Tema 11

Función inversa

Como segundo resultado fundamental del cálculo diferencial, estudiaremos ahora la posible existencia, así como la diferenciabilidad, de la función inversa de una función diferenciable. Empezamos probando una *regla de diferenciación de la función inversa*, para funciones entre espacios normados arbitrarios, que generaliza casi al pie de la letra la que conocemos para funciones reales de una variable real.

Pasamos entonces al problema mucho más interesante que empieza por discutir la propia existencia de la función inversa, esto es, la inyectividad de una función diferenciable. Claro está que, una vez asegurada la inyectividad, también nos preguntamos si la función inversa es diferenciable. En el caso conocido de las funciones reales de una variable real, este problema se resolvía mediante el llamado "teorema de la función inversa" para tales funciones, que tiene una versión "local" y otra "global". A la hora de su generalización a funciones de varias variables, estos dos resultados se comportan de manera muy diferente. La versión local se generaliza de forma casi literal, obteniendo el *teorema de la función inversa* en $\mathbb{R}^N$. Como consecuencia de este teorema fundamental, que como se ha dicho tiene carácter "local", deducimos fácilmente un resultado que puede entenderse como una versión "global" del teorema, pero no generaliza lo que conocemos para funciones de una variable, es un resultado mucho más débil.

11.1. Diferencial de la función inversa

Recordemos la regla de derivación de la inversa, para funciones de una variable:

- *Sea A un subconjunto no vacío de $\mathbb{R}$ y $f : A \to \mathbb{R}$ una función inyectiva. Sea $B = f(A)$ y consideremos la función inversa $f^{-1} : B \to \mathbb{R}$. Si f es derivable en un punto $a \in A \cap A'$, entonces $b = f(a) \in B'$ y las siguientes afirmaciones son equivalentes:*

  (i) *f^{-1} es derivable en el punto b*

  (ii) *f^{-1} es continua en b y $f'(a) \neq 0$*

  Caso de que se cumplan (i) y (ii), se tiene: $\left(f^{-1}\right)'(b) = \dfrac{1}{f'(a)}$.

Con vistas a generalizar este resultado, de forma que pueda aplicarse a una función entre dos espacios normados arbitrarios, conviene aclarar dos cuestiones.

Tendremos un abierto Ω de un espacio normado X y una función inyectiva $f : \Omega \to Y$, donde Y es otro espacio normado, que será diferenciable en un punto $a \in \Omega$. Surge aquí la única pequeña diferencia con el caso particular que acabamos de recordar: para poder hablar de la diferenciabilidad de f^{-1} en el punto $b = f(a)$, necesitamos que b sea, no sólo punto de acumulación del conjunto $B = f(\Omega)$, sino punto interior del mismo. Las hipótesis sobre f no permiten asegurar que esto ocurra, luego deberemos suponerlo. Al fin y al cabo, cuando $b \notin B^\circ$, no tiene sentido plantearse el problema que estamos estudiando.

La otra cuestión, más relevante, se refiere a la forma de generalizar la afirmación (ii), más concretamente, la condición $f'(a) \neq 0$ que en ella aparece. Es una ingenuidad pensar que la condición vaya a ser $Df(a) \neq 0$. Recordemos que, en el caso particular, la diferencial $Df(a)$ consiste en multiplicar por $f'(a)$. Por tanto, decir que $f'(a) \neq 0$, equivale a decir que $Df(a)$ es biyectiva, y entonces, la expresión de $\big(f^{-1}\big)'(b)$ toma la forma $Df^{-1}(b) = Df(a)^{-1}$. En general, si queremos tener la misma expresión para la diferencial de la función inversa, no sólo necesitamos que $Df(a)$ sea biyectiva, sino también que su inversa $Df(a)^{-1}$ sea continua. Como se trata de una aplicación lineal de Y en X, su continuidad será automática cuando Y tenga dimensión finita, pero no en general. Esto motiva la siguiente definición.

Si X e Y son espacios normados, llamamos **homeomorfismo lineal** de X sobre Y a toda aplicación lineal y continua $T : X \to Y$, que sea biyectiva y tal que T^{-1} también sea continua. Más brevemente, estamos exigiendo que $T \in L(X,Y)$ sea biyectiva con $T^{-1} \in L(Y,X)$. La nomenclatura es coherente con la noción de homeomorfismo que se usa en Topología: biyección continua entre dos espacios topológicos, cuya función inversa también es continua. En nuestro caso, un homeomorfismo lineal entre dos espacios normados X e Y no es más que lo que su nombre indica: un homeomorfismo de X sobre Y que además es lineal, lo que obviamente hace que la función inversa también sea lineal. Notemos que, cuando dos espacios normados son *linealmente homeomorfos*, es decir, existe un homeomorfismo lineal de uno en otro, son idénticos como espacios vectoriales (pues tenemos una biyección lineal entre ellos) y también como espacios topológicos (pues son homeomorfos). Podemos por tanto pensar que se trata de un mismo espacio vectorial, en el que tenemos dos normas que pueden ser distintas, pero son equivalentes, pues dan lugar a la misma topología.

Los comentarios hechos anteriormente hacen ver que el siguiente resultado generaliza, de manera casi literal, la regla de derivación de la función inversa que acabamos de recordar.

Teorema (Regla de diferenciación de la inversa). *Sean X e Y espacios normados, Ω un abierto de X, $f : \Omega \to Y$ una función inyectiva y $f^{-1} : B \to X$ su inversa, donde $B = f(\Omega)$. Supongamos que f es diferenciable en un punto $a \in \Omega$ y que $b = f(a) \in B^\circ$. Entonces, las siguientes afirmaciones son equivalentes:*

(i) *f^{-1} es diferenciable en el punto b.*
(ii) *f^{-1} es continua en b y $Df(a)$ es un homeomorfismo lineal de X sobre Y.*

En caso de que se cumplan (i) y (ii), se tiene: $Df^{-1}(b) = Df(a)^{-1}$.

