El teorema de la función inversa
Óscar R.S.

Uno de los problemas elementales que nos plantea el álgebra lineal es resolver sistemas lineales de la forma $Ax = b$, donde $A \in M_n(\mathbb{R})$ y $b \in \mathbb{R}^n$ son dados y buscamos hallar un $x \in R^n$ que cumpla las condiciones. Sabemos que esto lo podemos hacer si y solo si A tiene determinante no nulo, y que en ese caso la función inversa viene dada simplemente por la inversa de la matriz, esto es, dado $b \in \mathbb{R}^n$, $x = A^{-1}b$ es el único vector que cumple $Ax = b$.

Ahora bien, cuando trascendemos el mundo lineal la situación puede ser más complicada: sea $F : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $F(x, y) = (e^x \cos y, e^x \sin y)$. Se tiene que

$$DF(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}.$$

Esta función, si pensamos en $\mathbb{R}^2$ como $\mathbb{C}$, es simplemente la exponencial, cuya inversa es el logaritmo. Pero no se trata de una inversa global, básicamente porque esta función no es inyectiva: la exponencial compleja es $2\pi i$ periódica; puedo por tanto definir una inversa local, pero no global.

La aproximación que realizaremos aquí al teorema de la función inversa será entonces meramente local. Buscamos unas condiciones bajo las cuales podamos garantizar que existe una inversa local de la función. El primer requerimiento ha de ser necesariamente la inyectividad, pues sino el concepto no tendría sentido (esto es lo mismo que pedir que sea biyectiva con la imagen). Esta condición de inyectividad la codificará la diferencial: este objeto ha de pensarse como una aplicación lineal entre los espacios tangentes en p y en $F(p)$ así que lo que pediremos es que sea de rango máximo (por tanto biyectiva). Recuérdese que por Taylor,

$$F(p + tv) = F(p) + tDF(p)v + o(t).$$

Esto indica que a primer orden, el movimiento en una cierta dirección v pasa a ser en dirección $DF(p)v$ así que la biyectividad de la diferencial me indica que cada dirección a la salida la alcanzaré de una manera (única) a partir de una dirección en la entrada. Dicho informalmente, queremos que sea equivalente que F sea un difeomorfismo local (biyectiva y tanto función como inversa diferenciables) a que DF sea un isomorfismo. Pero esto de entrada no será cierto; hemos de pedir que la función sea C^1 para que todo funcione, o al menos que haya un entorno del punto donde el determinante de la diferencial no se anule (demostrarlo con esta condición más débil es más complicado). Intuitivamente, si hubiese puntos arbitrariamente próximos donde la diferencial se anula podría no tener definida la inversa.

Teorema 1. *Sea $F : A \subset \mathbb{R}^n \to \mathbb{R}^n \in C^1(A)$ (A abierto). Sea $p \in A$ de forma que $\det JF(p) \neq 0$. Entonces, existe un abierto U que contiene a p de forma tal que $V = F(U)$ es un abierto y se puede definir $F^{-1} : V \to U \in C^1$. Se cumple además que $DF^{-1}(y) = (DF(F^{-1}(y)))^{-1}$.*

Antes de comenzar con la demostración notemos que al ser la diferencial una aplicación lineal simplemente haciendo un cambio de base en el espacio de llegada podemos

suponer que es la identidad. En efecto, escribimos $DF^{-1}(F(p)) = T^{-1}$ y de esta manera $g = T^{-1} \circ F$ cumple que $DG(p) = \mathrm{Id}$. Supongamos el teorema cierto cuando la jacobiana es la identidad. Entonces, puedo definir $g^{-1} = (T^{-1} \circ F)^{-1}$. Definiendo ahora $F^{-1} = g^{-1} \circ T^{-1}$, tenemos una inversa para F dado que $F \circ F^{-1} = F^{-1} \circ F = \mathrm{Id}$. Supondremos ahora en todo momento que la jacobiana en el punto p vale 0.
Pasamos ahora a la parte más sutil que es la construcción de la inversa. El teorema que me aporta existencia y unicidad para la solución de ciertas ecuaciones es el del punto fijo de Banach; para ello, tengo que estar en un cerrado de un espacio de Banach y tener una función que lo aplica en él mismo. El problema de la F podría ser que creciese rápido y que por tanto no hubiese ningún entorno del origen que se aplicase en él mismo. Es por eso que definiré

$$g(x) = x - F(x)$$

cuya diferencial en el origen es 0 y por ser la función C^1 es continua, con lo que en un entorno del origen (pongamos $||x|| < r$) se cumple que $||Dg_i(x)|| < \frac{1}{2n}$ para todo i. Esto implicará que $g(\bar{B}(0,r)) \subset \bar{B}(0, r/2)$, simplemente por el teorema del valor medio dado que

$$||f(x)|| \leq \sum_{i=1}^{n} ||g_i(x)|| = \sum_{i=1}^{n} ||Dg_i(c_i)(x)|| \leq \sum_{i=1}^{n} ||Dg_i(c_i)|| \cdot ||x|| < \frac{||x||}{2} < \frac{r}{2}.$$

