del eje OZ.

97. Considera la curva dada por las ecuaciones paramétricas $x(t)=\mathrm{e}^t+\cos t$, $y(t)=\mathrm{e}^{-t}+\operatorname{sen} t$. Calcula la ecuación de la recta tangente en el punto $(x(0),y(0))$.

98. Calcula, para los siguientes campos escalares, el vector normal en P_0 a la curva de nivel que pasa por dicho punto.

1. $f(x,y) = \operatorname{arc\,tg}\left(\dfrac{y}{\sqrt{1+x^2+y^2}}\right) \qquad P_0 = (1,1).$
2. $f(x,y) = \dfrac{\operatorname{sen}(x+y)}{2+\cos(x-y)} \qquad P_0 = (\pi/2, \pi/4).$

99. Calcula la derivada de $h(x,y) = \dfrac{x-y}{1+\log(1+x^2y^2)}$ en el punto $(-1,-1)$ en la dirección dada por el vector ortogonal (de norma 1) en el punto $(1,1)$ a la curva de nivel del campo $f(x,y) = xy^3 + x^3y$ que pasa por dicho punto.

100. Calcula la ecuación de la recta tangente y de la recta normal a la elipse de ecuación

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

en un punto (u,v) de la misma.

101. Calcula las ecuaciones del plano tangente y de la recta normal a cada una de las siguientes superficies en el punto P_o indicado.

$$\begin{aligned}
z^2 - 2x^2 - 2y^2 - 12 = 0, &\quad P_o(1,-1,4);\\
z - \log(x^2+y^2) = 0, &\quad P_o(1,0,0)\\
x^2+y^2+z^3-2x+4y+3z+1 = 0, &\quad P_o(3,4,-3);\\
4 - x^2 - 4z^2 = y, &\quad P_o(0,0,1)\\
z(xy-1)-(x+y) = 0, &\quad P_o(1,2,3);\\
z + \mathrm{e}^z + 2x + 2y - x^2 - y^2 - 3 = 0, &\quad P_o(1, 1+\sqrt{e}, 1)
\end{aligned}$$

102. Halla la ecuación de la tangente a la curva dada como intersección del elipsoide $x^2 + 4y^2 + 2z^2 = 27$ y el hiperboloide $x^2 + y^2 - 2z^2 = 11$ en el punto $(3,-2,1)$.

103. Calcula la ecuación de la recta tangente a la curva definida por la intersección de las superficies $z = xy$, $x^2 + y^2 - 2z = 4$ en el punto $(3,1,3)$. Comprueba el resultado expresando la curva por sus ecuaciones paramétricas.

104. Calcula la ecuación de la recta tangente a la curva definida por la intersección de las superficies $4xz = (x+z)y$, $3z^2 + y = 5x$ en el punto $(1,2,1)$.

105. Estudia la diferenciabilidad de los siguientes campos escalares

$$\text{a) } f(x,y) = \frac{\mathrm{e}^{x^2+y^2}-1}{x^2+y^2},\ f(0,0) = 1 \qquad \text{b) } f(x,y) = \frac{\log(1+x^2+y^2)}{x^2+y^2},\ f(0,0) = 1$$

106. Sea

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

a) Estudia la continuidad y diferenciabilidad de f.

b) Calcula $D_{\mathbf{u}} f(1,0)$ donde $\mathbf{u} = (1/\sqrt{2}, 1/\sqrt{2})$.

c) Calcula la dirección y el valor de la derivada direccional mínima de f en $(0,0)$.

d) Calcula las ecuaciones del plano tangente y de la recta normal a la superficie $z = f(x,y)$ en el origen.

107. Estudia si los siguientes campos escalares definidos por $f(0,0) = 0$ y para $(x,y) \neq (0,0)$ por:

$$\text{a) } f(x,y) = \frac{4x^3}{x^2+y^2}, \quad \text{b) } f(x,y) = (x^2+y^2)\operatorname{arctg}\frac{1}{x^2+y^2}$$

son continuos, diferenciables o de clase $\mathcal{C}^1$ en $\mathbb{R}^2$. Calcula también en cada caso la derivada $D_{\mathbf{u}} f(0,0)$ donde $\mathbf{u} = (1/\sqrt{2}, 1/\sqrt{2})$.

108. Estudia si los siguientes campos escalares definidos por $f(0,0) = 0$ y para $(x,y) \neq (0,0)$ por:

$$\text{a) } f(x,y) = \frac{\operatorname{sen}(x^2y)}{\sqrt{x^2+y^2}}, \text{ b) } f(x,y) = \frac{\operatorname{sen}x \operatorname{sen}y - xy}{\sqrt{x^2+y^2}}, \ c) \ f(x,y) = \frac{1-\cos x\cos y}{\sqrt{x^2+y^2}}$$

son de clase $\mathcal{C}^1$ en $\mathbb{R}^2$.

109. Estudia si los siguientes campos escalares definidos por $f(0,0) = 0$ y para $(x,y) \neq (0,0)$ por:

$$\text{a) } f(x,y) = xy\frac{x^2-y^2}{x^2+y^2}, \text{ b) } f(x,y) = \frac{y\operatorname{arctg}x - x\operatorname{arctg}y}{x^2+y^2}, \text{ c) } f(x,y) = \frac{\operatorname{arctg}x\operatorname{sen}y - xy}{x^2+y^2}$$

$$\text{d) } f(x,y) = xy\frac{\cos x - \cos y}{x^2+y^2}, \text{ e) } f(x,y) = \frac{y\operatorname{tg}x - x\operatorname{tg}y}{x^2+y^2}, \text{ f) } f(x,y) = \frac{x\cos y - y\cos x - x + y}{\sqrt{x^2+y^2}}$$

son de clase $\mathcal{C}^1$ en $\mathbb{R}^2$. Estudia en cada caso $D_{12} f(0,0)$ y $D_{21} f(0,0)$ e indica si son de clase $\mathcal{C}^2$ en $\mathbb{R}^2$.

110. Estudia la derivabilidad de las normas $\mathbf{x} \mapsto \|\mathbf{x}\|_2$, $\mathbf{x} \mapsto \|\mathbf{x}\|_1$ y $\mathbf{x} \mapsto \|\mathbf{x}\|_\infty$ y calcula sus vectores gradiente.

111. Justifica que un campo escalar de clase $\mathcal{C}^k$ también es de clase $\mathcal{C}^q$ para $q \leqslant k$.

112. Justifica que la suma, el producto y el cociente de campos escalares de clase $\mathcal{C}^k$ también son campos escalares de clase $\mathcal{C}^k$. Deduce que toda función racional de n variables es de clase $\mathcal{C}^\infty$ en su dominio natural de definición.

113. Generaliza las proposiciones 2.8 y 2.9 para campos de clase $\mathcal{C}^k$.

114. Un campo escalar $f : \mathbb{R}^n \to \mathbb{R}$ se dice que es *homogéneo* de grado $n \in \mathbb{N}$, si para todo $\mathbf{x} \in \mathbb{R}^n$ y para todo $t \in \mathbb{R}$ se verifica que $f(t\mathbf{x}) = t^n f(\mathbf{x})$.

Supuesto que f es homogéneo de grado n y derivable, prueba que:

a) $\displaystyle\sum_{j=1}^{n} x_j D_j f(\mathbf{x}) = n f(\mathbf{x}) \qquad \forall \mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$

b) Si $n \geqslant 2$, las derivadas parciales $D_j f$ son funciones homogéneas de grado $n-1$.

c) Si $n = 1$ entonces f es una forma lineal.

115. Prueba que si f es de clase $\mathcal{C}^2$ y es homogéneo de grado 2 se verifica que:

$$f(\mathbf{x}) = \frac{1}{2}\sum_{i,j=1}^{n} D_{ij}f(0)x_i x_j \qquad \forall \mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$$

116. Sea $f : \Omega \to \mathbb{R}$ un campo escalar de clase $\mathcal{C}^1$. Supongamos que Ω es un conjunto convexo y que $\mathbf{0} \in \Omega$. Prueba que para todo $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \Omega$ se verifica que:

$$f(\mathbf{x}) = f(\mathbf{0}) + \sum_{j=1}^{n} x_j \int_0^1 D_j f(t\mathbf{x})\,\mathrm{d}t$$

2.17 Observación. ¿Qué notación es mejor para las derivadas parciales? ¿$\dfrac{\partial f}{\partial x}$ o $D_1 f$?. Quizás un ejemplo te ayude a decidir: consideremos $f(x,y) = x\mathrm{e}^y$. Tenemos que $D_1 f(x,y) = \mathrm{e}^y$ y $\dfrac{\partial f}{\partial x}(x,y) = \mathrm{e}^y$. Si ahora permutamos x por y en estas igualdades obtenemos $D_1 f(y,x) = \mathrm{e}^x$ y $\dfrac{\partial f}{\partial y}(y,x) = \mathrm{e}^x$. La igualdad $D_1 f(y,x) = \mathrm{e}^x$ es correcta porque, efectivamente, e^x es el valor de la derivada parcial de $f(x,y)$ respecto a la primera variable evaluada en el punto (y,x). La igualdad $\dfrac{\partial f}{\partial y}(y,x) = \mathrm{e}^x$ no es correcta porque $\dfrac{\partial f}{\partial y}(x,y) = x\mathrm{e}^y$ y si la evaluamos en (y,x) el resultado es $y\mathrm{e}^x$.

Este ejemplo pone de manifiesto que es mucho mejor la notación de subíndices $D_1 f$, $D_2 f$ para las derivadas parciales que las notaciones $\dfrac{\partial f}{\partial x}$ y $\dfrac{\partial f}{\partial y}$. La razón es que la notación D_1 no tiene nunca ambigüedad: indica derivación respecto a la primera variable y no depende para nada de la forma en que representemos dicha variable, mientras que la notación $\dfrac{\partial f}{\partial x}$ nos dice que dicha variable la hemos representado con la letra x. Pero con frecuencia hay que cambiar las letras con las que representamos las variables.

2.4. Teorema de Taylor. Extremos relativos

2.18 Definición. Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$. Se dice que f tiene un **máximo relativo** (resp. **mínimo relativo**) en un punto $\mathbf{a} \in E$, si $\mathbf{a}$ es un punto interior de E y existe un número $r > 0$ tal que $B(\mathbf{a}, r) \subset E$ y $f(\mathbf{x}) \leqslant f(\mathbf{a})$ (resp. $f(\mathbf{a}) \leqslant f(\mathbf{x})$) para todo $\mathbf{x} \in B(\mathbf{a}, r)$. Cuando las desigualdades anteriores se verifican de forma estricta para todo $\mathbf{x} \in B(\mathbf{a}, r)$ con $\mathbf{x} \neq \mathbf{a}$ se dice que el máximo o el mínimo relativo es **estricto**.

Los puntos en los que f tiene un máximo o un mínimo relativos se llaman **extremos relativos** de f.

2.19 Proposición (Condición necesaria de extremo relativo). *Sea f un campo escalar definido en un conjunto $E \subset \mathbb{R}^n$ y supongamos que f tiene un extremo relativo en un punto $\mathbf{a} \in E$ y además que el vector gradiente de f en $\mathbf{a}$ está definido. Entonces se verifica que $\nabla f(\mathbf{a}) = \mathbf{0}$. Es decir, las derivadas parciales de primer orden de f en $\mathbf{a}$ son todas nulas.*

Demostración. Supongamos que f tiene un máximo relativo en $\mathbf{a}$ y sea $r>0$ tal que $B(\mathbf{a},r)\subset E$ y $f(\mathbf{x})\leqslant f(\mathbf{a})$ para todo $\mathbf{x}\in B(\mathbf{a},r)$. Definamos $\varphi:]-r,r[\to\mathbb{R}$ por $\varphi(t)=f(\mathbf{a}+t\mathbf{e_k})$. La función φ está definida en el intervalo $]-r,r[$ pues para todo $t\in]-r,r[$ se tiene que $\|\mathbf{a}+t\mathbf{e_k}-\mathbf{a}\|=|t|<r$ por lo que $\mathbf{a}+t\mathbf{e_k}\in B(\mathbf{a},r)\subset E$. Además, para todo $t\in]-r,r[$ se tiene que $\varphi(t)=f(\mathbf{a}+t\mathbf{e_k})\leqslant f(\mathbf{a})=\varphi(0)$. Luego φ tiene en $t=0$ un máximo relativo. Además como, por hipótesis, existe $D_k f(\mathbf{a})$, tenemos que φ es derivable en $t=0$. Luego $\varphi'(0)=0$, pero $\varphi'(0)=D_k f(\mathbf{a})$. □

2.20 Definición. Los puntos donde se anula el gradiente de un campo escalar f se llaman **puntos críticos** o **puntos estacionarios** de f. Los puntos críticos de un campo escalar que no son extremos relativos se llaman **puntos de silla**.

Si f es un campo escalar diferenciable, en los puntos críticos el hiperplano tangente es "horizontal".

La condición necesaria de extremo relativo no es suficiente. Por ejemplo, el campo escalar $f(x,y)=x^2-y^2$ tiene un punto crítico en $(0,0)$, pero no tiene extremo relativo en dicho punto pues en toda bola centrada en $(0,0)$ toma valores positivos y negativos.

Al igual que para funciones de una variable, la derivada segunda proporciona una condición suficiente de extremo relativo, para campos escalares de varias variables las derivadas parciales de segundo orden nos van a permitir dar una condición suficiente de extremo relativo.

Necesitaremos para ello generalizar el teorema de Taylor para campos escalares de n variables. Para expresar dicho teorema de una forma sencilla es muy conveniente introducir una notación apropiada.

Si f es un campo escalar y $\mathbf{a},\mathbf{x}\in\mathbb{R}^n$ definimos el operador:

$$\mathrm{d}^k(f,\mathbf{a},\mathbf{x})=(x_1D_1+x_2D_2+\cdots+x_nD_n)^k f(\mathbf{a}) \tag{2.9}$$

Donde se entiende que la potencia k-ésima se desarrolla de forma simbólica. Un par de ejemplos aclarará esto.

$$\mathrm{d}^2(f,\mathbf{a},\mathbf{x})=(x_1D_1+x_2D_2+\cdots+x_nD_n)^2 f(\mathbf{a})=\left(\sum_{i,j=1}^{n}x_ix_jD_iD_j\right)f(\mathbf{a})=\sum_{i,j=1}^{n}x_ix_jD_{ij}f(\mathbf{a})$$

$$\mathrm{d}^3(f,\mathbf{a},\mathbf{x})=\left(\sum_{j=1}^{n}x_jD_j\right)^3 f(\mathbf{a})=\left(\sum_{i,j,k=1}^{n}x_ix_jx_kD_iD_jD_k\right)f(\mathbf{a})=\sum_{i,j,k=1}^{n}x_ix_jx_kD_{ijk}f(\mathbf{a})$$

Observa que $\mathrm{d}^1(f,\mathbf{a},\mathbf{x})=\langle\nabla f(\mathbf{a})|\mathbf{x}\rangle$.

2.21 Teorema (Taylor). *Sea f un campo escalar de clase $\mathcal{C}^{k+1}$ en un abierto Ω. Sean $\mathbf{a}\in\Omega$, $\mathbf{x}\in\mathbb{R}^n$ tales que el segmento $[\mathbf{a},\mathbf{a}+\mathbf{x}]$ está contenido en Ω. Entonces existe un punto $\mathbf{c}\in[\mathbf{a},\mathbf{a}+\mathbf{x}]$ tal que se verifica la igualdad:*

$$f(\mathbf{a}+\mathbf{x})=f(\mathbf{a})+\sum_{j=1}^{k}\frac{1}{j!}\mathrm{d}^j(f,\mathbf{a},\mathbf{x})+\frac{1}{(k+1)!}\mathrm{d}^{k+1}(f,\mathbf{c},\mathbf{x}) \tag{2.10}$$

Demostración. Tenemos que $[\mathbf{a},\mathbf{a}+\mathbf{x}]=\{\mathbf{a}+t\mathbf{x}:0\leqslant t\leqslant 1\}\subset\Omega$. Como el conjunto $\{t\in\mathbb{R}:\mathbf{a}+t\mathbf{x}\in\Omega\}$ es abierto (es la imagen inversa de Ω por la función continua $t\mapsto\mathbf{a}+t\mathbf{x}$)

y contiene al intervalo $[0,1]$, hay un intervalo abierto $I \supset [0,1]$ tal que $\mathbf{a}+t\mathbf{x}\in\Omega$ para todo $t\in I$. Definamos $\gamma: I\to\mathbb{R}^n$ por $\gamma(t)=\mathbf{a}+t\mathbf{x}$. Consideremos la función $h: I\to\mathbb{R}$ dada por $h(t)=(f\circ\gamma)(t)=f(\gamma(t))$. Por la proposición 2.8 dicha función es derivable y su derivada viene dada por:

$$h'(t)=\big\langle\nabla f(\gamma(t))\big|\gamma'(t)\big\rangle=\big\langle\nabla f(\gamma(t))\big|\mathbf{x}\big\rangle=\sum_{j=1}^{n}D_j f(\gamma(t))x_j=\mathrm{d}^1(f,\gamma(t),\mathbf{x})$$

Puesto que las funciones $D_j f$ tienen derivadas parciales continuas son diferenciables. Podemos calcular la derivada de la función $t\mapsto\big((D_j f)\circ\gamma\big)(t)$ volviendo a usar la igualdad (2.9) como hemos hecho antes pero con f sustituido por $D_j f$. Tenemos que:

$$\big((D_j f)\circ\gamma\big)'(t)=\big\langle\nabla(D_j f)(\gamma(t))\big|\mathbf{x}\big\rangle=\sum_{k=1}^{n}D_k(D_j f)(\gamma(t))x_k=\sum_{k=1}^{n}D_{kj}f(\gamma(t))x_k$$

Por tanto:

$$h''(t)=\sum_{j=1}^{n}\left(\sum_{k=1}^{n}D_{kj}f(\gamma(t))x_k\right)x_j=\sum_{j,k=1}^{n}D_{kj}f(\gamma(t))x_kx_j=\mathrm{d}^2(f,\gamma(t),\mathbf{x})$$

Análogamente se comprueba que $h'''(t)=\mathrm{d}^3(f,\gamma(t),\mathbf{x})$. En general, se tiene que $h^{(j)}(t)=\mathrm{d}^j(f,\gamma(t),\mathbf{x})$. Como las derivadas de orden k tienen derivadas parciales continuas son diferenciables, esto nos dice que h es $k+1$ veces derivable en I con derivada de orden $k+1$ continua en I. Podemos aplicar el teorema de Taylor a la función h para obtener que hay algún $\lambda\in[0,1]$ tal que:

$$h(1)=h(0)+\sum_{j=1}^{k}\frac{1}{j!}h^{(j)}(0)+\frac{1}{(k+1)!}h^{(k+1)}(\lambda)$$

Como $h(1)=f(\mathbf{a}+\mathbf{x})$, $h(0)=f(\mathbf{a})$, $h^{(j)}(0)=\mathrm{d}^j(f,\mathbf{a},\mathbf{x})$ y $h^{(k+1)}(\lambda)=\mathrm{d}^{k+1}(f,\mathbf{a}+\lambda\mathbf{x},\mathbf{x})$, la igualdad anterior es la misma (2.10) con $\mathbf{c}=a+\lambda\mathbf{x}\in[\mathbf{a},\mathbf{x}]$. □

2.22 Teorema (Taylor–Young). *Sea f un campo escalar definido en un abierto $\Omega\subset\mathbb{R}^n$ y supongamos que f tiene derivadas parciales de orden k continuas en Ω. Sea $\mathbf{a}\in\Omega$ y $r>0$ tal que $B(\mathbf{a},r)\subset\Omega$. Entonces para todo $\mathbf{x}$ con $\|\mathbf{x}\|<r$ se tiene que*

$$f(\mathbf{a}+\mathbf{x})=f(\mathbf{a})+\sum_{j=1}^{k}\frac{1}{j!}\mathrm{d}^j(f,\mathbf{a},\mathbf{x})+\|\mathbf{x}\|^k\varphi(\mathbf{x})\quad\textit{con}\ \lim_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0 \tag{2.11}$$

Demostración. Por comodidad de notación haremos la demostración para $k=2$. Tenemos que probar que:

$$f(\mathbf{a}+\mathbf{x})=f(\mathbf{a})+\sum_{k=1}^{n}D_kf(\mathbf{a})x_k+\frac{1}{2}\sum_{j,k=1}^{n}D_{jk}f(\mathbf{a})x_jx_k+\|\mathbf{x}\|^2\varphi(\mathbf{x})\quad\text{con}\ \lim_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0 \tag{2.12}$$

Para cada $\mathbf{x}$ con $0<\|\mathbf{x}\|<r$ se tiene que $\mathbf{a}+t\mathbf{x}\in B(\mathbf{a},r)$ para $0\leqslant t\leqslant 1$, por lo que $[\mathbf{a},\mathbf{a}+\mathbf{x}]\subset B(\mathbf{a},r)\subset\Omega$. Por el teorema de Taylor hay algún punto $\mathbf{c}\in[\mathbf{a},\mathbf{a}+\mathbf{x}]$ tal que

$$f(\mathbf{a}+\mathbf{x})=f(\mathbf{a})+\sum_{k=1}^{n}D_kf(\mathbf{a})x_k+\frac{1}{2}\sum_{j,k=1}^{n}D_{jk}f(\mathbf{c})x_jx_k$$

Lógicamente, el punto $\mathbf{c}$ dependerá de $\mathbf{x}$ y será de la forma $\mathbf{c}=\mathbf{a}+\lambda\mathbf{x}$ donde λ es un número en $[0,1]$ que dependerá de $\mathbf{x}$. Tenemos así que:

$$f(\mathbf{a}+\mathbf{x})=f(\mathbf{a})+\sum_{k=1}^{n}D_k f(\mathbf{a})x_k+\frac{1}{2}\sum_{j,k=1}^{n}D_{jk}f(\mathbf{a})x_jx_k+\frac{1}{2}\sum_{j,k=1}^{n}\big(D_{jk}f(\mathbf{c})-D_{jk}f(\mathbf{a})\big)x_jx_k$$

Definamos

$$\varphi(\mathbf{x})=\frac{1}{2\|x\|^2}\sum_{j,k=1}^{n}\big(D_{jk}f(\mathbf{c})-D_{jk}f(\mathbf{a})\big)x_jx_k$$

Con lo que, evidentemente, tenemos que:

$$f(\mathbf{a}+\mathbf{x})=f(\mathbf{a})+\sum_{k=1}^{n}D_k f(\mathbf{a})x_k+\frac{1}{2}\sum_{j,k=1}^{n}D_{jk}f(\mathbf{a})x_jx_k+\|\mathbf{x}\|^2\varphi(\mathbf{x})$$

Podemos suponer que la norma que usamos es la euclídea, con lo que $|x_j|\leqslant\|\mathbf{x}\|$ y tenemos:

$$|\varphi(\mathbf{x})|\leqslant\frac{1}{2}\sum_{j,k=1}^{n}\big|D_{jk}f(\mathbf{c})-D_{jk}f(\mathbf{a})\big|\frac{|x_i||x_j|}{\|x\|^2}\leqslant\frac{1}{2}\sum_{j,k=1}^{n}\big|D_{jk}f(\mathbf{c})-D_{jk}f(\mathbf{a})\big|$$

Cuando $\mathbf{x}\to\mathbf{0}$ se tiene que $\mathbf{c}=\mathbf{a}+\lambda\mathbf{x}\to\mathbf{a}$, por lo que, en virtud de la continuidad de las derivadas parciales de segundo orden, concluimos que $\lim\limits_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0$. □

El siguiente resultado, consecuencia directa del teorema anterior, es muy útil para calcular límites.

2.23 Corolario. *Sea f un campo escalar definido en un abierto $\Omega\subset\mathbb{R}^n$ y supongamos que f tiene derivadas parciales de orden k continuas en Ω. Supongamos también que todas las derivadas parciales de f de orden menor o igual que k se anulan en un punto $a\in\Omega$. Entonces se verifica que:*

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

2.24 Definición. Sea f un campo escalar de n variables que tiene derivadas parciales de segundo orden continuas en Ω y sea $\mathbf{a}\in\Omega$. La matriz $n\times n$

$$H(f,\mathbf{a})=\big(D_{ij}f(\mathbf{a})\big)_{1\leqslant i,j\leqslant n}$$

se llama **matriz hessiana** de f en $\mathbf{a}$.

Observa que la matriz hessiana es simétrica porque $D_{ij}f(\mathbf{a})=D_{ji}f(\mathbf{a})$. En consecuencia, dicha matriz define una forma cuadrática, que representaremos por $Q(f,\mathbf{a})$, que viene dada para todo $\mathbf{x}=(x_1,x_2,\dots,x_n)\in\mathbb{R}^n$ por

$$Q(f,\mathbf{a})(\mathbf{x})=\mathbf{x}.H(f,\mathbf{a}).\mathbf{x}^t=\sum_{j,k=1}^{n}D_{jk}f(\mathbf{a})x_jx_k$$

donde el punto "." indica producto matricial y $\mathbf{x}^t$ es el vector columna $\mathbf{x}$. Con esta notación podemos escribir la igualdad (2.12) en la forma

$$f(\mathbf{a}+\mathbf{x})=f(\mathbf{a})+\big\langle\nabla f(\mathbf{a})\big|\mathbf{x}\big\rangle+\frac{1}{2}Q(f,\mathbf{a})(\mathbf{x})+\|\mathbf{x}\|^2\varphi(\mathbf{x})\quad\text{donde}\quad\lim_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0\qquad(2.13)$$

Si suponemos que $\mathbf{a}$ es un punto crítico de f podemos escribir

$$f(\mathbf{a}+\mathbf{x}) = f(\mathbf{a}) + \frac{1}{2}Q(f,\mathbf{a})(\mathbf{x}) + \|\mathbf{x}\|^2\varphi(\mathbf{x}) \quad \text{donde} \quad \lim_{\mathbf{x}\to\mathbf{0}} \varphi(\mathbf{x}) = 0 \tag{2.14}$$

De donde se sigue que

$$\frac{f(\mathbf{a}+\mathbf{x}) - f(\mathbf{a})}{\|\mathbf{x}\|^2} = \frac{1}{2\|\mathbf{x}\|^2}Q(f,\mathbf{a})(\mathbf{x}) + \varphi(\mathbf{x}) \quad \text{donde} \quad \lim_{\mathbf{x}\to\mathbf{0}} \varphi(\mathbf{x}) = 0$$

Teniendo en cuenta que las formas cuadráticas son polinomios homogéneos de grado 2, es decir, $Q(f,\mathbf{a})(\lambda\mathbf{x}) = \lambda^2 Q(f,\mathbf{a})(\mathbf{x})$, se tiene que $\dfrac{1}{2\|\mathbf{x}\|^2}Q(f,\mathbf{a})(\mathbf{x}) = \dfrac{1}{2}Q(f,\mathbf{a})(\mathbf{x}/\|\mathbf{x}\|)$. Resulta así la igualdad

$$\frac{f(\mathbf{a}+\mathbf{x}) - f(\mathbf{a})}{\|\mathbf{x}\|^2} = \frac{1}{2}Q(f,\mathbf{a})(\mathbf{x}/\|\mathbf{x}\|) + \varphi(\mathbf{x}) \quad \text{donde} \quad \lim_{\mathbf{x}\to\mathbf{0}} \varphi(\mathbf{x}) = 0 \tag{2.15}$$

2.25 Definición. Una forma cuadrática $Q(\mathbf{x}) = \sum_{i,j=1}^{n} \alpha_{ij} x_i x_j$ se llama:

- **Definida positiva** si $Q(\mathbf{x}) > 0$ para todo $\mathbf{x} \in \mathbb{R}^n$ con $\mathbf{x} \neq \mathbf{0}$.
- **Semidefinida positiva** si $Q(\mathbf{x}) \geqslant 0$ para todo $\mathbf{x} \in \mathbb{R}^n$.
- **Definida negativa** si $Q(\mathbf{x}) < 0$ para todo $\mathbf{x} \in \mathbb{R}^n$ con $\mathbf{x} \neq \mathbf{0}$.
- **Semidefinida negativa** si $Q(\mathbf{x}) \leqslant 0$ para todo $\mathbf{x} \in \mathbb{R}^n$.
- **No definida** o **indefinida** si toma valores positivos y negativos.

