Capítulo 6

Funciones dos veces diferenciables

Derivadas parciales de segundo orden. Aplicaciones dos veces diferenciables: Teorema de Young sobre permutabilidad del orden de las derivaciones. Desarrollo de Taylor de orden 2. Extremos relativos. Funciones convexas.

Este capítulo está dedicado a las funciones dos veces diferenciables y sus principales aplicaciones: La discusión de la naturaleza de los extremos locales y el estudio de las funciones convexas.

La noción de aplicación diferenciable dos veces la formulamos sólo en el caso de funciones de varias variables reales, tomando como base la diferenciabilidad de las derivadas parciales. El primer resultado básico es el teorema de Young 6.4 sobre la simetría de las derivadas parciales de segundo orden, con el que se obtiene que la diferencial segunda $d^2f(\mathbf{a})(\mathbf{h},\mathbf{k})$ es una aplicación bilineal simétrica que restringida a la diagonal produce una forma cuadrática $Q_{\mathbf{a}}(\mathbf{h}) = d^2f(\mathbf{a})(\mathbf{h},\mathbf{h})$ con la que se consigue una aproximación local de la función mejor que la proporcionada por la diferencial primera. Esta aproximación se establece de modo preciso mediante el desarrollo de Taylor de segundo orden (un anticipo de la versión general que se considerará en el capítulo 7) con el que se aborda la discusión de la naturaleza de los extremos locales y el estudio de las funciones convexas.

Como material complementario, en el apéndice **??** el lector interesado puede ver, como alternativa razonable al teorema de Young, el clásico teorema de Schwarz, sobre la igualdad de las derivadas mixtas. También puede encontrar allí una formulación equivalente de la noción de función dos veces diferenciable en la que no intervienen las derivadas parciales, lo que permite extender la definición al caso de funciones cuyo dominio no es finito dimensional. En el apéndice F se repasan y amplían los resultados básicos sobre las funciones convexas de una variable que se suelen estudiar en el curso de Análisis Matemático I. También se completa el estudio de las funciones convexas de varias variables demostrando que toda función convexa definida en un subconjunto abierto y convexo de $\mathbb{R}^n$ es continua.

6.1. Funciones dos veces diferenciables.

Sea $\mathbf{f} : \Omega \to F$ una aplicación definida en un abierto Ω del espacio normado $(E, \| \|)$, con valores en el espacio normado $(F, \| \|)$. Supongamos que $\mathbf{a} \in \Omega$ posee un entorno abierto $V_\mathbf{a} \subset \Omega$ tal que para todo $\mathbf{x} \in V_\mathbf{a}$ existe la derivada $D_\mathbf{v}\mathbf{f}(\mathbf{x})$ según el vector $\mathbf{v} \in E$. Si la aplicación $\mathbf{x} \to D_\mathbf{v}\mathbf{f}(\mathbf{x})$, definida en $V_\mathbf{a}$ con valores en F, es derivable en $\mathbf{a}$ según el vector $\mathbf{u} \in E$, entonces la derivada $D_\mathbf{u}(D_\mathbf{v}\mathbf{f})(\mathbf{a})$, denotada $D_{\mathbf{uv}}\mathbf{f}(\mathbf{a})$, se llama *derivada segunda* de $\mathbf{f}$ en $\mathbf{a}$ según el par de vectores $(\mathbf{u}, \mathbf{v}) \in E^2$:

$$D_{\mathbf{uv}}\mathbf{f}(\mathbf{a}) = D_\mathbf{u} D_\mathbf{v}\mathbf{f}(\mathbf{a})$$

Si $E = \mathbb{R}^n$, cuando $\mathbf{u} = \mathbf{e}_i$, y $\mathbf{v} = \mathbf{e}_j$ son vectores de la base canónica, como caso particular resulta la definición de las *derivadas parciales de segundo orden*, $D_{\mathbf{e}_i\mathbf{e}_j}\mathbf{f}(\mathbf{a})$, para las que se utilizan las notaciones

$$D_{ij}\mathbf{f}(\mathbf{a}) = \frac{\partial^2 \mathbf{f}}{\partial x_i \partial x_j}(\mathbf{a}) \quad \text{si } i \neq j, \quad D_{ii}\mathbf{f}(\mathbf{a}) = \frac{\partial^2 \mathbf{f}}{\partial x_i^2}(\mathbf{a}).$$

Habitualmente las derivadas parciales de segundo orden se pueden calcular fácilmente haciendo uso de las reglas del cálculo diferencial de funciones de una variable: Para obtener $D_{ij}\mathbf{f}(\mathbf{a})$ basta derivar respecto a la variable x_i (en el punto $\mathbf{a}$) la función de n variables reales $(x_1, x_2, \cdots x_n) \to D_j\mathbf{f}(x_1, x_2, \cdots x_n)$.

Si $E = \mathbb{R}^3$ es costumbre que (x, y, z) designe un punto genérico de $\mathbb{R}^3$, y las derivadas parciales segundas de una función de tres variables reales $\mathbf{f}(x, y, z)$ se escriben en la forma

$$\frac{\partial^2 \mathbf{f}}{\partial x^2}, \ \frac{\partial^2 \mathbf{f}}{\partial y^2}, \ \frac{\partial^2 \mathbf{f}}{\partial z^2}, \ \frac{\partial^2 \mathbf{f}}{\partial x \partial y}, \ \frac{\partial^2 \mathbf{f}}{\partial y \partial x}, \ \frac{\partial^2 \mathbf{f}}{\partial x \partial z}, \ \frac{\partial^2 \mathbf{f}}{\partial z \partial x}, \ \frac{\partial^2 \mathbf{f}}{\partial y \partial z}, \ \frac{\partial^2 \mathbf{f}}{\partial z \partial y}.$$

donde, para simplificar la escritura, hemos omitido el punto donde se evalúan las derivadas. A veces es más cómodo utilizar la notación $\mathbf{f}_x$, $\mathbf{f}_y$, $\mathbf{f}_z$ para las derivadas parciales primeras y $\mathbf{f}_{xx}$, $\mathbf{f}_{xy}$, $\mathbf{f}_{xz}$, $\mathbf{f}_{yx}$, $\mathbf{f}_{yy}$, $\mathbf{f}_{yz}$, $\mathbf{f}_{zx}$, $\mathbf{f}_{zy}$, $\mathbf{f}_{zz}$ para las derivadas parciales segundas. Cuando se usa esta notación, $\mathbf{f}_{xy}$ es una abreviatura de $(\mathbf{f}_x)_y$, luego

$$\mathbf{f}_{xy} = \frac{\partial^2 \mathbf{f}}{\partial y \partial x}$$

de manera que se invierte el orden de la x y la y en las dos notaciones.

Más adelante veremos que bajo ciertas condiciones se puede asegurar la coincidencia de las llamadas *derivadas mixtas* $\mathbf{f}_{xy} = \mathbf{f}_{yx}$, $\mathbf{f}_{xz} = \mathbf{f}_{zx}$, $\mathbf{f}_{yz} = \mathbf{f}_{zy}$, de modo que en el caso de funciones de tres variables sólo habrá 3 derivadas mixtas distintas.

Se introducen notaciones análogas para funciones de distinto número de variables. Así por ejemplo, si $\mathbf{f}(x, y, z, t)$ es una función de cuatro variables, hay 16 derivadas parciales segundas: $\mathbf{f}_{xx}$, $\mathbf{f}_{xy}$, $\mathbf{f}_{xz}$, $\mathbf{f}_{xt}$,..... $\mathbf{f}_{tx}$, $\mathbf{f}_{ty}$, $\mathbf{f}_{tz}$, $\mathbf{f}_{tt}$. y cuando haya coincidencia de las derivadas mixtas sólo habrá 6 derivadas mixtas distintas.

Definición 6.1 *Una aplicación $\mathbf{f} : \Omega \to F$, definida en un abierto $\Omega \subset \mathbb{R}^n$, es diferenciable dos veces en el punto $\mathbf{a} \in \Omega$ cuando se verifican las condiciones:*

i) $\mathbf{f}$ *es diferenciable en un entorno abierto de* $\mathbf{a}$*,* $V_{\mathbf{a}} \subset \Omega$*.*

ii) *Las derivadas parciales* $D_i\mathbf{f}(\mathbf{x})$*,* $1 \leq i \leq n$*, que están definidas en* $V_{\mathbf{a}}$*, son diferenciables en* $\mathbf{a}$*.*

Si $\mathbf{f}$ *es diferenciable dos veces en cada* $\mathbf{x} \in \Omega$ *se dice que* $\mathbf{f}$ *es diferenciable dos veces en* Ω*. Si en cada* $\mathbf{x} \in \Omega$ *existen todas las derivadas parciales segundas y son continuas, se dice que* $\mathbf{f}$ *es de clase* C^2 *en* Ω *y se escribe* $\mathbf{f} \in C^2(\Omega, F)$*.*

OBSERVACIONES:

a) Si en cada $\mathbf{x} \in \Omega$ existen las derivadas parciales $D_i\mathbf{f}(\mathbf{x})$, $1 \leq i \leq n$, y son diferenciables en Ω entonces $\mathbf{f}$ es diferenciable dos veces en Ω: Basta tener en cuenta que la condición i) de 6.1 se cumple en todo punto de Ω ya que $\mathbf{f}$ tiene derivadas parciales continuas (porque son diferenciables).

b) Toda función de clase $C^2(\Omega, F)$ es diferenciable dos veces en Ω (las derivadas parciales $D_i\mathbf{f}(\mathbf{x})$, $1 \leq i \leq n$, son diferenciables porque tienen derivadas parciales continuas) y toda función diferenciable dos veces en Ω es de clase $C^1(\Omega, F)$ ya que las derivadas parciales $D_i\mathbf{f}(\mathbf{x})$, $1 \leq i \leq n$, son continuas (por ser diferenciables).

Comenzamos obteniendo una condición suficiente, de tipo local, para que una función sea diferenciable dos veces en un punto, similar a la condición suficiente de diferenciabilidad.

Proposición 6.2 *Sea* $\mathbf{f} : \Omega \to F$ *definida en un abierto* $\Omega \subset \mathbb{R}^n$*, con valores en el espacio normado* $(F, \| \|)$*. Se supone que* $\mathbf{a} \in \Omega$ *posee un entorno abierto* $V_{\mathbf{a}} \subset \Omega$ *donde existen las derivadas parciales segundas* $D_{ij}\mathbf{f}(\mathbf{x})$*,* $1 \leq i, j \leq n$*, y todas son continuas en* $\mathbf{a}$*. Entonces* $\mathbf{f}$ *es diferenciable dos veces en* $\mathbf{a}$*.*

DEM: Veamos en primer lugar que se cumple la condición i) de la definición 6.1), es decir, que existe un entorno abierto de $\mathbf{a}$, $U_{\mathbf{a}} \subset V_{\mathbf{a}}$ donde $\mathbf{f}$ es diferenciable.

Como todas las derivadas parciales segundas $D_{ij}\mathbf{f} = D_i(D_j\mathbf{f})$, que están definidas en $V_{\mathbf{a}}$, son continuas en $\mathbf{a}$, existe un entorno abierto $U_{\mathbf{a}} \subset V_{\mathbf{a}}$ donde todas ellas están acotadas. Como cada derivada primera $D_j\mathbf{f}$ tiene derivadas parciales acotadas en $U_{\mathbf{a}}$, y las funciones con derivadas parciales acotadas son continuas (5.5) podemos asegurar que las derivadas parciales $D_i\mathbf{f} : U_{\mathbf{a}} \to F$, $1 \leq i \leq n$, son continuas en $U_{\mathbf{a}}$ y por lo tanto $\mathbf{f}$ es diferenciable en $U_{\mathbf{a}}$ (condición suficiente de diferenciabilidad 5.16).

Según la hipótesis, las funciones $D_i\mathbf{f}(\mathbf{x})$, $1 \leq i \leq n$, que están definidas en $U_{\mathbf{a}}$, tienen derivadas parciales continuas $\mathbf{a}$, y por lo tanto son diferenciables en $\mathbf{a}$, de modo que también se cumple la condición ii) de la definición 6.1. ∎

El siguiente ejemplo muestra que en general no se puede asegurar la igualdad $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$.

