Notas de Cálculo

DERIVABILIDAD Y DIFERENCIABILIDAD
DE FUNCIONES DE VARIAS VARIABLES
(Bloque 1, Temas 2 y 3)

Grado en Ingeniería de Organización Industrial
UPCT−CUD San Javier
Curso 2020-21 (v. 1.3)

Índice

Introducción

Estas notas para el seguimiento de las clases de Cálculo desarrollan los siguientes temas del primer bloque de la asignatura:

- Tema 2: Derivabilidad y diferenciabilidad de funciones escalares.
- Tema 3: Derivabilidad y diferenciabilidad de funciones vectoriales.

Las unidades didácticas se componen del resumen de la teoría matemática (definiciones, proposiciones y notas prácticas), ejemplos resueltos (sobre los conceptos teóricos y los métodos de resolución de problemas) y ejercicios propuestos (que incluyen la solución para permitir la autocomprobación en el trabajo individual del alumno).

Dado que muchos aspectos del estudio de la derivabilidad y diferenciablidad de las funciones escalares y vectoriales son comunes, la división de unidades didácticas en estas notas de clase no sigue la estructura de temas anteriores, sino de contenido, distinguiendo las generalidades de la derivación y diferenciación de funciones y las aplicaciones del cálculo diferencial.

Algunas lecturas recomendadas para esta parte del temario, incluidas en la bibliografía de la asignatura, son:

- Marsden, J. E. y Tromba, A. J. 2018. Cálculo vectorial, 6ª ed. Pearson.
- García, A. López, A., Rodríguez, G., Romero, S., de la Villa, A. 2002. Teoría y problemas de funciones de varias variables. Clagsa.
- Burgos, J. 2007. Cálculo infinitesimal de varias variables. McGraw-Hill.

T. Baenas
tomas.baenas@cud.upct.es
Febrero, 2021

Parte 1

Derivación y diferenciación

1.1. Derivadas direccionales y parciales

Nota 1.1. En el caso de funciones rales de variable real, $f : A \subseteq \mathbb{R} \longrightarrow \mathbb{R}$, $a \in \operatorname{Int} A$, se define la derivada de la función f en a, $f'(a)$, si existe y es finito el límite

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}. \tag{1.1}$$

Esta misma idea se puede generalizar al caso de $\mathbf{f} : A \subseteq \mathbb{R} \longrightarrow \mathbb{R}^p$ $(p \geq 1)$, $a \in \operatorname{Int} A$, definiendo la derivada de la función $\mathbf{f}$ en a, $\mathbf{f}'(a)$, si existe y es finitio el límite

$$\mathbf{f}'(a) = \lim_{h \to 0} \frac{\mathbf{f}(a+h) - \mathbf{f}(a)}{h}, \tag{1.2}$$

de manera que, además, $\mathbf{f}'(a) = \left(f_1'(a), \ldots, f_p'(a)\right)$ con f_i las funciones componente de $\mathbf{f}$. La ambigüedad en la generalización de este concepto surge si $A \subseteq \mathbb{R}^n$, ya que en ese caso hay que definir el dirección en la que nos aproximamos al punto $\mathbf{a} \in \operatorname{Int} A$.

Definición 1.1 (Derivadas direccionales)**.** Sean $\mathbf{f} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$ y $\mathbf{a} \in \operatorname{Int} A$. Llamaremos **derivada direccional** de $\mathbf{f}$ en $\mathbf{a}$ con dirección $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, y la representaremos por $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$, al límite siguiente, siempre que exista y sea finito,

$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = \lim_{h \to 0} \frac{\mathbf{f}(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a})}{h}. \tag{1.3}$$

Nota 1.2 (Interpretación geométrica)*.* Sean $f : A \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$ y $\mathbf{a} \in \operatorname{Int} A$, $\mathbf{v} \in \mathbb{R}^n \setminus \{(0,0)\}$ (situación que podemos representar gráficamente, ya que $\operatorname{graf} f \subset \mathbb{R}^3$). Sea $R \subset \mathbb{R}^2$ la recta formada por los puntos $\mathbf{a} + h\mathbf{v}$, con $h \in \mathbb{R}$.

Definimos la función de una variable $\sigma : \mathbb{R} \longrightarrow \mathbb{R} \ / \ \sigma(h) = f(\mathbf{a}+h\mathbf{v})$ y obtenemos su derivada en el punto 0, esto es,

$$\sigma'(0) = \lim_{h\to 0} \frac{\sigma(0+h) - \sigma(0)}{h} = \lim_{h\to 0} \frac{f(\mathbf{a}+h\mathbf{v}) - f(\mathbf{a})}{h},$$

por lo tanto, $\sigma'(0) = D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$. Dado que conocemos la interpretación geométrica de la derivada de una función de una variable (ec. 1.1), vemos que $D_{\mathbf{v}}f(\mathbf{a})$ nos da el valor de la pendiente de la recta R' que es tangente a $f(R)$ (curva sobre graf f).

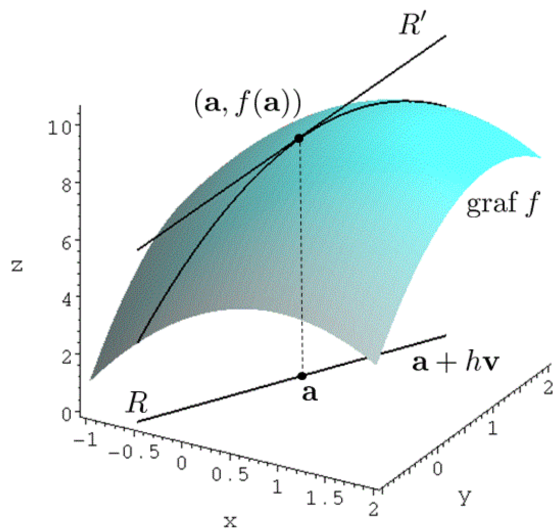


Figura 1.1: Interpretación geométrica de la derivada direccional (función escalar de dos variables).

Nota 1.3. Si cambiamos la recta R, esto es, la dirección $\mathbf{v}$, cambiamos el valor de $\sigma'(0)$ y por lo tanto de la pendiente. En el caso de una variable, la derivada nos da la mejor aproximación a una curva mediante una recta (la recta tangente en el punto, que es única), sin embargo esta propiedad no se conserva por $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$. Veremos que la existencia de $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ tampoco asegura la continuidad de la función $\mathbf{f}$ en $\mathbf{a}$. Por todo ello, la derivada direccional no sirve como generalización del concepto de derivada para funciones de varias variables.

Proposición 1.1. Sean $\mathbf{f} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$, $\mathbf{a} \in \operatorname{Int} A$ y $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Si existe $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$, también existe $D_{\lambda\mathbf{v}}\mathbf{f}(\mathbf{a})$ para cualquier $\lambda \neq 0$ y se tiene que $D_{\lambda\mathbf{v}}\mathbf{f}(\mathbf{a}) = \lambda D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$.

Nota 1.4. En efecto,

$$D_{\lambda\mathbf{v}}\mathbf{f}\left(\mathbf{a}\right)=\lim_{h\to 0}\frac{\mathbf{f}\left(\mathbf{a}+h\lambda\mathbf{v}\right)-f\left(\mathbf{a}\right)}{h}=\lambda\lim_{\lambda h\to 0}\frac{\mathbf{f}\left(\mathbf{a}+h\lambda\mathbf{v}\right)-f\left(\mathbf{a}\right)}{h\lambda}=\lambda D_{\mathbf{v}}\mathbf{f}\left(\mathbf{a}\right).$$

Como $D_{\mathbf{v}}\mathbf{f}\left(\mathbf{a}\right)$ da una medida de la variación de f en la dirección de $\mathbf{v}$, y como vemos depende proporcionalmente del módulo, cuando se quiere comparar medidas de variación en diferentes direcciones se acostumbra a fijar el módulo del vector con $\|\mathbf{v}\|=1$.

Definición 1.2 (Derivadas parciales)**.** Sean el campo escalar $\mathbf{f}:A\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^p$ y $\mathbf{a}\in\operatorname{Int}A$. Consideramos los vectores $\mathbf{e}_i$ $(i=1,\ldots,n)$ de la base canónica de $\mathbb{R}^n$. Se define la **derivada parcial** $i-$ésima de $\mathbf{f}$ en $\mathbf{a}$, y se representa por $\partial\mathbf{f}/\partial x_i\left(\mathbf{a}\right)$ como

$$\frac{\partial\mathbf{f}}{\partial x_i}\left(\mathbf{a}\right)=D_{\mathbf{e}_i}\mathbf{f}\left(\mathbf{a}\right),\tag{1.4}$$

o equivalentemente, si el límite existe y es finito,

$$\frac{\partial\mathbf{f}}{\partial x_i}\left(\mathbf{a}\right)=\lim_{h\to 0}\frac{\mathbf{f}\left(\mathbf{a}+h\mathbf{e}_i\right)-\mathbf{f}\left(\mathbf{a}\right)}{h}=\lim_{h\to 0}\frac{\mathbf{f}\left(a_1,\ldots,a_i+h,\ldots,a_n\right)-\mathbf{f}\left(\mathbf{a}\right)}{h}.\tag{1.5}$$

Ejemplo 1. Estudiamos la existencia y valor de $D_{\mathbf{v}}f\left(0,0\right)$ de $f:\mathbb{R}^2\longrightarrow\mathbb{R}$,

$$f\left(x,y\right)=\begin{cases}\dfrac{x^2y}{x^4+y^2} & \text{si}\quad (x,y)\neq(0,0)\\ 0 & \text{si}\quad (x,y)=(0,0)\end{cases}.$$

Consideramos $\mathbf{v}=\left(v_1,v_2\right)\in\mathbb{R}^n\setminus\{(0,0)\}$. Aplicamos la definición 1.1,

$$\begin{aligned}D_{\mathbf{v}}f\left(0,0\right)&=\lim_{h\to 0}\frac{f\left(h\mathbf{v}\right)-f\left(\mathbf{0}\right)}{h}=\lim_{h\to 0}\frac{f\left(hv_1,hv_2\right)}{h}=\lim_{h\to 0}\frac{1}{h}\frac{\left(hv_1\right)^2hv_2}{\left(hv_1\right)^4+\left(hv_2\right)^2}\\&=\lim_{h\to 0}\frac{v_1^2v_2}{h^2v_1^4+v_2^2}=\frac{v_1^2}{v_2},\text{ si } v_2\neq 0.\end{aligned}$$

Si $v_2=0$ se tiene que $D_{\mathbf{v}}f\left(0,0\right)=0$, ya que el numerador es nulo en todo caso. Por lo tanto, existe $D_{\mathbf{v}}f\left(0,0\right)$ en cualquier dirección $\mathbf{v}\neq\mathbf{0}$.

Estudiemos ahora la continuidad de f en $(0,0)$. Consideramos el límite

$$\lim_{(x,y)\to(0,0)}f\left(x,y\right)=\lim_{(x,y)\to(0,0)}\frac{x^2y}{x^4+y^2},$$

y resolvemos el límite restringido a $S = \{(x,y) \in \mathbb{R}^2 \ / \ y = mx^2,\ m \in \mathbb{R}\}$, esto es,

$$\lim_{\substack{(x,y)\to(0,0)\\(x,y)\in S}} f(x,y) = \lim_{x\to 0} \frac{mx^4}{x^4 + m^2x^4} = \frac{m}{1+m^2},$$

que depende de m, luego no existe $\lim_{(x,y)\to(0,0)} f(x,y)$. Por lo tanto f no es continua en $(0,0)$. Vemos entonces que la existencia de derivadas direccionales no garantiza la continuidad de la función.

Ejemplo 2. Estudiamos la existencia y valor de las derivadas parciales de $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ en $(0,0)$, y de la derivada direccional en el mismo punto según el vector $\mathbf{v} = (1,1)$ con

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{si} \quad (x,y) \neq (0,0) \\ 0 & \text{si} \quad (x,y) = (0,0) \end{cases}.$$

Aplicamos la definición 1.2,

$$\begin{aligned} \frac{\partial f}{\partial x}(0,0) &= \lim_{h\to 0} \frac{f(\mathbf{0} + h(1,0)) - f(\mathbf{0})}{h} = \lim_{h\to 0} \frac{f(h,0)}{h} = 0, \\ \frac{\partial f}{\partial y}(0,0) &= \lim_{h\to 0} \frac{f(\mathbf{0} + h(0,1)) - f(\mathbf{0})}{h} = \lim_{h\to 0} \frac{f(0,h)}{h} = 0, \end{aligned}$$

por lo tanto ambas derivadas parciales existen y su valor es 0. Estudiamos $D_{(1,1)}f(0,0)$:

$$D_{(1,1)}f(0,0) = \lim_{h\to 0} \frac{f(\mathbf{0} + h(1,1)) - f(\mathbf{0})}{h} = \lim_{h\to 0} \frac{f(h,h)}{h} = \lim_{h\to 0} \frac{1}{2h}.$$

Dado que el límite no existe (distinto signo en $h \to 0^{\pm}$), no existe la derivada direccional. Vemos entonces que la existencia de derivadas parciales no garantiza la existencia de direccionales (al contrario sí ya que, evidentemente, son un caso particular).

Nota 1.5 (Cálculo de derivadas parciales). Cuando existe $\partial f/\partial x_i(\mathbf{a})$ es equivalente a $dg/dx_i(x_i)$ (derivada ordinaria) donde $g(x_i) = f(a_1, \ldots, x_i, \ldots, a_n)$, ya que a efectos prácticos se puede considerar que todas las variables salvo x_i son constantes, lo que define la función de una variable $g(x_i)$. Por lo tanto podemos calcular las derivadas parciales usando las **reglas de derivación** de funciones reales de variable real. P. ej.: si $f(x,y) = x^2 + 2\operatorname{sen} y$, entonces $\partial f/\partial x = 2x$, $\partial f/\partial y = 2\cos y$.

Nota 1.6 (Interpretación geométrica). Las derivadas parciales admiten una interpretación geométrica similar a la de la derivada ordinaria de funciones de una variable. Es el caso de la figura 1.1 cuando la recta R es paralela a los ejes coordenados x ó y. Esto es, la gráfica de una función $f(x,y)$ que admite derivadas parciales en todo punto de su dominio es una superficie en $\mathbb{R}^3$,

$$\operatorname{graf} f = \left\{(x, y, f(x,y)) \in \mathbb{R}^3 \ / \ (x,y) \in \operatorname{dom} f\right\},$$

y la intersección de $\operatorname{graf} f$ con el plano $y = y_0$ es una curva dada por la gráfica de la función $g(x) = f(x, y_0)$. La pendiente de la recta tangente a dicha curva en el punto $(x_0, y_0, g(x_0))$ es la derivada parcial de f respecto a x en el punto (x_0, y_0). De forma similar, intersectando con el plano $x = x_0$, se interpreta la derivada parcial de f respecto a y en el punto (x_0, y_0).

1.2. Diferencial de una función

Nota 1.7. Los problemas mostrados en los ejemplos 1 y 2 anteriores muestran que el concepto de derivada direccional no generaliza las propiedades de la derivada de funciones de una variable al caso de varias. El siguiente candidato es el concepto de **diferencial de una función**.

Nota 1.8. En el caso de funciones rales de variable real, $f : A \subseteq \mathbb{R} \longrightarrow \mathbb{R}$, $a \in \operatorname{Int} A$, se dice que f es diferenciable en a si existe una aplicación lineal, $df_a(h) : \mathbb{R} \longrightarrow \mathbb{R}$ (que llamamos diferencial de f en a) tal que

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

Sabemos que la existencia de $df_a(h)$ asegura la derivabilidad de f en a. En el caso de funciones de una variable, este concepto no aporta ventajas operativas ya que se tiene la relación directa $df_a(h) = hf'(a)$. Estudiemos su generalización al caso de funciones de varias variables.

Definición 1.3. Sean $\mathbf{f} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$ y $\mathbf{a} \in \operatorname{Int} A$. Se dice que $\mathbf{f}$ es **diferenciable** en $\mathbf{a}$ si existe una aplicación lineal, $d\mathbf{f}_{\mathbf{a}}(\mathbf{h}) : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ (que llamamos diferencial de $\mathbf{f}$ en $\mathbf{a}$) tal que

$$\lim_{\mathbf{h}\to \mathbf{0}} \frac{\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) - d\mathbf{f}_{\mathbf{a}}(\mathbf{h})}{\|\mathbf{h}\|} = 0, \tag{1.6}$$

o equivalentemente, si existe $\boldsymbol{\xi} : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ con $\lim_{\mathbf{h}\to\mathbf{0}} \boldsymbol{\xi}(\mathbf{h}) = \mathbf{0}$ tal que

$$\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) = d\mathbf{f}_{\mathbf{a}}(\mathbf{h}) + \|\mathbf{h}\|\, \boldsymbol{\xi}(\mathbf{h}). \tag{1.7}$$

Nota 1.9. Nótese que

$$\lim_{\mathbf{h}\to\mathbf{0}} \mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) = \lim_{\mathbf{h}\to\mathbf{0}} \left[d\mathbf{f}_{\mathbf{a}}(\mathbf{h}) + \|\mathbf{h}\|\, \boldsymbol{\xi}(\mathbf{h})\right] = \mathbf{0},$$

esto es, $\lim_{\mathbf{h}\to\mathbf{0}} \mathbf{f}(\mathbf{a}+\mathbf{h}) = \mathbf{f}(\mathbf{a})$, luego la función es continua en $\mathbf{a}$ (basta sustituir $\mathbf{x} = \mathbf{a}+\mathbf{h}$ para recuperar la definición). Lo enunciamos.

Proposición 1.2. Sean $\mathbf{f} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$ y $\mathbf{a} \in \operatorname{Int} A$. Si $\mathbf{f}$ es diferenciable en $\mathbf{a}$ entonces $\mathbf{f}$ es continua en $\mathbf{a}$.

Proposición 1.3. Sean $\mathbf{f} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$, $\mathbf{a} \in \operatorname{Int} A$ y f_i $(i = 1, \ldots, p)$ las funciones componente de $\mathbf{f}$. Entonces

1. Existe $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ para todo $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ si, y sólo si, existen $D_{\mathbf{v}} f_i(\mathbf{a})$ $(i = 1, \ldots, p)$, y además
$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = \left(D_{\mathbf{v}} f_1(\mathbf{a}), \ldots, D_{\mathbf{v}} f_p(\mathbf{a})\right). \tag{1.8}$$

2. $\mathbf{f}$ es diferenciable en $\mathbf{a}$ si, y sólo si, f_i $(i = 1, \ldots, p)$ son diferenciables en $\mathbf{a}$, y además
$$d\mathbf{f}_{\mathbf{a}} = \left(df_{1\mathbf{a}}, \ldots, df_{p\mathbf{a}}\right). \tag{1.9}$$

3. Si $\mathbf{f}$ es diferenciable en $\mathbf{a}$ entonces $d\mathbf{f}_{\mathbf{a}}$ es única.

4. Si $\mathbf{f}$ es diferenciable en $\mathbf{a}$ entonces existe $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ para todo $\mathbf{v} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Además
$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = d\mathbf{f}_{\mathbf{a}}(\mathbf{v}). \tag{1.10}$$

5. Si $\mathbf{f}$ es diferenciable en $\mathbf{a}$ entonces para todo $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ se tiene que
$$D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = v_1 \frac{\partial \mathbf{f}}{\partial x_1}(\mathbf{a}) + \ldots + v_n \frac{\partial \mathbf{f}}{\partial x_n}(\mathbf{a}). \tag{1.11}$$

Nota 1.10. La importante propiedad 4 anterior, que nos relaciona la diferencial con las derivadas direccionales (si $\mathbf{f}$ es diferenciable en el punto), se sigue de la definición de $D_{\mathbf{v}}\mathbf{f}(\mathbf{a})$ y de $d\mathbf{f}_{\mathbf{a}}(\mathbf{h})$ a través de (1.7). En efecto,

$$\begin{aligned} D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) &= \lim_{h\to 0} \frac{\mathbf{f}(\mathbf{a}+h\mathbf{v}) - \mathbf{f}(\mathbf{a})}{h} = \lim_{h\to 0} \frac{d\mathbf{f}_{\mathbf{a}}(h\mathbf{v}) + \|h\mathbf{v}\|\, \boldsymbol{\xi}(h\mathbf{v})\,.}{h} = \\ &= \lim_{h\to 0} \frac{d\mathbf{f}_{\mathbf{a}}(h\mathbf{v})}{h} \overset{\text{(lineal)}}{=} \lim_{h\to 0} \frac{h}{h} d\mathbf{f}_{\mathbf{a}}(\mathbf{v}) = d\mathbf{f}_{\mathbf{a}}(\mathbf{v}), \end{aligned}$$

donde se ha usado que $d\mathbf{f}_{\mathbf{a}}(h\mathbf{v})$ es una aplicación lineal.

Nota 1.11. Incidamos sobre la propiedad 5. Si la función es diferenciable en un punto, entonces sus derivadas direccionales se pueden escribir en términos de las derivadas parciales (que actúan a modo de "base de derivadas"). Veamos que esto es también consecuencia de la linealidad de la diferencial y de resultados previos:

$$\begin{aligned} D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) \overset{(1.10)}{=} d\mathbf{f}_{\mathbf{a}}(\mathbf{v}) &= d\mathbf{f}_{\mathbf{a}}\left(\sum_{i=1}^{n} v_i \mathbf{e}_i\right) \overset{(\text{lineal})}{=} \sum_{i=1}^{n} v_i d\mathbf{f}_{\mathbf{a}}(\mathbf{e}_i) \\ &\overset{(1.10)}{=} \sum_{i=1}^{n} v_i D_{\mathbf{e}_i}\mathbf{f}(\mathbf{a}) \overset{(1.4)}{=} \sum_{i=1}^{n} v_i \frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{a}). \end{aligned}$$

Proposición 1.4 (Condición suficiente para la diferenciabilidad)**.** Sean $\mathbf{f} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$ y $\mathbf{a} \in \operatorname{Int} A$. Si existen las derivadas parciales $\partial \mathbf{f}/\partial x_i$ en un entorno de $\mathbf{a}$ para $i = 1, \ldots, n$, y son continuas en $\mathbf{a}$, entonces $\mathbf{f}$ es diferenciable en $\mathbf{a}$.

Ejemplo 3. Los polinomios $p(x_1, \ldots, x_n)$ son funciones diferenciables en todo punto, ya que en todo punto existen sus derivadas parciales y son continuas.

Ejemplo 4 (Contraejemplo)**.** La función

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

cumple que es continua y diferenciable en $(0, 0)$, sus derivadas parciales existen en un entorno de $(0, 0)$, pero no son continuas en dicho punto.

1.3. Funciones diferenciables

Nota 1.12. Dado que una aplicación lineal referida a las bases canónicas se puede representar matricialmente, podemos hallar la forma de estas matrices para la diferencial.

1.3.1. Funciones escalares

Nota 1.13. Sean $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ y $\mathbf{a} \in \operatorname{Int} A$, con f diferenciable en $\mathbf{a}$. Sea $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n \setminus \{\mathbf{0}\}$. Se tiene

$$\begin{aligned} D_{\mathbf{v}} f(\mathbf{a}) &= df_{\mathbf{a}}(\mathbf{v}) = \sum_{i=1}^{n} v_i \frac{\partial f}{\partial x_i}(\mathbf{a}) \\ &= \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{a})\right) \cdot (v_1, \ldots, v_n), \end{aligned} \tag{1.12}$$

donde $\cdot$ es el producto escalar canónico.

Definición 1.4 (Gradiente)**.** La matriz asociada a la aplicación lineal $df_{\mathbf{a}}$ respecto de las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}$ se denomina **gradiente** de f en $\mathbf{a}$, se representa por $\boldsymbol{\nabla} f(\mathbf{a})$, y viene dada por

$$\boldsymbol{\nabla} f(\mathbf{a}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{a}) \right).$$

Nota 1.14*.* Entonces, en las condiciones de la nota 1.13 se tiene que

$$D_{\mathbf{v}} f(\mathbf{a}) = df_{\mathbf{a}}(\mathbf{v}) = \boldsymbol{\nabla} f(\mathbf{a}) \cdot \mathbf{v},$$

donde matricialmente $\boldsymbol{\nabla} f(\mathbf{a})$ es una matriz fila $1 \times n$ y $\mathbf{v}$ se representa por una matriz columna $n \times 1$, de modo que el producto escalar viene dado por el producto matricial usual, y se denota por $\boldsymbol{\nabla} f(\mathbf{a})\, \mathbf{v}$.

Nota 1.15*.* También se utiliza la notación $\boldsymbol{\nabla} f(\mathbf{a}) = f'(\mathbf{a})$, incidiendo en que es la generalización de derivada en un punto para funciones escalares de varias variables.

Nota 1.16*.* Si existe la diferencial de f en $\mathbf{a}$ entonces viene dada por el gradiente, esto es, $df_{\mathbf{a}} = \boldsymbol{\nabla} f(\mathbf{a})$ (aunque puede ser más claro escribir $df_{\mathbf{a}} = \boldsymbol{\nabla} f(\mathbf{a}) \,\cdot$, indicando así la manera en la que opera la aplicación lineal). Pero en cualquier caso hay que verificar la existencia de $df_{\mathbf{a}}$ mediante su definición (1.3) o proposiciones que la garanticen, ya que aun teniendo un vector gradiente, si el límite no existe, no existe la diferencial.

Ejemplo 5. Vemos el ejemplo de una función para la que existen las derivadas direccionales en $(0,0)$, y por lo tanto el gradiente, pero en cambio no es diferenciable en $(0,0)$. Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$,

$$f(x,y) = \begin{cases} \dfrac{x^3 y}{x^4 + y^2} & \text{si} \quad (x,y) \neq (0,0) \\ 0 & \text{si} \quad (x,y) = (0,0) \end{cases}.$$

Consideramos $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^n \setminus \{(0,0)\}$. Aplicamos la definición 1.1,

$$\begin{aligned} D_{\mathbf{v}} f(0,0) &= \lim_{h \to 0} \frac{f(h\mathbf{v}) - f(\mathbf{0})}{h} = \lim_{h \to 0} \frac{f(hv_1, hv_2)}{h} = \lim_{h \to 0} \frac{1}{h} \frac{(hv_1)^3\, hv_2}{(hv_1)^4 + (hv_2)^2} \\ &= \lim_{h \to 0} \frac{hv_1^3 v_2}{h^2 v_1^4 + v_2^2} = 0, \text{ si } v_2 \neq 0. \end{aligned}$$

Si $v_2 = 0$ se tiene que $D_{\mathbf{v}} f(0,0) = 0$, ya que el numerador es nulo en todo caso. Por lo tanto, existe $D_{\mathbf{v}} f(0,0)$ en cualquier dirección $\mathbf{v} \neq \mathbf{0}$. El gradiente es entonces

$$\boldsymbol{\nabla} f(0,0) = \left(\frac{\partial f}{\partial x}(0,0), \frac{\partial f}{\partial y}(0,0) \right) = (0,0).$$

Si f fuese diferenciable en $(0,0)$, entonces la diferencial sería $df_{(0,0)} = \nabla f(0,0) = (0,0)$ (que es lineal de forma trivial). Tomamos $df_{(0,0)}$ como candidato para diferencial y aplicamos la definición 1.3. Esto es,

$$\begin{aligned}
& \lim_{(h_1,h_2)\to(0,0)} \frac{f((0,0)+(h_1,h_2)) - f((0,0)) - df_{(0,0)}(h_1,h_2)}{\|(h_1,h_2)\|} \\
= & \lim_{(h_1,h_2)\to(0,0)} \frac{f(h_1,h_2) - 0 - (0,0)\cdot(h_1,h_2)}{\sqrt{h_1^2+h_2^2}} \\
= & \lim_{(h_1,h_2)\to(0,0)} \frac{1}{\sqrt{h_1^2+h_2^2}} \frac{h_1^3 h_2}{h_1^4+h_2^2}.
\end{aligned}$$

Consideramos el límite restringido a $S = \{(h_1,h_2)\in\mathbb{R}^2 \ / \ h_2 = mh_1^2,\, m\in\mathbb{R}\}$, esto es,

$$\begin{aligned}
\lim_{\substack{(h_1,h_2)\to(0,0)\\(h_1,h_2)\in S}} \frac{1}{\sqrt{h_1^2+h_2^2}} \frac{h_1^3 h_2}{h_1^4+h_2^2} &= \lim_{h_1\to 0} \frac{1}{\sqrt{h_1^2+(mh_1^2)^2}} \frac{h_1^3 m h_1^2}{h_1^4+(mh_1^2)^2} \\
&= \lim_{h_1\to 0} \frac{1}{|h_1|\sqrt{1+m^2h_1^2}} \frac{h_1^5 m}{h_1^4(1+m^2)}
\end{aligned}$$

cuyo valor no está definido, y por lo tanto no existe el límite. Concluimos entonces que f no es diferenciable en $(0,0)$.