Demostración. $(i) \Rightarrow (ii)$. Como f^{-1} es diferenciable en b, también es continua en b. Si ahora $\mathrm{Id}_X \in L(X,X)$ e $\mathrm{Id}_Y \in L(Y,Y)$ son las funciones identidad en X e Y respectivamente, vemos que $f^{-1} \circ f$ es la restricción a Ω de Id_X. Sabemos entonces que $f^{-1} \circ f$ es diferenciable en el punto a y su diferencial es la propia Id_X. Análogamente, la diferencial de $f \circ f^{-1}$ en el punto b es Id_Y. Aplicando entonces dos veces la regla de la cadena, obtenemos

$$Df^{-1}(b) \circ Df(a) = \mathrm{Id}_X \qquad \text{y} \qquad Df(a) \circ Df^{-1}(b) = \mathrm{Id}_Y$$

Esto prueba que $Df(a)$ es biyectiva con $Df(a)^{-1} = Df^{-1}(b) \in L(Y,X)$, luego $Df(a)$ es un homeomorfismo lineal de X sobre Y, como queríamos. De paso, hemos probado ya la última afirmación del enunciado: $Df^{-1}(b) = Df(a)^{-1}$.

$(ii) \Rightarrow (i)$. Para abreviar escribimos $T = Df(a) \in L(X,Y)$ y, por hipótesis, sabemos que T es biyectiva con $T^{-1} \in L(Y,X)$. Como ya hemos usado otras veces, la diferenciabilidad de f en el punto a equivale a la continuidad en a de la función $\Phi : \Omega \to \mathbb{R}$ definida por

$$\Phi(x) = \frac{\big\| f(x) - f(a) - T(x-a) \big\|}{\|x-a\|} \quad \forall x \in \Omega \setminus \{a\} \quad \text{y} \quad \Phi(a) = 0$$

Por tanto Φ es continua en el punto a y evidentemente verifica que

$$\big\| f(x) - f(a) - T(x-a) \big\| = \Phi(x) \big\| x-a \big\| \qquad \forall x \in \Omega \tag{1}$$

Definimos entonces $\Lambda : B \to \mathbb{R}$ por

$$\Lambda(y) = \big\| f^{-1}(y) - f^{-1}(b) - T^{-1}(y-b) \big\| \qquad \forall y \in B$$

de modo que bastará probar la igualdad

$$\lim_{y \to b} \frac{\Lambda(y)}{\|y-b\|} = 0 \tag{2}$$

Fijamos $y \in B$ y escribimos $x = f^{-1}(y) \in \Omega$, con lo cual tenemos

$$\begin{aligned} \Lambda(y) &= \big\| x - a - T^{-1}(y-b) \big\| = \big\| T^{-1}\big(T(x-a) - (y-b)\big) \big\| \\ &\leqslant \big\| T^{-1} \big\| \, \big\| T(x-a) - \big(f(x) - f(a)\big) \big\| \\ &= \big\| T^{-1} \big\| \, \Phi(x) \, \big\| x-a \big\| \end{aligned} \tag{3}$$

donde hemos usado la definición de la norma en $L(Y,X)$ y la igualdad (1). Por otra parte, tenemos también

$$\begin{aligned} \|x-a\| &= \big\| T^{-1}\big(T(x-a)\big) \big\| \leqslant \big\| T^{-1} \big\| \, \big\| T(x-a) \big\| \\ &\leqslant \big\| T^{-1} \big\| \Big(\big\| T(x-a) - \big(f(x) - f(a)\big) \big\| + \big\| f(x) - f(a) \big\| \Big) \\ &= \big\| T^{-1} \big\| \Big(\Phi(x) \big\| x-a \big\| + \big\| y-b \big\| \Big) \end{aligned}$$

o lo que es lo mismo,

$$\big(1 - \|T^{-1}\| \, \Phi(x)\big) \, \|x-a\| \leqslant \|T^{-1}\| \, \|y-b\| \tag{4}$$

Queremos deducir de (4) una estimación de $\|x-a\|$ que podamos usar en (3). Para ello necesitamos que $\|T^{-1}\|\,\Phi(x) < 1$, lo cual es cierto para y suficientemente próximo a b, como vamos a ver. En primer lugar, por ser Φ continua en el punto a con $\Phi(a)=0$, existe $\rho>0$ tal que, para $z\in\Omega$ con $\|z-a\|<\rho$ se tiene $\|T^{-1}\|\,\Phi(z) < 1$. Como ahora f^{-1} es continua en b, existe $\delta>0$ tal que, para $w\in B$ con $\|w-b\|<\delta$ se tiene $\|f^{-1}(w)-f^{-1}(b)\|<\rho$. Volviendo al punto $y\in B$ que habíamos fijado, pero suponiendo además que $\|y-b\|<\delta$, usamos lo anterior con $w=y$ para obtener $\|f^{-1}(y)-f^{-1}(b)\|<\rho$, es decir, $\|x-a\|<\rho$. Esto permite tomar $z=x$ para concluir que $\|T^{-1}\|\,\Phi(x) < 1$ como queríamos.