2.26 Teorema. *Sea f un campo escalar definido en un abierto $\Omega \subset \mathbb{R}^n$ y supongamos que f tiene derivadas parciales de segundo orden continuas en Ω. Sea $\mathbf{a} \in \Omega$ un punto crítico de f y sea $Q(f,\mathbf{a})$ la forma cuadrática asociada a la matriz hessiana de f en $\mathbf{a}$.*

$$Q(f,\mathbf{a})(\mathbf{x}) = \mathbf{x}.H(f,\mathbf{a}).\mathbf{x}^t = \sum_{j,k=1}^{n} D_{jk} f(\mathbf{a}) x_j x_k$$

a) Si la forma cuadrática $Q(f,\mathbf{a})$ es definida positiva entonces f tiene en $\mathbf{a}$ un mínimo relativo estricto.

b) Si la forma cuadrática $Q(f,\mathbf{a})$ es definida negativa entonces f tiene en $\mathbf{a}$ un máximo relativo estricto.

c) Si f tiene un máximo relativo en $\mathbf{a}$ entonces la forma cuadrática $Q(f,\mathbf{a})$ es semidefinida negativa.

d) Si f tiene un mínimo relativo en $\mathbf{a}$ entonces la forma cuadrática $Q(f,\mathbf{a})$ es semidefinida positiva.

e) Si la forma cuadrática $Q(f,\mathbf{a})$ es no definida entonces f tiene un punto de silla en $\mathbf{a}$.

Demostración. Como $Q(f,\mathbf{a})$ es una función polinómica y, por tanto, continua, y la esfera unidad de $\mathbb{R}^n$, $S(\mathbf{0},1) = \{\mathbf{u} \in \mathbb{R}^n : \|\mathbf{u}\| = 1\}$, es un conjunto compacto, dicha función alcanza un mínimo valor y un máximo valor en $S(\mathbf{0},1)$. Sea

$$m = \text{mín}\{Q(f,\mathbf{a})(\mathbf{u}) : \|\mathbf{u}\| = 1\}, \quad M = \text{máx}\{Q(f,\mathbf{a})(\mathbf{u}) : \|\mathbf{u}\| = 1\}$$

a) Supongamos que $Q(f,\mathbf{a})$ es definida positiva. Entonces se tiene que $m>0$. y, por la igualdad (2.15), tenemos que

$$\frac{f(\mathbf{a}+\mathbf{x})-f(\mathbf{a})}{\|\mathbf{x}\|^2}=\frac{1}{2}Q(f,\mathbf{a})(\mathbf{x}/\|\mathbf{x}\|)+\varphi(\mathbf{x})\geqslant\frac{m}{2}+\varphi(\mathbf{x})\quad\text{donde}\quad\lim_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0$$

La condición $\lim\limits_{\mathbf{x}\to\mathbf{0}}\varphi(\mathbf{x})=0$ garantiza la existencia de un número $s>0$ tal que $|\varphi(\mathbf{x})|<m/4$ siempre que $0<\|\mathbf{x}\|<s$. En consecuencia, si en la desigualdad anterior suponemos que $0<\|\mathbf{x}\|<s$, se tiene

$$\frac{f(\mathbf{a}+\mathbf{x})-f(\mathbf{a})}{\|\mathbf{x}\|^2}\geqslant\frac{m}{2}+\varphi(\mathbf{x})>\frac{m}{2}-\frac{m}{4}=\frac{m}{4}>0$$

Deducimos que $f(\mathbf{a}+\mathbf{x})-f(\mathbf{a})>0$ para todo $\mathbf{x}$ con $0<\|\mathbf{x}\|<s$. O, lo que es igual, $f(\mathbf{z})-f(\mathbf{a})>0$ para todo $\mathbf{z}$ tal que $0<\|\mathbf{z}-\mathbf{a}\|<s$. Lo que prueba que f tiene en $\mathbf{a}$ un mínimo relativo estricto.

Los demás puntos se prueban de forma parecida. □

Para poder usar el resultado anterior hay que saber clasificar una forma cuadrática. Hay varios procedimientos sencillos para ello. Los dos que siguen a continuación son los que me parecen más cómodos.

Clasificación de formas cuadráticas

Sean $\mathcal{A}=\left(a_{ij}\right)_{1\leqslant i,j\leqslant n}$ una matriz simétrica de números reales y

$$Q_{\mathcal{A}}(\mathbf{x})=\mathbf{x}.\mathcal{A}.\mathbf{x}^t=\sum_{i,j=1}^{n}a_{ij}x_ix_j \tag{2.16}$$

la forma cuadrática definida por $\mathcal{A}$. Los *valores propios* de $\mathcal{A}$ son las raíces del polinomio característico $p(\lambda)$, que se define como el determinante de la matriz $\mathcal{A}-\lambda I$:

$$p(\lambda)=\left|\mathcal{A}-\lambda I\right|$$

Es sabido que, en la situación que estamos considerando, las raíces de dicho polinomio son todas reales.

Sean λ_j $(1\leqslant j\leqslant n)$ los valores propios de $\mathcal{A}$. Se demuestra que hay una base $\mathbf{B}=\{\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_n\}$ en $\mathbb{R}^n$ tal que para todo vector $\mathbf{x}\in\mathbb{R}^n$ se tiene que

$$Q_{\mathcal{A}}(\mathbf{x})=\sum_{j=1}^{n}\lambda_jx_j^2$$

donde $(x_1,x_2,\dots,x_n)$ son los coordenadas del vector $\mathbf{x}$ en la base $\mathbf{B}$. De aquí se siguen los siguientes criterios.

- La forma cuadrática $Q_{\mathcal{A}}$ es definida positiva si, y sólo si, todos los valores propios de $\mathcal{A}$ son positivos.
- La forma cuadrática $Q_{\mathcal{A}}$ es definida negativa si, y sólo si, todos los valores propios de $\mathcal{A}$ son negativos.

- La cuadrática $Q_{\mathcal{A}}$ es no definida si, y sólo si, $\mathcal{A}$ tiene valores propios positivos y negativos.
- La forma cuadrática $Q_{\mathcal{A}}$ es semidefinida positiva si, y sólo si, todos los valores propios de $\mathcal{A}$ son mayores o iguales que 0.
- La forma cuadrática $Q_{\mathcal{A}}$ es semidefinida negativa si, y sólo si, todos los valores propios de $\mathcal{A}$ son menores o iguales que 0.

Para aplicar estos criterios no es preciso calcular los valores propios de $\mathcal{A}$ sino solamente saber cuántos de ellos son positivos, negativos o nulos. Afortunadamente, hay un criterio que nos proporciona esta información sin más que observar los coeficientes del polinomio característico.

2.27 Proposición (Regla de los signos de Descartes). *Sea*

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

un polinomio con coeficientes reales y cuyas raíces son todas números reales. Se verifica entonces que:

a) El número de raíces positivas de f (contando multiplicidades) es igual al número de cambios de signo en la sucesión $(a_n, a_{n-1}, \ldots, a_1, a_0)$ de los coeficientes de f.

b) El número de raíces negativas de f (contando multiplicidades) es igual al número de cambios de signo en la sucesión $((-1)^n a_n, (-1)^{n-1} a_{n-1}, \ldots, -a_1, a_0)$ de los coeficientes de $f(-x)$.

Para contar los cambios de signo en la sucesión de coeficientes se saltan los coeficientes nulos. Por ejemplo, si $f(x) = 2x^6 + x^5 - x^3 + x^2 - 5$, la sucesión de coeficientes de f es $(2, 1, 0, -1, 1, 0, -1)$ cuyo número de cambios de signo es 3.

2.28 Corolario. *Sea $p(\lambda)$ el polinomio característico de la matriz hessiana de f en* **a**. *Entonces.*

- *Si $p(\lambda)$ tiene grado n, todos sus coeficientes son distintos de cero y tienen igual signo, se verifica que f tiene un máximo relativo estricto en* **a**.
- *Si $p(\lambda)$ tiene grado n, todos sus coeficientes son distintos de cero y van alternando su signo, se verifica que f tiene un mínimo relativo estricto en* **a**.

Otro criterio para estudiar el carácter de la forma cuadrática (2.16) es el *criterio de Sylvester* se basa en la sucesión de signos de los *menores principales* de la matriz $\mathcal{A}$. El menor principal de orden k es el determinante $\Delta_k = \left| a_{i,j} \right|_{1 \leqslant i,j \leqslant k}$. Se verifica que:

- *La forma cuadrática es definida positiva si, y sólo si, $\Delta_k > 0$ para $1 \leqslant k \leqslant n$.*
- *La forma cuadrática es definida negativa si, y sólo si, $(-1)^k \Delta_k > 0$ para $1 \leqslant k \leqslant n$.*

Observa que cuando la dimensión n es par, si el determinante de la matriz $\mathcal{A}$ es negativo entonces la forma es no definida.

Podemos particularizar este criterio para el caso de dos dimensiones.

Sea $A \subset \mathbb{R}^2$ un conjunto abierto y sea f un campo escalar definido en A que tiene derivadas parciales de segundo orden continuas. Supongamos que $(a,b)\in A$ es un punto crítico de f y sea

$$H(f,(a,b))=\begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2}(a,b) & \dfrac{\partial^2 f}{\partial x\partial y}(a,b) \\ \dfrac{\partial^2 f}{\partial x\partial y}(a,b) & \dfrac{\partial^2 f}{\partial y^2}(a,b) \end{pmatrix}$$

la matriz hessiana de f en (a,b) y notemos $\det H(f,(a,b))$ su determinante.

- Si $\det H(f,(a,b))>0$ y $\dfrac{\partial^2 f}{\partial x^2}(a,b)>0$ entonces f tiene en (a,b) un mínimo relativo estricto.
- Si $\det H(f,(a,b))>0$ y $\dfrac{\partial^2 f}{\partial x^2}(a,b)<0$ entonces f tiene en (a,b) un máximo relativo estricto.
- Si $\det H(f,(a,b))<0$ entonces f no tiene extremo relativo en (a,b). Se dice que (a,b) es un punto de silla de f.
- Cuando $\det H(f,(a,b))=0$ el conocimiento de la matriz hessiana no permite decidir si hay o no hay extremo relativo en (a,b). Cuando esto sucede puede ser interesante estudiar el comportamiento de las curvas $f(a,t+b)$ y $f(a+t,b)$. Si alguna de dichas curvas no tiene extremo relativo o tienen extremos relativos de distinta naturaleza en $t=0$, podemos concluir que en (a,b) no hay extremo relativo de f.

Ejercicios propuestos

117. Generaliza el teorema de Taylor – Young para funciones de clase $\mathcal{C}^k$.

118. Clasificar los puntos críticos de los campos escalares:

$$\begin{array}{ll}
f(x,y)=2xy-2x^3y-xy^2+x^3y^2; & f(x,y)=-2x^3-6xy^2+3x^2+3y^2; \\
f(x,y)=x^3+3xy^2-15x-12y; & f(x,y)=x^3+y^3-xy^2-x+16; \\
f(x,y)=-xy+2x^3y-xy^2-2x^3y^2; & f(x,y)=\cos(x)\cos(y) \\
f(x,y)=2x+y+x^2+xy+y^3; & f(x,y)=x^2y^2-x^2-y^2; \\
f(x,y)=\operatorname{sen}x+\operatorname{sen}y+\cos(x+y),\ x,y\in]0,2\pi[& f(x,y)=2x^4+y^4-4x^2-2y^2; \\
f(x,y,z)=x^2+y^2+z^2+3xyz-x-y-z; & f(x,y,z)=(x+y+z)\,\mathrm{e}^{1-x^2-y^2-z^2}
\end{array}$$

119. Trazar un plano que pase por el punto $(1,2,3)$ y que forme con los ejes coordenados un tetraedro de volumen mínimo (el volumen del tetraedro es un tercio del área de la base por la altura).

120. Calcula un punto (u,v,w) de coordenadas positivas de la esfera unidad $x^2+y^2+z^2=1$ tal que el plano tangente a la esfera en dicho punto determine con los ejes coordenados un tetraedro de volumen mínimo.

121. **Recta de mínimos cuadrados**. Dados n puntos $(x_i, y_i) \in \mathbb{R}^2$, determinar los números α y β para que la cantidad $\displaystyle\sum_{i=1}^{n} \left(y_i - \alpha x_i - \beta \right)^2$ sea mínima.

122. Dados m puntos $\mathbf{a}_i \in \mathbb{R}^n$, calcular el valor mínimo de la función $f(\mathbf{x}) = \displaystyle\sum_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|_2^2$.

123. Sea $f : \Omega \to \mathbb{R}$ un campo escalar de clase $\mathcal{C}^2$ en un abierto $\Omega \subset \mathbb{R}^n$. Sea $\mathbf{a} \in \Omega$ un punto crítico de f y supongamos que el determinante de la matriz hessiana de f en $\mathbf{a}$ es distinto de cero. Prueba que hay un abierto U que contiene a $\mathbf{a}$ tal que el único punto crítico de f en U es $\mathbf{a}$.

Capítulo 3

Derivación de campos vectoriales

3.1. Derivada de un campo vectorial. Matriz jacobiana

Una función vectorial es cualquier función que toma valores en un espacio vectorial de dimensión mayor que 1. Las curvas en el plano o en el espacio son funciones vectoriales de una variable. Ahora nos interesa considerar funciones vectoriales de varias variables.

La siguiente definición expresa la idea básica del cálculo diferencial: aproximar localmente una función por una aplicación lineal.

3.1 Definición. Sea $\mathbf{F}:\Omega\to\mathbb{R}^m$, donde $\Omega\subset\mathbb{R}^n$ es un abierto, y sea $\mathbf{a}\in\Omega$. Se dice que $\mathbf{F}$ es **diferenciable** o **derivable** en $\mathbf{a}$ si hay una aplicación lineal $\mathbf{T}:\mathbb{R}^n\to\mathbb{R}^m$ tal que:

$$\lim_{\mathbf{x}\to\mathbf{a}}\frac{\mathbf{F}(\mathbf{x})-\mathbf{F}(\mathbf{a})-\mathbf{T}(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|}=\mathbf{0} \tag{3.1}$$

Dicha aplicación lineal, si existe, es única; se llama **diferencial** o **derivada** de $\mathbf{F}$ en el punto $\mathbf{a}$ y se representa por $D\mathbf{F}(\mathbf{a})$.

La siguiente proposición pone las cosas mucho más claras y, de paso, probamos la unicidad de la aplicación lineal $\mathbf{T}$ que verifica (3.1).

3.2 Proposición. *Sea $\mathbf{F}=(f_1,f_2,\dots,f_m):\Omega\to\mathbb{R}^m$, donde $\Omega\subset\mathbb{R}^n$ es un abierto, una función vectorial de n variables y m componentes y sea $\mathbf{a}\in\Omega$. Equivalen las siguientes afirmaciones:*

a) $\mathbf{F}$ *es diferenciable en* $\mathbf{a}$.

b) Los campos escalares $f_1,f_2,\dots,f_m$ *componentes de* $\mathbf{F}$ *son diferenciables en* $\mathbf{a}$.

*Supuesto que se verifican a) y b), la matriz de la aplicación lineal $D\mathbf{F}(\mathbf{a})$ en las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$ es la matriz cuyas filas son los vectores gradiente $\nabla f_i(\mathbf{a})$, esto es la matriz de m filas y n columnas $\big(D_j f_i(\mathbf{a})\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$. Dicha matriz se llama **matriz jacobiana** de $\mathbf{F}$ en $\mathbf{a}$ y se representará por $J_{\mathbf{F}}(\mathbf{a})$.*

La aplicación lineal $D\mathbf{F}(\mathbf{a}):\mathbb{R}^n\to\mathbb{R}^m$ está definida para todo $\mathbf{x}\in\mathbb{R}^n$ por

$$D\mathbf{F}(\mathbf{a})(\mathbf{x}) = J_{\mathbf{F}}(\mathbf{a}).\mathbf{x}^t$$

*donde "**.**" indica producto matricial y $\mathbf{x}^t$ es el vector columna $\mathbf{x}$.*

Observa que en el caso particular de que $m=1$, es decir, $\mathbf{F}=f$ es un campo escalar, la aplicación derivada es una forma lineal sobre $\mathbb{R}^n$ que viene dada por:

$$Df(\mathbf{a})(\mathbf{x}) = \big\langle \nabla f(\mathbf{a}) \big| \mathbf{x} \big\rangle \qquad \forall \mathbf{x}\in\mathbb{R}^n$$

En el caso en que $n=1$, es decir, $\mathbf{F}=\gamma$ es una curva en $\mathbb{R}^m$, la aplicación derivada es una aplicación lineal de $\mathbb{R}$ en $\mathbb{R}^m$ que viene dada por:

$$D\gamma(a)(t) = t\gamma'(a) \qquad \forall t\in\mathbb{R}$$

Y en el caso en que $n=m=1$, esto es $\mathbf{F}=h$ es una función real de variable real, la aplicación derivada de h en a es la aplicación lineal de $\mathbb{R}$ en $\mathbb{R}$ dada por:

$$Dh(a)(t) = th'(a) \qquad \forall t\in\mathbb{R}$$

La idea básica del cálculo diferencial es aproximar (localmente) una función por su derivada, que es una aplicación lineal. Puesto que la aplicación derivada es una buena aproximación local de la función, cabe esperar que la función herede localmente algunas de las propiedades de su aplicación derivada. Las propiedades de una aplicación lineal son muy fáciles de estudiar y lo que se hace en casi todos los resultados del cálculo diferencial es trasladar localmente las propiedades de la aplicación derivada a la función. Un ejemplo aclarará esto. Considera una función real h derivable en un intervalo I; la aplicación derivada de h en cada punto $a\in I$ es la aplicación lineal de $\mathbb{R}$ en $\mathbb{R}$ $Dh(a)(t)=th'(a)$. Si para $a\in I$ es $h'(a)>0$ dicha aplicación es creciente, pero entonces también la función h es creciente en I. Este ejemplo puedes llevarlo a otras situaciones más complicadas: si la aplicación derivada es una biyección lineal ¿será la función localmente inversible? Claro está, la respuesta que se obtiene a este tipo de preguntas es siempre de tipo local porque la derivada es una buena aproximación local y solamente proporciona información local sobre el comportamiento de la función.

3.1.1. El espacio normado $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$

Antes de seguir conviene que consideremos el conjunto $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ de todas las aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$. Con las operaciones usuales de suma y producto por un número real, dicho conjunto es un espacio vectorial real de dimensión nm. Supongamos que en $\mathbb{R}^n$ y en $\mathbb{R}^m$ tenemos fijadas sendas normas que notaremos igual, $\|.\|$. Para $\mathbf{T}\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ definimos:

$$\|\mathbf{T}\| = \text{máx}\{\|\mathbf{T}(\mathbf{x})\| : \|\mathbf{x}\|=1\} \tag{3.2}$$

Se verifica entonces que la aplicación $\mathbf{T}\mapsto\|\mathbf{T}\|$ es una norma en $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$. Dicha norma se llama *norma de operadores asociada a las normas fijadas en* $\mathbb{R}^n$ *y en* $\mathbb{R}^m$. Es importante que te des cuenta de que en la igualdad 3.2 intervienen tres normas: la definida en $\mathbb{R}^n$, que aparece en la expresión $\|\mathbf{x}\|=1$, la definida en $\mathbb{R}^m$, que aparece en la expresión $\|\mathbf{T}(\mathbf{x})\|$ y la norma de operadores correspondiente $\|\mathbf{T}\|$. En adelante consideraremos siempre $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ como espacio normado con la norma de operadores asociada a las normas fijadas en $\mathbb{R}^n$ y en $\mathbb{R}^m$.

Ejercicios propuestos

124. Prueba que la la igualdad 3.2 define una norma en $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$.

125. Prueba que para todo $\mathbf{x}\in\mathbb{R}^n$ se verifica que

$$\|\mathbf{T}(\mathbf{x})\|\leqslant\|\mathbf{T}\|\|\mathbf{x}\| \tag{3.3}$$

126. Prueba que para $\mathbf{T}\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ se verifica la igualdad:

$$\|\mathbf{T}\|=\text{mín}\left\{M\in\mathbb{R}^+_o : \|\mathbf{T}(\mathbf{x})\|\leqslant M\|\mathbf{x}\|\ \forall\mathbf{x}\in\mathbb{R}^n\right\} \tag{3.4}$$

127. Supongamos que en $\mathbb{R}^n$ R^m y R^q tenemos fijadas normas y sean $\mathbf{S}\in\mathcal{L}(\mathbb{R}^m,\mathbb{R}^q)$ y $\mathbf{T}\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$. Prueba la desigualdad

$$\|\mathbf{S}\circ\mathbf{T}\|\leqslant\|\mathbf{S}\|\|\mathbf{T}\| \tag{3.5}$$

Todos los conceptos propios de espacios normados tienen perfecto sentido en el espacio $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$. De hecho, como tal espacio es de dimensión finita en él son válidos los teoremas 1.22, 1.24, 1.25 y 1.28. Por supuesto, tiene perfecto sentido considerar la convergencia de sucesiones en $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$.

Representaremos por $GL(\mathbb{R}^n)$ el conjunto de todas las aplicaciones lineales biyectivas (isomorfismos lineales) de $\mathbb{R}^n$ en $\mathbb{R}^n$. Notaremos por $\mathcal{L}(\mathbb{R}^n)$ el espacio vectorial de los operadores lineales de $\mathbb{R}^n$ en $\mathbb{R}^n$.

3.3 Proposición. *Sea* $\mathbf{T}\in\mathcal{L}(\mathbb{R}^n)$. *Equivalen las siguientes afirmaciones:*

a) $\mathbf{T}\in GL(\mathbb{R}^n)$.

b) $\mathbf{T}$ *es inyectivo.*

c) $\mathbf{T}$ *es sobreyectivo.*

3.4 Proposición.

a) $GL(\mathbb{R}^n)$ *es un conjunto abierto en* $\mathcal{L}(\mathbb{R}^n)$. *Concretamente:*

Para todo $\mathbf{T}\in GL(\mathbb{R}^n)$ se verifica que:

$$\left\{\mathbf{S}\in\mathcal{L}(\mathbb{R}^n) : \|\mathbf{T}-\mathbf{S}\| < \frac{1}{\|\mathbf{T}^{-1}\|}\right\} \subset GL(\mathbb{R}^n)$$

En particular, si $\mathbf{I}$ es la identidad en $\mathbb{R}^n$ entonces $B(\mathbf{I},1)\subset GL(\mathbb{R}^n)$.

b) La aplicación $\mathbf{T}\mapsto\mathbf{T}^{-1}$ de $GL(\mathbb{R}^n)$ en $GL(\mathbb{R}^n)$ es continua.

Ejercicios propuestos

128. Sea $\mathbf{T}\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$. ¿Cuál es la derivada de $\mathbf{T}$ en un punto $\mathbf{a}\in\mathbb{R}^n$?

129. Prueba que una función vectorial diferenciable en un punto $\mathbf{a}$ es continua en $\mathbf{a}$.

130. Prueba que un campo vectorial derivable en un dominio con derivada nula en todo punto del mismo es constante.

131. Supongamos que $\{\mathbf{S}_n\}\to\mathbf{S}$ en el espacio $\mathcal{L}(\mathbb{R}^m,\mathbb{R}^q)$ y que $\{\mathbf{T}_n\}\to\mathbf{T}$ en el espacio $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$. Prueba que $\{\mathbf{S}_n\circ\mathbf{T}_n\}\to\mathbf{S}\circ\mathbf{T}$ en el espacio $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^q)$.

132. Sea $\mathbf{a}\in\mathbb{R}^n$ y $\alpha:\mathbb{R}^n\to\mathbb{R}$ la aplicación lineal dada por $\alpha(\mathbf{x})=\langle\mathbf{a}|\mathbf{x}\rangle$. Calcula $\|\alpha\|$ cuando en $\mathbb{R}^n$ se consideran las normas $\|\ \|_2$, $\|\ \|_1$ y $\|\ \|_\infty$.

133. Sea $\mathbf{T}:\mathbb{R}^2\to\mathbb{R}^2$ la aplicación lineal de $(\mathbb{R}^2,\|\ \|_1)$ en sí mismo, cuya matriz en las bases canónicas es $\begin{pmatrix} a & b\\ c & d\end{pmatrix}$. Prueba que

$$\|\mathbf{T}\| = \text{máx}\{|a|+|c|,|b|+|d|\}$$

Si $\mathbf{F}$ es diferenciable en $\mathbf{a}$, definiendo:

$$R(\mathbf{a},\mathbf{x}) = \frac{\mathbf{F}(\mathbf{x})-\mathbf{F}(\mathbf{a})-D\mathbf{F}(\mathbf{a})(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|}$$

La igualdad

$$\lim_{\mathbf{x}\to\mathbf{0}} \frac{\mathbf{F}(\mathbf{x})-\mathbf{F}(\mathbf{a})-D\mathbf{F}(\mathbf{a})(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|} = \mathbf{0} \tag{3.6}$$

puede escribirse de forma equivalente como:

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{a})+D\mathbf{F}(\mathbf{a})(\mathbf{x}-\mathbf{a})+\|\mathbf{x}-\mathbf{a}\|R(\mathbf{a},\mathbf{x}) \quad \text{con} \quad \lim_{\mathbf{x}\to\mathbf{a}} R(\mathbf{a},\mathbf{x})=\mathbf{0} \tag{3.7}$$

3.5 Definición. Se dice que un campo vectorial es de clase $\mathcal{C}^k$ ($k\in\mathbb{N}\cup\{\infty\}$) en un abierto Ω si lo son sus campos escalares componentes.

3.6 Proposición. *Sea $\mathbf{F}=(f_1,f_2,\dots,f_n):\Omega\to\mathbb{R}^m$ donde Ω es un abierto en $\mathbb{R}^n$. Equivalen las afirmaciones:*

a) $\mathbf{F}$ es de clase $\mathcal{C}^1$ en Ω.

b) $\mathbf{F}$ es diferenciable en Ω y la aplicación $\mathbf{x}\mapsto D\mathbf{F}(\mathbf{x})$ de Ω en $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ es continua.

Además, como consecuencia de lo visto en el ejercicio 111, todo campo vectorial de clase $\mathcal{C}^k$ también es de clase $\mathcal{C}^q$ para $q\leqslant k$.

3.1.2. Regla de la cadena. Derivadas parciales de funciones compuestas

3.7 Teorema (Regla de la cadena). *Sean* $\mathbf{G}: A \to \mathbb{R}^n$ *y* $\mathbf{F}: B \to \mathbb{R}^m$*, donde* $A \subset \mathbb{R}^q$ *y* $B \subset \mathbb{R}^n$ *son conjuntos abiertos y* $\mathbf{G}(A) \subset B$*. Supongamos que* $\mathbf{G}$ *es diferenciable en un punto* $\mathbf{a} \in A$ *y que* $\mathbf{F}$ *es diferenciable en el punto* $\mathbf{G}(\mathbf{a}) \in B$*. Entonces se verifica que la función compuesta* $\mathbf{H} = \mathbf{F} \circ \mathbf{G}: A \to \mathbb{R}^m$ *es diferenciable en* $\mathbf{a}$*, y su diferencial viene dada como la composición de las respectivas diferenciales:*

$$D\mathbf{H}(\mathbf{a}) = D\mathbf{F}(\mathbf{G}(\mathbf{a})) \circ D\mathbf{G}(\mathbf{a}) \tag{3.8}$$

Además, si $\mathbf{F}$ *y* $\mathbf{G}$ *son de clase* $\mathcal{C}^k$ *entonces* $\mathbf{H}$ *también es de clase* $\mathcal{C}^k$*.*

Observa que $D\mathbf{G}(\mathbf{a}): \mathbb{R}^q \to \mathbb{R}^n$ y $D\mathbf{F}(\mathbf{G}(\mathbf{a})): \mathbb{R}^n \to \mathbb{R}^m$, por lo que la composición es una aplicación lineal de $\mathbb{R}^q$ a $\mathbb{R}^m$, como debe ser, pues $\mathbf{H}$ es una función vectorial de q variables y m componentes.

La expresión de la igualdad (3.8) por medio de matrices jacobianas es:

$$J_{\mathbf{H}}(\mathbf{a}) = J_{\mathbf{F}}(\mathbf{G}(\mathbf{a})).J_{\mathbf{G}}(\mathbf{a}) \tag{3.9}$$

Poniendo $\mathbf{H} = (h_1, h_2, \dots, h_m)$, $\mathbf{F} = (f_1, f_2, \dots, f_m)$, $\mathbf{G} = (g_1, g_2, \dots, g_n)$; notando las variables por $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, $\mathbf{y} = (y_1, y_2, \dots, y_q) \in \mathbb{R}^q$, y escribiendo $\mathbf{b} = \mathbf{G}(\mathbf{a})$, tenemos que:

$$\left(\frac{\partial h_i}{\partial y_j}(\mathbf{a})\right)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant q}} = \left(\frac{\partial f_i}{\partial x_k}(\mathbf{b})\right)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant k\leqslant n}} \cdot \left(\frac{\partial g_k}{\partial y_j}(\mathbf{a})\right)_{\substack{1\leqslant k\leqslant n\\ 1\leqslant j\leqslant q}} \qquad \mathbf{b} = \mathbf{G}(\mathbf{a})$$

De donde se sigue:

$$\frac{\partial h_i}{\partial y_j}(\mathbf{a}) = \sum_{k=1}^{n} \frac{\partial f_i}{\partial x_k}(\mathbf{b}) \frac{\partial g_k}{\partial y_j}(\mathbf{a}) \qquad \mathbf{b} = \mathbf{G}(\mathbf{a}) \qquad (1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant q) \tag{3.10}$$

Esta igualdad constituye la **regla de la cadena para derivadas parciales** y es importante que aprendas a aplicarla y que entiendas lo que dice. Voy a intentar facilitarte las cosas.