Ejemplo 6. Las funciones lineales

$$f(\mathbf{x}) = \alpha_1 x_1 + \ldots + \alpha_n x_n = \sum_{i=1}^{n} \alpha_i x_i = \boldsymbol{\alpha}\cdot\mathbf{x}$$

son diferenciables en todo punto (polinomios). Para estas, además, la diferencial coincide con la propia función. En efecto, calculamos el gradiente

$$\nabla f = \left(\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_n}\right) = (\alpha_1,\ldots,\alpha_n) = \boldsymbol{\alpha},$$

cuyo valor es independiente del punto. Por lo tanto su diferencial en $\mathbf{a}\in\mathbb{R}^n$ vendrá dada por

$$df_{\mathbf{a}}(\mathbf{h}) = \nabla f(\mathbf{a})\cdot\mathbf{h} = \boldsymbol{\alpha}\cdot\mathbf{h} = f(\mathbf{h}).$$

Proposición 1.5. Sean $f : A\subseteq\mathbb{R}^n \longrightarrow \mathbb{R}$ y $\mathbf{a}\in \operatorname{Int} A$, con f diferenciable en $\mathbf{a}$. Entonces, si $\nabla f(\mathbf{a}) \neq \mathbf{0}$ indica la dirección y sentido en que la derivada direccional $D_{\mathbf{u}} f(\mathbf{a})$ con $\|\mathbf{u}\| = 1$ es máxima. En tal dirección y sentido se tiene que $D_{\mathbf{u}} f(\mathbf{a}) = \|\nabla f(\mathbf{a})\|$.

Nota 1.17. En efecto, si consideramos $D_{\mathbf{u}}f(\mathbf{a})$ con $\|\mathbf{u}\| = 1$ para comparar direcciones (nota 1.4), resulta

$$D_{\mathbf{u}}f(\mathbf{a}) = \boldsymbol{\nabla} f(\mathbf{a}) \cdot \mathbf{u} = \|\boldsymbol{\nabla} f(\mathbf{a})\| \, \|\mathbf{u}\| \cos\theta$$

con θ el ángulo formado por los dos vectores, entonces es máxima si $\theta = 0$. En esa dirección, $D_{\mathbf{u}}f(\mathbf{a}) = \|\boldsymbol{\nabla} f(\mathbf{a})\|$. Por lo tanto, $\boldsymbol{\nabla} f(\mathbf{a})$ indica la dirección y sentido de **máximo crecimiento** de la función f.

Ejemplo 7. Dada $f(x,y,z) = x^3 - 3xy + z^3$, hallamos $D_{\mathbf{v}}f(-1,0,0)$ con $\mathbf{v} = (1,-2,1)$. Dado que f es un polinomio, es diferenciable en $\mathbb{R}^3$. Calculamos $\boldsymbol{\nabla} f(x,y,z) = (3x^2 - 3y, -3x, 3z^2)$, entonces

$$D_{\mathbf{v}}f(-1,0,0) = \boldsymbol{\nabla} f(-1,0,0) \cdot \mathbf{v} = (3,3,0) \cdot (1,-2,1) = -3.$$

Por la proposición 1.5, la dirección de máxima derivada direccional es la del vector gradiente $\boldsymbol{\nabla} f(-1,0,0) = (3,3,0)$, y el valor de este máximo es $\|(3,3,0)\| = 3\sqrt{2}$.

Nota 1.18. El ejemplo 9 que estudiaremos más adelante muestra gráficamente esta propiedad del gradiente.

Proposición 1.6 (Reglas de diferenciación)**.** Sean $f, g : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ y $\mathbf{a} \in \operatorname{Int} A$, diferenciables en $\mathbf{a}$. Entonces se cumple que

1. $f + g$ es diferenciable en $\mathbf{a}$ y $d(f+g)_{\mathbf{a}} = df_{\mathbf{a}} + dg_{\mathbf{a}}$.
2. fg es diferenciable en $\mathbf{a}$ y $d(fg)_{\mathbf{a}} = g(\mathbf{a})\,df_{\mathbf{a}} + f(\mathbf{a})\,dg_{\mathbf{a}}$.
3. si $g(\mathbf{a}) \neq 0$, f/g es diferenciable en $\mathbf{a}$ y $d(f/g)_{\mathbf{a}} = \dfrac{g(\mathbf{a})\,df_{\mathbf{a}} - f(\mathbf{a})\,dg_{\mathbf{a}}}{g(\mathbf{a})^2}$.

Nota 1.19 (Interpretación geométrica). Sean $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ y $\mathbf{a} \in \operatorname{Int} A$, con f diferenciable en $\mathbf{a}$. Si en la definición 1.7 sustituimos $z = f(\mathbf{a} + \mathbf{h}) - \|\mathbf{h}\|\,\xi(\mathbf{h})$, nos queda

$$z - f(\mathbf{a}) = df_{\mathbf{a}}(\mathbf{h}) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(\mathbf{a}).$$

Esto es, haciendo $\mathbf{x} = \mathbf{a} + \mathbf{h}$,

$$z - f(\mathbf{a}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})(x_i - a_i), \tag{1.13}$$

la ecuación de un hiperplano en $\mathbb{R}^n \times \mathbb{R}$ que pasa por el punto $(\mathbf{a}, f(\mathbf{a}))$ y que se denomina **hiperplano tangente** a la hipersuperficie $\operatorname{graf} f$ en el punto. El vector $(-\boldsymbol{\nabla} f(\mathbf{a}), 1)$ es normal al hiperplano tangente en el punto.

Nota 1.20. Si $n = 1$, la ecuación (1.13) es la ecuación de la *recta tangente* a graf f, que en general denominamos curva $y = f(x)$, en el punto $x = a$,

$$y - f(a) = f'(a)(x - a). \tag{1.14}$$

El vector $(-f'(a), 1)$ es normal a la recta tangente en el punto.

Si $n = 2$, la ecuación (1.13) es la ecuación del *plano tangente* a graf f, que en general denominamos superficie $z = f(x, y)$, en el punto $(x, y) = (a_1, a_2)$,

$$z - f(a_1, a_2) = \frac{\partial f}{\partial x}(a_1, a_2)(x - a_1) + \frac{\partial f}{\partial y}(a_1, a_2)(x - a_2). \tag{1.15}$$

El vector $(-\nabla f(a_1, a_2), 1)$ es normal al plano tangente en el punto.

Ejemplo 8. Dada $f(x, y) = 1 + (x-1)^2 + (y-1)^2$ obtenemos el plano tangente en $(1/2, 1/2)$. Dado que f es un polinomio, en virtud de la proposición 1.4 f es diferenciable en todo punto. Calculamos

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2(x-1) \rightarrow \frac{\partial f}{\partial x}\left(\frac{1}{2}, \frac{1}{2}\right) = -1, \\ \frac{\partial f}{\partial y} &= 2(y-1) \rightarrow \frac{\partial f}{\partial y}\left(\frac{1}{2}, \frac{1}{2}\right) = -1, \end{aligned}$$

por lo tanto $\nabla f(1/2, 1/2) = (-1, -1)$. Entonces, la ecuación del plano tangente es

$$z - f\left(\frac{1}{2}, \frac{1}{2}\right) = -1\left(x - \frac{1}{2}\right) - 1\left(y - \frac{1}{2}\right) \rightarrow x + y + z - \frac{5}{2} = 0.$$

Nota 1.21. Si la función f no es diferenciable en $\mathbf{a}$, pero tiene derivadas parciales, la ecuación (1.13) se puede construir igualmente pero no representa el plano tangente, sino simplemente a un plano que pasa por $(\mathbf{a}, f(\mathbf{a}))$.

Nota 1.22. En las condiciones de la nota 1.19, si se hace una traslación del sistema de referencia, llevando el origen al punto $(\mathbf{a}, f(\mathbf{a}))$, entonces la hipersuperficie graf f pasará por el origen, así como el hiperplano tangente, cuya ecuación será $z = df_{\mathbf{0}}(\mathbf{x})$, esto es, coincide con la diferencial de la función en el origen. Por lo tanto, *la actuación de la diferencial de la funcion f en el punto* $\mathbf{a}$ *es la ecuación del hiperplano tangente a* graf f *en* $(\mathbf{a}, f(\mathbf{a}))$, *referida a "coordenadas locales"*.

Proposición 1.7 (Gradiente y conjuntos de nivel)**.** Sean $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ y $\mathbf{a} \in \operatorname{Int} A$, con f diferenciable en $\mathbf{a}$, y L_k el conjunto de nivel definido por $f(\mathbf{x}) = k, k \in \mathbb{R}$. Entonces, el vector $\nabla f(\mathbf{a})$ es normal a L_k (en el sentido de que es normal a los vectores tangentes al conjunto de nivel en ese punto).

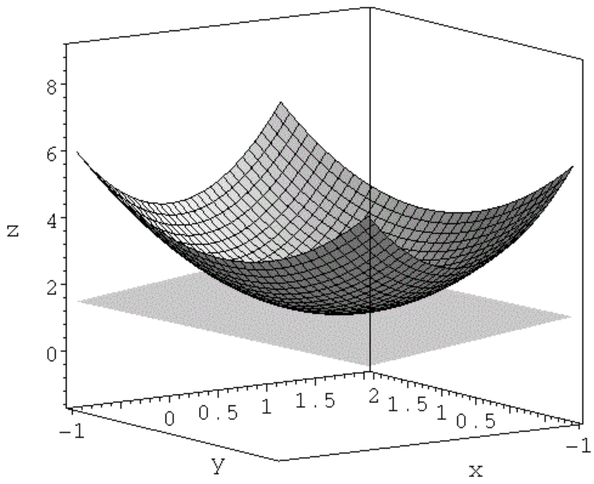


Figura 1.2: Gráfica de la función y plano tangente del ejemplo 8.

Nota 1.23. En efecto, dentro del conjunto de nivel de cierto valor k, la función se mantiene constante, es decir, si $\mathbf{v} \neq \mathbf{0}$ es un vector tangente al conjunto de nivel en el punto $\mathbf{a}$, entonces $D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = 0$. Dado que f es diferenciable en $\mathbf{a}$, $D_{\mathbf{v}}\mathbf{f}(\mathbf{a}) = \boldsymbol{\nabla} f(\mathbf{a}) \cdot \mathbf{v} = 0$, y por lo tanto el vector gradiente es perpendicular al vector tangente en el punto.

Ejemplo 9. La figura 1.3 muestra la gráfica de la función

$$f(x,y) = \left(x^2 + 3y^2\right) e^{1-x^2-y^2}.$$

Apreciamos que ésta alcanza dos valores máximos en los puntos $(0,1)$ y $(0,-1)$, y alcanza un valor mínimo en $(0,0)$ (la definición y métodos de cálculo de máximos y mínimos, relativos y absolutos, se estudian en el capítulo 3).

La figura 1.4 representa las curvas de nivel de $f(x,y)$ y el campo de vectores gradiente. Se aprecia que los vectores gradiente son perpendiculares a las curvas de nivel, y apuntan hacia los puntos de valores máximos de la función ('sumideros' de líneas de campo), alejándose del punto de valor mínimo ('fuente' de líneas de campo). Los puntos $(-1,0)$ y $(0,-1)$ también muestran un comportamiento característico del campo de vectores a su alrededor (las líneas de campo parecen esquivarlos), se denominan puntos de silla. En todos estos puntos, el valor del gradiente es nulo.

1.3.2. Funciones vectoriales

Nota 1.24. Sean $\mathbf{f} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$, $\mathbf{a} \in \operatorname{Int} A$ y f_i $(i = 1, \dots, p)$ las funciones componente de $\mathbf{f}$, con $\mathbf{f}$ diferenciable en $\mathbf{a}$. Sea $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^n \setminus \{\mathbf{0}\}$.

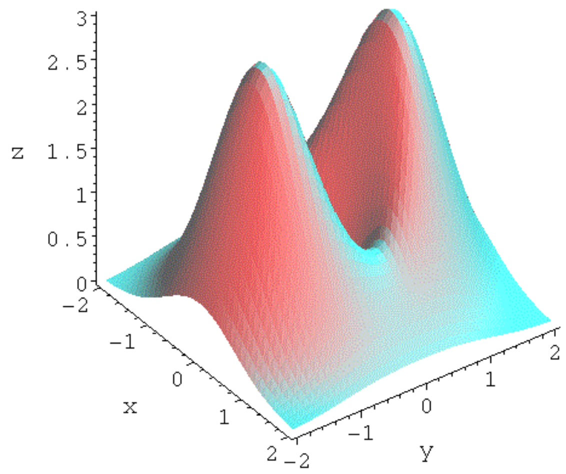


Figura 1.3: Gráfica de la función del ejemplo 9.

Se tiene

$$D_{\mathbf{v}}\mathbf{f}\left(\mathbf{a}\right)=d\mathbf{f}_{\mathbf{a}}\left(\mathbf{v}\right)=\left(D_{\mathbf{v}}f_1\left(\mathbf{a}\right),\ldots,D_{\mathbf{v}}f_p\left(\mathbf{a}\right)\right). \tag{1.16}$$

Definición 1.5 (Jacobiana)**.** La matriz asociada a la apliación lineal $d\mathbf{f}_{\mathbf{a}}$ respecto de las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^p$ se denomina **matriz jacobiana** de $\mathbf{f}$ en $\mathbf{a}$, se representa por $J\mathbf{f}\left(\mathbf{a}\right)$, y viene dada por

$$J\mathbf{f}\left(\mathbf{a}\right)=\begin{pmatrix}\boldsymbol{\nabla}f_1\left(\mathbf{a}\right)\\ \boldsymbol{\nabla}f_2\left(\mathbf{a}\right)\\ \vdots\\ \boldsymbol{\nabla}f_p\left(\mathbf{a}\right)\end{pmatrix}=\begin{pmatrix}\frac{\partial f_1}{\partial x_1}\left(\mathbf{a}\right) & \cdots & \frac{\partial f_1}{\partial x_n}\left(\mathbf{a}\right)\\ \vdots & & \vdots\\ \frac{\partial f_p}{\partial x_1}\left(\mathbf{a}\right) & \cdots & \frac{\partial f_p}{\partial x_n}\left(\mathbf{a}\right)\end{pmatrix}. \tag{1.17}$$

Nota 1.25*.* Entonces, en las condiciones de la nota 1.24 se tiene que

$$d\mathbf{f}_{\mathbf{a}}\left(\mathbf{v}\right)=J\mathbf{f}\left(\mathbf{a}\right)\mathbf{v}, \tag{1.18}$$

donde matricialmente $Jf\left(\mathbf{a}\right)$ es una matriz $p\times n$ y $\mathbf{v}$ se representa por una matriz columna $n\times 1$, y se aplica el producto matricial usual.

Nota 1.26*.* Si $p=1$ se recupera el caso escalar, $df_{\mathbf{a}}\left(\mathbf{v}\right)=Jf\left(\mathbf{a}\right)\mathbf{v}=\boldsymbol{\nabla}f\left(\mathbf{a}\right)\mathbf{v}$.

Nota 1.27*.* También se denota $J\mathbf{f}\left(\mathbf{a}\right)=\mathbf{f}'\left(\mathbf{a}\right)$, incidiendo en que es la generalización de derivada en un punto para funciones vectoriales de varias variables.

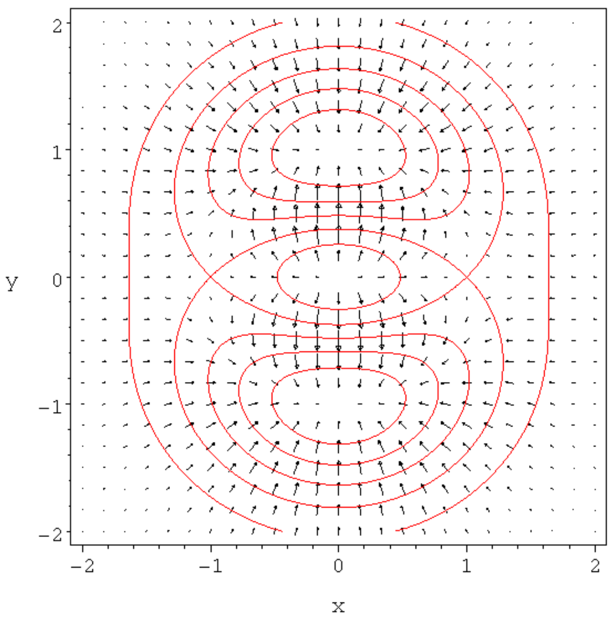


Figura 1.4: Curvas de nivel y campo de vectores gradiente del ejemplo 9.

Nota 1.28. Si existe la diferencial de $\mathbf{f}$ en $\mathbf{a}$ entonces viene dada por la jacobiana, esto es, $d\mathbf{f}_{\mathbf{a}} = J\mathbf{f}\,(\mathbf{a})$ (y la aplicación lineal opera mediante el producto matricial). Pero en cualquier caso hay que verificar la existencia de $d\mathbf{f}_{\mathbf{a}}$ mediante su definición 1.3 o proposiciones que la garanticen, ya que aun teniendo una matriz jacobiana, si el límite no existe, no existe la diferencial.

Ejemplo 10. Calculamos la diferencial de $\mathbf{f}\,(x, y) = (e^x \cos y, x^y)$. La función está definida en $\operatorname{dom} \mathbf{f} = \{(x, y) \in \mathbb{R}^2 \ / \ x > 0\}$, que es un conjunto abierto. Obtenemos la matriz jacobiana (1.17) en cualquier punto del dominio (x, y),

$$J\mathbf{f}\,(x,y) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\[2ex] \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} e^x \cos y & -e^x \operatorname{sen} y \\ y x^{y-1} & x^y \log x \end{pmatrix}.$$

Las derivadas parciales de las funciones componente f_1 y f_2 están definidas y son continuas en todo punto del dominio, entonces son diferenciables por la proposición 1.4, y $\mathbf{f}$ es diferenciable por la proposición 1.3. La diferencial de $\mathbf{f}$ resulta (ec. 1.18)

$$\begin{aligned} d\mathbf{f}_{(x,y)}(\mathbf{h}) &= \begin{pmatrix} e^x \cos y & -e^x \operatorname{sen} y \\ yx^{y-1} & x^y \log x \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \\ &= \begin{pmatrix} e^x \cos(y) h_1 - e^x \operatorname{sen}(y) h_2 \\ yx^{y-1} h_1 + x^y \log(x) h_2 \end{pmatrix}. \end{aligned}$$

1.4. Funciones de clase m

Nota 1.29. La proposición 1.4 establece una condición suficiente para la diferenciabilidad muy usada en la práctica. Sus hipótesis motivan la siguiente definición para caracterizar las funciones denominadas de "clase 1".

Definición 1.6. Sea $\mathbf{f} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$, con A conjunto abierto. Se dice que

a) $\mathbf{f}$ es de **clase** 0 en A, y se escribe $\mathbf{f} \in C^0(A)$, si $\mathbf{f}$ es continua en A.

b) $\mathbf{f}$ es de **clase** 1 en A, y se escribe $\mathbf{f} \in C^1(A)$, si existen sus derivadas parciales $\partial \mathbf{f}/\partial x_i$ $(i = 1, \dots, n)$ y son continuas en A.

Nota 1.30. Si $\mathbf{f} \in C^1(A)$, por aplicación de la proposición 1.4 $\mathbf{f}$ es diferenciable en A, entonces por la proposición 1.2 $\mathbf{f}$ es continua en A, esto es, $\mathbf{f} \in C^0(A)$. La implicación contraria no se da en general.

Nota 1.31. Sea $\mathbf{f} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p$. Las derivadas parciales en cada punto permiten construir la **función derivada parcial** $i-$**ésima**, es decir, si

$$A_i = \left\{ \mathbf{a} \in \operatorname{Int} A \text{ tal que existe } \frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{a}) \right\},$$

definimos

$$\frac{\partial \mathbf{f}}{\partial x_i} : A_i \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p, \quad \left(\frac{\partial \mathbf{f}}{\partial x_i} \right)(\mathbf{a}) = \frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{a}). \tag{1.19}$$

A las derivadas parciales de $\mathbf{f}$ las llamaremos de **primer orden**.

La función $\partial \mathbf{f}/\partial x_i$ puede admitir a su vez derivadas parciales respecto a cualquiera de las variables, que denominaremos **derivadas parciales de segundo orden**, esto es, si

$$A_{ij} = \left\{ \mathbf{a} \in \operatorname{Int} A_i \text{ tal que existe } \frac{\partial}{\partial x_j} \left(\frac{\partial \mathbf{f}}{\partial x_i} \right)(\mathbf{a}) \right\},$$

se define

$$\frac{\partial^2 \mathbf{f}}{\partial x_i \partial x_j} : A_{ij} \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^p, \quad \left(\frac{\partial^2 \mathbf{f}}{\partial x_i \partial x_j} \right)(\mathbf{a}) = \frac{\partial}{\partial x_j} \left(\frac{\partial \mathbf{f}}{\partial x_i} \right)(\mathbf{a}), \tag{1.20}$$

y así, sucesivamente.

Nota 1.32. Las funciones $\frac{\partial^2 \mathbf{f}}{\partial x_i \partial x_j}$ y $\frac{\partial^2 \mathbf{f}}{\partial x_j \partial x_i}$ $(i \neq j)$ se denominan derivadas **cruzadas** o **mixtas**. En general se tiene que

$$\frac{\partial^2 \mathbf{f}}{\partial x_i \partial x_j} \neq \frac{\partial^2 \mathbf{f}}{\partial x_j \partial x_i} \ (i \neq j),$$

aunque se pueden dar condiciones suficientes para asegurar la igualdad, que es la situación más frecuente. Aunque estamos trabajando con funciones vectoriales, este problema se puede tratar con funciones escalares de varias variables ya que como sabemos el carácter vectorial se establece coordenada a coordenada.

Teorema 1 (de Schwarz)**.** *Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A conjunto abierto, y $\mathbf{a} \in A$. Se supone que existen en A las funciones $\frac{\partial f}{\partial x_i}$, $\frac{\partial f}{\partial x_j}$ y $\frac{\partial^2 f}{\partial x_i \partial x_j}$ $(i \neq j)$ y que $\frac{\partial^2 f}{\partial x_i \partial x_j}$ es continua en $\mathbf{a}$. Entonces existe $\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{a})$ y además se tiene*

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

Nota 1.33. Obsérvese que en las hipótesis del teorema no se pide la continuidad de f ni las de sus derivadas parciales de primer orden en A.

Nota 1.34. El número de derivadas parciales de primer orden en $\mathbf{a}$, supuestas existentes, es n. El número de derivadas parciales de segundo orden es n^2, y en general el de orden m es n^m. Pero si se tiene en cuenta la igualdad de derivadas mixtas, el número de derivadas distintas de orden m son las combinaciones con repetición de n elementos tomados de m en m, es decir,

$$\binom{n+m-1}{m} = \frac{(n+m-1)!}{m!\,(n-1)!}.$$

Nota 1.35. Como las condiciones del teorema de Schwarz son suficientes, resulta que si en un entorno de $\mathbf{a}$ existen $\frac{\partial f}{\partial x_i}$, $\frac{\partial f}{\partial x_j}$, $\frac{\partial^2 f}{\partial x_i \partial x_j}$ y $\frac{\partial^2 f}{\partial x_i \partial x_j}$ $(i \neq j)$ y las de segundo orden son distintas en $\mathbf{a}$, forzosamente ambas son discontinuas en $\mathbf{a}$. Un caso ilustrativo de esta situación es el ejemplo 11, que además nos muestra el cálculo de derivadas parciales de segundo orden.

Ejemplo 11 (de Peano)**.** Sea la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$,

$$f(x,y) = \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2} & \text{si} \quad (x,y) \neq (0,0) \\ 0 & \text{si} \quad (x,y) = (0,0) \end{cases}.$$

Queremos obtener las derivadas parciales cruzadas de segundo orden en el punto $(0,0)$. Calcuamos

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0)=\lim_{h\to 0}\frac{1}{h}\left[\frac{\partial f}{\partial y}(h,0)-\frac{\partial f}{\partial y}(0,0)\right],$$

para ello necesitamos $\partial f/\partial y$ en puntos de la forma $(x,0)$, esto es,

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

Sustituyendo este resultado en el límite anterior,

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(0,0)=\lim_{h\to 0}\frac{h-0}{h}=1.$$

Repitiendo el proceso para la otra derivada mixta, se obtiene (se deja como ejercicio),

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(0,0)=\lim_{h\to 0}\frac{-h-0}{h}=-1.$$

Por lo tanto,

$$\frac{\partial^2 f}{\partial x\partial y}(0,0)\neq\frac{\partial^2 f}{\partial y\partial x}(0,0).$$

Esto ocurre porque ambas derivadas son discontinuas en $(0,0)$. En efecto, para todo $(x,y)\in\mathbb{R}^2\setminus\{(0,0)\}$, aplicando las reglas de derivación se tiene que

$$\frac{\partial^2 f}{\partial x\partial y}=\frac{\partial^2 f}{\partial y\partial x}=\frac{x^6+9x^4y^2-9x^2y^4-y^6}{\left(x^2+y^2\right)^3},$$

que no tiene límite en $(0,0)$ (lo que es fácil de ver usando límites restringidos a conjuntos con $x=0$ e $y=0$).

Definición 1.7 (Funciones de clase m)**.** Sea $\mathbf{f}:A\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^p$, con A conjunto abierto. Se dice que $\mathbf{f}$ es de **clase** m en A, y se escribe $\mathbf{f}\in C^m(A)$, si existen todas las derivadas parciales de orden m en A y son continuas en A.

Nota 1.36. Sea $f\in C^2(A)$, esto es, existen $\frac{\partial^2 f}{\partial x_i\partial x_j}$ $(i,j=1,\ldots,n)$ y son continuas en A. Dado que $\frac{\partial^2 f}{\partial x_i\partial x_j}$ son las derivadas parciales de $\frac{\partial f}{\partial x_i}$ entonces $\frac{\partial f}{\partial x_i}\in C^1(A)$. Por aplicación de la proposición 1.4, $\frac{\partial f}{\partial x_i}$ es diferenciable en A, entonces por la proposición 1.2 $\frac{\partial f}{\partial x_i}$ es continua en A, esto es, $f\in C^1(A)$. Este razonamiento se aplica sucesivamente, de modo que si $f\in C^m(A)$ entonces $f\in C^{m-1}(A)$. Como ya se ha indicado, el resultado es extensible a funciones vectoriales.

Nota 1.37. Si $\mathbf{f}$ admite en A derivadas parciales de cualquier orden, se dice que f es de **clase** ∞ en A, y se escribe $\mathbf{f} \in C^{\infty}(A)$.

Nota 1.38. Es frecuente encontrar el **teorema de Schwarz** enunciado con hipótesis más restrictivas, en la forma siguiente.