Por tanto, si $y\in B$ verifica que $\|y-b\|<\delta$, para $x=f^{-1}(y)$ tenemos, usando (4), que

$$\|x-a\| \leqslant \frac{\|T^{-1}\|\,\|y-b\|}{1-\|T^{-1}\|\,\Phi(x)}$$

y de (3) deducimos entonces que

$$\Lambda(y) \leqslant \frac{\|T^{-1}\|^2\,\Phi(x)\,\|y-b\|}{1-\|T^{-1}\|\,\Phi(x)}$$

En resumen, hemos encontrado $\delta>0$ tal que

$$y\in B,\ 0<\|y-b\|<\delta \implies 0\leqslant \frac{\Lambda(y)}{\|y-b\|} \leqslant \frac{\|T^{-1}\|^2\,\Phi\big(f^{-1}(y)\big)}{1-\|T^{-1}\|\,\Phi\big(f^{-1}(y)\big)} \tag{5}$$

Finalmente, como f^{-1} es continua en b y Φ es continua en $f^{-1}(b)=a$, tenemos que $\Phi\circ f^{-1}$ es continua en b, es decir,

$$\lim_{y\to b}\Phi\big(f^{-1}(y)\big) = \Phi\big(f^{-1}(b)\big) = \Phi(a) = 0$$

Esta igualdad, junto con (5), nos da directamente (2), que era la igualdad buscada. ■

La implicación fácil del teorema anterior tiene una consecuencia que merece la pena resaltar:

- *Sean U y V abiertos de sendos espacios normados X e Y. Si existe una aplicación biyectiva $f:U\to V$ que es diferenciable en un punto $a\in U$ y f^{-1} es diferenciable en el punto $f(a)$, entonces X e Y son linealmente homeomorfos. En particular, si X tiene dimensión finita $N\in\mathbb{N}$, también Y tendrá dimensión N.*

En efecto, por el teorema anterior, $Df(a)$ es un homeomorfismo lineal de X sobre Y. ■

A partir de ahora trabajamos en espacios de dimensión finita, y en vista de la observación anterior, de la misma dimensión, es decir, $X=Y=\mathbb{R}^N$ aunque las normas de X e Y pueden ser dos normas cualesquiera en $\mathbb{R}^N$. Al aplicar el teorema anterior, $Df(a)^{-1}$ será continua siempre que $Df(a)$ sea biyectiva, pues $Df(a)^{-1}$ es lineal y está definida en un espacio normado de dimensión finita. Por tanto, para asegurar que f^{-1} es diferenciable en b, bastará comprobar que es continua en b y que $Df(a)$ es biyectiva.

Para cada aplicación lineal $T \in L(\mathbb{R}^N, \mathbb{R}^N)$ recordemos la matriz cuadrada $A_T \in \mathcal{M}_{N\times N}$ que la representa en la base usual de $\mathbb{R}^N$, cuyo determinante denotamos por $\det A_T$. Sabemos que T es biyectiva si, y sólo si, $\det A_T \neq 0$. Si ahora Ω es un abierto de $\mathbb{R}^N$ y $f : \Omega \to \mathbb{R}^N$ una función diferenciable en un punto $a \in \Omega$, para $T = Df(a)$ sabemos que $A_T = Jf(a)$ es la matriz jacobiana de f en a y se dice que $\det Jf(a)$ es el **determinante jacobiano** de f en a. Según hemos dicho, $Df(a)$ es biyectiva si, y sólo si, $\det Jf(a) \neq 0$. Así pues, en nuestro caso particular, el teorema anterior toma la siguiente forma:

- *Sea Ω un abierto de $\mathbb{R}^N$, $f : \Omega \to \mathbb{R}^N$ una función inyectiva y $f^{-1} : B \to \mathbb{R}^N$ su inversa, donde $B = f(\Omega)$. Supongamos que f es diferenciable en $a \in \Omega$ y que $b = f(a) \in B^\circ$. Entonces, las siguientes afirmaciones son equivalentes:*

  (i) *f^{-1} es diferenciable en el punto b.*
  (ii) *f^{-1} es continua en b y $\det Jf(a) \neq 0$.*

  En caso de que se cumplan (i) y (ii), se tiene que $Df^{-1}(b) = Df(a)^{-1}$, o lo que es lo mismo, $Jf^{-1}(b) = Jf(a)^{-1}$.

Acerca del determinante jacobiano, resaltamos un hecho elemental que será útil enseguida. Es fácil ver que la aplicación $T \mapsto \det A_T$, definida en el espacio normado $L(\mathbb{R}^N, \mathbb{R}^N)$ y con valores en $\mathbb{R}$, es continua. En efecto, para cada $j,k \in \Delta_N$ sabemos que el coeficiente que aparece en la j-ésima fila y la k-ésima columna de la matriz A_T viene dado por

$$\alpha_{jk}(T) = \big(T\,e_k \,|\, e_j\big) \qquad \forall\, T \in L(\mathbb{R}^N, \mathbb{R}^N)$$

donde $\{e_1, e_2, \ldots e_N\}$ es la base usual de $\mathbb{R}^N$. Entonces la aplicación $\alpha_{jk} : L(\mathbb{R}^N, \mathbb{R}^N) \to \mathbb{R}$ es continua, puesto que es lineal y está definida en un espacio normado de dimensión finita. Por tanto, la función $T \mapsto \det A_T$ se puede expresar como combinación lineal de productos de funciones continuas, luego es una función continua. Apliquemos ahora esta observación al caso particular del determinante jacobiano de una función diferenciable:

- *Sea Ω un abierto de $\mathbb{R}^N$ y $f : \Omega \to \mathbb{R}^N$ una función diferenciable. Si Df es continua en un punto $a \in \Omega$, entonces la función $x \mapsto \det Jf(x)$, definida en Ω y con valores en $\mathbb{R}$, también es continua en a.*

En efecto, se trata de la composición de la función Df, de Ω en $L(\mathbb{R}^N, \mathbb{R}^N)$, que es continua en a, con la función $T \mapsto \det A_T$ de $L(\mathbb{R}^N, \mathbb{R}^N)$ en $\mathbb{R}$, que es continua, como hemos visto. ■

11.2. Teorema de la función inversa

En la regla de diferenciación de la función inversa que acabamos de estudiar, trabajamos con una función que de entrada se supone inyectiva. Sin embargo, igual que se hace con las funciones reales de una variable real, podemos intentar averiguar si una función diferenciable es inyectiva, usando su diferencial. El significado analítico de la diferencial invita a pensar que podemos tener éxito a nivel "local", como pasamos a explicar.