Ejemplo 6.3 *La función* $f : \mathbb{R}^2 \to \mathbb{R}$, *definida por*

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

es diferenciable en todo punto pero $D_{12}f(0,0) \neq D_{21}f(0,0)$.

Dem: Es fácil ver que las derivadas parciales $D_1f(x,y)$, $D_2f(x,y)$ existen en todo punto, y con un cálculo elemental se obtiene el valor

$$D_1f(x,y) = \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2} \quad \text{si} \quad (x,y) \neq 0, \quad D_1f(0,0) = 0$$

La función D_1f es continua en todo punto: La continuidad en los puntos $(x,y) \neq (0,0)$ es inmediata y la continuidad en $(0,0)$ se obtiene fácilmente considerando coordenadas polares, $x = r\cos\theta$, $y = r\operatorname{sen}\theta$, con las que se establece la acotación $|D_1f(x,y)| \leq 5r = 5\sqrt{x^2+y^2}$, de donde se sigue que

$$\lim_{(x,y) \to (0,0)} D_1f(x,y) = 0 = D_1f(0,0)$$

Análogamente $D_2f(x,y)$ es continua en cada $(x,y) \in \mathbb{R}^2$ y con el teorema 5.16 se concluye que f es diferenciable en todo punto. Como $D_1f(0,y) = -y$ se sigue que $D_{21}f(0,0) = -1$, y análogamente se obtiene que $D_{12}f(0,0) = 1$. ■

El siguiente teorema proporciona una condición suficiente para que se cumpla la igualdad $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$.

Teorema 6.4 [Young] *Sea* $\mathbf{f} : \Omega \to F$ *definida en un abierto* $\Omega \subset \mathbb{R}^n$. *Se supone que* $\mathbf{a} \in \Omega$ *posee un entorno* $V_\mathbf{a} \subset \Omega$ *donde existen las derivadas parciales* $D_i\mathbf{f}$, $D_j\mathbf{f}$ $(i \neq j)$ *y ambas son diferenciables en* $\mathbf{a}$. *Entonces se verifica* $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$.

En particular, si $\mathbf{f}$ *es diferenciable dos veces en* $\mathbf{a}$, *para cada par de índices* $1 \leq i < j \leq n$, *se cumple* $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$.

Dem: Como el teorema sólo involucra dos variables, basta considerar el caso $n = 2$. Para simplificar la notación suponemos $\mathbf{a} = (0,0)$ y $V_\mathbf{a} = (-r,r)\times(-r,r)$. Tomando $0 < h < r$ podemos asegurar que el cuadrado de vértices $(0,0)$, $(h,0)$, $(0,h)$, (h,h) está contenido en V. Demostraremos que $\Delta(h) = \mathbf{f}(h,h) - \mathbf{f}(h,0) - \mathbf{f}(0,h) + \mathbf{f}(0,0)$ verifica

$$D_{12}\mathbf{f}(0,0) = \lim_{h \to 0+} \frac{\Delta(h)}{h^2} = D_{21}\mathbf{f}(0,0)$$

Obsérvese que $\Delta(h) = \mathbf{g}(h) - \mathbf{g}(0)$, donde $\mathbf{g}(x) = \mathbf{f}(x,h) - \mathbf{f}(x,0)$ es derivable en cada $x \in [0,h]$, con $\mathbf{g}'(x) = D_1\mathbf{f}(x,h) - D_1\mathbf{f}(x,0)$.

Como $D_1\mathbf{f}$ es diferenciable en $(0,0)$ se tiene

i) $D_1\mathbf{f}(x,h) = D_1\mathbf{f}(0,0) + D_{11}\mathbf{f}(0,0)x + D_{21}\mathbf{f}(0,0)h + \|(x,h)\|_\infty \boldsymbol{\rho}(x,h)$.

ii) $D_1\mathbf{f}(x,0) = D_1\mathbf{f}(0,0) + D_{11}\mathbf{f}(0,0)x + \|(x,0)\|_\infty \boldsymbol{\rho}(x,0)$.

donde $\boldsymbol{\rho}(x, y)$ tiende hacia 0 cuando $\|(x, y)\|_\infty \to 0$.

Sustituyendo i) y ii) en la expresión $\mathbf{g}'(x) = D_1\mathbf{f}(x, h) - D_1\mathbf{f}(x, 0)$ resulta

$$\mathbf{g}'(x) = hD_{21}\mathbf{f}(0, 0) + \|(x, h)\|_\infty \boldsymbol{\rho}(x, h) - \|(x, 0)\|_\infty \boldsymbol{\rho}(x, 0)$$

La función $\boldsymbol{\varphi}(x) = \mathbf{g}(x) - D_{21}\mathbf{f}(0, 0)hx$ está definida para $0 \leq x \leq h$, y

$$\boldsymbol{\varphi}'(x) = \mathbf{g}'(x) - D_{21}\mathbf{f}(0, 0)h = \|(x, h)\|_\infty \boldsymbol{\rho}(x, h) - \|(x, 0)\|_\infty \boldsymbol{\rho}(x, 0)$$

Dado $\epsilon > 0$, usando la definición de límite, podemos encontrar $0 < r' < r$ tal que

$$\|(x, y)\|_\infty < r' \Rightarrow \|\boldsymbol{\rho}(x, y)\| < \epsilon/2$$

Si $0 \leq x \leq h < r'$ se cumple $\|(x, h)\|_\infty < r'$, $\|(x, 0)\|_\infty < r'$, luego

$$\|\boldsymbol{\varphi}'(x)\| < h\epsilon/2 + h\epsilon/2 = \epsilon h$$

y aplicando el teorema del incremento finito resulta $\|\boldsymbol{\varphi}(h) - \boldsymbol{\varphi}(0)\| \leq \epsilon h^2$, es decir

$$\left\|(\mathbf{f}(h, h) - \mathbf{f}(h, 0) - D_{21}\mathbf{f}(0, 0)h^2) - (\mathbf{f}(0, h) - \mathbf{f}(0, 0))\right\| < \epsilon h^2$$

$$\left\|\Delta(h) - D_{21}\mathbf{f}(0, 0)h^2\right\| < \epsilon h^2$$

Hemos demostrado así que

$$0 < h < r' \Rightarrow \left\|\frac{\Delta(h)}{h^2} - D_{21}\mathbf{f}(0, 0)\right\| < \epsilon$$

es decir:

$$\lim_{h \to 0+} \frac{\Delta(h)}{h^2} = D_{21}\mathbf{f}(0, 0)$$

Análogamente, considerando la función auxiliar $\mathbf{g}(y) = \mathbf{f}(h, y) - \mathbf{f}(0, y)$, que cumple $\Delta(h) = \mathbf{g}(h) - \mathbf{g}(0)$, usando ahora que $D_2\mathbf{f}$ es diferenciable en $(0, 0)$, y estimando la derivada $\mathbf{g}'(y) = D_2\mathbf{f}(h, y) - D_2\mathbf{f}(0, y)$, con un razonamiento paralelo se llega a

$$\lim_{h \to 0+} \frac{\Delta(h)}{h^2} = D_{12}\mathbf{f}(0, 0)$$

Finalmente, si $\mathbf{f}$ es diferenciable dos veces en $\mathbf{a}$, según la definición 6.1 existe $V_\mathbf{a} \subset \Omega$, entorno abierto de $\mathbf{a}$, donde $\mathbf{f}$ es diferenciable y además todas las derivadas parciales $D_i\mathbf{f}(x)$, $1 \leq i \leq n$, son diferenciables en $\mathbf{a}$. Por lo tanto podemos aplicar el resultado demostrado a todas las derivadas mixtas $D_{ij}\mathbf{f}(\mathbf{a})$. ■

Corolario 6.5 *Sea* $\mathbf{f} : \Omega \to F$ *definida en un abierto* $\Omega \subset \mathbb{R}^n$*. Si* $\mathbf{a} \in \Omega$ *posee un entorno abierto* $V_\mathbf{a} \subset \Omega$ *donde existen todas las derivadas parciales* $D_{ij}\mathbf{f}(\mathbf{x})$*,* $1 \leq i, j \leq n$*, y son continuas en* $\mathbf{a}$*, entonces* $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$*.*

Dem: Es consecuencia directa de la proposición 6.2 y del teorema de Young 6.4. ■

Corolario 6.6 *Sea* $\mathbf{f} : \Omega \to F$ *definida en un abierto* $\Omega \subset \mathbb{R}^2$. *Si* $\mathbf{a} \in \Omega$ *posee un entorno* $V_{\mathbf{a}} \subset \Omega$ *donde existen las cuatro derivadas parciales* $D_{ij}\mathbf{f}$, $1 \leq i, j \leq 2$ *y* $D_{12}\mathbf{f}$ *y* $D_{21}\mathbf{f}$ *son continuas en* $\mathbf{a}$ *entonces* $D_{12}\mathbf{f}(\mathbf{a}) = D_{21}\mathbf{f}(\mathbf{a})$.

DEM: $D_1\mathbf{f}$ es diferenciable en $\mathbf{a}$ porque en un entorno de $\mathbf{a}$ existen sus dos derivadas parciales y una de ellas es continua en $\mathbf{a}$. Lo mismo se puede decir de $D_2\mathbf{f}$ y por lo tanto se puede aplicar el teorema de Young 6.4. ■

Proposición 6.7 *Sea* $\mathbf{f} : \Omega \to F$ *definida en un abierto* $\Omega \subset \mathbb{R}^n$. *Si* $\mathbf{f}$ *es diferenciable dos veces en* $\mathbf{a} \in \Omega$ *entonces para cada par* $(\mathbf{u}, \mathbf{v}) \in \mathbb{R}^n \times \mathbb{R}^n$ *existen y son iguales las derivadas segundas* $D_{\mathbf{uv}}\mathbf{f}(\mathbf{a})$, $D_{\mathbf{vu}}\mathbf{f}(\mathbf{a})$, *que vienen dadas por*

$$D_{\mathbf{uv}}\mathbf{f}(\mathbf{a}) = D_{\mathbf{vu}}\mathbf{f}(\mathbf{a}) = \sum_{i,j=1}^{n} D_{ij}\mathbf{f}(\mathbf{a})u_i v_j$$

DEM: Según la definición 6.1 $\mathbf{f}$ es diferenciable en un entorno abierto de $\mathbf{a}$, $V_{\mathbf{a}} \subset \Omega$, luego en cada $\mathbf{x} \in V_{\mathbf{a}}$ existe la derivada $D_{\mathbf{v}}\mathbf{f}(\mathbf{x}) = d\mathbf{f}(\mathbf{x})\mathbf{v}$, que viene dada por

$$D_{\mathbf{v}}\mathbf{f}(\mathbf{x}) = D_1\mathbf{f}(\mathbf{x})v_1 + D_2\mathbf{f}(\mathbf{x})v_2 + \cdots + D_n\mathbf{f}(\mathbf{x})v_n$$

La función $\mathbf{g}(\mathbf{x}) = D_{\mathbf{v}}\mathbf{f}(\mathbf{x})$, definida en $V_{\mathbf{a}}$, es diferenciable en $\mathbf{a}$ (por ser suma de funciones diferenciables en $\mathbf{a}$) y por lo tanto existe su derivada según el vector $\mathbf{u}$

$$D_{\mathbf{u}}\mathbf{g}(\mathbf{a}) = \sum_{i=1}^{n} D_i\mathbf{g}(\mathbf{a})u_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} D_i(D_j\mathbf{f})(\mathbf{a})v_j \right) u_i$$

Es decir

$$D_{\mathbf{u}}(D_{\mathbf{v}}\mathbf{f})(\mathbf{a}) = \sum_{i,j=1}^{n} D_{ij}\mathbf{f}(\mathbf{a})u_i v_j$$

Usando esta fórmula y la simetría de las derivadas segundas $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$ (teorema 6.4) se concluye que $D_{\mathbf{uv}}\mathbf{f}(\mathbf{a}) = D_{\mathbf{vu}}\mathbf{f}(\mathbf{a})$. ■

Definición 6.8 *Si* $\mathbf{f} : \Omega \to F$, *definida en un abierto* $\Omega \subset \mathbb{R}^n$, *es diferenciable dos veces en* $\mathbf{a} \in \Omega$ *la diferencial segunda de* $\mathbf{f}$ *en* $\mathbf{a}$ *es la aplicación bilineal simétrica* $d^2\mathbf{f}(\mathbf{a}) : \mathbb{R}^n \times \mathbb{R}^n \to F$, *definida por*

$$d^2 f(\mathbf{a})(\mathbf{u}, \mathbf{v}) = D_{\mathbf{uv}}\mathbf{f}(\mathbf{a}) = \sum_{i,j=1}^{n} D_{ij}\mathbf{f}(\mathbf{a})u_i v_j$$

En las condiciones de 6.8, en virtud de la simetría, de las n^2 derivadas segundas $D_{ij}\mathbf{f}(\mathbf{a})$ hay muchas que son iguales; hay a lo más $n(n-1)/2$ derivadas parciales segundas que pueden ser diferentes. Cuando $f : \Omega \to \mathbb{R}$ es una función con valores reales diferenciable dos veces en $\mathbf{a}$, la matriz simétrica formada con las derivadas parciales segundas, $(D_{ij}f(\mathbf{a}))$, se llama matriz *Hessiana* de f en el punto $\mathbf{a}$.