Primero, lo más frecuente es que $\mathbf{F}$ sea un campo escalar. Supongamos, pues, que en lo anterior, $\mathbf{F} = f$ es un campo escalar, en cuyo caso $h = f \circ \mathbf{G}$ también es un campo escalar. La igualdad (3.10) queda ahora:

$$\frac{\partial h}{\partial y_j}(\mathbf{a}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{b}) \frac{\partial g_k}{\partial y_j}(\mathbf{a}) \qquad \mathbf{b} = \mathbf{G}(\mathbf{a}) \qquad (1 \leqslant j \leqslant q) \tag{3.11}$$

En esta igualdad se interpreta que la función $\mathbf{G}: A \to B \subset \mathbb{R}^n$ lo que hace es un *"cambio de variables"*. Hablando familiarmente, podemos decir, que las "variables antiguas" de la función f, esto es las $\mathbf{x} = (x_1, x_2, \dots, x_n) \in B$ se han sustituido por "variable nuevas" $\mathbf{y} = (y_1, y_2, \dots, y_q) \in A$ y la función f se ha "expresado en estas nuevas variables" dando lugar a la función h. La relación entre unas variables y otras viene dada por:

$$x_k = g_k(y_1, y_2, \dots, y_q), \qquad 1 \leqslant k \leqslant n \tag{3.12}$$

De esta manera podemos interpretar la igualdad (3.11) en la forma siguiente:

> Para derivar la función nueva h, respecto a una nueva variable y_j, se deriva la función antigua f respecto a cada una de sus variables x_k y se multiplica por la derivada de cada una de ellas $x_k = g_k(y_1, y_2, \dots, y_q)$ respecto a la variable y_j.

Ya se ve que la situación está pidiendo que hagamos algunas simplificaciones que, además, son las que se hacen siempre en la práctica porque, aunque son algo confusas, facilitan mucho los cálculos.

Lo primero que se hace es identificar las funciones g_k que introducen las nuevas coordenadas con las coordenadas antiguas x_k, es decir, vemos las coordenadas antiguas como funciones de las nuevas y esto lo escribimos en la forma siguiente.

$$x_k = x_k(y_1, y_2, \dots, y_q), \qquad 1 \leqslant k \leqslant n \tag{3.13}$$

Con esta notación, la igualdad (3.11) queda como sigue.

$$\frac{\partial h}{\partial y_j}(\mathbf{a}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{b}) \frac{\partial x_k}{\partial y_j}(\mathbf{a}) \qquad \mathbf{b} = \mathbf{G}(\mathbf{a}) \qquad (1 \leqslant j \leqslant q) \tag{3.14}$$

Observa el doble papel que desempeña a la derecha de esta igualdad la letra x_k; cuando se deriva respecto de ella representa una variable y cuando ella se deriva respecto de una variable nueva representa una función.

La igualdad (3.14) ya es bastante fácil de recordar pero todavía se siguen haciendo en la práctica, sobre en todo en los textos de Física que suelen usar notaciones muy desafortunadas, algunas simplificaciones adicionales (y peligrosas). A saber: no se distingue entre la función f y la función h porque, como suele decirse en esos textos aludidos, son "*la misma función expresada en distintas variables*". Haciendo la identificación de f con h nos queda lo siguiente.

$$\frac{\partial f}{\partial y_j}(\mathbf{a}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{b}) \frac{\partial x_k}{\partial y_j}(\mathbf{a}) \qquad \mathbf{b} = \mathbf{G}(\mathbf{a}) \qquad (1 \leqslant j \leqslant q) \tag{3.15}$$

Aquí la letra f desempeña un doble papel: a la izquierda es la función compuesta y a la derecha es la función dada en sus variable iniciales.

Todavía suele darse un pasito más que consiste en representar la función f con una letra que suele usarse para representar variables; a saber, la letra z. Esto es frecuente también en textos de Física. Vamos a hacerlo así.

$$\frac{\partial z}{\partial y_j}(\mathbf{a}) = \sum_{k=1}^{n} \frac{\partial z}{\partial x_k}(\mathbf{b}) \frac{\partial x_k}{\partial y_j}(\mathbf{a}) \qquad \mathbf{b} = \mathbf{G}(\mathbf{a}) \qquad (1 \leqslant j \leqslant q) \tag{3.16}$$

Todavía hay algo que podemos simplificar. Habrás observado que siempre indico la relación que hay entre los puntos **b** y **a**. Eso es muy importante para entender lo que se hace. Hay que saber dónde se evalúan las derivadas parciales de cada función. Pues bien, eso no se indica *jamás* en textos de Física. Nunca se indica en dónde se evalúan las derivadas parciales. Así que vamos a suprimirlo.

$$\frac{\partial z}{\partial y_j} = \sum_{k=1}^{n} \frac{\partial z}{\partial x_k} \frac{\partial x_k}{\partial y_j} \qquad (1 \leqslant j \leqslant q) \tag{3.17}$$

Debes de familiarizarte con esta igualdad y saber reconocer en ella la igualdad de partida. Y no olvides la forma en que se evalúa esta igualdad. Lo vuelvo a poner.

$$\frac{\partial z}{\partial y_j}(\mathbf{y}) = \sum_{k=1}^{n} \frac{\partial z}{\partial x_k}(\mathbf{G}(\mathbf{y}))\frac{\partial x_k}{\partial y_j}(\mathbf{y}) \qquad (1 \leqslant j \leqslant q) \tag{3.18}$$

Si tuviéramos que volver a derivar en esta igualdad respecto a una variable y_k se derivaría como de costumbre: la derivada de una suma es la suma de las derivadas y para derivar un producto se aplica la regla usual. Pero hay un detalle muy importante y es que la función $\frac{\partial z}{\partial x_k}(\mathbf{G}(\mathbf{y}))$ *vuelve a ser la función compuesta del campo escalar* $\frac{\partial z}{\partial x_k}$ *con la función* $\mathbf{G}$. *Por tanto, para derivarla, hay que aplicarle la misma regla que hemos aplicado para derivar* z como función compuesta y que nos ha llevado a la igualdad anterior. Es por eso que el cálculo de derivadas parciales de segundo orden en funciones compuestas suele ser bastante engorroso y es fácil equivocarse si no se sabe lo que se hace.

3.8 Ejemplo. Vamos a calcular $\frac{\partial z}{\partial x}$ siendo $z = u^2 + v^5 + 3uv$ donde $u = x^2 + y^2$, $v = \operatorname{sen}(xy)$.

Así es como suelen enunciarse estos ejercicios y debes entender bien el enunciado. Nos están dando una función de las variables (u,v) a la que llaman z. Esto es la letra z representa una función, a saber, $z = u^2 + v^5 + 3uv$. Nos están dando un *cambio de variables* por medio de las igualdades $u = x^2 + y^2$, $v = \operatorname{sen}(xy)$. Y nos piden calcular $\frac{\partial z}{\partial x}$. Esto último ya nos dice claramente que debemos ver z como función de x e y, es decir, la letra z en $\frac{\partial z}{\partial x}$ es la función que nos dan *después de sustituir en ella las nuevas variables*, o sea, la función compuesta de $z = u^2 + v^5 + 3uv$ con $\mathbf{G}(x,y) = (x^2 + y^2, \operatorname{sen}(xy))$.

Sabemos que

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x} = (2u + 3v)2x + (5v^4 + 3u)y\cos(xy)$$

Si lo dejamos así escrito parece que $\frac{\partial z}{\partial x}$ depende de 4 variables. Pero no es así porque en la igualdad anterior las variables son x e y (las nuevas variables) mientras que u y v (las antiguas variables) vienen dadas por $u = x^2 + y^2$, $v = \operatorname{sen}(xy)$. Por tanto, es mejor hacer la sustitución, con lo que resulta

$$\frac{\partial z}{\partial x} = (2(x^2 + y^2) + 3\operatorname{sen}(xy))2x + (5\operatorname{sen}^4(xy) + 3x^2 + y^2)y\cos(xy)$$

que nos da el valor de la derivada parcial de la función compuesta en un punto (x,y). En este caso es muy sencillo calcular la función compuesta. Hazlo y comprueba el resultado obtenido. ♦

Ejercicios propuestos

Consideremos una función de dos variables x e y, $z = z(x,y)$, y supongamos que expresamos x e y en función de nuevas variables u y v, lo que indicamos en la forma $x = x(u,v)$,

$y = y(u,v)$. De esta forma la función z es función (función compuesta) de las "variables libres" u y v, a través de las "variables dependientes" x e y. Se trata de calcular las derivadas parciales de z respecto de las nuevas variables u y v. La regla para hacerlo es la siguiente: para derivar una función

$$z = z(x,y), \quad x = x(u,v),\ y = y(u,v)$$

respecto de una nueva variable, se deriva z respecto de cada una de las antiguas variables y se multiplica por la derivada de cada antigua variable respecto de la nueva variable. Se entiende mejor si lo escribimos simbólicamente

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}$$

En esta igualdad debes darte cuenta de que a la izquierda, como estamos derivando respecto a u, la letra z representa a la función compuesta $z = z(x(u,v), y(u,v))$ y la derivada está calculada en un punto (u,v). En la parte derecha de la igualdad la letra z representa la función dada $z = z(x,y)$ y las letras x e y representan variables (cuando se deriva respecto de ellas) y funciones (cuando se derivan respecto de u). Debe entenderse que cuando se sustituye un valor de (u,v) en la igualdad los valores de x e y deben sustituirse por $x = x(u,v)$, $y = y(u,v)$.

135. Sea $z = \cos(xy) + \mathrm{e}^{y-1}\cos x$ donde $x = u^2 + v$, $y = u - v^2$. Calcula $\dfrac{\partial z}{\partial u}$ en el punto $(u,v) = (1,1)$.

136. Sea $u = (x+y)^4 + y^2(z+x)^3$ donde $x = rs\,\mathrm{e}^{-t}$, $y = rs\log(1+t^2)$, $z = r^2 s\cos t$. Calcula $\dfrac{\partial u}{\partial s}$ cuando $r = 2$, $s = 1$, $t = 0$.

137. Sea $z = f(x,y)$, y pongamos $x = u^2 + v^2$, $y = u/v$. Calcula las derivadas parciales de de z respecto de las nuevas variables u y v en función de las derivadas parciales de z respecto de x e y.

138. Sea $u = x^4 y + y^2 z^3 + \varphi(x/y)$, donde

$$\begin{cases} x = 1 + rs\,\mathrm{e}^{t} \\ y = rs^2\,\mathrm{e}^{-t} \\ z = r^2 s \operatorname{sen} t \end{cases}$$

Calcula $\dfrac{\partial u}{\partial s}$ cuando $r = 2$, $s = 1$, $t = 0$, sabiendo que $\varphi'(3/2) = -1$.

139. Sea $z = f(x,y)$ donde $x = s^4 + r^4$, $y = 2rs^2$. Calcula $\dfrac{\partial z}{\partial x}(2,2)$ y $\dfrac{\partial z}{\partial y}(2,2)$. Siendo $\dfrac{\partial z}{\partial r}(1,1) = -2$ y $\dfrac{\partial z}{\partial s}(1,1) = 3$.

140. Prueba que la función $F(x,y) = f(\frac{y}{x^2-y^2})$, donde f es una función real derivable, verifica la igualdad

$$(x^2+y^2)\frac{\partial F}{\partial x} + 2xy\frac{\partial F}{\partial y} = 0$$

141. Prueba que la función $F(u,v) = f(uv,(u^2-v^2)/2)$, donde $f:\mathbb{R}^2\to\mathbb{R}$ es una función diferenciable, verifica la igualdad

$$(u^2+v^2)\left[\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2\right]=\left(\frac{\partial F}{\partial u}\right)^2+\left(\frac{\partial F}{\partial v}\right)^2$$

142. Sea $z = f(x,y)$, donde $x=\rho\cos\vartheta$, $y=\rho\operatorname{sen}\vartheta$. Calcula $\partial z/\partial\rho$ y $\partial z/\partial\vartheta$ y prueba que

$$\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2=\left(\frac{\partial z}{\partial \rho}\right)^2+\frac{1}{\rho^2}\left(\frac{\partial z}{\partial \vartheta}\right)^2$$

143. Sea $g(s,t) = f(s^2-t^2,t^2-s^2)$. Prueba la igualdad $t\dfrac{\partial g}{\partial s}+s\dfrac{\partial g}{\partial t}=0$.

144. Sea $u = f(x,y)$ donde $x=\mathrm{e}^s\cos t$, $y=\mathrm{e}^s\operatorname{sen} t$. Justifica que

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=\mathrm{e}^{-2s}\left(\frac{\partial^2 u}{\partial s^2}+\frac{\partial^2 u}{\partial t^2}\right)$$

145. Sea $z = f(x,y)$, donde $x=\rho\cos\vartheta$, $y=\rho\operatorname{sen}\vartheta$. Prueba que

$$\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial y^2}=\frac{\partial^2 z}{\partial \rho^2}+\frac{1}{\rho^2}\frac{\partial^2 z}{\partial \vartheta^2}+\frac{1}{\rho}\frac{\partial z}{\partial \rho}$$

146. Sea $z = f(x,y)$ donde $x=x(u,v)$, $y=y(u,v)$. Prueba que

$$\frac{\partial^2 z}{\partial u^2}=\frac{\partial^2 z}{\partial x^2}\left(\frac{\partial x}{\partial u}\right)^2+2\frac{\partial^2 z}{\partial x\partial y}\frac{\partial x}{\partial u}\frac{\partial y}{\partial u}+\frac{\partial^2 z}{\partial y^2}\left(\frac{\partial y}{\partial u}\right)^2+\frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial u^2}+\frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial u^2}$$

E indica la forma en que se evalúan estas funciones.

147. Sea $h(x,y) = f(u,v)$ donde f es de clase $\mathcal{C}^2$ y $u=u(x,y)$, $v=v(x,y)$ son funciones de clase $\mathcal{C}^2$ verificando que $\dfrac{\partial u}{\partial x}=\dfrac{\partial v}{\partial y}$ y $\dfrac{\partial u}{\partial y}=-\dfrac{\partial v}{\partial x}$. Prueba que:

$$\frac{\partial^2 h}{\partial x^2}+\frac{\partial^2 h}{\partial y^2}=\left(\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial v}{\partial x}\right)^2\right)\left(\frac{\partial^2 f}{\partial u^2}+\frac{\partial^2 f}{\partial v^2}\right)$$

148. Sean las funciones $f(x,y,z)=(\mathrm{e}^x+y^2,\lambda\mathrm{e}^z+y)$, $g(u,v)=\log u+v^2$ para $(u,v)\in\mathbb{R}^+\times\mathbb{R}$. ¿Qué valor debe tener λ para que la derivada direccional máxima de $g\circ f$ en $(0,0,0)$ sea igual a 1?

149. Sea $\mathbf{F}:\Omega\to\mathbb{R}^n$ un campo vectorial diferenciable e inyectivo en un abierto $\Omega\subset\mathbb{R}^n$. Sea $W=\mathbf{F}(\Omega)$ y supongamos que W es abierto y $\mathbf{F}^{-1}$ es diferenciable en W. Para $\mathbf{y}\in W$ expresa $D\mathbf{F}^{-1}(\mathbf{y})$ en función de la derivada de $\mathbf{F}$.

3.1.3. Teorema del valor medio para campos vectoriales

En el teorema 2.11 vimos una generalización del teorema del valor medio para campos escalares que se expresaba por la igualdad:

$$f(\mathbf{b})-f(\mathbf{a})=\big\langle\nabla f(\mathbf{c})\big|\mathbf{b}-\mathbf{a}\big\rangle \tag{3.19}$$

No existe una generalización del mismo tipo para funciones vectoriales. Por ejemplo, sea $\gamma(t)=(\cos t,\operatorname{sen} t)$. Tenemos que $\gamma'(t)=(-\operatorname{sen} t,\cos t)$ es un vector unitario y, como $\gamma(2\pi)-\gamma(0)=(0,0)$, no existe ningún $t\in\mathbb{R}$ que verifique la igualdad $\gamma(2\pi)-\gamma(0)=\gamma'(t)2\pi$, es decir *no podemos expresar el incremento de la función igual al valor de la derivada en un punto intermedio por el incremente de la variable*. Lo que se generaliza para funciones vectoriales no es una igualdad análoga a (3.19) sino una *desigualdad*: si tomamos valores absolutos en (3.19) obtenemos:

$$|f(\mathbf{b})-f(\mathbf{a})|\leqslant\|\nabla f(\mathbf{c})\|_2\|\mathbf{b}-\mathbf{a}\|_2 \tag{3.20}$$

Es esta desigualdad la que se generaliza para campos vectoriales con independencia, además, de las normas que se usen. Recuerda que la principal utilidad del teorema del valor medio es que permite acotar el incremento de una función por el incremento de la variable cuando se conoce una cota de la derivada, por ello nada se pierde con la generalización que vamos a obtener.

3.9 Lema. *Consideremos en $\mathbb{R}^n$ la norma euclídea. Dado un vector $\mathbf{u}\neq\mathbf{0}$ hay una aplicación lineal $\varphi:\mathbb{R}^n\to\mathbb{R}$ que verifica que $\varphi(\mathbf{u})=\|\mathbf{u}\|_2$ y $\|\varphi\|=1$.*

Demostración. Basta definir $\varphi:\mathbb{R}^n\to\mathbb{R}$ por

$$\varphi(\mathbf{x})=\left\langle\frac{\mathbf{u}}{\|\mathbf{u}\|_2}\middle|\mathbf{x}\right\rangle.$$

□

3.10 Teorema (Del valor medio para funciones vectoriales)**.** *Sea $\mathbf{F}:\Omega\to\mathbb{R}^m$ una función vectorial diferenciable en un abierto $\Omega\subset\mathbb{R}^n$. Consideraremos en $\mathbb{R}^n$ y en $\mathbb{R}^m$ las normas euclídeas. Supongamos que $[\mathbf{a},\mathbf{b}]\subset\Omega$ y que $\|D\mathbf{F}(\mathbf{c})\|\leqslant M$ para todo $\mathbf{c}\in[\mathbf{a},\mathbf{b}]$. Entonces se verifica que*

$$\|\mathbf{F}(\mathbf{b})-\mathbf{F}(\mathbf{a})\|_2\leqslant M\|\mathbf{b}-\mathbf{a}\|_2 \tag{3.21}$$

Demostración. Si $\mathbf{F}(\mathbf{b})-\mathbf{F}(\mathbf{a})=\mathbf{0}$ nada hay que probar. Supondremos que $\mathbf{F}(\mathbf{b})-\mathbf{F}(\mathbf{a})\neq\mathbf{0}$. Por el lema anterior sabemos que hay una aplicación lineal $\varphi:\mathbb{R}^m\to\mathbb{R}$ que verifica $\|\varphi\|=1$ y $\varphi\big(\mathbf{F}(\mathbf{b})-\mathbf{F}(\mathbf{a})\big)=\|\mathbf{F}(\mathbf{b})-\mathbf{F}(\mathbf{a})\|_2$. La aplicación $h=\varphi\circ\mathbf{F}:\Omega\to\mathbb{R}$ es un campo escalar diferenciable en Ω al que podemos aplicar el teorema 2.11 para obtener que existe $\mathbf{c}\in[\mathbf{a},\mathbf{b}]$ tal que

$$h(\mathbf{b})-h(\mathbf{a})=\big\langle\nabla h(\mathbf{c})\big|\mathbf{b}-\mathbf{a}\big\rangle=D(\varphi\circ\mathbf{F})(\mathbf{c})\big(\mathbf{b}-\mathbf{a}\big)=\varphi\circ D\mathbf{F}(\mathbf{c})(\mathbf{b}-\mathbf{a})$$

Tomando normas deducimos que

$$\|\mathbf{F}(\mathbf{b})-\mathbf{F}(\mathbf{a})\|_2\leqslant M\|\mathbf{b}-\mathbf{a}\|_2.$$

□

3.11 Corolario. *Un campo vectorial derivable en un dominio con derivada nula en todo punto del mismo es constante.*

3.2. Teorema de la función inversa

Un resultado que debes conocer de primer curso es el siguiente:

Teorema de inversión global para funciones reales. *Si $f: I \to \mathbb{R}$ es una función derivable en un intervalo I cuya derivada no se anula en ningún punto de I, entonces f es una biyección de I sobre el intervalo $J = f(I)$, y su inversa $g = f^{-1}: J \to \mathbb{R}$ es derivable en J siendo*

$$g'(y) = \frac{1}{f'(f^{-1}(y))} \qquad \forall y \in J$$

Recuerda que si f es una función real de una variable real derivable en un punto a, la aplicación derivada de f en a es la aplicación lineal de $\mathbb{R}$ en $\mathbb{R}$ dada por $t \mapsto f'(a)t$. Que dicha aplicación lineal sea invertible equivale a que $f'(a) \neq 0$.

También debes saber que si $\mathbf{T} \in \mathcal{L}(\mathbb{R}^n)$, las matrices que representan a $\mathbf{T}$ en distintas bases de $\mathbb{R}^n$ tienen el mismo determinante, dicho determinante se llama el *determinante de la aplicación lineal* $\mathbf{T}$ y lo representaremos por $\det(\mathbf{T})$. La aplicación lineal $\mathbf{T}$ es invertible equivale a que $\det(\mathbf{T}) \neq 0$.

Si $\mathbf{F}: \Omega \to \mathbb{R}^n$ es un campo vectorial definido en un abierto $\Omega \subset \mathbb{R}^n$ y diferenciable en un punto $\mathbf{a} \in \Omega$, el determinante de la derivada, $\det(D\mathbf{F}(\mathbf{a}))$, se llama el **jacobiano** de $\mathbf{F}$ en $\mathbf{a}$. Teniendo en cuenta lo antes dicho, el determinante jacobiano de $\mathbf{F}$ en $\mathbf{a}$ es el determinante de la matriz jacobiana $J_{\mathbf{F}}(\mathbf{a})$. La derivada $D\mathbf{F}(\mathbf{a})$ es invertible si, y sólo si, $\det(J_{\mathbf{F}}(\mathbf{a})) \neq 0$.

Pues bien, es fácil dar ejemplos de campos vectoriales de más de una variable que son diferenciables y cuya derivada en todo punto es invertible pero ellos no lo son. Es decir, el anterior teorema de inversión global para funciones reales de una variable real no puede generalizarse literalmente para campos vectoriales diferenciables de varias variables.

3.12 Ejemplo. El campo vectorial de $\mathbb{R}^2$ en $\mathbb{R}^2$ dado por $\mathbf{F}(x,y) = (\mathrm{e}^x \cos y, \mathrm{e}^x \operatorname{sen} y)$ es, evidentemente, diferenciable (es de clase $\mathcal{C}^\infty$) en $\mathbb{R}^2$. Su matriz jacobiana en un punto (x,y) es

$$J_{\mathbf{F}}(x,y) = \begin{pmatrix} \mathrm{e}^x \cos y & -\mathrm{e}^x \operatorname{sen} y \\ \mathrm{e}^x \operatorname{sen} y & \mathrm{e}^x \cos y \end{pmatrix}$$

cuyo determinante es igual a $\mathrm{e}^{2x} > 0$. Sin embargo, dicho campo no es inyectivo en $\mathbb{R}^2$ pues $\mathbf{F}(x, y+2\pi) = \mathbf{F}(x,y)$, por lo que la función $\mathbf{F}$ no tiene inversa. Observa, sin embargo, que la función $\mathbf{F}$ es inyectiva en $\mathbb{R} \times]0, 2\pi[$. ♦

Una versión local del teorema de inversión global es la siguiente.

Teorema de inversión local para funciones reales. *Si $f: I \to \mathbb{R}$ es una función derivable en un intervalo abierto I cuya derivada no se anula en un punto $a \in I$ y es continua en dicho punto. Entonces hay un intervalo abierto $H \subset I$ tal que $a \in H$, la restricción de f a H es una biyección de H sobre el intervalo $J = f(H)$, y su inversa $g = f^{-1}: J \to \mathbb{R}$ es derivable en J siendo*

$$g'(y) = \frac{1}{f'(f^{-1}(y))} \qquad \forall y \in J$$

Como vamos a ver, este teorema de inversión local sí puede generalizarse para campos vectoriales.

Sea, pues, $\mathbf{F}=(f_1,f_2,\dots,f_n):\Omega\to\mathbb{R}^n$ un campo vectorial diferenciable en un abierto $\Omega\subset\mathbb{R}^n$. Queremos dar condiciones que garanticen que para cada $\mathbf{y}=(y_1,y_2,\dots,y_n)\in\mathbf{F}(\Omega)$ el sistema de ecuaciones:

$$\begin{array}{rcl} f_1(x_1,x_2,\dots,x_n) & = & y_1 \\ f_2(x_1,x_2,\dots,x_n) & = & y_2 \\ \dots\dots\dots\dots\dots & & \dots \\ f_n(x_1,x_2,\dots,x_n) & = & y_n \end{array} \tag{3.22}$$

tiene solución única en $\mathbf{x}=(x_1,x_2,\dots,x_n)\in\Omega$. Dicho de otra forma, queremos saber si dicho sistema permite "despejar" $x_1,x_2,\dots,x_n$ en función de $\mathbf{y}=(y_1,y_2,\dots,y_n)$. Además, queremos que las soluciones sean funciones diferenciables de $\mathbf{y}$. Así planteado, la pregunta que nos estamos haciendo es si $\mathbf{F}$ es inyectiva en Ω. Acabamos de ver con un ejemplo que en general esto no es así: los campos vectoriales no suelen ser funciones inyectivas en dominios muy grandes. Además, las herramientas que vamos a usar proporcionan información sobre el comportamiento local de la función. Por eso, lo primero que hacemos es *localizar* el problema. Dado $\mathbf{a}\in\Omega$, nos preguntamos si hay conjuntos abiertos $U\subset\Omega$ y $V\subset\mathbb{R}^n$ con $a\in U$ tales que para todo $\mathbf{y}\in V$ las ecuaciones 3.22 tienen solución única $\mathbf{x}\in U$.

Naturalmente, las funciones f_j pueden ser de muy variada naturaleza y sería una ingenuidad suponer que hay algún tipo de algoritmo que permita resolver dicho sistema. Hay un caso en que sí sabemos resolverlo: cuando las ecuaciones son lineales. Pues bien, en situaciones como ésta lo que se hace es *linealizar el problema*, es decir, sustituimos la función $\mathbf{F}$ por su derivada y trasladamos el problema a dicha aplicación derivada, con lo cual hemos convertido nuestro problema inicial en un problema de álgebra lineal mucho más simple. En nuestro caso, la linealización del problema conduce al sistema de ecuaciones lineales $D\mathbf{F}(\mathbf{a})(\mathbf{x})=\mathbf{y}$:

$$\begin{array}{rcl} D_1f_1(\mathbf{a})x_1+D_2f_1(\mathbf{a})x_2+\cdots+D_nf_1(\mathbf{a})x_n & = & y_1 \\ D_1f_2(\mathbf{a})x_1+D_2f_2(\mathbf{a})x_2+\cdots+D_nf_2(\mathbf{a})x_n & = & y_2 \\ \dots\dots\dots\dots\dots & & \dots \\ D_1f_n(\mathbf{a})x_1+D_2f_n(\mathbf{a})x_2+\cdots+D_nf_n(\mathbf{a})x_n & = & y_n \end{array}$$

Para cada $\mathbf{y}\in\mathbb{R}^n$ dicho sistema tiene solución única en $\mathbf{x}$ si, y sólo si, el determinante jacobiano $\det(J_{\mathbf{F}}(\mathbf{a}))\neq 0$.

Un **entorno abierto** de un punto es un conjunto abierto que contiene a dicho punto.