Hablando informalmente, una función f que sea diferenciable en un punto a, se aproxima bien, "cerca" de a, por una función afín y continua g, que se obtiene componiendo $Df(a)$ con una traslación. Cuando $Df(a)$ es inyectiva, está claro que g también lo es, y es lógico esperar que f sea inyectiva donde sabemos que debe comportarse de forma parecida a como lo hace g, es decir, en un entorno suficientemente pequeño de a. Con las hipótesis adecuadas, probaremos rigurosamente que esta idea intuitiva es correcta, obteniendo así el que puede considerarse como el resultado más importante del cálculo diferencial en varias variables: el teorema de la función inversa en $\mathbb{R}^N$. Para que se comprenda mejor el enunciado de dicho teorema, recordemos lo que ocurría en el caso $N = 1$, es decir, la versión local del teorema de la función inversa para funciones reales de variable real:

- *Sea $I \subset \mathbb{R}$ un intervalo no trivial y $f : I \to \mathbb{R}$ una función derivable en I. Supongamos que f' es continua en un punto $a \in I$, con $f'(a) \neq 0$. Entonces existe un $\delta > 0$ tal que f es inyectiva en el intervalo $I_\delta = I \cap]a-\delta, a+\delta[$. Además, si llamamos φ a la restricción de f a I_δ y consideramos el intervalo $J_\delta = f(I_\delta)$, se tiene que φ^{-1} es derivable en J_δ con $\big(\varphi^{-1}\big)'\big(f(x)\big) = 1/f'(x)$ para todo $x \in I_\delta$.*

Vamos a conseguir una generalización casi literal del teorema anterior, válida para funciones de N variables. El único cambio se debe a que sólo podemos usar la diferenciabilidad de una función en puntos interiores de su conjunto de definición. Por tanto, el papel de los intervalos que aparecen en el enunciado anterior, lo harán convenientes abiertos de $\mathbb{R}^N$. La demostración es muy diferente de la que se hace en $\mathbb{R}$, basada en la noción de monotonía. En $\mathbb{R}^N$ las cosas son más complicadas, como se podrá comprobar.

Teorema de la función inversa (local). *Sea Ω un abierto de $\mathbb{R}^N$ y $f : \Omega \to \mathbb{R}^N$ una función diferenciable, es decir, $f \in D(\Omega, \mathbb{R}^N)$. Supongamos que la función diferencial Df es continua en un punto $a \in \Omega$ y que $\det Jf(a) \neq 0$. Entonces existe un abierto U de $\mathbb{R}^N$, con $a \in U \subset \Omega$, para el que se verifican las siguientes afirmaciones:*

(i) *f es inyectiva en U*
(ii) *$V = f(U)$ es un abierto de $\mathbb{R}^N$*
(iii) *$\det Jf(x) \neq 0$ para todo $x \in U$*
(iv) *Si $\varphi = f\big|_U$, entonces $\varphi^{-1} \in D(V, \mathbb{R}^N)$, con $D\varphi^{-1}\big(f(x)\big) = Df(x)^{-1}$ para todo $x \in U$.*

Demostración. Usaremos cualquier norma en $\mathbb{R}^N$, y empezamos trabajando en un caso particular más cómodo, del que al final se deducirá fácilmente el caso general. Concretamente suponemos por ahora que $a = f(a) = 0$ y que $Df(0) = \mathrm{Id}$ es la identidad en $\mathbb{R}^N$.

En este caso, la idea intuitiva en la que se basa el teorema está muy clara: en un entorno del origen, f debe comportarse de forma similar a como lo hace la identidad, una función que cumple obviamente todas las afirmaciones del teorema. Así pues, procede trabajar con la diferencia entre ambas, es decir, la función $g : \Omega \to \mathbb{R}^N$ dada por

$$g(x) = x - f(x) \qquad \forall x \in \Omega$$

Es claro que $g(0) = 0$, así como que g es diferenciable con

$$Dg(x) = \mathrm{Id} - Df(x) \qquad \forall x \in \Omega$$

Como, por hipótesis, Df es continua en el origen, vemos claramente que Dg también lo es, con $Dg(0) = 0$, luego existe $\delta_1 > 0$ tal que $B(0, \delta_1) \subset \Omega$ y

$$\| Dg(x) \| \leqslant 1/2 \qquad \forall x \in B(0, \delta_1)$$

Por otra parte, recordemos que la continuidad de Df en el origen también implica la de la función $x \mapsto \det Jf(x)$, siendo $\det Jf(0) = 1$, luego existe $\delta_2 > 0$ tal que $B(0, \delta_2) \subset \Omega$ y

$$\det Jf(x) \neq 0 \qquad \forall x \in B(0, \delta_2)$$

Tomamos ahora $r = (1/3) \min\{\delta_1, \delta_2\}$ con lo cual tenemos:

$$x \in \mathbb{R}^N, \ \|x\| < 3r \implies x \in \Omega, \ \|Dg(x)\| \leqslant 1/2 \ \text{ y } \ \det Jf(x) \neq 0 \tag{6}$$

Como $B(0, 3r)$ es un subconjunto abierto y convexo de $\mathbb{R}^N$, un corolario de la desigualdad del valor medio nos asegura que g es lipschitziana en dicho abierto, más concretamente:

$$x, z \in B(x, 3r) \implies \| g(x) - g(z) \| \leqslant (1/2) \| x - z \| \tag{7}$$

En particular, tomando $z = 0$ tenemos

$$x \in B(x, 3r) \implies \| g(x) \| \leqslant (1/2) \| x \| \tag{8}$$

Llegamos al paso clave de la demostración, que consiste en usar el teorema del punto fijo de Banach, para probar lo siguiente:

(*) *Para cada $y_0 \in \overline{B}(0, r)$ existe un único $x_0 \in \overline{B}(0, 2r)$ tal que $f(x_0) = y_0$. Además, si de hecho $\| y_0 \| < r$, se tiene también $\| x_0 \| < 2r$.*

Fijado $y_0 \in \overline{B}(0, r)$, para $x \in \overline{B}(0, 2r)$ se tiene

$$f(x) = y_0 \iff g(x) = x - y_0 \iff g(x) + y_0 = x \tag{9}$$

luego buscamos un punto fijo de la función $x \mapsto g(x) + y_0$. Así pues, tomamos $E = \overline{B}(0, 2r)$, que es un espacio métrico completo, subconjunto cerrado de $\mathbb{R}^N$, y definimos $h(x) = g(x) + y_0$ para todo $x \in E$. Usando (8) tenemos

$$\| h(x) \| \leqslant \| g(x) \| + \| y_0 \| \leqslant (1/2) \| x \| + \| y_0 \| \leqslant 2r \qquad \forall x \in E$$

luego $h(E) \subset E$. Además, h es contractiva, pues de (7) deducimos claramente que,

$$\| h(x) - h(z) \| = \| g(x) - g(z) \| \leqslant (1/2) \| x - z \| \qquad \forall x, z \in E$$

Por el teorema del punto fijo, existe un único $x_0 \in E$ tal que $h(x_0) = x_0$ y, en vista de (9), x_0 es el único punto de $\overline{B}(0, 2r)$ tal que $f(x_0) = y_0$. Por último, si $\| y_0 \| < 2r$, razonando como antes tenemos $\| x_0 \| = \| h(x_0) \| \leqslant (1/2) \| x_0 \| + \| y_0 \| \leqslant r + \| y_0 \| < 2r$. Queda así comprobada la afirmación **(*)**, y el resto de la demostración se obtendrá ya sin dificultad.

Concretamente tomamos

$$U = B(0,2r) \cap f^{-1}\big(B(0,r)\big)$$

Como f es continua, el conjunto $f^{-1}\big(B(0,r)\big)$ es abierto, luego U también es abierto, y es claro que $0 \in U \subset \Omega$. Comprobamos ahora todas las afirmaciones del teorema.

(i). Si $x,z \in U$ verifican que $f(x) = f(z)$, tomando $y_0 = f(x) \in B(0,r)$, de **(*)** deducimos que existe un único $x_0 \in B(0,2r)$ tal que $f(x_0) = y_0$, pero $x,z \in B(0,2r)$, luego $x = z = x_0$ y hemos probado que f es inyectiva en U.

(ii). Tomando $V = f(U)$, es claro que $V \subset B(0,r)$ y comprobamos enseguida que se da la igualdad, con lo que V es una bola abierta. Para cada $y_0 \in B(0,r)$, usando de nuevo **(*)** tenemos un $x_0 \in B(0,2r)$ tal que $f(x_0) = y_0$, pero entonces $x_0 \in U$, luego $y_0 \in f(U)$ como queríamos.

(iii). Para $x \in U$ tenemos $\|x\| < 3r$ y (6) nos dice que $\det Jf(x) \neq 0$.

(iv). Usando la regla de diferenciación de la función inversa, y teniendo en cuenta (iii), basta probar que φ^{-1} es continua, y de hecho veremos que es lipschitziana.

Para $y_1, y_2 \in V$, sean $x_1 = \varphi^{-1}(y_1)$ y $x_2 = \varphi^{-1}(y_2)$, con lo que $x_1, x_2 \in U$, $f(x_1) = y_1$ y también $f(x_2) = y_2$. Por definición de g, tenemos $x_1 = y_1 + g(x_1)$ y $x_2 = y_2 + g(x_2)$. Por tanto, usando (7) obtenemos

$$\|x_1 - x_2\| \leqslant \|y_1 - y_2\| + \|g(x_1) - g(x_2)\| \leqslant \|y_1 - y_2\| + (1/2)\|x_1 - x_2\|$$

Deducimos claramente que

$$\|\varphi^{-1}(y_1) - \varphi^{-1}(y_2)\| = \|x_1 - x_2\| \leqslant 2\|y_1 - y_2\| \qquad \forall y_1, y_2 \in V$$

Queda así probado el teorema cuando $a = f(a) = 0$ y $Df(a) = \mathrm{Id}$. Completamos ahora la demostración, viendo que el caso general se deduce del que ya tenemos resuelto.

Sea $\Omega_0 = \{z \in \mathbb{R}^N : z + a \in \Omega\}$ que es un abierto de $\mathbb{R}^N$, como imagen inversa de Ω por una traslación, y verifica que $0 \in \Omega_0$. Como, por hipótesis, $\det Jf(a) \neq 0$, tenemos que $T = Df(a)$ es biyectiva y usaremos su inversa. Consideramos entonces la función $f_0 : \Omega_0 \to \mathbb{R}^N$ dada por

$$f_0(z) = T^{-1}\big(f(z+a) - f(a)\big) \qquad \forall z \in \Omega_0 \tag{10}$$

que verifica $f_0(0) = 0$. La regla de la cadena nos dice que f_0 es diferenciable con

$$Df_0(z) = T^{-1} \circ Df(z+a) \qquad \forall z \in \Omega_0 \tag{11}$$

y en particular $Df_0(0) = T^{-1} \circ T = \mathrm{Id}$. También deducimos que Df_0 en continua en el origen, pues para $z \in U_0$, se tiene

$$\big\|Df_0(z) - Df_0(0)\big\| = \big\|T^{-1} \circ \big(Df(z+a) - T\big)\big\| \leqslant \big\|T^{-1}\big\| \, \big\|Df(z+a) - Df(a)\big\|$$

y basta tener en cuenta que Df es continua en a. En resumen, f_0 verifica las hipótesis del teorema, en el caso particular que ya hemos resuelto.

