Capítulo 8

Función inversa y función implícita

Aplicaciones localmente inyectivas y aplicaciones abiertas. Inversión local de aplicaciones diferenciables. Funciones implícitas. Cambio de variable y técnicas de cálculo con funciones inversas e implícitas.

El objetivo del cálculo diferencial es el estudio del comportamiento local de una función en el entorno de un punto. Si la aproximación local de primer orden proporcionada por la diferencial tiene cierta propiedad, cabe esperar que la función también tenga esa propiedad localmente. Resultados de este tipo son los que se tratan en este capítulo al estudiar la inversión local de aplicaciones diferenciables y la existencia de funciones definidas implícitamente.

Si $\mathbf{f} : \Omega \to F$ es diferenciable en $\mathbf{a} \in \Omega$ y su diferencial $d\mathbf{f}(\mathbf{a}) : E \to F$ es una aplicación lineal invertible cabe esperar que $\mathbf{f}$ sea localmente invertible en $\mathbf{a}$, lo que significa que existe algún entorno $U \subset \Omega$ de $\mathbf{a}$ tal que $\mathbf{f}|_U$ tiene inversa. Los resultados de naturaleza local que conciernen al teorema de la función inversa los presentamos desdoblados en dos tipos de resultados: Los que garantizan la inyectividad local y los que aseguran que la aplicación es abierta y con ello la continuidad de la inversa.

Después de estudiar la diferenciabilidad de las funciones inversas se introducen los C^m-difeomorfismos, que son los cambios de variable naturales en problemas de cálculo diferencial donde intervienen funciones de clase C^m.

Los problemas de existencia de funciones definidas implícitamente se enmarcan en el siguiente planteamiento: Dado un sistema de $m = n - k$ ecuaciones con n incógnitas

$$g_1(x_1, \cdots, x_n) = 0; \quad g_2(x_1, \cdots, x_n) = 0; \quad g_m(x_1, \cdots, x_n) = 0;$$

se trata de resolverlo localmente con el fin de expresar, en el entorno de un punto, a las m variables $x_{k+1}, x_{k+2} \cdots x_n$ en función de las restantes variables $x_1, x_2, \cdots x_k$.

Los resultados que sobre este asunto se exponen aquí son versiones no lineales de resultados bien conocidos en el ámbito lineal. En ellos se asume que la diferencial de cierta aplicación cumple la hipótesis del caso lineal y se demuestra, bajo las condiciones naturales, que esta propiedad se transmite localmente a la función.

Este capítulo finaliza exponiendo con detalle las técnicas de cálculo con funciones definidas implícitamente, y en particular con funciones inversas. Finalmente se consideran los problemas de cambio de variable en el contexto del cálculo diferencial. Con varios ejemplos se explica la técnica sistemática para realizarlos y se aplica para resolver algunas ecuaciones funcionales sencillas.

Los teoremas de la función inversa y de la función implícita intervienen en el siguiente capítulo para establecer la equivalencia de las diferentes formas de definir las subvariedades diferenciables de $\mathbb{R}^n$. Como tema complementario directamente relacionado con el material de este capítulo el lector interesado puede ver en el apéndice H las nociones de dependencia e independencia funcional, otro asunto interesante para el que son esenciales los teoremas de la función inversa y de la función implícita.

Las demostraciones que se ofrecen aquí para estos teoremas son de naturaleza finito dimensional. No obstante, hay otras técnicas más generales que permiten extenderlos al caso de aplicaciones entre espacios normados completos arbitrarios.

8.1. Aplicaciones con inversa local

Para que una función $\mathbf{f} : \Omega \to \mathbb{R}^n$, definida en un abierto $\Omega \subset \mathbb{R}^n$, se pueda invertir localmente en $\mathbf{a} \in \Omega$, hay que encontrar un entorno abierto U de $\mathbf{a}$ tal que $\mathbf{f}|_U$ sea inyectiva y abierta. Esta segunda propiedad garantizará que $\mathbf{f}(U) = V$ es abierto y la continuidad de la inversa $(\mathbf{f}|_U)^{-1} : V \to U$. Comenzamos obteniendo condiciones suficientes para la inyectividad local.

Aplicaciones localmente inyectivas. El siguiente ejemplo muestra que la hipótesis de que la diferencial $d\mathbf{f}(\mathbf{a})$ sea inyectiva no garantiza que $\mathbf{f}|_U$ sea inyectiva en algún entorno U de $\mathbf{a}$:

Ejemplo 8.1 *La función $f : \mathbb{R} \to \mathbb{R}$, definida por $f(x) = x/2+x^2 \operatorname{sen}(1/x)$ si $x \neq 0$, $f(0) = 0$, es derivable en todo $x \in \mathbb{R}$ y $f'(0) \neq 0$, luego su diferencial $h \to f'(0)h$ es inyectiva. Sin embargo f no es inyectiva en los intervalos de la forma $(-\epsilon, \epsilon)$ porque en ellos no es estrictamente monótona (ya que $f'(1/(n\pi)) = 1/2 - (-1)^n$) y por lo tanto f' cambia de signo en estos intervalos)*

En este ejemplo se aprecia que la discontinuidad de f' en $x = 0$ es la que permite que f no sea inyectiva en los entornos de 0. Para impedir situaciones como esta, en la proposición 8.2 y en el teorema de la función inversa 8.13 interviene la hipótesis de que la diferencial $d\mathbf{f}(\mathbf{x})$ exista en un entorno de $\mathbf{a}$, y sea continua en $\mathbf{a}$. Veremos que con esta hipótesis, en el caso finito dimensional, el problema de la inversión local tiene solución satisfactoria y se consigue una inversa local, definida y continua en un entorno de $\mathbf{b} = \mathbf{f}(\mathbf{a})$, que resulta de clase C^m si $\mathbf{f}$ es de clase C^m.

En lo que sigue, si $\mathbf{f} : \Omega \to \mathbb{R}^n$ está definida en un abierto $\Omega \subset \mathbb{R}^n$ y en el punto $\mathbf{a} \in \Omega$ existen todas las derivadas parciales $D_i f_j(\mathbf{a})$, $1 \leq i, j \leq n$, escribiremos

$$\det \mathbf{f}'(\mathbf{a}) = \det \left[D_i f_j(\mathbf{a})\right]$$

para denotar el valor del determinante de la matriz Jacobiana en el punto $\mathbf{a}$.

Proposición 8.2 *Sea* $\mathbf{f} : \Omega \to \mathbb{R}^n$ *diferenciable en un abierto* $\Omega \subset \mathbb{R}^n$ *con derivadas parciales continuas en* $\mathbf{a} \in \Omega$*. Si* $\det \mathbf{f}'(\mathbf{a}) \neq 0$ *entonces existe* $B(\mathbf{a}, r) \subset \Omega$ *tal que* $\mathbf{f}|_{B(\mathbf{a},r)}$ *es inyectiva.*

Dem: La función $h(\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n) = \det[D_i f_k(\mathbf{z}_k)]$, está definida en

$$\Omega^n \subset \mathbb{R}^n \times \overset{n}{\cdots} \times \mathbb{R}^n = \mathbb{R}^{n^2}$$

Como las funciones $D_i f_k$ son continuas en $\mathbf{a}$ se sigue que h es continua en $(\mathbf{a}, \mathbf{a} \cdots, \mathbf{a})$. Es claro que $h(\mathbf{a}, \mathbf{a}, \cdots \mathbf{a}) = \det \mathbf{f}'(\mathbf{a}) \neq 0$, y si suponemos que $\det \mathbf{f}'(\mathbf{a}) \neq 0$, la continuidad de h en $(\mathbf{a}, \mathbf{a}, \cdots, \mathbf{a})$ permite asegurar que existe un entorno de $(\mathbf{a}, \mathbf{a} \cdots \mathbf{a})$, de la forma

$$V = B(\mathbf{a}, r) \times \overset{n}{\cdots} \times B(\mathbf{a}, r) \subset \Omega^n$$

tal que $h(\mathbf{z}_1, \mathbf{z}_2, \cdots \mathbf{z}_n) \neq 0$ si $(\mathbf{z}_1, \mathbf{z}_2, \cdots \mathbf{z}_n) \in V$.
Demostramos a continuación que si $\mathbf{x}, \mathbf{y} \in B(\mathbf{a}, r)$, y $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{y})$, entonces $\mathbf{x} = \mathbf{y}$. Como el segmento $[\mathbf{x}, \mathbf{y}] = \{t\mathbf{x} + (1-t)\mathbf{y} : 0 \leq t \leq 1\}$ está contenido en $B(\mathbf{a}, r) \subset \Omega$, para $k \in \{1, 2, \cdots n\}$, están definidas la funciones reales de variable real $\varphi_k(t) = f_k(\mathbf{z}(t))$ donde $\mathbf{z}(t) = t\mathbf{x} + (1-t)\mathbf{y}$. En virtud de la regla de la cadena estas funciones son derivables en $[0, 1]$, con derivada

$$\varphi_k'(t) = df_k(\mathbf{z}(t))\mathbf{z}'(t) = df_k(\mathbf{z}(t))(\mathbf{x} - \mathbf{y}) = \sum_{j=1}^{n} D_j f_k(\mathbf{z}(t))(x_j - y_j)$$

Según el teorema del valor medio existe $\theta_k \in (0, 1)$ tal que

$$0 = f_k(\mathbf{x}) - f_k(\mathbf{y}) = \varphi_k(1) - \varphi_k(0) = \varphi_k'(\theta_k)$$

es decir, los puntos $\mathbf{z}_k = \mathbf{z}(\theta_k)$ verifican

$$\sum_{j=1}^{n} D_j f_k(\mathbf{z}_k)(x_j - y_j) = \varphi_k'(\theta_k) = 0$$

Como $\mathbf{z}_k = \mathbf{z}(\theta_k) \in [\mathbf{x}, \mathbf{y}] \subset B(\mathbf{a}, r)$, podemos asegurar que $\det[D_i f_k(\mathbf{z}_k)] \neq 0$. Entonces, considerando

$$\sum_{j=1}^{n} D_j f_k(\mathbf{z}_k)(x_j - y_j) = 0, \quad 1 \leq k \leq n$$

como un sistema lineal en las incógnitas $x_j - y_j$, cuyo determinante no es nulo, se concluye que $x_j - y_j = 0$ para $1 \leq j \leq n$, es decir, $\mathbf{x} = \mathbf{y}$. ■

Obsérvese que el resultado que se obtiene con la proposición 8.2 es de naturaleza local. La razón de esto se debe a que las propiedades de la diferencial, que aproxima localmente a la función en un punto, sólo pueden propiciar propiedades de la función de tipo local. El siguiente ejemplo muestra que, en general, con la proposición 8.2 no se pueden conseguir un resultado de tipo global

Ejemplo 8.3 *La aplicación* $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$, *definida por* $\mathbf{f}(x, y) = e^x(\cos y, \operatorname{sen} y)$, *es localmente inyectiva en todo punto, pero no es globalmente inyectiva aunque su diferencial* $d\mathbf{f}(\mathbf{a})$ *lo es en todo punto* $\mathbf{a} \in \mathbb{R}^2$.

El lector que sólo esté interesado en la demostración del teorema de la función inversa puede omitir el siguiente resultado que completa el que acabamos de obtener.

Teorema 8.4 *Sea* $\mathbf{f} : \Omega \to \mathbb{R}^n$ *diferenciable en un abierto* $\Omega \subset \mathbb{R}^k$, *con derivadas parciales continuas en* $\mathbf{a} \in \Omega$. *Si* $k \leq n$ *y* $d\mathbf{f}(\mathbf{a}) : \mathbb{R}^k \to \mathbb{R}^n$ *es inyectiva (e.d si la matriz jacobiana* $\mathbf{f}'(\mathbf{a}) = (D_i f_j(\mathbf{a}))_{1\leq i\leq k, 1\leq j\leq n}$ *tiene rango* k*) entonces existe una bola abierta* $B(\mathbf{a}, r) \subset \Omega$ *tal que* $\mathbf{f}|_{B(\mathbf{a},r)}$ *es inyectiva.*

Dem: Después de la proposición 8.2 sólo tenemos que considerar el caso $k < n$. Como la matriz $\mathbf{f}'(\mathbf{a}) = (D_i f_j(\mathbf{a}))_{1\leq i\leq k, 1\leq j\leq n}$ tiene rango k suponemos, para simplificar la notación, que no es nulo el determinante de la matriz $(D_i f_j(\mathbf{a}))_{1\leq i\leq k, 1\leq j\leq k}$. Entonces la función $\mathbf{g}(\mathbf{x}) = (f_1, f_2, \cdots f_k)$ cumple que $\det \mathbf{g}'(\mathbf{a}) \neq 0$, y con la proposición 8.2 se obtiene una bola abierta abierta $B(\mathbf{a}, r) \subset \Omega$ tal que $\mathbf{g}|_{B(\mathbf{a},r)}$ es inyectiva lo que implica que $\mathbf{f}|_{B(\mathbf{a},r)}$ también lo es ■

Ejemplo 8.5

La aplicación $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $\mathbf{f}(x, y, z) = (x + y, x^2 - y, y^4)$ es localmente inyectiva en cada punto $(x, y) \neq (-1/2, 0)$ ya que la matriz Jacobiana

$$\begin{pmatrix} 1 & 2x & 0 \\ 1 & -1 & 4y^3 \end{pmatrix}$$

tiene rango dos en todo $(x, y) \neq (-1/2, 0)$. El punto $(-1/2, 0)$, donde el rango de la matriz es 1, no tiene ningún entorno sobre el que $\mathbf{f}$ sea inyectiva: Basta observar que para todo $\epsilon > 0$ se cumple $\mathbf{f}(-1/2 + \epsilon, -\epsilon) = \mathbf{f}(-1/2 - \epsilon, +\epsilon)$. ■

Aplicaciones abiertas. Recordemos que una transformación espacios topológicos se dice que es abierta cuando transforma abiertos en abiertos, lo que equivale a que transforma cada entorno de un punto de su dominio en un entorno del punto imagen. El siguiente lema proporciona un ingrediente básico para la demostración del teorema de la aplicación abierta.

Lema 8.6 *Sea* $\mathbf{f} : \overline{B} \to \mathbb{R}^n$ *una aplicación continua, donde* $B = B(\mathbf{a}, r) \subset \mathbb{R}^n$ *una bola abierta para la norma euclídea. Se supone que*
i) Para cada $\mathbf{x} \in B(\mathbf{a}, r)$ *existen las derivadas* $D_i f_k(\mathbf{x})$, $1 \leq i, k \leq n$, *y* $\det \mathbf{f}'(\mathbf{x}) \neq 0$.
ii) $f(\mathbf{a}) \neq f(\mathbf{x})$ *si* $\|\mathbf{x} - \mathbf{a}\|_2 = r$.
Entonces existe $\rho > 0$ *tal que* $B(\mathbf{f}(\mathbf{a}), \rho) \subset \mathbf{f}(B(\mathbf{a}, r))$.