3.13 Teorema (de la función inversa). *Sea* $\mathbf{F}=(f_1,f_2,\dots,f_n):\Omega\to\mathbb{R}^n$ *un campo vectorial de clase* $\mathcal{C}^1$ *en un abierto* $\Omega\subset\mathbb{R}^n$*. Sea* $\mathbf{a}\in\Omega$ *y supongamos que* $\det(J_{\mathbf{F}}(\mathbf{a}))\neq 0$*. Entonces existe un entorno abierto de* $\mathbf{a}$*,* $U\subset\Omega$*, tal que* $\mathbf{F}(U)=V$ *es abierto,* $\mathbf{F}$ *es una biyección de* U *sobre* V*, la biyección inversa* $\mathbf{G}:V\to U\subset\mathbb{R}^n$ *es de clase* $\mathcal{C}^1$ *en* V *y para todo* $\mathbf{y}\in V$ *es*

$$D\mathbf{G}(\mathbf{y})=\big(D\mathbf{F}(\mathbf{x})\big)^{-1}\quad \textit{donde } \mathbf{x}=\mathbf{G}(\mathbf{y}) \tag{3.23}$$

Si $\mathbf{F}$ *es de clase* $\mathcal{C}^k$ *en* Ω *entonces* $\mathbf{G}$ *es de clase* $\mathcal{C}^k$ *en* V*.*

Demostración. Supondremos fijada la norma euclídea en $\mathbb{R}^n$. Pongamos $\mathbf{L} = D\mathbf{F}(\mathbf{a})$. Por hipótesis, la aplicación $D\mathbf{F} : \Omega \to \mathcal{L}(\mathbb{R}^n)$ es continua. En particular, dicha aplicación es continua en $\mathbf{a} \in \Omega$. Por tanto, existe $r > 0$ tal que $B(\mathbf{a}, r) \subset \Omega$ y para todo $\mathbf{z} \in B(\mathbf{a}, r)$ se verifica que:

$$\|D\mathbf{F}(\mathbf{z}) - D\mathbf{F}(\mathbf{a})\| \leqslant \frac{1}{2\|\mathbf{L}^{-1}\|} \tag{3.24}$$

Aplicamos ahora el teorema del valor medio al campo vectorial $\mathbf{x} \mapsto \mathbf{F}(\mathbf{x}) - D\mathbf{F}(\mathbf{a})(\mathbf{x})$, cuya derivada en un punto $\mathbf{z} \in \Omega$ es $D\mathbf{F}(\mathbf{z}) - D\mathbf{F}(\mathbf{a})$, para obtener que para todos $\mathbf{x}, \mathbf{y} \in B(\mathbf{a}, r)$ se verifica que

$$\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y}) - D\mathbf{F}(\mathbf{a})\big(\mathbf{x} - \mathbf{y}\big)\|_2 \leqslant \frac{1}{2\|\mathbf{L}^{-1}\|}\|\mathbf{x} - \mathbf{y}\|_2$$

Deducimos que

$$\|\mathbf{L}\big(\mathbf{x} - \mathbf{y}\big)\|_2 - \|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})\|_2 \leqslant \frac{1}{2\|\mathbf{L}^{-1}\|}\|\mathbf{x} - \mathbf{y}\|_2.$$

Y teniendo en cuenta que $\dfrac{1}{\|\mathbf{L}^{-1}\|}\|\mathbf{x} - \mathbf{y}\|_2 \leqslant \|\mathbf{L}\big(\mathbf{x} - \mathbf{y}\big)\|_2$, obtenemos

$$\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y})\|_2 \geqslant \frac{1}{2\|\mathbf{L}^{-1}\|}\|\mathbf{x} - \mathbf{y}\|_2 \quad \forall \mathbf{x}, \mathbf{y} \in B(\mathbf{a}, r) \tag{3.25}$$

lo que implica que $\mathbf{F}$ es inyectiva en $U = B(\mathbf{a}, r)$.

Observa también que, en virtud de la proposición 3.4 la desigualdad 3.24 implica que la aplicación lineal $D\mathbf{F}(\mathbf{z})$ es inversible para todo $\mathbf{z} \in U$ y, por tanto, su determinante jacobiano no es nulo.

Seguidamente probaremos que la imagen por $\mathbf{F}$ de todo conjunto abierto contenido en U es un conjunto abierto en $\mathbb{R}^n$.

Sean $W \subset U$ un abierto y $\mathbf{v} \in \mathbf{F}(W)$. Sea $\mathbf{c} \in W$ tal que $\mathbf{F}(\mathbf{c}) = \mathbf{v}$. Tomemos $s > 0$ tal que la bola *euclídea* cerrada $\overline{B}(\mathbf{c}, s) \subset W$. Sea $\Delta = \mathrm{Fr}(\overline{B}(\mathbf{c}, s)) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{c}\|_2 = s\}$. La aplicación $h : \Delta \to \mathbb{R}$ dada por $h(\mathbf{x}) = \|\mathbf{F}(\mathbf{x}) - \mathbf{v}\|_2$ verifica que $h(\mathbf{x}) > 0$ para todo $\mathbf{x} \in \Delta$ y es continua. Por compacidad, h alcanza un mínimo absoluto en un punto $\mathbf{x}_0 \in \Delta$. Sea $m = h(\mathbf{x}_0) > 0$. Probaremos que $B(\mathbf{v}, m/2) \subset \mathbf{F}(B(\mathbf{c}, s))$ donde las bolas son bolas euclídeas.

Sea $\mathbf{y} \in B(\mathbf{v}, m/2)$ *fijo* en lo que sigue. Definimos la función $g : \overline{B}(\mathbf{c}, s) \to \mathbb{R}$ por $g(\mathbf{x}) = \|\mathbf{F}(\mathbf{x}) - \mathbf{y}\|_2$. Dicha aplicación es continua y, por compacidad, alcanza su mínimo absoluto en un punto $\mathbf{z} \in \overline{B}(\mathbf{c}, s)$. Como $g(\mathbf{c}) = \|\mathbf{v} - \mathbf{y}\|_2 < m/2$, tenemos que $g(\mathbf{z}) < m/2$. Para $\mathbf{x} \in \Delta$ se verifica que:

$$g(\mathbf{x}) = \|\mathbf{F}(\mathbf{x}) - \mathbf{y}\|_2 \geqslant \|\mathbf{F}(\mathbf{x}) - \mathbf{v}\|_2 - \|\mathbf{v} - \mathbf{y}\|_2 > m - \frac{m}{2} = \frac{m}{2}$$

Luego $g(\mathbf{x})$ no es el mínimo de g, es decir, $\mathbf{x} \neq \mathbf{z}$. Por tanto $\mathbf{z} \in B(\mathbf{c}, s)$.

Como g es una función que toma valores positivos, es claro que g^2 alcanza su mínimo en $\mathbf{z}$. Como

$$g^2(\mathbf{x}) = \|\mathbf{F}(\mathbf{x}) - \mathbf{y}\|_2^2 = \sum_{i=1}^{n}\big(f_i(\mathbf{x}) - y_i\big)^2$$

es una función diferenciable dicho punto debe ser un punto crítico de g^2. Deducimos que:

$$D_j g^2(\mathbf{z}) = 2\sum_{i=1}^{n} D_j f_i(\mathbf{z})(f_i(\mathbf{z}) - y_i) = 0 \qquad 1 \leqslant j \leqslant n$$

Igualdades que podemos expresar matricialmente como:

$$\big(J_{\mathbf{F}}(\mathbf{z})\big)^t . (\mathbf{F}(\mathbf{z}) - \mathbf{y})^t = 0$$

Como $\det(J_{\mathbf{F}}(\mathbf{z})) \neq 0$, concluimos que debe ser $\mathbf{F}(\mathbf{z}) = \mathbf{y}$. Como $\mathbf{y}$ es cualquier punto en $B(\mathbf{v}, m/2)$ y $\mathbf{z} \in B(\mathbf{c}, s)$, hemos probado que $\mathbf{F}(B(\mathbf{c}, s)) \supset B(\mathbf{v}, m/2)$ lo que implica que $B(\mathbf{v}, m/2) \subset \mathbf{F}(W)$. Por tanto $\mathbf{F}(W)$ es un conjunto abierto.

En particular $V = \mathbf{F}(U)$ es un conjunto abierto. La aplicación $\mathbf{F}$ es una biyección de U sobre V. Sea $\mathbf{G} : V \to U \subset \mathbb{R}^n$ la biyección inversa definida por:

$$\mathbf{G}(\mathbf{F}(\mathbf{x})) = \mathbf{x} \quad \forall \mathbf{x} \in U, \qquad \mathbf{F}(\mathbf{G}(\mathbf{y})) = \mathbf{y} \qquad \forall \mathbf{y} \in V$$

La desigualdad 3.25 nos dice ahora que:

$$\|\mathbf{G}(\mathbf{u}) - \mathbf{G}(\mathbf{v})\|_2 \leqslant 2\|\mathbf{L}^{-1}\| \|\mathbf{u} - \mathbf{v}\|_2 \quad \forall \mathbf{u}, \mathbf{v} \in V \tag{3.26}$$

lo que implica que $\mathbf{G}$ es continua en V.

Probaremos que $\mathbf{G}$ es diferenciable en V. Sea $\mathbf{v} \in V$ y $\mathbf{c} = \mathbf{G}(\mathbf{v}) \in U$. Pongamos $\mathbf{T} = D\mathbf{F}(\mathbf{c})$ y $\mathbf{S} = \mathbf{T}^{-1}$. La función $\varphi : \Omega \to \mathbb{R}$ definida por $\varphi(\mathbf{c}) = 0$, y para $\mathbf{x} \in U$, $\mathbf{x} \neq \mathbf{c}$, por la igualdad:

$$\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{c}) - \mathbf{T}(\mathbf{x} - \mathbf{c})\| = \varphi(\mathbf{x}) \|\mathbf{x} - \mathbf{c}\| \tag{3.27}$$

es continua en $\mathbf{c}$ porque $\mathbf{F}$ es diferenciable en $\mathbf{c}$. Pongamos $\mathbf{u} = \mathbf{F}(\mathbf{x}) \in V$ donde $\mathbf{x} \in U$, con lo que $\mathbf{G}(\mathbf{u}) = \mathbf{x}$. Tenemos:

$$\begin{aligned}
\|\mathbf{G}(\mathbf{u}) - \mathbf{G}(\mathbf{v}) - \mathbf{S}(\mathbf{u} - \mathbf{v})\|_2 &= \|\mathbf{S}\big(\mathbf{u} - \mathbf{v} - \mathbf{T}(\mathbf{G}(\mathbf{u}) - \mathbf{G}(\mathbf{v}))\big)\|_2 \leqslant \\
&\leqslant \|\mathbf{S}\| \|\mathbf{u} - \mathbf{v} - \mathbf{T}(\mathbf{G}(\mathbf{u}) - \mathbf{G}(\mathbf{v}))\|_2 = \text{ por (3.27) } = \\
&= \|\mathbf{S}\| \varphi(\mathbf{G}(\mathbf{u})) \|\mathbf{G}(\mathbf{u}) - \mathbf{G}(\mathbf{v})\|_2 \leqslant \text{ por (3.26) } \leqslant \\
&\leqslant \|\mathbf{S}\| \varphi(\mathbf{G}(\mathbf{u})) 2\|\mathbf{L}^{-1}\| \|\mathbf{u} - \mathbf{v}\|_2
\end{aligned}$$

Como, $\lim\limits_{\mathbf{u} \to \mathbf{v}} \varphi(\mathbf{G}(\mathbf{u})) = \varphi(\mathbf{c}) = 0$, deducimos que $\mathbf{G}$ es diferenciable en $\mathbf{v}$ y su derivada viene dada por $D\mathbf{G}(\mathbf{v}) = (D\mathbf{F}(\mathbf{c}))^{-1}$.

Como esto es válido para todo $\mathbf{v} \in V$ hemos probado que $\mathbf{G}$ es diferenciable en V y para todo punto $\mathbf{y} \in V$ es:

$$D\mathbf{G}(\mathbf{y}) = (D\mathbf{F}(\mathbf{G}(\mathbf{y})))^{-1} \qquad \forall \mathbf{y} \in V \tag{3.28}$$

Como la aplicación $\mathbf{y} \mapsto \mathbf{G}(\mathbf{y})$ de V en $U \subset \mathbb{R}^n$ es continua, la aplicación $\mathbf{x} \mapsto D\mathbf{F}(\mathbf{x})$ de U en $GL(\mathbb{R}^n)$ es continua y la aplicación $\mathbf{T} \mapsto \mathbf{T}^{-1}$ de $GL(\mathbb{R}^n)$ en $\mathcal{L}(\mathbb{R}^n)$ es continua, deducimos que la aplicación $\mathbf{y} \mapsto D\mathbf{G}(\mathbf{y})$ de V en $\mathcal{L}(\mathbb{R}^n)$ es continua en V, es decir $\mathbf{G}$ es de clase $\mathcal{C}^1$ en V.

Finalmente, probaremos que si $\mathbf{F}$ es de clase $\mathcal{C}^k$ también lo es $\mathbf{G}$. Acabamos de ver que esto es cierto para $k = 1$. Supongamos que es cierto para k y veamos que en tal caso también lo es para $k + 1$. Sea $\mathbf{F} = (f_1, f_2, \dots, f_n)$ de clase $\mathcal{C}^{k+1}$. Eso quiere decir que los campos escalares componentes de $\mathbf{F}$, esto es las funciones f_i, tienen derivadas parciales continuas de orden $k+1$

y, por tanto, las funciones $D_j f_i$ tienen derivadas parciales continuas de orden k. Teniendo en cuenta la igualdad (3.28) y la regla de Cramer para calcular la inversa de una matriz cuadrada, deducimos que las funciones $D_j g_i$ (donde g_i, $1 \leqslant i \leqslant n$, son los campos escalares componentes de $\mathbf{G}$) son funciones racionales de las funciones $D_j f_i \circ \mathbf{G}$ cuyo denominador es siempre el determinante de $J_{\mathbf{F}}(\mathbf{G}(\mathbf{y}))$ que sabemos no se anula para $\mathbf{y} \in V$. Como, por la hipótesis de inducción, $\mathbf{G}$ es de clase $\mathcal{C}^k$, las funciones $D_j f_i \circ \mathbf{G}$ son también de clase $\mathcal{C}^k$ y, siendo las funciones racionales de clase $\mathcal{C}^\infty$, concluimos, por la regla de la cadena, que las funciones $D_j g_i$ son de clase $\mathcal{C}^k$ en V, es decir $\mathbf{G}$ es de clase $\mathcal{C}^{k+1}$ en V. □

Se dice que una aplicación entre espacios métricos es una **aplicación abierta** si la imagen por dicha aplicación de cualquier abierto es un abierto.

Sean U, V abiertos en $\mathbb{R}^n$. Un **difeomorfismo** de U sobre V es una biyección de U sobre V que es diferenciable en U y cuya inversa es diferenciable en V. Si ambas aplicaciones son de clase $\mathcal{C}^k$ se dice que el difeomorfismo es de clase $\mathcal{C}^k$.

Sea $\mathbf{F} : \Omega \to \mathbb{R}^n$ un campo vectorial diferenciable en un abierto $\Omega \subset \mathbb{R}^n$. Se dice que $\mathbf{F}$ es un **difeomorfismo local** en Ω si para todo punto $\mathbf{x} \in \Omega$ hay un entorno abierto $U_{\mathbf{x}} \subset \Omega$ de dicho punto tal que $\mathbf{F}(U_{\mathbf{x}}) = V_{\mathbf{x}}$ es un abierto y $\mathbf{F}$ es un difeomorfismo de $U_{\mathbf{x}}$ sobre $V_{\mathbf{x}}$.

3.14 Corolario. *Sea* $\mathbf{F} : \Omega \to \mathbb{R}^n$ *un campo vectorial de clase* $\mathcal{C}^k$ *en un abierto* $\Omega \subset \mathbb{R}^n$ *y supongamos que* $\det J_{\mathbf{F}}(\mathbf{x}) \neq 0$ *para todo* $\mathbf{x} \in \Omega$. *Entonces* $\mathbf{F}$ *es un difeomorfismo local de clase* $\mathcal{C}^k$ *en* Ω *y es una aplicación abierta.*

Si, además, $\mathbf{F}$ *es inyectivo en* Ω *entonces es un difeomorfismo de clase* $\mathcal{C}^k$ *de* Ω *sobre* $\mathbf{F}(\Omega)$.

Un difeomorfismo $\mathbf{F}$ de un abierto $U \subset \mathbb{R}^n$ sobre un abierto $V \subset \mathbb{R}^n$ suele interpretarse como una función que define nuevas coordenadas en V; de hecho, a veces se llama a los difeomorfismos *cambios de variables* o *cambios de coordenadas*: las coordenadas según $\mathbf{F}$ de un punto $\mathbf{y} \in V$ son las coordenadas cartesianas del punto $\mathbf{x} \in U$ tal que $\mathbf{F}(\mathbf{x}) = \mathbf{y}$.

3.15 Ejemplo (Coordenadas polares en el plano)**.** El campo vectorial $\mathbf{F} : \mathbb{R}^+ \times]-\pi, \pi[\to \mathbb{R}^2$ dado por $\mathbf{F}(\rho, \vartheta) = (\rho\cos\vartheta, \rho\operatorname{sen}\vartheta)$, es inyectivo en $\Omega = \mathbb{R}^+ \times]-\pi, \pi[$. Como $\mathbf{F}$ es de clase $\mathcal{C}^1$ (es, de hecho, de clase $\mathcal{C}^\infty$) y $\det J_{\mathbf{F}}(\rho, \vartheta) = \rho > 0$, concluimos que $\mathbf{F}$ es un difeomorfismo de clase $\mathcal{C}^1$ de Ω sobre $\mathbf{F}(\Omega) = \mathbb{R}^2 \setminus \{(x, 0) : x \leqslant 0\}$. Dado $(x, y) \in \mathbf{F}(\Omega)$ el par $(\rho, \vartheta) \in \Omega$ tal que $\mathbf{F}(\rho, \vartheta) = (x, y)$ son las *coordenadas polares* de (x, y). ♦

Una consecuencia interesante del teorema de la función inversa es la siguiente.

3.16 Proposición. *Sea* $\mathbf{F} : \Omega \to \mathbb{R}^m$ *una función de clase* $\mathcal{C}^1$ *en un abierto* $\Omega \subset \mathbb{R}^n$. *Supongamos que la derivada de* $\mathbf{F}$ *en todo punto de* Ω *es sobreyectiva. Entonces* $\mathbf{F}$ *es una aplicación abierta.*

Demostración. Sean $\mathbf{a} \in \Omega$ y $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m\}$ la base canónica de $\mathbb{R}^m$. Como $D\mathbf{F}(\mathbf{a}) : \mathbb{R}^n \to \mathbb{R}^m$ es sobreyectiva, existen vectores $\mathbf{u}_i \in \mathbb{R}^n$ tales que $D\mathbf{F}(\mathbf{a})(\mathbf{u}_i) = \mathbf{e}_i$, $1 \leqslant i \leqslant m$. Definamos $\mathbf{S} : \mathbb{R}^m \to \mathbb{R}^n$ por:

$$\mathbf{S}\left(\sum_{i=1}^{m} \lambda_i \mathbf{e}_i\right) = \sum_{i=1}^{m} \lambda_i \mathbf{u}_i$$

Claramente, $\mathbf{S}$ es lineal y $D\mathbf{F}(\mathbf{a}) \circ \mathbf{S}(\mathbf{e}_i) = \mathbf{e}_i$, $1 \leqslant i \leqslant m$. Por tanto $D\mathbf{F}(\mathbf{a}) \circ \mathbf{S} = \mathbf{I}_{\mathbb{R}^m}$ (la identidad en $\mathbb{R}^m$). Definamos $\mathbf{S}_{\mathbf{a}} = \tau_{\mathbf{a}} \circ \mathbf{S}$ donde $\tau_{\mathbf{a}} : \mathbb{R}^n \to \mathbb{R}^n$ es la traslación de vector $\mathbf{a}$: $\tau_{\mathbf{a}}(\mathbf{x}) = \mathbf{a} + \mathbf{x}$

para todo $\mathbf{x} \in \mathbb{R}^n$. Como $\mathbf{a} = \mathbf{S_a}(\mathbf{0}) \in \Omega$, el conjunto $W_\mathbf{a} = \mathbf{S}_\mathbf{a}^{-1}(\Omega)$ es un entorno abierto de $\mathbf{0}$ en $\mathbb{R}^m$.

Podemos aplicar ahora el teorema de la función inversa a la función $\mathbf{G_a} : W_\mathbf{a} \to \mathbb{R}^m$ definida por $\mathbf{G_a}(\mathbf{x}) = \mathbf{F} \circ \mathbf{S_a}(\mathbf{x})$ para todo $\mathbf{x} \in W_\mathbf{a}$. Claramente $\mathbf{G_a}$ es de clase $\mathcal{C}^1$ en $W_\mathbf{a}$ y $D\mathbf{G_a}(\mathbf{0}) = D\mathbf{F}(\mathbf{a}) \circ \mathbf{S} = \mathbf{I}_{\mathbb{R}^m}$. Obtenemos así que hay un entorno abierto $V_\mathbf{a} \subset W_\mathbf{a}$ de $\mathbf{0}$ tal que $\mathbf{G_a}$ es un difeomorfismo de $V_\mathbf{a}$ sobre el abierto $\mathbf{G_a}(V_\mathbf{a})$.

Sea $U \subset \mathbf{\Omega}$ un conjunto abierto y sea $\mathbf{a} \in U$. El conjunto $U_\mathbf{a} = \mathbf{S}_\mathbf{a}^{-1}(U) \cap V_\mathbf{a}$ es un entorno abierto de $\mathbf{0}$ contenido en $V_\mathbf{a}$, por lo que $\mathbf{G_a}\big(U_\mathbf{a}\big)$ es abierto y $\mathbf{G_a}(\mathbf{0}) = \mathbf{F}(\mathbf{a}) \in \mathbf{G_a}\big(U_\mathbf{a}\big)$. Como:

$$\mathbf{G_a}\big(U_\mathbf{a}\big) \subset \mathbf{G_a}\big(\mathbf{S}_\mathbf{a}^{-1}(U)\big) = (\mathbf{F} \circ \mathbf{S_a})\big(\mathbf{S}_\mathbf{a}^{-1}(U)\big) = \mathbf{F}\left(\mathbf{S_a}\big(\mathbf{S}_\mathbf{a}^{-1}(U)\big)\right) \subset \mathbf{F}(U)$$

Se sigue que $\mathbf{F}(\mathbf{a})$ es un punto interior de $\mathbf{F}(U)$. Luego todos los puntos de $\mathbf{F}(U)$ son interiores y $\mathbf{F}(U)$ es abierto. □

3.3. Teorema de la función implícita

Sea $f(x,y)$ una función de dos variables con derivadas parciales de primer orden continuas y consideremos la ecuación $f(x,y) = 0$. Las soluciones de dicha ecuación representan una curva en el plano. Bueno, hablando con propiedad pueden representar algo más general que una curva. Para que te convenzas de ello basta que consideres la ecuación

$$f(x,y) = (x^2 + y^2 - 1)(2(x-1)^2 + 3(y-2)^2 - 1)(y - x^2) = 0$$

la función f se anula en los puntos de la circunferencia $x^2 + y^2 = 1$, de la parábola $y = x^2$ y de la elipse $2(x-1)^2 + 3(y-2)^2 = 1$. Por tanto la ecuación $f(x,y) = 0$ representa la unión de todas esas curvas.

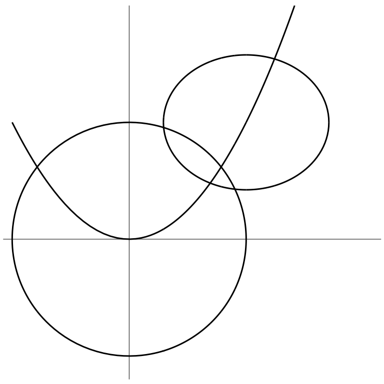


Figura 3.1. Conjunto dado por $f(x,y) = 0$

Ese conjunto (ver figura (3.1)) no es exactamente una curva pero ***localmente*** se parece a una curva. La palabra "localmente" quiere decir que si fijamos un punto (a,b) tal que $f(a,b) = 0$ entonces hay una bola abierta $B((a,b),r)$ tal que el corte de dicha bola con el conjunto de puntos $M = \{(x,y) : f(x,y) = 0\}$ es una curva. De hecho, no es cierto que la condición anterior se verifique para todos los puntos (a,b) tales que $f(a,b) = 0$. Dicha condición falla en los puntos donde se cortan dos de las curvas cuya unión forma M, pues es claro que en dichos puntos el conjunto M no parece localmente una curva. Pues bien, en dichos puntos se anula el vector gradiente de f y en ellos la recta tangente no está definida. Este ejemplo te ayudará a entender lo que sigue.

Volvamos al caso general de una función de dos variables $f(x,y)$ con derivadas parciales continuas de primer orden. Consideremos ahora la ecuación $f(x,y)=0$ desde otro punto de vista. Intuitivamente, *una* ecuación es *una* condición que debe ligar a *una* de las variables, es decir, que si en la igualdad $f(x,y)=0$ se fija un valor de x entonces el valor de y queda determinado de manera única por dicho valor de x. A veces esto es verdad como en el siguiente ejemplo. Consideremos

$$f(x,y)=y^3+y\,\mathrm{e}^x+\operatorname{sen} x$$

Fijado un valor de x la ecuación $f(x,y)=0$ es un polinomio de tercer grado en y que tiene una única solución real pues su derivada respecto de y es $3y^2+\mathrm{e}^x$ que no se anula. Es decir, en este caso es cierto que la igualdad

$$y^3+y\,\mathrm{e}^x+\operatorname{sen} x=0 \tag{3.29}$$

define de manera única a y como función de x, en el sentido de que fijado un valor de x, hay un único $y=\varphi(x)$ que verifica dicha igualdad, esto es, la función $\varphi(x)$ está definida por la condición:

$$\varphi(x)^3+\varphi(x)\,\mathrm{e}^x+\operatorname{sen} x=0 \tag{3.30}$$

Se dice que la función φ **está implícitamente definida** por la igualdad (3.29). Puedes calcular con *wxMaxima* el valor de dicha función y comprobarás que es bastante complicada. El hecho es que la mejor forma de trabajar con la función φ es la igualdad (3.30) que la define. Por ejemplo, si queremos calcular la derivada de φ en un punto basta con que derivemos dicha igualdad para obtener

$$3\varphi'(x)\varphi(x)^2+\varphi'(x)\,\mathrm{e}^x+\varphi(x)\,\mathrm{e}^x+\cos x=0$$

lo que permite calcular $\varphi'(x)$ en función de $\varphi(x)$.

En general, no es cierto que una igualdad de la forma $f(x,y)=0$ permita despejar una variable en función de la otra. Para convencerte, considera el primer ejemplo que pusimos. Ni tan siquiera una igualdad tan sencilla como $x^2+y^2-1=0$ permite despejar una variable como función de la otra pues es claro que para cada valor que fijemos de una variable (comprendido entre -1 y 1) hay *dos* posibles valores de la otra que verifican dicha igualdad.

Relacionemos ahora los dos puntos de vista que hemos considerado. Pongamos

$$\Gamma=\left\{(x,y)\in\mathbb{R}^2 : f(x,y)=0\right\}$$

Si la igualdad $f(x,y)=0$ permitiera despejar y en función de x, es decir, definiera una función $y=\varphi(x)$ por la condición

$$f(x,y)=0 \Longleftrightarrow y=\varphi(x)$$

entonces se tendría que (llamando I al intervalo donde está definida φ)

$$\Gamma=\left\{(x,y)\in\mathbb{R}^2 : f(x,y)=0\right\}=\{(x,\varphi(x)) : x\in I\}$$

es decir, el conjunto Γ sería la gráfica de φ, que, como sabemos, es un tipo muy particular de curva. Pero ya hemos visto que el conjunto Γ puede ser una “curva” mucho más general que la gráfica de una función. Pero incluso en este caso, dicha “curva” es *localmente*, excepto en los puntos donde se anula el gradiente, una gráfica de una función.

Las consideraciones anteriores se pueden llevar al caso de una función de tres variables $f(x,y,z)$ considerando ahora la “superficie” definida por la ecuación $f(x,y,z)=0$. La pregunta ahora es si fijados un valor de x y otro de y queda determinado de manera única un valor de $z=\varphi(x,y)$ que verifica dicha ecuación. En caso afirmativo tendríamos que la superficie de ecuación $f(x,y,z)=0$ coincidiría con la gráfica de φ. Ya puedes suponer que esto no es cierto en general pues la mayoría de las “superficies” no son gráficas de funciones.