Proposición 1.8. Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A conjunto abierto, $f \in C^m(A)$. Entonces no cambia el valor de las derivadas parciales en $\mathbf{a} \in A$ al variar el orden de derivación.

Nota 1.39. Es decir, con funciones de clase m en A sólo importa el número de veces que se deriva respecto a cada variable, y no el orden en que se haga hasta el orden m de derivación parcial. P. ej., si $f \in C^3(A)$,

$$\frac{\partial^3 f}{\partial x \partial x \partial y} = \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial y \partial x \partial x} \equiv \frac{\partial^3 f}{\partial^2 x \partial y}.$$

1.5. Notación clásica de la diferencial

Nota 1.40. En ciencias e ingenierías es habitual expresar la diferencial con una notación distinta a la empleada en este tema. P. ej, para una función de tres variables, $f(x, y, z)$, es común encontrar expresiones del tipo

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz.$$

Esta **notación clásica** de la diferencial, es predominante en el ámbito de las ecuaciones diferenciales ordinarias, como veremos en este curso.

Nota 1.41. Para aclarar esta notación, recordemos las **funciones proyección** $i-$ésima estudiadas en el Tema 1 (Bloque 1), definidas como los campos escalares $\pi_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ tal que $\pi_i(\mathbf{x}) = x_i$ $(i = 1, \ldots, n.)$. Estas funciones son diferenciables (polinomios) en todo su dominio. Calculamos

$$\boldsymbol{\nabla} \pi_i = \left(\frac{\partial \pi_i}{\partial x_1}, \ldots, \frac{\partial \pi_i}{\partial x_i}, \ldots, \frac{\partial \pi_i}{\partial x_n} \right) = (0, \ldots, 1, \ldots, 0) = \mathbf{e}_i.$$

Así pues, su diferencial vendrá dada por

$$d\pi_{i\mathbf{a}}(\mathbf{h}) = \boldsymbol{\nabla} \pi_i(\mathbf{a}) \cdot \mathbf{h} = \mathbf{e}_i \cdot \mathbf{h} = h_i$$

donde h_i es la coordenada $i-$ésima de $\mathbf{h}$. Por lo tanto,

$$d\pi_{i\mathbf{a}}(\mathbf{h}) = h_i = \pi_i(\mathbf{h}),$$

esto es, la diferencial de π_i coincide con la propia función (lo que ocurre con las funciones lineales, véase el ejemplo 6) y es independiente del punto $\mathbf{a}$. En tal caso es innecesario emplear el punto en la notación, y ésta se aligera escribiendo

$$d\pi_i = \pi_i. \tag{1.21}$$

Nota 1.42. Como además $\pi_i(\mathbf{x}) = x_i$, es usual escribir $d\pi_i \equiv dx_i$. Con esta notación se tiene que

$$dx_i(\mathbf{h}) = h_i. \tag{1.22}$$

Nota 1.43. Podemos reescribir entonces la expresión de la diferencial de un campo escalar $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ diferenciable en un punto $\mathbf{a} \in \operatorname{Int} A$ (ec. 1.12),

$$df_{\mathbf{a}}(\mathbf{h}) = \nabla f(\mathbf{a}) \cdot \mathbf{h} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})\, h_i,$$

en la forma (igualdad entre aplicaciones lineales):

$$df_{\mathbf{a}} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})\, dx_i = \frac{\partial f}{\partial x_1}(\mathbf{a})\, dx_1 + \ldots + \frac{\partial f}{\partial x_n}(\mathbf{a})\, dx_n. \tag{1.23}$$

Por abuso de notación también es frecuente omitir el punto de cálculo de las derivadas, $\mathbf{a}$, recuperándose así expresiones del tipo señalado en la nota 1.40.

Ejercicios

Ejercicio 1. Sea $f(x, y) = \frac{\alpha(|x|+|y|)}{\sqrt{x^2+y^2}}$, si $(x, y) \neq (0, 0)$; $f(0, 0) = \beta$. Hallar una relación entre α y β para que existan las derivadas parciales de f en el origen.

Sol.: $\alpha = \beta$

Ejercicio 2. Estudiar la continuidad, existencia de derivadas direccionales y diferenciabilidad en el origen de las siguientes funciones

a) $f(x, y) = \dfrac{x^3 y}{x^4 + y^2}$ si $(x, y) \neq (0, 0)$, $f(0, 0) = 0$.

b) $f(x, y) = \sqrt{|xy|}$.

c) $f(x, y) = \dfrac{x^4 - y^4}{x^2 + xy + y^2}$ si $(x, y) \neq (0, 0)$, $f(0, 0) = 0$.

d) $f(x, y) = \left(\sqrt[3]{x} + \sqrt[3]{y}\right)^3$.

e) $f(x, y) = \dfrac{x^2 y}{x^4 + y^2}$ si $(x, y) \neq (0, 0)$, $f(0, 0) = 0$.

f) $f(x, y) = \dfrac{2x^2 y + 5x^3}{x^2 + y^4}$ si $(x, y) \neq (0, 0)$, $f(0, 0) = 0$.

Sol.: a) Continua, existen direccionales, no diferenciable. b) Continua, existen parciales (únicamente), no diferenciable. c) Continua, existen direccionales, diferenciable. d) Continua, existen direccionales, no diferenciable. e) No continua, existen direccionales. f) Continua, existen direccionales, no diferenciable.

Ejercicio 3. Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por

$$f(x, y) = \begin{cases} \frac{x^3 - x^2 y - y^3}{x^2 + y^2} & \text{si} \quad (x, y) \neq (0, 0) \\ 0 & \text{si} \quad (x, y) = (0, 0) \end{cases}.$$

Se pide: a) Estudiar la continuidad de f en todo punto. b) Calcular las derivadas parciales de f en todo punto. c) Calcular $D_{\mathbf{v}} f(0, 0)$ con $\mathbf{v} = (\cos\theta, \operatorname{sen}\theta)$, $\theta \in \mathbb{R}$. d) Como consecuencia de c) deducir que f no es diferenciable en $(0, 0)$.

Sol.: a) Continua en $\mathbb{R}^2$. b) $\frac{\partial f}{\partial x} = \frac{x^4+3x^2y^2}{(x^2+y^2)^2}$ si $(x,y) \neq 0$, $\frac{\partial f}{\partial x}(0,0) = 1$, $\frac{\partial f}{\partial y} = -\frac{x^4+y^4+2x^2y^2+2x^3y}{(x^2+y^2)^2}$ si $(x,y) \neq 0$, $\frac{\partial f}{\partial y}(0,0) = -1$. c) $D_{\mathbf{v}}f(0,0) = \cos^3\theta - \operatorname{sen}\theta$. d) Suponga que f es diferenciable en $(0,0)$ y proceda por reducción al absurdo.

Ejercicio 4. Sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ tal que $|f(\mathbf{x})| \leq M \|\mathbf{x}\|^2$. Demostrar que es diferenciable en $\mathbf{0}$. Aplicar el resultado para $n = 2$ a la siguiente función para estudiar su diferenciabilidad en $(0,0)$,

$$f(x,y) = \begin{cases} Ax^2 \cos\frac{1}{x} + By^2 \operatorname{sen}\frac{1}{y} & x \neq 0,\ y \neq 0 \\ Cx^2 \operatorname{sen}\frac{1}{x} & x \neq 0,\ y = 0 \\ Dy^2 \cos\frac{1}{y} & x = 0,\ y \neq 0 \end{cases}.$$

Ejercicio 5. Hallar la derivada direccional de la función $f(x,y,z) = x^2 - 2xy + z^3$ en el punto $(1,-1,2)$ según la dirección $(-1,3,1)$. Calcular en qué dirección unitaria será máxima y cuál será el valor de ese máximo.

Sol.: a) $D_{(-1,3,1)}f(1,-1,2) = 2$, $\mathbf{u} = (2,-1,6)/\sqrt{41}$, $D_{\mathbf{u}}f(1,-1,2) = 2\sqrt{41}$.

Ejercicio 6. Sea $f(x,y) = e^{ax+by}\cos(x+y)$, con $a, b \in \mathbb{R}$. Se pide: a) Calcular $\boldsymbol{\nabla} f(0,0)$. b) Determinar a y b para que la derivada direccional de f en el punto $(0,0)$, en la dirección unitaria de la bisectriz del primer cuadrante tome un valor máximo de $3\sqrt{2}$.

Sol.: a) (a,b). b) $a = b = 3$.

Ejercicio 7. Determinar los 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 $(1,2,-1)$ y en la dirección $(0,0,1)$ tome un valor máximo de 64.

Sol.: $a = 6$, $b = 24$, $c = -8$.

Ejercicio 8. Hallar la ecuación del plano tangente a la gráfica de las siguientes funciones en el punto que se indica,

a) $f(x,y) = x^2 + y^2$, en el punto $(0,0,0)$.
b) $f(x,y) = \sqrt{x^2 - y^2 + 2xy + 1}$, en el punto $\left(1,1,\sqrt{3}\right)$.
c) $f(x,y) = xe^{xy-2}$, en el punto $(2,1,2)$.

Sol.: a) $z = 0$. b) $2x - \sqrt{3}z + 1 = 0$. c) $z = 3x + 4y - 8$.

Ejercicio 9. Hallar la ecuación del plano tangente la superficie $z = x^2 + 2y^2$ y paralelo al plano $x + 2y - z = 10$.

Sol.: $x + 2y - z - 3/4 = 0$.

Ejercicio 10. Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por

$$f(x,y) = \begin{cases} \cos\left(x^2+y^2\right) \dfrac{1-\cos\left(x^3+y^3\right)}{x^2+y^2} & \text{si} \quad (x,y) \neq (0,0) \\ 0 & \text{si} \quad (x,y) = (0,0) \end{cases}.$$

Se pide: a) Estudiar la continuidad de f en $(0,0)$. b) Estudiar la diferenciabilidad de f en $(0,0)$. c) Calcular $D_{\mathbf{v}} f(0,0)$ con $\mathbf{v} = (-2,1)$. d) Determinar $df_{\left(\sqrt[3]{\pi}, -\sqrt[3]{\pi}\right)}$. e) Obtener el plano tangente a la gráfica de $f(x,y)$ en $\left(\sqrt[3]{\pi}, -\sqrt[3]{\pi}\right)$.

Sol.: a) Continua. b) Diferenciable. c) 0. d) 0 (aplicación lineal nula). e) $z = 0$.

Ejercicio 11. Sea $\mathbf{f} : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ tal que $\mathbf{f}(x,y) = (x, y, y^2)$. Pruebe con la definición que es diferenciable en todo $\mathbb{R}^2$.

Ejercicio 12. Dada $f(x,y) = \frac{1-e^{x(y-2)}}{xy-2x}$, se pide: a) Estudiar la continuidad en $\mathbb{R}^2$ y prolongar por continuidad cuando sea posible. b) Hallar las derivadas parciales de la función prolongada por continuidad.

Sol.: a) La función prolongada por continuidad es: $g(x,y) = f(x,y)$ si $x \neq 0$ e $y \neq 2$; $g(x,y) = -1$ si $x = 0$ o $y = 2$. b) si $x \neq 0$ e $y \neq 2$, $\frac{\partial g}{\partial x} = -\frac{1+e^{x(y-2)}[x(y-2)-1]}{x^2(y-2)}$, $\frac{\partial g}{\partial y} = -\frac{1+e^{x(y-2)}[x(y-2)-1]}{x(y-2)^2}$; si $x = 0$ e $y \neq 2$, $\frac{\partial g}{\partial x} = -\frac{b-2}{2}$, $\frac{\partial g}{\partial y} = 0$; si $x \neq 0$ e $y = 2$, $\frac{\partial g}{\partial x} = 0$, $\frac{\partial g}{\partial y} = -\frac{a}{2}$; si $x = 0$ e $y = 2$, $\frac{\partial g}{\partial x} = 0$, $\frac{\partial g}{\partial y} = 0$.

Ejercicio 13. Calcule la matriz jacobiana de las siguientes funciones en todo punto de su dominio.

a) $\mathbf{f}(x,y,z) = (x+y, z, z-x, 2z)$.
b) $\mathbf{f}(t) = (t, t^2, \operatorname{sen} t)$.
c) $\mathbf{f}(r,\theta,z) = (r\cos\theta, r\operatorname{sen}\theta, z)$

Sol.: a) $J\mathbf{f}(x,y,z) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. b) $J\mathbf{f}(t) = \begin{pmatrix} 1 \\ 2t \\ \cos t \end{pmatrix}$.

c) $\mathbf{f}(r,\theta,z) = \begin{pmatrix} \cos\theta & -r\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Ejercicio 14. Bajo ciertas hipótesis, la temperatura de una varilla aislada en un punto de abscisa x puede describirse en cada instante t mediante la función $\omega(x,t)$ suficientemente regular que verifica la ecuación en derivadas parciales

$$\frac{\partial \omega}{\partial t} = k \frac{\partial^2 \omega}{\partial x^2} \ ,$$

donde $k > 0$ depende de las propiedades del material (ecuación de fusión o ecuación unidimensional del calor). Compruebe que $\omega(x,t) = \frac{1}{2\sqrt{\pi t}} e^{-\frac{x^2}{4t}}$ es solución con $k = 1$.

Ejercicio 15. Calcular las derivadas parciales de primer orden y comprobar la igualdad de las derivadas parciales cruzadas de segundo orden de las siguientes funciones:

a) $f(x,y) = \operatorname{arcsen}(x + y^2)$.

b) $f(x,y,z) = e^{x/y} + e^{z/y}$.

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

d) $f(x,y) = \sqrt[3]{y + \log x}$.

e) $f(x,y) = e^{\operatorname{sen}(y/x)}$.

Sol: a) $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{2y\left(x+y^2\right)}{\left[1-(x+y^2)^2\right]^{3/2}}$.

b) $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -\frac{e^{x/y}}{y^2}\left(1 - \frac{x}{y}\right)$, $\frac{\partial^2 f}{\partial x \partial z} = \frac{\partial^2 f}{\partial z \partial x} = 0$, $\frac{\partial^2 f}{\partial y \partial y} = \frac{\partial^2 f}{\partial z \partial y} = -\frac{e^{z/y}}{y^2}\left(1 - \frac{z}{y}\right)$.

c) $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{2y}{x} x^{y^2}\left(1 + y^2 \log x\right)$.

d) $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -\frac{2}{9x}(y + \log x)$.

e) $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = e^{\operatorname{sen}(y/x)}\left(\frac{y}{x^3} \operatorname{sen} \frac{y}{x} - \frac{y}{x^3} \cos^2 \frac{y}{x} - \frac{1}{x^2} \cos \frac{y}{x}\right)$.

Parte 2

Aplicaciones del cálculo diferencial (I)

2.1. Regla de la cadena

Proposición 2.1 (Regla de la cadena)**.** Sean $\mathbf{f} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$, diferenciable en un punto $\mathbf{a} \in A$, y $\mathbf{g} : B \subseteq \mathbb{R}^m \longrightarrow \mathbb{R}^p$, diferenciable en $\mathbf{f}(\mathbf{a}) \in B$. Entonces $\mathbf{g} \circ \mathbf{f}$ es diferenciable en $\mathbf{a}$ y se tiene

$$d(\mathbf{g} \circ \mathbf{f})_{\mathbf{a}} = d\mathbf{g}_{\mathbf{f}(\mathbf{a})} \circ d\mathbf{f}_{\mathbf{a}}. \tag{2.1}$$

Nota 2.1*.* En términos de matrices jacobianas, la relación (2.1) se escribe como el producto matricial

$$J\mathbf{g} \circ \mathbf{f}(\mathbf{a}) = J\mathbf{g}(\mathbf{f}(\mathbf{a}))\, J\mathbf{f}(\mathbf{a}),$$

o explicitando sus componentes

$$\begin{pmatrix} \frac{\partial(\mathbf{g}\circ\mathbf{f})_1}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial(\mathbf{g}\circ\mathbf{f})_1}{\partial x_n}(\mathbf{a}) \\ \vdots & & \vdots \\ \frac{\partial(\mathbf{g}\circ\mathbf{f})_p}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial(\mathbf{g}\circ\mathbf{f})_p}{\partial x_n}(\mathbf{a}) \end{pmatrix} \tag{2.2}$$

$$= \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(\mathbf{f}(\mathbf{a})) & \cdots & \frac{\partial g_1}{\partial y_m}(\mathbf{f}(\mathbf{a})) \\ \vdots & & \vdots \\ \frac{\partial g_p}{\partial y_1}(\mathbf{f}(\mathbf{a})) & \cdots & \frac{\partial g_p}{\partial y_m}(\mathbf{f}(\mathbf{a})) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix},$$

donde hemos denotado por $(x_1, \ldots, x_n)$ las coordenadas en $A \subseteq \mathbb{R}^n$, y por $(y_1, \ldots, y_m)$ las coordenadas en $B \subseteq \mathbb{R}^m$. El producto de matrices tiene dimensiones $(p \times n) = (p \times m) \times (m \times n)$.

Ejemplo 12. Sea $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$, denotada por $g\,(x, y)$, diferenciable en $\mathbb{R}^2$. Sea $\mathbf{f} : A \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por

$$\mathbf{f}\,(r, \theta) = (r\cos\theta, r\,\mathrm{sen}\,\theta)\,,$$

con $A = \{(r, \theta) \in \mathbb{R}^2 \setminus r > 0,\ \theta \in \,]0, 2\pi[\}$ (conjunto abierto). Consideramos la composición $g \circ \mathbf{f}$,

$$g \circ \mathbf{f} : A \subseteq \underbrace{\overbrace{\mathbb{R}^2}^{(r,\theta)} \longrightarrow}_{\mathbf{f}} \underbrace{\overbrace{\mathbb{R}^2}^{(x,y)} \longrightarrow \mathbb{R}}_{g},$$

que viene dada por $(g \circ \mathbf{f})\,(r, \theta) = g\,(\mathbf{f}\,(r, \theta)) = g\,(r\cos\theta, r\,\mathrm{sen}\,\theta)$ (esto es, expresa g en coordenadas polares). Queremos estudiar $d\,(g \circ \mathbf{f})$ en un punto arbitrario de su dominio. Como $f_1\,(r, \theta) = r\cos\theta$, y $f_2\,(r, \theta) = r\,\mathrm{sen}\,\theta$ son de clase 1 en A, son diferenciables. La función $g\,(x, y)$ es diferenciable por hipótesis. Obtenemos las matrices jacobianas involucradas en la composición:

$$\begin{aligned}
J\mathbf{f}\,(r, \theta) &= \begin{pmatrix} \dfrac{\partial f_1}{\partial r}\,(r, \theta) & \dfrac{\partial f_1}{\partial \theta}\,(r, \theta) \\[2ex] \dfrac{\partial f_2}{\partial r}\,(r, \theta) & \dfrac{\partial f_2}{\partial \theta}\,(r, \theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & r\cos\theta \end{pmatrix}, \\
Jg\,(x, y) &= \begin{pmatrix} \dfrac{\partial g}{\partial x}\,(x, y) & \dfrac{\partial g}{\partial y}\,(x, y) \end{pmatrix}, \\
Jg \circ \mathbf{f}\,(r, \theta) &= \begin{pmatrix} \dfrac{\partial g \circ \mathbf{f}}{\partial r}\,(r, \theta) & \dfrac{\partial g \circ \mathbf{f}}{\partial \theta}\,(r, \theta) \end{pmatrix}.
\end{aligned}$$

Para simplificar la escritura es habitual renombrar la función compuesta, por ejemplo usaremos $\tilde{g}\,(r, \theta) \equiv (g \circ \mathbf{f})\,(r, \theta)$. La regla de la cadena (proposición 2.1) nos dice que $\tilde{g}\,(r, \theta)$ es diferenciable en (r, θ) y nos establece la relación entre matrices jacobianas,

$$\begin{aligned}
J\tilde{g}\,(r, \theta) &= \begin{pmatrix} \dfrac{\partial \tilde{g}}{\partial r}\,(r, \theta) & \dfrac{\partial \tilde{g}}{\partial \theta}\,(r, \theta) \end{pmatrix} = Jg\,(\mathbf{f}\,(r, \theta))\,J\mathbf{f}\,(r, \theta) = \\
&= \left.\begin{pmatrix} \dfrac{\partial g}{\partial x}\,(x, y) & \dfrac{\partial g}{\partial y}\,(x, y) \end{pmatrix}\right|_{(f_1(r,\theta), f_2(r,\theta))} \begin{pmatrix} \cos\theta & -r\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & r\cos\theta \end{pmatrix},
\end{aligned}$$

de donde, se obtienen las relaciones,

$$\begin{aligned}
\frac{\partial \tilde{g}}{\partial r}\,(r, \theta) &= \frac{\partial g}{\partial x}\,(r\cos\theta, r\,\mathrm{sen}\,\theta)\cos\theta + \frac{\partial g}{\partial y}\,(r\cos\theta, r\,\mathrm{sen}\,\theta)\,\mathrm{sen}\,\theta, \qquad (2.3) \\
\frac{\partial \tilde{g}}{\partial \theta}\,(r, \theta) &= -\frac{\partial g}{\partial x}\,(r\cos\theta, r\,\mathrm{sen}\,\theta)\,r\,\mathrm{sen}\,\theta + \frac{\partial g}{\partial y}\,(r\cos\theta, r\,\mathrm{sen}\,\theta)\,r\cos\theta,
\end{aligned}$$

que permiten calcular la diferencial de $g \circ \mathbf{f}$ en un punto cualquiera de su dominio.

Nota 2.2. Es común aligerar la notación omitiendo los puntos de cálculo de las derivadas, e incluso, por abuso de notación, llamar igual a la función compuesta que a la segunda función de la composición, esto es, $\tilde{g} \equiv g$, En general no hay confusión porque las distingue el nombre de las variables respecto a las que se deriva. En tal caso, las expresiones (2.3) se escriben, de forma más simplificada,

$$\begin{aligned} \frac{\partial g}{\partial r}(r,\theta) &= \frac{\partial g}{\partial x}\cos\theta + \frac{\partial g}{\partial y}\operatorname{sen}\theta, \\ \frac{\partial g}{\partial \theta}(r,\theta) &= -\frac{\partial g}{\partial x} r\operatorname{sen}\theta + \frac{\partial g}{\partial y} r\cos\theta, \end{aligned} \tag{2.4}$$

donde las derivadas respecto de r y θ hacen referencia a la función compuesta $\tilde{g}$, y las derivadas respecto de x e y a la función g.

Nota 2.3 (Árbol de dependencias). Con la notación simplificada se pueden obtener las relaciones entre derivadas de funciones compuestas mediante la regla práctica denominada **árbol de dependencias**. Las ecuaciones (2.4) del ejemplo anterior se construyen mediante el esquema de la figura 2.1.

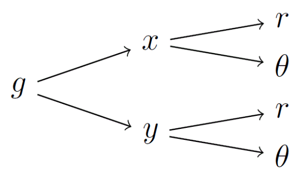


Figura 2.1: Árbol de dependencias del ejemplo 2.2.

La interpretación del árbol de dependencias es la siguiente: La derivada parcial de la función del primer nodo respecto de una de las variables de los nodos finales se obtiene formando para cada camino que las conecta el producto de las derivadas de cada función respecto de la variable que inmediatamente le sigue en el camino, y sumando todos los productos obtenidos.

Ejemplo 13. Dada $f\left(x^2+y^2\right)$ con f diferenciable en todo punto, calculamos la expresión

$$y\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial y}. \tag{2.5}$$

En términos de composición de funciones, la situación es la siguiente,

$$f \circ g : I \subseteq \underbrace{\overbrace{\mathbb{R}^2}^{(x,y)} \longrightarrow \overbrace{\mathbb{R}}^{z}}_{g} \underbrace{\phantom{\mathbb{R}} \longrightarrow \mathbb{R}}_{f},$$

esto es,

$$(f \circ g)(x, y) = f(g(x, y)) = f\left(x^2 + y^2\right) \equiv \tilde{f}(x, y).$$

Dado que $g(x, y) = x^2 + y^2$ es diferenciable en todo punto, la aplicación de la regla de la cadena (proposición 2.1) nos dice que $f \circ g$ es diferenciable en (x, y) y nos establece la relación entre las diferenciales,

$$J\tilde{f}(x, y) = Jf(g(x, y))\, Jg(x, y).$$

Las matrices jacobianas son,

$$\begin{aligned} Jg(x, y) &= \begin{pmatrix} \dfrac{\partial g}{\partial x}(x, y) & \dfrac{\partial g}{\partial y}(x, y) \end{pmatrix}, \\ Jf(z) &= \frac{df}{dz}, \\ J\tilde{f}(x, y) &= \begin{pmatrix} \dfrac{\partial \tilde{f}}{\partial x}(x, y) & \dfrac{\partial \tilde{f}}{\partial y}(x, y) \end{pmatrix}. \end{aligned}$$

Por lo tanto, se tiene que

$$\begin{aligned} \begin{pmatrix} \dfrac{\partial \tilde{f}}{\partial x}(x, y) & \dfrac{\partial \tilde{f}}{\partial y}(x, y) \end{pmatrix} &= \frac{df}{dz}(g(x, y)) \begin{pmatrix} \dfrac{\partial g}{\partial x}(x, y) & \dfrac{\partial g}{\partial y}(x, y) \end{pmatrix} \\ &= \frac{df}{dz}\left(x^2 + y^2\right) \begin{pmatrix} 2x & 2y \end{pmatrix} \\ &= \left(2x\frac{df}{dz}\left(x^2 + y^2\right), 2y\frac{df}{dz}\left(x^2 + y^2\right)\right). \end{aligned}$$

Para calcular la expresión (2.5) hemos de tener en cuenta que las derivadas de f son parciales, respecto a x e y, lo que denota que se trata de la función compuesta $\tilde{f}(x, y)$, dado que f es función de una única variable. Entonces

$$\begin{aligned} y\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial y} &= y\frac{\partial \tilde{f}}{\partial x}(x, y) - x\frac{\partial \tilde{f}}{\partial y}(x, y) \\ &= 2xy\frac{df}{dz}\left(x^2 + y^2\right) - 2xy\frac{df}{dz}\left(x^2 + y^2\right) = 0. \end{aligned}$$

$$
\begin{tikzcd}[row sep=tiny]
 & & x \\
f \arrow[r] & z \arrow[ur] \arrow[dr] & \\
 & & y
\end{tikzcd}
$$

Figura 2.2: Árbol de dependencias del ejemplo 13.

Nota 2.4. El árbol de dependencias en el ejemplo anterior es el de la figura 2.2,. que equivale, en notación simplificada, a las relaciones

$$\frac{\partial f}{\partial x} = \frac{df}{dz}\frac{\partial z}{\partial x}, \quad \frac{\partial f}{\partial y} = \frac{df}{dz}\frac{\partial z}{\partial y}.$$

Tales expresiones son totalmente operativas, si se sustituye en ellas $z = x^2+y^2$ se recupera el resultado obtenido previamente con el cálculo matricial.

2.2. Fórmula de Taylor

2.2.1. Polinomio y fórmula de Taylor

Nota 2.5. En el caso de funciones rales de variable real, $f : A \subseteq \mathbb{R} \longrightarrow \mathbb{R}$, $a \in \operatorname{Int} A$, con f $m+1$ veces derivable en $]a-k, a+k[\subseteq A$ $(k > 0)$, la **fórmula de Taylor** establece que si $0 < |h| < k$, existe $\theta \in]0,1[$ tal que

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \frac{1}{2!}f''(a)h^2 + \ldots + \frac{1}{m!}f^{(m)}(a)h^2 \\ &\quad + \frac{1}{(m+1)!}f^{(m+1)}(a+\theta h)h^{m+1}. \end{aligned} \tag{2.6}$$

La primera fila de (2.6) es el llamado *polinomio de Taylor* de grado m en a y la segunda es el *resto de Lagrange*. La siguiente proposición generaliza este resultado para campos escalares. Por el momento eludiremos la forma explícita del resto de Lagrange, lo que simplifica el enunciado de la proposición.