Dem: La función continua $\mathbf{x} \to \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\|_2$ alcanza un mínimo absoluto sobre el compacto $S_r = \{\mathbf{x} : \|\mathbf{x} - \mathbf{a}\|_2 = r\}$, luego existe $\mathbf{z} \in S_r$ tal que

$$M = \min\{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\|_2 : \|\mathbf{x} - \mathbf{a}\|_2 = r\} = \|\mathbf{f}(\mathbf{z}) - \mathbf{f}(\mathbf{a})\|_2$$

Según la hipótesis ii), $\mathbf{f}(\mathbf{z}) \neq \mathbf{f}(\mathbf{a})$, y así podemos asegurar que $M > 0$. Vamos a demostrar que con $\rho = M/2$ se cumple la inclusión $B(\mathbf{f}(\mathbf{a}), \rho) \subset \mathbf{f}(B(\mathbf{a}, r))$.

Dado $\mathbf{y} \in B(\mathbf{f}(\mathbf{a}), \rho)$, la función continua $h(\mathbf{x}) = \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|_2$ alcanza en un punto $\mathbf{e}$ del compacto $\overline{B}$ el mínimo absoluto

$$\min\{h(\mathbf{x}) : \|\mathbf{x} - \mathbf{a}\|_2 \leq r\} = h(\mathbf{e})$$

Si $\|\mathbf{x} - \mathbf{a}\|_2 = r$ se cumple $h(\mathbf{x}) > h(\mathbf{e})$, ya que

$$h(\mathbf{x}) = \|\mathbf{f}(\mathbf{x}) - \mathbf{y}\|_2 = \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a}) - (\mathbf{y} - \mathbf{f}(\mathbf{a}))\|_2 \geq$$

$$\geq \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\|_2 - \|\mathbf{y} - \mathbf{f}(\mathbf{a})\|_2 \geq M - \rho = \rho > \|\mathbf{f}(\mathbf{a}) - \mathbf{y}\|_2 = h(\mathbf{a}) \geq h(\mathbf{e})$$

y podemos asegurar así que $\mathbf{e} \in B(\mathbf{a}, r)$. En definitiva, en la bola abierta $B(\mathbf{a}, r)$, la función diferenciable $h(\mathbf{x})^2 = \sum_{j=1}^n (f_j(\mathbf{x}) - y_j)^2$ alcanza un mínimo (absoluto) en $\mathbf{e} \in B(\mathbf{a}, r)$, y por lo tanto $D_k h^2(\mathbf{e}) = 0$ para $1 \leq k \leq n$, es decir

$$2\sum_{j=1}^n (f_j(\mathbf{e}) - y_j) D_k f_j(\mathbf{e}) = 0$$

Por hipótesis $\det[D_k f_j(\mathbf{e})] = \det \mathbf{f}'(\mathbf{e}) \neq 0$, luego el sistema homogéneo de ecuaciones lineales asociado a la matriz $D_k f_j(\mathbf{e})$ sólo tiene la solución trivial. Por lo tanto $f_j(\mathbf{e}) - y_j = 0$ para cada $1 \leq j \leq n$, y queda demostrado que $\mathbf{y} = \mathbf{f}(\mathbf{e}) \in \mathbf{f}(B(\mathbf{a}, r))$.
■

Proposición 8.7 *Sea* $\mathbf{f} : \Omega \to \mathbb{R}^n$ *continua e inyectiva en un abierto* $\Omega \subset \mathbb{R}^n$ *tal que en cada* $\mathbf{x} \in \Omega$ *existen las derivadas parciales* $D_i f_k(\mathbf{x})$, $1 \leq i, k \leq n$, *y* $\det \mathbf{f}'(\mathbf{x}) \neq 0$. *Entonces* $\mathbf{f}$ *es abierta, es decir,* $\mathbf{f}(V)$ *es abierto para cada abierto* $V \subset \Omega$.

DEM: Dado $\mathbf{a} \in V$, sea $r > 0$ tal que $\overline{B(\mathbf{a}, r)} \subset V$. Sobre la bola cerrada $\overline{B(\mathbf{a}, r)}$ se satisfacen las hipótesis del lema 8.6, luego existe $\rho > 0$ tal que

$$B(\mathbf{f}(\mathbf{a}), \rho) \subset \mathbf{f}(B(\mathbf{a}, r)) \subset \mathbf{f}(V)$$

■

Obsérvese que en la proposición 8.7 no se ha supuesto que $\mathbf{f}$ sea de clase C^1. Para funciones de clase C^1 el resultado proporcionado por esta proposición queda cubierto por el siguiente teorema, donde no se supone que la función sea inyectiva.

Teorema 8.8 [Aplicación abierta] *Si* $\mathbf{f} : \Omega \to \mathbb{R}^n$ *es de clase* $C^1(\Omega)$ *en un abierto* $\Omega \subset \mathbb{R}^n$ *y* $\det \mathbf{f}'(\mathbf{x}) \neq 0$ *para todo* $\mathbf{x} \in \Omega$ *entonces* $\mathbf{f}$ *es abierta.*

DEM: Si $U \subset \Omega$ es abierto y $\mathbf{a} \in U$, aplicando la proposición 8.2, se obtiene una bola abierta $B_{\mathbf{a}} \subset U$, centrada en $\mathbf{a}$, tal que $\mathbf{f}|_{B_{\mathbf{a}}}$ es inyectiva. En virtud de la proposición 8.7 cada $\mathbf{f}(B_{\mathbf{a}})$ es abierto, luego $\mathbf{f}(U) = \bigcup_{\mathbf{a} \in U} \mathbf{f}(B_{\mathbf{a}})$ es abierto. ■

NOTA: En el ejercicio resuelto 8.19 se muestra que el teorema 8.8 sigue valiendo con una hipótesis más débil: Basta suponer que $\mathbf{f}$ es diferenciable en cada $\mathbf{x} \in \Omega$ y que $\det \mathbf{f}'(\mathbf{x}) \neq 0$ para todo $\mathbf{x} \in \Omega$.

El lector que sólo esté interesado en el teorema de la función inversa también puede omitir el siguiente teorema y su corolario, que completan el resultado anterior.

Teorema 8.9 *Sea $\mathbf{f} : \Omega \to \mathbb{R}^m$ de clase C^1 en un abierto $\Omega \subset \mathbb{R}^n$, donde $m \leq n$ y $\mathbf{a} \in \Omega$. Si $d\mathbf{f}(\mathbf{a}) : \mathbb{R}^n \to \mathbb{R}^m$ es sobreyectiva (e.d si la matriz jacobiana $\mathbf{f}'(\mathbf{a})$ es de rango m) y $A \subset \Omega$ es entorno de $\mathbf{a}$ entonces $\mathbf{f}(A)$ es entorno de $\mathbf{f}(\mathbf{a})$. Más aún, $\mathbf{a}$ posee un entorno abierto $U \subset \Omega$ tal que $\mathbf{f}|_U$ es abierta.*

Dem: Como el caso $n = m$ ya ha sido demostrado en el teorema 8.8, sólo tenemos que considerar el caso $m < n$. Suponemos, para simplificar el la notación, que no se anula el determinante $\Delta(\mathbf{a})$ de la matriz cuadrada $(D_i f_j(\mathbf{a}))_{1 \leq i,j \leq m}$. Como el determinante $\Delta(\mathbf{x})$ de la matriz $(D_i f_j(\mathbf{x}))_{1 \leq i,j \leq m}$ es una función continua de $\mathbf{x}$ podemos asegurar que existe $B(\mathbf{a}, r) \subset A$ tal que $\Delta(\mathbf{x}) \neq 0$ para cada $\mathbf{x} \in B(\mathbf{a}, r)$. Si consideramos la función auxiliar $\mathbf{g} : \Omega \to \mathbb{R}^n$, definida por

$$\mathbf{g}(x_1, x_2, \cdots, x_n) = (f_1(x_1, \cdots x_n), f_2(x_1, \cdots x_n) \cdots, f_m(x_1, \cdots x_n), x_{m+1}, \cdots x_n)$$

es fácil comprobar que $\det \mathbf{g}'(\mathbf{x}) = \Delta(\mathbf{x})$. Como la función $\mathbf{g}$ es de clase $C^1(\Omega)$ y $\det \mathbf{g}'(\mathbf{x}) \neq 0$ para todo $\mathbf{x} \in B(\mathbf{a}, r)$, con el teorema 8.8 obtenemos que $\mathbf{g}$ es abierta sobre $U = B(\mathbf{a}, r)$. Como la proyección

$$\pi : \mathbb{R}^n \to \mathbb{R}^m, \quad \pi(x_1, x_2, \cdots, x_m, x_{m+1}, \cdots x_n) = (x_1, x_2, \cdots, x_m)$$

transforma abiertos en abiertos se sigue que $\mathbf{f}|_U = \pi \circ \mathbf{g}|_U : U \to \mathbb{R}^m$ es abierta, luego $\mathbf{f}(U)$ es un subconjunto abierto de $\mathbb{R}^m$ contenido en $\mathbf{f}(A)$, y por lo tanto $\mathbf{f}(A)$ es entorno de $\mathbf{f}(\mathbf{a})$. ∎

Corolario 8.10 *Sea $\mathbf{f} : \Omega \to \mathbb{R}^m$ de clase C^1 en un abierto $\Omega \subset \mathbb{R}^n$, donde $m \leq n$. Si Si $d\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$ es sobreyectiva en todo punto $\mathbf{x} \in \Omega$ (e.d. si la matriz jacobiana $\mathbf{f}'(\mathbf{x})$ es de rango m en cada $\mathbf{x} \in \Omega$) entonces $\mathbf{f}$ es abierta.*

Dem: Es consecuencia directa del teorema 8.9 ya que se cumplen sus hipótesis en cada punto $\mathbf{a} \in \Omega$. ∎

Funciones inversas. Teorema de inversión local. Antes de demostrar el teorema de inversión local demostraremos dos resultados preliminares que nos dicen que, bajo las hipótesis naturales, cuando una función $\mathbf{f}$ tiene inversa continua las propiedades de diferenciabilidad de $\mathbf{f}$ las hereda la inversa.

Teorema 8.11 *Sea $\mathbf{f} : A \to B$ una biyección entre dos abiertos $A, B \subset \mathbb{R}^n$ tal que $\mathbf{f}$ es diferenciable en $\mathbf{a} \in A$ y su inversa $\mathbf{g} : B \to A$ es continua en $\mathbf{b} = \mathbf{f}(\mathbf{a})$.*

Si $\det \mathbf{f}'(\mathbf{a}) \neq 0$ entonces $\mathbf{g} = \mathbf{f}^{-1} : B \to A$ es diferenciable en $\mathbf{b} = \mathbf{f}(\mathbf{a})$ y su diferencial $d\mathbf{g}(\mathbf{b})$ es la inversa de la aplicación lineal $d\mathbf{f}(\mathbf{a})$.

Dem: La diferencial $L = d\mathbf{f}(\mathbf{a})$ es una aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^n$ que satisface

$$\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}) = L(\mathbf{h}) + \|\mathbf{h}\|\,\epsilon(\mathbf{h})\text{donde} \lim_{\mathbf{h}\to\mathbf{0}} \epsilon(\mathbf{h}) = \mathbf{0}$$

En lo que sigue suponemos definido $\epsilon(\mathbf{0}) = \mathbf{0}$ de modo que $\epsilon(\mathbf{h})$ es continua en $\mathbf{h} = \mathbf{0}$ y cumple la igualdad anterior incluso en el caso $\mathbf{h} = \mathbf{0}$.

Si $\mathbf{k} \in \mathbb{R}^n$, con $\mathbf{b} + \mathbf{k} \in B$, entonces el incremento $\mathbf{h} = \mathbf{g}(\mathbf{b}+\mathbf{k}) - \mathbf{g}(\mathbf{b}) \in \mathbb{R}^n$ cumple que $\mathbf{a} + \mathbf{h} \in A$ y además $\mathbf{k} = \mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a})$, luego

$$\mathbf{k} = L(\mathbf{h}) + \|\mathbf{h}\|\,\epsilon(\mathbf{h})$$

La aplicación lineal L tiene inversa porque $\det(L) = \det \mathbf{f}'(\mathbf{a}) \neq 0$, y aplicando L^{-1} a los dos miembros de la última igualdad resulta

$$L^{-1}(\mathbf{k}) = \mathbf{h} + \|\mathbf{h}\|\,L^{-1}(\epsilon(\mathbf{h}))$$

y sustituyendo $\mathbf{h} = \mathbf{g}(\mathbf{b}+\mathbf{k}) - \mathbf{g}(\mathbf{b})$ se obtiene

$$\mathbf{g}(\mathbf{b}+\mathbf{k}) - \mathbf{g}(\mathbf{b}) = L^{-1}(\mathbf{k}) - \|\mathbf{h}\|\,L^{-1}(\epsilon(\mathbf{h}))$$

En la fórmula anterior, y en lo que sigue, se considera siempre que $\mathbf{h}$ es función de $\mathbf{k}$, es decir, se supone efectuada la sustitución $\mathbf{h} = \mathbf{g}(\mathbf{b}+\mathbf{k}) - \mathbf{g}(\mathbf{b})$. Obsérvese que el incremento $\mathbf{h} = \mathbf{g}(\mathbf{b}+\mathbf{k}) - \mathbf{g}(\mathbf{b})$ tiende hacia $\mathbf{0}$ cuando $\mathbf{k} \to \mathbf{0}$ porque $\mathbf{g}$ es continua en $\mathbf{b}$. Se sigue que $\epsilon(\mathbf{h})$ también tiende hacia $\epsilon(\mathbf{0}) = \mathbf{0}$ cuando $\mathbf{k}$ tiende hacia $\mathbf{0}$.

Para demostrar que $\mathbf{g}$ es diferenciable en $\mathbf{b}$, con $d\mathbf{g}(\mathbf{b}) = L^{-1}$, basta ver que

$$\|\mathbf{h}\|\,L^{-1}(\epsilon(\mathbf{h})) = o(\|\mathbf{k}\|)$$

Como L^{-1} es continua y $\epsilon(\mathbf{h})$ tiende hacia $\mathbf{0}$ cuando $\mathbf{k}$ tiende hacia $\mathbf{0}$, existe $\delta > 0$ tal que $\|\mathbf{k}\| < \delta \Rightarrow \|L^{-1}(\epsilon(\mathbf{h}))\| < 1/2$, luego

$$\left\|L^{-1}(\mathbf{k})\right\| = \left\|\mathbf{h} + \|\mathbf{h}\|\,L^{-1}(\epsilon(\mathbf{h}))\right\| \geq \|\mathbf{h}\| - \|\mathbf{h}\|\left\|L^{-1}(\epsilon(\mathbf{h}))\right\| \geq \frac{1}{2}\|\mathbf{h}\|$$

Entonces, cuando $\|\mathbf{k}\| < \delta$, se cumple

$$\|\mathbf{h}\|\left\|L^{-1}(\epsilon(\mathbf{h}))\right\| \leq 2\left\|L^{-1}(\mathbf{k})\right\|\left\|L^{-1}(\epsilon(\mathbf{h}))\right\| \leq \|\mathbf{k}\|\left\|L^{-1}\right\|^2\|\epsilon(\mathbf{h})\|$$

y se sigue de esta desigualdad que $\|\mathbf{h}\|\,\|L^{-1}(\epsilon(\mathbf{h}))\| / \|\mathbf{k}\|$ tiende hacia 0 cuando $\mathbf{k}$ tiende hacia $\mathbf{0}$. ■

observación. En el teorema anterior la hipótesis $\det \mathbf{g}'(\mathbf{a}) \neq 0$ es crucial para conseguir la diferenciabilidad de $\mathbf{g}$ en $\mathbf{b} = f(\mathbf{a})$: La función $f(x) = x^3$ es derivable en $a = 0$ pero su inversa $g(x) = \sqrt[3]{x}$, que es continua en $b = 0$, no es derivable en este punto.