Por tanto, tenemos $0 \in U_0 = U_0^\circ \subset \Omega_0$, con f_0 inyectiva en U_0, el conjunto $V_0 = f_0(U_0)$ es abierto, $\det Jf_0(z) \neq 0$ para todo $z \in U_0$ y $\varphi_0^{-1} \in D(V_0, \mathbb{R}^N)$ donde $\varphi_0 = f_0 \big|_{U_0}$. Sólo queda hacer el camino de vuelta, traduciendo toda esta información en términos de f.

Para ello, sea $U = \{z + a : z \in U_0\}$ que claramente es un abierto de $\mathbb{R}^N$ con $a \in U \subset \Omega$. A partir de (10) deducimos claramente que

$$f(x) = T\big(f_0(x-a)\big) + f(a) \qquad \forall x \in \Omega \tag{12}$$

mientras que (11) se traduce en

$$Df(x) = T \circ Df_0(x-a) \qquad \forall x \in \Omega \tag{13}$$

Comprobamos ya, de forma bastante rutinaria, que f tiene en U las propiedades requeridas.

(i). Si $f(x_1) = f(x_2)$ con $x_1, x_2 \in U$, usando (12) tenemos $T\big(f_0(x_1 - a)\big) = T\big(f_0(x_2 - a)\big)$, pero T es biyectiva, luego $f_0(x_1 - a) = f_0(x_2 - a)$. Como $x_1 - a, x_2 - a \in U_0$ y f_0 es inyectiva en U_0, concluimos que $x_1 = x_2$, luego f es inyectiva en U.

(ii). Veamos que $V = f(U)$ es abierto. Como V_0 es abierto y T es un homeomorfismo, $T(V_0)$ es abierto, y de (12) se deduce que $V = \{v \in \mathbb{R}^N : v - f(a) \in T(V_0)\}$, luego V es abierto.

(iii). Para $x \in U$ tenemos $x - a \in U_0$, luego $\det Jf_0(x-a) \neq 0$, así que $Df_0(x-a)$ es biyectiva. Como T también lo es, deducimos de (13) que $Df(x)$ es biyectiva, es decir, $\det Jf(x) \neq 0$.

(iv). Si $\varphi = f\big|_U$, veamos la relación entre φ^{-1} y φ_0^{-1}. Para $y \in V$ se tiene $y - f(a) \in T(V_0)$, luego $T^{-1}\big(y - f(a)\big) \in V_0$, lo que permite tomar $x = \big(\varphi_0^{-1} \circ T^{-1}\big)\big(y - f(a)\big) + a \in U$, y usando (12) vemos claramente que $f(x) = y$. Esto prueba que

$$\varphi^{-1}(y) = \big(\varphi_0^{-1} \circ T^{-1}\big)\big(y - f(a)\big) + a \qquad \forall y \in V$$

donde vemos claramente que φ^{-1} es continua. Como φ es diferenciable y $D\varphi(x) = Df(x)$ es biyectiva para todo $x \in U$, la regla de diferenciación de la función inversa nos dice que φ^{-1} es diferenciable con $D\varphi^{-1}\big(f(x)\big) = Df(x)^{-1}$ para todo $x \in U$. ■

El teorema anterior se resume en la existencia y diferenciabilidad de la función φ^{-1}, a la que conviene dar un nombre. No es inversa de f, pues f puede no ser inyectiva, es la inversa de la restricción de f a un entorno abierto de a. Por ello se dice que φ^{-1} es una **inversa local** de f en el punto a. Nótese que φ^{-1} está definida en un entorno abierto de $f(a)$.

11.3. Coordenadas polares, cilíndricas y esféricas

Veamos ejemplos sencillos que muestran la utilidad del teorema anterior. Consideremos el abierto $\Omega = \mathbb{R}^+ \times \mathbb{R} \subset \mathbb{R}^2$ y la función $f : \Omega \to \mathbb{R}^2$ definida por

$$f(\rho, \theta) = \big(\rho \cos \theta, \rho \operatorname{sen} \theta\big) \qquad \forall (\rho, \theta) \in \Omega$$

Es fácil ver que $f(\Omega) = \mathbb{R}^2 \setminus \{(0,0)\}$, pero f no es inyectiva. Para cada $(\rho, \theta) \in \Omega$, se dice que el punto $(x, y) = f(\rho, \theta)$ tiene **coordenadas polares** (ρ, θ), pero no son únicas. De hecho, el *radio polar* ρ es único, pero el *ángulo polar* θ no lo es.

Se tiene $f \in C^1(\Omega, \mathbb{R}^2)$ con $\det Jf(\rho,\theta) = \rho \neq 0$ para todo $(\rho,\theta) \in \Omega$, luego existe una inversa local de f en cada punto de Ω. Esto significa que, en un entorno de cada punto del plano, excluido el origen, podemos definir sin ambigüedad coordenadas polares, y obtenerlas a partir de las cartesianas, mediante una función diferenciable.

Esto puede comprobarse directamente, pues f es una función sencilla. Dado $(\rho_0,\theta_0) \in \Omega$, es fácil encontrar un abierto concreto U, con $(\rho_0,\theta_0) \in U \subset \Omega$, tal que f es inyectiva en U, ver que $V = f(U)$ es abierto, e incluso, si $\varphi = f\big|_U$, calcular explícitamente φ^{-1}, ver que es diferenciable y calcular su diferencial en cada punto de V. Sin embargo, debemos pensar que el teorema seguiría siendo útil, aunque f fuese más complicada, y no pudiésemos encontrar explícitamente el abierto U y la función φ^{-1}, debiendo contentarnos con saber que existen. Por tanto, para mostrar esa utilidad mediante f, debemos extraer alguna consecuencia del teorema que no precise calcular explícitamente U y φ^{-1}, como puede ser la siguiente:

- *Sea* $g : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ *un campo escalar y consideremos la función* $h : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ *definida por:*

$$h(\rho,\theta) = g\big(\rho\cos\theta\,,\,\rho\operatorname{sen}\theta\big) \qquad \forall\,(\rho,\theta) \in \mathbb{R}^+ \times \mathbb{R}$$

Entonces g *es diferenciable si, y sólo si, lo es* h.