Si $\mathbf{f}$ es diferenciable dos veces en $\mathbf{a} \in \Omega$, la expresión

$$Q_{\mathbf{a}}(\mathbf{h}) = d^2\mathbf{f}(\mathbf{a})(\mathbf{h}, \mathbf{h}) = \sum_{i,j=1}^{n} D_{ij}\mathbf{f}(\mathbf{a})h_i h_j$$

que a veces se suele denotar $d^2\mathbf{f}(\mathbf{a})\mathbf{h}^2$ es un polinomio homogéneo de segundo grado en las variables $(h_1, h_2, \cdots h_n)$ (cuyos coeficientes son vectores de F). Para una función de dos variables reales se tiene:

$$Q(\mathbf{h}) = D_{11}\mathbf{f}(\mathbf{a})h_1^2 + 2D_{12}\mathbf{f}(\mathbf{a})h_1h_2 + D_{22}\mathbf{f}(\mathbf{a})h_2^2$$

y para una función de tres variables reales

$$D_{11}\mathbf{f}(\mathbf{a})h_1^2 + D_{22}\mathbf{f}(\mathbf{a})h_2^2 + D_{33}\mathbf{f}(\mathbf{a})h_3^2 + 2D_{12}\mathbf{f}(\mathbf{a})h_1h_2 + 2D_{13}\mathbf{f}(\mathbf{a})h_1h_3 + 2D_{23}\mathbf{f}(\mathbf{a})h_2h_3$$

Teorema 6.9 *Sea $\mathbf{f} : \Omega \to F$, definida en un abierto $\Omega \subset \mathbb{R}^n$, y diferenciable dos veces en $\mathbf{a}$. Entonces*

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) = \mathbf{f}(\mathbf{a}) + d\mathbf{f}(\mathbf{a})\mathbf{h} + \frac{1}{2}d^2\mathbf{f}(\mathbf{a})\mathbf{h}^2 + \mathbf{R}(\mathbf{h})$$

donde $\mathbf{R}(\mathbf{h}) = o(\|\mathbf{h}\|^2)$ (es decir, $\mathbf{R}(\mathbf{h}) = \|\mathbf{h}\|^2 \mathbf{r}(\mathbf{h})$, con $\lim_{\mathbf{h} \to 0} \mathbf{r}(\mathbf{h}) = 0$),

DEM: Como $\mathbf{f}$ es diferenciable en una bola $B(\mathbf{a}, r) \subset \Omega$, podemos asegurar que

$$\mathbf{R}(\mathbf{h}) = \mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) - \sum_{i=1}^{n} D_i\mathbf{f}(\mathbf{a})h_i - \frac{1}{2}\sum_{i,j=1}^{n} D_{ij}\mathbf{f}(\mathbf{a})h_ih_j$$

es diferenciable en $B(\mathbf{0}, r)$, donde existirán sus derivadas parciales. Para calcularlas observemos que, en virtud de la simetría $D_{ij}\mathbf{f}(\mathbf{a}) = D_{ji}\mathbf{f}(\mathbf{a})$, las derivadas parciales de $Q(\mathbf{h}) = \sum_{i,j=1}^{n} D_{ij}\mathbf{f}(\mathbf{a})h_ih_j$ vienen dadas por $D_kQ(\mathbf{h}) = 2\sum_{i=1}^{n} D_{ik}\mathbf{f}(\mathbf{a})h_i$, luego

$$D_k\mathbf{R}(\mathbf{h}) = D_k\mathbf{f}(\mathbf{a} + \mathbf{h}) - D_k\mathbf{f}(\mathbf{a}) - \sum_{i=1}^{n} D_{ik}\mathbf{f}(\mathbf{a})h_i$$

Cada $D_k\mathbf{f}$ es diferenciable en $\mathbf{a}$, luego $D_k\mathbf{R}(\mathbf{h}) = o(\|\mathbf{h}\|)$ es decir

$$D_k\mathbf{R}(\mathbf{h}) = \|\mathbf{h}\| \boldsymbol{\rho}_k(\mathbf{h}) \text{ con } \lim_{\mathbf{h} \to 0} \boldsymbol{\rho}_k(\mathbf{h}) = 0$$

Dado $\epsilon > 0$ existe $\delta > 0$ tal que, para todo $k \in \{1, \cdots n\}$ se cumple

$$\|\mathbf{h}\| < \delta \Rightarrow \|\boldsymbol{\rho}_k(\mathbf{h})\| < \epsilon/n$$

$\boldsymbol{\varphi}(t) = \mathbf{R}(t\mathbf{h})$ es derivable en $[0, 1]$ y su derivada $\boldsymbol{\varphi}'(t) = \sum_{j=1}^{n} D_j\mathbf{R}(t\mathbf{h})h_j$ verifica

$$\|\boldsymbol{\varphi}'(t)\| \leq \sum_{j=1}^{n} \left\|\boldsymbol{\rho}_j(t\mathbf{h})\right\| \|t\mathbf{h}\| |h_j| \leq \epsilon \|\mathbf{h}\|^2$$

Aplicando el teorema del incremento finito a la función $\boldsymbol{\varphi}$ se obtiene

$$\|\boldsymbol{\varphi}(1) - \boldsymbol{\varphi}(0)\| \leq \epsilon \|\mathbf{h}\|^2, \quad \text{si} \quad \|\mathbf{h}\| < \delta$$

es decir

$$\|\mathbf{R}(\mathbf{h})\| \leq \epsilon \|\mathbf{h}\|^2, \quad \text{si} \quad \|\mathbf{h}\| < \delta.$$

■

6.2. Extremos relativos

Sea $f : \Omega \to \mathbb{R}$ una función diferenciable de n variables reales definida en un abierto $\Omega \subset \mathbb{R}^n$. Ya hemos visto en el capítulo 5 que si f presenta un extremo relativo en $\mathbf{a} \in \Omega$ entonces $D_j f(\mathbf{a}) = 0$ para todo $j \in \{1, 2, \cdots n\}$, lo que se expresa brevemente diciendo que $\mathbf{a}$ es un punto estacionario de f.

Los siguientes ejemplos muestran que, en lo referente a extremos locales, las funciones de varias variables pueden tener comportamientos que no ocurren en el caso de las funciones de una variable (la comprobación de las afirmaciones que en ellos se hacen se hacen se dejan al cuidado del lector).

Ejemplos 6.10
a) El único punto estacionario de la función $f(x, y) = x^2(1+y)^3+y^2$ es $(0, 0)$ donde f presenta un mínimo relativo que no es mínimo absoluto.
b) La función $g(x, y) = (x^2y-x-1)^2+(x^2-1)^2$ sólo tiene dos puntos estacionarios $(1, 2)$ y $(-1, 0)$, y ambos son puntos de mínimo absoluto.

En esta sección proseguimos con el estudio de los extremos locales de las funciones dos veces diferenciables, obteniendo condiciones necesarias y condiciones suficientes de segundo orden (en términos de las derivadas parciales segundas) para la existencia de máximo relativo, o mínimo relativo, en un punto estacionario.

Proposición 6.11 *Sea $f : \Omega \to \mathbb{R}$, definida en un abierto $\Omega \subset \mathbb{R}^n$, diferenciable dos veces en $\mathbf{a} \in \Omega$. Si f presenta en $\mathbf{a} \in \Omega$ un mínimo relativo se cumple*

$$d^2f(\mathbf{a})\mathbf{h}^2 \geq 0 \quad \textit{para todo} \quad \mathbf{h} \in \mathbb{R}^n$$

Si f presenta en $\mathbf{a} \in \Omega$ un máximo relativo se cumple

$$d^2f(\mathbf{a})\mathbf{h}^2 \leq 0 \quad \textit{para todo} \quad \mathbf{h} \in \mathbb{R}^n$$

DEM: Por hipótesis $\mathbf{a}$ posee un entorno abierto $V_\mathbf{a} \subset \Omega$ donde f es diferenciable. Dado $\mathbf{h} \in \mathbb{R}^n$, sea $r > 0$ tal que $|t| < r \Rightarrow \mathbf{a} + t\mathbf{h} \in V_\mathbf{a}$. En virtud de la regla de la cadena, $\varphi(t) = f(\mathbf{a} + t\mathbf{h})$ está definida y es derivable en $(-r, r)$. Su derivada viene dada por

$$\varphi'(t) = df(\mathbf{a} + t\mathbf{h})\mathbf{h} = \sum_{j=1}^{n} D_j f(\mathbf{a} + t\mathbf{h})h_j$$

Por hipótesis, las derivadas parciales $D_j f : V_\mathbf{a} \to \mathbb{R}$ son diferenciables en $\mathbf{a}$, y aplicando otra vez la regla de la cadena, obtenemos que las funciones $\alpha_j(t) = D_j f(\mathbf{a}+t\mathbf{h})$ son derivables en $t = 0$ y que la derivada vale $\alpha_j'(0) = \sum_{i=1}^{n} D_i D_j f(\mathbf{a})h_i$. Entonces $\varphi'(t)$ es derivable en $t = 0$ y

$$\varphi''(0) = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} D_i D_j f(\mathbf{a})h_i \right) h_j = \sum_{i,j=1}^{n} D_{ij} f(\mathbf{a})h_i h_j = d^2f(\mathbf{a})(\mathbf{h}, \mathbf{h})$$

Si f presenta en $\mathbf{a}$ un mínimo (resp. máximo) relativo entonces φ presenta en $t = 0$ un extremo local del mismo tipo y se concluye que

$$d^2 f(\mathbf{a})(\mathbf{h}, \mathbf{h}) = \varphi''(0) \geq 0, \quad (\text{resp. } \leq 0).$$

■

Corolario 6.12 *Sea $f : \Omega \to \mathbb{R}$ definida en un abierto $\Omega \subset \mathbb{R}^n$ y diferenciable dos veces en $\mathbf{a} \in \Omega$. Si existen $\mathbf{h}, \mathbf{k} \in \mathbb{R}^n$ tales que $d^2 f(\mathbf{a})\mathbf{k}^2 < 0 < d^2 f(\mathbf{a})\mathbf{h}^2$ entonces $\mathbf{f}$ no presenta en $\mathbf{a}$ ni máximo relativo ni mínimo relativo.*

Dem: Es consecuencia directa de la proposición 6.11 ■

El siguiente ejemplo muestra que la condición $d^2 f(\mathbf{a})(\mathbf{h}, \mathbf{h}) \geq 0$ para todo $\mathbf{h} \in \mathbb{R}^n$, no garantiza que f presente en $\mathbf{a}$ un mínimo relativo. (Si $d^2 f(\mathbf{a})(\mathbf{h}, \mathbf{h})$ no es idénticamente nula, esta condición sólo permite asegurar que f no presenta máximo relativo).