Teorema 8.12 *Sea $\mathbf{f} : A \to B$ una biyección entre dos abiertos $A, B \subset \mathbb{R}^n$, diferenciable en cada $\mathbf{x} \in A$ con $\det \mathbf{f}'(\mathbf{x}) \neq 0$. Si la inversa $\mathbf{g} = \mathbf{f}^{-1} : B \to A$ es continua, entonces es diferenciable en cada $\mathbf{y} \in B$. Si $\mathbf{f}$ es de clase $C^m(A)$ entonces $\mathbf{g}$ es de clase $C^m(B)$.*

DEM: Por lo que se acaba de demostrar en el teorema 8.11 la inversa $\mathbf{g} : B \to A$ es diferenciable en cada $\mathbf{y} \in B$ y su diferencial es $d\mathbf{g}(\mathbf{y}) = [d\mathbf{f}(\mathbf{g}(\mathbf{y}))]^{-1}$, es decir, la matriz jacobiana de $\mathbf{g}'(\mathbf{y})$ se obtiene invirtiendo la matriz jacobiana $\mathbf{f}'(\mathbf{x})$ y sustituyendo luego $\mathbf{x} = \mathbf{g}(\mathbf{y})$.

Para demostrar que $\mathbf{g}$ es de clase $C^m(B)$ cuando $\mathbf{f}$ es de clase $C^m(A)$ consideramos el espacio $\mathcal{M}$ formado por las matrices cuadradas $n \times n$ de números reales, que se supone identificado con $\mathbb{R}^{n^2}$ dotado de la topología usual. Con esta topología la aplicación $\det : \mathcal{M} \to \mathbb{R}$ que asocia a cada matriz $M \in \mathcal{M}$ su determinante es continua y por lo tanto el conjunto de las matrices invertibles

$$\Gamma = \{M \in \mathcal{M} : \det(M) \neq 0\}$$

es un subconjunto abierto en $\mathcal{M}$. La aplicación $\mathrm{Inv} : \Gamma \to \Gamma$ que asocia a cada matriz $M = (m_{ij}) \in \Gamma$ su matriz inversa $\mathrm{Inv}(M) = M^{-1}$ es de clase C^∞ (basta tener en cuenta que cada elemento de la matriz inversa M^{-1} es una función racional de las variables m_{ij} cuyo denominador $\det[m_{ij}] \neq 0$ no se anula). Si $\mathbf{f}$ es de clase C^1, como $\mathbf{g}'$ se obtiene componiendo las aplicaciones continuas

$$\mathbf{g}' : B \xrightarrow{\mathbf{g}} A \xrightarrow{\mathbf{f}'} \Gamma \xrightarrow{\mathrm{Inv}} \Gamma$$

obtenemos que $\mathbf{g}'$ es continua, lo que significa que $\mathbf{g}$ es de clase C^1. Razonado por inducción sobre m se demuestra que $\mathbf{g}$ es de clase C^m si $\mathbf{f}$ lo es:

Ya hemos visto que el resultado es cierto para $m = 1$. Si $\mathbf{f}$ es de clase C^m y el resultado se supone cierto para funciones de clase C^{m-1}, esta hipótesis de inducción conduce a que $\mathbf{g}$ es de clase C^{m-1}. Como $\mathbf{f}'$ también es de clase C^{m-1}, y lo mismo le ocurre a Inv resulta que $\mathbf{g}'$ es la composición de tres aplicaciones de clase C^{m-1}. La proposición 7.4 permite concluir que $\mathbf{g}'$ es de clase C^{m-1}, lo que significa que $\mathbf{g}$ es de clase C^m. ■

Teorema 8.13 [Función inversa] *Sea* $\mathbf{f} : \Omega \to \mathbb{R}^n$ *diferenciable en un abierto* $\Omega \subset \mathbb{R}^n$ *y* $\mathbf{a} \in \Omega$ *un punto donde las derivadas* $D_i f_j$, $1 \leq i, j \leq n$ *son continuas.*
Si $\det \mathbf{f}'(\mathbf{a}) \neq 0$, *existen abiertos* $A \subset \Omega$, $B \subset \mathbb{R}^n$, *con* $\mathbf{a} \in A$, $\mathbf{b} = \mathbf{f}(\mathbf{a}) \in B$, *verificando*

i) $\mathbf{f}|_A$ *es inyectiva y* $\mathbf{f}(A) = B$.

ii) $\mathbf{g} = (\mathbf{f}|_A)^{-1} : B \to A$ *es diferenciable en* B.

Si $\mathbf{f}$ *es de clase* $C^m(A)$ *entonces* $\mathbf{g}$ *sea de clase* $C^m(B)$.

DEM: La hipótesis sobre las derivadas parciales garantiza que la función $\mathbf{x} \to \det \mathbf{f}'(\mathbf{x})$ es continua en $\mathbf{a}$, luego la condición $\det \mathbf{f}'(\mathbf{a}) \neq 0$ permite asegurar que existe $B(\mathbf{a}, \rho) \subset \Omega$ tal que $\det \mathbf{f}'(\mathbf{x}) \neq 0$ para todo $\mathbf{x} \in B(\mathbf{a}, \rho)$. Según la proposición 8.2 existe una bola abierta $A = B(\mathbf{a}, r) \subset B(\mathbf{a}, \rho)$ tal que $\mathbf{f}|_A$ es inyectiva. En el abierto A la función $\mathbf{f}$ cumple las hipótesis de la proposición 8.7 luego $\mathbf{f}|_A$ es abierta, es decir $B = f(A)$ es abierto y $\mathbf{f}$ transforma cada abierto $U \subset A$ en un abierto $V = f(U) \subset B$, lo que significa que la inversa $\mathbf{g} = (\mathbf{f}|_A)^{-1} : B \to A$ de la aplicación inyectiva $\mathbf{f}|_A$ es continua en B. Queda establecido así que los abiertos $A = B(\mathbf{a}, r)$,

$B = \mathbf{f}(A)$ cumplen i), y que la inversa $g : B \to A$ es continua. Como $\det \mathbf{f}'(\mathbf{x}) \neq 0$ para todo $\mathbf{x} \in A \subset B(\mathbf{a}, \rho)$, con el teorema 8.12 se concluye que $\mathbf{g}$ es de diferenciable en cada $\mathbf{y} \in B$, y de clase $C^m(B)$ si $\mathbf{f}$ es de clase $C^m(A)$. ■

Cuando $\mathbf{f} : U \to V$ es una biyección de clase C^m entre dos abiertos $U, V \subset \mathbb{R}^n$ y su inversa $\mathbf{g} = \mathbf{f}^{-1} : V \to U$ también es de clase C^m se dice que $\mathbf{f}$ es un *difeomorfismo* de clase C^m (o un C^m-*difeomorfismo*) entre los abiertos U y V. En este caso se dice que U y V son C^m-difeomorfos.

Corolario 8.14 *Sea $\mathbf{f} : U \to V$ una biyección de clase C^m, entre dos abiertos, $U, V \subset \mathbb{R}^n$. Una condición necesaria y suficiente para que $\mathbf{f}$ sea un C^m-difeomorfismo es que para todo $\mathbf{x} \in U$ sea $\det \mathbf{f}'(\mathbf{x}) \neq 0$.*

Dem: La condición es suficiente: En un entorno B de cada $\mathbf{b} = \mathbf{f}(\mathbf{a}) \in V$ la inversa $\mathbf{g}$ de $\mathbf{f}$ coincide con la inversa local de $\mathbf{f}$ proporcionada por el teorema de la función inversa 8.13, que es de clase C^m, luego $\mathbf{g}$ es de clase $C^m(V)$.
La condición es necesaria: Dado $\mathbf{x} \in U$, por hipótesis $\mathbf{f}$ es diferenciable en $\mathbf{x}$ y $\mathbf{g}$ es diferenciable en $\mathbf{y} = \mathbf{f}(\mathbf{x})$. Como $\mathbf{g}(\mathbf{f}(\mathbf{x})) = \mathbf{x}$ para todo $\mathbf{x} \in U$, en virtud de la regla de la cadena $d\mathbf{g}(\mathbf{y}) \circ d\mathbf{f}(\mathbf{x}) = I$ (identidad), luego la aplicación lineal $d\mathbf{f}(\mathbf{x})$ es invertible y por lo tanto $\det \mathbf{f}'(\mathbf{x}) \neq 0$. ■

8.2. Funciones implícitas

El problema de la función implícita, en su forma más simple, se plantea en los siguientes términos: Si g es una función real de dos variables reales definida en un abierto $\Omega \subset \mathbb{R}^2$, se considera la ecuación $g(x, y) = 0$. Si no es vacío el conjunto de sus soluciones $S = \{(x, y) \in \Omega : g(x, y) = 0\}$, se trata de decidir cuándo y en qué sentido esta ecuación determina a la variable y como función de la variable x, es decir, bajo qué condiciones queda definida una función $y = f(x)$ tal que $g(x, f(x)) = 0$.

Esto ocurrirá con seguridad cuando $S = \{(x, y) \in \Omega : g(x, y) = 0\}$ sea la gráfica de una función $f : A \to \mathbb{R}$, definida en $A = \pi_1(\Omega)$. En este caso, la función f, que asigna a cada $x \in A$ la única solución y de la ecuación $g(x, y) = 0$ se dice que está definida implícitamente por dicha ecuación. Esto es lo que ocurre, por ejemplo, con $g(x, y) = x^2 + y^2 - 1$, que define en el abierto $\Omega = \{(x, y) \in \mathbb{R} : y > 0\}$, la función implícita $f(x) = \sqrt{1 - x^2}$, con dominio $A = (-1, +1)$.

Es raro que se presente una situación tan sencilla como la anterior pues puede ocurrir que la ecuación $g(x, y) = 0$ no tenga solución, o que el conjunto de sus soluciones se reduzca a un punto (p.e. si $g(x, y) = x^2 + y^2$). Aún suponiendo que este no es el caso, puede ocurrir que para algunos valores $x \in A = \pi_1(S)$ la ecuación $g(x, y) = 0$ tenga varias o infinitas soluciones. En estos casos, para que quede determinada una función implícita $f : A \to \mathbb{R}$, será preciso considerar alguna condición adicional que garantice que para cada $x \in A$ hay un único y que satisface la ecuación $g(x, y) = 0$ y la condición propuesta.

Las condición adicional que se suele proponer para determinar una función implícita es de tipo local: Dado un punto $(a,b) \in S$ se trata de encontrar un entorno $A \times B \subset \Omega$ de (a,b) con la propiedad de que para cada $x \in A$ haya un único $y \in B$ que sea solución de la ecuación $g(x,y) = 0$ de modo que $S \cap (A \times B)$ es la gráfica de una función $f : A \to B$. En este caso se suele decir que f es la función implícita que $A \times B$ determina en la ecuación $g(x,y) = 0$.

En el caso particular que estamos considerando la interpretación geométrica de la hipótesis que habrá que considerar para garantizar la existencia de función implícita es sencilla y natural: En el plano se tiene una curva S de ecuación $g(x,y) = 0$ que se pretende expresar, en un entorno $A \times B$ de $(a,b) \in S$, en la forma explícita habitual, como gráfica de una función real de variable real $f : A \to B$. Para ello debemos descartar los puntos (a,b) de la curva donde hay tangente vertical porque en ellos puede ocurrir que sea imposible despejar localmente a la variable y como función implícita de la variable x. Un ejemplo típico de esta situación lo proporciona la circunferencia $C = \{(x,y) : x^2 + y^2 - 1 = 0\}$ cuando se considera el punto $(a,b) = (1,0)$: En todo entorno $A \times B$ de $(1,0)$, por pequeño que sea, es imposible expresar $C \cap (A \times B)$ como la gráfica de una función $f : A \to B$. Este ejemplo revela, en el caso de funciones de dos variables reales, el papel que desempeña la hipótesis $D_2 g(a,b) \neq 0$ para conseguir que la ecuación $g(x,y) = 0$ defina, en un entorno de (a,b), a la variable y como función implícita de la variable x.

Más generalmente, dado un sistema de m ecuaciones con $n > m$ incógnitas

$$g_1(x_1, x_2, \cdots x_k, x_{k+1}, \cdots x_n) = 0$$

$$g_2(x_1, x_2, \cdots x_k, x_{k+1}, \cdots x_n) = 0$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$g_m(x_1, x_2, \cdots x_k, x_{k+1}, \cdots x_n) = 0$$

del que se conoce una solución particular $(a_1, a_2, \cdots, a_n)$, se plantea el problema análogo de estudiar cuando es posible despejar localmente $m = n - k$ variables, por ejemplo $(x_{k+1}, \cdots x_n)$ en función de las restantes $(x_1, x_2, \cdots x_k)$.

En lo que sigue, si $k \in \{1, 2, \cdots n-1\}$, y $m = n - k$, conviene considerar $\mathbb{R}^n$ identificado con $\mathbb{R}^k \times \mathbb{R}^m$, y así un elemento genérico de $\mathbb{R}^n$ se escribirá en la forma $(\mathbf{x}, \mathbf{y})$ con $\mathbf{x} = (x_1, x_2, \cdots x_k) \in \mathbb{R}^k$, $\mathbf{y} = (y_1, y_2, \cdots y_m)$. Con estos convenios de notación, usando notación vectorial, $\mathbf{g} = (g_1, g_2, \cdots, g_m)$, $\mathbf{x} = (x_1, \cdots x_k)$, $\mathbf{y} = (x_{k+1}, \cdots x_n)$, $\mathbf{a} = (a_1, \cdots a_k)$, $\mathbf{b} = (a_{k+1}, \cdots a_n)$, el problema de existencia de función implícita se plantea así: Dada la ecuación $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$, de la que se conoce una solución particular $\mathbf{g}(\mathbf{a}, \mathbf{b}) = 0$, se trata de estudiar cuando es posible despejar localmente, en un entorno de $(\mathbf{a}, \mathbf{b})$, a la variable vectorial $\mathbf{y}$ en función de la variable vectorial $\mathbf{x}$.

Definición 8.15 *Si* $\mathbf{g} : \Omega \to \mathbb{R}^m$ *está definida en un abierto* $\Omega \subset \mathbb{R}^k \times \mathbb{R}^m$ *y* $(\mathbf{a}, \mathbf{b}) \in \Omega$ *es un punto que cumple* $\mathbf{g}(\mathbf{a}, \mathbf{b}) = 0$, *se dice que la ecuación* $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$ *define, en un entorno de* $(\mathbf{a}, \mathbf{b})$, *a la variable* $\mathbf{y} \in \mathbb{R}^m$ *como función implícita de la variable* $\mathbf{x} \in \mathbb{R}^k$ *si ocurre lo siguiente: Existe un entorno abierto* A *de* $\mathbf{a}$, *y un entorno abierto*

B de $\mathbf{b}$, tales que $A \times B \subset \Omega$, y para cada $\mathbf{x} \in A$ hay un único $\mathbf{y} = \mathbf{f}(\mathbf{x}) \in B$ que verifica $\mathbf{g}(\mathbf{x}, \mathbf{f}(\mathbf{x})) = 0$. En este caso se dice que $\mathbf{f} : A \to B$ es la función implícita que $A \times B$ determina en la ecuación $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$.