Consideremos ahora el caso general en que tenemos una función vectorial $\mathbf{F}:\Omega\to\mathbb{R}^m$ de clase $\mathcal{C}^1$ definida en un abierto $\Omega\subset\mathbb{R}^{n+m}$. Estamos interesados en estudiar el conjunto donde se anula $\mathbf{F}$, es decir, el conjunto $M=\{\mathbf{x}\in\Omega:\mathbf{F}(\mathbf{x})=\mathbf{0}\}$. Dicho conjunto está determinado por el sistema de m ecuaciones:

$$\begin{array}{rcl} f_1(x_1,x_2,\dots,x_n,x_{n+1},\dots,x_{n+m}) & = & 0\\ f_2(x_1,x_2,\dots,x_n,x_{n+1},\dots,x_{n+m}) & = & 0\\ \dots\dots\dots\dots\dots\dots\dots\dots\dots & & \dots\\ f_m(x_1,x_2,\dots,x_n,,x_{n+1},\dots,x_{n+m}) & = & 0 \end{array}\tag{3.31}$$

En Mecánica estas ecuaciones se interpretan como “ligaduras” que deben satisfacer las variables. Tenemos en total m ligaduras que, intuitivamente, deben “atar” a m de las variables en el sentido de que el valor de las mismas quede determinado por el valor de las otras n. Por comodidad, consideremos las últimas m variables. La pregunta que nos hacemos es si el sistema de ecuaciones (3.31) permite “despejar” $(x_{n+1},x_{n+2},\dots,x_{n+m})$ en función de $(x_1,x_2,\dots,x_n)$. Para responder a esta pregunta lo primero que hacemos, siguiendo una estrategia ya conocida, es localizar el problema. La siguiente notación es cómoda: si $\mathbf{x}=(x_1,x_2,\dots,x_n)\in\mathbb{R}^n$ y $\mathbf{y}=(y_1,y_2,\dots,y_m)\in\mathbb{R}^m$, pondremos $(\mathbf{x},\mathbf{y})=(x_1,x_2,\dots,x_n,y_1,y_2,\dots,y_m)\in\mathbb{R}^{n+m}$, es decir, identificamos $\mathbb{R}^{n+m}$ con $\mathbb{R}^n\times\mathbb{R}^m$. Supongamos que $(\mathbf{a},\mathbf{b})\in\Omega$ satisface el sistema de ecuaciones (3.31), es decir $\mathbf{F}(\mathbf{a},\mathbf{b})=\mathbf{0}$. Nos preguntamos si hay entornos abiertos, $A\subset\mathbb{R}^n$ de $\mathbf{a}$ y $B\subset\mathbb{R}^m$ de $\mathbf{b}$, tales que para todo $\mathbf{x}\in A$ existe un único $\mathbf{y}\in B$ tal que $(\mathbf{x},\mathbf{y})$ sea solución de (3.31). Cuando esto es así, existe una función $\varphi:A\to B$ que a cada $\mathbf{x}\in A$ hace corresponder el único $\mathbf{y}\in B$ que verifica que es solución de (3.31), es decir, $\mathbf{F}(\mathbf{x},\varphi(\mathbf{x}))=\mathbf{0}$. En tal caso, también se tiene que $M\cap(A\times B)=\{(\mathbf{x},\varphi(\mathbf{x})):\mathbf{x}\in A\}$, es decir, el conjunto $M\cap(A\times B)$ es la gráfica de la función φ.

El siguiente paso es linealizar el problema, esto es, sustituir la función $\mathbf{F}$ por su derivada en $(\mathbf{a},\mathbf{b})$ y trasladar el problema que nos ocupa a dicha derivada. El sistema de ecuaciones lineales que resulta escrito en forma matricial es:

$$J_{\mathbf{F}}(\mathbf{a},\mathbf{b}).(\mathbf{x},\mathbf{y})^t=\mathbf{0}$$

Se trata de un sistema de m ecuaciones lineales con $n+m$ incógnitas y lo que queremos saber es si dicho sistema permite despejar $\mathbf{y}$ en función de $\mathbf{x}$. La respuesta es conocida: basta exigir que el determinante de las últimas m columnas de la matriz $J_{\mathbf{F}}(\mathbf{a},\mathbf{b})$ sea distinto de cero. Llamando J_1 la matriz de las primeras n columnas de $J_{\mathbf{F}}(\mathbf{a},\mathbf{b})$ y J_2 a la matriz de las últimas m columnas tenemos que:

$$J_{\mathbf{F}}(\mathbf{a},\mathbf{b}).(\mathbf{x},\mathbf{y})^t = J_1.\mathbf{x}^t + J_2.\mathbf{y}^t = \mathbf{0} \Longleftrightarrow \mathbf{y}^t = -(J_2^{-1}.J_1).\mathbf{x}^t$$

Después de lo que antecede, el teorema que sigue es lo que se espera.

3.17 Teorema (de la función implícita). *Sea* $\mathbf{F} = (f_1, f_2, \dots, f_m) : \Omega \to \mathbb{R}^m$ *una función de clase* $\mathcal{C}^1$ *definida en un abierto* $\Omega \subset \mathbb{R}^{n+m}$. *Sea* $(\mathbf{a},\mathbf{b}) \in \Omega$ *tal que* $\mathbf{F}(\mathbf{a},\mathbf{b}) = \mathbf{0}$ *y*

$$\det\big(D_{n+j} f_i(\mathbf{a},\mathbf{b})\big)_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant m}} \neq 0$$

Entonces existen un entorno abierto $W \subset \Omega$ *de* $(\mathbf{a},\mathbf{b})$, *un entorno abierto* U *de* $\mathbf{a}$ *y una función* $\varphi : U \to \mathbb{R}^m$ *de clase* $\mathcal{C}^1$ *tales que:*

$$\{(\mathbf{x},\mathbf{y}) \in W : \mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{0}\} = \{(\mathbf{x},\varphi(\mathbf{x})) : \mathbf{x} \in U\}$$

Además, para todo $\mathbf{x} \in U$ *se verifica que*

$$\det\big(D_{n+j} f_i(\mathbf{x},\varphi(\mathbf{x}))\big)_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant m}} \neq 0$$

y

$$J_{\varphi}(\mathbf{x}) = -\left(\big(D_{n+j} f_i(\mathbf{x},\varphi(\mathbf{x}))\big)_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant m}}\right)^{-1} . \big(D_j f_i(\mathbf{x},\varphi(\mathbf{x}))\big)_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}}$$

Si $\mathbf{F}$ *es de clase* $\mathcal{C}^k$ *entonces* φ *es de clase* $\mathcal{C}^k$.

Demostración. Definimos el campo vectorial $\mathbf{H} : \Omega \to \mathbb{R}^{n+m}$ por:

$$H(\mathbf{x},\mathbf{y}) = (\mathbf{x},\mathbf{F}(\mathbf{x},\mathbf{y})) \qquad \forall (\mathbf{x},\mathbf{y}) \in \Omega.$$

$\mathbf{H}$ es de clase $\mathcal{C}^1$ y su matriz jacobiana en $(\mathbf{x},\mathbf{y}) \in \Omega$ viene dada por:

$$J_{\mathbf{H}}(\mathbf{x},\mathbf{y}) = \begin{pmatrix} & & \mathbf{I}_{n\times n} & & \mathbf{0}_{n\times m} & \\ & & & & & \\ D_1 f_1(\mathbf{x},\mathbf{y}) & \dots & D_n f_1(\mathbf{x},\mathbf{y}) & D_{n+1} f_1(\mathbf{x},\mathbf{y}) & \dots & D_{n+m} f_1(\mathbf{x},\mathbf{y}) \\ \dots\dots\dots & \dots & \dots\dots\dots & \dots\dots\dots & \dots & \dots\dots\dots \\ D_1 f_m(\mathbf{x},\mathbf{y}) & \dots & D_n f_m(\mathbf{x},\mathbf{y}) & D_{n+1} f_m(\mathbf{x},\mathbf{y}) & \dots & D_{n+m} f_m(\mathbf{x},\mathbf{y}) \end{pmatrix}$$

Tenemos que:

$$\det J_{\mathbf{H}}(\mathbf{x},\mathbf{y}) = \det\big(D_{n+j} f_i(\mathbf{x},\mathbf{y})\big)_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant m}}$$

Por lo que $\det J_{\mathbf{H}}(\mathbf{a},\mathbf{b}) \neq 0$. Podemos aplicar el teorema de la función inversa al campo vectorial $\mathbf{H}$ para obtener que hay un entorno abierto de $\mathbf{H}(\mathbf{a},\mathbf{b}) = (\mathbf{a},\mathbf{0})$, que podemos tomarlo de la forma $U \times V$ donde $U \subset \mathbb{R}^n$ es un entorno abierto de $\mathbf{a}$ y $V \subset \mathbb{R}^m$ es un entorno abierto de $\mathbf{0}$, y un entorno abierto de $(\mathbf{a},\mathbf{b})$, $W \subset \Omega$, tal que $\mathbf{H}$ es un difeomorfismo de clase $\mathcal{C}^1$ de W sobre $U \times V$.

Sea $\mathbf{G} : U \times V \to W \subset \mathbb{R}^{n+m}$ el difeomorfismo inverso de $\mathbf{H}_{|W}$. Teniendo en cuenta la definición de $\mathbf{H}$, claramente $\mathbf{G}$ es de la forma:

$$\mathbf{G}(\mathbf{x},\mathbf{z}) = (\mathbf{x},\alpha(\mathbf{x},\mathbf{z})) \qquad \forall (\mathbf{x},\mathbf{z}) \in U \times V$$

donde $\alpha : U \times V \to \mathbb{R}^m$ es una función de clase $\mathcal{C}^1$. Definimos $\varphi : U \to \mathbb{R}^m$ por:

$$\varphi(\mathbf{x}) = \alpha(\mathbf{x}, \mathbf{0}) \qquad \forall \mathbf{x} \in U.$$

La función φ es de clase $\mathcal{C}^1$. Para todo $\mathbf{x} \in U$ se tiene que $(\mathbf{x}, \mathbf{0}) \in U \times V$ y $\mathbf{G}(\mathbf{x}, \mathbf{0}) = (\mathbf{x}, \varphi(\mathbf{x}))$. Por lo que $(\mathbf{x}, \mathbf{0}) = \mathbf{H}(\mathbf{G}(\mathbf{x}, \mathbf{0})) = \mathbf{H}(\mathbf{x}, \varphi(\mathbf{x})) = (\mathbf{x}, \mathbf{F}(\mathbf{x}, \varphi(\mathbf{x})))$, luego $\mathbf{F}(\mathbf{x}, \varphi(\mathbf{x})) = \mathbf{0}$. Observa que para todo $\mathbf{x} \in U$ se tiene que $(\mathbf{x}, \varphi(\mathbf{x})) \in W$. Luego hemos probado que:

$$\{(\mathbf{x}, \mathbf{y}) \in W : \mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}\} \supset \{(\mathbf{x}, \varphi(\mathbf{x})) : \mathbf{x} \in U\}$$

Supongamos ahora que $(\mathbf{x}, \mathbf{y}) \in W$ y $\mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$. Entonces $\mathbf{H}(\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{0}) \in U \times V$, luego $\mathbf{x} \in U$ y, por tanto, $(\mathbf{x}, \varphi(\mathbf{x})) = \mathbf{G}(\mathbf{x}, \mathbf{0}) \in W$. Puesto que $\mathbf{H}(\mathbf{x}, \varphi(\mathbf{x})) = (\mathbf{x}, \mathbf{0})$, y $\mathbf{H}$ es inyectiva en W ha de ser $\mathbf{y} = \varphi(\mathbf{x})$. Hemos probado así que:

$$\{(\mathbf{x}, \mathbf{y}) \in W : \mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}\} \subset \{(\mathbf{x}, \varphi(\mathbf{x})) : \mathbf{x} \in U\}.$$

Queda probado que:

$$\{(\mathbf{x}, \mathbf{y}) \in W : \mathbf{F}(\mathbf{x}, \mathbf{y}) = \mathbf{0}\} = \{(\mathbf{x}, \varphi(\mathbf{x})) : \mathbf{x} \in U\}$$

En particular, se tiene que $\varphi(\mathbf{a}) = \mathbf{b}$.

Por el teorema de la inversa, tenemos que para todo $(\mathbf{x}, \mathbf{y}) \in W$ es $\det J_{\mathbf{H}}(\mathbf{x}, \mathbf{y}) \neq 0$; en particular, para todo $\mathbf{x} \in U$ es:

$$\det J_{\mathbf{H}}(\mathbf{x}, \varphi(\mathbf{x})) = \det\big(D_{n+j} f_i(\mathbf{x}, \varphi(\mathbf{x}))\big)_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant m}} \neq 0$$

Para cada $i = 1, 2, \dots, m$ y para todo $\mathbf{x} \in U$ se verifica que:

$$f_i(\mathbf{x}, \varphi(\mathbf{x})) = 0$$

Derivando esta identidad respecto a la variable k-ésima por la regla de la cadena, poniendo $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_m)$, tenemos que:

$$D_k f_i(\mathbf{x}, \varphi(\mathbf{x})) + \sum_{j=1}^{m} D_{n+j} f_i(\mathbf{x}, \varphi(\mathbf{x})) D_k \varphi_j(\mathbf{x}) = 0$$

Para cada valor fijo de $\mathbf{x} \in U$ y de $k = 1, 2, \dots, n$ éste es un sistema lineal de m ecuaciones y m incógnitas: las $D_k \varphi_j(\mathbf{x})$ para $1 \leqslant j \leqslant m$. El determinante de la matriz del sistema es:

$$\det\big(D_{n+j} f_i(\mathbf{x}, \varphi(\mathbf{x}))\big)_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant m}} \neq 0$$

Y, por tanto, dicho sistema tiene solución única que viene dada por:

$$\begin{pmatrix} D_k \varphi_1(\mathbf{x}) \\ \cdots\cdots \\ \cdots\cdots \\ D_k \varphi_m(\mathbf{x}) \end{pmatrix} = - \begin{pmatrix} D_{n+1} f_1(\mathbf{x}, \varphi(\mathbf{x})) & \dots & D_{n+m} f_1(\mathbf{x}, \varphi(\mathbf{x})) \\ \cdots\cdots\cdots\cdots & \cdots & \cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots & \cdots & \cdots\cdots\cdots\cdots \\ D_{n+1} f_m(\mathbf{x}, \varphi(\mathbf{x})) & \dots & D_{n+m} f_m(\mathbf{x}, \varphi(\mathbf{x})) \end{pmatrix}^{-1} \cdot \begin{pmatrix} D_k f_1(\mathbf{x}, \varphi(\mathbf{x})) \\ \cdots\cdots \\ \cdots\cdots \\ D_k f_m(\mathbf{x}, \varphi(\mathbf{x})) \end{pmatrix}$$

Tenemos en total n sistemas de ecuaciones, uno para cada valor de $k=1,2,\dots,n$. Cada uno de ellos tiene como solución la columna k-ésima de la matriz jacobiana $J_\varphi(\mathbf{x})$. Las soluciones de todos ellos podemos escribirlas matricialmente en la forma:

$$J_\varphi(\mathbf{x}) = -\left(\big(D_{n+j}f_i(\mathbf{x},\varphi(\mathbf{x}))\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant m}}\right)^{-1} . \big(D_j f_i(\mathbf{x},\varphi(\mathbf{x}))\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$$

Finalmente, si la función $\mathbf{F}$ es de clase $\mathcal{C}^k$ esta propiedad se traslada a las funciones $\mathbf{H}$, $\mathbf{G}$, α y, por tanto, a la función implícita φ. □

3.18 Observación. Con las notaciones usadas en la demostración del teorema, como W es un entorno abierto de $(\mathbf{a},\mathbf{b})$, hay entornos abiertos $A_1\subset\mathbb{R}^n$ de $\mathbf{a}$ y $B\subset\mathbb{R}^m$ de $\mathbf{b}$ tales que $A_1\times B\subset W$. Como $\varphi:U\to\mathbb{R}^m$ es continua y U es abierto, el conjunto $A=\varphi^{-1}(B)\cap A_1$ es un entorno abierto de $\mathbf{a}$. Para todo $\mathbf{x}\in A$ tenemos que $\varphi(\mathbf{x})\in B$ por lo que $\{(\mathbf{x},\varphi(\mathbf{x})):\mathbf{x}\in A\}\subset A\times B$. Además $A\times B\subset W$. Deducimos fácilmente que:

$$(A\times B)\cap\mathbf{F}^{-1}(\mathbf{0}) = \{(\mathbf{x},\mathbf{y})\in A\times B:\mathbf{F}(\mathbf{x},\mathbf{y})=\mathbf{0}\} = \{(\mathbf{x},\varphi(\mathbf{x}):\mathbf{x}\in A\}$$

En la práctica el teorema de la función implícita se aplica en la forma que te explico en los siguientes ejemplos.

3.19 Ejemplo. Comprueba que la ecuación

$$xyz+\operatorname{sen}(z-6)-2(x+y+x^2y^2)=0$$

define a z como función implícita de (x,y) en un entorno de $(1,1)$, con $z(1,1)=6$. Comprueba que $(1,1)$ es un punto crítico de la función $z=z(x,y)$.

Solución. Pongamos $f(x,y,z)=xyz+\operatorname{sen}(z-6)-2(x+y+x^2y^2)$ que tiene derivadas parciales continuas de todo orden. Tenemos que $\dfrac{\partial f}{\partial z}=xy+\cos(z-6)$, por lo que $\dfrac{\partial f}{\partial z}(1,1,6)=2\neq 0$. Como, además, $f(1,1,6)=0$, el teorema de la función implícita garantiza que hay una función con derivadas parciales continuas, $(x,y)\mapsto z(x,y)$, definida en un entorno, U, de $(1,1)$ tal que $z(1,1)=6$, y

$$f(x,y,z(x,y))=0 \ \text{ para todo } \ (x,y)\in U.$$

Derivando esta identidad tenemos que:

$$\begin{aligned}
\frac{\partial f}{\partial x}+\frac{\partial f}{\partial z}\frac{\partial z}{\partial x} &= yz-2(1+2xy^2)+(xy+\cos(z-6))\frac{\partial z}{\partial x}=0\\
\frac{\partial f}{\partial y}+\frac{\partial f}{\partial z}\frac{\partial z}{\partial y} &= xz-2(1+2x^2y)+(xy+\cos(z-6))\frac{\partial z}{\partial y}=0
\end{aligned}$$

Donde las derivadas parciales de la función implícita $z=z(x,y)$ están calculadas en un punto $(x,y)\in U$ y las de f están calculadas en el punto $(x,y,z(x,y))$. Haciendo $x=y=1$, $z=z(1,1)=6$, en las igualdades anteriores, se obtiene que $\dfrac{\partial z}{\partial x}(1,1)=\dfrac{\partial z}{\partial y}(1,1)=0$, esto es, $(1,1)$ es un punto crítico de $z=z(x,y)$. ♦

El ejemplo anterior es todavía demasiado explícito, nos dice muy claramente lo que hay que hacer. Lo más frecuente es que nos encontremos con ejercicios como el siguiente.

3.20 Ejemplo. Sabiendo que

$$y\cos(xz) + x^3 \mathrm{e}^{zy} - z + 1 = 0 \tag{3.32}$$

Calcula $\dfrac{\partial z}{\partial x}(x,y)$ y particulariza para el punto $(x,y) = (0,0)$.

Solución. En un ejercicio como este lo más fácil es que en la igualdad (3.32) sustituyas mentalmente $z = z(x,y)$ y la veas como

$$y\cos\big(xz(x,y)\big) + x^3 \mathrm{e}^{z(x,y)y} - z(x,y) + 1 = 0 \tag{3.33}$$

es decir, supones que has calculado para valores de x e y dados la solución respecto a z de la igualdad (3.32). Esta solución (que de hecho no es posible expresar de forma explícita, esto es, que no puede calcularse) la representamos por $z = z(x,y)$ y es la función implícita definida por la igualdad (3.32) (el teorema de la función implícita *que es un teorema de existencia* garantiza que dicha función existe cuando se cumplen las hipótesis del mismo). Ahora derivamos en la igualdad (3.33) respecto a x para obtener

$$-y\,\mathrm{sen}\big(xz(x,y)\big)\left(z(x,y) + x\frac{\partial z}{\partial x}(x,y)\right) + 3x^2 \mathrm{e}^{z(x,y)y} + x^3 y\frac{\partial z}{\partial x}(x,y)\,\mathrm{e}^{z(x,y)y} - \frac{\partial z}{\partial x}(x,y) = 0$$

de donde

$$\frac{\partial z}{\partial x}(x,y) = \frac{yz(x,y)\,\mathrm{sen}\big(xz(x,y)\big) - 3x^2 \mathrm{e}^{z(x,y)y}}{x^3 y\,\mathrm{e}^{z(x,y)y} - xy\,\mathrm{sen}(xz(x,y)) - 1}$$

Naturalmente, esta igualdad tiene sentido siempre que el denominador de la fracción sea distinto de cero. Puedes comprobar que si llamas $f(x,y,z) = y\cos(xz) + x^3 \mathrm{e}^{zy} - z + 1$ entonces la igualdad anterior es precisamente

$$\frac{-\dfrac{\partial f}{\partial x}(x,y,z)}{\dfrac{\partial f}{\partial z}(x,y,z)}$$

calculada en el punto $(x,y,z(x,y))$. Para $(x,y) = (0,0)$ se tiene que $z(0,0)$ viene dado por la ecuación que se obtiene haciendo $x = 0$ e $y = 0$ en la igualdad (3.32) de donde se sigue $z(0,0) = 1$. Además

$$\frac{\partial f}{\partial z}(0,0,z(0,0)) = \frac{\partial f}{\partial z}(0,0,1) = -1 \neq 0$$

Por lo que

$$\frac{\partial z}{\partial x}(0,0) = \frac{0}{-1} = 0$$

♦

Ejercicios propuestos

150. Particulariza el teorema de la función implícita:

a) Para campos escalares de dos y de tres variables.

b) Para funciones vectoriales de $\mathbb{R}^3$ en $\mathbb{R}^2$ y para funciones vectoriales de $\mathbb{R}^4$ en $\mathbb{R}^2$.

151. Calcula las derivadas parciales de primer orden de la función $z = z(x,y)$ definida implícitamente por $yz^4 + x^2z^3 - \mathrm{e}^{xyz} = 0$. Particularizar para el punto $(x,y) = (1,0)$.

152. Calcula las derivadas parciales de primer orden de la función $z = z(x,y)$ definida implícitamente por $z^3 + z\mathrm{e}^x + \cos y = 0$.

153. Calcula las derivadas parciales de primer orden de la función $z = z(x,y)$ dada implícitamente por $3x^2y^2 + 2z^2xy - 2zx^3 + 4zy^3 - 4 = 0$, en el punto $(2,1)$ siendo $z(2,1) = 2$.

154. Supongamos que la igualdad

$$\int_{xy}^{y+z} g(t)dt + \int_{3x+y}^{z^2} h(t)dt = 0$$

donde g y h son funciones reales derivables, define a z como función implícita de x,y. Calcula las derivadas parciales de primer orden de $z = z(x,y)$.

155. Sabiendo que g es una función continua con $g(0) = 3$, justifica que

$$\int_{x^2-1}^{xy} g(t)\,\mathrm{d}t + x^2y = 0$$

define a y como función implícita de x en un entorno del punto $(1,0)$. Calcula $y'(1)$.

156. Supongamos que la igualdad $F(x,y,z) = 0$ determina implícitamente funciones diferenciables $x = x(y,z)$, $y = y(x,z)$, $z = z(x,y)$. Probar que $\dfrac{\partial x}{\partial y}\dfrac{\partial y}{\partial z}\dfrac{\partial z}{\partial x} = -1$.

157. Calcula la derivada de la función $y = y(x)$ definida implícitamente por

$$x^y + 3x^2 - 2y^2 - 2y = 0$$

Particularizar para $x = 1$ sabiendo que $y(1) = 1$.

158. Calcula la derivada de la función $y = y(x)$ definida implícitamente por

$$y\log(x^2 + y^2) - 2xy = 0$$

Particularizar para $x = 0$ sabiendo que $y(0) = 1$.

159. Sea $\mathbf{F} = (u,v) : \Omega \to \mathbb{R}^2$ un campo vectorial de clase $\mathcal{C}^1$ en un abierto $\Omega \subset \mathbb{R}^2$ cuya derivada no se anula en Ω. Supongamos que $\mathbf{F}$ satisface las ecuaciones de Cauchy-Riemann en Ω, es decir, que para todo punto $(x,y) \in \Omega$ se verifican las igualdades:

$$\frac{\partial u}{\partial x}(x,y) = \frac{\partial v}{\partial y}(x,y), \quad \frac{\partial u}{\partial y}(x,y) = -\frac{\partial v}{\partial x}(x,y)$$

Prueba que $\mathbf{F}$ es un difeomorfismo local en Ω. Si, además, $\mathbf{F}$ es inyectiva, prueba que su inversa también satisface las ecuaciones de Cauchy-Riemann.

160. Justifica la existencia de una función φ de clase $\mathcal{C}^\infty$ definida en un entorno U de cero en $\mathbb{R}$ con valores en $\mathbb{R}$ tal que:

$$1 - \varphi(x)\,\mathrm{e}^x + x\,\mathrm{e}^{\varphi(x)} = 0 \qquad \forall x \in U$$

Calcula el polinomio de Taylor de segundo orden de φ en $x = 0$.

161. Prueba que las ecuaciones:

$$\begin{aligned} xy^5 + yu^5 + zv^5 &= 1 \\ x^5y + y^5u + z^5v &= 1 \end{aligned}$$

definen a u y a v como funciones implícitas de (x,y,z) en un entorno de $(0,1,1,1,0)$. Calcula $\dfrac{\partial u}{\partial y}(0,1,1)$ y $\dfrac{\partial^2 u}{\partial y^2}(0,1,1)$.

162. Justifica la existencia de abiertos difeomorfos A y B de $\mathbb{R}^2$ verificando que $(1,1) \in A$, $(0,1) \in B$ y que para cada $(x,y) \in A$ existe un único $(u,v) \in B$ tal que:

$$\begin{aligned} x\,\mathrm{e}^u + y\,\mathrm{e}^v &= 1 + \mathrm{e} \\ u\,\mathrm{e}^x + v\,\mathrm{e}^y &= \mathrm{e} \end{aligned}$$

Calcula la matriz jacobiana en $(1,1)$ de la aplicación $(x,y) \mapsto (u(x,y), v(x,y))$.

Calcula la matriz jacobiana en $(0,1)$ de la aplicación $(u,v) \mapsto (x(u,v), y(u,v))$.

163. Sea $z = z(x,y)$ la función dada implícitamente por

$$3x^2y^2 + 2z^2xy - 2zx^3 + 4zy^3 - 4 = 0.$$

Calcula $\dfrac{\partial^2 z}{\partial x \partial y}$ en el punto $(2,1)$ siendo $z(2,1) = 2$.

164. Comprueba que la ecuación

$$xyz + \operatorname{sen}(z-6) - 2(x + y + x^2y^2) = 0$$

define a z como función implícita de (x,y) en un entorno de $(1,1)$, con $z(1,1) = 6$. Comprobar que $(1,1)$ es un punto crítico de la función $z = z(x,y)$ y estudiar si se trata de un máximo relativo, mínimo relativo o punto de silla.

165. Sea $z = z(x,y)$ la función dada implícitamente por $yz^4 + x^2z^3 - \mathrm{e}^{xyz} = 0$. Calcula $\dfrac{\partial^2 z}{\partial x \partial y}$ en el punto $(x,y) = (1,0)$.

166. Calcula y clasifica los puntos críticos de la función $z = z(x,y)$ definida implícitamente por la igualdad $2x^2 + 2y^2 + z^2 + 8xz - z + 8 = 0$.

167. Prueba que es posible despejar de manera diferenciable u y v en función de x, y, z en las ecuaciones:

$$\begin{cases} xy^2 + xzu + yv^2 &= 3 \\ u^3xy + 2xv - u^2v^2 &= 2 \end{cases}$$

En un entorno de $(x,y,z) = (1,1,1)$ y $(u,v) = (1,1)$. Calcula $\dfrac{\partial u}{\partial z}(1,1,1)$ y $\dfrac{\partial v}{\partial z}(1,1,1)$.

168. Justifica que existen dos campos escalares de clase $\mathcal{C}^\infty$ u y v definidos en un entorno abierto U del punto $(1,-1)$ verificando que $u(1,-1) = v(1,-1) = (1,1)$ y para todo $(x,y) \in U$:

$$\begin{cases} x^3 + y^3 + u(x,y)^3 - v(x,y)^3 &= 0 \\ x^2 + y^2 - 2u(x,y)v(x,y) &= 0 \end{cases}$$

Sea $(s,t) \mapsto f(s,t)$ un campo escalar de clase $\mathcal{C}^2$ definido en un entorno abierto de $(1,1)$ y definamos $h(x,y) = f(u(x,y), v(x,y))$. Calcula $\dfrac{\partial^2 h}{\partial x \partial y}(1,-1)$ sabiendo que $D_1 f(1,1) = 2$, $D_2 f(1,1) = 2$, $D_{12} f(1,1) = -1$.