Ejemplo 6.13 *En el punto $(0, 0)$, que es estacionario para $f(x, y) = x^2 y + y^2$, se cumple la condición*

$$d^2 f(0, 0)(\mathbf{h}, \mathbf{h}) = 2h_2^2 \geq 0 \quad \textit{para todo} \quad \mathbf{h} \in \mathbb{R}^2$$

y sin embargo f no presenta en este punto un extremo relativo: $f(0, 0) = 0$ y es claro que en todo entorno de $(0, 0)$ hay puntos donde la función $f(x, y) = y(x^2 + y)$ es positiva y puntos donde es negativa.

Una *forma cuadrática* real de n variables reales es una aplicación $Q : \mathbb{R}^n \to \mathbb{R}$ de la forma

$$Q(\mathbf{x}) = Q(x_1, x_2, \cdots x_n) = \sum_{i,j=1}^{n} \alpha_{ij} x_i x_j$$

Si $Q(\mathbf{x}) > 0$ (resp. < 0) para todo $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ se dice que Q es una forma cuadrática *definida positiva* (resp. *negativa*). Si $Q(\mathbf{x}) \geq 0$ (resp. ≤ 0) para todo $\mathbf{x} \in \mathbb{R}^n$ se dice que Q es una forma cuadrática *definida no negativa* (resp. *no positiva*). Finalmente, si existen $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ tales que $Q(\mathbf{x}) < 0 < Q(\mathbf{y})$ se dice que Q es una forma cuadrática *indefinida*.

Lema 6.14 *Si la forma cuadrática $Q : \mathbb{R}^n \to \mathbb{R}$ es definida positiva (resp. negativa) entonces existe $m > 0$ (resp. $m < 0$) tal que $Q(\mathbf{x}) \geq m \|\mathbf{x}\|^2$ (resp. $Q(\mathbf{x}) \leq m \|\mathbf{x}\|^2$) para todo $\mathbf{x} \in \mathbb{R}^n$.*

Dem: Como Q es continua y $S = \{\mathbf{x} \in \mathbb{R}^n : \|x\| = 1\}$ es compacto (cerrado y acotado) existe $\mathbf{a} \in S$ tal que $Q(\mathbf{a}) = \min\{Q(\mathbf{z}) : \mathbf{z} \in S\}$. Si Q es definida positiva podemos asegurar que $Q(\mathbf{a}) > 0$ y tomando $m = Q(\mathbf{a})$ se cumple la condición requerida: Es evidente si $\mathbf{x} = 0$ y cuando $\mathbf{x} \neq 0$ basta considerar $\mathbf{z} = \mathbf{x} / \|\mathbf{x}\| \in S$, y tener en cuenta que Q es homogénea de grado 2, para obtener

$$Q(\mathbf{x}) = Q(\|\mathbf{x}\| \mathbf{z}) = \|\mathbf{x}\|^2 Q(\mathbf{z}) \geq m \|\mathbf{x}\|^2$$

Análogamente se razona, cuando Q es definida negativa, considerando el máximo de Q sobre S, que es negativo. ■

Teorema 6.15 *Sea $f : \Omega \to \mathbb{R}$, definida en un abierto $\Omega \subset \mathbb{R}^n$, y diferenciable dos veces en $\mathbf{a} \in \Omega$. Si $\mathbf{a}$ es punto estacionario de f y la forma cuadrática*

$$d^2 f(\mathbf{a})\mathbf{h}^2 = \sum_{i,j=1}^{n} D_{ij} f(\mathbf{a}) h_i h_j$$

es definida positiva (resp. negativa) entonces $\mathbf{f}$ *presenta en* $\mathbf{a}$ *un mínimo (resp. un máximo) relativo estricto*

DEM: Sea $Q_{\mathbf{a}}(\mathbf{h}) = \frac{1}{2} d^2 f(\mathbf{a})\mathbf{h}^2$. En virtud del teorema 6.9

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + df(\mathbf{a})\mathbf{h} + Q_{\mathbf{a}}(\mathbf{h}) + \mathbf{r}(\mathbf{h}) \|\mathbf{h}\|^2$$

donde $\lim_{\mathbf{h} \to \mathbf{0}} \mathbf{r}(\mathbf{h}) = \mathbf{0}$. Según el lema 6.14 existe $m > 0$ tal que $Q_{\mathbf{a}}(\mathbf{h}) \geq m \|\mathbf{h}\|^2$. y teniendo en cuenta que $df(\mathbf{a}) = 0$ resulta

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) \geq (m + \mathbf{r}(\mathbf{h})) \|\mathbf{h}\|^2$$

Por la definición de límite, existe $B(\mathbf{a}, \delta) \subset \Omega$ tal que $\|\mathbf{h}\| < \delta \Rightarrow |\mathbf{r}(\mathbf{h})| < m/2$, luego

$$\|\mathbf{h}\| < \delta \Rightarrow f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) \geq (m - m/2) \|\mathbf{h}\|^2 > 0$$

Hemos demostrado así que $\mathbf{x} \in B(\mathbf{a}, \delta) \Rightarrow f(\mathbf{x}) > f(\mathbf{a})$, y por lo tanto f presenta en $\mathbf{a}$ un mínimo relativo estricto. Análogamente se razona, en el caso alternativo, para el máximo relativo esticto. ■

Si f es diferenciable dos veces en un punto estacionario $\mathbf{a}$ y la forma cuadrática $d^2 f(\mathbf{a})\mathbf{h}^2 = \sum_{i,j=1}^{n} D_{ij} f(\mathbf{a}) h_i h_j$ es indefinida, según el corolario 6.12, f no presenta ni mínimo ni máximo relativo en este punto. Utilizando el lema 6.14 se consigue una descripción más precisa de lo que ocurre en este caso

Proposición 6.16 *Se supone que $f : \Omega \to \mathbb{R}$, definida en un abierto $\Omega \subset \mathbb{R}^n$, es diferenciable dos veces en $\mathbf{a} \in \Omega$. Si $\mathbf{a}$ es punto estacionario de f y la forma cuadrática*

$$d^2 f(\mathbf{a})\mathbf{h}^2 = \sum_{i,j=1}^{n} D_{ij} f(\mathbf{a}) h_i h_j$$

es indefinida entonces existen dos rectas que pasan por $\mathbf{a}$ de modo que a lo largo de una de ellas f presenta en $\mathbf{a}$ un mínimo relativo estricto y a lo largo de la otra un máximo relativo estricto.

DEM: Razonando como en la demostración del teorema 6.9, existe $B(\mathbf{a}, r) \subset \Omega$ tal que si $\mathbf{x} \in B(\mathbf{a}, r)$ y $\mathbf{h} = \mathbf{x} - \mathbf{a}$ se cumple $f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = Q_{\mathbf{a}}(\mathbf{h}) + \epsilon(\mathbf{h}) \|\mathbf{h}\|^2$, donde $Q_{\mathbf{a}}(\mathbf{h}) = \frac{1}{2}d^2 f(\mathbf{a})\mathbf{h}^2$, y $\lim_{\mathbf{h} \to \mathbf{0}} \epsilon(\mathbf{h}) = \mathbf{0}$. Por hipótesis existen $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ tales que $Q(\mathbf{u}) < 0 < Q(\mathbf{v})$. Cuando $\mathbf{h} = t\mathbf{u}$, la diferencia

$$f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a}) = Q_{\mathbf{a}}(t\mathbf{u}) + \epsilon(t\mathbf{u}) \|t\mathbf{u}\|^2 = t^2(Q_{\mathbf{a}}(\mathbf{u}) + \epsilon(t\mathbf{u}) \|\mathbf{u}\|^2)$$

tiene el mismo signo que $Q_{\mathbf{a}}(\mathbf{u}) + \epsilon(t\mathbf{u}) \|\mathbf{u}\|^2$. Cuando $t \to 0$ esta expresión tiene límite $Q_{\mathbf{a}}(\mathbf{u}) < 0$, luego existe $\rho > 0$ tal que $Q_{\mathbf{a}}(\mathbf{u}) + \epsilon(t\mathbf{u}) \|\mathbf{u}\|^2 < 0$ si $|t| < \rho$. Vemos así que a lo largo de la recta $\mathbf{a} + t\mathbf{u}$ la función f presenta en $\mathbf{a}$ un máximo relativo estricto: $|t| < \rho \Rightarrow f(\mathbf{a} + t\mathbf{u}) - f(\mathbf{a}) < 0$.

Con un razonamiento similar se obtiene que a lo largo de la recta $\mathbf{a} + t\mathbf{v}$ la función f presenta en $\mathbf{a}$ un mínimo relativo estricto. ■

A la vista de los resultados expuestos en los teorema 6.15 y 6.16, es claro el interés que tienen, para el asunto que nos ocupa, los criterios que permiten asegurar que una forma cuadrática concreta es definida positiva, definida negativa o indefinida. Un resultado de esta naturaleza es el siguiente.

Proposición 6.17 *Sea* $Q(\mathbf{x}) = \sum_{i,j=1}^n \alpha_{ij} x_i x_j$ *la forma cuadrática asociada a una matriz simétrica* $\alpha_{ij} = \alpha_{ji}$, *y* Δ_k *el determinante de la matriz* $\{\alpha_{ij} : 1 \leq i, j \leq k\}$.
Q es definida positiva si y sólo si $\Delta_k > 0$ *para todo* $k \in \{1, 2, \cdots, n\}$
Q es definida negativa si y sólo si $(-1)^k \Delta_k > 0$ *para todo* $k \in \{1, 2, \cdots, n\}$.

Dem: De momento damos una demostración elemental para el caso $n = 2$. El lector interesado puede onsultar más adelante, en la sección I.1, una demostración del caso general. En el caso particular que nos ocupa escribimos

$$Q(x, y) = Ax^2 + 2Bxy + Cy^2$$

Si $AC - B^2 > 0$ debe ser $A \neq 0$, y completando cuadrados se obtiene

$$Q(x, y) = \frac{1}{A}[(Ax + By)^2 + (AC - B^2)y^2]$$

Cuando $(x, y) \neq (0, 0)$, teniendo en cuenta que $x \neq 0$ si $y = 0$, se obtiene que Q es definida positiva si $A > 0$ y definida negativa si $A < 0$.

Si $AC - B^2 < 0$, y $A \neq 0$, completando cuadrados se puede expresar $Q(x, y)$ como producto de dos factores lineales distintos:

$$Q(x, y) = \frac{1}{A}[(Ax + By)^2 - M^2 y^2] = \frac{1}{A}[Ax + (B + M)y][Ax + (B - M)y]$$

donde $M = \sqrt{B^2 - AC}$. Si consideremos los semiplanos

$$U^+ = \{(x, y) : Ax + (B + M)y > 0\}, \quad U^- = \{(x, y) : Ax + (B + M)y < 0\}$$

$$V^+ = \{(x, y) : Ax + (B - M)y > 0\}, \quad V^- = \{(x, y) : Ax + (B - M)y < 0\}$$

es claro que sobre las regiones angulares $U^+ \cap V^+$, $U^- \cap V^-$ la forma cuadrática tiene el signo de A, mientras que sobre las regiones angulares $U^+ \cap V^-$, $U^- \cap V^+$ tiene el signo de $-A$, luego se trata de una forma cuadrática indefinida.

Cuando $AC - B^2 < 0$, y $A = 0$, debe ser $B \neq 0$. En este caso $Q(x, y)$ también se expresa como producto de dos factores lineales distintos, $Q(x, y) = y(2Bx + Cy)$ y razonando en forma similar se concluye que la forma cuadrática es indefinida. ■

En el ejercicio 9.17 volveremos obtener, con las interpretaciones geométricas oportunas, los resultados anteriores sobre formas cuadráticas de dos variables. Combinando el teorema 6.15 con la proposición 6.17 se llega al siguiente resultado

Corolario 6.18 *Se supone que $f : \Omega \to \mathbb{R}$ es diferenciable dos veces en $\mathbf{a} \in \Omega \subset \mathbb{R}^2$ y que $D_1 f(\mathbf{a}) = D_2 f(\mathbf{a}) = 0$, $D_{11} f(\mathbf{a}) = A$, $D_{12} f(\mathbf{a}) = B$, $D_{22} f(\mathbf{a}) = C$. Se verifica:*
$AC - B^2 > 0$, $A > 0$ $\Rightarrow$ $\mathbf{a}$ *es punto de mínimo relativo estricto.*
$AC - B^2 > 0$, $A < 0$ $\Rightarrow$ $\mathbf{a}$ *es punto de máximo relativo estricto.*
$AC - B^2 < 0$, $\Rightarrow$ $\mathbf{a}$ *es punto de silla (ni máximo ni mínimo)*
Cuando $AC - B^2 = 0$, la función f puede presentar en $\mathbf{a}$ un máximo relativo, un mínimo relativo o ninguna de las dos cosas (caso dudoso).