Según esta definición existirá función implícita cuando se pueda garantizar la existencia de un entorno abierto $A \times B \subset \Omega$ de $(\mathbf{a}, \mathbf{b})$ tal que $\{(\mathbf{x}, \mathbf{y}) \in A \times B : \mathbf{g}(\mathbf{x}, \mathbf{y}) = 0\}$ sea la gráfica de alguna función $\mathbf{f} : A \to B$.

Antes de formular y demostrar el teorema de existencia de funciones implícitas 8.16 una sencilla reflexión preliminar revelará el papel que desempeña la hipótesis central de este teorema (la no anulación, en el punto correspondiente, del jacobiano de $\mathbf{g}$ respecto a las variables que se desean despejar).

En el caso particularmente simple de que las ecuaciones sean lineales

$$g_i(x_1, x_2, \cdots x_k, x_{k+1}, \cdots x_n) = \sum_{j=1}^{k} a_{ij}x_j + \sum_{j=k+1}^{n} a_{ij}x_j, \quad 1 \leq i \leq n-k$$

para poder resolver el sistema respecto a las variables $x_{k+1}, x_{k+2} \cdots x_n$ sabemos que hay que requerir que la matriz cuadrada $\{a_{i,k+j} : 1 \leq i, j \leq m\}$ tenga determinante no nulo. En este caso, si $\mathbf{g} : \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^m$ es la aplicación lineal de componentes $(g_1 \cdots g_m)$, fijado $\mathbf{x} \in \mathbb{R}^k$, la diferencial de la aplicación afín $\mathbf{y} \to g_{\mathbf{x}}(\mathbf{y}) = g(\mathbf{x}, \mathbf{y})$ es una aplicación lineal invertible $L : \mathbb{R}^m \to \mathbb{R}^m$ ya que el determinante de su matriz $(a_{ik+j})_{1 \leq i,j \leq m})$ no es nulo.

Cuando el sistema no es lineal, para que quede defina una función implícita en un entorno de $(\mathbf{a}, \mathbf{b})$ se deberá reemplazar la condición anterior por la condición de que la matriz cuadrada $(D_{k+i}\mathbf{g}_j(\mathbf{a}, \mathbf{b}))_{1 \leq i,j \leq m}$ tenga determinante no nulo, lo que significa que la función parcial $\mathbf{g}_{\mathbf{a}} : \mathbf{y} \to \mathbf{g}(\mathbf{a}, \mathbf{y})$ tiene en $\mathbf{b}$ una diferencial invertible.

De un modo heurístico e intuitivo podemos pensar así: Si $\mathbf{g}$ es diferenciable en $(\mathbf{a}, \mathbf{b})$ y su diferencial $L = d\mathbf{g}(\mathbf{a}, \mathbf{b})$ tiene la propiedad de que en la ecuación $L(\mathbf{x}, \mathbf{y}) = 0$ se puede despejar a $\mathbf{y}$ en función $\mathbf{x}$, cabe esperar que ocurra lo mismo en la ecuación original $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$, después de restringir esta ecuación a un entorno suficientemente pequeño de $(\mathbf{a}, \mathbf{b})$. Como la ecuación lineal $L(\mathbf{x}, \mathbf{y}) = 0$ se escribe en forma de sistema lineal

$$\sum_{i=1}^{n} D_i g_j(\mathbf{a}, \mathbf{b}) x_i = 0; \quad 1 \leq j \leq m$$

la no anulación del determinante de la matriz $(D_{k+i}g_j(\mathbf{a}, \mathbf{b}))_{1 \leq i,j \leq m}$ es la hipótesis que permite despejar en este sistema lineal a las variables $(x_{k+1}, \cdots, x_n)$ en función de las variables $(x_1, x_2, \cdots, x_k)$. Esta es la hipótesis crucial que interviene en el teorema de la función implícita.

Teorema 8.16 *Sea $\mathbf{g} : \Omega \to \mathbb{R}^m$ una función de clase $C^p(\Omega)$, $(p \geq 1q1)$ definida en un abierto $\Omega \subset \mathbb{R}^k \times \mathbb{R}^m$ y $(\mathbf{a}, \mathbf{b}) \in \Omega$ un punto que satisface $\mathbf{g}(\mathbf{a}, \mathbf{b}) = 0$. Si el determinante jacobiano de las componentes de $\mathbf{g}$ respecto a las variables $(y_1, y_2 \cdots y_m)$ no se anula en el punto $(\mathbf{a}, \mathbf{b})$*

$$\frac{D(g_1, g_2, \cdots g_m)}{D(y_1, y_2, \cdots y_m)}(\mathbf{a}, \mathbf{b}) \neq 0$$

entonces la ecuación $\mathbf{g}(\mathbf{x},\mathbf{y}) = 0$ *define, en un entorno de* $(\mathbf{a},\mathbf{b})$*, a la variable* $\mathbf{y} = (y_1, y_2, \cdots y_m)$ *como función implícita de la variable* $\mathbf{x} = (x_1, x_2, \cdots x_k)$*, es decir existe* $A \times B \subset \Omega$*, entorno abierto de* $(\mathbf{a},\mathbf{b})$*, tal que para cada* $\mathbf{x} \in A$ *hay un único* $\mathbf{y} = \mathbf{f}(\mathbf{x}) \in B$ *que verifica* $\mathbf{g}(\mathbf{x},\mathbf{y}) = 0$*. La función implícita* $\mathbf{f} : A \to B$ *es de clase* $C^p(A)$

DEM: La función $\mathbf{G} : \Omega \to \mathbb{R}^k \times \mathbb{R}^m$, definida por $\mathbf{G}(\mathbf{x},\mathbf{y}) = (\mathbf{x}, \mathbf{g}(\mathbf{x},\mathbf{y}))$ es de clase $C^p(\Omega)$ y verifica

$$\det \mathbf{G}'(\mathbf{a},\mathbf{b}) = \frac{D(g_1, g_2, \cdots g_m)}{D(y_1, y_2, \cdots y_m)}(\mathbf{a},\mathbf{b}) \neq 0$$

Aplicando a la función $\mathbf{G}$ el teorema de la función inversa 8.13, se obtienen abiertos $U, V \subset \mathbb{R}^n$ con $(\mathbf{a},\mathbf{b}) \in U$ y $\mathbf{G}(\mathbf{a},\mathbf{b}) = (\mathbf{a},\mathbf{0}) \in V$, tales que $\mathbf{G}|_U : U \to V$ es una biyección con inversa $\mathbf{F} : V \to U$ de clase C^p. No hay inconveniente suponer que $U = A_0 \times B_0$, donde A_0 es un entorno abierto de $\mathbf{a}$, y B_0 un entorno abierto de $\mathbf{b}$.

En lo que sigue, para cada $(\mathbf{u},\mathbf{v}) \in V$ escribimos $\mathbf{F}(\mathbf{u},\mathbf{v}) = (\mathbf{F}_1(\mathbf{u},\mathbf{v}), \mathbf{F}_2(\mathbf{u},\mathbf{v}))$ con $\mathbf{F}_1(\mathbf{u},\mathbf{v}) \in \mathbb{R}^k$, $\mathbf{F}_2(\mathbf{u},\mathbf{v}) \in \mathbb{R}^m$. Como $\mathbf{G}$ deja fijas las primeras variables $(x_1, x_2, \cdots x_k)$, lo mismo le ocurre a su inversa $\mathbf{F}$, luego $\mathbf{F}_1(\mathbf{u},\mathbf{v}) = \mathbf{u}$.

Para cada $(\mathbf{u},\mathbf{v}) \in V$ se cumple

$$(\mathbf{u},\mathbf{v}) = \mathbf{G}(\mathbf{F}(\mathbf{u},\mathbf{v})) = \mathbf{G}(\mathbf{u}, \mathbf{F}_2(\mathbf{u},\mathbf{v})) = (\mathbf{u}, \mathbf{g}(\mathbf{u}, \mathbf{F}_2(\mathbf{u},\mathbf{v}))$$

luego $\mathbf{g}(\mathbf{u}, \mathbf{F}_2(\mathbf{u},\mathbf{v})) = \mathbf{v}$ para cada $(\mathbf{u},\mathbf{v}) \in V$.

Entonces la función $\mathbf{f} : A \to \mathbb{R}^m$, definida en $A = \{\mathbf{x} \in \mathbb{R}^k : (\mathbf{x},\mathbf{0}) \in V\}$ por $\mathbf{f}(\mathbf{x}) = \mathbf{F}_2(\mathbf{x},\mathbf{0})$ cumple que $\mathbf{g}(\mathbf{x},\mathbf{f}(\mathbf{x})) = 0$ para todo $\mathbf{x} \in A$. Obsérvese que A es un entorno abierto de $\mathbf{a}$ (porque la función $\mathbf{x} \to (\mathbf{x}, 0)$ es continua y $(\mathbf{a},\mathbf{0}) \in V$). Además, $A \subset A_0$ y $\mathbf{f}(A) \subset B_0$, ya que

$$\mathbf{x} \in A \Rightarrow (\mathbf{x},\mathbf{0}) \in V \Rightarrow \mathbf{F}(\mathbf{x},\mathbf{0}) = (\mathbf{x},\mathbf{f}(\mathbf{x})) \in U = A_0 \times B_0$$

Con $B = B_0$ se cumple que $\mathbf{f}(A) \subset B$, y $\mathbf{g}(\mathbf{x},\mathbf{f}(\mathbf{x})) = 0$ para todo $\mathbf{x} \in A$. Para concluir la demostración basta ver que para cada $\mathbf{x} \in A$ hay un único $\mathbf{y} \in B$ que verifica $\mathbf{g}(\mathbf{x},\mathbf{y}) = \mathbf{0}$. Efectivamente, si $\mathbf{y} \in B$ y $\mathbf{g}(\mathbf{x},\mathbf{y}) = 0$ se cumple

$$\mathbf{G}(\mathbf{x},\mathbf{y}) = (\mathbf{x},\mathbf{g}(\mathbf{x},\mathbf{y})) = (\mathbf{x},\mathbf{0}), \quad \mathbf{G}(\mathbf{x},\mathbf{f}(\mathbf{x})) = (\mathbf{x},\mathbf{g}(\mathbf{x},\mathbf{f}(\mathbf{x}))) = (\mathbf{x},\mathbf{0})$$

Entonces, teniendo en cuenta que $\mathbf{G}|_U$ es inyectiva, y que los puntos $(\mathbf{x},\mathbf{y})$, $(\mathbf{x},\mathbf{f}(\mathbf{x})$ pertenecen a $A \times B \subset A_0 \times B_0 = U$ se concluye que $\mathbf{y} = \mathbf{f}(\mathbf{x})$. Finalmente $\mathbf{f}$ es de clase $C^p(A)$ porque $\mathbf{F}$ es de clase $C^p(V)$. ■

NOTA: En el teorema 8.16 se ha supuesto para simplificar la notación, la hipótesis apropiada para obtener a las últimas m variables como funciones implícitas de las k primeras. Análogamente, si $(i_1, i_2, \cdots, i_m)$ es un subconjunto de $\{1, 2, \cdots n\}$, con m elementos y se supone que

$$\frac{D(g_1, g_2, \cdots g_m)}{D(x_{i_1}, x_{i_2}, \cdots x_{i_m})}(\mathbf{p}) \neq 0$$

entonces la ecuación vectorial $\mathbf{g}(x_1, x_2, \cdots, x_n) = \mathbf{0}$ define, en un entorno de $\mathbf{p}$, a las variables $(x_{i_1}, x_{i_2}, \cdots x_{i_m})$ como funciones implícitas de las restantes.

8.3. Cálculo con funciones implícitas e inversas

Derivadas parciales de funciones implícitas. En las condiciones del teorema 8.16 cuando $\mathbf{g}$ es de clase C^p, aunque no se conozca una fórmula explícita para la función implícita $\mathbf{f}$ es posible calcular sus derivadas parciales sucesivas (hasta las de orden p) en el punto concreto $(\mathbf{a}, \mathbf{b})$. Para simplificar la exposición del método consideremos el caso particular $n = 4$ y $k = 2$, denotando (x, y, u, v) a las variables de la función $\mathbf{g}$. Si f_1, f_2 son las componentes de la función implícita: $\mathbf{f} : A \to B$ en lo que sigue resultará cómodo designarlas con la notación más flexible $u(x, y) = f_1(x, y)$, $v(x, y) = f_2(x, y)$. Para todo $(x, y) \in A$ se cumple

$$g_1(x, y, u(x, y), v(x, y)) = 0, \qquad g_2(x, y, u(x, y), v(x, y)) = 0.$$

Derivando las dos ecuaciones respecto a la variable x y respecto a la variable y se obtienen las identidades i), ii), iii), iv).

$$\text{i)}\ \frac{\partial g_1}{\partial x} + \frac{\partial g_1}{\partial u}u_x + \frac{\partial g_1}{\partial v}v_x = 0; \qquad \text{iiii)}\ \frac{\partial g_1}{\partial y} + \frac{\partial g_1}{\partial u}u_y + \frac{\partial g_1}{\partial v}v_y = 0;$$

$$\text{ii)}\ \frac{\partial g_2}{\partial x} + \frac{\partial g_2}{\partial u}u_x + \frac{\partial g_2}{\partial v}v_x = 0; \qquad \text{iv)}\ \frac{\partial g_2}{\partial y} + \frac{\partial g_2}{\partial u}u_y + \frac{\partial g_2}{\partial v}v_y = 0;$$

donde las derivadas parciales de g_1 y g_2 se suponen evaluadas $(x, y, u(x, y), v(x, y))$, y las derivadas parciales u_x, u_y, v_x, v_y en el punto (x, y). Cuando $(x, y) = \mathbf{a}$ se tiene $(u(\mathbf{a}), v(\mathbf{a})) = \mathbf{b}$ y si las ecuaciones i), ii) se particularizan en el punto $\mathbf{p} = (\mathbf{a}, \mathbf{b})$ resulta

$$\text{i')}\ \frac{\partial g_1}{\partial x}(\mathbf{p}) + \frac{\partial g_1}{\partial u}(\mathbf{p})u_x(\mathbf{a}) + \frac{\partial g_1}{\partial v}(\mathbf{p})v_x(\mathbf{a}) = 0;$$

$$\text{ii')}\ \frac{\partial g_2}{\partial x}(\mathbf{p}) + \frac{\partial g_2}{\partial u}(\mathbf{p})u_x(\mathbf{a}) + \frac{\partial g_2}{\partial v}(\mathbf{p})v_x(\mathbf{a}) = 0;$$

Este sistema lineal de dos ecuaciones con dos incógnitas $u_x(\mathbf{a})$, $v_x(\mathbf{a})$, con determinante $\frac{D(g_1,g_2)}{D(u,v)}(\mathbf{p}) \neq 0$, permite calcular los valores $u_x(\mathbf{a})$, $v_x(\mathbf{a})$. Análogamente, usando las ecuaciones iii) y iv), se pueden calcular $u_y(\mathbf{a})$ y $v_y(\mathbf{a})$.