Nótese que $h = g \circ f$ donde f es la función con la que veníamos trabajando. La regla de la cadena nos da una implicación: h será diferenciable siempre que g lo sea. El recíproco sería igual de fácil si f fuese biyectiva y f^{-1} fuese diferenciable, pues escribiríamos $g = h \circ f^{-1}$ y volveríamos a usar la regla de la cadena. Pero f no es inyectiva, no tiene "inversa global". Sin embargo, queremos comprobar una propiedad local, la diferenciabilidad de g, y para ello podemos usar una inversa local de f.

Fijado $(x_0,y_0) \in \mathbb{R}^2 \setminus \{(0,0)\}$, tomamos $(\rho_0,\theta_0) \in \Omega = \mathbb{R}^+ \times \mathbb{R}$ tal que $f(\rho_0,\theta_0) = (x_0,y_0)$. Entonces existe un abierto U de $\mathbb{R}^2$, con $(\rho_0,\theta_0) \in U \subset \Omega$, tal que f es inyectiva en U, el conjunto $V = f(U)$ es abierto y φ^{-1} es diferenciable, donde $\varphi = f\big|_U$. Para $(x,y) \in V$ se tiene que $\varphi^{-1}(x,y) \in U$ luego $f\big(\varphi^{-1}(x,y)\big) = \varphi\big(\varphi^{-1}(x,y)\big) = (x,y)$, de donde deducimos que

$$g(x,y) = g\Big(f\big(\varphi^{-1}(x,y)\big)\Big) = h\big(\varphi^{-1}(x,y)\big) \qquad \forall\,(x,y) \in V$$

Así pues, tenemos $g\big|_V = h \circ \varphi^{-1}$. Como h y φ^{-1} son diferenciables, la regla de la cadena nos dice que $g\big|_V$ es diferenciable. Por el carácter local de la diferenciabilidad, g es diferenciable en (x_0,y_0) como queríamos demostrar. ■

En $\mathbb{R}^3$ podemos definir coordenadas muy similares a las polares del plano, pero ahora hay que excluir, no sólo el origen, sino uno de los ejes de coordenadas. Para concretar excluimos el tercero, y trabajamos con el abierto $G = \mathbb{R}^3 \setminus \{(0,0,z) : z \in \mathbb{R}\}$. Se tiene entonces $G = f(\Omega)$, donde $\Omega = \mathbb{R}^+ \times \mathbb{R}^2$ y $f : \Omega \to \mathbb{R}^3$ viene dada por

$$f\big(\rho\,,\theta\,,z\big) = \big(\rho\cos\theta\,,\,\rho\operatorname{sen}\theta\,,\,z\big) \qquad \forall\,(\rho,\theta,z) \in \Omega$$

con lo que f no es inyectiva. Para cada $(\rho,\theta,z) \in \Omega$, decimos que el punto $f(\rho,\theta,z) \in G$ tiene **coordenadas cilíndricas** (ρ,θ,z), que no son únicas.

Es claro que $f \in C^1(\Omega, \mathbb{R}^3)$ con $\det Jf(\rho, \theta)| = \rho \neq 0$ para todo $(\rho, \theta) \in \Omega$. En un entorno de cada punto de G, podemos definir coordenadas cilíndricas, como función diferenciable de las cartesianas. Razonando como en el plano, vemos que un campo escalar definido en G es diferenciable en coordenadas cartesianas si, y sólo si, lo es en coordenadas cilíndricas.

Veamos otras coordenadas en G, más útiles que las cilíndricas. Tenemos ahora $G = f(\Omega)$, donde tomamos $\Omega = \{(r, \theta, \varphi) \in \mathbb{R}^3 : r > 0,\ |\varphi| < \pi/2\}$ y $f : \Omega \to \mathbb{R}^3$ viene dada por

$$f(r, \theta, \varphi) = \big(r \cos\theta \cos\varphi,\, r \operatorname{sen}\theta \cos\varphi,\, r \operatorname{sen}\varphi \big) \qquad \forall (r, \theta, \varphi) \in \Omega$$

Para cada $(r, \theta, \varphi) \in \Omega$, decimos que $f(r, \theta, z)$ tiene **coordenadas esféricas** (r, θ, φ), que no son únicas.

Es fácil ver que $f \in C^1(\Omega, \mathbb{R}^3)$, con $\det Jf(r, \theta, \varphi) = r^2 \cos\varphi \neq 0$ para todo $(r, \theta, \varphi) \in \Omega$. Todo lo dicho sobre las coordenadas cilíndricas es también válido para las esféricas.