Proposición 2.2 (Fórmula de Taylor)**.** Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A abierto, $f \in C^m(A)$. Sean $\mathbf{a}$ y $\mathbf{a}+\mathbf{h}$, tales que $\bar{L}(\mathbf{a}, \mathbf{a}+\mathbf{h}) \subseteq A$. Entonces existe una función $\varepsilon : \mathbb{R}^n \longrightarrow \mathbb{R}$ con $\lim_{\mathbf{h}\to\mathbf{0}} \varepsilon(\mathbf{h}) = 0$ tal que

$$f(\mathbf{a}+\mathbf{h}) = P_{m,\mathbf{a}}(\mathbf{a}+\mathbf{h}) + \varepsilon(\mathbf{h}) \|\mathbf{h}\|^m \tag{2.7}$$

donde

$$\begin{aligned}P_{m,\mathbf{a}}(\mathbf{a}+\mathbf{h}) &= f(\mathbf{a})+\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(\mathbf{a})\,h_i+\frac{1}{2!}\sum_{i_1,i_2=1}^{n}\frac{\partial^2 f}{\partial x_{i_1}\partial x_{i_2}}(\mathbf{a})\,h_{i_1}h_{i_2}+\dots\\ &\quad+\frac{1}{m!}\sum_{i_1\dots i_m=1}^{n}\frac{\partial^m f}{\partial x_{i_1}\dots\partial x_{i_m}}(\mathbf{a})\,h_{i_1}\dots h_{i_m}\end{aligned} \tag{2.8}$$

se denomina **Polinomio de Taylor** de grado m en $\mathbf{a}$.

Nota 2.6. Dado que el polinomio de Taylor de grado m no tiene por qué ser un polinomio de grado m (por la anulación de los coeficientes de mayor grado), es frecuente la denominación polinomio de Taylor de *orden* m.

Ejemplo 14. Sean $f(x,y)=y^2e^{\operatorname{sen}x}+x^2/y$ y $\mathbf{a}=(0,1)$. Obtenemos $P_{2,\mathbf{a}}(\mathbf{x})$, con $\mathbf{x}=\mathbf{a}+\mathbf{h}$. En este caso, $\operatorname{dom}f$ es $\mathbb{R}^2$ salvo los puntos con $y=0$. La función f es de clase ∞ en su dominio. La forma del polinomio de Taylor es

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

$$+\frac{1}{2!}\left[\frac{\partial^2 f}{\partial x^2}(0,1)(x-0)^2+2\frac{\partial^2 f}{\partial x\partial y}(0,1)(x-0)(y-1)+\frac{\partial^2 f}{\partial y^2}(0,1)(y-1)^2\right].$$

Tenemos $f(0,1)=1$, y calculamos las derivadas parciales de primer orden,

$$\begin{aligned}\frac{\partial f}{\partial x} &= y^2e^{\operatorname{sen}x}\cos x+\frac{2x}{y}\to\frac{\partial f}{\partial x}(0,1)=1,\\ \frac{\partial f}{\partial y} &= 2ye^{\operatorname{sen}x}-\frac{x^2}{y^2}\to\frac{\partial f}{\partial y}(0,1)=2,\end{aligned}$$

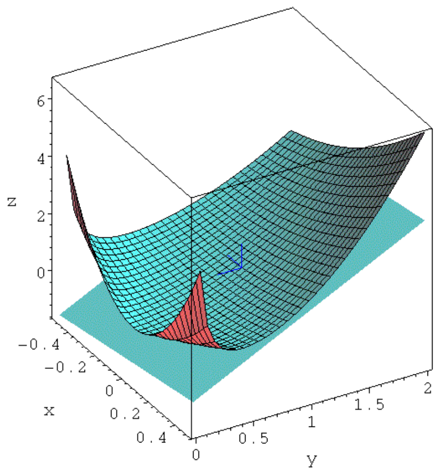
y de segundo,

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)=y^2e^{\operatorname{sen}x}\left(\cos^2 x-\operatorname{sen}x\right)+\frac{2}{y}\to\frac{\partial^2 f}{\partial x^2}(0,1)=3,\\ \frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)=2e^{\operatorname{sen}x}+\frac{2x^2}{y^3}\to\frac{\partial^2 f}{\partial y^2}(0,1)=2,\\ \frac{\partial^2 f}{\partial x\partial y} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)=2ye^{\operatorname{sen}x}\cos x-\frac{2x}{y^2}\to\frac{\partial^2 f}{\partial x\partial y}(0,1)=2.\end{aligned}$$

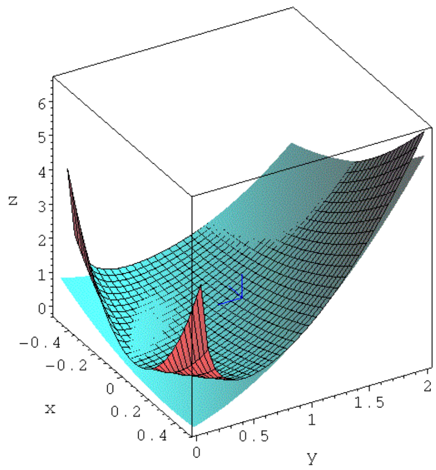
Por lo tanto,

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

Evidentemente, $P_{1,(0,1)}(x,y)=1+x+2(y-1)$. Las siguientes figuras muestran la superficie $\operatorname{graf}f$ y los grafos correspondientes a los polinomios $P_{1,(0,1)}$ y $P_{2,(0,1)}$.



Gráficas de f y del polinomio de Taylor de f de grado 1 en el punto $(0, 1)$ (plano tangente).



Gráficas de f y del polinomio de Taylor de f de grado 2 en el punto $(0, 1)$.

Nota 2.7. Teniendo en cuenta el expresión general 2.8, nótese que, si $\mathbf{x} = \mathbf{a} + \mathbf{h}$, el polinomio de grado 1 es

$$P_{1,\mathbf{a}}(\mathbf{x}) = f(\mathbf{a}) + \boldsymbol{\nabla} f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) = f(\mathbf{a}) + df_{\mathbf{a}}(\mathbf{x} - \mathbf{a}), \qquad (2.9)$$

esto es, viene definido por la diferencial de la función f en el punto $\mathbf{a}$, y es equivalente a la ecuación (1.13) del hiperplano tangente a graf f en el punto $(\mathbf{a}, f(\mathbf{a}))$. Esto ya ocurre con el caso del polinomio de Taylor en funciones de una variable, donde $P_{m,a}(x) = f(a) + f'(a)(x - a)$ y por lo tanto viene definido por $df_a(x - a) = f'(a)(x - a)$ (nota 1.8) y coincide con la ecuación de la recta tangente a graf f en el punto $(a, f(a))$.

2.2.2. Diferenciales de orden superior

Nota 2.8. Es posible generalizar la forma (2.9) de escribir el polinomio de Taylor para grados superiores, y por lo tanto órdenes superiores de derivación parcial. Ello induce la siguiente definición.

Definición 2.1 (Diferenciales de orden superior)**.** Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A abierto, $f \in C^k(A)$ y $\mathbf{a} \in A$. Se denomina **diferencial de orden** k de f en A a la aplicación $d^k f_{\mathbf{a}} : \mathbb{R}^n \longrightarrow \mathbb{R}$ tal que

$$d^k f_{\mathbf{a}}(\mathbf{h}) = \sum_{i_1 \ldots i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(\mathbf{a})\, h_{i_1} \ldots h_{i_k}. \qquad (2.10)$$

Nota 2.9. Esto es, una combinación lineal de las derivadas parciales de orden k, que es lo que aparece en el polinomio de Taylor. La aplicación $d^k f_{\mathbf{a}}(\mathbf{h})$ es también un polinomio en las variables h_1, h_2,... h_n en el que los monomios tienen grado k. Evidentemente, $d^1 f_{\mathbf{a}}(\mathbf{h}) = df_{\mathbf{a}}(\mathbf{h})$.

Nota 2.10 (Cálculo de diferenciales de orden k). El cálculo de $d^k f_{\mathbf{a}}(\mathbf{h})$ implica determinar todas las combinaciones lineales con derivadas parciales de orden k. Dado que trabajamos con funciones de clase k, en virtud de la proposición (1.8) (teorema de Schwarz), no importa el orden de las variables en la derivación, lo que permite reducir el número de sumandos. Una regla práctica para el cálculo de las diferenciales de orden k es la denominada *potencia simbólica*, que denotamos como $[k]$, y consiste en tratar el orden de derivación como si fuese una potencia, entendiendo que

$$\begin{aligned} \left(\frac{\partial f}{\partial x_1}\right)^{[i_1]} \dots \left(\frac{\partial f}{\partial x_n}\right)^{[i_n]} &= \frac{\partial^k f}{\partial x_i^{i_1} \dots \partial x_n^{i_n}} \\ h_1^{[i_1]} \dots h_n^{[i_n]} &= h_1^{i_1} \dots h_n^{i_n}\ , \text{con } i_1 + \dots + i_n = k. \end{aligned}$$

Ejemplo 15. Aplicamos la potencia simbólica [2] para calcular $d^2 f_{\mathbf{x}}(\mathbf{h})$ con $f(x, y, z)$ de clase 2 (omitimos por simplicidad el punto de cálculo de las derivadas),

$$\begin{aligned} d^2 f_{\mathbf{x}}(\mathbf{h}) &= [df_{\mathbf{x}}(\mathbf{h})]^{[2]} = \left(\frac{\partial f}{\partial x}h_1 + \frac{\partial f}{\partial y}h_2 + \frac{\partial f}{\partial z}h_3\right)^{[2]} = \\ &= \frac{\partial^2 f}{\partial x^2}h_1^2 + \frac{\partial^2 f}{\partial y^2}h_2^2 + \frac{\partial^2 f}{\partial z^2}h_2^2 + \\ &\quad +2\frac{\partial^2 f}{\partial x \partial y}h_1 h_2 + 2\frac{\partial^2 f}{\partial x \partial z}h_1 h_3 + 2\frac{\partial^2 f}{\partial y \partial z}h_2 h_3. \end{aligned}$$

Nota 2.11. Mediante la definición de $d^k f_{\mathbf{a}}$ el polinomio de Taylor (2.8), $P_{m,\mathbf{a}}(\mathbf{x})$, se escribe en forma compacta según la siguiente proposición, que también incluye la forma explícita del **resto de Lagrange**, $R_{m,\mathbf{a}}(\mathbf{x})$.

2.2.3. Fórmula de Taylor con resto de Lagrange

Proposición 2.3 (Fórmula de Taylor con resto de Lagrange)**.** Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A abierto, $f \in C^{m+1}(A)$. Sean $\mathbf{a}$ y $\mathbf{x}$, tales que $\bar{L}(\mathbf{a}, \mathbf{x}) \subseteq A$. Entonces existe un punto $\mathbf{c} \in L(\mathbf{a}, \mathbf{x})$ tal que

$$f(\mathbf{x}) = P_{m,\mathbf{a}}(\mathbf{x}) + R_{m,\mathbf{a}}(\mathbf{x})$$

donde

$$
\begin{aligned}
P_{m,\mathbf{a}}(\mathbf{x}) &= f(\mathbf{a}) + \sum_{i=1}^{m} \frac{1}{k!} d^k f_{\mathbf{a}}(\mathbf{x}-\mathbf{a}), \\
R_{m,\mathbf{a}}(\mathbf{x}) &= \frac{1}{m+1} d^{m+1} f_{\mathbf{c}}(\mathbf{x}-\mathbf{a}). \qquad (2.11)
\end{aligned}
$$

Nota 2.12. La combinación de los enunciados (2.2) y (2.3) permite escribir que

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{R_{m,\mathbf{a}}(\mathbf{x})}{\|\mathbf{x}-\mathbf{a}\|^m} = \lim_{\mathbf{x}\to\mathbf{a}} \varepsilon(\mathbf{x}-\mathbf{a}) = 0. \qquad (2.12)$$

Por lo tanto el polinomio de Taylor mejora la aproximación $f(\mathbf{x}) \simeq P_{m,\mathbf{a}}(\mathbf{x})$ cuanto más nos acercamos a $\mathbf{a}$.

Nota 2.13. Dado que en general el punto $\mathbf{c} \in L(\mathbf{a}, \mathbf{x})$ es desconocido, también lo es valor de $R_{m,\mathbf{a}}(\mathbf{x})$ y sólo es posible acotar $|R_{m,\mathbf{a}}(\mathbf{x})|$ superiormente para controlar el error que se comete al aproximar $f(\mathbf{x})$ por un polinomio de Taylor de cierto grado. Se puede demotrar que, en las condiciones de la proposición 2.3, si $d^{m+1} f_{\mathbf{x}}$ está acotada en A, es decir, $|d^{m+1} f_{\mathbf{x}}(\mathbf{x}-\mathbf{a})| \leq \lambda$ para todo $\mathbf{x} \in A$, se tiene que

$$|R_{m,\mathbf{a}}(\mathbf{x})| \leq \lambda \frac{\|\mathbf{x}-\mathbf{a}\|^{m+1}}{(m+1)!}. \qquad (2.13)$$

2.2.4. Expresiones matriciales

Nota 2.14. Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A abierto, $f \in C^1(A)$. Sean $\mathbf{a}$ y $\mathbf{x}$, tales que $\bar{L}(\mathbf{a}, \mathbf{x}) \subseteq A$. El polinomio de Taylor de grado 1,

$$P_{1,\mathbf{a}}(\mathbf{x}) = f(\mathbf{a}) + df_{\mathbf{a}}(\mathbf{x}-\mathbf{a}), \qquad (2.14)$$

admite representación matricial mediante la matriz jacobiana,

$$P_{1,\mathbf{a}}(\mathbf{x}) = f(\mathbf{a}) + Jf(\mathbf{a})(\mathbf{x}-\mathbf{a}) \qquad (2.15)$$

Donde recordamos que $\mathbf{x}-\mathbf{a}$ representa tanto el vector como la matriz columna asociada en la base canónica.

Nota 2.15. Si $f \in C^2(A)$, el polinomio de Taylor de grado 2,

$$P_{2,\mathbf{a}}(\mathbf{x}) = f(\mathbf{a}) + df_{\mathbf{a}}(\mathbf{x}-\mathbf{a}) + \frac{1}{2} d^2 f_{\mathbf{a}}(\mathbf{x}-\mathbf{a}), \qquad (2.16)$$

también puede representarse matricialmente mediante la siguiente definición.

Definición 2.2 (Matriz hessiana)**.** Sean $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A abierto, $f \in C^2(A)$, y $\mathbf{a} \in A$. Se denomina **matriz hessiana** de f en $\mathbf{a}$, y se representa por $Hf\left(\mathbf{a}\right)$, a la matriz simétrica

$$Hf\left(\mathbf{a}\right) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}\left(\mathbf{a}\right) & \dfrac{\partial^2 f}{\partial x_1 \partial x_2}\left(\mathbf{a}\right) & \dots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}\left(\mathbf{a}\right) \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1}\left(\mathbf{a}\right) & \dfrac{\partial^2 f}{\partial x_2^2}\left(\mathbf{a}\right) & \dots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n}\left(\mathbf{a}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}\left(\mathbf{a}\right) & \dfrac{\partial^2 f}{\partial x_n \partial x_2}\left(\mathbf{a}\right) & \dots & \dfrac{\partial^2 f}{\partial x_n^2}\left(\mathbf{a}\right) \end{pmatrix}. \tag{2.17}$$

Nota 2.16. Utilizando la matriz hessiana, la diferencial de segundo orden se escribe como

$$d^2 f_{\mathbf{a}}\left(\mathbf{h}\right) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}\left(\mathbf{a}\right) h_i h_j = \left(h_1, \dots, h_n\right) Hf\left(\mathbf{a}\right) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \mathbf{h}^t Hf\left(\mathbf{a}\right)\mathbf{h}, \tag{2.18}$$

donde el superíndice t representa trasposición. Por lo tanto, el polinomio de Taylor de grado 2 se escribe,

$$P_{2,\mathbf{a}}\left(\mathbf{x}\right) = f\left(\mathbf{a}\right) + Jf\left(\mathbf{a}\right)\left(\mathbf{x} - \mathbf{a}\right) + \frac{1}{2}\left(\mathbf{x} - \mathbf{a}\right)^t Hf\left(\mathbf{a}\right)\left(\mathbf{x} - \mathbf{a}\right). \tag{2.19}$$

Ejemplo 16. Dada la función $f\left(x, y, z\right) = xyz$, con dominio $\mathbb{R}^3$ y de clase ∞ en el dominio, obtenemos la matriz hessiana y la diferencial segunda en cualquier punto (omitimos el punto de cálculo en las derivadas):

$$Hf\left(\mathbf{a}\right) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial x \partial z} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial^2 f}{\partial y \partial z} \\ \dfrac{\partial^2 f}{\partial z \partial x} & \dfrac{\partial^2 f}{\partial z \partial y} & \dfrac{\partial^2 f}{\partial z^2} \end{pmatrix} = \begin{pmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{pmatrix}. \tag{2.20}$$

Por lo tanto,

$$\begin{aligned} d^2 f_{\mathbf{x}}(\mathbf{h}) &= \mathbf{h}^t H f(\mathbf{x})\,\mathbf{h} = \begin{pmatrix} h_1 & h_2 & h_3 \end{pmatrix} \begin{pmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix} = \\ &= 2xh_2h_3 + 2yh_1h_3 + 2zh_1h_2. \end{aligned} \tag{2.21}$$

Puede comprobarse que es el mismo resultado que se obtendría aplicando la potencia simbólica del ejemplo 15.

Ejercicios

Ejercicio 1. Sean $\mathbf{f}(x,y) = (x^2+y, xy-1)$ y $g(x,y) = \sqrt{1+|x|\,y^2}$. Estudiar la diferenciabilidad de $g\circ\mathbf{f}$ en $(-1,-1)$ y hallar $dg\circ\mathbf{f}_{\,(-1,-1)}$.

Sol.: $g\circ\mathbf{f}$ es diferenciable en $(-1,-1)$. $dg\circ\mathbf{f}_{\,(-1,-1)} = \begin{pmatrix} 0 & 0 \end{pmatrix}$.

Ejercicio 2. Calcular $d\mathbf{g}\circ\mathbf{f}_{(0,0)}$ en los casos siguientes:
a) $\mathbf{f}(x,y) = (y+\cos x, x+e^y)$, $g(u,v) = u+v$.
b) $\mathbf{f}(x,y) = (e^{3x+y}, y-x+1, x^2+3)$, $\mathbf{g}(u,v,w) = (\operatorname{sen}(u+w), v^u)$.

Sol.: a) $dg\circ\mathbf{f}_{(0,0)} = \begin{pmatrix} 1 & 2 \end{pmatrix}$. b) $d\mathbf{g}\circ\mathbf{f}_{(0,0)} = \begin{pmatrix} 3\cos 4 & \cos 4 \\ -1 & 1 \end{pmatrix}$.

Ejercicio 3. Probar que $\omega = f(u,v)$ con $u = x+at$, $v = y+bt$ (a y b constantes) satisface

$$\frac{\partial\omega}{\partial t} = a\frac{\partial\omega}{\partial x} + b\frac{\partial\omega}{\partial y}.$$

Ejercicio 4. Obtener du/dt donde $u = xyz$, con $x = t^2+1$, $y = \log t$, $z = \tan t$, mediante la regla de la cadena de funciones de varias variables. Comprobar el resultado mediante derivación ordinaria en una variable de la función compuesta u.

Sol.: $\frac{du}{dt} = 2t\log t\tan t + \frac{t^2+1}{t}\tan t + \left(t^2+1\right)\log t\frac{1}{\cos^2 t}$.

Ejercicio 5. Sea $f:\mathbb{R}\longrightarrow\mathbb{R}$ una función derivable, y sea $z = f(x^2 y)$. Se pide comprobar que

$$x\frac{\partial z}{\partial x} = 2y\frac{\partial z}{\partial y}.$$

Ejercicio 6. Obtener el polinomio de Taylor de grado m en el punto $\mathbf{a}$, $P_{m,\mathbf{a}}(\mathbf{x})$, de las siguientes funciones:

a) $f(x,y) = y^2 e^{\operatorname{sen} x} + x^2/y$, $m = 2$, $\mathbf{a} = (0,1)$.
b) $f(x,y) = x^y$, $m = 2$, $\mathbf{a} = (1,2)$.
c) $f(x,y) = e^x + y^3$, $m = 3$, $\mathbf{a} = (0,0)$.
d) $f(x,y) = \log(x/y)$, $m = 3$, $\mathbf{a} = (1,1)$.
Utilice el apartado b) para dar una aproximación cuadrática de $(0.95)^{2.01}$.

Sol.: a) $P_{2,(0,1)}(x,y) = 1 + x + 2(y-1) + \frac{3}{2}x^2 + 2x(y-1) + (y-1)^2$.
b) $P_{2,(1,2)}(x,y) = 1 + 2(x-1) + (x-1)^2 + (x-1)(y-2)$.
c) $P_{3,(0,0)}(x,y) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}\left(x^3 + 6y^3\right)$.
d) $P_{3,(x,y)}(x,y) = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - (y-1) + \frac{(y-1)^2}{2} - \frac{(y-1)^3}{3}$.
$(0.95)^{2.01} \simeq P_{2,(1,2)}(0.95, 2.01) = 0.9020$.

Ejercicio 7. Expresar los polinomios $P(x,y) = x^2 + xy + y^2$, $Q(x,y) = x^3 + y^2 + xy^2$ en potencias de $(x-1)$ y de $(y-2)$.

Sol.: $P(x,y) = 7 + 4(x-1) + 5(y-2) + (x-1)^2 + (x-1)(y-2) + (y-2)^2$, $Q(x,y) = 9 + 7(x-1) + 8(y-2) + 3(x-1)^2 + 2(y-2)^2 + 4(y-2)(x-1) + (x-1)^3 + (x-1)(y-2)^2$.

Ejercicio 8. Obtener el polinomio de McLaurin de grado m de $f(x,y) = x \operatorname{sen} y + y \operatorname{sen} x$.

Sol.: $P_{m,(0,0)}(x,y) = 2xy - \frac{x^3y + xy^3}{3!} + \frac{x^5y + xy^5}{5!} - \ldots + (-1)^m \frac{x^{2m+1}y + xy^{2m+1}}{(2m+1)!}$.

Parte 3

Aplicaciones del cálculo diferencial (II)

3.1. Extremos relativos libres

Definición 3.1. Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$. Se dice que $\mathbf{a} \in A$ es un **máximo relativo** de f si existe un entorno $B(\mathbf{a}, \delta)$ tal que $f(\mathbf{a}) \geq f(\mathbf{x})$, para todo $\mathbf{x} \in B(\mathbf{a}, \delta) \cap A$. Análogamente, $\mathbf{a} \in A$ es un **mínimo relativo** de f si existe un entorno $B(\mathbf{a}, \delta)$ tal que $f(\mathbf{a}) \leq f(\mathbf{x})$, para todo $\mathbf{x} \in B(\mathbf{a}, \delta) \cap A$. A los máximos o mínimos relativos los denominaremos **extremos relativos**, **locales** o **libres**.

Nota 3.1. En rigor, un máximo o mínimo relativo de f hace referencia al valor $f(\mathbf{a})$ y no al punto $\mathbf{a}$, siendo más correcto decir que f *alcanza* en $\mathbf{a}$ un extremo. No obstante es común la licencia de referirse al propio punto como el extremo, tal y como se sigue en estas notas.

Definición 3.2. Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A conjunto abierto y f diferenciable en A. Se dice que $\mathbf{a} \in A$ es un **punto crítico** o **estacionario** de f si $\boldsymbol{\nabla} f(\mathbf{a}) = \mathbf{0}$. Los puntos estacionarios que no son extremos locales se denominan **puntos de silla.**

Nota 3.2. Los puntos estacionarios son las soluciones del sistema de ecuaciones

$$\frac{\partial f}{\partial x_1}(\mathbf{a}) = 0,\ \frac{\partial f}{\partial x_2}(\mathbf{a}) = 0,\ ...,\frac{\partial f}{\partial x_n}(\mathbf{a}) = 0. \tag{3.1}$$

Como exigimos en la definición que f sea diferenciable en el abierto A, en todos los puntos de A existen las derivadas parciales primeras, y los puntos estacionarios son los únicos en los que f puede alcanzar un extremo local, como se verá a continuación. Algunos textos emplean definiciones menos

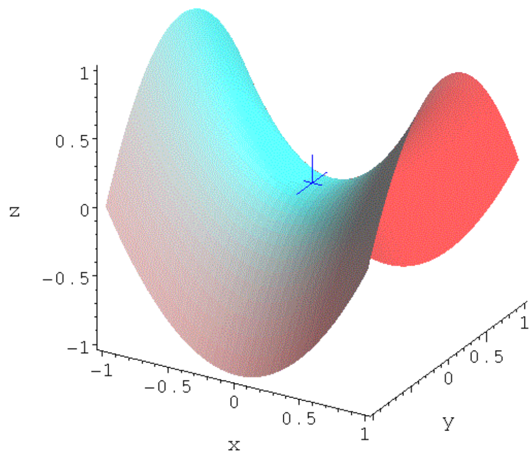


Figura 3.1: Gráfica de la función $f(x,y) = x^2 - y^2$, que presenta un punto de silla en el origen (señalado con la cruz azul).

restrictivas, exigiendo solamente la existencia de derivadas parciales primeras en el punto.

Ejemplo 17. Los puntos estacionarios de la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$

$$f(x,y) = \left(x^2 + y^2\right) e^{-x^2-y^2}$$

se obtienen del sistema

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= 2xe^{-x^2-y^2}\left(1 - x^2 - y^2\right) = 0, \\ \frac{\partial f}{\partial y}(x,y) &= 2ye^{-x^2-y^2}\left(1 - x^2 - y^2\right) = 0. \end{aligned}$$

Por lo tanto son estacionarios el punto $(0,0)$ y aquéllos que verifican $x^2+y^2 = 1$. El punto $(0,0)$ es un mínimo relativo ya que $f(0,0) = 0$ y $f(x,y) \geq 0$. Se puede ver que los puntos de la circunferencia $x^2 + y^2 = 1$ son máximos relativos con el cambio de variable $w = x^2+y^2$ que permite definir la función real de una variable $\tilde{f}(w) = we^{-w}$ para la que se comprueba que $\tilde{f}'(w) = 0 \rightarrow w = 1$, $\tilde{f}''(1) < 0$ (véase la figura 3.2).

Proposición 3.1 (Condición necesaria de extremo relativo)**.** Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A conjunto abierto y f diferenciable en A. Si $\mathbf{a} \in A$ es un extremo relativo de f, entonces $\mathbf{a}$ es punto estacionario de f.

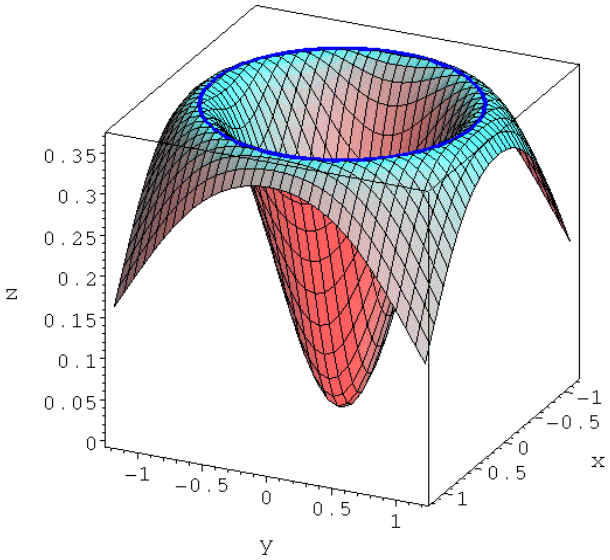


Figura 3.2: Gráfica de la función del ejemplo 17. La circunferencia de máximos relativos se ha marcado en azul.