Dem: Para completar la demostración sólo queda mostrar ejemplos de lo que puede ocurrir en el caso dudoso $AC - B^2 = 0$:

El origen $(0, 0)$ es punto estacionario de $f(x, y) = x^2 + y^4$, (resp. $-(x^2 + y^4)$) donde se cumple $AC - B^2 = 0$, y es inmediato que f presenta en $(0, 0)$ un mínimo (resp. un máximo) relativo.

El origen $(0, 0)$ también es punto estacionario de $g(x, y) = x^2 y + y^2$. En este caso $AC - B^2 = 0$ y en $(0, 0)$ no hay ni máximo ni mínimo relativo porque $(x^2 + y)y$ cambia de signo en todo entorno de $(0, 0)$. ■

6.3. Funciones convexas

En esta sección se extienden, al contexto de las funciones diferenciables de varias variables, las caracterizaciones habituales de las funciones convexas derivables que se suelen estudiar en el curso de Análisis Matemático I. El lector interesado puede acudir al apéndice F donde se repasan y amplían los resultados básicos referentes a funciones convexas de una variable.

Recordemos que un conjunto $\Omega \subset \mathbb{R}^n$ es convexo cuando, para cada par de puntos $\mathbf{x}, \mathbf{y} \in \Omega$, el segmento $[\mathbf{x}, \mathbf{y}] := \{(1 - t)\mathbf{x} + t\mathbf{y}) : 0 \leq t \leq 1\}$ está contenido en Ω.

Definición 6.19 *Una función $f : \Omega \to \mathbb{R}$, definida en un conjunto convexo $\Omega \subset \mathbb{R}^n$ se dice que es convexa cuando para todo segmento $[\mathbf{x}, \mathbf{y}] \subset \Omega$ y todo $t \in (0, 1)$ se cumple:*

$$f((1 - t)\mathbf{x} + t\mathbf{y}) \leq (1 - t)f(\mathbf{x}) + tf(\mathbf{y})$$

Es inmediato que cada bola (abierta o cerrada) de centro $\mathbf{a} \in \mathbb{R}^n$ y radio $r > 0$, es convexa y que en esa bola la función $f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\|$ es convexa. Combinando este hecho con la siguiente proposición, de carácter elemental, se obtienen buenos ejemplos de funciones convexas:

Proposición 6.20 *Si $f : \Omega \to \mathbb{R}$ es convexa en un conjunto convexo $\Omega \subset \mathbb{R}^n$ y $\varphi : I \to \mathbb{R}$ es creciente y convexa en un intervalo $I \supset f(\Omega)$, entonces la función $\varphi \circ f : \Omega \to \mathbb{R}$ es convexa en Ω. (Brevemente, la composición de una función convexa con una función creciente y convexa es convexa.)*

Dem: Es una comprobación rutinaria que se deja al cuidado del lector ■

Ejemplos 6.21 *Las siguientes funciones son convexas en los conjuntos convexos que se indican:*

a) $f(\mathbf{x}) = \|\mathbf{x}\|^p$, *en* $\mathbb{R}^n$, *con* $p \geq 1$.

b) $f(\mathbf{x}) = (1 + \|\mathbf{x}\|^2)^{\|\mathbf{x}\|^2}$, *en* $\mathbb{R}^n$.

c) $f(\mathbf{x}) = e^{-\sqrt{1-\|x\|^2}}$, *en la bola* $B(\mathbf{0}, 1)$.

Dem: a) Basta aplicar la proposición 6.20 con $f(\mathbf{x}) = \|\mathbf{x}\|$, pues $\varphi(t) = t^p$ es creciente y convexa en $I = [0, +\infty)$, ya que para todo $t \in I$ se cumple $\varphi'(t) = pt^{p-1} \geq 0$ y $\varphi''(t) = p(p-1)t^{(p-1)(p-2)} \geq 0$.
b) $f(\mathbf{x}) = e^{g(\mathbf{x})}$ donde $g(\mathbf{x}) = \|\mathbf{x}\|^2 \log(1 + \|\mathbf{x}\|^2)$. Como la función exponencial es creciente y convexa, según la proposición 6.20, basta ver que $\mathbf{g}$ es convexa en $\mathbb{R}^n$. Esto también se puede justificar con la proposición 6.20: Según a) la función $\|\mathbf{x}\|^2$ es convexa en $\mathbb{R}^n$, y la función $\varphi(t) = t \log(1+t)$ es creciente y convexa en $[0, +\infty)$ pues, para todo $t \geq 0$, se cumple $\varphi'(t) = \log(1+t) + t/(1+t) \geq 0$, $\varphi''(t) = t/(1+t) \geq 0$.
c) Razonando como en b) basta ver que la función $g(\mathbf{x}) = -\sqrt{1 - \|\mathbf{x}\|^2}$ es convexa en la bola abierta $B(\mathbf{0}, 1)$, donde está definida. Esto también se puede justificar con la proposición 6.20 considerando la función $\varphi(t) = -\sqrt{1-t}$ que es creciente y convexa en $[0, +\infty)$, ya que para todo $t \in [0, 1)$ se cumple $\varphi'(t) = \frac{1}{2}(1-t)^{-1/2} \geq 0$, y $\varphi''(t) = \frac{1}{4}(1-t)^{-3/2} \geq 0$. ■

Si $\Omega \subset \mathbb{R}^n$ es convexo, dados $\mathbf{a} \in \Omega$ y $\mathbf{h} \in \mathbb{R}$ el conjunto

$$I_{\mathbf{a},\mathbf{h}} = \{t \in \mathbb{R} : \mathbf{a} + t\mathbf{h} \in \Omega\}$$

es un intervalo porque tiene la propiedad de que dados $x, y \in I_{\mathbf{a},\mathbf{h}}$, cualquier punto intermedio $z = \alpha x + \beta y$, con $\alpha > 0, \beta > 0$, $\alpha + \beta = 1$, está en $I_{\mathbf{a},\mathbf{h}}$: Basta observar que $\mathbf{a} + z\mathbf{h} = \alpha(\mathbf{a} + x\mathbf{h}) + \beta(\mathbf{a} + y\mathbf{h})$ pertenece al conjunto convexo Ω porque $\mathbf{a} + x\mathbf{h}$ y $\mathbf{a} + y\mathbf{h}$ son puntos de Ω. Cuando $\Omega \subset \mathbb{R}^n$ sea un abierto convexo, el intervalo $I_{\mathbf{a},\mathbf{h}}$ será abierto por ser la preimagen de Ω mediante la función continua $t \to \mathbf{a} + t\mathbf{h}$.

Proposición 6.22 *Sea $\Omega \subset \mathbb{R}^n$ un conjunto convexo. Una función $f : \Omega \to \mathbb{R}$, es convexa si y sólo si para cada $\mathbf{a} \in \Omega$ y cada $\mathbf{h} \in \mathbb{R}^n$ la función $\varphi_{\mathbf{a},\mathbf{h}}(t) = f(\mathbf{a} + t\mathbf{h})$, es convexa en el intervalo $I_{\mathbf{a},\mathbf{h}} = \{t \in \mathbb{R} : \mathbf{a} + t\mathbf{h} \in \Omega\}$.*

Dem: Una comprobación rutinaria permite establecer que si f es convexa entonces las funciones $\varphi = \varphi_{\mathbf{a},\mathbf{h}}$ también cumplen la condición de convexidad, es decir

$$\varphi(\alpha x + \beta y) \leq \alpha\varphi(x) + \beta\varphi(y)$$

siempre que $x, y \in I_{\mathbf{a},\mathbf{h}}$, $\alpha > 0$, $\beta > 0$, y $\alpha + \beta = 1$.

Para demostrar el recíproco, consideramos un segmento arbitrario $[\mathbf{x}, \mathbf{y}] \subset \Omega$, al que asociamos la función $\varphi(t) = \varphi_{\mathbf{x},\mathbf{h}}(t) = f(\mathbf{x} + t\mathbf{h})$, con $\mathbf{h} = \mathbf{y} - \mathbf{x}$, que está definida en un intervalo $I \supset [0, 1]$. Esta función es convexa por hipótesis, luego dados $\alpha > 0$ y $\beta > 0$, con $\alpha + \beta = 1$, se cumple:

$$f(\alpha\mathbf{x} + \beta\mathbf{y}) = f(\mathbf{x} + \beta\mathbf{h}) = \varphi(\beta) =$$

$$= \varphi(\alpha 0 + \beta 1) \leq \alpha\varphi(0) + \beta\varphi(1) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$$

y queda demostrado que f cumple la condición de convexidad. ■

Como en el caso de las funciones de una variable se cumple que toda función convexa definida en un abierto convexo $\Omega \subset \mathbb{R}^n$ es continua. El lector interesado en la demostración puede consultar el apéndice F, o [13] (Teor.3.5, pág.110).

Aquí sólo nos ocuparemos de la caracterización de las funciones convexas diferenciables definidas en un abierto convexo $\Omega \subset \mathbb{R}^n$. La razón de considerar dominios abiertos se debe a que para $n > 1$ no hay una alternativa razonable a la noción de derivada lateral que interviene, cuando $n = 1$, en los extremos del intervalo.

Como en el caso de las funciones de una variable la convexidad se puede caracterizar por la condición de que la gráfica quede siempre por encima del plano tangente.

Proposición 6.23 *Si $f : \Omega \to \mathbb{R}$ es diferenciable en un abierto convexo $\Omega \subset \mathbb{R}^n$, son equivalentes*

a) f *es convexa.*

b) $f(\mathbf{y}) \geq f(\mathbf{x}) + df(\mathbf{x})(\mathbf{y} - \mathbf{x})$ *para todo par de puntos* $\mathbf{x}, \mathbf{y} \in \Omega$.