11.4. Función inversa global

Como hicimos con la versión local, recordemos el teorema de la función inversa global para funciones reales de variable real:

- *Sea $I \subset \mathbb{R}$ un intervalo no trivial y $f : I \to \mathbb{R}$ una función derivable en I con $f'(x) \neq 0$ para todo $x \in I$. Entonces f es inyectiva, $J = f(I)$ es un intervalo y f^{-1} es derivable en J con $\big(f^{-1}\big)'\big(f(x)\big) = 1/f'(x)$ para todo $x \in I$.*

A diferencia de lo que ocurrió con el teorema local, la generalización del resultado anterior que cabría esperar, no es cierta. Concretamente, sustituyendo como es natural el intervalo I por un abierto $\Omega \subset \mathbb{R}^N$, que se puede suponer conexo o incluso convexo, podríamos pensar que una función diferenciable $f : \Omega \to \mathbb{R}^N$ verificando que $\det Jf(x) \neq 0$ para todo $x \in \Omega$, tenga que ser inyectiva, pero esto es falso: basta considerar cualquiera de las funciones que acabamos de usar para estudiar distintos tipos de coordenadas. Todas ellas son inyectivas en un entorno de cada punto de un abierto convexo Ω, podríamos decir que son localmente inyectivas, pero ninguna es inyectiva en Ω. Debemos pues aceptar que, para funciones de varias variables, la versión global del teorema de la función inversa va a ser mucho más débil que en el caso de una variable, pues la inyectividad de la función no será parte de la tesis, sino que habrá de suponerse como hipótesis. Enunciamos ya el resultado global que podemos conseguir.

Teorema de la función inversa (global). *Sea Ω un abierto de $\mathbb{R}^N$ y $f \in C^1(\Omega, \mathbb{R}^N)$. Supongamos que f es inyectiva y que $\det Jf(x) \neq 0$ para todo $x \in \Omega$. Entonces $W = f(\Omega)$ es abierto y f^{-1} es diferenciable con $Df^{-1}\big(f(x)\big) = Df(x)^{-1}$ para todo $x \in \Omega$. De hecho se tiene que $f^{-1} \in C^1(W, \mathbb{R}^N)$.*

Demostración. Sea $y \in W$ y $x = f^{-1}(y) \in \Omega$. Como Df es continua en x y $\det Jf(x) \neq 0$, podemos aplicar el teorema local, obteniendo un abierto U de $\mathbb{R}^N$, con $x \in U \subset \Omega$, verificando las propiedades que vamos a comentar.

En primer lugar, que f sea inyectiva en U no es nada nuevo, pues f es inyectiva en Ω. Pero también tenemos que $V = f(U)$ es abierto, con $y = f(x) \in f(U) = V \subset W$, luego $y \in W^\circ$. Antes de seguir con el razonamiento, como $y \in W$ era arbitrario, tenemos ya probado que W es abierto. De vuelta al punto $y \in W$ antes fijado, si $\varphi = f\big|_U$, está claro que $\varphi^{-1} = f^{-1}\big|_V$. Como el teorema local nos dice que φ^{-1} es diferenciable, por el carácter local de la diferenciabilidad, deducimos que f^{-1} es diferenciable en el punto y. Otra vez $y \in W$ era arbitrario, luego f^{-1} es diferenciable. La regla de derivación de la función inversa nos da la diferencial de f^{-1} en cada punto de W. Para ver que Df^{-1} es continua, comprobaremos previamente el siguiente resultado.

- *Sea $\mathcal{G}$ el subconjunto de $L(\mathbb{R}^N, \mathbb{R}^N)$ formado por las aplicaciones lineales biyectivas. Entonces, la aplicación $T \mapsto T^{-1}$, de $\mathcal{G}$ en sí mismo, es continua.*

Empezamos comprobando la continuidad de la función $\alpha : \mathcal{G} \to \mathbb{R}^+$ definida por

$$\alpha(T) = \frac{1}{\|T^{-1}\|} \qquad \forall\, T \in \mathcal{G}$$

De hecho, para $S, T \in \mathcal{G}$ tenemos

$$\begin{aligned} |\,\alpha(T) - \alpha(S)\,| &= \frac{\big|\,\|S^{-1}\| - \|T^{-1}\|\,\big|}{\|T^{-1}\|\,\|S^{-1}\|} \leqslant \frac{\|S^{-1} - T^{-1}\|}{\|T^{-1}\|\,\|S^{-1}\|} = \frac{\|T^{-1} \circ \big(T - S\big) \circ S^{-1}\|}{\|T^{-1}\|\,\|S^{-1}\|} \\ &\leqslant \frac{\|T^{-1}\|\,\|T - S\|\,\|S^{-1}\|}{\|T^{-1}\|\,\|S^{-1}\|} = \|T - S\| \end{aligned}$$

Así pues, vemos que de hecho α es no expansiva, luego continua. Deducimos obviamente que la función $1/\alpha$, es decir, la función $T \mapsto \|T^{-1}\|$, de $\mathcal{G}$ en $\mathbb{R}^+$, también es continua.

Finalmente, dada $T \in \mathcal{G}$ y una sucesión $\{T_n\}$ de elementos de $\mathcal{G}$ con $\{T_n\} \to T$, bastará comprobar que $\{T_n^{-1}\} \to T^{-1}$. En efecto, para todo $n \in \mathbb{N}$ se tiene

$$\|T_n^{-1} - T^{-1}\| = \|T^{-1} \circ (T - T_n) \circ T_n^{-1}\| \leqslant \|T^{-1}\|\,\|T - T_n\|\,\|T_n^{-1}\|$$

Como ya sabemos que $\big\{\|T_n^{-1}\|\big\} \to \|T^{-1}\|$, vemos que $\big\{\|T_n^{-1} - T^{-1}\|\big\} \to 0$. □

Para terminar la demostración del teorema de la función inversa global, teníamos ya

$$Df^{-1}(y) = \big[Df\big(f^{-1}(y)\big)\big]^{-1} \qquad \forall\, y \in W$$

y vemos que $Df^{-1} = J \circ Df \circ f^{-1}$, donde J es la función $T \mapsto T^{-1}$, de $\mathcal{G}$ en sí mismo, cuya continuidad acabamos de comprobar. Como f^{-1} es continua, porque es diferenciable, y Df es continua por hipótesis, concluimos que Df^{-1} es continua, como composición de tres funciones continuas. Así pues, $f^{-1} \in C^1(W, \mathbb{R}^N)$ como queríamos demostrar. ■