Nota 3.3. Recordando la ecuación (1.13) del hiperplano tangente, en un extremo relativo en el que la función sea diferenciable, el hiperplano tangente es $z = f(\mathbf{a})$ (en el caso de dos variables, un plano horizontal).

Nota 3.4. El recíproco del teorema no es cierto. Por ejemplo, la función

$$f(x, y) = \left(y - x^2\right)\left(y - 2x^2\right)$$

es diferenciable en $(0, 0)$ y tiene un punto de silla en $(0, 0)$. Esto se puede ver considerando que $f(0, 0) = 0$ y en todo entorno del punto $(0, 0)$ hay puntos que están por encima o por debajo de las parábolas $y = x^2$, $y = 2x^2$, donde la función es positiva por ser producto de dos factores positivos o negativos, y puntos entre ambas parábolas (zona sombreada en la figura 3.3), donde la función es negativa por ser producto de un factor positivo y otro negativo. Por lo tanto $(0, 0)$ no puede ser un extremo relativo.

Nota 3.5. Los puntos de no diferenciabilidad no están contemplados en la proposición. En estos puntos se podría alcanzar un extremo relativo. P. ej. la función

$$f(\mathbf{x}) = |x_1| + |x_2| + \ldots + |x_n|$$

tiene un mínimo relativo en $\mathbf{0}$, ya que $f(\mathbf{0}) = 0$, $f(\mathbf{x}) \geq 0$, pero no existen las derivadas parciales en este punto.

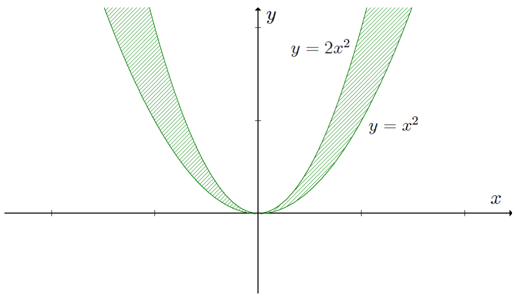


Figura 3.3: Regiones de la nota 3.4.

Nota 3.6. Para las funciones de clase 2 es posible dar un criterio para decidir en qué puntos estacionarios (candidatos) se alcanza un extremo relativo.

Proposición 3.2 (Condición suficiente de extremo relativo)**.** Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A conjunto abierto y $f \in C^2(A)$. Sea $\mathbf{a} \in A$ un punto estacionario de f y $Hf(\mathbf{a})$ la matriz hessiana de f en $\mathbf{a}$. Entonces se verifica:

1. Si $Hf(\mathbf{a})$ es una matriz definida positiva, $\mathbf{a}$ es un mínimo relativo.
2. Si $Hf(\mathbf{a})$ es una matriz definida negativa, $\mathbf{a}$ es un máximo relativo.
3. Si $Hf(\mathbf{a})$ es una matriz indefinida, $\mathbf{a}$ es un punto de silla.

Nota 3.7. Este resultado es consecuencia de la aproximación local de f en el punto $\mathbf{a}$ mediante el polinomio de Taylor de segundo grado. Si en $\mathbf{a}$ las derivadas primeras son nulas, con $\|\mathbf{h}\|$ suficientemente pequeño, se tiene que

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) \simeq \frac{1}{2} d^2 f_{\mathbf{a}}(\mathbf{h}) = \frac{1}{2}(h_1, \ldots, h_n) Hf(\mathbf{a}) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}. \tag{3.2}$$

Esta expresión permite evaluar el signo de $f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})$ con $\mathbf{h}$ en un entorno de $\mathbf{0}$, o equivalentemente, si existirá un extremo local de f en $\mathbf{a}$.

Nota 3.8. Las expresiones de la foma matricial de la diferencial de segundo orden, miembro derecho de (3.2), se denominan **formas cuadráticas**, esto es, $X^t M X$ donde X es una matriz columna y M una matriz **simétrica** (matriz asociada a la forma cuadrática). El estudio del signo de la forma cuadrática viene determinado por el carácter definido positivo, definido negativo o indefinido de la matriz M, en nuestro caso $Hf(\mathbf{a})$. Con este propósito son de aplicación los siguientes criterios:

- *Criterio de autovalores*: Sea M una matriz real simétrica. Se verifica que M es
  1. Definida positiva si, y sólo si, sus autovalores son mayores que cero.
  2. Definida negativa si, y sólo si, sus autovalores son menores que cero.
  3. Indefinida si existen autovalores mayores y menores que cero.
- *Criterio de Sylvester*: Sea M una matriz real simétrica de dimensión n, y sea $\triangle_k$ el menor principal de orden k de la matriz M (determinante de orden k formado por las k primeras filas y las k primeras columnas de M). Entonces M es
  1. Definida positiva si, y sólo si, $\triangle_k > 0$, $\forall k = 1, \ldots, n$.
  2. Definida negativa si, y sólo si, $(-1)^k \triangle_k > 0$, $\forall k = 1, \ldots, n$.
  3. Indefinida si no se verifican 1 y 2 y $\det M \neq 0$.

Ejemplo 18. Determinamos los extremos relativos de

$$f(x, y) = x^3 + y^3 - 3x - 12y + 25.$$

La función es polinómica, definida en $\mathbb{R}^2$ y de clase ∞. los puntos estacionarios se obtienen de

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= 3x^2 - 3 = 0 \to x = \pm 1, \\ \frac{\partial f}{\partial y}(x, y) &= 3y^2 - 12 = 0 \to y = \pm 2, \end{aligned}$$

y son $(1, 2)$, $(1, -2)$, $(-1, 2)$ y $(-1, -2)$. La matriz hessiana resulta

$$Hf(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}.$$

Dado que en este caso $Hf(x, y)$ es diagonal, los elementos de la diagonal son sus valores propios, lo que nos permite aplicar directamente el criterio de autovalores:

- $Hf(1, 2)$ es definida positiva, $(1, 2)$ es mínimo relativo.

- $Hf(1,-2)$ es indefinida, $(1,-2)$ es punto de silla.
- $Hf(-1,2)$ es indefinida, $(-1,2)$ es punto de silla.
- $Hf(-1,-2)$ es definida negativa, $(-1,-2)$ es máximo relativo.

Ejemplo 19. Obtenemos los extremos relativos de

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

La función es polinómica, definida en $\mathbb{R}^2$ y de clase ∞. los puntos estacionarios se obtienen de

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= 2x - 2y + 4x^3y^4 = 0, \\ \frac{\partial f}{\partial y}(x,y) &= -2x + 2y + 4x^4y^3 = 0. \end{aligned}$$

Sumando ambas ecuaciones se llega a que el único punto crítico es $(0,0)$. La matriz hessiana en este punto es

$$Hf(0,0) = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}.$$

Dado que $\det Hf(0,0) = 0$ ninguno de los criterios de la nota 3.8 anteriores es aplicable (un autovalor es nulo). No obstante, si estudiamos la función vemos que

$$\begin{aligned} f(0,0) &= 0, \\ f(x,y) &= (x-y)^2 + x^4y^4 \geq 0, \end{aligned}$$

por lo tanto $(0,0)$ es, por definición, un mínimo relativo de f.

Ejemplo 20. Dada $f(x,y) = -2xy$ se pide estudiar el punto crítico $(0,0)$. La función es polinómica, definida en $\mathbb{R}^2$ y de clase ∞. La matriz hessiana en $(0,0)$ es

$$Hf(0,0) = \begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix}.$$

Si aplicamos el criterio de Sylvester se tiene $\triangle_1 = 0$, $\triangle_2 = \det Hf(0,0) = -4$, la matriz es indefinida y $(0,0)$ es un punto de silla.

En efecto, $f(0,0) = 0$, si nos aproximamos a $(0,0)$ con puntos de la forma (x,x), se tiene que $f(x,x) = -2x^2 < 0$ $(x \neq 0)$. Si nos aproximamos a $(0,0)$ con puntos de la forma $(x,-x)$, se tiene que $f(x,-x) = 2x^2 > 0$ $(x \neq 0)$. Por lo tanto en cualquier entorno de $(0,0)$ hay puntos del dominio de la función donde ésta es positiva y puntos en los que es negativa, de manera que $(0,0)$ no puede ser un extremo local (véase la figura 3.4).

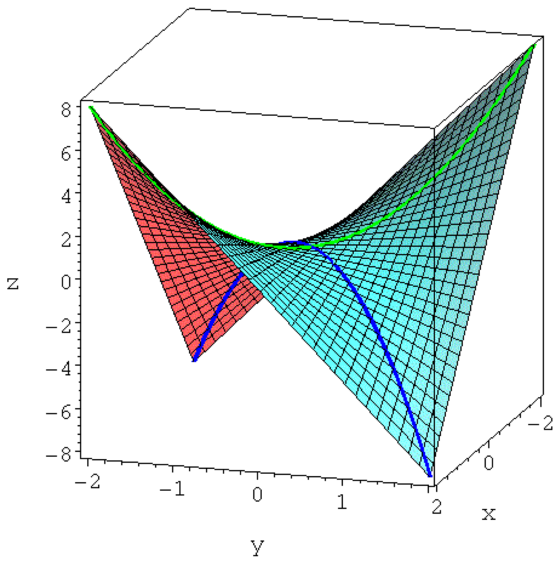


Figura 3.4: Gráfica de la función del ejemplo 20 con punto de silla en $(0,0)$. En azul los puntos de aproximación a $(0,0)$ con $y=x$, en verde con $y=-x$.

3.2. Extremos relativos condicionados

Nota 3.9. Sean $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ y M un subconjunto de A. Se dice que $\mathbf{a} \in M$ es un **máximo relativo** de f **condicionado por** M si existe un entorno $B(\mathbf{a}, \delta)$ tal que $f(\mathbf{a}) \geq f(\mathbf{x})$, para todo $\mathbf{x} \in B(\mathbf{a}, \delta) \cap M$. Análogamente, $\mathbf{a} \in A$ es un **mínimo relativo** de f **condicionado por** M si existe un entorno $B(\mathbf{a}, \delta)$ tal que $f(\mathbf{a}) \leq f(\mathbf{x})$, para todo $\mathbf{x} \in B(\mathbf{a}, \delta) \cap M$. A los máximos o mínimos relativos condicionados los denominaremos **extremos condicionados**.

Nota 3.10. Conceptualmente los extremos de f condicionados por M son los extremos libres de la restricción f_M. Por lo tanto, si $\mathbf{a} \in \operatorname{Int} M$, el estudio se puede realizar con los métodos estudiados para extremos libres, que están desarrollados para dominios abiertos. Pero tales métodos no son de aplicación si $\mathbf{a} \in \operatorname{Fr} M$.

Nota 3.11. En general nos restringiremos al caso en que $M \subset A$ esté definido a través de cierta función vectorial $\mathbf{g} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ $(m < n)$ en la forma

$$S = \{\mathbf{x} \in M \ / \ \mathbf{g}(\mathbf{x}) = \mathbf{0}\},$$

A las ecuaciones $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ las denominaremos **ligaduras**, restricciones o condiciones.

Nota 3.12. Una función puede carecer de extremos relativos en A, pero tener extremos relativos condicionados cuando se considera la función restringida al subconjunto $M \subset A$. Por ejemplo, la función $f \in C^2(\mathbb{R}^2)$ definida por

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

no tiene puntos estacionarios ya que sus derivadas parciales no se anulan en ningún punto, por lo tanto f no tiene extremos relativos en $\mathbb{R}^2$. Nótese que la gráfica de f es un plano. Sin embargo, si la restringimos al conjunto

$$M = \left\{(x, y) \in \mathbb{R}^2 \ / \ g(x, y) = x^2 + y^2 - 1 = 0\right\},$$

que define un cilindro que intersecta al plano gráfica de f, se tiene que la gráfica de f restringida a M es una elipse inscrita en el plano gráfica de f. Al no ser éste un plano horizontal, sobre la elipse existe un valor máximo y otro mínimo de la función, que se corresponden con el máximo y mínimo relativos de f condicionados por M (véase la figura 3.5).

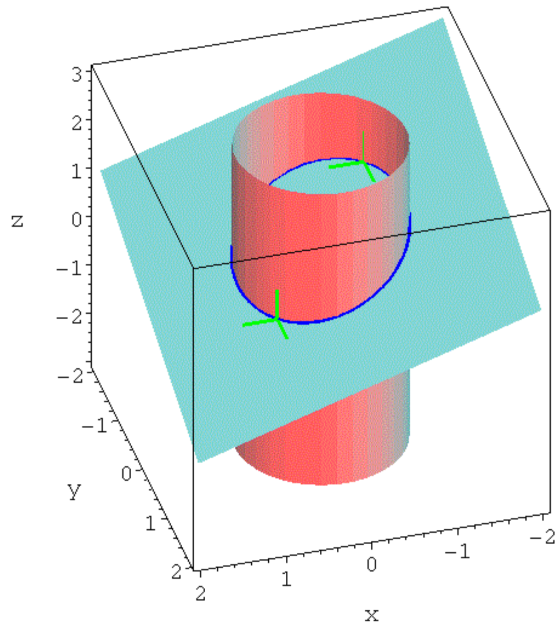


Figura 3.5: Gráfica de las funciones de la nota 3.12. Las cruces de color verde indican los extremos relativos condicionados.

Nota 3.13. Las ligaduras $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ son un conjunto de m ecuaciones con n incógnitas $(x_1, x_2, \ldots, x_n)$. Si podemos sustituir las m ecuaciones en la función f, ésta queda reducida a $n - m$ variables independientes, y el estudio de extremos relativos condicionados consiste en el estudio de extremos relativos libres. Dado que en general las restricciones vendrán definidas por funciones implícitas, este procedimiento sólo está garantizado si las ecuaciones son lineales respecto a m de las variables. Veamos un ejemplo.

Ejemplo 21. Obtenemos los extremos relativos de $f(x, y) = x^2 + y^2$ condicionados por $x + y - 1 = 0$.

En este caso $n = 2$ y $m = 1$, con

$$M = \left\{(x, y) \in \mathbb{R}^2 \;/\; g(x, y) = x + y - 1 = 0\right\}.$$

Dado que en la ligadura podemos despejar una de las variables (de hecho, ambas), $y = 1 - x$, el problema se reduce entonces al estudio de extremos relativos de la función de una variable

$$h(x) = x^2 + (1 - x)^2 = 2x^2 - 2x + 1.$$

Resolviendo $h'(x) = 0$ se tiene que $x = 1/2$ es un punto crítico, y dado que $h''(1/2) > 0$, entonces $x = 1/2$ es un mínimo relativo de h. Por lo tanto, como $x = 1/2 \to y = 1/2$, el punto $(1/2, 1/2)$ es un mínimo relativo de f condicionado por M.

3.2.1. Multiplicadores de Lagrange

Nota 3.14. Nos proponemos resolver el ejemplo 21 sin reducirlo a una variable, ya que no siempre es posible despejar alguna de las variables de la ecuación $g(x, y) = 0$.

Para ello vamos a considerar las curvas de nivel de la función $f(x, y) = x^2 + y^2$, esto es, $x^2 + y^2 = k$ donde $k \geq 0$. Esto son circunferencias de centro el origen y radio k. La primera circunferencia que toca a la recta de ecuación $x + y - 1 = 0$ (la ligadura del problema) lo hará en el punto (x_0, y_0) perteneciente a la curva de nivel con menor valor de k (menor radio) y a la restricción, en consecuencia será un mínimo relativo condicionado.

Geométricamente se trata de encontrar el punto (x_0, y_0) donde la curva de nivel $f(x, y) = k$ es tangente a la curva de ecuación $g(x, y) = 0$. Ambas curvas son tangentes si sus vectores normales (supuestos no nulos) son paralelos, esto es, los vectores gradiente han de ser proporcionales o, equivalentemente, existirá un $\lambda \in \mathbb{R}$ tal que

$$\nabla f(x, y) + \lambda \nabla g(x, y) = \mathbf{0}. \tag{3.3}$$

Resolvamos, $\nabla f(x, y) = (2x, 2y)$, $\nabla g(x, y) = (1, 1)$, por lo tanto, $2x + \lambda = 0$, $2y + \lambda = 0$, y podemos hacer uso también de la ecuación de ligadura,

$x + y - 1 = 0$. En este caso se tiene un sistema lineal (en general no tiene por qué serlo) de tres ecuaciones con tres incógnitas, cuya solución es $x_0 = y_0 = 1/2$, $\lambda = -1$. Por lo tanto la función $f(x, y)$ alcanza un mínimo relativo condicionado en el punto $(1/2, 1/2)$.

La generalización de este problema da lugar al método de múltiplicadores de Lagrange que traslada el estudio de extremos condicionados al estudio de extremos libres de una función auxiliar.

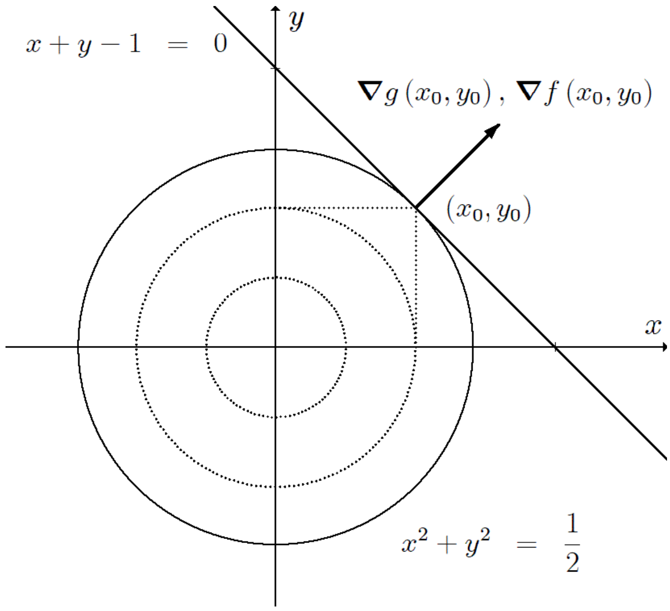


Figura 3.6: Construcción geométrica de la nota 3.14. La flecha indica la dirección de los vectores gradiente en el punto de contacto entre la recta (ligadura) y la circunferencia (curva de nivel).

Proposición 3.3 (Condición necesaria de extremo condicionado)**.** Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A conjunto abierto, $f \in C^1(A)$, y $\mathbf{g} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ $(m < n)$, $g \in C^1(A)$, Consideremos el conjunto $M = \{\mathbf{x} \in A \ / \ \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$ y un punto $\mathbf{a} \in M$ en el que el rango de $J\mathbf{g}(\mathbf{a})$ es m.
Entonces, si f tiene en $\mathbf{a}$ un extremo relativo condicionado por M, existen m números reales $\lambda_1, \lambda_2, \ldots, \lambda_m$ tales que $\mathbf{a}$ es un punto estacionario de la función

$$F(\mathbf{x}) = f(\mathbf{x}) + \lambda_1 g_1(\mathbf{x}) + \ldots + \lambda_m g_m(\mathbf{x}). \tag{3.4}$$

Definición 3.3. A los m números reales $\lambda_1, \lambda_2, \ldots, \lambda_m$ se les denomina **multiplicadores de Lagrange** en el punto $\mathbf{a}$, y a la función F **función** (auxiliar) **de Lagrange** (o función Lagrangeana).

Nota 3.15. Si $\mathbf{a}$ es un punto estacionario de F se tiene

$$\boldsymbol{\nabla} F(\mathbf{a}) = \boldsymbol{\nabla} f(\mathbf{a}) + \sum_{j=1}^{m} \lambda_j \boldsymbol{\nabla} g_j(\mathbf{a}) = \mathbf{0}. \tag{3.5}$$

Nota 3.16. La proposición no implica que en un punto estacionario de la función de Lagrange necesariamente la función f deba alcanzar un extremo relativo condicionado por M, sino que los candidatos a extremos relativos condicionados estarán entre las soluciones del sistema algebraico de $n+m$ ecuaciones

$$\boldsymbol{\nabla} F(\mathbf{a}) = \mathbf{0}, \ \ \mathbf{g}(\mathbf{a}) = \mathbf{0}, \tag{3.6}$$

o en componentes,

$$\begin{aligned} \frac{\partial F}{\partial x_1}(\mathbf{a}) &= 0, \ldots, \frac{\partial F}{\partial x_n}(\mathbf{a}) = 0, \\ g_1(\mathbf{a}) &= 0, \ldots, g_m(\mathbf{a}) = 0. \end{aligned} \tag{3.7}$$

Las incógnitas de este sistema de ecuaciones son las n coordenadas de $\mathbf{a}$ y los m multiplicadores de Lagrange $\lambda_1, \lambda_2, \ldots, \lambda_m$, de forma que cada valor de $\mathbf{a}$ lleva asociado un valor concreto de los multiplicadores. En general los sistemas de ecuaciones anteriores no son lineales, por lo que no existe un procedimiento sistemático para resolverlos.

Proposición 3.4 (Condición suficiente de extremo condicionado)**.** Sean $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, con A conjunto abierto, $f \in C^2(A)$, y $\mathbf{g} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ $(m < n)$, $g \in C^2(A)$, Sean el conjunto $M = \{\mathbf{x} \in A \ / \ \mathbf{g}(\mathbf{x}) = \mathbf{0}\}$ y $\mathbf{a}$ un punto estacionario de la función de Lagrange $F(\mathbf{x})$ en el que el rango de $J\mathbf{g}(\mathbf{a})$ es m. Consideremos la diferencial segunda de F en $\mathbf{a}$,

$$d^2 F_{\mathbf{a}}(\mathbf{h}) = \mathbf{h}^t HF(\mathbf{a})\,\mathbf{h}, \tag{3.8}$$

con $\mathbf{h} \in \mathbb{R}^n$, $\mathbf{h} \neq \mathbf{0}$, sujeta a las condiciones

$$J\mathbf{g}(\mathbf{a})\,\mathbf{h} = \mathbf{0}. \tag{3.9}$$

Entonces,

1. Si $d^2F_{\mathbf{a}}(\mathbf{h}) > 0$, $\mathbf{a}$ es un mínimo relativo de f condicionado por M.
2. Si $d^2F_{\mathbf{a}}(\mathbf{h}) < 0$, $\mathbf{a}$ es un máximo relativo de f condicionado por M.
3. Si $d^2F_{\mathbf{a}}(\mathbf{h})$ toma valores positivos y negativos, $\mathbf{a}$ no es un extremo relativo de f condicionado por M.

Nota 3.17. La diferencial segunda (3.8) sujeta a las condiciones (3.9) se puede escribir en la forma

$$(h_1 \dots h_{n-m})\, S(\mathbf{a}) \begin{pmatrix} h_1 \\ \vdots \\ h_{n-m} \end{pmatrix}$$

donde $S(\mathbf{a})$ es una matriz real y simétrica de dimensión $n-m$. Entonces, si $S(\mathbf{a})$ es definida positiva, el punto **a** es un mínimo relativo de f condicionado por M; si $S(\mathbf{a})$ es definida negativa, el punto **a** es un máximo relativo de f condicionado por M; si $S(\mathbf{a})$ es indefinida, **a** no es un extremo relativo.

Nota 3.18. No siempre es necesario hacer uso de las condiciones (3.9), ya que si directamente la matriz hessiana $HF(\mathbf{a})$ es definida positiva, el punto **a** es un mínimo relativo de f condicionado por M; y si $HF(\mathbf{a})$ es definida negativa, el punto **a** es un máximo relativo de f condicionado por M.

Ejemplo 22. Obtengamos los extremos relativos de $f(x,y) = x^2 + y^2 - xy$ condicionados por $x^2 + y^2 = 2$.

Buscamos los extremos relativos de f condicionados por el conjunto

$$M = \left\{(x,y) \in \mathbb{R}^2 \ / \ g(x,y) = x^2 + y^2 - 2 = 0\right\}.$$

Tanto f como g son polinomios, con dominio $\mathbb{R}^2$ y de clase ∞. La función de Lagrange (3.4) es

$$F(x,y) = x^2 + y^2 - xy + \lambda\left(x^2 + y^2 - 2\right)$$

y sus puntos estacionarios se obtienen del sistema de ecuaciones (3.7), de 3 ecuaciones con 3 incógnitas (x, y, λ), esto es,

$$\begin{aligned} \frac{\partial F}{\partial x}(x,y) &= 2x - y + 2\lambda x = 0, \\ \frac{\partial F}{\partial y}(x,y) &= 2y - x + 2\lambda y = 0, \\ g(x,y) &= x^2 + y^2 - 2 = 0. \end{aligned}$$

El sistema es no lineal. Se puede resolver comenzando con la sustitución de la segunda ecuación en la primera,

$$y = 2x(1+\lambda),\ x = 2y(1+\lambda) \longrightarrow (1+\lambda)^2 = \frac{1}{4} \longrightarrow \lambda = -\frac{1}{2}, -\frac{3}{2}.$$

Con los valores de λ se obtienen directamente los de x e y, esto es, el conjunto de puntos críticos

$$\left\{\left(1,1,-\frac{1}{2}\right), \left(-1,-1,-\frac{1}{2}\right), \left(1,-1,-\frac{3}{2}\right), \left(-1,1,-\frac{3}{2}\right)\right\}.$$

Para que estos puntos sean candidatos a extremos relativos condicionados es necesario que el rango de la matriz jacobiana de g calculada en cada uno de ellos sea $m = 1$ (proposición 3.3). Dado que

$$\text{rango } \; Jg\,(x,y) = \text{rango } \; (2x, 2y) = 1 \text{ si } \; (x,y) \neq (0,0)\,,$$

la condición se verifica para todos los puntos. Estudiamos entonces las condiciones suficientes de extremo condicionado (proposición 3.4).
La matriz hessiana de F es

$$HF\,(x,y) = \begin{pmatrix} 2+2\lambda & -1 \\ -1 & 2+2\lambda \end{pmatrix},$$

que evaluada en los puntos anteriores resulta

$$\begin{aligned} HF\,(1,1) &= HF\,(-1,-1) \overset{\lambda=-\frac{1}{2}}{=} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \\ HF\,(1,-1) &= HF\,(-1,1) \overset{\lambda=-\frac{3}{2}}{=} \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}. \end{aligned}$$

En todos los casos $\det HF = 0$, por lo que no es definida positiva ni negativa. Por lo tanto, dado $\mathbf{h} = (h_1, h_2) \neq \mathbf{0}$ se requiere estudiar $d^2F_{(x_i,y_i)}\,(h_1,h_2)$, con (x_i, y_i) los candidatos a extremo ($i = 1,2,3,4$), sujeta a la condición

$$Jg\,(x_i,y_i) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = (2x_i, 2y_i) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = 0.$$

En cada uno de los puntos críticos esta condición se traduce en

$$\begin{aligned} \mathbf{a}_1 &= (1,1) \to (2,2) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = 0 \to h_1 + h_2 = 0 \to \mathbf{h} = (h_1, -h_1)\,, \\ \mathbf{a}_2 &= (-1,-1) \to (-2,-2) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = 0 \to h_1 + h_2 = 0 \to \mathbf{h} = (h_1, -h_1)\,, \\ \mathbf{a}_3 &= (1,-1) \to (2,-2) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = 0 \to h_1 - h_2 = 0 \to \mathbf{h} = (h_1, h_1)\,, \\ \mathbf{a}_4 &= (-1,1) \to (-2,2) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = 0 \to -h_1 + h_2 = 0 \to \mathbf{h} = (h_1, h_1)\,. \end{aligned}$$

Por lo tanto,

$$\begin{aligned} d^2F_{(1,1)}\,(h_1,h_2) &= d^2F_{(-1,-1)}\,(h_1,h_2) = \\ &= (h_1, -h_1) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} h_1 \\ -h_1 \end{pmatrix} = 4h_1^2 > 0 \; (\mathbf{h} \neq \mathbf{0}), \\ d^2F_{(1,-1)}\,(h_1,h_2) &= d^2F_{(-1,1)}\,(h_1,h_2) = \\ &= (h_1, h_1) \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} h_1 \\ h_1 \end{pmatrix} = -4h_1^2 < 0 \; (\mathbf{h} \neq \mathbf{0}). \end{aligned}$$

Por aplicación de la proposición 3.4 concluimos que los puntos $(1,1)$ y $(-1,-1)$ son mínimos relativos de f condicionados por M, y los puntos $(-1,1)$ y $(1,-1)$ son máximos relativos de f condicionados por M.