169. Sea $\mathbf{f} : \mathbb{R}^3 \to \mathbb{R}^3$ el campo vectorial definido por:

$$\mathbf{f}(x,y,z) = (x - xy, xy - xyz, xyz)$$

Sea $\Omega = \{(x,y,z) \in \mathbb{R}^3 : xy \neq 0\}$. Justifica que Ω es abierto y que $\mathbf{f}$ es un difeomorfismo de Ω sobre $\mathbf{f}(\Omega)$. Calcula la matriz jacobiana del difeomorfismo inverso en el punto $(0,0,1)$.

170. Prueba que hay un abierto $U \subset \mathbb{R}^2$ con $(1,0) \in U$ y un difeomorfismo $\mathbf{f} = (f_1, f_2)$ de U sobre $V = \mathbf{f}(U)$ que verifica $\mathbf{f}(1,0) = (0,1)$ y para todo $(x,y) \in U$:

$$\begin{cases} x^2 f_1(x,y) + 2y f_2(x,y)^2 + xy &= 0 \\ y f_1(x,y)^2 + 3 f_1(x,y) f_2(x,y) + f_2(x,y)x^2 &= 1 \end{cases}$$

171. Justifica que las igualdades:

$$\begin{cases} xu - yv + e^u \cos v &= 1 \\ xv + yu + e^u \operatorname{sen} v &= 0 \end{cases}$$

definen localmente a u y a v como funciones de x e y en un entorno de $(0,0)$ en el cual se verifica que $\dfrac{\partial u}{\partial x} = \dfrac{\partial v}{\partial y}$ y $\dfrac{\partial u}{\partial y} = -\dfrac{\partial v}{\partial x}$.

172. Sea $A = \{(x,y,z) \in \mathbb{R}^3 : x = y \vee y = z \vee z = x\}$. Prueba que A es un conjunto cerrado y que la función $\mathbf{f} : \Omega \to \mathbb{R}^3$ dada por:

$$\mathbf{f}(x,y,z) = (x + y + z, xy + yz + zx, xyz)$$

es un difeomorfismo local en $\Omega = \mathbb{R}^3 \backslash A$.

Sea $h : \mathbb{R}^3 \to \mathbb{R}$ dada por $h(u,v,w) = u + e^v + \operatorname{sen} w$. Prueba que existe un abierto $U \subset \mathbb{R}^3$ con $(0,-1,0) \in U$ y un campo escalar $g : U \to \mathbb{R}$ de clase $\mathcal{C}^1$ en U con $g(0,-1,0) = 1 + e^{-1}$ tal que $g \circ \mathbf{f} = h$. Calcula $\nabla g(0,-1,0)$.

173. Prueba que las igualdades:

$$\begin{cases} x\,\mathrm{e}^v + yu - u^2 & = & 0 \\ y\cos v + x^2 - u^2 & = & 1 \end{cases}$$

definen a u y a v como funciones implícitas de x e y en un entorno del punto $(2,1)$ siendo $u(2,1)=2$, $v(2,1)=0$. Calcula las derivadas parciales de segundo orden de u en $(2,1)$.

174. Sea $\mathbf{F}:\mathbb{R}^n\to\mathbb{R}^n$ un campo vectorial de clase $\mathcal{C}^1$ y supongamos que hay un número $\rho<1$ tal que para todo $\mathbf{x}\in\mathbb{R}^n$ es $\|D\mathbf{F}(\mathbf{x})\|\leqslant\rho$. Sea $\mathbf{G}:\mathbb{R}^n\to\mathbb{R}^n$ el campo vectorial definido para todo $\mathbf{x}\in\mathbb{R}^n$ por $\mathbf{G}(\mathbf{x})=\mathbf{x}+\mathbf{F}(\mathbf{x})$. Prueba que $\mathbf{G}$ es un difeomorfismo de $\mathbb{R}^n$ sobre $\mathbb{R}^n$.

175. Sea $\mathbf{F}:\mathbb{R}^n\to\mathbb{R}^n$ un campo vectorial de clase $\mathcal{C}^1$ y supongamos que para todos $\mathbf{x},\mathbf{y}\in\mathbb{R}^n$ se verifica que:

$$\|\mathbf{x}-\mathbf{y}\|\leqslant\|\mathbf{F}(\mathbf{x})-\mathbf{F}(\mathbf{y})\|$$

Prueba que $D\mathbf{F}(\mathbf{a})$ es invertible en todo punto $\mathbf{a}\in\mathbb{R}^n$, que $\mathbf{F}(\mathbb{R}^n)$ es cerrado y que $\mathbf{F}$ es un difeomorfismo de $\mathbb{R}^n$ sobre $\mathbb{R}^n$.

176. Sea $f:\mathbb{R}\to\mathbb{R}$ una función de clase $\mathcal{C}^1$ en $\mathbb{R}$ y supongamos que hay un número $\alpha<1$ tal que $|f'(t)|\leqslant\alpha$ para todo $t\in\mathbb{R}$. Definamos $\mathbf{F}:\mathbb{R}^2\to\mathbb{R}^2$ por

$$\mathbf{F}(x,y)=(x+f(y),y+f(x))$$

Prueba que $\mathbf{F}$ es un difeomorfismo de $\mathbb{R}^2$ sobre $\mathbb{R}^2$.

3.4. Variedades diferenciables en $\mathbb{R}^n$

La palabra *“variedad”* se usa en matemáticas para describir espacios topológicos que *localmente* son parecidos al espacio $\mathbb{R}^k$ con la topología de la norma para algún $k\in\mathbb{N}$ llamado la *dimensión* de la variedad. Como ya habrás estudiado en la asignatura de Topología I, el concepto de identidad matemática entre espacios topológicos es el de homeomorfismo: dos espacios topológicos son *homeomorfos* cuando entre ellos existe una biyección continua y con inversa continua. Una tal biyección se llama un *homeomorfismo*. Recuerda también que todo subconjunto no vacío, M, de un espacio métrico (E,d) se considera siempre como espacio métrico con la distancia inducida, y también como espacio topológico con la topología dada por dicha distancia en la cual los abiertos son los abiertos relativos: intersecciones de los abiertos del espacio total E con M. Esta topología se llama la *topología inducida* en M o la *topología relativa* de M.

Si $A\subset\mathbb{R}^n$ y $B\subset\mathbb{R}^k$ son conjuntos no vacíos, diremos que son **homeomorfos** si lo son como espacios topológicos con las respectivas topologías relativas, es decir, existe una biyección $\sigma:A\to B$ continua cuya inversa también es continua. Una tal biyección, cuando existe, se llama un **homeomorfismo** de A sobre B.

Ejercicios propuestos

177. Prueba que todas las bolas abiertas en un espacio normado $(X, \| \ \|)$ son homeomorfas a la bola abierta unidad, $B(\mathbf{0},1)$, y que ésta es homeomorfa a X.

178. Prueba que la esfera euclídea unidad en $\mathbb{R}^3$ no es homeomorfa a $\mathbb{R}^2$.

Consideremos, pues, un conjunto no vacío $M \subset \mathbb{R}^n$. Dicho conjunto es un espacio topológico y será *localmente* parecido al espacio $\mathbb{R}^k$ si para todo punto $\mathbf{a} \in M$ existe un abierto relativo $V_\mathbf{a} \subset M$ que lo contiene que es homeomorfo a $\mathbb{R}^k$. Como toda bola abierta en $\mathbb{R}^k$ es homeomorfa a $\mathbb{R}^k$, y la restricción de un homeomorfismo sigue siendo un homeomorfismo, no se pierde nada, y se gana agilidad en la exposición, si en vez de exigir que $V_\mathbf{a}$ sea homeomorfo a $\mathbb{R}^k$ exigimos que sea homeomorfo a un abierto $U \subset \mathbb{R}^k$. Supongamos, pues, que $\wp : U \to V_\mathbf{a}$ es un homeomorfismo. Podemos interpretar que U es un mapa "plano" que representa $V_\mathbf{a}$ y $\wp$ es la aplicación que lleva cada punto de dicho mapa a su correspondiente en $V_\mathbf{a}$. Digamos que lo que hace la aplicación $\wp$ es "curvar y pegar" el mapa plano U sobre un trozo $V_\mathbf{a}$ de M.

También queremos que nuestras "variedades" tengan un "espacio tangente" en cada punto que tenga la misma dimensión que la variedad. Para ello vamos a exigir que $\wp$ sea al menos de clase $\mathcal{C}^1$ y que su diferencial tenga rango k en todo punto. Naturalmente, $V_\mathbf{a}$ será de la forma $V_\mathbf{a} = W_\mathbf{a} \cap M$ donde $W_\mathbf{a}$ es un entorno abierto de $\mathbf{a}$ en $\mathbb{R}^n$.

3.21 Definición. Un conjunto no vacío $M \subset \mathbb{R}^n$ es una **variedad** (diferenciable de clase $\mathcal{C}^q$, $1 \leqslant q \leqslant \infty$) de dimensión k ($1 \leqslant k \leqslant n$) si para todo punto $\mathbf{a} \in M$ hay un entorno abierto $W_\mathbf{a} \subset \mathbb{R}^n$ de $\mathbf{a}$, un abierto $U \subset \mathbb{R}^k$ y una aplicación $\wp : U \to \mathbb{R}^n$ (de clase $\mathcal{C}^q$, $1 \leqslant q \leqslant \infty$) tal que $\operatorname{rango} J_\wp(\mathbf{u}) = k$ para todo $\mathbf{u} \in U$ y $\wp$ es un homeomorfismo de U sobre $W_\mathbf{a} \cap M$.

Diremos que un par $(\wp, U)$ en las condiciones de la definición anterior es **una carta local** de M en $\mathbf{a}$. El par $(\wp^{-1}, \wp(U))$ se llama un *sistema de coordenadas locales*. Las coordenadas de un punto $\mathbf{x} \in \wp(U)$ son las coordenadas cartesianas de $\wp^{-1}(\mathbf{x})$. Esta terminología no es universal y muchos textos llaman "carta local" a lo que hemos llamado "sistema de coordenadas locales". Tanto da, pues conocer $(\wp, U)$ es lo mismo que conocer $(\wp^{-1}, \wp(U))$. También se dice que el par $(\wp, U)$ o la aplicación $\wp$ es una **parametrización local** de M en $\mathbf{a}$. Una parametrización $(\wp, U)$ tal que la aplicación $\wp^{-1}$ sea una proyección en k-coordenadas, es decir, $\wp^{-1}(x_1, x_2, \ldots, x_n) = (x_{i_1}, x_{i_2}, \ldots, x_{i_k})$ donde $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$, se llama una **parametrización cartesiana**.

Las expresiones escritas entre paréntesis en la definición anterior se sobrentenderán en todo lo que sigue: nuestras variedades serán siempre de clase $\mathcal{C}^q$ con $1 \leqslant q \leqslant \infty$.

Las variedades de dimensión 1 se suelen llamar *curvas regulares* y las de dimensión 2 *superficies regulares*. Las variedades de dimensión $n-1$ que están en $\mathbb{R}^n$ con $n > 3$ suelen llamarse *hipersuperficies*.

Con el convenio $\mathbb{R}^0 = \{0\}$ (espacio vectorial nulo), puede incluirse en la definición anterior el caso $k = 0$.

Ejercicios propuestos

179. Prueba que las variedades de dimensión n contenidas en $\mathbb{R}^n$ son los conjuntos abiertos de $\mathbb{R}^n$.

180. Prueba que las variedades de dimensión 0 son conjuntos de puntos aislados.

En lo que sigue, para no caer en los casos extremos considerados en los ejercicios anteriores, se supondrá que $1 \leqslant k < n$.

En la definición 3.21 las variedades aparecen como subconjuntos de $\mathbb{R}^n$ que localmente son imagen homeomorfa de abiertos de $\mathbb{R}^k$ mediante homeomorfismos diferenciables de rango k. Podemos dar ejemplos sencillos de variedades.

3.22 Ejemplo. Sea $\wp : U \to \mathbb{R}^n$ una función vectorial de clase $\mathcal{C}^q$ en un abierto $U \subset \mathbb{R}^k$ tal que $\operatorname{rango} J_\wp(\mathbf{u}) = k$ para todo $\mathbf{u} \in U$ y $\wp$ es un homeomorfismo de U sobre $\wp(U)$. Entonces $\wp(U)$ es una variedad de clase $\mathcal{C}^q$ y dimensión k. Se dice que dicha variedad está globalmente determinada por la parametrización $\wp$.

En particular, si $U \subset \mathbb{R}^k$ es abierto y $\mathbf{F} : U \to \mathbb{R}^{n-k}$ es una función de clase $\mathcal{C}^q$, entonces la aplicación $\wp : U \to \mathbb{R}^n$ dada por $\wp(\mathbf{u}) = (\mathbf{u}, \mathbf{F}(\mathbf{u}))$ para todo $\mathbf{u} \in U$, es de clase $\mathcal{C}^q$, $\operatorname{rango} J_\wp(\mathbf{u}) = k$ para todo $\mathbf{u} \in U$ y es un homeomorfismo de U sobre $\wp(U)$. Por lo que el conjunto $\wp(U) = \{(\mathbf{u}, \mathbf{F}(\mathbf{u})) : \mathbf{u} \in U\} \subset \mathbb{R}^n$ es una variedad de clase $\mathcal{C}^q$ y dimensión k. Observa que dicha variedad es la gráfica de $\mathbf{F}$ y que $\wp$ es una parametrización cartesiana global de la misma. ♦

El ejemplo anterior nos dice que la gráfica de toda función de clase $\mathcal{C}^q$ es una variedad de clase $\mathcal{C}^q$. Veremos enseguida que toda variedad es *localmente* la gráfica de una función.

El teorema de la función implícita nos dice que el conjunto donde se anula un campo vectorial diferenciable de rango máximo es localmente una gráfica funcional y proporciona interesantes ejemplos de variedades.

Seguiremos aquí los convenios de notación que usamos en el teorema de la función implícita: haremos la identificación $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^{n-k}$. Una expresión de la forma $(\mathbf{x}, \mathbf{y})$ donde $\mathbf{x} \in \mathbb{R}^k$, $\mathbf{y} \in \mathbb{R}^{n-k}$, se interpreta como un vector de $\mathbb{R}^n$ cuyas primeras k componentes son las de $\mathbf{x}$ y cuyas últimas $n-k$ componentes son las de $\mathbf{y}$. Notaremos por $\pi_1 : \mathbb{R}^n \to \mathbb{R}^k$ la proyección en las primeras k coordenadas, esto es, $\pi_1(\mathbf{x}, \mathbf{y}) = \mathbf{x}$. Notaremos por $\pi_2 : \mathbb{R}^n \to \mathbb{R}^{n-k}$ la proyección en las últimas $n-k$ coordenadas, esto es, $\pi_1(\mathbf{x}, \mathbf{y}) = \mathbf{y}$. Observa que estas proyecciones son aplicaciones abiertas.

3.23 Proposición. *Sea* $\mathbf{F} : \Omega \to \mathbb{R}^{n-k}$ *una función de clase* $\mathcal{C}^q$ *en un abierto* $\Omega \subset \mathbb{R}^n$. *Sea*

$$M = \{\mathbf{x} \in \Omega : \mathbf{F}(\mathbf{x}) = \mathbf{0}\}$$

y supongamos que $\operatorname{rango} J_\mathbf{F}(\mathbf{x}) = n-k$ *para todo* $\mathbf{x} \in M$. *Entonces se verifica que* M *es una variedad de clase* $\mathcal{C}^q$ *y dimensión* k. *Se dice que la variedad* M *está determinada por* $\mathbf{F}$ *y que* $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ *son las* ***ecuaciones implícitas*** *de la variedad.*

Demostración. Sea $\mathbf{x}_0 \in M$. Entonces $\mathbf{F}(\mathbf{x}_0) = \mathbf{0}$ y $\operatorname{rango} J_{\mathbf{F}}(\mathbf{x}_0) = n-k$. No es restrictivo suponer que las últimas $n-k$ columnas de la matriz $J_{\mathbf{F}}(\mathbf{x}_0)$ son linealmente independientes[1]. Pongamos $\mathbf{a} = \pi_1(\mathbf{x}_0)$, $\mathbf{b} = \pi_2(\mathbf{x}_0)$. El teorema de la función implícita nos dice que hay un entorno abierto $W_{\mathbf{x}_0} \subset \Omega$ de $\mathbf{x}_0 = (\mathbf{a}, \mathbf{b})$, un entorno abierto $U \subset \mathbb{R}^k$ de $\mathbf{a}$ y una aplicación de clase $\mathcal{C}^q$, $\varphi : U \to \mathbb{R}^{n-k}$ tal que:

$$\{\mathbf{x} \in W_{\mathbf{x}_0} : \mathbf{F}(\mathbf{x}) = \mathbf{0}\} = W_{\mathbf{x}_0} \cap M = \{(\mathbf{u}, \varphi(\mathbf{u})) : \mathbf{u} \in U\}$$

Basta tener en cuenta ahora que la aplicación $\wp : U \to \mathbb{R}^n$ dada por $\wp(\mathbf{u}) = (\mathbf{u}, \varphi(\mathbf{u}))$ para todo $\mathbf{u} \in U$, es de clase $\mathcal{C}^q$, $\operatorname{rango} J_{\wp}(\mathbf{u}) = k$ para todo $\mathbf{u} \in U$ y es un homeomorfismo de U sobre $W_{\mathbf{x}_0} \cap M$. □

3.24 Proposición. *Sea $M \subset \mathbb{R}^n$. Las siguientes afirmaciones son equivalentes:*

a) M es una variedad de clase $\mathcal{C}^q$ y dimensión k.

b) Para todo punto $\mathbf{a} \in M$ existen, previa una reordenación de las variables en $\mathbb{R}^n$, entornos abiertos $W_{\mathbf{a}} \subset \mathbb{R}^n$ de $\mathbf{a}$, $U \subset \mathbb{R}^k$ de $\pi_1(\mathbf{a})$, y una función $\varphi : U \to \mathbb{R}^{n-k}$ de clase $\mathcal{C}^q$ tal que $W_{\mathbf{a}} \cap M = \{(\mathbf{u}, \varphi(\mathbf{u})) : \mathbf{u} \in U\}$.

c) Para todo punto $\mathbf{a} \in M$ existen, previa una reordenación de las variables en $\mathbb{R}^n$, un entorno abierto $G_{\mathbf{a}} \subset \mathbb{R}^n$ de $\mathbf{a}$ y una función $\mathbf{F} : G_{\mathbf{a}} \to \mathbb{R}^{n-k}$ de clase $\mathcal{C}^q$ tales que $\operatorname{rango} J_{\mathbf{F}}(\mathbf{x}) = n-k$ para todo $\mathbf{x} \in G_{\mathbf{a}}$ y

$$G_{\mathbf{a}} \cap M = \{\mathbf{x} \in G_{\mathbf{a}} : \mathbf{F}(\mathbf{x}) = \mathbf{0}\}$$

Se dice que la función **F** ***determina localmente*** *a M en* **a** *y también que $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ son las* ***ecuaciones implícitas locales*** *de M en* **a**.

Demostración. Probaremos que $a) \Longrightarrow b)$. Sea $\mathbf{a} \in M$. Por hipótesis hay un abierto $\Omega \subset \mathbb{R}^k$, una aplicación $\wp : \Omega \to \mathbb{R}^n$ de clase $\mathcal{C}^q$ tal que $\operatorname{rango} J_{\wp}(\mathbf{u}) = k$ para todo $\mathbf{u} \in \Omega$, $\wp$ es un homeomorfismo de Ω sobre $\wp(\Omega) \subset M$ y $\wp(\Omega)$ es un entorno abierto relativo de $\mathbf{a}$ en M. Sea $\mathbf{u}_0 \in \Omega$ tal que $\wp(\mathbf{u}_0) = \mathbf{a}$. Como $\operatorname{rango} J_{\wp}(\mathbf{u}_0) = k$, reordenando las variables si fuera preciso, podemos suponer que las primeras k filas de $J_{\wp}(\mathbf{u}_0)$ tienen determinante distinto de cero. Esto es tanto como decir que la aplicación $\wp_1 = \pi_1 \circ \wp : \Omega \to \mathbb{R}^k$, cuya matriz jacobiana está formada precisamente por las primeras k filas de $J_{\wp}(\mathbf{u}_0)$, verifica que $\det J_{\wp_1}(\mathbf{u}_0) \neq 0$.

El teorema de la función inversa afirma que hay un entornos abiertos $V \subset \Omega$ de $\mathbf{u}_0$ y U de $\wp_1(\mathbf{u}_0) = \pi_1(\mathbf{a})$ tal que $\wp_{1|V} : V \to U$ es un difeomorfismo de clase $\mathcal{C}^q$ de V sobre U.

Sea $\wp_{1|V}^{-1} : U \to V$ el difeomorfismo inverso de $\wp_{1|V}$. Observa que lo que hace este difeomorfismo es invertir las primeras k funciones componentes de $\wp$. Definamos $\mathbf{h} : U \to \mathbb{R}^n$ por $\mathbf{h}(\mathbf{u}) = \wp(\wp_{1|V}^{-1}(\mathbf{u}))$.

Pongamos $\wp_2 = \pi_2 \circ \wp : \Omega \to \mathbb{R}^{n-k}$, con lo que para todo $\mathbf{u} \in \Omega$ es $\wp(\mathbf{u}) = (\wp_1(\mathbf{u}), \wp_2(\mathbf{u}))$.

Para todo $\mathbf{u} \in U$ es:

$$\mathbf{h}(\mathbf{u}) = \wp(\wp_{1|V}^{-1}(\mathbf{u})) = (\wp_1(\wp_{1|V}^{-1}(\mathbf{u})), \wp_2(\wp_{1|V}^{-1}(\mathbf{u}))) = (\mathbf{u}, \wp_2(\wp_{1|V}^{-1}(\mathbf{u}))$$

[1]Esto siempre puede conseguirse componiendo **F** con una conveniente permutación de ejes en $\mathbb{R}^n$.

El conjunto $\mathbf{h}(U) = \wp(\wp_{1|V}^{-1}(U)) = \wp(V)$ es un entorno abierto en M de $\wp(\mathbf{u}_0) = \mathbf{a}$ por lo que hay un abierto $W_\mathbf{a} \subset \mathbb{R}^n$ tal que $\wp(V) = W_\mathbf{a} \cap M$.

Definamos $\varphi : U \to \mathbb{R}^{n-k}$ por $\varphi(\mathbf{u}) = (\wp_2 \circ \wp_{1|V}^{-1})(\mathbf{u})$. Dicha aplicación es de clase $\mathcal{C}^q$ y:

$$\mathbf{h}(U) = \{\mathbf{h}(\mathbf{u}) : \mathbf{u} \in U\} = \{(\mathbf{u}, \varphi(\mathbf{u})) : \mathbf{u} \in U\} = W_\mathbf{a} \cap M \tag{3.34}$$

Probaremos que $b) \Longrightarrow c)$. Dado $\mathbf{a} \in M$, por hipótesis, previa una reordenación de las variables en $\mathbb{R}^n$, existen entornos abiertos $W_\mathbf{a} \subset \mathbb{R}^n$ de $\mathbf{a}$, $U \subset \mathbb{R}^k$ de $\pi_1(\mathbf{a})$, y una función $\varphi : U \to \mathbb{R}^{n-k}$ de clase $\mathcal{C}^q$ tal que $W_\mathbf{a} \cap M = \{(\mathbf{u}, \varphi(\mathbf{u})) : \mathbf{u} \in U\}$. Sea $G_\mathbf{a} = W_\mathbf{a} \cap (U \times \mathbb{R}^{n-k})$ y definamos $\mathbf{F} : G_\mathbf{a} \to \mathbb{R}^{n-k}$ por $\mathbf{F}(\mathbf{x}) = \pi_2(\mathbf{x}) - \varphi(\pi_1(\mathbf{x}))$ para todo $\mathbf{x} \in G_\mathbf{a}$. Observa que:

$$\mathbf{F}(x_1, x_2, \dots, x_n) = (x_{k+1} - \varphi_{k+1}(x_1, \dots, x_k), \dots, x_n - \varphi_n(x_1, \dots, x_k)).$$

Deducimos que $\operatorname{rango} J_\mathbf{F}(\mathbf{x}) = n - k$ para todo $\mathbf{x} \in G_\mathbf{a}$. Además, para todo $\mathbf{x} \in G_\mathbf{a} \cap M$ se tiene que $\mathbf{x} \in W_\mathbf{a} \cap M$ por lo que $\mathbf{x} = (\mathbf{u}, \varphi(\mathbf{u}))$ para algún $\mathbf{u} \in U$ por lo que

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{u}, \varphi(\mathbf{u})) = \varphi(\mathbf{u}) - \varphi(\mathbf{u}) = \mathbf{0}$$

Recíprocamente, si $\mathbf{x} \in G_\mathbf{a}$ y $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ entonces $\pi_1(\mathbf{x}) = \mathbf{u} \in U$ y $\pi_2(\mathbf{x}) == \varphi(\mathbf{u})$ por lo que $\mathbf{x} = (\pi_1(\mathbf{x}), \pi_2(\mathbf{x})) = (\mathbf{u}, \varphi(\mathbf{u})) \in M$. Hemos probado así que $G_\mathbf{a} \cap M = \{\mathbf{x} \in G_\mathbf{a} : \mathbf{F}(\mathbf{x}) = \mathbf{0}\}$.

La implicación $c) \Longrightarrow a)$ es consecuencia de la proposición 3.23 y del carácter local del concepto de variedad. $\square$

Observa que la función $\mathbf{h} : U \to \mathbb{R}^n$ definida en la demostración anterior, $\mathbf{h}(\mathbf{u}) = (\mathbf{u}, \varphi(\mathbf{u}))$, es una aplicación de clase $\mathcal{C}^q$ y $\operatorname{rango} J_\mathbf{h}(\mathbf{u}) = k$ para todo $\mathbf{u} \in U$. Además $\mathbf{h}$ es una biyección continua de U sobre $W_\mathbf{a} \cap M$ cuya biyección inversa, $(\mathbf{u}, \varphi(\mathbf{u})) \mapsto \mathbf{u}$, es la proyección en las primeras k-coordenadas y, por tanto, continua. Esto es, el par $(\mathbf{h}, U)$ es una parametrización cartesiana local de M en $\mathbf{a}$. Hemos probado así el siguiente resultado.

3.25 Corolario. *Toda variedad tiene parametrizaciones cartesianas en el entorno de cada uno de sus puntos.*

Hemos visto que una variedad $M \subset \mathbb{R}^n$ de dimensión k puede venir dada (localmente) por medio de sus ecuaciones paramétricas, es decir como la imagen de una parametrización, o por medio de sus ecuaciones implícitas, es decir como el conjunto de puntos donde se anula una función de n variables y $n-k$ componentes y de rango máximo. Es instructivo considerar los casos más simples de variedades: los subespacios afines de $\mathbb{R}^n$. Como debes saber, un subespacio afín de dimensión k de $\mathbb{R}^n$ puede venir dado por sus ecuaciones implícitas:

$$\left\{\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & c_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & c_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots & \dots & \dots\dots \\ a_{n-k1}x_1 + a_{n-k2}x_2 + \cdots + a_{n-kn}x_n & = & c_{n-k} \end{array}\right.$$

Dichas ecuaciones representan la intersección de $n-k$ hiperplanos y deben ser linealmente independientes, es decir, el rango de la matriz de los coeficientes del sistema debe ser igual a $n-k$. Definiendo $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^{n-k}$ por:

$$\mathbf{F}(x_1, x_2, \dots, x_n) = \left(\sum_{j=1}^n a_{1j}x_j - c_1, \sum_{j=1}^n a_{2j}x_j - c_2, \dots, \sum_{j=1}^n a_{n-kj}x_j - c_{n-k} \right)$$

Se tiene que $\mathbf{F}$ es de clase $\mathcal{C}^\infty$ y su matriz jacobiana, que es la matriz de los coeficientes del sistema, tiene rango máximo $n-k$, y el subespacio afín viene dado globalmente como el conjunto de los ceros de $\mathbf{F}$.

Si en la situación anterior suponemos que las últimas $n-k$ columnas de la matriz de los coeficientes del sistema son linealmente independientes, podemos resolver dicho sistema y su solución, poniendo $\mathbf{A}=(a_{ij})_{\substack{1\leqslant i\leqslant n-k\\1\leqslant j\leqslant k}}$, $\mathbf{B}=(a_{ij})_{\substack{1\leqslant i\leqslant n-k\\k+1\leqslant j\leqslant n}}$ y $\mathbf{c}=(c_1,c_2,\dots,c_{n-k})$, está dada por:

$$(x_{k+1},x_{k+2},\dots,x_n)^t=\mathbf{B}^{-1}.\mathbf{c}^t-\mathbf{B}^{-1}.\mathbf{A}.(x_1,x_2,\dots,x_k)^t$$

Poniendo

$$\wp(x_1,x_2,\dots,x_k)=\big(x_1,x_2,\dots,x_k,\mathbf{B}^{-1}.\mathbf{c}^t-\mathbf{B}^{-1}.\mathbf{A}.(x_1,x_2,\dots,x_k)^t\big)$$

se tiene que $\wp:\mathbb{R}^k\to\mathbb{R}^n$ es una parametrización que determina globalmente al subespacio afín. Observa que la imagen de $\wp(\mathbb{R}^k)$ viene dada en función de los parámetros $x_1,x_2,\dots,x_k$ que toman valores reales cualesquiera.