Si suponemos que $p \geq 2$ podemos seguir derivando respecto a x y respecto a y las identidades i), ii), iii) y iv). Para calcular, en el punto $\mathbf{a}$, las derivadas parciales segundas u_{xx}, u_{xy}, u_{yy}, v_{xx}, v_{xy}, v_{yy}, se derivan las identidades i), ii), iii) y iv) respecto a las variables x e y, se particulariza el resultado para $(x, y) = \mathbf{a}$, $(u, v) = \mathbf{b}$ y se obtienen sistemas de ecuaciones lineales que permiten calcular los valores particulares de estas derivadas segundas. Así por ejemplo, derivando i) y ii) respecto a x

$$\text{a)}\ \frac{\partial^2 g_1}{\partial x^2} + \frac{\partial^2 g_1}{\partial u \partial x}u_x + \frac{\partial^2 g_1}{\partial v \partial x}v_x + \left[\frac{\partial^2 g_1}{\partial x \partial u} + \frac{\partial^2 g_1}{\partial^2 u}u_x + \frac{\partial^2 g_1}{\partial v \partial u}v_x\right] u_x + \frac{\partial g_1}{\partial u}u_{xx} +$$

$$+ \left[\frac{\partial^2 g_1}{\partial x \partial v} + \frac{\partial^2 g_1}{\partial u \partial v}u_x + \frac{\partial^2 g_1}{\partial^2 v}v_x\right] v_x + \frac{\partial g_1}{\partial v}v_{xx} = 0$$

b) $$\frac{\partial^2 g_2}{\partial x^2} + \frac{\partial^2 g_2}{\partial u \partial x} u_x + \frac{\partial^2 g_2}{\partial v \partial x} v_x + \left[\frac{\partial^2 g_2}{\partial x \partial u} + \frac{\partial^2 g_2}{\partial^2 u} u_x + \frac{\partial^2 g_2}{\partial v \partial u} v_x \right] u_x + \frac{\partial g_2}{\partial u} u_{xx} +$$

$$+ \left[\frac{\partial^2 g_2}{\partial x \partial v} + \frac{\partial^2 g_2}{\partial u \partial v} u_x + \frac{\partial^2 g_2}{\partial^2 v} v_x \right] v_x + \frac{\partial g_2}{\partial v} v_{xx} = 0$$

donde las derivadas parciales de g_1 y g_2 están evaluadas en $(x, y, u(x, y), v(x, y))$, y las derivadas parciales de u y v están evaluadas en (x, y).

Sustituyendo $(x, y) = \mathbf{a}$, $(u, v) = \mathbf{b}$, $(x, y, u, v) = \mathbf{p}$, y utilizando los valores ya calculados $u_x(\mathbf{a})$, $v_x(\mathbf{a})$ se llega a un sistema lineal de dos ecuaciones con dos incógnitas $u_{xx}(\mathbf{a})$, $v_{xx}(\mathbf{a})$, que tiene determinante $\frac{D(g_1,g_2)}{D(u,v)}(\mathbf{p}) \neq 0$, y su solución proporciona las derivadas segundas $u_{xx}(\mathbf{a})$, $v_{xx}(\mathbf{a})$.

Procediendo en forma similar, derivando las identidades i) y ii) respecto a la variable y, al particularizar el resultado en el punto $\mathbf{p}$, se llega a otro sistema lineal, con el mismo determinante no nulo, que permite calcular $u_{xy}(\mathbf{a})$, $v_{xy}(\mathbf{a})$ (estos valores también se pueden calcular derivando respecto a x las identidades iii) y iv), sustituyendo luego $(x, y, u, v) = \mathbf{p}$ y resolviendo el correspondiente sistema lineal). Finalmente $u_{yy}(\mathbf{b})$, $v_{yy}(\mathbf{b})$ se calculan con el mismo método, derivando respecto a y las identidades iii) y iv).

Cuando $p \geq 3$ se puede continuar con el procedimiento: Se deriva respecto a x y respecto a y cada una de las identidades obtenidas en la etapa anterior, se particulariza el resultado en el punto $\mathbf{p}$ y se resuelven luego los correspondientes sistemas lineales de dos ecuaciones con dos incógnitas, cuyo determinante siempre es el mismo, $\frac{D(g_1,g_2)}{D(u,v)}(\mathbf{p}) \neq 0$. ∎

Derivadas parciales de funciones inversas. En las condiciones del teorema 8.12, cuando $\mathbf{f}$ es de clase C^m, aunque no se conozca una fórmula explícita para la función inversa $\mathbf{g}$ también es posible calcular sus derivadas parciales sucesivas (hasta las de orden m) en un punto concreto $\mathbf{b} = \mathbf{f}(\mathbf{a})$. La técnica que acabamos de exponer para calcular derivadas parciales de funciones implícitas también se puede utilizar ahora ya que una inversa local de la función $\mathbf{y} = \mathbf{f}(\mathbf{x})$ se puede considerar como función implícita definida por la ecuación $\mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, donde $\mathbf{g}(\mathbf{x}, \mathbf{y}) = \mathbf{f}(\mathbf{x}) - \mathbf{y}$.

Como conviene adquirir destreza en estas técnicas de cálculo insistimos con ella en el contexto de las funciones inversas. Para simplificar la exposición lo hacemos en el caso particular $n = 2$.

Si f_1, f_2 son las componentes de $\mathbf{f}$, se suele decir que las ecuaciones $u = f_1(x, y)$, $v = f_2(x, y)$ establecen una transformación de un abierto A del plano de las variables (x, y) sobre un abierto B del plano de las variables (u, v). La función inversa $\mathbf{g}$, transforma cada $(u, v) \in B$ en el punto $(x, y) \in A$ dado por $x = g_1(u, v)$, $y = g_2(u, v)$.

En lo que sigue designamos las componentes de $\mathbf{g}$ con la notación más cómoda $x(u, v) = g_1(u, v)$, $y(u, v) = g_2(u, v)$. Para todo $(u, v) \in B$ se cumple

$$f_1(x(u, v), y(u, v)) = u, \qquad f_2(x(u, v), y(u, v)) = v.$$

Derivando las dos ecuaciones respecto a la variable u y respecto a la variable v se obtienen las identidades i), ii), iii), iv).

$$\text{i)}\quad \frac{\partial f_1}{\partial x}x_u + \frac{\partial f_1}{\partial y}y_u = 1; \qquad \text{iii)}\quad \frac{\partial f_1}{\partial x}x_v + \frac{\partial f_1}{\partial y}y_v = 0;$$

$$\text{ii)}\quad \frac{\partial f_2}{\partial x}x_u + \frac{\partial f_2}{\partial y}y_u = 0; \qquad \text{iv)}\quad \frac{\partial f_2}{\partial x}x_v + \frac{\partial f_2}{\partial y}y_v = 1;$$

donde las derivadas parciales de f_1 y f_2 se suponen evaluadas en $(x(u,v), y(u,v))$, y las derivadas parciales x_u, y_u en (u,v). Sustituyendo los valores $(u,v) = \mathbf{b}$, $(x(u,v), y(u,v)) = \mathbf{a}$ las ecuaciones i),ii) se concretan en

$$\text{i')}\quad \frac{\partial f_1}{\partial x}(\mathbf{a})x_u(\mathbf{b}) + \frac{\partial f_1}{\partial y}(\mathbf{a})y_u(\mathbf{b}) = 1$$

$$\text{ii')}\quad \frac{\partial f_2}{\partial x}(\mathbf{a})x_u(\mathbf{b}) + \frac{\partial f_2}{\partial y}(\mathbf{a})y_u(\mathbf{b}) = 0$$

Este sistema lineal de dos ecuaciones con dos incógnitas $x_u(\mathbf{b})$, $y_u(\mathbf{b})$, con determinante $\det \mathbf{f}'(\mathbf{a}) \neq 0$, permite calcular los valores $x_u(\mathbf{b})$, $y_u(\mathbf{b})$. Análogamente, usando las ecuaciones iii) y iv), se pueden calcular $x_v(\mathbf{b})$ y $y_v(\mathbf{b})$.

Si suponemos que $m \geq 2$ podemos seguir derivando respecto a u y respecto a v las identidades i), ii), iii) y iv). Para calcular las derivadas parciales segundas $x_{uu}(\mathbf{b})$, $y_{uu}(\mathbf{b})$ se derivan las identidades i) y ii) respecto a la variable u y se obtiene

$$\text{a)}\quad \left[\frac{\partial^2 f_1}{\partial x^2}x_u + \frac{\partial^2 f_1}{\partial y \partial x}y_u\right]x_u + \frac{\partial f_1}{\partial x}x_{uu} + \left[\frac{\partial^2 f_1}{\partial x \partial y}x_u + \frac{\partial^2 f_1}{\partial y^2}y_u\right]y_u + \frac{\partial f_1}{\partial y}y_{uu} = 0$$

$$\text{b)}\quad \left[\frac{\partial^2 f_2}{\partial x^2}x_u + \frac{\partial^2 f_2}{\partial y \partial x}y_u\right]x_u + \frac{\partial f_2}{\partial x}x_{uu} + \left[\frac{\partial^2 f_2}{\partial x \partial y}x_u + \frac{\partial^2 f_2}{\partial y^2}y_u\right]y_u + \frac{\partial f_2}{\partial y}y_{uu} = 0$$

donde las derivadas parciales de f_1 y f_2 están evaluadas en $(x(u,v), y(u,v))$, y las derivadas parciales x_u, y_u, x_{uu}, y_{uu} están evaluadas en (u,v).

Sustituyendo $(u,v) = \mathbf{b}$, $(x(u,v), y(u,v)) = \mathbf{a}$, y utilizando los valores ya calculados $x_u(\mathbf{b})$, $y_u(\mathbf{b})$ se obtiene un sistema lineal de determinante $\det \mathbf{f}'(\mathbf{a}) \neq 0$ cuya solución proporciona las derivadas segundas $x_{uu}(\mathbf{b})$, $y_{uu}(\mathbf{b})$.

Análogamente, si se derivan las identidades i) y ii) respecto a la variable v, y se concreta el resultado en el punto $\mathbf{b}$, resulta otro sistema lineal, con determinante $\det \mathbf{f}'(\mathbf{a}) \neq 0$, que permite calcular $x_{uv}(\mathbf{b})$, $y_{uv}(\mathbf{b})$ (los valores $x_{vu}(\mathbf{b}) = x_{uv}(\mathbf{b})$, $y_{vu}(\mathbf{b}) = y_{uv}(\mathbf{b})$ también se pueden obtener derivando respecto a u las identidades iii) y iv), sustituyendo luego $(u,v) = \mathbf{b}$ y resolviendo el correspondiente sistema lineal).

Finalmente $x_{vv}(\mathbf{b})$, $y_{vv}(\mathbf{b})$ se calculan usando el mismo método, empezando con las derivadas parciales respecto a v de las identidades iii) y iv).

Cuando $m \geq 3$ se puede continuar con el procedimiento. Así por ejemplo, para calcular las derivadas terceras

$$x_{uuv}(\mathbf{b}) = \frac{\partial^3 x}{\partial v \partial u^2}(\mathbf{b}), \qquad y_{uuv}(\mathbf{b}) = \frac{\partial^3 y}{\partial v \partial u^2}(\mathbf{b})$$

habría que derivar las identidades a), b) respecto a la variable v, sustituir en el resultado los valores, en el punto $\mathbf{b}$, de las funciones $x(u,v)$, $y(u,v)$ y de sus derivadas parciales primeras y segundas (calculados en las etapas anteriores) y resolver finalmente un sistema lineal, cuyo determinante seguirá siendo $\det \mathbf{f}'(\mathbf{a}) \neq 0$.

8.4. Cambio de variable en el cálculo diferencial

Dado un C^m-difeomorfismo $\mathbf{P} : U \to \Omega$ entre dos abiertos $U, \Omega \subset \mathbb{R}^n$ cada punto $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in \Omega$ es la imagen de un único $\mathbf{u} = (u_1, u_2, \cdots, u_n)$ y se suele decir que $(u_1, u_2, \cdots, u_n)$ son las coordenadas curvilíneas de $\mathbf{x} = \mathbf{P}(\mathbf{u})$ en el sistema de coordenadas curvilíneas asociado a $\mathbf{P}$.

Ejemplos típicos son las coordenadas polares en el plano, y también las coordenadas cilíndricas y las coordenadas esféricas del espacio $\mathbb{R}^3$.
En el plano $\mathbb{R}^2$ las *coordenadas polares* son las asociadas a $\mathbf{P}(r, \theta) = (r\cos\theta, r\,\mathrm{sen}\,\theta)$.
En el espacio $\mathbb{R}^3$ las *coordenadas cilíndricas* son las asociadas a

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

y las *coordenadas esféricas* las asociadas a

$$\mathbf{P}(\rho, \theta, \varphi) = (\rho\cos\varphi\cos\theta, \rho\cos\varphi\sin\theta, \rho\sin\varphi)$$

En cada caso se supone $\mathbf{P}$ definida en un abierto U sobre el que es inyectiva.

Dada una función $f : \Omega \to \mathbb{R}$ de la variable $\mathbf{x} \in \Omega$ si efectuamos la sustitución $\mathbf{x} = \mathbf{P}(\mathbf{u})$ se obtiene la expresión $F = f \circ \mathbf{P}$ de la función f en términos de las coordenadas curvilíneas $(u_1, u_2, \cdots u_n) \in U$. Aunque desde el punto de vista formal F y f son funciones distintas, sin embargo, desde el punto de vista de las posibles aplicaciones e interpretaciones conviene considerarlas como diferentes expresiones analíticas de la misma aplicación, que surgen al adoptar distintos sistemas de coordenadas curvilíneas para representar numéricamente los puntos de su dominio. Así, de acuerdo con este convenio se suele decir que $F(u_1, u_2, \cdots, u_n)$ es la expresión analítica de f en el sistema de coordenadas curvilíneas asociado a $\mathbf{P}$.

En esta situación, un problema que se plantea con frecuencia en el cálculo diferencial es el de calcular las derivadas parciales sucesivas de f, en un punto genérico, en términos de las derivadas parciales de la nueva función F. Cuando conocemos explícitamente las ecuaciones (fórmulas) para la función inversa $\mathbf{P}^{-1}$ el cálculo se puede hacer fácilmente usando la regla de la cadena ya que $f = F \circ \mathbf{P}^{-1}$. Sin embargo, ocurre a menudo que no se conocen fórmulas explícitas para la inversa, o se conocen pero son engorrosas de manejar.

Para abordar este asunto exponemos a continuación, en el caso particular de los cambios de variable a coordenadas polares, una técnica sistemática que se puede aplicar, con modificaciones obvias, a otros casos.