Ejemplo 23. Consideramos el problema planteado en la nota 3.12. Veamos que, en efecto, tal y como se razonó geométricamente, el problema tiene dos extremos relativos (máximo y mínimo) de f condicionados a M. La función de Lagrange es

$$F(x,y) = 1 - \frac{x}{2} - \frac{y}{2} + \lambda\left(x^2 + y^2 - 1\right),$$

a la que corresponden los puntos críticos (queda como ejercicio)

$$\left\{\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{4}\right), \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{4}\right)\right\},$$

que son candidatos a extremo en virtud de que

$$\text{rango } Jg(x,y) = \text{rango } (2x, 2y) = 1 \text{ si } (x,y) \neq (0,0).$$

En este caso, la matriz hessiana de F en todo punto es

$$HF(x,y) = \begin{pmatrix} 2\lambda & 0 \\ 0 & 2\lambda \end{pmatrix},$$

que es definida positiva si $\lambda = \sqrt{2}/4$ y definida negativa si $\lambda = -\sqrt{2}/4$. Por lo tanto, el punto $\left(\sqrt{2}/2, \sqrt{2}/2\right)$ es un mínimo relativo de f condicionado por M, y el punto $\left(-\sqrt{2}/2, -\sqrt{2}/2\right)$ es un máximo relativo de f condicionado por M (véase la figura 3.5).

Nota 3.19. El problema de determinar si una forma cuadrática es definida positiva o negativa para aquellos vectores que cumplen ciertas condiciones de tipo lineal es de naturaleza algebraica. En las condiciones de la proposición 3.4, los puntos críticos se pueden clasificar alternativamente construyendo el polinomio $p(\alpha)$ de grado $n-m$ dado por (matriz expresada por bloques)

$$p(\alpha) = \det\left(\begin{array}{c|c} HF(\mathbf{a}) - \alpha I_n & J\mathbf{g}(\mathbf{a})^t \\ \hline J\mathbf{g}(\mathbf{a}) & 0_m \end{array}\right),$$

donde I_n es la matriz identidad de dimensión n y 0_m es la matriz nula cuadrada de dimensión m. Entonces,

1. Si todas las raíces de $p(\alpha)$ son positivas, $\mathbf{a}$ es un mínimo relativo de f condicionado por M.

2. Si todas las raíces de $p(\alpha)$ son negativas, **a** es un máximo relativo de f condicionado por M.

En otro caso el criterio no clasifica.

Ejemplo 24. Con los datos del ejemplo 22 para el punto crítico $(1,1)$ se tiene

$$p(1,1) = \det \begin{pmatrix} 1-\alpha & -1 & 2 \\ -1 & 1-\alpha & 2 \\ 2 & 2 & 0 \end{pmatrix} = -8 - 8(1-\alpha),$$

que es un polinimo de grado 1 ($n - m = 2 - 1 = 1$). Su única raíz es $\alpha = 2$, positiva, luego $(1,1)$ es un mínimo relativo de f condicionado por M.

3.3. Extremos absolutos

Definición 3.4. Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$. Se dice que $\mathbf{a} \in A$ es un **máximo absoluto** de f en A si $f(\mathbf{a}) \geq f(\mathbf{x})$, para todo $\mathbf{x} \in A$. Análogamente, $\mathbf{a} \in A$ es un **mínimo absoluto** de f en A si $f(\mathbf{a}) \leq f(\mathbf{x})$, para todo $\mathbf{x} \in A$. A los máximos o mínimos absolutos los denominaremos **extremos absolutos** o **globales**.

Nota 3.20. Si $\mathbf{a} \in \operatorname{Int} A$ es un extremo absoluto, también es extremo relativo. El concepto de extremo absoluto ya se introdujo al enunciar el *teorema de Weierstrass* para funciones escalares, que nos aseguraba la existencia de mínimo y máximo absolutos si f es continua en A y A es un conjunto compacto.

Proposición 3.5. Sea $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, continua en A y A conjunto compacto. Consideremos los siguientes conjuntos:

1. El conjunto $S_{\operatorname{Int} A}$ de los puntos estacionarios de f en $\operatorname{Int} A$ (supuesto no vacío).

2. El conjunto $S_{\operatorname{Fr} A}$ de los puntos donde f podría alcanzar un extremo relativo condicionado por el conjunto $\operatorname{Fr} A$.

Entonces, el máximo (resp. mínimo) absoluto de f en A se alcanza en los puntos del conjunto $S_{\operatorname{Int} A} \cup S_{\operatorname{Fr} A}$ donde f tome el mayor (resp. menor) valor.

Nota 3.21. Las situaciones más frecuentes serán en la que el conjunto $\operatorname{Fr} A$ esté definido por un conjunto de nivel de una función diferenciable, $g(\mathbf{x}) = k$, $k \in \mathbb{R}$, entonces los posibles extremos relativos condicionados se pueden encontrar mediante el *método de los multiplicadores de Lagrange*; o bien $\operatorname{Fr} A = A$ (y por tanto con interior vacío), en tal caso los candidatos sólo provienen de los posibles extremos relativos condicionados por $\operatorname{Fr} A$.

Nota 3.22. Al contrario de lo que ocurre con los extremos relativos, para determinar los extremos absolutos sólo es necesario evaluar directamente la función en los puntos candidatos en el conjunto $S_{\operatorname{Int} A} \cup S_{\operatorname{Fr} A}$.

Proposición 3.6. Sean $f : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ y $\mathbf{a} \in A$. Sean $B \subseteq \mathbb{R}$ con $f(A) \subseteq B$, y $h : B \longrightarrow \mathbb{R}$ una función real de variable real. Consideremos la función compuesta $h \circ f$, entonces se verifica que:

1. Si h es estrictamente creciente en $f(A)$, entonces f alcanza un máximo (resp. mínimo) relativo o absoluto en el punto $\mathbf{a}$ si, y sólo si, $h \circ f$ alcanza un máximo (resp. mínimo) relativo o absoluto en el punto $\mathbf{a}$.

2. Si h es estrictamente decreciente en $f(A)$, entonces f alcanza un máximo (resp. mínimo) relativo o absoluto en el punto $\mathbf{a}$ si, y sólo si, $h \circ f$ alcanza un mínimo (resp. máximo) relativo o absoluto en el punto $\mathbf{a}$.

Nota 3.23. Esto es, bajo las condiciones de la proposición, las funciones f y $h \circ f$ alcanzan los extremos locales o globales en los mismos puntos. Con las anteriores herramientas, nos centraremos en el estudio del problema de cálculo de distancias entre conjuntos.

3.3.1. Cálculo de distancias de puntos a curvas

Nota 3.24. Consideremos un punto $\mathbf{a} = (a, b) \in \mathbb{R}^2$. La distancia entre este punto y un punto cualquiera $\mathbf{x} = (x, y)$ perteneciente a un conjunto $A \subseteq \mathbb{R}^2$ viene dada por la función (distancia euclídea)

$$d_{\mathbf{a}}(x, y) = \sqrt{(x-a)^2 + (y-b)^2}.$$

Los extremos, relativos o absolutos, de $d_{\mathbf{a}}(x, y)$ son los mismos que los de la función

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

por aplicación de la proposición 3.6. En efecto, si definimos $h(t) = \sqrt{t}$, se tiene que

$$d_{\mathbf{a}}(x, y) = (h \circ f)(x, y) = h(f(x, y)) = \sqrt{(x-a)^2 + (y-b)^2}.$$

Dado que $h'(t) = t^{-1/2}/2 > 0$ para todo $t > 0$, h es estrictamente creciente para $t > 0$, luego las funciones $d_{\mathbf{a}}$ y f alcanzan sus extremos en los mismos puntos. En general es más conveniente hacer uso de f para realizar cálculos ya que esta función es un polinomio.

Ejemplo 25. Calculamos la distancia máxima y mínima del punto $(2,0)$ a la curva $x^2+y^2=1$.

La función para optimizar es

$$d_{(2,0)}(x,y)=\sqrt{(x-2)^2+y^2}.$$

Los puntos (x,y) deben pertenecer a la curva $x^2+y^2=1$ (circunferencia de centro el origen y radio 1). Por lo tanto se buscan los **extremos absolutos** de $d_{(2,0)}(x,y)$ en el conjunto compacto

$$M=\left\{(x,y)\in\mathbb{R}^2 \;/\; g(x,y)=x^2+y^2-1=0\right\}.$$

En virtud de la nota 3.24, utilizaremos la función polinómica

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

en lugar de $d_{(2,0)}(x,y)$. Para encontrar estos puntos, dado que $\operatorname{Int} M$ es vacío, $\operatorname{Fr} M=M$, únicamente es necesario determinar los posibles extremos relativos de f condicionados por M. Es decir, el conjunto de puntos candidatos es $S_{\operatorname{Fr} M}=S_M$.
Obtenemos en primer lugar los puntos estacionarios de la función de Lagrange

$$F(x,y)=f(x,y)+\lambda g(x,y)=(x-2)^2+y^2+\lambda\left(x^2+y^2-1\right).$$

Los candidatos a extremos relativos han de verificar rango $Jg(x,y)=1$. Dado que

$$\text{rango}\, Jg(x,y)=\text{rango}\,(2x,2y)=1 \text{ si } (x,y)\neq(0,0),$$

y $(0,0)\notin M$, esta condición se cumple en cualquier caso.
El sistema de ecuaciones a resolver es

$$\begin{aligned}
\frac{\partial F}{\partial x}(x,y) &= 2x-4+2\lambda x=0,\\
\frac{\partial F}{\partial y}(x,y) &= 2y+2\lambda y=0,\\
g(x,y) &= x^2+y^2-1=0.
\end{aligned}$$

De las dos primeras ecuaciones se obtiene

$$x(1+\lambda)=2,\; y(1+\lambda)=0$$

de donde $y=0$ (nótese que $1+\lambda\neq 0$ necesariamente) y por lo tanto, de la tercera ecuación, $x^2=1$, resulta $x=\pm1$. Esto es, las soluciones son los puntos $(-1,0)$ y $(1,0)$. Por la proposición 3.5, en este conjunto de puntos necesariamente f alcanza un extremo absoluto. Así que se tiene que

$$S_M = \{(-1,0), (1,0)\}.$$

Evaluando directamente la función obtenemos $f(1,0) = 1$ y $f(-1,0) = 9$, por lo tanto $(-1,0)$ es el *máximo absoluto* de f en M, y $(1,0)$ el *mínimo absoluto* de f en M.
De este resultado se deduce que las distancias mínima y máxima de la circunferencia al punto $(2,0)$ son

$$\begin{aligned} d_{(2,0)}(1,0) &= \sqrt{f(1,0)} = 1, \\ d_{(2,0)}(-1,0) &= \sqrt{f(-1,0)} = 3. \end{aligned}$$

Nota 3.25. Nótese en el ejemplo anterior que la ventaja de usar la proposición 3.5 es que no se requiere aplicar la condición suficiente de extremo relativo condicionado (estudio del carácter de la diferencial segunda).

Nota 3.26. Existen extensiones del teorema de Weierstrass para casos con M conjunto cerrado y no acotado, y f función continua en M, que aseguran la **existencia** de uno o más puntos donde la función alcanza el **mínimo absoluto**. En la práctica es suficiente comprobar si los puntos $\mathbf{x} \in M$ cuya distancia al origen tienda a infinito, proporcionan valores positivos de f no acotados, esto es,

$$\lim_{\substack{\|\mathbf{x}\| \to \infty \\ \mathbf{x} \in M}} f(\mathbf{x}) = +\infty.$$

Por lo tanto, si se cumple esta condición y se tiene un único candidato a *extremo relativo* de f *condicionado* por M este será un mínimo absoluto. Si hay más de uno, en el caso general, resolveríamos aplicando también la condición suficiente de extremos relativos condicionados por M.

Ejemplo 26. Hallar los puntos de la curva $x^2 + 14xy + y^2 = 1$ más cercanos al origen de coordenadas.

En este caso no podemos aplicar la proposición 3.5 debido a que el conjunto

$$M = \left\{(x,y) \in \mathbb{R}^2 \ / \ g(x,y) = x^2 + 14xy + y^2 - 1 = 0\right\}$$

no es acotado, y por lo tanto no es compacto (se trata de una hipérbola).

Para el cálculo de distancias siempre se tiene $f(x,y) = x^2 + y^2 \geq 0$. Si $(x,y) \in M$, entonces $x^2 + y^2 = 1 - 14xy \geq 0$, por lo tanto

$$\lim_{\substack{\|(x,y)\| \to \infty \\ (x,y) \in M}} f(x,y) = \lim_{\substack{\|(x,y)\| \to \infty \\ xy < 0}} (1 - 14xy) = +\infty,$$

lo que asegura la existencia de un punto (al menos) donde la función alcanza el mínimo absoluto.
Para encontrar los candidatos a extremo relativo usamos el método de los multiplicadores de Lagrange (ejercicio), que arroja los puntos $(1/4, 1/4)$ y $(-1/4, -1/4)$. En este caso, ambos son mínimos absolutos porque la función alcanza en ambos su valor mínimo, es decir,

$$d_{(0,0)}\left(\pm\frac{1}{4}, \pm\frac{1}{4}\right) = \sqrt{\left(\pm\frac{1}{4}\right)^2 + \left(\pm\frac{1}{4}\right)^2} = \sqrt{\frac{1}{8}} = \frac{\sqrt{2}}{4}.$$

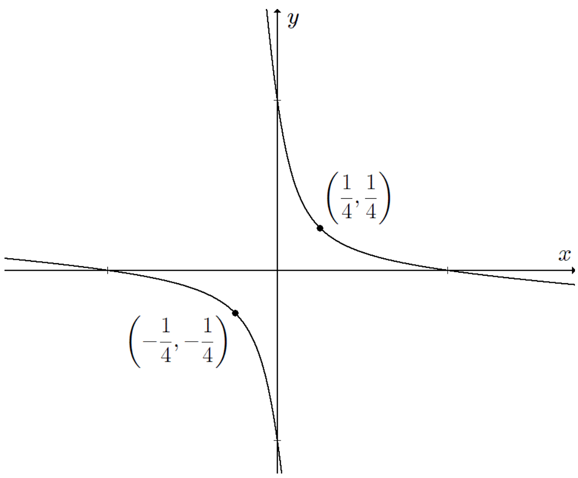


Figura 3.7: Curva y puntos de distancia mínima al origen del ejemplo 26.

Ejercicios

Ejercicio 1. Obtener los extremos relativos de las siguientes funciones:
a) $f(x,y) = 2(x^2+y^2)e^{-x^2-y^2}$.
b) $f(x,y) = x^3+y^3-3x-12y+25$.
c) $f(x,y) = x^2+y^2+xy+3y-5$.
d) $f(x,y,z) = x^2+y^2+z^2+2\alpha xy$ ($\alpha > 0$, $\alpha \neq 1$).
e) $f(x,y) = (x^2+y^2)^2 - 2a^2(x^2-y^2)$ ($a \neq 0$).
f) $f(x,y) = x^2-4xy^2+4y^4$.
g) $f(x,y,z) = e^{\sqrt[5]{x^2-2x+y^2}}$.

Sol.: a) $(0,0)$ mínimo, $\{(x_0,y_0) \in \mathbb{R}^2 \;/\; x_0^2+y_0^2 = 1\}$ máximos. b) $(1,2)$ mínimo, $(-1,-2)$ máximo, $\{(1,-2),(-1,2)\}$ puntos de silla. c) $(1,-2)$ mínimo. d) $(0,0,0)$ si $0 < \alpha < 1$: mínimo; si $\alpha > 1$ punto de silla. d) $(0,0)$ punto de silla, $(a,0)$ mínimo, $(-a,0)$ mínimo. f) Infinitos mínimos en puntos con $x = y^2$. g) $(1,0)$ mínimo (use proposición 3.6).

Ejercicio 2. Estudiar el punto crítico origen de las siguientes funciones:
a) $f(x,y) = -2xy$.
b) $f(x_1,x_2,\ldots,x_n) = |x_1|+|x_2|+\ldots+|x_n|$.
c) $f(x,y) = (x-y^2)(x-y^3)$.
d) $f(x,y) = x^2-2xy+y^2+x^4y^4$.
e) $f(x,y) = \frac{x^2}{2p}+\frac{x^2}{2q}$, con $p,q \neq 0$.

Sol.: a) Punto de silla. b) Mínimo. c) Punto de silla. d) Mínimo. e) Si $\operatorname{signo} p \neq \operatorname{signo} q$, punto de silla. Si $\operatorname{signo} p = \operatorname{signo} q$: mínimo si $p > 0$, máximo $p > 0$.

Ejercicio 3. Calcular los extremos relativos de f condicionados por M en los siguientes casos:
a) $f(x,y) = x^3-y^2+1$, $M = \{(x,y) \in \mathbb{R}^2 \;/\; x+y^2 = 1\}$.
b) $f(x,y) = 15-6x-8y$, $M = \{(x,y) \in \mathbb{R}^2 \;/\; x^2+y^2 = 16\}$.
c) $f(x,y) = x^{-1}+y^{-1}$, $M = \{(x,y) \in \mathbb{R}^2 \;/\; x^2y^2 = 1\}$.
d) $f(x,y) = x^2+y^2-xy$, $M = \{(x,y) \in \mathbb{R}^2 \;/\; x^2+y^2 = 2\}$.

Sol.: a) $(1,0)$ máximo. b) $(12/5,16/5)$ mínimo, $(-12/5,-16/5)$ máximo. c) $(1,1)$ mínimo, $(-1,-1)$ máximo. d) $\{(1,1),(-1,-1)\}$ mínimos, $\{(1,-1),(-1,1)\}$ máximos.

Ejercicio 4. Considerando el estudio de los extremos relativos de f condicionados por M, determinar el carácter de los puntos críticos de la función de Lagrange que se indica en los siguientes casos:

a) $f(x,y,z) = xz + \log y + 2$, $\mathbf{x}_0 = (1,1,1)$,
$M = \{(x,y,z) \in \mathbb{R}^3 \ / \ z^2 + 2xy = 3\}$.

b) $f(x,y,z) = xz + yz$, $\mathbf{x}_0 = (2,0,1)$,
$M = \{(x,y,z) \in \mathbb{R}^3 \ / \ x^3 - y^3 - 4z - 4 = 0,\ x + y + z^2 - 3 = 0\}$.

---

a) $(1,1,1)$ no es extremo. b) $(2,0,1)$ es máximo..

Ejercicio 5. Dados los números complejos $z_1 = e^{i\alpha}$, $z_2 = e^{i\beta}$ y $z_3 = 1$, con $0 < \alpha$, $\beta < 2\pi$, se pide determinar α y β para que sea máxima la suma

$$S = |z_1 - z_2|^2 + |z_2 - z_3|^2 + |z_3 - z_1|^2 .$$

Sol.: Valor máximo $S = 9$, para $\alpha = 4\pi/3$, $\beta = 2\pi/3$.

Ejercicio 6. a) Hallar un punto en el interior de un triángulo dado tal que la suma de los cuadrados de las distancias a los tres vértices sea mínima.

b) Determinar el triángulo de área máxima inscrito en una circunferencia.

Sol.: a) $x = (x_1 + x_2 + x_3)/3$, $y = (y_1 + y_2 + y_3)/3$, donde (x_i, y_i) son las coordenadas del vértice $i-$ésimo. b) Triángulo equilatero.

Ejercicio 7. Hallar las distancias máxima y mínima de los puntos a las curvas siguientes:

a) Punto $(1,0)$, curva $y^2 = 4x$,
b) Punto $(2,1)$, curva $x^2 - 2x + y^2 = 7$,
c) Punto $(-1,-3)$, curva $9x^2 + 18x + y^2 - 6y + 9 = 0$,

Sol.: a) $(0,0)$ mínimo, $d = 1$. b) $(3,2)$ mínimo, $d = \sqrt{2}$; $(-1,-2)$ máximo, $d = 3\sqrt{2}$. c) $(-1,0)$ mínimo, $d = 3$; $(-1,6)$ máximo, $d = 9$.

Ejercicio 8. Hallar la distancia mínima del origen a la superficie

$$x^2 - yz + y = 5.$$

Sol.: a) $\left(\pm\sqrt{37}/3, 2/3, -1/3\right)$, $d = \sqrt{14/3}$.

Parte 4

Aplicaciones del cálculo diferencial (III)

4.1. Teorema de la función implícita

Nota 4.1. Supongamos la ecuación $f(x,y) = 0$ definida con una función real de dos variables reales. Buscamos condiciones suficientes que aseguren para cada valor de x en un cierto conjunto $X \subseteq \mathbb{R}$, la existencia de un valor y de otro conjunto $Y \subseteq \mathbb{R}$ de manera que el par (x,y) verifique la ecuación. Es el caso más simple, pero como veremos ya incluye todas las circunstancias que aparecen en el caso general, y admite una interpretación geométrica.
En efecto, si consideramos la función real dada por $z = f(x,y)$, definida en un abierto $A \subseteq \mathbb{R}^2$, continua en A, como sabemos graf $f \subset \mathbb{R}^3$ es una superficie, que llamaremos S. El objetivo del **teorema de la función implícita** es determinar las condiciones para que la superficie S corte al plano xy (de ecuación $z = 0$) y además que la intersección entre ambos sea la gráfica, contenida en el plano xy, de una función $y = g(x)$, esto es, que permita "despejar" la variable y como función de x. Tal función $g(x)$ se denomina **implícita**, ya que está definida implícitamente a través de $f(x,y)$.

Para que lo anterior suceda, primero será necesario que S intersecte al plano xy. Supongamos un punto (a,b) que pertenece tanto al plano xy como a la superficie S, y que S atraviesa al plano en un entorno de este punto. Lo mismo ocurrirá a las curvas trazadas sobre la superficie. En particular, la curva $z = f(a,y)$ deberá atravesar el plano xy. Una condición suficiente para que esto ocurra es que la tangente a la curva $z = f(a,y)$ en el punto

(a, b) tenga pendiente no nula. Esta condición se expresa por

$$\frac{\partial f}{\partial y}(a, b) \neq 0,$$

ya que entonces la curva es estrictamente monótona en el punto (a, b), y en un entorno de él, solo puede cortar al plano xy en dicho punto. Nótese que la derivada parcial es respecto a la variable "que se quiere despejar". No ahondaremos en el resto de hipótesis del teorema con carácter técnico para asegurar la existencia de la función implícita, $y = g(x)$, y su diferenciabilidad. Enunciamos a continuación un primer teorema para este caso particular, $f(x, y) = 0$, por su uso frecuente.

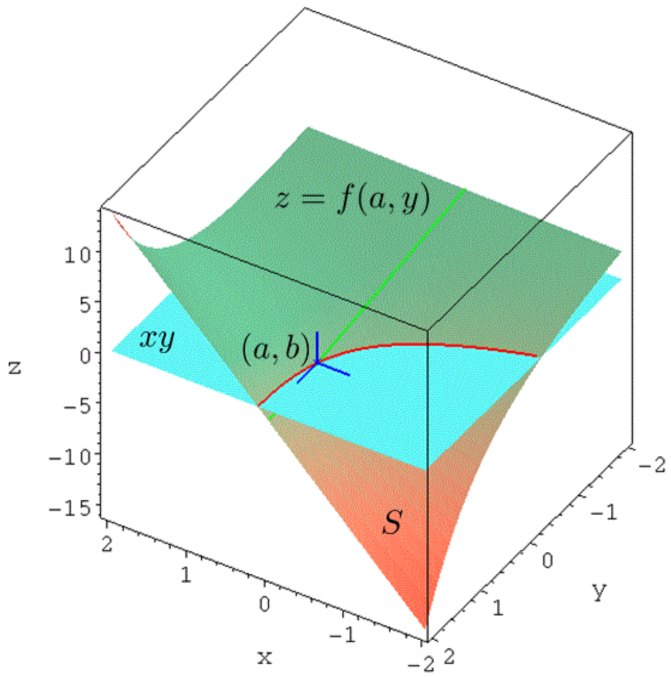


Figura 4.1: Gráfico de la nota 4.1. La curva $z = f(a, y)$ (en verde) atraviesa el plano xy (gris) en el punto (a, b) (cruz azul). La intersección de S con el plano xy en un entorno de (a, b) define la función implícita $y = g(x)$ (en rojo). Las gráficas representadas se corresponden con el ejemplo 27.

Teorema 2 (de la función implícita, caso $f(x, y) = 0$)**.** *Sea $f : A \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$, con A conjunto abierto y $f \in C^1(A)$. Sea un punto $(a, b) \in A$ tal que $f(a, b) = 0$. Si*

$$\frac{\partial f}{\partial y}(a, b) \neq 0, \tag{4.1}$$

entonces existen un intervalo abierto I que contiene a a, y una función única $g : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}$, $g \in C^1(I)$, tal que

$$g(a) = b, \;\; f(x, g(x)) = 0 \;\; \forall x \in I,$$

y además

$$g'(x) = -\frac{\dfrac{\partial f}{\partial x}(x, g(x))}{\dfrac{\partial f}{\partial y}(x, g(x))}, \; x \in I. \tag{4.2}$$

Si la función f es de clase r en A, la función g es también de clase r en I.

Nota 4.2. Nótese que la fórmula de la derivada de la función g se obtiene anulando la derivada de la función compuesta $f(x, g(x))$, que resulta

$$\frac{\partial f}{\partial x}(x, g(x)) + \frac{\partial f}{\partial y}(x, g(x))\, g'(x) = 0 \tag{4.3}$$

de donde se despeja $g'(x)$. La ecuación (4.3) procede de la aplicación de la regla de la cadena, a través del producto de las matrices jacobianas involucradas. La utilidad del teorema radica en que podemos calcular la derivada de la función implícita (y con carácter general, sus derivadas parciales en el caso de más de dos variables) aun desconociendo la expresión explícita de g. Su aplicación reiterada permitiría obtener derivadas de órdenes superiores de la función implícita.

Nota 4.3. Usualmente, por abuso de notación, se emplea $y = y(x)$ para la función implícita, en lugar de g.

Ejemplo 27. Sea $f(x, y) = xe^y - y + 1$, definida en $\mathbb{R}^2$ y de clase ∞.

La ecuación $f(x, y) = 0$, $xe^y - y + 1 = 0$, se verifica en el punto $(0, 1)$. Calculamos

$$\frac{\partial f}{\partial y}(x, y) = xe^y - 1, \; \frac{\partial f}{\partial y}(0, 1) = -1 \neq 0,$$

luego existe una función implícita $y = g(x)$ definida por la ecuación dada, tal que $1 = g(0)$ (véase la figura 4.1). En un entorno de $x = 0$, su derivada es

$$g'(x) = -\frac{e^y}{xe^y - 1} \equiv y',$$

en particular

$$g'(0) = -\frac{e^1}{-1} = e.$$

La expresión anterior de la derivada también se obtiene por derivación implícita (4.3), es decir, a partir de $xe^y - y + 1 = 0$ se tiene

$$e^y + xe^y y' - y' = 0 \to y' = -\frac{e^y}{xe^y - 1}.$$

Igualmente podemos obtener la derivada segunda, partiendo de esta última,

$$y'' = -\frac{e^y y' \left(xe^y - 1\right) - e^y \left(e^y + xe^y y'\right)}{\left(xe^y - 1\right)^2} = \frac{e^y \left(y' + e^y\right)}{\left(xe^y - 1\right)^2}$$

y sustituyendo el valor de y' resulta

$$y'' = \frac{e^{2y} \left(xe^y - 2\right)}{\left(xe^y - 1\right)^3}.$$

Nota 4.4. La generalización del teorema 2 a **funciones escalares** de varias variables, esto es, el caso de una ecuación de $n+1$ variables reales,

$$f\left(x_1, \ldots, x_n, y\right) = 0,$$

de la que se quiere despejar la variable y, es semejante al caso anterior, donde la hipótesis (4.1) vendrá dada por

$$\frac{\partial f}{\partial y}\left(a_1, \ldots, a_n, b\right) \neq 0. \tag{4.4}$$

La generalización del resto de hipótesis y tesis del teorema es inmediata. Además esta situación queda englobada en el teorema que enunciaremos a continuación para funciones vectoriales de varias variables.