A la vista de estos ejemplos, queda claro que el concepto de variedad que hemos definido refleja una vez más la idea de "linealización" básica del cálculo diferencial.

3.5. Espacios tangente y normal

En lo que sigue $M\subset\mathbb{R}^n$ será una variedad de dimensión k ($1\leqslant k<n$), clase $\mathcal{C}^q$ y $\mathbf{a}\in M$ un punto de M.

Se dice que $\mathbf{u}\in\mathbb{R}^n$ es un **vector tangente** a M en $\mathbf{a}$ si existe una curva contenida en M, $\gamma:]-\delta,\delta[\to M$, que pasa por $\mathbf{a}$, $\gamma(0)=\mathbf{a}$, y que tiene tangente en $\mathbf{a}$ con dirección $\mathbf{u}$, $\gamma'(0)=\mathbf{u}$.

Se define el **espacio tangente** a M en $\mathbf{a}$ como el conjunto de todos los vectores tangentes a M en $\mathbf{a}$. Lo notaremos $T_M(\mathbf{a})$ y veremos enseguida que es un subespacio vectorial de $\mathbb{R}^n$.

La variedad afín $\mathbf{a}+T_M(\mathbf{a})$ se llama **variedad afín tangente** a M en $\mathbf{a}$ (recta tangente si $k=1$, plano tangente si $k=2$, hiperplano tangente si $k=n-1$).

El complemento ortogonal en $\mathbb{R}^n$ del espacio tangente a M en $\mathbf{a}$, que notaremos $T_M^\perp(\mathbf{a})$, se llama **espacio normal** a M en $\mathbf{a}$:

$$T_M^\perp(\mathbf{a})=\big\{\mathbf{x}\in\mathbb{R}^n:\big\langle\mathbf{x}\big|\mathbf{u}\big\rangle=0,\ \forall\mathbf{u}\in T_M(\mathbf{a})\big\}$$

La variedad afín $\mathbf{a}+T_M^\perp(\mathbf{a})$ se llama **variedad afín normal** a M en $\mathbf{a}$.

Cuando se conoce una parametrización o las ecuaciones implícitas que representan localmente a M en $\mathbf{a}$ es muy fácil calcular los espacios tangente y normal.

3.26 Proposición. *Sean* $\mathbf{F}(\mathbf{x})=\mathbf{0}$ *las ecuaciones implícitas locales de* M *en* $\mathbf{a}$ *y* $(\wp,\Omega)$ *una parametrización local de* M *en* $\mathbf{a}$ *con* $\wp(\mathbf{c})=\mathbf{a}$*. Entonces se verifica que:*

$$T_M(\mathbf{a})=\operatorname{Ker}D\mathbf{F}(\mathbf{a})=\operatorname{Im}D\wp(\mathbf{c})$$

Demostración. Como las matrices jacobianas de $\mathbf{F}$ y $\wp$ tienen respectivamente rango $n-k$ y k, los espacios vectoriales $\operatorname{Ker}D\mathbf{F}(\mathbf{a})$ e $\operatorname{Im}D\wp(\mathbf{c})$ tiene dimensión k. Bastará, por ello, probar que:

$$\operatorname{Im}D\wp(\mathbf{c})\subset T_M(\mathbf{a})\subset\operatorname{Ker}D\mathbf{F}(\mathbf{a})$$

Sea $\mathbf{h}\in \operatorname{Im} D\wp(\mathbf{c})$ y sea $\mathbf{v}\in\mathbb{R}^k$ tal que $D\wp(\mathbf{c})(\mathbf{v})=\mathbf{h}$. Sea $\delta>0$ tal que el segmento $]\mathbf{c}-\delta\mathbf{v},\mathbf{c}+\delta\mathbf{v}[\subset\Omega$ y definamos $\gamma:]-\delta,\delta[\to M$ por $\gamma(t)=\wp(\mathbf{c}+t\mathbf{v})$. Tenemos que $\gamma(0)=\mathbf{a}$ y $\gamma'(0)=D\wp(\mathbf{c})(\mathbf{v})=\mathbf{h}$. Lo que prueba que $\mathbf{h}\in T_M(\mathbf{a})$.

Sea ahora $\mathbf{u}\in T_M(\mathbf{a})$ y sea $\gamma:]-\delta,\delta[\to M$ una curva con $\gamma(0)=\mathbf{a}$ y $\gamma'(0)=\mathbf{u}$. Tomando δ suficientemente pequeño para que la curva γ esté contenida en la parte de M determinada por $\mathbf{F}$ se tiene que $\sigma(t)=(\mathbf{F}\circ\gamma)(t)=0$ y por tanto $\sigma'(0)=D\mathbf{F}(\mathbf{a}).\mathbf{u}=0$, lo que prueba que $\mathbf{u}\in\operatorname{Ker} D\mathbf{F}(\mathbf{a})$. □

La proposición anterior tiene una simetría que merece destacar: si representamos localmente M en $\mathbf{a}$ de forma implícita por los *ceros* de una función $\mathbf{F}$, entonces el espacio tangente $T_M(\mathbf{a})$ viene dado por los *ceros* de $D\mathbf{F}(\mathbf{a})$; y si representamos localmente M en $\mathbf{a}$ como la *imagen* de una parametrización $\wp$, entonces el espacio tangente $T_M(\mathbf{a})$ viene dado como la *imagen* de $D\wp(\mathbf{c})$ siendo $\wp(\mathbf{c})=\mathbf{a}$.

3.27 Corolario. *Sean $\mathbf{F}(\mathbf{x})=\mathbf{0}$ las ecuaciones implícitas locales de M en $\mathbf{a}$ y pongamos $\mathbf{F}=(F_1,F_2,\dots,F_{n-k})$, y sea $(\wp,\Omega)$ una parametrización local de M en $\mathbf{a}$ con $\wp(\mathbf{c})=\mathbf{a}$ y pongamos $\wp=(\wp_1,\wp_2,\dots,\wp_n)$. Entonces se verifica que:*

a) Los vectores fila de la matriz jacobiana $J_{\mathbf{F}}(\mathbf{a})$, es decir, los vectores $\nabla F_i(\mathbf{a})$, $1\leqslant i\leqslant n-k$, forman una base de $T_M^{\perp}(\mathbf{a})$ y la variedad afín tangente a M en $\mathbf{a}$ viene dada por:

$$\mathbf{a}+T_M(\mathbf{a})=\left\{\mathbf{x}\in\mathbb{R}^n:\left\langle\nabla F_i(\mathbf{a})\middle|\mathbf{x}-\mathbf{a}\right\rangle=0,\ 1\leqslant i\leqslant n-k\right\}\tag{3.35}$$

Naturalmente, el espacio afín normal a M en $\mathbf{a}$ viene dado por:

$$\mathbf{a}+T_M^{\perp}(\mathbf{a})=\left\{\mathbf{x}\in\mathbb{R}^n:\mathbf{x}=\mathbf{a}+\sum_{i=1}^{n-k}\mu_i\nabla F_i(\mathbf{a}),\ \ \mu_i\in\mathbb{R},\ \ 1\leqslant i\leqslant n-k\right\}\tag{3.36}$$

b) Los vectores columna de la matriz jacobiana $J_{\wp}(\mathbf{c})$, es decir, los vectores:

$$D_j\wp(\mathbf{c})=(D_j\wp_1(\mathbf{c}),D_j\wp_2(\mathbf{c}),\dots,D_j\wp_n(\mathbf{c}))\qquad 1\leqslant j\leqslant k$$

son una base del espacio $T_M(\mathbf{a})$ y la variedad afín normal a M en $\mathbf{a}$ viene dada por:

$$\mathbf{a}+T_M^{\perp}(\mathbf{a})=\left\{\mathbf{x}\in\mathbb{R}^n:\left\langle D_j\wp(\mathbf{c})\middle|\mathbf{x}-\mathbf{a}\right\rangle=0,\ 1\leqslant j\leqslant k\right\}\tag{3.37}$$

Naturalmente, el espacio afín tangente a M en $\mathbf{a}$ viene dado por:

$$\mathbf{a}+T_M(\mathbf{a})=\left\{\mathbf{x}\in\mathbb{R}^n:\mathbf{x}=\mathbf{a}+\sum_{j=1}^{k}\lambda_jD_j\wp(\mathbf{c}),\ \ \lambda_j\in\mathbb{R},\ \ 1\leqslant j\leqslant k\right\}\tag{3.38}$$

Observa que (3.35) y (3.37) son, respectivamente, las ecuaciones implícitas de la variedad afín tangente y del espacio afín normal, mientras que (3.36) y (3.38) son, respectivamente, las ecuaciones paramétricas del espacio afín normal y de la variedad afín tangente.

Ejercicios propuestos

181. Comprueba que las rectas y planos tangentes a curvas y superficies estudiados en la sección 2.2 son casos particulares de los resultados generales para variedades.

182. Interpreta en el contexto de la teoría de variedades los ejercicios 100 – 104.

183. Prueba que cada uno de los siguientes subconjuntos de $\mathbb{R}^3$ es una variedad y calcula el espacio afín tangente y normal en el punto P que se indica.

a) $M = \left\{(x,y,z)\mathbb{R}^3 : \dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1\right\}, \quad P=(a,b-c).$

b) $M = \left\{(x,y,z)\mathbb{R}^3 : z=x^2-y^2\right\}, \quad P=(0,0,0).$

c) $M = \left\{(x,y,z)\mathbb{R}^2\times\mathbb{R}^+ : x^2+y^2+z^2=1 \wedge \left(x-\dfrac{1}{2}\right)^2+y^2=\dfrac{1}{4}\right\}, \; P=(0,0,1).$

d) $M = \left\{(x,y,z)\mathbb{R}^3 : (\sqrt{x^2+y^2}-2)^2+z^2=1\right\}, \quad P=(3,0,0).$

184. Prueba que $M=\left\{(x,y,z)\in\mathbb{R}^3 : x^2+y^2-2z^2=0 \wedge x+y+z=1\right\}$ es una curva regular y calcula su recta tangente en $(1,1,-1)$.

185. Prueba que $M=\left\{(x,y,z)\in\mathbb{R}^3 : x^2+y^2+z^2=1 \wedge z=xy\right\}$ es una curva regular y calcula su recta tangente en $(a,b,c)\in M$.

186. Prueba que la esfera euclídea unidad en $\mathbb{R}^n$:

$$S^{n-1}=\{\mathbf{x}\in\mathbb{R}^n : \|\mathbf{x}\|_2=1\}$$

es una variedad de dimensión $n-1$ y calcula el hiperplano tangente y la recta normal en un punto $\mathbf{a}\in S^{n-1}$.

187. Sean $M\subset\mathbb{R}^n$ y $N\subset\mathbb{R}^m$ variedades y $\mathbf{f}:\mathcal{U}\to\mathbb{R}^m$ una función de clase $\mathcal{C}^1$ en un abierto $\mathcal{U}\supset M$ tal que $\mathbf{f}(M)\subset N$. Prueba que $D\mathbf{f}(\mathbf{a})(T_M(\mathbf{a}))\subset T_N(\mathbf{f}(\mathbf{a}))$.

188. Calcula bases del espacio tangente y del espacio normal en el punto $(0,1,1,0)$ del toro:

$$M=\left\{(x_1,x_2,x_3,x_4)\in\mathbb{R}^4 : x_1^2+x_2^2=1 \wedge x_3^2+x_4^2=1\right\}$$

3.6. Extremos condicionados

En la teoría de extremos relativos se supone que las variables pueden tomar valores en cualquier punto de un conjunto abierto, es decir, pueden *“moverse libremente”* en dicho conjunto. En muchos, por no decir que en la mayoría, de los problemas reales las variables no tienen tanta libertad y están obligadas a satisfacer ciertas condiciones que en Física suelen llamarse *“ligaduras”*. Por ejemplo, supongamos que un móvil se mueve en una curva Γ dada por la intersección de dos superficies; para cada punto $(x,y,z)\in\Gamma$ la energía cinética del móvil viene dada por una función conocida $f(x,y,z)$ y queremos calcular los puntos de la trayectoria donde dicha energía es máxima o mínima. En esta situación las variables x,y,z no son libres sino que deben satisfacer la condición $(x,y,z)\in\Gamma$. Otro ejemplo; supongamos que la temperatura en un punto (x,y,z) de la superficie terrestre viene dada por una función $T(x,y,z)$ y queremos calcular los puntos de mayor y menor temperatura. Aquí las variables tampoco son libres pues deben verificar una condición de la forma $x^2+y^2+z^2=R^2$ donde R es el radio de la Tierra. Igualmente, en problemas de optimización de costes o beneficios las variables están siempre sometidas a restricciones que dependen de las condiciones de producción o del mercado.

Es importante que comprendas la diferencia entre un problema de extremos relativos "libres" y un problema de extremos condicionados. Considera el siguiente ejemplo.

3.28 Ejemplo. La función $f(x,y)=xy\mathrm{e}^{x^2+y^2}$ tiene un único punto crítico, el origen, que es un punto de silla. Por tanto dicha función no tiene extremos relativos en $\mathbb{R}^2$. Supongamos que imponemos a las variables la condición $x^2+y^2=1$ y queremos calcular el máximo valor de $f(x,y)$ cuando se verifique que $x^2+y^2=1$. Fíjate en que el problema es completamente distinto. Ahora solamente nos interesan los valores que toma la función $f(x,y)$ en el conjunto

$$K=\left\{(x,y)\in\mathbb{R}^2 : x^2+y^2=1\right\}$$

Sabemos que dicho conjunto es un conjunto compacto; además la función f es continua, por tanto podemos asegurar, de entrada, que tiene que haber algún punto $(a,b)\in K$ en el cual la función f alcanza su mayor valor en K (y tiene que haber otro donde alcance su menor valor en K). Calcular dicho punto es, en este caso, muy sencillo pues para $(x,y)\in K$ se tiene que $f(x,y)=\mathrm{e}xy$. Como para $(x,y)\in K$ se tiene que $y=\pm\sqrt{1-x^2}$ y los valores negativos de f no nos interesan porque queremos calcular el mayor valor que toma en K, se sigue que

$$\text{máx}\{f(x,y):(x,y)\in K\}=\text{máx}\left\{\mathrm{e}x\sqrt{1-x^2} : 0\leqslant x\leqslant 1\right\}$$

Nuestro problema se ha convertido en algo que ya sabes hacer: calcular el máximo absoluto de la función $h(x)=\mathrm{e}x\sqrt{1-x^2}$ para $0\leqslant x\leqslant 1$.

Por cierto, por la desigualdad de las medias se tiene que:

$$x\sqrt{1-x^2}=\sqrt{x^2(1-x^2)}\leqslant\frac{1}{2}$$

Y la igualdad se da si, y sólo si, $x^2=1-x^2$, esto es en $x=1/\sqrt{2}$ se alcanza el máximo absoluto de f en K. ♦

De hecho, tú has resuelto ejercicios de extremos condicionados aunque puede que no seas consciente de ello. Por ejemplo, seguro que alguna vez has resuelto el siguiente ejercicio.

3.29 Ejemplo. Entre todos los rectángulos cuyo perímetro es igual a 16 calcular el que tiene área máxima.

Este ejercicio puedes plantearlo como sigue. Sea $f(x,y)=xy$ la función que da el área de un rectángulo cuyos lados tienen longitudes x e y. Se trata de calcular el máximo de $f(x,y)$ cuando las variables verifican la condición $2x+2y=16$. Por tanto, es un problema de extremos condicionados. Seguro que ahora recuerdas algunos otros ejercicios parecidos a este que has hecho sin saber que estabas haciendo problemas de extremos condicionados. La razón es clara: la condición que nos dan es tan sencilla que permite despejar una variable en función de la otra, $y = 8 - x$, con lo que nuestra función se convierte en $xy = x(8 - x)$ y el problema queda reducido a calcular el mayor valor de $x(8 - x)$ cuando $0 \leqslant x \leqslant 8$.

Por cierto, por la desigualdad de las medias se tiene que:

$$\sqrt{x(8-x)} \leqslant 4$$

Y la igualdad se da si, y sólo si, $x = 8 - x$, esto es en $x = 4$ y el rectángulo de área máxima es el cuadrado. ♦

Los ejemplos anteriores ponen de manifiesto que *los problemas de extremos condicionados de funciones de dos variables en los que puede utilizarse la condición que nos dan para despejar una variable en función de otra, se reducen fácilmente a problemas de extremos de funciones de una variable*. Pero supongamos ahora que cambiamos la condición del ejemplo 1 por la siguiente:

$$x - \mathrm{e}^x + y + \mathrm{e}^y + \sin(1 + xy) = 2$$

La cosa se complica porque ahora es imposible usar la condición impuesta para despejar una variable en función de la otra. Ahora sí tenemos un auténtico problema de extremos condicionados.

Lo antes dicho para funciones de dos variables puedes generalizarlo para funciones de tres variables. Por ejemplo el problema de calcular las dimensiones de un ortoedro de volumen igual a 8 para que su superficie lateral sea mínima, puedes plantearlo como sigue: calcular el mínimo de

$$f(x,y,z) = 2xy + 2xz + 2yz$$

(la función que da la superficie lateral de un ortoedro cuyos lados tiene longitudes x, y, z) con la condición $xyz = 8$. Se trata de un problema de extremos condicionados, pero la condición dada permite despejar una variable en función de las otras dos, $z = 8/(xy)$, con lo que nuestra función queda $2xy + 2xz + 2yz = xy + 16/y + 16/x$, función de la que hay que calcular su mínimo absoluto cuando $0 < x$, $0 < y$. Hemos convertido así el problema inicial de extremos condicionados en uno de extremos libres porque ahora las variables (x,y) se mueven en el abierto $\mathbb{R}^+ \times \mathbb{R}^+$.

Por cierto, por la desigualdad de las medias se tiene que:

$$\sqrt[3]{(xy)(xz)(yz)} = \sqrt[3]{(xyz)^2} = \sqrt[3]{64} \leqslant \frac{xy + xz + yz}{3}$$

Y la igualdad se da si, y sólo si, $xy = xz = yz$ lo que implica que $x = y = z = 2$. El ortoedro buscado es el cubo.

Pero si cambiamos la condición anterior por la siguiente

$$x^2yz^3 + \operatorname{sen}(1 + xz) + y - \mathrm{e}^{yx} = 1$$

o bien, si imponemos dos condiciones como las siguientes:

$$\log(1 + x^2 + y^2) + \sin(1 + xz) - 1 = 0, \qquad \mathrm{e}^{1+y+x+z} + \cos(xyz) + x^2z^2 - 3 = 0$$

entonces no podemos usar esa condición (o condiciones) para despejar una variable (o dos variables) en función de las otras (de la otra).

Ocurre que muchos de los ejercicios que se proponen en los libros de texto para ser resueltos utilizando la teoría de extremos condicionados que vamos a desarrollar seguidamente, se resuelven con frecuencia de forma más sencilla usando otro tipo de técnicas o algunas desigualdades como la de las medias geométrica y aritmética o la de Cauchy - Schwarz. Pero aunque su utilidad práctica pueda ser discutida, la teoría de extremos condicionados tiene un gran interés teórico, especialmente en la formulación Lagrangiana de la Mecánica.

En términos generales, el problema clásico de extremos condicionados puede enunciarse como sigue: dados un abierto $\Omega \subset \mathbb{R}^n$, un campo escalar $f : \Omega \to \mathbb{R}$ de clase $\mathcal{C}^q$ y una función $\mathbf{g} : \Omega \to \mathbb{R}^{n-k}$ de clase $\mathcal{C}^q$ verificando que el conjunto:

$$M = \{\mathbf{x} \in \Omega : \mathbf{g}(\mathbf{x}) = \mathbf{0} \wedge \operatorname{rango} D\mathbf{g}(\mathbf{x}) = n-k\}$$

no es vacío, se trata de calcular los *extremos locales* de la función $f_{|M}$, **restricción** de f a M. Es decir, calcular los máximos y mínimos locales de $f(\mathbf{x})$ cuando las variables, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, están obligadas a satisfacer las condiciones dadas por:

$$\mathbf{g}(\mathbf{x}) = \mathbf{0} \Longleftrightarrow g_i(x_1, x_2, \dots, x_n) = 0 \quad 1 \leqslant i \leqslant n-k \tag{3.39}$$

donde $\mathbf{g} = (g_1, g_2, \dots, g_{n-k})$.

En Mecánica clásica, las ecuaciones (3.39) representan las ligaduras (que los físicos llaman *holónomas*) que obligan a un punto material o a un sistema más complejo a desplazarse en una determinada región del espacio, y el problema de optimizar ciertos campos escalares asociados al movimiento del sistema responde al problema de extremos condicionados.

El planteamiento del problema pone de manifiesto que M es una variedad en $\mathbb{R}^n$ de dimensión k y que se trata de determinar los extremos de la **restricción** de f a esa variedad.

3.30 Definición. Sea $f : \Omega \to \mathbb{R}$, donde $\Omega \subset \mathbb{R}^n$ y sea $M \subset \Omega$. Se dice que f tiene un *máximo* (resp. *mínimo*) *local condicionado* por M en un punto $\mathbf{a} \in M$, si hay un número $r > 0$ tal que para todo $\mathbf{x} \in B(\mathbf{a}, r) \cap M$ se verifica que $f(\mathbf{a}) \geqslant f(\mathbf{x})$ (resp. $f(\mathbf{a}) \leqslant f(\mathbf{x})$). Se dice que f tiene en $\mathbf{a}$ un *extremo local condicionado* por M cuando f tiene en $\mathbf{a}$ un máximo o un mínimo local condicionado por M. Cuando las desigualdades anteriores son estrictas para todo $\mathbf{x} \in B(\mathbf{x}, r)$ con $\mathbf{x} \neq \mathbf{a}$ se dice que el extremo local condicionado correspondiente es estricto.

El siguiente resultado es importante porque permite reducir un problema de extremos condicionados a un problema de extremos relativos ordinario.

3.31 Proposición. *Sean $f : \Omega \to \mathbb{R}$ un campo escalar de clase $\mathcal{C}^q$ en un abierto $\Omega \subset \mathbb{R}^n$ y $M \subset \Omega$ una variedad de clase $\mathcal{C}^q$ y dimensión k. Sean $\mathbf{a} \in M$, $(\wp, U)$ una parametrización local de M en $\mathbf{a}$ y $\mathbf{c}$ el vector de coordenadas de $\mathbf{a}$, es decir, $\wp(\mathbf{c}) = \mathbf{a}$. Se verifica entonces que f tiene en $\mathbf{a}$ un extremo local (estricto) condicionado por M si, y sólo si, el campo escalar $f \circ \wp : U \to \mathbb{R}$ tiene un extremo relativo (estricto) del mismo tipo en $\mathbf{c}$.*

Parece que todo está resuelto, pero no es así porque en la práctica pocas veces se conocen las parametrizaciones locales de una variedad.

3.32 Proposición. *Sean $f:\Omega\to\mathbb{R}$ un campo escalar de clase $\mathcal{C}^q$ en un abierto $\Omega\subset\mathbb{R}^n$, $M\subset\Omega$ una variedad de clase $\mathcal{C}^q$ y dimensión k y $\mathbf{a}\in M$. Supongamos que f tiene en $\mathbf{a}$ un extremo local condicionado por M. Entonces se verifica que $\nabla f(\mathbf{a})\in T_M^{\perp}(\mathbf{a})$.*

Demostración. Sea $(\wp,U)$ una parametrización local de M en $\mathbf{a}$ y sea $\wp(\mathbf{c})=\mathbf{a}$. En virtud de la proposición anterior, la hipótesis hecha es lo mismo que afirmar que $f\circ\wp$ tiene un extremo relativo en $\mathbf{c}$, en cuyo caso, como es un campo escalar diferenciable en $\mathbf{c}\in U$ y U es abierto, $\mathbf{c}$ debe ser un punto crítico de $f\circ\wp$, es decir, $D(f\circ\wp)(\mathbf{c})=\mathbf{0}$, esto es, $Df(\mathbf{a})\circ D\wp(\mathbf{c})=\mathbf{0}$. Por tanto, para todo $\mathbf{h}\in\mathbb{R}^k$ es

$$\big(Df(\mathbf{a})\circ D\wp(\mathbf{c})\big)(\mathbf{h})=\big\langle\nabla f(\mathbf{a})\big|D\wp(\mathbf{c})(\mathbf{h})\big\rangle=0$$

Como esto vale para todo $\mathbf{h}\in\mathbb{R}^k$, resulta que el vector $\nabla f(\mathbf{a})$ es ortogonal a $D\wp(\mathbf{c})(\mathbb{R}^k)=T_M(\mathbf{a})$. □

En los dos resultados anteriores no se ha supuesto que la variedad M venga dada de alguna forma concreta, sin embargo el resultado siguiente requiere en su enunciado conocer las ecuaciones implícitas locales de M en $\mathbf{a}$, por lo que de aquí en adelante vamos a suponer que la variedad M es la variedad de dimensión $k=n-m$ dada por una función $\mathbf{g}:\Omega\to\mathbb{R}^m$ de clase $\mathcal{C}^q$ en un abierto $\Omega\subset\mathbb{R}^n$ por:

$$M=\{\mathbf{x}\in\Omega:\mathbf{g}(\mathbf{x})=\mathbf{0}\wedge\operatorname{rango}D\mathbf{g}(\mathbf{x})=m\}$$

Notaremos $\mathbf{g}=(g_1,g_2,\dots,g_m)$ los campos escalares componentes de $\mathbf{g}$. Observa que esta hipótesis no es realmente restrictiva porque todas las variedades vienen dadas localmente de esta forma y el problema que nos ocupa es justamente un problema local. En lo que sigue $f:\Omega\to\mathbb{R}$ será un campo escalar de clase $\mathcal{C}^q$ en Ω.

El siguiente resultado proporciona una *condición necesaria* para que f tenga en $\mathbf{a}\in M$ un extremo local condicionado por M. Para enunciarlo conviene definir la llamada "función de Lagrange" $\mathcal{L}:\Omega\times\mathbb{R}^m\to\mathbb{R}$ dada por:

$$\mathcal{L}(\mathbf{x},\lambda)=f(\mathbf{x})+\sum_{i=1}^{m}\lambda_i g_i(\mathbf{x}),\quad \mathbf{x}=(x_1,x_2,\dots,x_n)\in\Omega,\ \lambda=(\lambda_1,\lambda_2,\dots,\lambda_m)\in\mathbb{R}^m \tag{3.40}$$

Observa que es una función de $n+m$ variables.

3.33 Teorema (de Lagrange). *Supongamos que f tiene en $\mathbf{a}\in M$ un extremo local condicionado por M. Entonces existe un único vector $\alpha\in\mathbb{R}^m$ tal que $(\mathbf{a},\alpha)$ es un punto crítico de $\mathcal{L}$, es decir, $\nabla\mathcal{L}(\mathbf{a},\alpha)=\mathbf{0}$.*

Demostración. Por el corolario 3.27 sabemos que los vectores $\nabla g_i(\mathbf{a})$, $1\leqslant i\leqslant m$ forman una base de $T_M^{\perp}(\mathbf{a})$. Por la proposición 3.32 sabemos que $\nabla f(\mathbf{a})\in T_M^{\perp}(\mathbf{a})$. Por tanto, hay números $\alpha_i\in\mathbb{R}$, $1\leqslant i\leqslant m$, únicos, tales que $\nabla f(\mathbf{a})+\sum_{i=1}^{m}\alpha_i\nabla g_i(\mathbf{a})=\mathbf{0}$. Poniendo $\alpha=(\alpha_1,\alpha_2,\dots,\alpha_m)$, esta igualdad vectorial equivale a las igualdades:

$$D_j f(\mathbf{a})+\sum_{i=1}^{m}\alpha_i D_j g_i(\mathbf{a})=0\Longleftrightarrow D_j\mathcal{L}(\mathbf{a},\alpha)=0\quad 1\leqslant j\leqslant n \tag{3.41}$$

Puesto que $D_{n+j}\mathcal{L}(\mathbf{a},\alpha)=g_j(\mathbf{a})=0$ para $1\leqslant j\leqslant m$ porque $\mathbf{a}\in M$, se sigue que las igualdades (3.41) equivalen a que $(\mathbf{a},\alpha)$ sea un punto crítico de $\mathcal{L}$. □

Los números $\alpha_1,\alpha_2,\dots,\alpha_m$ en el teorema anterior se llaman los *multiplicadores de Lagrange* del punto $\mathbf{a}$. Obtenemos así una regla práctica para proceder en un problema de extremos condicionados: se empieza por resolver el "sistema de Lagrange":

$$\left.\begin{array}{rcl} D_j f(x_1,x_2,\dots,x_n)+\displaystyle\sum_{i=1}^{m}\lambda_i D_j g_i(x_1,x_2,\dots,x_n) & = & 0 \\ g_j(x_1,x_2,\dots,x_n) & = & 0 \end{array}\right\}\quad 1\leqslant j\leqslant n+m$$

Es un sistema de $n+m$ ecuaciones en general no lineales para el que no puede haber técnicas generales de resolución. De hecho, puede resolverse de forma exacta en muy pocos casos.