Sea $f(x, y)$ una función real de dos variables reales definida en un abierto $\Omega \subset \mathbb{R}^2$. Efectuando la sustitución $x = r\cos\theta$, $y = r\,\mathrm{sen}\,\theta$, se obtiene la expresión de f en coordenadas polares:

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

es decir $F = f \circ \mathbf{P}$, donde $\mathbf{P} : \mathbb{R}^2 \to \mathbb{R}^2$ es la transformación $\mathbf{P}(r, \theta) = (r\cos\theta, r\,\mathrm{sen}\,\theta)$.

Se supone que $\Omega = \mathbf{P}(U)$, donde $U \subset \mathbb{R}^2$ es un abierto tal que $\mathbf{g} = \mathbf{P}|_U$ es inyectiva (lo que lleva implícito que $(0, 0) \notin \Omega$). Así se puede asegurar que las coordenadas polares (r, θ) de un punto $(x, y) \in \Omega$ quedan unívocamente determinadas por la condición $(r,\theta) \in U$, lo que permite recuperar f a partir de F con el cambio de variable inverso $(r,\theta) = \mathbf{P}^{-1}(x,y)$.

Si suponemos que f es diferenciable en Ω, en virtud de la regla de la cadena aplicada a la composición $F = f \circ \mathbf{P}$, se tendrá que las derivadas parciales f_x, f_y deben satisfacer el sistema lineal

$$F_r = \cos\theta f_x + \operatorname{sen}\theta f_y, \quad F_\theta = -r\operatorname{sen}\theta f_x + r\cos\theta f_y$$

cuyo determinante es $r > 0$ (ya que $(0,0) \notin \Omega$). Resolviendo el sistema se obtienen

$$f_x = \cos\theta F_r - \frac{\operatorname{sen}\theta}{r}F_\theta, \quad f_y = \operatorname{sen}\theta F_r + \frac{\cos\theta}{r}F_\theta$$

Vemos así que la regla para calcular f_x consiste en aplicar el operador diferencial

$$A = \cos\theta\frac{\partial}{\partial r} - \frac{\operatorname{sen}\theta}{r}\frac{\partial}{\partial\theta}$$

a la función que resulta de expresar f en coordenadas polares. Análogamente, la regla para calcular f_y consiste en aplicar el operador diferencial

$$B = \operatorname{sen}\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}$$

a la función que resulta de expresar f en coordenadas polares.

Ejemplo 8.17 *El operador de Laplace en coordenadas polares*

Se trata de expresar en coordenadas polares la laplaciana $\Delta f = f_{xx} + f_{yy}$ de una función f que es dos veces diferenciable en un abierto $\Omega \subset \mathbb{R}^2$. Para obtener f_{xx} en coordenadas polares debemos aplicar el operador A a la expresión de f_x en coordenadas polares obtenida arriba, es decir

$$f_{xx} = \left[\cos\theta\frac{\partial}{\partial r} - \frac{\operatorname{sen}\theta}{r}\frac{\partial}{\partial\theta}\right]\left(\cos\theta F_r - \frac{\operatorname{sen}\theta}{r}F_\theta\right)$$

Efectuando las operaciones indicadas se obtiene

$$f_{xx} = (\cos^2\theta)F_{rr} + \frac{\operatorname{sen}^2\theta}{r^2}F_{\theta\theta} - \alpha F_{\theta r} + \beta F_\theta + \frac{\operatorname{sen}^2\theta}{r}F_r$$

donde

$$\alpha = \frac{2\operatorname{sen}\theta\cos\theta}{r}, \quad \beta = \frac{2\operatorname{sen}\theta\cos\theta}{r^2}.$$

Aplicando el operador B a la expresión de f_y en coordenadas polares se obtiene:

$$f_{yy} = \left[\operatorname{sen}\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\right]\left(\operatorname{sen}\theta F_r + \frac{\cos\theta}{r}F_\theta\right)$$

y efectuando los cálculos se llega a

$$f_{yy} = (\operatorname{sen}^2\theta)F_{rr} + \frac{\cos^2\theta}{r^2}F_{\theta\theta} + \alpha F_{\theta r} - \beta F_\theta + \frac{\cos^2\theta}{r}F_r$$

donde α y β son los mismos valores que aparecieron en el cálculo de f_{xx}.

Sumando las expresiones obtenidas para f_{xx} y f_{yy} se obtiene la fórmula para el operador de Laplace Δf en coordenadas polares:

$$\Delta(f) = \frac{\partial^2 F}{\partial r^2} + \frac{1}{r}\frac{\partial F}{\partial r} + \frac{1}{r^2}\frac{\partial^2 F}{\partial \theta^2}$$

■

Una función $f : \Omega \to \mathbb{R}$ que satisface la ecuación de Laplace $\Delta f = 0$ se dice que es *armónica*. Usando el cálculo anterior es fácil ver que son armónicas las funciones que en coordenadas polares se expresan en la forma $r^n \cos n\theta$, $r^n \operatorname{sen} n\theta$.

Ejemplo 8.18 *Ecuación de la cuerda vibrante*

La ecuación en derivadas parciales

$$\frac{\partial^2 f}{\partial x^2} = \frac{1}{\alpha^2}\frac{\partial^2 f}{\partial t^2}$$

llamada ecuación de onda unidimensional fue considerada por John Bernoulli alrededor de 1727 y varios años después por Jean Le Ron d'Alembert al estudiar el movimiento de una cuerda vibrante: $f(x,t)$ representa el desplazamiento vertical de un punto de abscisa x en el instante t de una cuerda que vibra. Con un cambio de variable lineal $x = Au + Bv$, $t = Cu + Dv$ la función $f(x,t)$ se transforma en $F(u,v) = f(Au + Bv, Cu + Dv)$. Usando la regla de la cadena se obtiene

$$F_{uv} = ABf_{xx} + (AD + BC)f_{xt} + CDf_{tt}$$

La expresión anterior se simplifica eligiendo A, B, C, D de modo que se anule el coeficiente de f_{xt}. Esto se consigue con $A = B = 1/2$, y $C = -D = 1/(2\alpha)$. Así, con el cambio de variable

$$x = \frac{1}{2}(u+v), \quad y = \frac{1}{2\alpha}(u-v)$$

se obtiene

$$4F_{uv} = f_{xx} - \frac{1}{\alpha^2}\frac{\partial^2 f}{\partial t^2}$$

y la ecuación del enunciado se transforma en $F_{uv} = 0$ cuyas soluciones son las funciones de la forma $F(u,v) = \varphi(u) + \psi(v)$, donde φ, ψ son funciones de una variable real dos veces derivables. Deshaciendo el cambio de variable con la sustitución $u = x + \alpha t$, $v = x - \alpha t$, se llega a que las soluciones de la ecuación original son las funciones de la forma

$$f(x,t) = \varphi(x + \alpha t) + \psi(x - \alpha t)$$

8.5. Ejercicios resueltos

Con el siguiente ejercicio se mejora el teorema de la aplicación abierta expuesto en 8.8:

Ejercicio 8.19 *Sea* $\mathbf{f} : \Omega \to \mathbb{R}^n$ *diferenciable en* $\mathbf{x} \in \Omega$ *con* $\det \mathbf{f}'(\mathbf{x}) \neq 0$, *y* $\delta(\mathbf{x}) = \frac{1}{2} \|d\mathbf{f}(\mathbf{x})^{-1}\|^{-1}$.
Demuestre que existe $\rho(\mathbf{x}) > 0$ *tal que* $\overline{B(\mathbf{x}, \rho(\mathbf{x}))} \subset \Omega$ *y*

$$\|\mathbf{y} - \mathbf{x}\| < \rho(\mathbf{x}) \Rightarrow \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \geq \delta(\mathbf{x}) \|\mathbf{y} - \mathbf{x}\|$$

Deduzca de ello que si $\mathbf{f} : \Omega \to \mathbb{R}^n$ *es diferenciable en un abierto* $\Omega \subset \mathbb{R}^n$ *y det* $\mathbf{f}'(\mathbf{x}) \neq 0$ *para todo* $\mathbf{x} \in \Omega$ *entonces* $\mathbf{f}$ *es abierta. Demuestre también que* $f^{-1}(\mathbf{b})$ *es un conjunto discreto para cada* $\mathbf{b} \in \mathbb{R}^n$.

SOLUCIÓN

Por la diferenciabilidad de f en $\mathbf{x}$ existe $\rho(\mathbf{x}) > 0$ tal que $\overline{B(\mathbf{x}, \rho(\mathbf{x}))} \subset \Omega$ y

$$\|\mathbf{y} - \mathbf{x}\| < \rho(\mathbf{x}) \Rightarrow \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) - d\mathbf{f}(\mathbf{x})(\mathbf{y} - \mathbf{x})\| < \delta(\mathbf{x}) \|\mathbf{y} - \mathbf{x}\|$$

Como $T = d\mathbf{f}(\mathbf{x})$ es lineal y $\det T \neq 0$, dado $\mathbf{v} = T(\mathbf{u})$ se verifica

$$\|\mathbf{u}\| \leq \left\|T^{-1}(\mathbf{v})\right\| \leq \left\|T^{-1}\right\| \|\mathbf{v}\|, \quad \text{luego} \quad \|T(\mathbf{u})\| \geq \left\|T^{-1}\right\|^{-1} \|\mathbf{u}\|$$

Así se obtiene que

$$\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \geq \|d\mathbf{f}(\mathbf{x})(\mathbf{y} - \mathbf{x})\| - \|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) - d\mathbf{f}(\mathbf{x})(\mathbf{y} - \mathbf{x})\| \geq$$

$$\geq \left\|d\mathbf{f}(\mathbf{x})^{-1}\right\|^{-1} \|\mathbf{y} - \mathbf{x}\| - \delta(\mathbf{x}) \|\mathbf{y} - \mathbf{x}\| = \delta(\mathbf{x}) \|\mathbf{y} - \mathbf{x}\|$$

Sea $V \subset \Omega$ abierto y $\mathbf{a} \in V$. Según a) cualquier bola $\overline{B(\mathbf{a}, r)} \subset V$, de radio $0 < r < \rho(\mathbf{a})$ cumple la condición $\mathbf{f}(\mathbf{a}) \notin \mathbf{f}(\partial[B(\mathbf{a}, r)])$ y con el lema 8.6 se obtiene que $\mathbf{f}(V) \supset \mathbf{f}(B(\mathbf{a}, r))$ es entorno de $\mathbf{f}(\mathbf{a})$. Finalmente, si $\mathbf{a} \in \mathbf{f}^{-1}(\mathbf{b})$, en virtud de la implicación establecida $0 < \|\mathbf{x} - \mathbf{a}\| < \rho(\mathbf{a}) \Rightarrow \mathbf{f}(\mathbf{x}) \neq \mathbf{b}$, es decir, $B(\mathbf{a}, \rho(\mathbf{a})) \cap \mathbf{f}^{-1}(\mathbf{b}) = \{\mathbf{a}\}$ luego $f^{-1}(\mathbf{b})$ es un conjunto discreto. ∎

Ejercicio 8.20 *Sea* $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ *una aplicación diferenciable tal que existe* $C > 0$ *verificando:* $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})\| \geq C \|\mathbf{x} - \mathbf{y}\|$ *para cada* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.
Demuestre que $\det \mathbf{f}'(\mathbf{x}) \neq 0$ *para todo* $\mathbf{x} \in \mathbb{R}^n$ *y deduzca de ello que* $\mathbf{f}(\mathbb{R}^n) = \mathbb{R}^n$.

SOLUCIÓN

Demostraremos la primera afirmación por reducción al absurdo. Si para algún $\mathbf{a} \in \mathbb{R}^n$ ocurriese que $\det \mathbf{f}'(\mathbf{a}) = 0$, entonces la diferencial $d\mathbf{f}(\mathbf{a})$ no sería inyectiva y existiría $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{u} \neq 0$ con $d\mathbf{f}(\mathbf{a})\mathbf{u} = 0$. Según la definición de diferencial

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) - \mathbf{f}(\mathbf{a}) = d\mathbf{f}(\mathbf{a})\mathbf{h} + \|\mathbf{h}\| \rho(\mathbf{h})$$

donde $\lim_{\mathbf{h} \to 0} \rho(\mathbf{h}) = 0$. Como $d\mathbf{f}(\mathbf{a})$ se anula sobre los vectores $\mathbf{h} \in \{t\mathbf{u} : t \in \mathbb{R}\}$, para estos vectores se cumpliría la desigualdad

$$C \|\mathbf{h}\| \leq \|\mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a})\| = \|\mathbf{h}\| \|\rho(\mathbf{h})\|$$

luego $C \leq \|\rho(t\mathbf{u})\|$, lo que es absurdo porque $\lim_{t \to 0} \rho(t\mathbf{u}) = 0$.

La condición del enunciado implica que $\mathbf{f}$ es inyectiva, y con la proposición 8.7 se obtiene que $\mathbf{f}$ es abierta, luego $\mathbf{f}(\mathbb{R}^n)$ es un subconjunto abierto no vacío de $\mathbb{R}^n$. Entonces, en virtud de la conexión de $\mathbb{R}^n$, basta demostrar que el conjunto $F = \mathbf{f}(\mathbb{R}^n)$ es cerrado para concluir que $\mathbf{f}(\mathbb{R}^n) = \mathbb{R}^n$.

Dado $\mathbf{y} \in \overline{F}$ existe una sucesión $\mathbf{y}_k \in \mathbf{f}(\mathbb{R}^n)$ convergente hacia $\mathbf{y}$. Cada $\mathbf{y}_k$ es de la forma $\mathbf{y}_k = \mathbf{f}(\mathbf{x}_k)$ para algún $\mathbf{x}_k \in \mathbb{R}^n$. La desigualdad

$$C \|\mathbf{x}_p - \mathbf{x}_q\| \leq \|\mathbf{f}(\mathbf{x}_p) - \mathbf{f}(\mathbf{x}_q)\| = \|\mathbf{y}_p - \mathbf{y}_q\|$$

válida para cada $p, q \in \mathbb{N}$ implica que $(\mathbf{x}_k)$ es una sucesión de Cauchy en $\mathbb{R}^n$ y por lo tanto convergente hacia un punto $\mathbf{x} = \lim_k \mathbf{x}_k$. En virtud de la continuidad

$$\mathbf{f}(\mathbf{x}) = \lim_k \mathbf{f}(\mathbf{x}_k) = \lim_k \mathbf{y}_k = \mathbf{y}$$

luego $\mathbf{y} \in \mathbf{f}(\mathbb{R}^n)$, y queda demostrado que $\mathbf{f}(\mathbb{R}^n)$ es cerrado.