Nota 4.5. La versión general del teorema 2 se aplica a **funciones vectoriales** de varias variables, esto es, el caso de un sistema de m ecuaciones, con $n+m$ $(m < n)$ incógnitas,

$$\begin{aligned} f_1\left(x_1, \ldots, x_n, y_1, \ldots, y_m\right) &= 0, \\ &\vdots \\ f_m\left(x_1, \ldots, x_n, y_1, \ldots, y_m\right) &= 0. \end{aligned} \tag{4.5}$$

Ahora se han de dar las condiciones de las funciones componente f_i para que las $(n+m)$ −tuplas que verifican el sistema vengan dadas por las n variables $x_1, \ldots, x_n$ elegidas arbitrariamente en un conjunto $I \subseteq \mathbb{R}^n$, y las m variables $y_1, \ldots, y_m$ estén definidas por m funciones g_i en la forma

$$y_1 = g_1\left(x_1, \ldots, x_n\right), \ \ldots, \ y_m = g_m\left(x_1, \ldots, x_n\right). \tag{4.6}$$

Si existen estas funciones, se dice que están definidas implícitamente por el sistema de ecuaciones dado.

Nota 4.6. Vectorialmente el sistema (4.5) se escribe

$$\mathbf{f}(\mathbf{x},\mathbf{y})=\mathbf{0} \text{ con } \mathbf{f}: A\subseteq\mathbb{R}^{n+m}\longrightarrow\mathbb{R}^m$$

y las funciones (4.6)

$$\mathbf{y}=\mathbf{g}(\mathbf{x}) \text{ con } \mathbf{g}: I\subseteq\mathbb{R}^{n}\longrightarrow\mathbb{R}^m,$$

donde $\mathbf{g}$ es la **función** (vectorial) **implícita**.

Teorema 3 (de la función implícita, caso general)**.** *Sea* $\mathbf{f}: A\subseteq\mathbb{R}^{n+m}\longrightarrow\mathbb{R}^m$*, con* A *conjunto abierto y* $\mathbf{f}\in C^1(A)$*. Sea un punto* $(\mathbf{a},\mathbf{b})\in A$ *(*$\mathbf{a}\in\mathbb{R}^n$*,* $\mathbf{b}\in\mathbb{R}^m$*) tal que* $\mathbf{f}(\mathbf{a},\mathbf{b})=\mathbf{0}$*. Si*

$$\det\frac{\partial(f_1,f_2,\ldots,f_m)}{\partial(y_1,y_2,\ldots,y_m)}(\mathbf{a},\mathbf{b})\neq 0, \tag{4.7}$$

entonces existen un conjunto abierto $I\subseteq\mathbb{R}^n$ *que contiene a* $\mathbf{a}$*, y una función única* $\mathbf{g}: I\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^m$*,* $\mathbf{g}\in C^1(I)$*, tal que*

$$\mathbf{g}(\mathbf{a})=\mathbf{b},\ \ \mathbf{f}(\mathbf{x},\mathbf{g}(\mathbf{x}))=\mathbf{0}\ \ \forall\mathbf{x}\in I,$$

y además

$$\begin{aligned}\frac{\partial\mathbf{g}}{\partial x_j}(\mathbf{x}) &= -\left[\frac{\partial(f_1,f_2,\ldots,f_m)}{\partial(y_1,y_2,\ldots,y_m)}(\mathbf{x},\mathbf{g}(\mathbf{x}))\right]^{-1}\left[\frac{\partial\mathbf{f}}{\partial x_j}(\mathbf{x},\mathbf{g}(\mathbf{x}))\right], \\ j &= 1,\ldots,n,\ \mathbf{x}\in I.\end{aligned} \tag{4.8}$$

Si la función $\mathbf{f}$ *es de clase* r *en* A*, la función* $\mathbf{g}$ *es también de clase* r *en* I*.*

Nota 4.7. Nótese que en la ecuación (4.7) se tiene el determinante de las m últimas columnas de la **matriz jacobiana** de $\mathbf{f}$ en el punto $(\mathbf{a},\mathbf{b})$. Se emplea esta notación porque explicita las variables involucradas. Igualmente, la expresión (4.8) es una ecuación matricial con las siguientes dimensiones

$$(m\times 1)=(m\times m)\times(m\times 1).$$

Nota 4.8. La ecuación (4.8) se obtiene anulando la diferencial de la función compuesta $\mathbf{f}(\mathbf{x},\mathbf{g}(\mathbf{x}))$, y procede de la aplicación de la regla de la cadena, a través del producto de las matrices jacobianas involucradas.

Nota 4.9. Usualmente, por abuso de notación, se emplea $\mathbf{y}=\mathbf{y}(\mathbf{x})$ para la función implícita y sus derivadas, en lugar de $\mathbf{g}$.

Nota 4.10. Si la función es escalar, $f : A \subseteq \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$, la ecuación (4.8) se escribe

$$\frac{\partial g}{\partial x_j}(\mathbf{x}) = -\frac{\dfrac{\partial f}{\partial x_j}(\mathbf{x}, g(\mathbf{x}))}{\dfrac{\partial f}{\partial y}(\mathbf{x}, g(\mathbf{x}))},\ j = 1, \ldots, n,\ \mathbf{x} \in I, \tag{4.9}$$

Ejemplo 28. Sea la función escalar

$$f(x_1, x_2, y) = \frac{x_1^2}{a^2 + y} + \frac{x_2^2}{b^2 + y} - 2,$$

con $a \neq 0$, $b \neq 0$, definida en $\mathbb{R}^3$ salvo para $y = -a^2$, $y = -b^2$. La función es de clase ∞ en su dominio. La ecuación

$$f(x_1, x_2, y) = 0$$

se cumple en el punto $(x_1, x_2, y) = (a, b, 0)$. Verificamos la condición (4.7), que se escribe como (4.4) para el caso de función escalar,

$$\frac{\partial f}{\partial y}(x_1, x_2, y) = -\frac{x_1^2}{(a^2 + y)^2} - \frac{x_2^2}{(b^2 + y)^2},$$

de donde

$$\frac{\partial f}{\partial y}(a, b, 0) = -\left(\frac{1}{a^2} + \frac{1}{b^2}\right) \neq 0.$$

Por lo tanto existe una función implícita $y = g(x_1, x_2)$, en un entorno de (a, b) tal que $0 = g(a, b)$. Obtenemos las derivadas parciales de g usando (4.8), esto es, (4.9) para el caso de función escalar:

$$\begin{aligned}
\frac{\partial g}{\partial x_1}(x_1, x_2) &= -\frac{\dfrac{\partial f}{\partial x_1}(x_1, x_2, y)}{\dfrac{\partial f}{\partial y}(x_1, x_2, y)} = \frac{\dfrac{2x_1}{a^2 + y}}{\dfrac{x_1^2}{(a^2 + y)^2} + \dfrac{x_2^2}{(b^2 + y)^2}}, \\
\frac{\partial g}{\partial x_2}(x_1, x_2) &= -\frac{\dfrac{\partial f}{\partial x_2}(x_1, x_2, y)}{\dfrac{\partial f}{\partial y}(x_1, x_2, y)} = \frac{\dfrac{2x_2}{b^2 + y}}{\dfrac{x_1^2}{(a^2 + y)^2} + \dfrac{x_2^2}{(b^2 + y)^2}}.
\end{aligned}$$

Si actuáramos con derivación implícita, mediante la aplicación de la regla de la cadena en la ecuación $f(x_1, x_2, y) = 0$, se tendría

$$\frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial x_i} = 0,\ \ i = 1, 2,$$

de donde despejando $\partial y/\partial x_i$ se obtienen las expresiones anteriores (recuérdese el empleo de la notación $y \equiv g$). Dado que g es una función escalar, su diferencial en los puntos (x_1, x_2) del entorno de (a, b) es

$$\begin{aligned} dg_{(x_1,x_2)}(h_1, h_2) &= \nabla g(x_1, x_2) \cdot (h_1, h_2) = \\ &= \left(\frac{x_1^2}{(a^2+y)^2} + \frac{x_2^2}{(b^2+y)^2}\right)^{-1} \left(\frac{2x_1}{a^2+y} h_1 + \frac{2x_2}{b^2+y} h_2\right). \end{aligned}$$

Ejemplo 29. Sea el sistema de ecuaciones

$$\begin{aligned} f_1(x, y, z) &= x^2 - y - 2z = 0, \\ f_2(x, y, z) &= 2x - y + z^2 = 0. \end{aligned}$$

Las funciones f_1 y f_2 son de clase ∞ en $\mathbb{R}^3$ (polinomios). Una solución del sistema es $(x, y, z) = (0, 0, 0)$. La condición (4.7) resulta

$$\det \frac{\partial (f_1, f_2)}{\partial (y, z)}(0,0,0) = \begin{vmatrix} \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial z} \\ \\ \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial z} \end{vmatrix}_{(0,0,0)} = \begin{vmatrix} -1 & -2 \\ -1 & 2z \end{vmatrix}_{(0,0,0)} = -2 \neq 0.$$

Existen por lo tanto dos funciones $y = g_1(x)$ y $z = g_2(x)$ que verifican el sistema en un entorno del punto $(0, 0, 0)$. Si realizamos derivación implícita en las ecuaciones de sistema, se tiene

$$\begin{aligned} 2x - y' - 2z' &= 0, \\ 2 - y' + 2zz' &= 0, \end{aligned}$$

de donde resolviendo en y' y z' obtenemos

$$y' = 2\frac{xz+1}{z+1}, \; z' = \frac{x-1}{z+1}.$$

Nota 4.11. En ejercicios como el anterior, suele ser más práctico el uso de la derivación implícita que la fórmula (4.8). No obstante, se puede comprobar que se obtiene el mismo resultado. En efecto,

$$\begin{aligned} \begin{pmatrix} y' \\ z' \end{pmatrix} &\equiv \begin{pmatrix} g_1'(x) \\ g_2'(x) \end{pmatrix} = -\left[\frac{\partial (f_1, f_2)}{\partial (y, z)}\right]^{-1} \begin{pmatrix} \partial f_1/\partial x \\ \partial f_2/\partial x \end{pmatrix} = \\ &= -\begin{pmatrix} -1 & -2 \\ -1 & 2z \end{pmatrix}^{-1} \begin{pmatrix} 2x \\ 2 \end{pmatrix} = -\frac{1}{z+1} \begin{pmatrix} -z & -1 \\ -1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 2x \\ 2 \end{pmatrix} = \\ &= \begin{pmatrix} 2\dfrac{xz+1}{z+1} \\ \\ \dfrac{x-1}{z+1} \end{pmatrix}. \end{aligned}$$

Ejemplo 30. Consideremos la ecuación $f\left(\sqrt{x\cos y}, x^2 z\right) = 0$ que define a z como función implícita de x e y de clase 1 en un cierto subconjunto de $\mathbb{R}^2$. Se pide obtener de la forma más simplificada posible la expresión

$$x\frac{\partial z}{\partial x} + \frac{\cos y}{\operatorname{sen} z}\frac{\partial z}{\partial y}.$$

La función implícita (cuya existencia impone el enunciado) viene definida por la ecuación $f(x, y, z) = 0$ donde f es cierta función no indicada. Esta función no depende de x, y, z de forma arbitraria, sino a través de las combinaciones $\sqrt{x\cos y}$ y $x^2 z$. Esto es, una composición de funciones donde f depende de dos variables u y v, que vienen dadas por $u(x, y) = \sqrt{x\cos y}$, y $v(x, z) = x^2 z$,

$$f\left(\sqrt{x\cos y}, x^2 z\right) = f\left(u(x, y), v(x, z)\right).$$

Para obtener sus derivadas aplicamos la regla de la cadena (véase la figura 4.2), que conduce a

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} = \frac{\partial f}{\partial u}\frac{\cos y}{2\sqrt{x\cos y}} + \frac{\partial f}{\partial v}2xz,\\
\frac{\partial f}{\partial y} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} = \frac{\partial f}{\partial u}\left(\frac{-x\operatorname{sen} y}{2\sqrt{x\cos y}}\right),\\
\frac{\partial f}{\partial z} &= \frac{\partial f}{\partial v}\frac{\partial v}{\partial z} = \frac{\partial f}{\partial v}x^2.
\end{aligned}$$

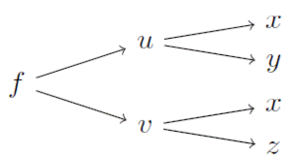


Figura 4.2: Árbol de dependencias del ejemplo 30.

En los puntos de no anulación de $\partial f/\partial z$, la expresión (4.9) nos permite escribir,

$$\frac{\partial z}{\partial x}(x, y) = -\frac{\dfrac{\partial f}{\partial x}(x, y, z(x, y))}{\dfrac{\partial f}{\partial z}(x, y, z(x, y))}, \quad \frac{\partial z}{\partial y}(x, y) = -\frac{\dfrac{\partial f}{\partial y}(x, y, z(x, y))}{\dfrac{\partial f}{\partial z}(x, y, z(x, y))},$$

donde, sustituyendo el valor de las derivadas calculadas anteriormente, se tiene

$$\begin{aligned}
\frac{\partial z}{\partial x}(x,y) &= -\frac{\dfrac{\partial f}{\partial u}\dfrac{\cos y}{2\sqrt{x\cos y}}+\dfrac{\partial f}{\partial v}2xz}{\dfrac{\partial f}{\partial v}x^2},\\
\frac{\partial z}{\partial y}(x,y) &= -\frac{\dfrac{\partial f}{\partial u}\left(\dfrac{-x\operatorname{sen} y}{2\sqrt{x\cos y}}\right)}{\dfrac{\partial f}{\partial v}x^2}.
\end{aligned}$$

Finalmente, tras la simplificación algebraica, la expresión del enunciado se puede escribir como

$$x\frac{\partial z}{\partial x}+\frac{\cos y}{\operatorname{sen} z}\frac{\partial z}{\partial y}=-2z.$$

4.2. Teorema de la función inversa

Nota 4.12. Supogamos una aplicación lineal $\mathbf{f}:\mathbb{R}^n\longrightarrow\mathbb{R}^m$, representada matricialmente en las bases canónicas de $\mathbb{R}^n$ y $\mathbb{R}^m$ por

$$\begin{pmatrix} y_1\\ \vdots\\ y_m\end{pmatrix}=\begin{pmatrix} a_{11} & \dots & a_{1n}\\ \vdots & & \vdots\\ a_{m1} & \dots & a_{mn}\end{pmatrix}\begin{pmatrix} x_1\\ \vdots\\ x_n\end{pmatrix},\quad \text{o equivalentemente, } \mathbf{y}=A\mathbf{x}.$$

La condición para que $\mathbf{f}$ sea **invertible** (o inversible) es de naturaleza algebraica y está resuelta por los teoremas de Cramer y Rouché-Frobenius: la aplicación $\mathbf{f}$ es invertible si, y sólo si, $m=n$ y además $\det A\neq 0$.

Sabemos también que las aplicaciones lineales son continuas en su dominio. Recordemos que la diferencial de una función en un punto es una aplicación lineal que tiene asociada la *matriz jacobiana* de la función en dicho punto.

Definición 4.1. Diremos que un campo vectorial $\mathbf{f}:A\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^n$, es **globalmente invertible** si es una aplicación inyectiva. En ese caso, existe una función **inversa** $\mathbf{f}^{-1}:\mathbf{f}(A)\longrightarrow\mathbb{R}^n$ tal que $\mathbf{f}^{-1}(\mathbf{y})=\mathbf{x}$ si, y sólo si, $\mathbf{f}(\mathbf{x})=\mathbf{y}$.

Ejemplo 31. Las aplicaciones lineales $\mathbf{f}:\mathbb{R}^n\longrightarrow\mathbb{R}^n$, $\mathbf{f}(\mathbf{x})=A\mathbf{x}$ con $\det A\neq 0$ son globalmente invertibles. Si $\mathbf{f}(\mathbf{x})=A\mathbf{x}$, entonces $\mathbf{f}^{-1}(\mathbf{y})=A^{-1}\mathbf{y}$.

Definición 4.2. Diremos que $\mathbf{f}: A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$, es **localmente invertible** en $\mathbf{a} \in A$ si existe un conjunto abierto $U \subseteq A$ tal que $\mathbf{a} \in U$ y la restricción de $\mathbf{f}$ a U, $\mathbf{f}_U : U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$, es globalmente invertible. La función (inversa) $\mathbf{f}_U^{-1}$ se denomina **inversa local** de $\mathbf{f}$.

Ejemplo 32. Sea $\mathbf{f} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $\mathbf{f}(x, y) = (e^x \cos y, e^x \operatorname{sen} y)$, $\mathbf{f}$ no es inyectiva en $\mathbb{R}^2$ ya que $(x, y) \neq (x, y + 2k\pi)$ con $k \in \mathbb{Z}$, pero $\mathbf{f}(x, y) = \mathbf{f}(x, y + 2k\pi)$. Por lo tanto no es globalmente invertible. Sin embargo, en todo entorno $B[(x, y), \delta]$ con radio $\delta < \pi$, $\mathbf{f}$ es inyectiva y, por tanto, localmente invertible en cada punto de $\mathbb{R}^2$.

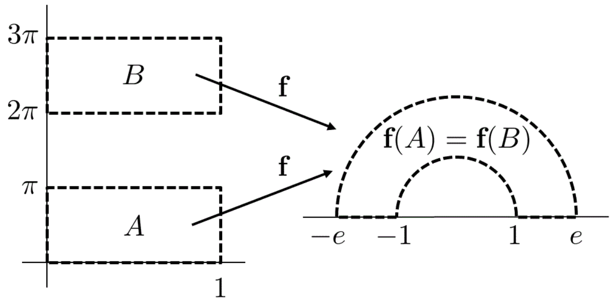


Figura 4.3: Esquema de la inversión local en el ejemplo 32.

Nota 4.13. En el siguiente teorema veremos que la inyectividad con carácter local de una función en el entorno de un punto, viene asegurada si el determinante de la matriz jacobiana, denominado **determinante jacobiano**, en el punto es no nulo.

Esto es fácil de ver en el caso particular de funciones reales de una variable. En efecto, supongamos $f : A \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ con A conjunto abierto, $f \in C^1(A)$. Si en un punto $a \in A$, se tiene que $f'(a) \neq 0$, como f' es continua en a, f' conserva el signo (positivo o negativo) en un entorno de a, por lo tanto f es estrictamente monótona (creciente o decreciente) en el mismo entorno, y necesariamente inyectiva en él.

Nota 4.14. La generalización a funciones vectoriales de varias variables no es trivial. No obstante, la idea subyacente es que si una función vectorial de varias variables $\mathbf{f}$ es diferenciable en un punto $\mathbf{a}$, entonces en un entorno $B(\mathbf{a}, \delta)$ con δ suficientemente pequeño, se puede aproximar por la diferencial en el punto, esto es, con $\|\mathbf{h}\| < \delta$,

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) \simeq d\mathbf{f}_{\mathbf{a}}(\mathbf{h}),$$

y sabemos que $d\mathbf{f}_{\mathbf{a}}(\mathbf{h})$ es una aplicación lineal. Cabe esperar que si $d\mathbf{f}_{\mathbf{a}}(\mathbf{h})$ es localmente invertible, también lo sea $\mathbf{f}$. Las aplicaciones lineales son invertibles si, y sólo si, el determinante de su matriz asociada (el determinante jacobiano) es no nulo, esto es $\det J\mathbf{f}(\mathbf{a}) \neq 0$.

Teorema 4 (de la función inversa)**.** *Sea* $\mathbf{f} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$*, con* A *conjunto abierto,* $\mathbf{f} \in C^r(A)$*, y sea* $\mathbf{a} \in A$*. Si* $\det J\mathbf{f}(\mathbf{a}) \neq 0$*, entonces* $\mathbf{f}$ *es localmente invertible de clase* r *en* $\mathbf{a}$*. Además, si* $\mathbf{f}^{-1}$ *es la inversa local de* $\mathbf{f}$*, se tiene*

$$d\mathbf{f}_{\mathbf{y}}^{-1} = (d\mathbf{f}_{\mathbf{x}})^{-1}, \text{ con } \mathbf{x} = \mathbf{f}^{-1}(\mathbf{y}). \tag{4.10}$$

Nota 4.15. La expresión (4.10) en términos de matrices jacobianas se escribe

$$J\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x})) = [J\mathbf{f}(\mathbf{x})]^{-1}. \tag{4.11}$$

Nota 4.16. Tanto (4.10) como (4.11) proceden de aplicar la regla de la cadena a la composición de funciones, $\mathbf{f}^{-1} \circ \mathbf{f} = \mathbf{i}$, siendo $\mathbf{i}$ la función identidad, $\mathbf{i}(\mathbf{x}) = \mathbf{x}$. Esto es, $d\mathbf{f}_{\mathbf{f}(\mathbf{x})}^{-1} \circ d\mathbf{f}_{\mathbf{x}} = d\mathbf{i}_{\mathbf{x}}$.

Nota 4.17. Para $n = 1$, la ecuación (4.11) recupera la conocida expresión para funciones reales de una variable, esto es,

$$\left(f^{-1}\right)'(f(x)) = \frac{1}{f'(x)}, \text{ con } f'(x) \neq 0. \tag{4.12}$$

Nota 4.18. Pese a la notación empleada no hay que olvidar que $\mathbf{f}^{-1}$ hace referencia a la *inversa local*, sólo está definida en un entorno del punto $\mathbf{a}$. Aunque el teorema de la función inversa no nos permite conocer una inversa local de la función $\mathbf{f}$, sí permite calcular la matriz jacobiana de $\mathbf{f}^{-1}$ en $\mathbf{f}(\mathbf{a})$, siempre y cuando conozcamos la matriz jacobiana de $\mathbf{f}$ en $\mathbf{a}$.

Ejemplo 33. En la función del ejemplo 32, $\mathbf{f}(x,y) = (e^x \cos y, e^x \operatorname{sen} y)$, de clase ∞ en $\mathbb{R}^2$, se tiene

$$\det J\mathbf{f}(x,y) = \begin{vmatrix} e^x \cos y & -e^x \operatorname{sen} y \\ e^x \operatorname{sen} y & e^x \cos y \end{vmatrix} = e^{2x}\left(\cos^2 y + \operatorname{sen}^2 y\right) = e^{2x} \neq 0.$$

Dado que el jacobiano no se anula nunca, $\mathbf{f}$ es localmente invertible y su inversa local es de clase ∞ en $\mathbb{R}^2$.
Si restringimos el dominio al conjunto abierto $A = \{(x,y) \in \mathbb{R}^2,\ y \in]0, 2\pi[\}$, la inversa es global y de hecho la podemos calcular. En efecto, tomando

$$e^x \cos y = u,\ e^x \operatorname{sen} y = v,$$

elevando al cuadrado y sumando ambas ecuaciones tenemos

$$e^{2x}\left(\cos^2 y + \operatorname{sen}^2 y\right) = u^2 + v^2 \to x = \frac{1}{2}\log\left(u^2+v^2\right), \ (u,v) \neq (0,0).$$

Si dividimos ambas ecuaciones entre sí se tiene

$$\frac{v}{u} = \tan y \to y = \arctan\left(\frac{v}{u}\right), \ u > 0, \ v \geq 0$$

(nos restringimos al primer cuadrante, para simplificar la expresión de y, sin pérdida de generalidad en lo que queremos ilustrar). Por lo tanto podemos definir

$$\mathbf{f}^{-1}(u,v) = (x(u,v), y(u,v)) = \left(\frac{1}{2}\log\left(u^2+v^2\right), \arctan\left(\frac{v}{u}\right)\right).$$

Se puede comprobar entonces que (ejercicio)

$$J\mathbf{f}^{-1}(u(x,y), v(x,y)) = \left[J\mathbf{f}(x,y)\right]^{-1} = \begin{pmatrix} e^{-x}\cos y & e^{-x}\operatorname{sen} y \\ -e^{-x}\operatorname{sen} y & e^{-x}\cos y \end{pmatrix}.$$

Ejemplo 34. Supongamos $\mathbf{f}:\mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $\mathbf{f}(x,y) = \left(-3x + y^3, -3y + x^3\right)$ y calculamos el jacobiano en el punto $(1,1)$,

$$\det J\mathbf{f}(1,1) = \left|\begin{matrix} -3 & 3y^2 \\ 3x^2 & -3 \end{matrix}\right|_{(1,1)} = 9\left(1 - x^2y^2\right)\Big|_{(1,1)} = 0.$$

Veamos que, en efecto, $\mathbf{f}$ no es inyectiva en un entorno de $(1,1)$. Si nos aproximamos a $(1,1)$ con la recta $y = x$, tenemos la función restringida de una variable dada por

$$\mathbf{f}_S:\mathbb{R} \longrightarrow \mathbb{R}, \ \mathbf{f}_S(x) = \left(-3x + x^3, -3x + x^3\right),$$

Se comprueba que esta función tiene un mínimo relativo en $x = 1$. Por lo tanto, siempre existirán dos puntos x_1 y x_2 tan próximos como se quiera a $x = 1$ tales que $x_1 \neq x_2$ pero $\mathbf{f}_S(x_1) = \mathbf{f}_S(x_2)$, o lo que es lo mismo, $\mathbf{f}(x_1,x_1) = \mathbf{f}(x_2,x_2)$. Por lo tanto $\mathbf{f}$ no es inyectiva en un entorno de $(1,1)$.

Ejemplo 35. Sea $f:\mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = x^3$. Aquí, $\det f'(x) = f'(x) = 3x^2 \neq 0$ si $x \neq 0$. Por lo tanto en $x = 0$ no se verifica que el jacobiano sea no nulo, hipótesis del teorema 4. Sin embargo, existe la función inversa $f^{-1}(y) = \sqrt[3]{y}$, que está definida en todo $\mathbb{R}$, y en particular, $f^{-1}(0) = 0$. Es decir, la condición de jacobiano no nulo es condición suficiente, pero no necesaria.

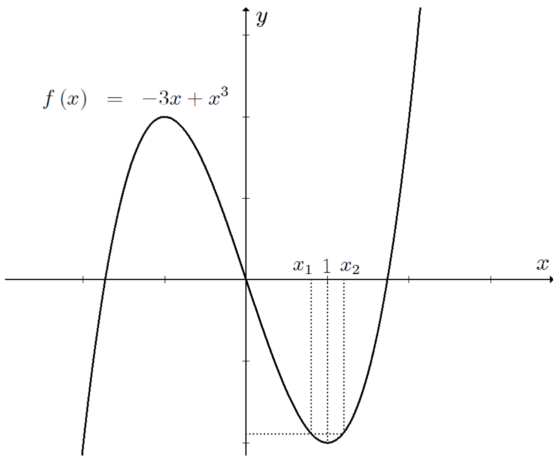


Figura 4.4: Función de una variable del ejemplo 34.

4.2.1. Cambios de variable

Definición 4.3. Sean $\mathbf{h} : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$, con A conjunto abierto, y $B = \mathbf{h}(A) \subseteq \mathbb{R}^n$ (conjunto imagen de $\mathbf{h}$). Se dice que $\mathbf{h}$ es un **cambio de variable**, o difeomorfismo, de clase 1 en A, si $\mathbf{h}$ es biyectiva y tanto $\mathbf{h}$ como su función inversa $\mathbf{h}^{-1}$ son funciones de clase 1 en A y B respectivamente.