El teorema de Lagrange, que no es más que una reformulación equivalente de la proposición 3.32, da una condición necesaria para que un punto $\mathbf{a}\in M$ sea extremo local de f condicionado por M. Daremos a continuación una condición suficiente. Suponemos en lo que sigue que nuestras funciones son de clase $\mathcal{C}^q$ con $q\geqslant 2$.

Dado $(\mathbf{a},\alpha)\in\Omega\times\mathbb{R}^m$, notaremos $\mathcal{Q}_{\mathcal{L}}(\mathbf{a},\alpha)$ la forma cuadrática sobre $\mathbb{R}^n$ cuya matriz en la base canónica es $\big(D_{ij}\mathcal{L}(\mathbf{a},\alpha)\big)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$. Se trata, pues, de la aplicación $\mathcal{Q}_{\mathcal{L}}(\mathbf{a},\alpha):\mathbb{R}^n\to\mathbb{R}$ definida por:

$$\mathcal{Q}_{\mathcal{L}}(\mathbf{a},\alpha)(\mathbf{x})=\sum_{i,j=1}^{n}D_{ij}\mathcal{L}(\mathbf{a},\alpha)x_ix_j\qquad\forall\mathbf{x}\in\mathbb{R}^n\tag{3.42}$$

Consideraremos también una parametrización local $(\wp,U)$ de M en $\mathbf{a}$ con $\wp(\mathbf{c})=\mathbf{a}$. En estas condiciones, sean $h=f\circ\wp:U\to\mathbb{R}$ y $\mathcal{Q}_h(\mathbf{c})$ la forma cuadrática sobre $\mathbb{R}^k$ dada por:

$$\mathcal{Q}_h(\mathbf{c})(\mathbf{u})=\sum_{i,j=1}^{k}D_{ij}h(\mathbf{c})u_iu_j\qquad\forall\mathbf{u}\in\mathbb{R}^k.$$

3.34 Proposición. *Sea $(\mathbf{a},\alpha)$ un punto crítico de la función de Lagrange. Con las notaciones anteriores, se verifica que:*

$$\mathcal{Q}_h(\mathbf{c})(\mathbf{u})=\mathcal{Q}_{\mathcal{L}}(\mathbf{a},\alpha)(D\wp(\mathbf{c})(\mathbf{u}))\qquad\forall\mathbf{u}\in\mathbb{R}^k\tag{3.43}$$

Demostración. Se trata de hacer un cálculo. La mayoría de los textos lo evitan. Vamos a hacerlo.

Notaremos por $\mathbf{e}_i$, $1\leqslant i\leqslant k$, los vectores de la base canónica de $\mathbb{R}^k$. Como h es un campo escalar, se verifica que $D_ih(\mathbf{c})=\langle\nabla h(\mathbf{c})|\mathbf{e}_i\rangle=Df(\mathbf{c})(\mathbf{e}_i)$. Para campos vectoriales, como $\wp$, esta relación también se mantiene si interpretamos que $D_i\wp(\mathbf{c})$ es la columna i-ésima de la matriz jacobiana $J_\wp(\mathbf{c})$. Dicha columna es el vector de $\mathbb{R}^n$ que se obtiene derivando las funciones componentes de $\wp$ con respecto a la variable i-ésima que notaremos por $D_i\wp(\mathbf{c})$. Por tanto $D_i\wp(\mathbf{c})=D\wp(\mathbf{c})(\mathbf{e}_i)$. Como hay que calcular derivadas parciales de segundo orden conviene calcular las derivadas parciales de primer orden en un punto genérico $\mathbf{u}\in U$ para volver a derivarlas usando la regla de la cadena. Para $\mathbf{u}\in U$ tenemos:

$$\begin{aligned} D_ih(\mathbf{u})&=D(f\circ\wp)(\mathbf{u})(\mathbf{e}_i)=\big(Df(\wp(\mathbf{u}))\circ D\wp(\mathbf{u})\big)(\mathbf{e}_i)=Df(\wp(\mathbf{u}))\big(D\wp(\mathbf{u})(\mathbf{e}_i)\big)=\\ &=\big\langle\nabla f(\wp(\mathbf{u}))\big|D_i\wp(\mathbf{u})\big\rangle=\big\langle(\nabla f)\circ\wp)(\mathbf{u})\big|D_i\wp(\mathbf{u})\big\rangle \end{aligned}\tag{3.44}$$

Y volviendo a derivar otra vez como se deriva un producto escalar:

$$D_{ij}h(\mathbf{u}) = D_j(D_i h(\mathbf{u})) = \left\langle D_j((\nabla f)\circ\wp)(\mathbf{u}) \middle| D_i\wp(\mathbf{u})\right\rangle + \left\langle \nabla f(\wp(\mathbf{u})) \middle| D_j(D_i\wp)(\mathbf{u})\right\rangle \tag{3.45}$$

Calculemos el primero de estos dos sumandos. Observa que $\nabla f : \mathbf{x} \mapsto \nabla f(\mathbf{x})$ es un campo vectorial cuyas funciones componentes son las derivadas parciales primeras de f, $(\nabla f)\circ\wp$ es el campo vectorial composición de ∇f y $\wp$, por lo que:

$$\begin{aligned} D_j((\nabla f)\circ\wp)(\mathbf{u}) &= D((\nabla f)\circ\wp)(\mathbf{u})(\mathbf{e}_j) = D(\nabla f)(\wp(\mathbf{u}))\circ D\wp(\mathbf{u})(\mathbf{e}_j) = \\ &= D(\nabla f)(\wp(\mathbf{u}))(D_j\wp(\mathbf{u})) \end{aligned} \tag{3.46}$$

Sustituyendo esta igualdad en (3.45) y haciendo $\mathbf{u} = \mathbf{c}$ tenemos:

$$D_{ij}h(\mathbf{c}) = \left\langle D(\nabla f)(\mathbf{a})(D_j\wp(\mathbf{c})) \middle| D_i\wp(\mathbf{c})\right\rangle + \left\langle \nabla f(\mathbf{a}) \middle| D_j(D_i\wp)(\mathbf{c})\right\rangle \tag{3.47}$$

Como $(\mathbf{a},\alpha)$ es un punto crítico de la función de Lagrange se verifica $\nabla f(\mathbf{a}) = -\sum_{l=1}^{m} \alpha_l \nabla g_l(\mathbf{a})$. Sustituyendo esta igualdad en (3.47) tenemos:

$$D_{ij}h(\mathbf{c}) = \left\langle D(\nabla f)(\mathbf{a})(D_j\wp(\mathbf{c})) \middle| D_i\wp(\mathbf{c})\right\rangle - \sum_{l=1}^{m} \alpha_l \left\langle \nabla g_l(\mathbf{a}) \middle| D_j(D_i\wp)(\mathbf{c})\right\rangle \tag{3.48}$$

Teniendo ahora en cuenta que $g_l\circ\wp(\mathbf{u}) = 0$ para todo $\mathbf{u}\in U$ y para $1 \leqslant l \leqslant m$, y usando la igualdad (3.47) con el campo escalar $h = f\circ\wp$ sustituido por $g_l\circ\wp$, obtenemos:

$$0 = D_{ij}(g_l\circ\wp)(\mathbf{c}) = \left\langle D(\nabla g_l)(\mathbf{a})(D_j\wp(\mathbf{c})) \middle| D_i\wp(\mathbf{c})\right\rangle + \left\langle \nabla g_l(\mathbf{a}) \middle| D_j(D_i\wp)(\mathbf{c})\right\rangle$$

de donde se sigue que:

$$\left\langle \nabla g_l(\mathbf{a}) \middle| D_j(D_i\wp)(\mathbf{c})\right\rangle = -\left\langle D(\nabla g_l)(\mathbf{a})(D_j\wp(\mathbf{c})) \middle| D_i\wp(\mathbf{c})\right\rangle$$

Sustituyendo esta igualdad en (3.48) resulta:

$$\begin{aligned} D_{ij}h(\mathbf{c}) &= \left\langle D(\nabla f)(\mathbf{a})(D_j\wp(\mathbf{c})) \middle| D_i\wp(\mathbf{c})\right\rangle + \left\langle \sum_{l=1}^{m} \alpha_l D(\nabla g_l)(\mathbf{a})(D_j\wp(\mathbf{c})) \middle| D_i\wp(\mathbf{c})\right\rangle = \\ &= \left\langle \left(D(\nabla f)(\mathbf{a}) + \sum_{l=1}^{m} \alpha_l D(\nabla g_l)(\mathbf{a}) \right)(D_j\wp(\mathbf{c})) \middle| D_i\wp(\mathbf{c})\right\rangle \end{aligned} \tag{3.49}$$

Pero esta igualdad es justamente lo que queríamos probar pues:

$$D_{ij}\mathcal{L}(\mathbf{a},\alpha) = D_{ij}f(\mathbf{a}) + \sum_{l=1}^{m} \alpha_l D_{ij}g_l(\mathbf{a})$$

Que coincide con el elemento ij de la matriz jacobiana de:

$$D(\nabla f)(\mathbf{a}) + \sum_{l=1}^{m} \alpha_l D(\nabla g_l)(\mathbf{a}) \tag{3.50}$$

Pues observa que la matriz jacobiana de $D(\nabla f)(\mathbf{a})$ es la matriz hessiana de f en $\mathbf{a}$ y que la matriz jacobiana de $D(\nabla g_l)(\mathbf{a})$ es la matriz hessiana de g_l en $\mathbf{a}$. Por tanto, la matriz jacobiana de (3.50) es justamente la matriz $\big(D_{ij}\mathcal{L}(\mathbf{a},\alpha)\big)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}$.

Lo que dice la igualdad (3.49) es que la forma cuadrática asociada a la matriz hessiana de $h = f\circ\wp$ en el punto $\mathbf{c}$ es igual a la **restricción** de la forma cuadrática asociada a la matriz $\big(D_{ij}\mathcal{L}(\mathbf{a},\alpha)\big)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$ al espacio tangente a M en $\mathbf{a}$. Para verlo más claro todavía, pongamos, por comodidad de escritura, $\mathcal{A} = \big(D_{ij}\mathcal{L}(\mathbf{a},\alpha)\big)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$. La forma cuadrática asociada a dicha matriz, es justamente la definida en (3.42), es decir, $\mathcal{Q}_{\mathcal{L}}(\mathbf{a},\alpha)$, y viene dada para todo $\mathbf{x}\in\mathbb{R}^n$ por:

$$\mathcal{Q}_{\mathcal{L}}(\mathbf{a},\alpha)(\mathbf{x}) = \mathbf{x}.\mathcal{A}.\mathbf{x}^t = \big\langle \mathcal{A}.\mathbf{x}^t \big| \mathbf{x}\big\rangle$$

Para todo $\mathbf{u} = (u_1,u_2,\dots,u_k)\in\mathbb{R}^k$ tenemos que:

$$D\wp(\mathbf{c})(\mathbf{u}) = \sum_{j=1}^{k} u_j D_j\wp(\mathbf{c})$$

es un vector del espacio tangente $T_M(\mathbf{a})$ y:

$$\begin{aligned}\mathcal{Q}_{\mathcal{L}}(\mathbf{a},\alpha)(D\wp(\mathbf{c})(\mathbf{u})) &= \left\langle \mathcal{A}.\left(\sum_{j=1}^{k} u_j D_j\wp(\mathbf{c})\right)^t \middle| \sum_{i=1}^{k} u_i D_i\wp(\mathbf{c})\right\rangle = \\ &= \sum_{i,j=1}^{k} u_i u_j \big\langle \mathcal{A}.(D_j\wp(\mathbf{c}))^t \big| D_i\wp(\mathbf{c})\big\rangle = \text{ por (3.49) } = \sum_{i,j=1}^{k} D_{ij}h(\mathbf{c})u_i u_j = \\ &= \mathcal{Q}_h(\mathbf{c})(\mathbf{u})\end{aligned}$$

Como queríamos probar. □

Naturalmente, para calcular la restricción de la forma cuadrática $\mathcal{Q}_{\mathcal{L}}(\mathbf{a},\alpha)$ al espacio $T_M(\mathbf{a})$, todo lo que se necesita es conocer una base de dicho espacio, pues si $\{\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_k\}$ es una tal base, la matriz de la forma cuadrática restringida a $T_M(\mathbf{a})$ es la matriz $(m_{ij})_{k\times k}$ cuyos elementos están dados por:

$$m_{ij} = \big\langle \mathcal{A}.\mathbf{a}_i^t \big| \mathbf{a}_j\big\rangle \qquad 1\leqslant i,j\leqslant k \tag{3.51}$$

Como consecuencia de la proposición 3.31, de la proposición anterior y del teorema 2.26, obtenemos condiciones suficientes para que f tenga en $\mathbf{a}$ un extremo local condicionado por M.

3.35 Teorema (Condición suficiente de extremo local condicionado)**.** *Con las notaciones y las hipótesis anteriores, supongamos que* $(\mathbf{a},\alpha)$ *es un punto crítico de la función de Lagrange y pongamos* $\mathcal{A} = \big(D_{ij}\mathcal{L}(\mathbf{a},\alpha)\big)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant n}}$. *Sea* $\{\mathbf{a}_1,\mathbf{a}_2,\dots,\mathbf{a}_k\}$ *una base del espacio* $T_M(\mathbf{a})$ *y sea* $(m_{ij})_{k\times k}$ *la matriz cuyos elementos* m_{ij} *vienen dados por* (3.51). *Sea* $\mathcal{Q}:\mathbb{R}^k\to\mathbb{R}$ *la forma cuadrática:*

$$\mathcal{Q}(\mathbf{u}) = \sum_{i,j=1}^{k} m_{ij}u_i u_j \qquad \forall \mathbf{u}\in\mathbb{R}^k$$

Se verifica entonces que:

a) Si la forma cuadrática $\mathcal{Q}$ *es definida positiva entonces* f *tiene en* $\mathbf{a}$ *un mínimo local estricto condicionado por* M.

b) Si la forma cuadrática $\mathcal{Q}$ es definida negativa entonces f tiene en **a** *un máximo local estricto condicionado por M.*

c) Si f tiene en **a** *un máximo local condicionado por M entonces la forma cuadrática $\mathcal{Q}$ es semidefinida negativa.*

d) Si f tiene en **a** *un mínimo local condicionado por M entonces la forma cuadrática $\mathcal{Q}$ es semidefinida positiva.*

e) Si la forma cuadrática $\mathcal{Q}$ es no definida entonces f no tiene en **a** *un extremo local condicionado por M.*

En general, la matriz $\mathcal{A} = \big(D_{ij}\mathcal{L}(\mathbf{a},\alpha)\big)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}$ depende de **a** y de α. Te digo esto porque es importante que calcules las soluciones del sistema de Lagrange en la forma $(\mathbf{a},\alpha)$ agrupando cada punto **a** con sus correspondientes multiplicadores de Lagrange. Cada **a** tiene asociado un único multiplicador α, pero puede ocurrir que diferentes **a** tengan asociado el mismo multiplicador porque distintos **a** pueden dar lugar al mismo $\nabla f(\mathbf{a})$. Los coeficientes m_{ij} dependen en general de **a** y de α y para cada par $(\mathbf{a},\alpha)$ resulta, en general, una forma cuadrática $\mathcal{Q}$ diferente cuyo carácter debes estudiar. Pese a todo, de hecho, en la práctica sucede con frecuencia que $\mathcal{A}$ no depende de α, en tal caso puedes olvidarte de los valores de α y considerar solamente los de **a**.

La forma de calcular rápidamente la matriz $(m_{ij})_{k\times k}$ es la siguiente: sea $\mathcal{B}_{n\times k}$ una matriz cuyas columnas sean los vectores base de $T_M(\mathbf{a})$, entonces:

$$(m_{ij})_{k\times k} = \mathcal{B}^t.\mathcal{A}.\mathcal{B} \tag{3.52}$$

3.6.1. Cálculo de extremos absolutos

Cuando la variedad M es un conjunto compacto, la existencia de extremos absolutos para f está garantizada por la propiedad de compacidad de las funciones continuas. En este tipo de problemas lo único que hay que hacer es calcular los puntos críticos de la función de Lagrange y evaluar en ellos la función f, lo que nos proporcionará directamente los puntos donde f alcanza su máximo y su mínimo absolutos en M. Claro está, se supone que hay un número finito de soluciones del sistema de Lagrange.

Situación diferente es cuando queremos calcular los extremos absolutos o relativos de una función f en un conjunto compacto $K \subset \mathbb{R}^n$ con interior no vacío y cuya frontera $M = \mathrm{Fr}(K)$ es una variedad. Por supuesto, la existencia de extremos absolutos está garantizada pero pueden alcanzarse en el interior o en la frontera. Por tanto, lo que hay que hacer es calcular los puntos críticos de f en el interior de K y las soluciones del sistema de Lagrange para f y M. Obtendremos así, en la práctica, un conjunto finito de puntos y todo lo que hay que hacer es evaluar f en cada uno de ellos, lo que nos proporcionará directamente los puntos donde f alcanza su máximo y su mínimo absolutos en K. Dicho sea de paso, con frecuencia dichos extremos absolutos se alcanzan en la frontera de K.

Sucede con frecuencia que la frontera de $K \subset \mathbb{R}^k$, $k = 2$ o $k = 3$, no es una variedad aunque se le parezca mucho; por ejemplo, la frontera puede ser un polígono o un ortoedro o una pirámide. En tales casos lo más sencillo es parametrizar las caras del polígono o del ortoedro y reducir el problema de calcular los extremos en cada parte de la frontera de K a un problema de extremos más sencillo en una o dos variables. Claro, también se puede descomponer la frontera de K como unión de variedades disjuntas: las caras o las aristas más los vértices y aplicar la teoría de extremos condicionados a cada una de ellas. No te lo aconsejo.

Ejercicios propuestos

189. Calcula, aplicando la teoría de extremos condicionados, el ortoedro de volumen igual a 8 y superficie lateral mínima.

Observación: este ejercicio lo hemos hecho en la introducción usando la desigualdad de las medias. La teoría de extremos condicionados te permitirá probar que el ortoedro solución solamente puede ser el cubo, pero dicha teoría no te dice si el mínimo alcanzado es absoluto.

190. Calcula la distancia mínima de un punto $P=(\alpha,\beta,\gamma)$ al plano $Ax+By+Cz=D$:

a) Usando la teoría de extremos condicionados.

b) Usando la desigualdad de Cauchy - Schwarz.

191. Calcula el punto de la superficie $4x^2+y^2-z^2=1$ que está más próximo al origen de coordendas.

192. Una curva está dada como la intersección del cilindro $x^2+y^2=4$ con el plano de ecuación $2x+3y+z=3$. Calcula el punto de la misma que está más próximo al origen.

193. Sea $\mathcal{A}$ una matriz simétrica $n\times n$. Aplica el método de los multiplicadores de Lagrange para calcular el máximo absoluto de la forma cuadrática asociada a $\mathcal{A}$ condicionado por la esfera euclídea unidad en $\mathbb{R}^n$, para probar que dicho máximo se alcanza en un punto de la forma $(\mathbf{a},\alpha)$ tal que $\mathcal{A}.\mathbf{a}^t=-\alpha\mathbf{a}$. Por tanto $-\alpha$ es un valor propio de $\mathcal{A}$.

194. Calcula el máximo absoluto de $f(x,y,z)=x^3y^2z$ en el conjunto:

$$M=\left\{(x,y,z)\in\mathbb{R}_o^+\times\mathbb{R}_o^+\times\mathbb{R}_o^+ :\ x+2y+3z=1\right\}$$

196. Utiliza el método de los multiplicadores de Lagrange para calcular un punto de la elipse de ecuación

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

tal que el segmento determinado por la intersección de la tangente a la elipse en dicho punto con los ejes coordenados tenga longitud mínima.

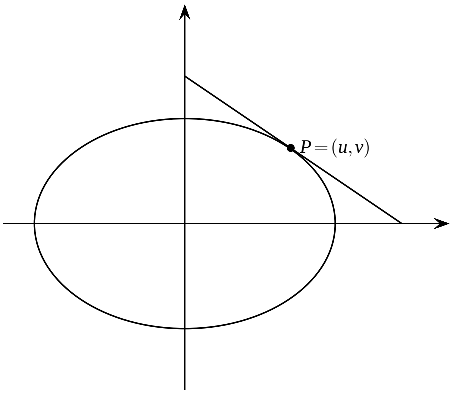
197. Prueba que el máximo absoluto de $f(x_1,x_2,\dots,x_n)=(x_1x_2\cdots x_n)^2$ sometido a la condición $\|(x_1,x_2,\dots,x_n)\|_2=r$ es igual a $(r^2/n)^n$. Deduce que:

$$\sqrt[n]{x_1^2x_2^2\cdots x_n^2}\leqslant\frac{1}{n}\sum_{i=1}^{n}x_i^2\qquad\forall(x_1,x_2,\dots,x_n)\in\mathbb{R}^n$$

195. Utiliza el método de los multiplicadores de Lagrange para calcular el área del triángulo de mínima área cuyos lados son los segmentos de los ejes coordenados positivos interceptados por la tangente a la elipse:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

en un punto $P=(u,v)$ de la misma situado en el primer cuadrante y dicho segmento de la tangente.



y deduce la desigualdad de las medias:

$$\sqrt[n]{x_1x_2\cdots x_n}\leqslant\frac{1}{n}\sum_{i=1}^{n}x_i \qquad \forall(x_1,x_2,\dots,x_n)\in(\mathbb{R}_o^+)^n$$

¿Cuándo se da la igualdad?

198. Utiliza el método de los multiplicadores de Lagrange para calcular el volumen del ortoedro de máximo volumen inscrito en el elipsoide de ecuación

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

¿Sabes resolver este ejercicio usando la desigualdad de las medias?

199. Prueba que el máximo absoluto de x^2y^2 sujeto a la condición

$$\frac{x^{2p}}{p}+\frac{y^{2q}}{q}=r^2$$

donde $p>0$, $q>0$ verifican que:

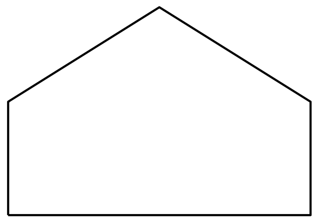
$$\frac{1}{p}+\frac{1}{q}=1$$

se alcanza cuando $x^{2p}=y^{2q}$ y es igual a r^2. Deduce que si $x>0$ e $y>0$ se verifica que:

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

¿Cuándo se da la igualdad?

200. Se quiere construir un pentágono colocando un triángulo isósceles sobre un rectángulo, como se muestra en la figura de la derecha. Si el pentágono tiene un perímetro fijo p, determinar las longitudes de los lados del pentágono que maximizan su área.



201. Calcula los puntos de la curva

$$\begin{cases} x^2 - xy + y^2 - z^2 = 1 \\ x^2 + y^2 = 1 \end{cases}$$

que están más próximos al origen de coordenadas.

202. Calcula la mínima distancia del origen a la superficie de ecuación $xy^2z^3 = 2$.

203. Calcula el mayor y el menor valor de la función $f(x,y,z) = x - 2y + 5z$ en la esfera $x^2 + y^2 + z^2 = 30$.

204. Sea Γ la curva intersección de la esfera $x^2 + y^2 + z^2 = 1$ y el plano $x + y + z = 1$. Calcula los puntos de Γ que están más cerca y más lejos del punto $(1,2,3)$. Justifica que los resultados obtenidos son valores máximos y mínimos absolutos.

205. Encontrar un punto P de coordenadas positivas perteneciente al elipsoide de ecuación $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$, tal que el plano tangente al elipsoide en P determine con los ejes coordenados un tetraedro de volumen mínimo.

El volumen del tetraedro con vértices en el origen y en los puntos $(a,0,0)$, $(0,b,0)$ y $(0,0,c)$ (donde a,b,c son números positivos), es igual a $\dfrac{1}{6}abc$.

206. Calcula los extremos absolutos de $f(x,y) = x^2 + y^2 - x - y$ en el conjunto

$$K = \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant 4 \wedge y \geqslant 0\right\}$$

207. Calcula el volumen máximo de una caja rectangular situada sobre el plano XY que tiene un vértice en el origen y el vértice opuesto, de coordenadas positivas, pertenece al paraboloide $z + x^2 + 4y^2 = 4$.

208. Calcula los extremos absolutos de $f(x,y) = x^2y^3 + x^2 + 6y^2$ en la bola cerrada $\overline{B}_{\| \|_2}((0,0),\sqrt{5})$.

209. Calcula los extremos absolutos de $f(x,y,z) = x^2 + yz$ en la bola euclídea unidad cerrada.

210. a) Clasifica los extremos relativos del campo escalar $f(x,y) = x^3 + y^3 - xy^2 - x + 16$.

b) Calcula el máximo y el mínimo absolutos de dicho campo escalar en el conjunto:

$$K = \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant 2,\ y \geqslant 0\right\}.$$

211. Se considera la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por:

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

a) Calcula y clasifica los puntos críticos de f en la bola abierta $B_{\| \|_2}((0,0),2)$.

b) Calcula los extremos absolutos de f en la bola cerrada $\overline{B}_{\| \|_2}((0,0),2)$.

212. Calcula los puntos del conjunto $A = \left\{(x,y)\in\mathbb{R}^2 : x^2+y^2 \leqslant 2x,\ y \geqslant 0\right\}$ donde el campo escalar $f(x,y) = x^2-2xy+y^2$ alcanza su máximo y mínimo absolutos.

213. Calcula los extremos absolutos de $f(x,y) = (x^2+2y^2)\,\mathrm{e}^{-x^2-y^2}$ en el disco $x^2+y^2 \leqslant 4$.

214. Calcula los valores máximos y mínimos absolutos de $f(x,y,z) = xy^2z^3$ en la bola $x^2+y^2+z^2 \leqslant 1$.

215. Calcula los extremos absolutos de $f(x,y) = x^2+3y^2$ en el círculo $x^2-2x+y^2-3 \leqslant 0$.

216. Calcula los extremos absolutos de la función $f(x,y) = x^2y^3(1-x-y)$ en el conjunto

$$K = \{(x,y) : |x|+|y| \leqslant 1\}$$

217. Calcula los extremos absolutos de $f(x,y) = x^2+y^2-xy-x-y$ en el conjunto

$$K = \left\{(x,y)\in\mathbb{R}^2 : x \geqslant 0, y \geqslant 0, x+y \leqslant 3\right\}$$

218. Calcula los extremos absolutos del campo escalar $f(x,y,z) = x+y+z$ en el conjunto

$$A = \left\{(x,y,z)\in\mathbb{R}^3 : x^2+y^2 \leqslant z \leqslant 1\right\}.$$

219. Calcula los valores máximo y mínimo de la función $f(x,y,z) = xyz$ cuando el punto (x,y,z) pertenece a la curva definida por la intersección del plano $x+y+z=0$ y la esfera $x^2+y^2+z^2-1=0$.

220. Calcula la mínima distancia entre la recta $x+y=4$ y la circunferencia $x^2+y^2=1$.

221. Calcula la mínima distancia entre la recta $x-y=2$ y la parábola $y=x^2$.

221. Calcula la distancia mínima entre la elipse $x^2+2y^2=6$ y la recta $x+y=5$.

222. El área de una caja rectangular sin tapa es de 108cm^2. Calcula sus dimensiones para que el volumen sea máximo.

223. Calcula el valor mayor y el valor menor que toma la función $f(x,y,z) = xyz$ en los puntos del elipsoide $x^2+4y^2+9z^2=3$.

224. Calcula el valor mayor y el valor menor que toma la función

$$f(x,y,z) = y^2+4z^2-4yz-2xz-2xy$$

en los puntos del elipsoide $2x^2+3y^2+6z^2=1$.

225. Calcula los puntos de la elipse $5x^2-6xy+5y^2=4$ para los cuales la distancia del origen a la tangente a la elipse en tales puntos es máxima o mínima.

226. Calcula y clasifica los extremos relativos de la función

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

Y calcula el mínimo valor que dicha función toma en la bola euclídea cerrada de centro el origen y radio 4.

227. Calcula por el método de los multiplicadores de Lagrange la mínima distancia entre la recta $x + y = 3$ y la elipse $3x^2 + y^2 = 3$.