Un procedimiento alternativo para demostrar que $\mathbf{f}(\mathbb{R}^n) = \mathbb{R}^n$, sin utilizar la proposición 8.7, es el siguiente: Se comienza demostrando que $F = \mathbf{f}(\mathbb{R}^n)$ es cerrado y luego se usa un resultado bien conocido de la topología de $\mathbb{R}^n$ que afirma que la distancia de un punto $\mathbf{p} \in \mathbb{R}^n$ a un cerrado $F = \mathbf{f}(\mathbb{R}^n)$,

$$d(\mathbf{p}, F) = \inf\{d(\mathbf{p}, \mathbf{y}) : \mathbf{y} \in F\}$$

se alcanza en algún $\mathbf{b} = \mathbf{f}(\mathbf{a}) \in F$ (véase el ejercicio 3.8.6 d)). Aplicando esta propiedad con la distancia euclídea $d_2(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2$ podemos afirmar que $\|\mathbf{y} - \mathbf{p}\|_2^2$ alcanza en F un mínimo absoluto en algún punto $\mathbf{b} = \mathbf{f}(\mathbf{a})$. Esto significa que

$$g(\mathbf{x}) = \|\mathbf{f}(\mathbf{x}) - \mathbf{p}\|_2^2 = \sum_{j=1}^{n} (f_j(\mathbf{x}) - p_j)^2$$

alcanza un mínimo absoluto en $\mathbf{a} \in \mathbb{R}^n$, luego para $k \in \{1, 2, \cdots n\}$ se cumple

$$0 = D_k g(\mathbf{a}) = 2 \sum_{j=1}^{n} (f_j(\mathbf{a}) - p_j) D_k f_j(\mathbf{a})$$

Es decir $f_j(\mathbf{a}) - p_j$, $1 \leq j \leq n$, es solución del sistema lineal homogéneo cuya matriz $D_k f_j(\mathbf{a})$ tiene determinante no nulo. La única solución de este sistema lineal es la trivial, luego $\mathbf{p} = \mathbf{f}(\mathbf{a}) \in \mathbf{f}(\mathbb{R}^n)$. Como $\mathbf{p} \in \mathbb{R}^n$ era arbitrario, queda demostrado que $\mathbf{f}(\mathbb{R}^n) = \mathbb{R}^n$. ■

Ejercicio 8.21 *Mediante un cambio de variable a coordenadas polares encuentre las funciones* $f : \Omega \to \mathbb{R}$, *diferenciables en* $\Omega = \mathbb{R}^2 \setminus \{(0, 0)\}$ *que satisfacen la ecuación* $x f_x(x, y) + y f_y(x, y) + 2(x^2 + y^2) = 0$

solución

Si $F(r,\theta) = f(r\cos\theta, r\operatorname{sen}\theta)$, la ecuación se transforma en $rF_r - 2r^2 = 0$. En los puntos de Ω es $r = \sqrt{x^2+y^2} > 0$, y la última ecuación es equivalente a $F_r = 2r$, luego $F(r,\theta) = r^2 + g(\theta)$ donde g es cualquier función periódica de periodo 2π.
Se sigue que las soluciones de la ecuación propuesta son las funciones de la forma

$$f(x,y) = x^2 + y^2 + G(x,y)$$

donde $G : \Omega \to \mathbb{R}$ es una función diferenciable que se mantiene constante sobre cada semirrecta que surge de 0 (pues $G(r\cos\theta, r\operatorname{sen}\theta) = g(\theta)$ sólo depende de θ). Obsérvese que si G tiene esta propiedad, para todo $t > 0$ y todo $(x,y) \in \Omega$ se cumple $G(tx,ty) = G(x,y)$. Derivando respecto a t se obtiene que para todo $t > 0$ se verifica $xD_1G(tx,ty) + yD_2G(tx,ty) = 0$, y en particular, $xD_1G(x,y) + yD_2G(x,y) = 0$. Con esta igualdad, que se satisface en todo $(x,y) \in \Omega$, es inmediato comprobar que $f(x,y) = x^2 + y^2 + G(x,y)$ satisface la ecuación propuesta. ■

Ejercicio 8.22 *Sea* $A = \{(u,v) : v > 0\}$, $B = \{(x,y) : y > 0\}$ *y* $\mathbf{g} : A \to B$ *la aplicación definida por* $\mathbf{g}(u,v) = (x,y)$ *donde* $x = (u^2+v^2)/2$, $y = u/v$. *Compruebe que* $\mathbf{g}$ *establece un* C^∞*-difeomorfismo entre* A *y* B *y obtenga las ecuaciones de la transformación inversa* $\mathbf{g}^{-1} : B \to A$. *Utilice el cambio de variable* $(x,y) = \mathbf{g}(u,v)$ *para encontrar las funciones diferenciables* $f : B \to \mathbb{R}$ *que satisfacen la ecuación* $2xf_x - y(1+y^2)f_y = 0$.

solución

$\mathbf{g}$ es inyectiva sobre A: Si $\mathbf{g}(u,v) = \mathbf{g}(s,t)$ con $v > 0$ y $t > 0$ se cumple, $u/v = s/t$, $u^2+v^2 = s^2+t^2$. Si $c = u/v$ resulta $(1+c^2)v^2 = (1+c^2)t^2$, luego $v = t$, y $u = s$.
$\mathbf{g}(A) = B$: Basta ver que para cada $(x,y) \in B$ el sistema de dos ecuaciones $u^2+v^2 = 2x$, $u/v = y$ tiene una única solución $(u,v) \in A$. Como $2x = v^2(1+y^2)$ resulta

$$v = \sqrt{\frac{2x}{1+y^2}}; \quad u = y\sqrt{\frac{2x}{1+y^2}}.$$

Como $\mathbf{g}$ es de clase C^∞ y en todo $(u,v) \in A$ es $\det \mathbf{g}'(u,v) = -(u^2+1)/v^2 \neq 0$, en virtud del corolario 8.14 podemos asegurar que $\mathbf{g}$ es un C^∞-difeomorfismo.

Con el cambio de variable propuesto f se transforma en $F(u,v) = f(\mathbf{g}(u,v))$ y usando la regla de la cadena para el cálculo de las derivadas parciales

$$F_u = f_x x_u + f_y y_u = uf_x + \frac{1}{v}f_y; \qquad F_v = f_x x_v + f_y y_v = vf_x - \frac{u}{v^2}f_y.$$

Resolviendo el sistema se obtiene

$$f_x = \frac{uF_u + vF_v}{u^2+v^2}, \quad f_y = \frac{vF_u - uF_v}{u^2+v^2}v^2$$

luego

$$2xf_x = uF_u + vF_v; \quad y(1+y^2)f_y = \frac{u}{v}\left(1 + \frac{u^2}{v^2}\right)f_y = uF_u - \frac{u^2}{v}F_v;$$

Restando miembro estas igualdades

$$2xf_x - y(1+y^2)f_y = \frac{u^2+v^2}{v}F_v$$

Obtenemos así que la ecuación del enunciado equivale a $F_v = 0$ cuyas soluciones son las funciones de la forma $F(u,v) = \varphi(u)$ donde φ es una función derivable de una variable real. Deshaciendo el cambio de variable se concluye que las soluciones de la ecuación propuesta son las funciones de la forma

$$f(x,y) = \varphi\left(y\sqrt{\frac{2x}{y^2+1}}\right)$$

donde φ es una función derivable. Se deja al cuidado del lector la comprobación de que las funciones de este tipo son soluciones de la ecuación propuesta. ■

Ejercicio 8.23 *Sea $g : \mathbb{R}^3 \to \mathbb{R}$ una función de clase C^2 tal que $g(1,1,0) = 0$ y $\nabla g(1,1,0) = (2,0,0)$, y sea $x = \varphi(y,z)$ la función implícita de clase C^2 definida por la ecuación $g(x,y,z) = 0$ en un entorno del punto $(1,1,0)$. Compruebe que $(1,0)$ es un punto estacionario de la función $F(y,z) = \varphi(y,z)^2 + y^2z^2$ y obtenga valores de $A = D_{11}\varphi(1,0)$, $B = D_{12}\varphi(1,0)$ y $C = D_{22}\varphi(1,0)$ para los que se pueda asegurar que F presenta en $(1,0)$ un mínimo relativo.*

SOLUCIÓN

La función implícita $x = \varphi(y,z)$ está definida en un entorno de $(1,0)$ y verifica $\varphi(1,0) = 1$. Comprobemos en primer lugar que $(1,0)$ es un punto estacionario de F: Derivando respecto a las variables y, z, en la identidad $g(\varphi(y,z),y,z) \equiv 0$ se obtiene

$$D_1g(\varphi(y,z),y,z)\varphi_y(y,z) + D_2g(\varphi(y,z),y,z) \equiv 0;$$

$$D_1g(\varphi(y,z),y,z)\varphi_z(y,z) + D_3g(\varphi(y,z),y,z) \equiv 0$$

Sustituyendo $y = 1$, $z = 0$, $x = \varphi(1,0) = 1$, y teniendo en cuenta que $D_1g(1,1,0) = 2$, $D_2g(1,1,0) = 0$, $D_3g(1,1,0) = 0$, resulta $\varphi_y(1,0) = \varphi_z(1,0) = 0$. Ahora podemos calcular las derivadas parciales primeras y segundas de F:

$$F_y(y,z) = 2\varphi(y,z)\varphi_y(y,z) + 2yz^2; \quad F_z(y,z) = 2\varphi(y,z)\varphi_z(y,z) + 2y^2z; \qquad [*]$$

Sustituyendo $y = 1$, $z = 0$, $\varphi(1,0) = 1$, $\varphi_y(1,0) = \varphi_z(1,0) = 0$, se obtiene que $F_y(1,0) = 0$, $F_z(1,0) = 0$, luego $(1,0)$ es un punto estacionario de F.

Volviendo a derivar en [*] respecto a las variables y, z se llega a

$$F_{yy}(y,z) = 2\varphi_y(y,z)^2 + 2\varphi(y,z)\varphi_{yy}(y,z) + 2z^2 = 0;$$
$$F_{zy}(y,z) = 2\varphi_y(y,z)\varphi_z(y,z) + 2\varphi(y,z)\varphi_{zy}(y,z) + 4yz = 0$$
$$F_{zz}(y,z) = 2\varphi_z(y,z)^2 + 2\varphi(y,z)\varphi_{zz}(y,z) + 2y^2 = 0;$$

Sustituyendo $y = 1$, $z = 0$, $\varphi(1,0) = 1,$, $\varphi_y(1,0) = \varphi_z(1,0) = 0$, se llega a

$$F_{yy}(1,0) = 2A, \quad F_{yz}(1,0) = 2B, \quad F_{zz}(1,0) = 2C + 2$$

luego F presentará un mínimo relativo en el punto $(1,0)$ cuando se verifique

$$A > 0; \quad \text{y} \quad \begin{vmatrix} 2A & 2B \\ 2B & 2C+2 \end{vmatrix} = 4[A(1+C) - B^2] > 0$$

∎

Ejercicio 8.24 [Unicidad de funciones implícitas] *Sea* $\mathbf{g} : \Omega \to \mathbb{R}^m$ *continua en un abierto* $\Omega \subset \mathbb{R}^k \times \mathbb{R}^m$*, tal que* $S = \{(\mathbf{x}, \mathbf{y}) \in \Omega : \mathbf{g}(\mathbf{x}, \mathbf{y}) = 0\}$ *no es vacío y cada* $(\mathbf{x}_0, \mathbf{y}_0) \in S$ *posee un entorno abierto* $U \times V$ *donde la ecuación* $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$ *define una función implícita* $\mathbf{f} : U \to V$.

Se supone que en un abierto conexo $G \subset \mathbb{R}^k$ *hay definidas dos funciones implícitas continuas* $\boldsymbol{\varphi}, \boldsymbol{\Psi} : G \to \mathbb{R}^m$*, (es decir, dos funciones continuas con gráfica contenida en* S*). Si* $\mathbf{a} \in G$ *y* $\boldsymbol{\Psi}(\mathbf{a}) = \boldsymbol{\varphi}(\mathbf{a})$ *demuestre que* $\boldsymbol{\Psi} = \boldsymbol{\varphi}$*. Muestre con un ejemplo que el resultado es falso cuando* G *no es conexo.*

SOLUCIÓN

El conjunto $G_0 = \{\mathbf{x} \in G : \boldsymbol{\varphi}(\mathbf{x}) = \boldsymbol{\Psi}(\mathbf{x})\}$ no es vacío, pues $\mathbf{a} \in G_0$. Como G es conexo basta demostrar que G_0 es un subconjunto abierto y cerrado de G, con su topología relativa, para concluir que $G = G_0$. En virtud de la continuidad de $\boldsymbol{\varphi}$ y $\boldsymbol{\Psi}$, el conjunto G_0 es cerrado relativo a G. Veamos que también es abierto:

Por hipótesis, para cada $\mathbf{x}_0 \in G_0$ el punto $(\mathbf{x}_0, \boldsymbol{\varphi}(\mathbf{x}_0)) \in S$ posee un entorno $U \times V$ donde la ecuación $\mathbf{g}(\mathbf{x}, \mathbf{y}) = 0$ define una función implícita $\mathbf{f} : U \to V$.

Como V es un entorno de $\boldsymbol{\varphi}(\mathbf{x}_0) = \boldsymbol{\Psi}(\mathbf{x}_0)$, y las funciones $\boldsymbol{\varphi}$, $\boldsymbol{\Psi}$ son continuas en $\mathbf{x}_0$, podemos suponer que el entorno U de $\mathbf{x}_0$ se ha elegido de modo que $\boldsymbol{\varphi}(U) \subset V$ y $\boldsymbol{\Psi}(U) \subset V$. Necesariamente $\boldsymbol{\varphi}|_U$ y $\boldsymbol{\Psi}|_U$ coinciden en U con la función implícita $\mathbf{f}$, luego U es un entorno abierto de $\mathbf{x}_0$ contenido en G_0.

El siguiente ejemplo muestra que el resultado es falso cuando G_0 no es conexo: $g(x,y) = x^2 + y^2 - 1$, $G = (-1,0) \cup (0,1)$, Es claro que las funciones $\varphi, \psi : G \to \mathbb{R}$ definidas por

- $\varphi(x) = \sqrt{1 - x^2}$ si $x \in G$.
- $\psi(x) = \sqrt{1 - x^2}$ si $-1 < x < 0$, $\psi(x) = -\sqrt{1 - x^2}$ si $0 < x < 1$,

son distintas y ambas satisfacen las condiciones del enunciado. ∎

8.6. Ejercicios propuestos

♢ **8.6.1** *Sea* $\mathbf{f}:\Omega\to\mathbb{R}^n$ *una aplicación diferenciable definida en un abierto convexo* $\Omega\subset\mathbb{R}^n$. *Demuestre que la condición*

$$\sum_{i,j=1}^{n} D_i f_j(\mathbf{x})h_i h_j > 0 \quad \textit{para todo} \quad \mathbf{x}\in\Omega \quad \textit{y todo} \;\; \mathbf{h}\in\mathbb{R}^n\setminus\{0\}$$

implica que $\mathbf{f}$ *es inyectiva.*
(Indic: Si $f(\mathbf{x})=f(\mathbf{y})$ *con* $\mathbf{h}=\mathbf{y}-\mathbf{x}\neq 0$, *aplique el teorema del valor medio a* $\varphi(t)=\langle f(\mathbf{x}+t\mathbf{h})-f(\mathbf{x})\mid\mathbf{h}\rangle$, *en* $[0,1]$*).*

♢ **8.6.2** *Demuestre que la función* $\mathbf{f}:\mathbb{R}^3\to\mathbb{R}^3$, *definida por*

$$\mathbf{f}(x,y,z)=(z\cos xy,\;\; z\operatorname{sen} xy,\;\; x+z)$$

admite inversa local en un entorno abierto A *de* $\mathbf{a}=(1,0,1)$. *Si* $\mathbf{g}=(\mathbf{f}|_A)^{-1}$ *y* $\mathbf{b}=\mathbf{f}(\mathbf{a})$, *obtenga* $d\mathbf{g}(\mathbf{b})$.