Nota 4.19. El teorema de la función inversa da las condiciones para que una función defina un **cambio de variable local**, esto es, un cambio de variable en un entorno de un punto $\mathbf{a}$, $B(\mathbf{a}, \delta)$, donde la función es de clase 1 y tal que $\det J\mathbf{h}(\mathbf{a}) \neq 0$.

Nota 4.20. Nótese que el papel de $\mathbf{h}$ y $\mathbf{h}^{-1}$ es intercambiable. Es decir, si $\mathbf{h}^{-1}$ es de clase C^1 en un entorno de $\mathbf{h}(\mathbf{a})$ y su jacobiano no se anula en $\mathbf{h}(\mathbf{a})$, la función $\mathbf{h}^{-1}$ define un cambio de variable local en un entorno abierto de $\mathbf{h}(\mathbf{a})$. Además, en virtud de (4.11), los jacobianos se relacionan según:

$$\det J\mathbf{h}^{-1}(\mathbf{h}(\mathbf{a})) = \frac{1}{\det J\mathbf{h}(\mathbf{a})}. \tag{4.13}$$

Nota 4.21. Cuando trabajamos con cambios de variables, supongamos de dos variables sin perdida de generalidad, escritos en la forma

$$x = x(u, v), \; y = y(u, v),$$

tales expresiones definen la función $\mathbf{h}^{-1}$, esto es, sus componentes son

$$x(u, v) = h_1^{-1}(u, v), \; y(u, v) = h_2^{-1}(u, v).$$

Por ejemplo, la relación entre determinantes jacobianos (4.13) se escribiría:

$$\det J\mathbf{h}^{-1}(u,v) = \left.\frac{1}{\det J\mathbf{h}(x,y)}\right|_{\substack{x=x(u,v)\\ y=y(u,v)}}. \tag{4.14}$$

Ejemplo 36. Consideramos las ecuaciones

$$x = e^u \cos v,\ y = e^u \operatorname{sen} v.$$

Nos preguntamos en qué puntos definen localmente un cambio de variable, y se pide expresar en términos de las variables u y v la ecuación

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

Las expresiones anteriores definen (nota 4.21) la función $\mathbf{h}^{-1}$ cuyas componentes, $h_1^{-1}(u,v)$ y $h_2^{-1}(u,v)$ se denotan como $x(u,v)$ e $y(u,v)$, esto es,

$$x(u,v) = e^u \cos v,\ y(u,v) = e^u \operatorname{sen} v.$$

Estas funciones están definidas en $\mathbb{R}^2$ y son de clase 1 en su dominio. Para que definan un cambio de variable local, es preciso que no se anule el jacobiano,

$$\det J\mathbf{h}^{-1}(u,v) = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} e^u \cos v & -e^u \operatorname{sen} v \\ e^u \operatorname{sen} v & e^u \cos v \end{vmatrix} = e^{2u}.$$

Por lo tanto el jacobiano es no nulo en todo punto de $\mathbb{R}^2$, y las ecuaciones dadas definen localmente un cambio de variable en los entornos de estos puntos.

Para transformar la ecuación que se pide, aplicamos la regla de la cadena en la composición de funciones

$$f(x(u,v), y(u,v)) = f(e^u \cos v, e^u \operatorname{sen} v),$$

esto es,

$$\begin{aligned}
\frac{\partial f}{\partial u} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} = \frac{\partial f}{\partial x}e^u \cos v + \frac{\partial f}{\partial y}e^u \operatorname{sen} v,\\
\frac{\partial f}{\partial v} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v} = \frac{\partial f}{\partial x}(-e^u \operatorname{sen} v) + \frac{\partial f}{\partial y}e^u \cos v.
\end{aligned}$$

No obstante, dado que la expresión que se desea transformar depende de las derivadas parciales $\partial f/\partial x$ y $\partial f/\partial y$, es necesario despejar éstas de las ecuaciones anteriores. El sistema se puede tratar algebraicamente, por ejemplo, mediante la regla de Cramer,

$$\begin{pmatrix} e^u \cos v & e^u \operatorname{sen} v \\ -e^u \operatorname{sen} v & e^u \cos v \end{pmatrix} \begin{pmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f}{\partial u} \\ \dfrac{\partial f}{\partial v} \end{pmatrix}.$$

El determinante de la matriz del sistema es e^{2u}, y por lo tanto,

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{1}{e^{2u}} \begin{vmatrix} \dfrac{\partial f}{\partial u} & e^u \operatorname{sen} v \\ \dfrac{\partial f}{\partial v} & e^u \cos v \end{vmatrix} = e^{-u} \left(\cos v \frac{\partial f}{\partial u} - \operatorname{sen} v \frac{\partial f}{\partial v} \right), \\ \frac{\partial f}{\partial y} &= \frac{1}{e^{2u}} \begin{vmatrix} e^u \cos v & \dfrac{\partial f}{\partial u} \\ -e^u \operatorname{sen} v & \dfrac{\partial f}{\partial v} \quad e^u \cos v \end{vmatrix} = e^{-u} \left(\cos v \frac{\partial f}{\partial v} + \operatorname{sen} v \frac{\partial f}{\partial u} \right). \end{aligned}$$

Sustituyendo estas derivadas en la expresión inicial, así como el cambio de variable en x e y, y tras la simplificación (ejercicio) resulta:

$$\frac{1}{x^2 + y^2} \left(x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} \right)^2 = 1 \to \left(\frac{\partial f}{\partial u} \right)^2 = e^{2u}.$$

Nota 4.22. En problemas como el anterior, en ocasiones puede resultar más directo obtener **h**, esto es, depejar las variables u y v de las ecuaciones para $x\,(u, v)$ e $y\,(u, v)$, de esta forma obtener directamente el valor de las derivadas $\partial f/\partial x$ y $\partial f/\partial y$ aplicando la regla de la cadena a la composición de funciones

$$f\left(u\,(x, y)\,, v\,(x, y)\right),$$

esto es,

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x}, \quad \frac{\partial f}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y}.$$

En el ejemplo anterior, invertir el sistema de ecuaciones se hace de forma análoga al caso del ejemplo 33, es decir, calculando $x^2 + y^2$,

$$x^2 + y^2 = e^{2u} \left(\cos^2 u + \cos^2 y \right) \to u = \frac{1}{2} \log \left(x^2 + y^2 \right), \quad (x, y) \neq (0, 0)\,,$$

y el cociente y/x,

$$\frac{y}{x} = \tan v \to v = \arctan\left(\frac{y}{x}\right), \quad x > 0, \; y \geq 0.$$

Por lo tanto,

$$f\left(u\left(x,y\right), v\left(x,y\right)\right) = f\left(\frac{1}{2}\log\left(x^2+y^2\right), \arctan\left(\frac{y}{x}\right)\right).$$

La comprobación de que se obtiene el mismo resultado queda como ejercicio.

4.2.1.1. Coordenadas polares

Nota 4.23. Recordemos que las **coordenadas polares** (r,θ) de un punto $(x,y) \in \mathbb{R}^2$ se definen por las relaciones con las coordenadas cartesianas

$$x = r\cos\theta, \;\; y = r\,\mathrm{sen}\,\theta, \tag{4.15}$$

con $r \geq 0$ y $\theta \in [0,2\pi[$. Estas relaciones definen un *cambio de variable* en el sentido de la definición 4.3 en virtud del resultado siguiente.

Proposición 4.1. La función $\mathbf{h} : A \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, donde A es el conjunto abierto

$$A = \left\{(r,\theta), \, r > 0, \, \theta \in \,]0,2\pi[\right\}$$

dada por

$$\mathbf{h}(r,\theta) = (r\cos\theta, r\,\mathrm{sen}\,\theta),$$

define un cambio de variable de clase ∞ en A.

Nota 4.24. En efecto, las componentes de $\mathbf{h}$ son funciones de clase ∞ en A. Verificamos que $\mathbf{h}$ es *inyectiva*. Para ello debe suceder que

$$\mathbf{h}(r_1,\theta_1) = \mathbf{h}(r_2,\theta_2) \to (r_1,\theta_1) = (r_2,\theta_2).$$

Si consideramos componente a componente

$$\begin{aligned} h_1(r_1,\theta_1) &= h_1(r_2,\theta_2) \to r_1\cos\theta_1 = r_2\cos\theta_2, \\ h_2(r_1,\theta_1) &= h_2(r_2,\theta_2) \to r_1\,\mathrm{sen}\,\theta_1 = r_2\,\mathrm{sen}\,\theta_2. \end{aligned}$$

Elevando ambas ecuaciones al cuadrado y sumando se tiene para la coordenada radial,

$$r_1^2 = r_2^2 \to |r_1| = |r_2| \to r_1 = r_2,$$

y para la angular,

$$\cos\theta_1 = \cos\theta_2, \quad \operatorname{sen}\theta_1 = \operatorname{sen}\theta_2. \tag{4.16}$$

Considerando la igualdad trigonométrica

$$\cos(\theta_1 - \theta_2) = \cos\theta_1 \cos\theta_2 + \operatorname{sen}\theta_1 \operatorname{sen}\theta_2,$$

combinada con (4.16) obtenemos

$$\cos(\theta_1 - \theta_2) = \cos^2\theta_1 + \operatorname{sen}^2\theta_1 = 1,$$

es decir, $\theta_1 - \theta_2 = 2k\pi$, $k \in \mathbb{Z}$. Dado que $\theta_1, \theta_2 \in \,]0, 2\pi[$, necesariamente $k = 0$, o equivalentemente, $\theta_1 = \theta_2$. Entonces $(r_1, \theta_1) = (r_2, \theta_2)$, con lo que $\mathbf{h}$ es una función inyectiva en A.

La biyectividad se garantiza entonces definiendo $\mathbf{h} : A \longrightarrow \operatorname{Im}(\mathbf{h})$, donde el conjunto imagen $\mathbf{h}$ es

$$\operatorname{im}(\mathbf{h}) = \mathbf{h}(A) = \mathbb{R}^2 \setminus \{(x, 0)\,,\, x \geq 0\}\,,$$

esto es, el plano $\mathbb{R}^2$ sin el semieje x positivo. Sin embargo, si añadimos al conjunto A el conjunto de puntos del semieje x positivo (salvo el origen), es decir, definimos $\mathbf{h}$ en el dominio

$$\tilde{A} = A \cup \{(r, 0)\,,\, r > 0\} = \{(r, \theta)\,,\, r > 0,\, \theta \in [0, 2\pi[\}\,,$$

la inyectividad de $\mathbf{h}$ se mantiene en $\tilde{A}$. Por lo tanto, la función

$$\mathbf{h} : \tilde{A} \longrightarrow \mathbb{R}^2 \setminus \{(0, 0)\}$$

es una aplicación *biyectiva.*

La función $\mathbf{h}$ es globalmente invertible, obtengamos $\mathbf{h}^{-1}$. Partiendo de

$$x = r\cos\theta, \;\; y = r\operatorname{sen}\theta, \quad (x, y) \in \mathbb{R}^2 \setminus \{(x, 0)\,,\, x \geq 0\} \tag{4.17}$$

Elevando al cuadrado las ecuaciones (4.17) y sumando se llega a

$$x^2 + y^2 = r^2 \overset{r>0}{\to} r = \sqrt{x^2 + y^2}.$$

Dividiendo las ecuaciones (4.17) se tiene

$$\tan\theta = \frac{y}{x} \;\; (x \neq 0).$$

Para obtener $\theta \in [0, 2\pi[$ recordamos que $\arctan \theta$ tiene por imagen el intervalo $]-\pi/2, \pi/2[$ (problema estudiado en los temas de variable compleja) de manera que el valor de θ viene dado por

$$\theta(x,y) = \begin{cases} \arctan\left(\dfrac{y}{x}\right) & \text{si} \quad x > 0, y \geq 0 \\ \pi + \arctan\left(\dfrac{y}{x}\right) & \text{si} \quad x < 0 \\ 2\pi + \arctan\left(\dfrac{y}{x}\right) & \text{si} \quad x > 0, y < 0 \end{cases}, \tag{4.18}$$

junto a $\theta = \pi/2$ si $x = 0, y > 0$, y $\theta = 3\pi/2$ si $x = 0, y < 0$. Finalmente, si $x = y = 0$ (origen), θ no está definido.
Dado que la función $\arctan x$ es de clase ∞ en $\mathbb{R}$,

$$\mathbf{h}^{-1}(x,y) = \left(\sqrt{x^2+y^2}, \theta(x,y)\right)$$

es de clase ∞ en $\mathbb{R}^2 \backslash \{(0,0)\}$, por lo tanto $\mathbf{h}$ define un cambio de variable de clase ∞ en $\tilde{A}$.

Nota 4.25. Otra forma de obtener $\mathbf{h}^{-1}$ es la siguiente. Dado que $\tan \alpha$ es inyectiva en $\alpha \in]-\pi/2, \pi/2[$, si tomamos el ángulo polar θ definido en el intervalo $]0, 2\pi[$, entonces $(\pi - \theta)/2$ está definido en $]-\pi/2, \pi/2[$ donde la función tangente es inyectiva. Podemos relacionar este ángulo con las variables cartesianas mediante manipulación trigonométrica:

$$\begin{aligned} y &= r \operatorname{sen} \theta = r \operatorname{sen}(\pi - \theta) \overset{[1]}{=} r 2 \operatorname{sen}\left(\frac{\pi-\theta}{2}\right) \cos\left(\frac{\pi-\theta}{2}\right) = \\ &= 2r \tan\left(\frac{\pi-\theta}{2}\right) \cos^2\left(\frac{\pi-\theta}{2}\right) \overset{[2]}{=} r \tan\left(\frac{\pi-\theta}{2}\right) [1 + \cos(\pi - \theta)] = \\ &= r \tan\left(\frac{\pi-\theta}{2}\right) (1 - \cos\theta) \overset{[3]}{=} \tan\left(\frac{\pi-\theta}{2}\right) (r - x) = \\ &= \tan\left(\frac{\pi-\theta}{2}\right) \left(\sqrt{x^2+y^2} - x\right), \end{aligned}$$

donde en [1] se ha usado que $\operatorname{sen} 2\alpha = 2 \operatorname{sen} \alpha \cos \alpha$, en [2] $\cos^2 \alpha = (1 + \cos 2\alpha)/2$, y en [3] $x = r \cos \theta$. Despejando,

$$\tan\left(\frac{\pi-\theta}{2}\right) = \frac{y}{\sqrt{x^2+y^2} - x} \to \theta = \pi - 2 \arctan \frac{y}{\sqrt{x^2+y^2} - x}$$

(expresión no definida en el origen). Por lo tanto, tenemos una expresión más compacta para $\mathbf{h}^{-1}$:

$$\mathbf{h}^{-1}(x,y) = \left(\sqrt{x^2+y^2}, \pi - 2\arctan\frac{y}{\sqrt{x^2+y^2}-x}\right). \tag{4.19}$$

Nota 4.26. Para todo punto $(r,\theta) \in A$,

$$\det J\mathbf{h}(r,\theta) = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\operatorname{sen}\theta \\ \operatorname{sen}\theta & r\cos\theta \end{vmatrix} = r > 0, \tag{4.20}$$

por lo tanto $\mathbf{h}$ es un cambio de variable local en A. Dado que sabemos que $\mathbf{h}$ es globalmente invertible, este aspecto del teorema de la función inversa no nos aporta más información, pero sí la relación (4.11) entre matrices jacobianas, que nos permite calcular,

$$\begin{aligned} J\mathbf{h}^{-1}(x,y) &= [J\mathbf{h}(r,\theta)]^{-1} = \begin{pmatrix} \cos\theta & -r\operatorname{sen}\theta \\ \operatorname{sen}\theta & r\cos\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ -\dfrac{1}{r}\operatorname{sen}\theta & \dfrac{1}{r}\cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{x}{\sqrt{x^2+y^2}} & \dfrac{y}{\sqrt{x^2+y^2}} \\ \\ -\dfrac{y}{x^2+y^2} & \dfrac{x}{x^2+y^2} \end{pmatrix}. \end{aligned}$$

Este resultado se puede verificar con el cálculo directo de la matriz jacobiana a partir de la expresión (4.19). Nótese que

$$\det J\mathbf{h}^{-1}(x,y) = \frac{1}{\sqrt{x^2+y^2}} = \frac{1}{r}, \tag{4.21}$$

tal y como se puede deducir directamente de la ecuación (4.14).

Ejercicios (parte III)

Ejercicio 1. Dadas las siguientes ecuaciones, y utilizando el teorema de la función implícita, hallar los puntos en los que la ecuación define a y como función implícita de x, $y = y(x)$, y calcular en ellos $y'(x)$.

a) $x^2 + y^2 = 1$.

b) $y^2(a - x) = x^2(a + x)$, $a > 0$.

Sol.: a) Todos los puntos de $\mathbb{R}^2$ salvo $(\pm 1, 0)$, $y'(x) = -x/y$. b) Todos los puntos de $\mathbb{R}^2$ salvo $(-a, 0)$, $(0, 0)$, $y'(x) = \left(y^2 + 2ax + 3x^2\right) / [2y(a - x)]$.

Ejercicio 2. Considerando que las siguientes ecuaciones definen a z como función implícita de x e y, $z = z(x, y)$, se pide simplificar las operaciones que se indica:

a) $yz\left(x^2 + y^2\right) + \arctan(yz) = 0$, $\dfrac{y}{z}\dfrac{\partial z}{\partial x} - \dfrac{\partial z}{\partial y}$.

b) $f\left(\sqrt{x\cos y}, x^2 z\right) = 0$, $x\dfrac{\partial z}{\partial x} + \dfrac{\cos y}{\operatorname{sen} y}\dfrac{\partial z}{\partial y}$.

Sol.: a) $\frac{y}{z}\frac{\partial z}{\partial x} - \frac{\partial z}{\partial y} = \frac{z}{y}$. b) $x\frac{\partial z}{\partial x} + \frac{\cos y}{\operatorname{sen} y}\frac{\partial z}{\partial y} = -2z$.

Ejercicio 3. Probar que la ecuación $x^2 + y^4 - 5 = 0$ define una variable como función implícita de la otra en un entorno del punto $(2, 1)$. a) Si $y = g(x)$ es una de estas funciones implícitas, hallar $g'(2)$ y $g''(2)$. b) ¿En qué puntos no puede expresarse una variable en función de la otra?

Sol.: a) $g'(2) = -1$, $g''(2) = -7/2$. b) En los puntos $\left(\pm\sqrt{5}, 0\right)$ y $\left(0, \pm\sqrt[4]{5}\right)$.

Ejercicio 4. Obtener los polinomios de Taylor de primer grado de la función $z(x, y)$ definida implícitamente por las siguientes ecuaciones, en un entorno de los puntos que se indica:

a) $2x^2 + 2y^2 + z^2 - 8xz - z + 8 = 0$, $(x_0, y_0, z_0) = (2, 0, 1)$.

b) $e^{x+z} + \cos(y + z) = 0$, $(x_0, y_0, z_0) = (0, \pi, 0)$.

Sol.: a) $P_{1,(2,0)}(x, y) = 1$. b) $P_{1,(0,\pi)}(x, y) = -x$.

Ejercicio 5. Hallar la recta tangente a la gráfica de $y = y(x)$, definida implícitamente por $x^2y - xy^2 + y^5 = 3$, en el punto $(x_0, y_0) = (-1, 1)$.

Sol.: $3x - 8y + 11 = 0$.

Ejercicio 6. Dada $f(x, y, z) = 3xy^2 + 12yz - 4z^2x^3 = 0$ se pide: a) Obtener el plano tangente a la gráfica de $y = y(x, z)$ definida implícitamente por $f(x, y, z) = 3xy^2 + 12yz - 4z^2x^3 = 0$, en el punto $(x_0, y_0, z_0) = (0, 0, 1)$. b) Comprobar que en ese punto $y = y(x, z)$ es la única función definida implícitamente por la ecuación.

Sol.: a) $y = 0$. $\frac{\partial f}{\partial x}(0, 0, 1) = 0$. b) $\frac{\partial f}{\partial y}(0, 0, 1) \neq 0$, $\frac{\partial f}{\partial z}(0, 0, 1) = 0$.

Ejercicio 7. Dada la ecuación

$$\operatorname{sen}(x + y + z) = e^{2y} - 1,$$

se pide a) demostrar que define una única función implícita $z = z(x, y)$ de clase ∞ en un entorno de $(1, 0)$ tal que $z(1, 0) = -1$. b) Obtener la ecuación del plano tangente a la superficie $z = z(x, y)$ en el punto $(1, 0)$. c) Obtener el polinomio de Taylor de grado 2 de la función $z(x, y)$ en el punto $(1, 0)$.

Sol.: a) Teorema de la función implícita, $\frac{\partial f}{\partial z}(1, 0, -1) \neq 0$. b) $x - y + z = 0$. c) $P_{2,(1,0)}(x, y) = -x + y + 2y^2$.

Ejercicio 8. Dada la ecuación

$$x \operatorname{sen} \frac{\pi}{y} - z^2 y + 1 = 0,$$

se pide a) demostrar que define una única función implícita $z = z(x, y)$ de clase ∞ en un entorno de $(1/\pi, 1)$ tal que $z(1/\pi, 1) = 1$. b) Demostrar que $(1/\pi, 1)$ es un punto crítico de la función z y deducir como consecuencia la ecuación del plano tangente a $z = z(x, y)$ en dicho punto. c) Demostrar que $D_{\mathbf{v}} z(1/\pi, 1)$ no depende del vector $\mathbf{v}$.

Sol.: a) Teorema de la función implícita, $\frac{\partial f}{\partial z}(1/\pi, 1, 1) \neq 0$. b) $\nabla z(1/\pi, 1) = \mathbf{0}$, $z = 1$. c) $D_{\mathbf{v}} z(1/\pi, 1) = 0\ \forall \mathbf{v} \neq \mathbf{0}$.

Ejercicio 9. Probar que el sistema de ecuaciones

$$\begin{cases} xz^2 + y^3u^2 = 4 \\ xy^2 + u^2z = 0 \end{cases}$$

define a x e y como funciones implícitas de z y u, $x = g_1(z,u)$, $y = g_2(z,u)$, en un entorno del punto $(x_0, y_0, z_0, u_0) = (0,1,0,2)$. Sea $\mathbf{g} = (g_1, g_2) : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, calcular $J\mathbf{g}(0,2)$..

Sol.: Teorema de la función implícita, det $\frac{\partial(f_1,f_2)}{\partial(x,y)}(0,1,0,2) = -12 \neq 0$; $J\mathbf{g}(0,2) = -\frac{1}{3}\begin{pmatrix} 12 & 0 \\ 0 & 1 \end{pmatrix}$.

Ejercicio 10. Comprobar que el sistema

$$\begin{cases} x \operatorname{sen} y + y \operatorname{sen} z + z \operatorname{sen} x = 0 \\ xy + yz + zx = -\pi^2 \end{cases}$$

define a y y z como funciones implícitas de x, $y = g_1(x)$, $z = g_2(x)$, en un entorno del punto $(x_0, y_0, z_0) = (\pi, \pi, -\pi)$. Calcular $g_1'(\pi)$ y $g_2'(\pi)$.

Sol.: Teorema de la función implícita, det $\frac{\partial(f_1,f_2)}{\partial(y,z)}(\pi,\pi,-\pi) = -2\pi^2 \neq 0$; $g_1'(\pi) = 1$, $g_2'(\pi) = 0$.

Ejercicio 11. Dado el sistema de ecuaciones

$$\begin{cases} x + y + t = 0 \\ xyt + \operatorname{sen}(xyt) = 0 \end{cases},$$

se pide: a) Demostrar que define las ecuaciones paramétricas $x = x(t)$, $y = y(t)$ de una curva plana en un entorno del punto $(-1,0)$ correspondiente a $t = 1$. b) Hallar el vector velocidad, $(x'(t), y'(t))$ en el punto $(-1,0)$ de dicha curva.

Sol.: a) Teorema de la función implícita, det $\frac{\partial(f_1,f_2)}{\partial(x,y)} \neq 0$. b)

Ejercicio 12. Determinar si las siguientes funciones son globalmente invertibles en su dominio y en el conjunto A. Obtener en qué puntos de su dominio son localmente invertibles. Calcular $\mathbf{f}^{-1}$ y comprobar la relación entre las matrices jacobianas de $\mathbf{f}$ y $\mathbf{f}^{-1}$ cuando corresponda.

a) $f(x) = x^3$, $A = \mathbb{R}$.

b) $\mathbf{f}(x,y) = (e^x + e^y, e^x - e^y)$, $A = \mathbb{R}^2$.

c) $\mathbf{f}(x,y) = \left(\frac{x^2+y^2}{2}, \frac{x^2-y^2}{2}\right)$, $A = \{(x,y) \in \mathbb{R}^2,\ x > 0,\ y > 0\}$.

d) $\mathbf{f}(x,y) = (x^2, y^2)$, $A = \{(x,y) \in \mathbb{R}^2,\ x < 0, y > 0\}$.

Sol.: a) Inversa global en $\mathbb{R}$, $f^{-1}(y) = \sqrt[3]{y}$, $y \in \mathbb{R}$. b) Inversa global en $\mathbb{R}^2$, $\mathbf{f}^{-1}(u,v) = \left(\log\left(\frac{u+v}{2}\right), \log\left(\frac{u-v}{2}\right)\right)$, $u > 0$, $|v| < u$. c) Inversa global en A, inversa local en $xy \neq 0$, $\mathbf{f}^{-1}(u,v) = \left(\sqrt{u+v}, \sqrt{u-v}\right)$, $u > 0, u > |v|$. d) Inversa global en A, inversa local en $xy \neq 0$, $\mathbf{f}^{-1}(u,v) = \left(-\sqrt{u}, \sqrt{v}\right)$, $u > 0, v > 0$.

Ejercicio 13. Dada las siguientes ecuaciones

$$x = e^{u+v},\ y = \operatorname{sen}(u-v),$$

se pide determinar en qué puntos definen un cambio de variable, y utilizarlas para transformar la siguiente relación donde $f \in C^1(\mathbb{R}^2)$:

$$4x\sqrt{1-y^2}\left(\frac{\partial f}{\partial x}\right)\left(\frac{\partial f}{\partial y}\right) = 1.$$

Sol.:$u \neq v + (2k+1)\frac{\pi}{2}$, $k \in \mathbb{Z}$; $\left(\frac{\partial f}{\partial u}\right)^2 - \left(\frac{\partial f}{\partial v}\right)^2 = 1$.

Ejercicio 14. Se considera la función

$$F(x,y) = x^2 + y^2 + xy + x^3 + ay.$$

Se pide:

a) Hallar los valores de a que garantizan que la ecuación $F(x,y) = 0$ define y como función implícita de x, $y = f(x)$, en un entorno de $(0,0)$.

b) Calcular a para que el polinomio de McLaurin de grado 2 de $f(x)$ tome el valor 1 en $x = 1$. Hallar los valores de a para los que la función $f(x)$ tiene un extremo relativo en $x = 0$.

c) Se define la función $\mathbf{g}(x,y) = \left(e^{x+y} + x^2 - 1, f(x) + y\cos x\right)$ con $x \in \operatorname{dom} f$. Probar que $\mathbf{g}$ admite inversa diferenciable en un entorno de $(0,0)$.

d) Probar que la función $\mathbf{h}(x,y) = \mathbf{g} \circ \mathbf{g}(x,y) + \mathbf{g}^{-1}(x,y)$ es diferenciable en $(0,0)$ y calcular en este punto su matriz jacobiana.

Sol.: a) Con $a \neq 0$, $\frac{\partial F}{\partial y}(0,0) \neq 0$. b) $a = -1$, $a < 0$ mínimo, $a > 0$ máximo. c) $\det J\mathbf{g}(0,0) \neq 0$. d) $J\mathbf{h}(0,0) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$.