Dem: a) $\Rightarrow$ b): Dados $\mathbf{x}, \mathbf{y} \in \Omega$, sea $\mathbf{h} = \mathbf{y} - \mathbf{x}$. Si f es convexa para cada $t \in [0, 1]$ se verifica:

$$f(\mathbf{x} + t\mathbf{h}) = f((1 - t)\mathbf{x} + t\mathbf{y}) \leq (1 - t)f(\mathbf{x}) + tf(\mathbf{y})$$

luego $f(\mathbf{x} + t\mathbf{h}) - f(\mathbf{x}) \leq t(f(\mathbf{y}) - f(\mathbf{x}))$. Dividiendo por $t > 0$ se conserva la desigualdad y se obtiene

$$\frac{f(\mathbf{x} + t\mathbf{h}) - \mathbf{f}(\mathbf{x})}{t} \leq f(\mathbf{y}) - f(\mathbf{x})$$

Como f es diferenciable en $\mathbf{x}$, existe la derivada $D_{\mathbf{h}}f(\mathbf{x})$, y pasando al límite cuando $t \to 0+$, resulta

$$df(\mathbf{x})\mathbf{h} = D_{\mathbf{h}}f(\mathbf{x}) \leq f(\mathbf{y}) - f(\mathbf{x})$$

es decir $f(\mathbf{y}) \geq f(\mathbf{x}) + df(\mathbf{x})(\mathbf{y} - \mathbf{x})$.
b) $\Rightarrow$ a): Si se cumple b), dados $\mathbf{x}, \mathbf{y} \in \Omega$ y $t \in [0, 1]$, consideremos el punto $\mathbf{x}_t = (1 - t)\mathbf{x} + t\mathbf{y}$. Aplicando la desigualdad b) a los puntos $\mathbf{x}, \mathbf{x}_t$ y a los puntos $\mathbf{y}, \mathbf{x}_t$ se obtiene

i) $f(\mathbf{x}) \geq f(\mathbf{x}_t) + df(\mathbf{x}_t)(\mathbf{x} - \mathbf{x}_t)$

ii) $f(\mathbf{y}) \geq f(\mathbf{x}_t) + df(\mathbf{x}_t)(\mathbf{y} - \mathbf{x}_t)$

Multiplicando la primera desigualdad por $(1-t) \geq 0$, la segunda por $t \geq 0$ sumando miembro a miembro, y teniendo en cuenta que $(1-t)(\mathbf{x}-\mathbf{x}_t)+t(\mathbf{y}-\mathbf{x}_t) = 0$ resulta $(1-t)f(\mathbf{x}) + tf(\mathbf{y}) \geq f(\mathbf{x}_t)$, es decir, se cumple la condición de convexidad. ■

Corolario 6.24 *Sea $f : \Omega \to \mathbb{R}$ una función convexa y diferenciable en el abierto convexo $\Omega \subset \mathbb{R}^n$. Si $\mathbf{a} \in \Omega$ es un punto estacionario de f (e.d. $D_k f(\mathbf{a}) = 0$ para $1 \leq k \leq n$) entonces f presenta en $\mathbf{a}$ un mínimo absoluto.*

Dem: Según 6.23, para todo $\mathbf{x} \in \Omega$ se cumple $f(\mathbf{x}) \geq f(\mathbf{a}) + df(\mathbf{a})(\mathbf{x} - \mathbf{a}) = f(\mathbf{a})$. ■

Proposición 6.25 *Si $f : \Omega \to \mathbb{R}$ es diferenciable dos veces en un abierto convexo $\Omega \subset \mathbb{R}^n$, son equivalentes*

a) f es convexa.

b) $\sum_{i,j=1}^n D_{ij} f(\mathbf{x}) u_i u_j \geq 0$, para todo $\mathbf{x} \in \Omega$ y todo $\mathbf{u} \in \mathbb{R}^n$.

Dem: a) $\Rightarrow$ b): A cada $\mathbf{x} \in \Omega$ le asociamos la forma cuadrática

$$Q_{\mathbf{x}}(\mathbf{h}) = \frac{1}{2} \sum_{i,j=1}^n D_{ij} f(\mathbf{x}) h_i h_j$$

Según 6.9, $f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) - df(\mathbf{x})\mathbf{h} = Q_{\mathbf{x}}(\mathbf{h}) + \|\mathbf{h}\|^2 \epsilon(\mathbf{h})$, donde $\lim_{\mathbf{h} \to \mathbf{0}} \epsilon(\mathbf{h}) = 0$. Dado $\mathbf{u} \in \mathbb{R}^n$ existe $r > 0$ tal que $0 < t < \delta \Rightarrow \mathbf{x} + t\mathbf{u} \in \Omega$, y aplicando 6.23 b) con $\mathbf{y} = \mathbf{x} + t\mathbf{u}$, se obtiene

$$Q_{\mathbf{x}}(t\mathbf{u}) + \|t\mathbf{u}\|^2 \epsilon(t\mathbf{u}) = f(\mathbf{x} + t\mathbf{u}) - f(\mathbf{x}) - df(\mathbf{x})(t\mathbf{u}) \geq 0$$

es decir

$$t^2 Q_{\mathbf{x}}(\mathbf{u}) + t^2 \|\mathbf{u}\|^2 \epsilon(t\mathbf{u}) \geq 0 \quad \text{para todo} \quad 0 < t < \delta.$$

Dividiendo por $t^2 > 0$ se conserva la desigualdad, y pasando al límite cuando $t \to 0+$ resulta $Q_{\mathbf{x}}(\mathbf{u}) \geq 0$, para todo $\mathbf{x} \in \Omega$ y todo $\mathbf{u} \in \mathbb{R}^n$.
b) $\Rightarrow$ a) Recíprocamente, si se cumple b), dados $\mathbf{x} \in \Omega$, $\mathbf{h} \in \mathbb{R}^n$, la función real de variable real $\varphi(t) = \mathbf{f}(\mathbf{x} + t\mathbf{h})$, que está definida en un intervalo abierto $I \subset \mathbb{R}$, es derivable dos veces en cada $t \in I$ y su derivada vale

$$\varphi''(t) = \sum_{i,j=1}^n D_{ij} \mathbf{f}(\mathbf{x} + t\mathbf{h}) h_1 h_j \geq 0$$

Un resultado bien conocido de la teoría de funciones reales de una variable (véase F.5) permite asegurar que φ es convexa sobre el intervalo I, y con la proposición 6.22 se concluye que f es convexa. ■

6.4. Ejercicios resueltos

Ejercicio 6.26 *Sea* $\Omega = \mathbb{R}^n \setminus \{\mathbf{0}\}$. *Obtenga las funciones* $f : \Omega \to \mathbb{R}$ *que son de la forma* $f(\mathbf{x}) = g(\|\mathbf{x}\|)$ *con* $g : (0, +\infty) \to \mathbb{R}$ *de clase* C^2, *y que satisfacen la ecuación de Laplace* $\sum_{j=1}^n D_{jj} f = 0$. *(*$\|\mathbf{x}\|$ *es la norma euclídea).*

SOLUCIÓN

$$\frac{\partial f(\mathbf{x})}{\partial x_j} = g'(\|\mathbf{x}\|)\frac{\partial \|\mathbf{x}\|}{\partial x_j} = g'(\|\mathbf{x}\|)\frac{x_j}{\|\mathbf{x}\|}$$

Derivando otra vez respecto a x_j:

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_j^2} = g''(\|\mathbf{x}\|)\frac{x_j^2}{\|\mathbf{x}\|^2} + \left(1 - \frac{x_j^2}{\|\mathbf{x}\|^2}\right)\frac{g'(\|\mathbf{x}\|)}{\|\mathbf{x}\|}$$

Sumando para $j = 1, 2 \cdots n$ resulta:

$$\Delta f(\mathbf{x}) = g''(\|\mathbf{x}\|) + (n-1)\frac{g'(\|\mathbf{x}\|)}{\|\mathbf{x}\|}$$

Si $r = \|\mathbf{x}\|$ la ecuación $\Delta f(\mathbf{x}) = 0$ se escribe en la forma

$$g''(r) + (n-1)\frac{g'(r)}{r} = 0$$

La derivada de $g'(r)e^{(n-1)\log r}$ es idénticamente nula luego existe una constante K tal que $g'(r) = Ke^{(1-n)logr} = Kr^{1-n}$, y se concluye que

$$g(r) = A + \frac{K}{(2-n)r^{n-2}} = A + \frac{B}{r^{n-2}} \quad \text{si} \quad n \geq 3.$$

$$g(r) = A + K\log r \quad \text{si} \quad n = 2.$$

■

Ejercicio 6.27 *Sea* $F = \varphi \circ f : \Omega \to \mathbb{R}$ *definida en un abierto* $\Omega \subset \mathbb{R}^n$, *donde* $f : \Omega \to \mathbb{R}$ *es diferenciable dos veces en* $\mathbf{a} \in \Omega$ *y* $\varphi : V \to \mathbb{R}$, *definida en un abierto* $V \subset \mathbb{R}$, *con* $V \supset f(\Omega)$, *es derivable dos veces en* $b = f(\mathbf{a})$.

Demuestre que es F *diferenciable dos veces en* $\mathbf{a}$ *y obtenga, para cada* $\mathbf{h} \in \mathbb{R}$, *el valor de* $d^2F(\mathbf{a})\mathbf{h}^2$ *en términos de las funciones* f *y* φ.

Si $\varphi'(b) > 0$ *y* $\varphi''(b) \geq 0$, *obtenga que* $d^2 f(\mathbf{a})\mathbf{h}^2 > 0 \Rightarrow d^2F(\mathbf{a})\mathbf{h}^2 > 0$.

SOLUCIÓN

La existencia de la derivada segunda $\varphi''(b)$ implica que hay un entorno de b, $V_b \subset V$ donde φ es derivable. Como f es continua en $\mathbf{a}$, hay un entorno abierto de $\mathbf{a}$, $U_\mathbf{a} \subset \Omega$, tal que $f(U_\mathbf{a}) \subset V_b$. De acuerdo con la definición de función dos veces diferenciable en $\mathbf{a}$ podemos suponer también que f es diferenciable en $U_{\mathbf{a}}$, donde estarán definidas las derivadas parciales $D_j f$, $1 \leq j \leq n$ que, por hipótesis, son diferenciables en $\mathbf{a}$.

Según la regla de la cadena la composición $F = \varphi \circ f$ es diferenciable en $U_{\mathbf{a}}$ donde sus derivadas parciales vienen dadas por

$$D_j F(\mathbf{x}) = \varphi'(f(\mathbf{x})) D_j f(\mathbf{x}),\ 1 \leq j \leq n$$

Como las dos funciones $\varphi'(f(\mathbf{x}))$, $D_j f(\mathbf{x})$ son diferenciables en $\mathbf{x} = \mathbf{a}$, también lo es su producto y por lo tanto cada derivada parcial $D_j F$ es diferenciable en $\mathbf{a}$. Queda demostrado así que F es diferenciable dos veces en $\mathbf{a}$. Las derivadas parciales $D_{ij}F(\mathbf{a})$ se pueden calcular con la regla para derivar un producto:

$$D_{ij}F(\mathbf{a}) = D_i D_j F(\mathbf{a}) = \varphi''(f(\mathbf{a}))) D_i f(\mathbf{a}) D_j f(\mathbf{a}) + \varphi'(f(\mathbf{a})) D_i D_j \mathbf{f}(\mathbf{a})$$

Entonces, para cada $\mathbf{h} \in \mathbb{R}^n$ se verifica

$$\sum_{i,j=1}^{n} D_{ij}F(\mathbf{a}) h_i h_j = \varphi''(b) \sum_{i,j=1}^{n} D_i f(\mathbf{a}) D_j f(\mathbf{a}) h_i h_j + \varphi'(b) \sum_{i,j=1}^{n} D_{ij} f(\mathbf{a}) h_i h_j$$

es decir

$$\sum_{i,j=1}^{n} D_{ij}F(\mathbf{a}) h_i h_j = \varphi''(b) \left[\sum_{i=1}^{n} D_i f(\mathbf{a}) h_i \right]^2 + \varphi'(b) \sum_{i,j=1}^{n} D_{ij} f(\mathbf{a}) h_i h_j$$

Si $\varphi'(b) > 0$ y $\varphi''(b) \geq 0$, manejando la última igualdad es inmediato que

$$d^2 f(\mathbf{a})\mathbf{h}^2 = \sum_{i,j=1}^{n} D_{ij} f(\mathbf{a}) h_i h_j > 0 \quad \Rightarrow \quad d^2 F(\mathbf{a})\mathbf{h}^2 = \sum_{i,j=1}^{n} D_{ij} F(\mathbf{a}) h_i h_j > 0$$

■

Ejercicio 6.28 *Determine los valores del parámetro $a \in \mathbb{R}$ para los que la función*

$$f_a(x, y) = a(2xy + y^2 + yx^2 + \cos(x + y)) + x^2(a^2 - y)$$

presenta un extremo relativo en el punto $(0, 0)$.

solución

Se comprueba fácilmente que para cualquier valor de $a \in \mathbb{R}$ el punto $(0, 0)$ es estacionario: $D_1 f(0, 0) = D_2 f(0, 0) = 0$. Un cálculo rutinario permite obtener la matriz Hessiana

$$\begin{pmatrix} D_{11}f(0,0) & D_{12}f(0,0) \\ D_{21}f(0,0) & D_{22}f(0,0) \end{pmatrix} = \begin{pmatrix} 2a^2 - a & a \\ a & a \end{pmatrix}$$

cuyo determinante vale $\Delta(a) = 2a^2(a - 1)$.

Si $a > 1$ se cumple $\Delta(a) > 0$ y $2a^2 - a > 0$ luego, en virtud de 6.18, f presenta en $(0, 0)$ un mínimo relativo.