♢ **8.6.3** *Demuestre que la aplicación* $\mathbf{f}:\mathbb{R}^3\to\mathbb{R}^3$, $\mathbf{f}(x,y,z)=(x^3,\;y^2-z^2,\;yz)$ *es abierta pero no es inyectiva.*

♢ **8.6.4** *Estudie los puntos donde es localmente inyectiva las función* $\mathbf{g}:\mathbb{R}^2\to\mathbb{R}^3$ *definid por* $\mathbf{g}(x,y)=(x+y,(x+y)^2,(x+y)^3)$.

♢ **8.6.5** *Para las transformaciones consideradas en los ejercicios 1.5.6,1.5.7, 1.5.8, 1.5.9 estudie su comportamiento local y global: Sobre qué abiertos son inyectivas, abiertas, difeomorfismos..)*

♢ **8.6.6** *Sea* $\mathbf{f}:\mathbb{R}^2\to\mathbb{R}^2$ *definida por* $\mathbf{f}(x,y)=(x+h(y),y+h(x))$ *donde* $h:\mathbb{R}\to\mathbb{R}$ *es derivable con derivada continua y* $|h'(t)|\leq c<1$ *para todo* $t\in\mathbb{R}$. *Demuestre que* $\mathbf{f}(\mathbb{R}^2)=\mathbb{R}^2$ *y que* $\mathbf{f}$ *establece un* C^1*-difeomorfismo de* $\mathbb{R}^2$ *sobre* $\mathbb{R}^2$.

♢ **8.6.7** *Demuestre que* $\mathbf{g}(x,y)=(x-a\cos y, y-a\cos x)$, *con* $0<a<1$, *es inyectiva y abierta.*

♢ **8.6.8** *Sea* $\mathbf{f}:A\to\mathbb{R}^k$ *de clase* C^2 *en un abierto* $A\subset\mathbb{R}^n$, *con* $n\geq k$, *tal que para cada* $\mathbf{x}\in A$ *los vectores* $\nabla f_1(\mathbf{x}),\nabla f_2(\mathbf{x}),\cdots\nabla f_k(\mathbf{x})$ *son linealmente independientes y sea* $\varphi:B\to\mathbb{R}$ *definida y continua en* $B=\mathbf{f}(A)$.

Demuestre que B *es abierto y que* φ *presenta un extremo relativo en* $\mathbf{b}=\mathbf{f}(\mathbf{a})$ *si y sólo si* $F=\varphi\circ\mathbf{f}$ *presenta un extremo relativo del mismo tipo en* $\mathbf{a}\in A$.

♢ **8.6.9** *Sea* $\mathbf{F}(x)=\nabla f(x)$ *donde* $f:\Omega\to\mathbb{R}$ *es de clase* C^k *(*$k\geq 2$*) en un abierto* $\Omega\subset\mathbb{R}^n$.

a) Dado $\mathbf{a}\in\Omega$, *obténga una condición (C) que garantice que* $\mathbf{F}$ *posee una inversa local* $\mathbf{G}$, *de clase* C^1, *definida en un entorno abierto de* $\mathbf{b}=F(\mathbf{a})$.

b) Obtenga $dh(\mathbf{x})$*, donde* $h(\mathbf{x}) = \langle \mathbf{x} \mid \mathbf{F}(\mathbf{x}) \rangle - f(\mathbf{x})$.

c) Si se cumple (C) demuestre que $\mathbf{G}$ *se deduce de* $\mathbf{g} = h \circ \mathbf{G} : B \to \mathbb{R}^n$ *de la misma forma que* $\mathbf{F}$ *se deduce de* f.

♢ **8.6.10** *Sea* $f : \mathbb{R}^2 \to \mathbb{R}$ *de clase* C^2 *de la que se conoce* $f(0,0) = 0$, $D_1 f(0,0) = 0$, $D_2 f(0,0) = 1$, $D_{11} f(0,0) = 1$, $D_{12} f(0,0) = 1$, $D_{22} f(0,0) = 2$.

Demuestre que la ecuación $\int_0^{f(x,y)} e^{t^2} dt = 0$ *define, en un entorno de* $(0,0)$*, a la variable* y *como función implícita de la variable* x.

♢ **8.6.11** *Compruebe que la siguiente función es de clase* C^1.

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

Demuestre que existe un entorno abierto U *de* $(1,0)$ *y una función* $g : U \to \mathbb{R}$*, de clase* $C^1(U)$*, tal que* $g(1,0) = 2$ *y*

$$f(xy, g(x,y) - 2x) = 0; \quad xD_1 g(x,y) - yD_2 g(x,y) = 2x \quad \text{para todo} \quad (x,y) \in U$$

♢ **8.6.12** *Sea* $g : \mathbb{R}^2 \to \mathbb{R}$ *de clase* C^2 *tal que* $g(b/a, c/a) = 0$ *y* $D_2 g(b/a, c/a) \neq 0$, $(a \neq 0)$*. Demuestre que existe un entorno abierto* $A \subset \mathbb{R}^2$ *de* (a,b)*, donde hay definida una única función* $f : A \to \mathbb{R}$*, de clase* $C^2(A)$*, verificando:*

$$f(a,b) = c, \quad y \quad g(y/x, f(x,y)/x)) = 0 \text{ para todo } (x,y) \in A.$$

Demuestre que f *satisface la ecuación:* $xD_1 f(x,y) + yD_2 f(x,y) = f(x,y)$.
Con un cambio de variable a coordenadas polares obtenga que $f(x,y)/\sqrt{x^2+y^2}$ *sólo depende de* y/x.

♢ **8.6.13** *Sea* $\psi : \mathbb{R} \longrightarrow \mathbb{R}$ *una función de clase* $\mathcal{C}^1$ *tal que* $\psi(0) = 0$ *y* $\psi'(0) = 1$*. Dados tres números reales* a*,* b *y* $c \neq 0$*, demuestre que existe un entorno* $U \subset \mathbb{R}^2$ *de* $(0,0)$ *y una función* $g : U \longrightarrow \mathbb{R}$ *de clase* $\mathcal{C}^1$ *que para todo* $(x,y) \in U$ *se cumplen las dos condiciones siguientes:*

a) $x^2 + y^2 + \big(g(x,y)\big)^2 = \psi\big(ax + by + cg(x,y)\big)$

b) $\big(cy - bg(x,y)\big) D_1 g(x,y) + \big(ag(x,y) - cx\big) D_2 g(x,y) = bx - ay$

Calcule $D_1 g(0,0)$ *y* $D_2 g(0,0)$.

♢ **8.6.14** *Compruebe que la ecuación* $z^3 + 2z + e^{z-x-y^2} + x + y^2 - \cos(x - y + z) = 0$ *define, en un entorno de* $(0,0,0)$*, a la variable* z *como función implícita de las variables* x, y*. Demuestre que la función implícita* $z = f(x,y)$*, definida en un entorno de* $(0,0)$*, presenta un máximo relativo en* $(0,0)$.

♢ **8.6.15** *Compruebe que la ecuación* $xyz + \operatorname{sen}(z - 6) - 2(x + y + x^2 y^2) = 0$ *define, en un entorno de* $(1,1,6)$*, una función implícita* $z = f(x,y)$ *que presenta un mínimo relativo en* $(1,1)$.

♢ **8.6.16** *Compruebe que la ecuación* $\operatorname{sen} z + (1 + x^2)^y + z + y^2 - 2y = 0$ *define, en un entorno de* $(0,1,0)$*, una función implícita* $z = f(x,y)$ *que verifica* $f(0,1) = 1$*. Estudie si* f *presenta un extremo local en* $(0,1)$.

♢ **8.6.17** *Compruebe que la ecuación* $x^3+y^3+z^3+x+y-z=4$ *define, en un entorno de* $\mathbf{p}=(1,1,1)$*, una función implícita* $z=f(x,y)$ *con la propiedad de que en un entorno de* $\mathbf{a}=(1,1)$ *la gráfica de* f *queda por debajo de su plano tangente en* $\mathbf{p}$.

♢ **8.6.18** *Justifique que la ecuación* $x^3+3z^2+8xy^2-3y^3z=2$ *define en un entorno de* $\mathbf{p}=(-1,0,1)$ *una función implícita* $z=f(x,y)$*. Demuestre que en un entorno de* $\mathbf{a}=(-1,0)$ *la gráfica de* f *queda por encima de su plano tangente en* $\mathbf{p}$.

♢ **8.6.19** *Sea* $g(x,y,z)=x^3+y^3+z^3+x^2+y^2+z^2+x+y-z$*. Compruebe que en un entorno de* $\mathbf{p}=(1,1,1)$*, la ecuación* $g(x,y,z)=7$ *define, una función implícita* $z=f(x,y)$*. Obtenga el polinomio de Taylor de grado* 2*, de* f *en el punto* $(1,1)$.

Si $M=\{(x,y,z)\in\mathbb{R}^3 : g(x,y,z)=7\}$ *demuestre que* $\mathbf{p}$ *posee un entorno* $V_{\mathbf{p}}$ *tal que* $M\cap V_{\mathbf{p}}$ *queda a un lado del plano tangente a* M *en* $\mathbf{p}$.

♢ **8.6.20** *Compruebe que las ecuaciones* $x\cos y+y\cos z+z\cos x=\pi$; $x^2+y^2+z^2-xy=\pi^2$ *definen, en un entorno de* $(x_0,y_0,z_0)=(0,0,\pi)$*, a las variables* y,z *como funciones implícitas de* x*:* $(y,z)=(f_1(x),f_2(x))$*. Calcule las derivadas* $f_i'(0)$, $f_i''(0)$, $i=1,2$.

♢ **8.6.21** *Compruebe que las ecuaciones*

$$yt+e^{xt}+e^{xy}=3;\quad y^x-y+t=1$$

definen en un entorno de $(x_0,y_0,t_0)=(0,1,1)$ *a las variables* x,y *como funciones implícitas de la variable* t*. Si* $x=f(t)$ $y=g(t)$ *son las funciones implícitas, calcule las derivadas* $f'(1)$,$g'(1)$, $f''(1)$, $g''(1)$.

♢ **8.6.22** *Compruebe que en un entorno de* $(x_0,y_0,u_0,v_0)=(\sqrt{2},0,1,-1)$*, el sistema de ecuaciones*

$$2uv+x^2-y^2=0,\quad u^2-v^2+2xy=0$$

define a las variables (u,v) *como funciones implícitas de las variables* (x,y).

Demuestre que la función implícita $(u,v)=\mathbf{f}(x,y)$ *es invertible en un entorno* U *de* $(\sqrt{2},0)$*. Si* $\mathbf{g}$ *es la inversa de* $\mathbf{f}|_U$*, obtenga* $d\mathbf{g}((1,-1))$.

♢ **8.6.23** *Compruebe que en un entorno de* $(1,1,1,1)$ *el sistema de ecuaciones*

$$x^2+y^2-u-v^2=0,\quad x^2-y^2+u^3-v^3=0$$

define funciones funciones implícitas $(u,v)=\mathbf{f}(x,y)$, $(x,y)=\mathbf{g}(u,v)$*. Explique la relación que hay entre las matrices* $\mathbf{f}'(1,1)$ *y* $\mathbf{g}'(1,1)$ *y calcule una de ellas.*

♢ **8.6.24** *Determine los valores de* $a\in\mathbb{R}$ *para los que el sistema*

$$xz^3+yu+ax=1,\quad 2xy^3+u^2z+ay=a$$

define, en un entorno de $(0,1)$ *una función implícita* $(x,y)=\varphi(z,u)$ *que cumple* $\varphi(0,1)=(0,1)$ *y estudie si* φ *es localmente invertible en* $(0,1)$.

♢ **8.6.25** *Compruebe que las ecuaciones*

$$xz^3 + y^2u^3 = 1; \quad 2xy^3 + u^2z = 0$$

definen, en un entorno de $(x_0, y_0, z_0, u_0) = (0, 1, 0, 1)$, *a las variables* x, y *como funciones implícitas de las variables* z, u. *Si* $x = g_1(z, u)$, $y = g_2(z, u)$ *son las funciones implícitas, definidas en un entorno de* $(0, 1)$, *demuestre que existe un entorno abierto de* $(0, 1)$, *donde* $g = (g_1, g_2)$ *es invertible con inversa indefinidamente derivable.*

♢ **8.6.26** *Obtenga el gradiente de una función diferenciable* $f : \mathbb{R}^3 \to \mathbb{R}$ *en términos de las coordenadas cilíndricas,* $x = r\cos\theta$, $y = r\operatorname{sen}\theta$, $z = t$, *y de las coordenadas esféricas,* $x = \rho\cos\varphi\cos\theta$, $y = \rho\cos\varphi\operatorname{sen}\theta$, $z = \rho\operatorname{sen}\varphi$.

♢ **8.6.27** *Efectuando el cambio de variable* $x = (u+v)/2$, $y = (u-v)/2$ *obtenga la forma general de las funciones de clase* C^2, $f : \mathbb{R}^2 \to \mathbb{R}$, *que satisfacen la ecuación en derivadas parciales*

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} - 2\frac{\partial^2 f}{\partial x \partial y} = 0$$

♢ **8.6.28** *Utilizando el cambio de variable* $x = v$, $y = uv$ *obtenga las funciones* f *de clase* C^2 *en* $\Omega = \{(x, y) : x > 0\}$ *que verifican*

$$x^2 D_{11}f(x, y) + 2xyD_{12}f(x, y) + y^2 D_{22}f(x, y) = 0, \quad \textit{para todo } (x, y) \in \Omega$$

♢ **8.6.29** *Utilizando el cambio de variable* $x = (u^2 + v^2)/2$, $y = u/v$ *obtenga las funciones* $f : \Omega \to \mathbb{R}$ *de clase* C^2 *en* $\Omega = \{(x, y) : x > 0\}$ *que verifican*

$$2xyD_1f(x, y) + (1 + y^2)D_2f(x, y) = 0 \ \textit{para todo } (x, y) \in \Omega$$

♢ **8.6.30** *Usando el cambio de variable* $x = u^2+v^2$, $y = u+v$ *obtenga las funciones* $f : \Omega \to \mathbb{R}$ *de clase* C^2 *en* $\Omega = \{(x, y) : 2x > y^2\}$ *que verifican*

$$2(y^2 - x^2)D_{11}f(x, y) + 2yD_{12}f(x, y) + D_{22}f(x, y) = y^2 - x^2 \ \textit{para todo } (x, y) \in \Omega$$

♢ **8.6.31** *Dado el abierto* $A = \{(x, y) : 0 < x < y\}$, *justifique la existencia de un abierto* $B \subset \mathbb{R}^2$ *tal que para cada* $f \in C^1(A)$ *existe una única* $F \in C^1(B)$ *que verifica :* $f(x, y) = F(x + y, xy)$ *para todo* $(x, y) \in A$.
Demuestre que son equivalentes

i) $D_1f(x, y) - D_2f(x, y) + 3(x - y)f(x, y) = 0$ *para todo* $(x, y) \in A$.

ii) $D_2F(u, v) = 3F(u, v)$ *para todo* $(u, v) \in B$.