Si $a < 1$ y $a \neq 0$ se cumple $\Delta(a) < 0$ luego, en virtud de 6.18, f no presenta en $(0, 0)$ ni máximo ni mínimo relativo (punto de silla).

Finalmente, cuando $a = 1$ ó $a = 0$ se cumple $\Delta(a) = 0$ por lo que no es aplicable el criterio 6.18 (caso dudoso) y debemos estudiar directamente la función.

Cuando $a = 1$ la función toma la forma

$$f(x, y) = (x + y)^2 + \cos(x + y) = \varphi(x + y)$$

donde $\varphi(t) = t^2 + \cos t$ presenta un mínimo relativo en $t = 0$, luego f también presenta un mínimo relativo en $(0, 0)$.

Cuando $a = 0$ la función se reduce a $f(x, y) = -x^2 y$ y es obvio que en $(0, 0)$ no hay ni máximo ni mínimo relativo.

Resumiendo: f presenta extremo relativo en $(0, 0)$ si y sólo si $a \geq 1$, y el extremo siempre es un mínimo relativo. ■

Ejercicio 6.29 *Estudie la existencia de extremos relativos de la función de tres variables* $f(x, y, z) = xy + yz + zx$.

SOLUCIÓN

Los puntos estacionarios, donde f puede presentar extremo local, son las soluciones del sistema $D_k f(x, y, z) = 0$, $k = 1, 2, 3$, es decir: $y + z = 0$; $x + z = 0$; $y + x = 0$.

La única solución de este sistema lineal es $x = 0$, $y = 0$, $z = 0$. En este punto las derivadas parciales segundas valen

$$D_{ij} f(0, 0, 0) = 1 \text{ si } i \neq j, \quad D_{ii} f(0, 0, 0) = 0$$

El criterio de la proposición 6.17 no sirve aquí para decidir si f presenta en $(0, 0, 0)$ un extremo local, por lo que debemos estudiar directamente la función. Para simplificar empezamos restringiendo f al plano $y = z$ (que contiene al punto $(0, 0, 0)$). Obtenemos así una función de dos variables

$$h(x, y) = f(x, y, y) = y(y + 2x)$$

en la que es fácil apreciar cambio de signo en todo entorno de $(0, 0)$: Para todo $t > 0$, se cumple $h(t, t) = f(t, t, t) = 3t^2 > 0$ y $h(-t, t) = f(-t, t, t) = -2t^2 < 0$, luego f no presenta un extremo local en $(0, 0, 0)$, pues $f(0, 0, 0) = 0$.

Ejercicio 6.30 *Encuentre los extremos relativos y absolutos de la función de tres variables* $g(x, y, z) = (x + z^2)e^{x(1+y^2+z^2)}$.

SOLUCIÓN

Empecemos calculando los puntos estacionarios, donde g puede presentar extremos locales. Para ello debemos resolver el sistema de las tres ecuaciones $D_k f(x, y, z) = 0$, $k = 1, 2, 3$, es decir:

i) $[(1 + y^2 + z^2)(x + z^2) + 1]e^{x(1+y^2+z^2)} = 0$;

ii) $2xy(x+z^2)e^{x(1+y^2+z^2)} = 0$;

iii) $[2z + 2xz(x+z^2)]e^{x(1+y^2+z^2)} = 0$;

El sistema equivale al siguiente

a) $(1+y^2+z^2)(x+z^2)+1 = 0$;

b) $xy(x+z^2) = 0$;

c) $z[1+x(x+z^2)] = 0$;

La ecuación b) conduce a tres alternativas

$$x = 0; \qquad x+z^2 = 0; \qquad y = 0$$

La ecuación a) permite descartar la dos primeras. Sustituyendo la tercera alternativa $y = 0$ en la otras dos ecuaciones obtenemos

a) $(1+z^2)(x+z^2)+1 = 0$;

c') $z(1+x(x+z^2)) = 0$

Ahora la ecuación c') conduce a dos posibilidades $z = 0$ ó $x(x+z^2) = -1$.

Con $z = 0$ obtenemos la solución $x = -1$, $y = 0$, $z = 0$.

La ecuación $x(x+z^2) = -1$ no proporciona soluciones: Si multiplicamos a) por x resulta $x + x(x+z^2)(1+z^2) = 0$, luego $x - (1+z^2) = 0$, y sustituyendo $x = 1+z^2$ en a) se obtiene una ecuación sin soluciones: $1 + (1+z^2)[1+z^2+z^2] = 0$.

Hemos obtenido así que $\mathbf{p} = (-1,0,0)$ es el un único punto estacionario de g. Un cálculo rutinario proporciona la matriz Hessiana

$$\begin{pmatrix} D_{11}g(\mathbf{p}) & D_{12}g(\mathbf{p}) & D_{13}g(\mathbf{p}) \\ D_{21}g(\mathbf{p}) & D_{22}g(\mathbf{p}) & D_{23}g(\mathbf{p}) \\ D_{31}g(\mathbf{p}) & D_{32}g(\mathbf{p}) & D_{33}g(\mathbf{p}) \end{pmatrix} = \begin{pmatrix} 1/e & 0 & 0 \\ 0 & 2/e & 0 \\ 0 & 0 & 4/e \end{pmatrix}$$

Como los determinantes de los sucesivos menores cumplen la condición

$$\Delta_1 > 0; \quad \Delta_2 > 0; \quad \Delta_3 > 0;$$

en virtud de 6.15 y 6.17 g presenta un mínimo relativo en $\mathbf{p} = (-1,0,0)$.

Para averiguar si este mínimo relativo $g(-1,0,0) = -1/e$ es un mínimo absoluto efectuamos un estudio directo sobre la función. Para simplificar la cuestión podemos restringir el estudio al abierto donde g es negativa $A := \{(x,y,z) : x+z^2 < 0\}$. En un punto de este abierto siempre es $x < 0$, luego $e^x \geq e^{x(1+y^2+z^2)}$. Multiplicando por x (< 0) resulta

$$xe^x \leq xe^{x(1+y^2+z^2)} \leq (x+z^2)e^{x(1+y^2+z^2)}$$

es decir, $g(x,0,0) \leq g(x,y,z)$ para todo $(x,y,z) \in A$.

Con los métodos del cálculo de funciones de una variable se obtiene fácilmente que la función $\varphi(x) = g(0,0,x) = xe^x$ alcanza un mínimo absoluto para $x = -1$ que vale $\varphi(-1) = -1/e$, es decir,

$$-1/e = g(-1,0,0) \leq g(x,y,z) \text{ para todo } (x,y,z) \in A$$

Como en $\mathbb{R}^3 \setminus A$ los valores de g son positivos, se sigue que $g(-1,0,0) = -1/e$ es el mínimo absoluto de g en todo $\mathbb{R}^3$. ■ .

Ejercicio 6.31 *Demuestre que la forma cuadrática asociada a una matriz simétrica:*

$$Q(\mathbf{x}) = \sum_{i,j=1}^{n} \alpha_{ij} x_i x_j \text{ con } \alpha_{ij} = \alpha_{ji}, \quad 1 \leq i,j \leq n$$

es convexa si y sólo si es definida no negativa e.d. $Q(\mathbf{x}) \geq 0$ *para todo* $\mathbf{x} \in \mathbb{R}^n$.

SOLUCIÓN

Es consecuencia inmediata de 6.25 ya que para todo $\mathbf{x} \in \mathbb{R}$ se verifica $D_{ij}Q(\mathbf{x}) = 2\alpha_{ij}$, luego $\sum_{i,j=1}^{n} D_{ij}Q(\mathbf{x})u_i u_j = 2Q(\mathbf{u})$. ■

Ejercicio 6.32 *Sea* $\alpha < 0 < \beta$ *y* $g : (\alpha, \beta) \to \mathbb{R}$ *una función de clase* C^2 *tal que* $g''(t) \leq 0$ *para todo* $t \in [0, \beta)$. *Se considera la función* $f(x,y) = g(x^2 + y^2)$, *que está definida en* $\{(x,y) : x^2 + y^2 < \beta\}$.

Si $0 < R^2 < \beta$, *demuestre que* f *es convexa sobre el disco* $x^2 + y^2 < R^2$ *si y sólo si* $g'(t) + 2tg''(t) \geq 0$ *para todo* $t \in [0, R^2]$.

Para cada una de las funciones $\log(1 + x^2 + y^2)$, $\operatorname{sen}(x^2 + y^2)$ *obtenga el mayor disco* $D_R = \{(x,y) : x^2 + y^2 < R^2\}$ *sobre el que la función es convexa.*

DEM: Como f es de clase C^2 para estudiar su convexidad podemos utilizar la proposición 6.25 donde interviene la forma cuadrática

$$Q_{\mathbf{p}}(\mathbf{h}) = D_{11}f(\mathbf{p})h_1^2 + 2D_{12}f(\mathbf{p})h_1h_2 + D_{22}f(\mathbf{p})h_2^2$$

asociada a un punto genérico $\mathbf{p}$ de su dominio. En lo que sigue conviene expresar $\mathbf{p} = (a,b)$ en coordenadas polares, $a = r\cos\theta$, $b = r\operatorname{sen}\theta$.

Calculando previamente las derivadas parciales

$$D_{11}f(\mathbf{p}) = 2g'(r^2) + 4a^2g''(r^2); \quad D_{12}f(\mathbf{p}) = 4abg''(r^2); \quad D_{22}f(\mathbf{p}) = 2g'(r^2) + 4b^2g''(r^2).$$

resulta

$$\begin{aligned} Q_{\mathbf{p}}(\mathbf{h}) &= 2g'(r^2)(h_1^2 + h_2^2) + 4g''(r^2)(ah_1 + bh_2)^2 = \\ &= 2g'(r^2)(h_1^2 + h_2^2) + 4r^2g''(r^2)(h_1\cos\theta + h_2\operatorname{sen}\theta)^2 \end{aligned}$$

Si suponemos que para todo $t \in [0, R^2]$ se verifica $g'(t) + 2tg''(t) \geq 0$, entonces $2g'(r^2) \geq -4r^2g''(r^2)$ para cada $r \in [0, R]$, luego para cada $\mathbf{p} \in D_R$ y cada $\mathbf{h} \in \mathbb{R}^2$ se cumple la desigualdad

$$Q_{\mathbf{p}}(\mathbf{h}) \geq -4r^2g''(r^2)(h_1^2 + h_2^2) + 4r^2g''(r^2)(h_1\cos\theta + h_2\operatorname{sen}\theta)^2 =$$

$$= -4r^2 g''(r^2)(h_1 \operatorname{sen}\theta + h_2 \cos\theta)^2$$

Por hipótesis $g''(r^2) \leq 0$, luego $Q_{\mathbf{p}}(\mathbf{h}) \geq 0$ para cada $\mathbf{p} \in D_R$ y cada $\mathbf{h} \in \mathbb{R}^2$, y con la proposición 6.25 concluimos que f es convexa en D_R

Recíprocamente, si f es convexa en D_R, según la proposición 6.25, podemos asegurar que para todo $r \in [0, R]$ y todo $\mathbf{h} \in \mathbb{R}^2$ se cumple

$$Q_{\mathbf{p}}(\mathbf{h}) = 2g'(r^2)(h_1^2 + h_2^2) + 4g''(r^2)(ah_1 + bh_2)^2 \geq 0$$

Aplicando esta desigualdad cuando $a = r < R$, $b = 0$, $h_1 > 0$, y $h_2 = 0$ se obtiene que $g'(r^2)h_1^2 + 2g''(r^2)r^2h_1^2 \geq 0$, luego $g'(r^2) + 2g''(r^2)r^2 \geq 0$ para todo $r \in [0, R]$.

De acuerdo con lo que se acaba de establecer, para determinar el mayor disco $D_R = \{(x, y) : x^2 + y^2 < R^2\}$ donde $\log(1 + x^2 + y^2)$ es convexa consideramos la función $g(t) = \log(1 + t)$, sobre el intervalo $(-1, +\infty)$, donde cumple $g''(t) = -1/(1 + t)^2 \leq 0$. En este caso la desigualdad $g'(t) + 2tg''(t) = (1 - t)/(1 + t)^2 \geq 0$ se satisface si y sólo si $t \leq 1$, luego la función $\log(1 + x^2 + y^2)$ es convexa en el disco D_1, y este es el mayor disco centrado en $(0, 0)$ donde es convexa.

Análogamente, para obtener el mayor disco D_R sobre el que $\operatorname{sen}(x^2 + y^2)$ es convexa consideramos la función $g(t) = \operatorname{sen} t$ en el intervalo $(-\pi/2, \pi/2)$, donde cumple que $g''(t) = -\operatorname{sen} t \leq 0$. Ahora $g'(t) + 2tg''(t) = \cos t - 2t \operatorname{sen} t$, y el mayor intervalo $[0, R^2]$ donde se cumple la desigualdad $\cos t - 2t \operatorname{sen} t \geq 0$ se consigue cuando R^2 es el único cero que tiene en $(0, \pi/2)$ la función $\varphi(t) = \cos t - 2t \operatorname{sen} t$ (obsérvese que φ es decreciente en $[0, \pi/2]$, y que $\varphi(0) = 1$, $\varphi(\pi/2) = -\pi$). ■

6.5. Ejercicios propuestos

♢ **6.5.1** *Sea* $g : \mathbb{R}^2 \to \mathbb{R}$ *definida por*

$$g(x,y) = \frac{x \operatorname{sen} y - y \operatorname{sen} x}{x^2+y^2} \quad \textit{si} \quad (x,y) \neq (0,0); \quad g(0,0) = 0$$

Compruebe que g *es diferenciable en* $(0,0)$*, y que existen y son distintas las derivadas parciales* $D_{12}g(0,0) \neq D_{21}g(0,0)$.

♢ **6.5.2** *Sea* $f(x,y) = xyg(x,y)$ *donde* $g : \mathbb{R}^2 \to \mathbb{R}$ *es una función acotada diferenciable en* $\mathbb{R}^2 \setminus \{(0,0)\}$*. Justifique las afirmaciones siguientes:*

i) Si $\lim_{x \to 0} g(x,0) \neq \lim_{y \to 0} g(0,y)$ *entonces* f *es diferenciable en* $(0,0)$ *pero no es diferenciable dos veces en* $(0,0)$.

ii) Si g *es continua en* $(0,0)$ *y sus derivadas parciales están acotadas entonces* f *es diferenciable dos veces en* $(0,0)$.

♢ **6.5.3** *En cada caso estudie si* $f : \mathbb{R}^2 \to \mathbb{R}$ *es dos veces diferenciabe en* $(0,0)$*:*

i) $f(x,y) = y^2 \operatorname{sen}(x/y^2)$ *si* $y \neq 0$, $\quad f(x,0) = 0$.

ii) $f(x,y) = \dfrac{x}{\sqrt{x^2+y^2}} \operatorname{sen}(x^2+y^2)$ *si* $(x,y) \neq 0$, $f(0,0) = 0$

iii) $f(x,y) = \dfrac{\operatorname{sen} x^2 \operatorname{sen} y^2}{x^2+y^2}$ *si* $(x,y) \neq (0,0)$ $\quad f(0,0) = 0$.

♢ **6.5.4** *Demuestre que la función*

$$f(x,y) = y^4/(x^2+y^2) \quad \textit{si} \quad (x,y) \neq (0,0); \quad f(0,0) = 0$$

es de clase C^1 *en* $\mathbb{R}^2$*. Estudie si* f *es diferenciable dos veces en* $(0,0)$.

♢ **6.5.5** *Sea* $f_\alpha : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *definida por*

$$f_\alpha(x,y) = \frac{x^\alpha}{x^2+y^2} \textit{ si } (x,y) \neq (0,0) \textit{ y } f_\alpha(0,0) = 0.$$

Estudie los valores $\alpha \in \mathbb{R}$, $\alpha > 2$*, para los que* f_α *es dos veces diferenciable en* $(0,0)$.

♢ **6.5.6** *Estudie los puntos* $(a,b) \in \mathbb{R}^2$ *donde es dos veces diferenciable la función* $f : \mathbb{R}^2 \to \mathbb{R}$ *definida por* $f(x,y) = yx^2 \operatorname{sen}(1/x)$ *si* $x \neq 0$; $\quad f(0,y) = 0$.

♢ **6.5.7** *Sean* $u, v : \mathbb{R}^2 \to \mathbb{R}$ *de clase* C^2 *verificando*

$$u_x(x,y) = v_y(x,y), \quad u_y(x,y) = -v_x(x,y) \textit{ para todo } (x,y) \in \mathbb{R}^2.$$

Si $g : \mathbb{R}^2 \to \mathbb{R}$ *es de clase* C^2 *y satisface la ecuación de Laplace* $D_{11}g + D_{22}g = 0$ *compruebe que* $G(x,y) = g(u(x,y), v(x,y))$ *también la satisface.*

♢ **6.5.8** *Sea $f : A \to \mathbb{R}$ de clase $C^2(A)$ en un abierto conexo $A \subset \mathbb{R}^n$, tal que $\nabla f(\mathbf{x}) \neq 0$ y $\sum_{i=1}^n D_{ii} f(\mathbf{x}) = 0$ para todo $\mathbf{x} \in A$. Demuestre que $B = f(A)$ es un intervalo abierto y obtenga la forma general de las funciones de clase $C^2(B)$, $g : B \to \mathbb{R}$ tales que $F = g \circ f$ verifica $\sum_{i=1}^n D_{ii} F(\mathbf{x}) = 0$.*

♢ **6.5.9** *Compruebe que la función $f(x, y, z) = \displaystyle\int_x^{y^2} \frac{\operatorname{sen} zt}{t} dt$ está bien definida y es de clase C^2 en todo $\mathbb{R}^3$. Calcule el máximo valor de la derivada direccional $D_{\mathbf{u}} f(1, 1, \pi/2)$, $\|\mathbf{u}\|_2 = 1$.*

♢ **6.5.10** *Justifique que la integral $F(x, y, z) = \displaystyle\int_y^z \frac{\operatorname{sen} tx}{t} dt$ define en $\mathbb{R}^3$ una función de clase C^2 que verifica $zF_z(x, y, z) = xF_x(x, y, z) + yF_y(x, y, z)$.*

♢ **6.5.11** *Sean A, $\Omega \subset \mathbb{R}^n$ abiertos y $\mathbf{g} : A \to \Omega$ un homeomorfismo. Dada $f : \Omega \to \mathbb{R}$ compruebe que $F = f \circ \mathbf{g} : A \to \mathbb{R}$ tiene un extremo relativo en $\mathbf{a} \in A$ si y sólo sí f tiene un extremo relativo (del mismo tipo) en $\mathbf{b} = \mathbf{g}(\mathbf{a})$. Aplique lo anterior para determinar los extremos relativos de la función*

$$f(x, y) = \left(\operatorname{arc\,tg} \frac{y}{x} - 1\right)^2 + \sqrt{(x^2 + y^2)}(x^2 + y^2 - 3)$$

definida en $\Omega = \{(x, y) \in \mathbb{R}^2 : x > 0,\ y > 0\}$.

♢ **6.5.12** *Sea $f : \Omega \to \mathbb{R}$ definida en un abierto $\Omega \subset \mathbb{R}^n$ y $\varphi : I \to \mathbb{R}$ una función creciente definida en un intervalo $I \supset f(\Omega)$. Compruebe que f y $F = \varphi \circ f$ tienen los mismos extremos relativos. ¿Qué ocurre cuando φ es decreciente?. Aplique lo anterior para determinar los extremos relativos de las funciones*

a) $F(x, y) = e^{x((\log x)^2 + y^2)}$, *definida en* $\{(x, y) : x > 0\}$;

b) $F(x, y) = \dfrac{x^3 y^3}{(x - a)(y - b)}$, *definida en* $\mathbb{R}^2 \setminus \{(a, b)\}$;

♢ **6.5.13** *Obtenga los extremos relativos de los polinomios*

$$x^3 + 3xy^2 - 15x - 12y; \quad 3x^4 - 4x^2 y + y^2;$$

$$(3 - x)(3 - y)(x + y - 3) \quad x^2 - y^2 + x^3 + x^2 y + y^3/3$$

Estudie en cada caso la existencia de extremos absolutos.

♢ **6.5.14** *Obtenga los extremos relativos de las funciones de dos variables*

$$(ax^2 + by^2)e^{-(x^2+y^2)}; \quad \frac{x - y}{x^2 + y^2 + 1}; \quad 1 - (x^2 + y^2)^{2/3}$$

Estudie en cada caso la existencia de extremos absolutos.

♢ **6.5.15** *Obtenga los extremos relativos de la función de dos variables*

$$f(x,y) = \log(1+x^2+y^2) - \int_0^x \frac{2t}{1+t^4}dt$$

♢ **6.5.16** *Demuestre que hay un entorno del punto* $(0,0,1)$ *en el que la gráfica de la función* $z = y^2 + y\cos x + e^{2x}\cos y$ *queda por encima de su plano tangente en este punto.*

♢ **6.5.17** *Estudie (según los valores de* $a, b \in \mathbb{R}$*) cuando la superficie*

$$z = e^{ax+y^2} + b\cos(x^2+y^2)$$

queda por encima o por debajo de su plano tangente en un entorno de $(0,0)$ *y cuando no ocurre ni una cosa ni la otra.*

♢ **6.5.18** *Sea* $\Omega \subset \mathbb{R}^n$ *un abierto convexo. Se dice que* $f : \Omega \to \mathbb{R}$ *es estrictamente convexa si*

$$f(t\mathbf{x} + (1-t)\mathbf{y}) < tf(\mathbf{x}) + (1-t)f(\mathbf{y})$$

para todo par de puntos distintos $\mathbf{x}, \mathbf{y} \in \Omega$ *y todo* $t \in (0,1)$*. Demuestre:*

i) Una función diferenciable $f : \Omega \to \mathbb{R}$ *es estrictamente convexa si y sólo si*

$$f(\mathbf{a}) + df(\mathbf{a})(\mathbf{x}-\mathbf{a}) < f(\mathbf{x}) \quad \text{para todo par de puntos distintos } \mathbf{a}, \mathbf{x} \in \Omega$$

ii) Una función diferenciable estrictamente convexa presenta a lo más un único extremo relativo que necesariamente es un mínimo absoluto.

iii) Una función de clase C^2, $f : \Omega \to \mathbb{R}$ *que verifica* $d^2f(\mathbf{x})\mathbf{h}^2 > 0$ *para todo* $\mathbf{h} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ *es estrictamente convexa.*

♢ **6.5.19** *Compruebe que la función* $x^3 + y^3 + 6xy$ *es convexa en el abierto* $A = \{(x,y) : xy > 1,\ x > 0\}$.

♢ **6.5.20** *Compruebe que la función* $f(x,y) = x^2 + y(y^3 - 4)$ *es convexa en* $\mathbb{R}^2$ *y estudie la existencia de mínimo absoluto global.*

♢ **6.5.21** *Obtenga una bola centrada en* $(0,0,0)$ *sobre la que sea convexa la función* $\log(1+x^2+y^2+z^2)$.

♢ **6.5.22** *Sea* $\varphi : \mathbb{R} \to \mathbb{R}$ *una función derivable con* $\varphi'(0) > 0$*. Demuestre que* $f(x,y) = \varphi(x^2+y^2)$ *presenta un mínimo relativo en* $(0,0)$*. Indique una clase de funciones* φ *para las que se pueda asegurar que* f *es convexa. en* $\mathbb{R}^2$.

♢ **6.5.23** *a) Demuestre que las siguientes funciones son convexas en* $\Omega = \{(x,y,z) : x > 0, y > 0, z > 0\}$.

$$f(x,y,z) = -\log(xyz^3); \quad g(x,y,z) = 1/(xyz^3)$$

b) Demuestre que la función $\log(xyz^3)$ *alcanza un máximo relativo sobre el trozo de esfera* $M = \{(x,y,z) \in \Omega : x^2+y^2+z^2 = 5r^2\}$*. Justifique que este máximo es absoluto (para ello se recomienda considerar la función* xyz^3*).*