Por tanto, lo que sé es que $\bar{B}(0,r)$ va de él mismo en él mismo al pasar por g. A partir de aquí, definiendo para todo $y \in B(0, r/2)$

$$g_y(x) = y + x - F(x)$$

tenemos que esta aplicación es 1/2-contractiva y esto nos permite concluir usando el teorema del punto fijo (miro ahora las dos hipótesis): cuando $y \in \bar{B}(0, r/2)$,

$$||g_y(x)|| = ||y + g(x)|| \leq ||y|| + ||g(x)|| < r.$$

Además, si $x_1, x_2 \in \bar{B}(0,r)$,

$$||g_y(x_1) - g_y(x_2)|| = ||g(x_1) - g(x_2)|| \leq \frac{1}{2}||x_1 - x_2||$$

Se cumplen entonces las hipótesis y existira un único punto $x \in \bar{B}(0,r)$ tal que $f(x) = y$. Concluimos que f tiene una inversa $f^{-1} : \bar{B}(0, r/2) \to \bar{B}(0,r)$.
Cuidado aquí con las r y las $r/2$: yo sé que $\bar{B}(0,r)$ va a caer dentro de $\bar{B}(0,r)$, pero para definir la inversa me estoy restringiendo en el espacio de llegada a puntos en $\bar{B}(0, r/2)$, cuya antiimagen será un cierto abierto dentro de $\bar{B}(0,r)$.

Lema 1. *f^{-1} es 2-Lipschitz y por tanto continua.*

Demostración. Esto equivale a probar

$$||x_1 - x_2|| \leq 2||f(x_1) - f(x_2)||$$

lo cual es consecuencia de que

$$||x_1 - x_2|| \leq ||f(x_1) - f(x_2)|| + ||g(x_1) - g(x_2)|| \leq ||f(x_1) - f(x_2)|| + \frac{1}{2}||x_1 - x_2||$$

Una vez tenemos esto, por existir localmente la inversa, ponemos $x_1 = f(y_1)$, $x_2 = f(y_2)$ y hemos acabado. $\square$

Lema 2. *Para un r suficientemente pequeño, la inversa es diferenciable en $B(0,r/2)$.*

Demostración. Como la función es de clase C^1, si sabemos que $Df(0)$ es invertible tenemos que en un entorno del 0, $(Df(x))^{-1}$ existirá (el determinante es una función polinómica de las parciales y por tanto continua). Tomemos entonces un r de forma tal que $B(0,r/2)$ está en el entorno donde existe la inversa de la diferencial; además, por estar en un compacto, la norma de la diferencial ha de estar acotada y se cumplirá que en todo $B(0,r/2)$, $||(Df(x))^{-1}|| \leq M||y||$.
Ahora todo se reduce a un cálculo; sean $y_1, y_2 \in B(0,r/2)$, $x_1 = f^{-1}(y_1), x_2 = f^{-1}(y_2)$, entonces

$$\frac{||f^{-1}(y_1) - f^{-1}(y_2) - (Df(x_2)^{-1}(y_1 - y_2)||}{||y_1 - y_2||} =$$

$$= \left[\frac{||x_1 - x_2||}{||f(x_1) - f(x_2)||}\right] \frac{||(Df(x_2))^{-1}(Df(x_2)(x_1 - x_2) - (f(x_1) - f(x_2)))||}{||x_1 - x_2||} \leq$$

$$\leq 2M \frac{||Df(x_2)(x_1 - x_2) - (f(x_1) - f(x_2))||}{||x_1 - x_2||}$$

cuyo límite es claramente cero cuando la norma de $x_1 - x_2$ tiende a 0 por la diferenciabilidad de f. De aquí se deduce ya que

$$(Df(x_2))^{-1} = (Df(f^{-1}(y_2)))^{-1}$$

□

Lema 3. *$f^{-1} : B(0,r/2) \to \mathbb{R}^n$ es de clase C^p.*

Demostración. Tenemos que probar que $Df^{-1}(y) = (Df(f^{-1}(y)))^{-1}$ es continua, pero lo es por ser composición de continuas, ya que tanto f^{-1} como Df lo son e invertir una matriz es una función racional, que de hecho es C^∞ allá donde el denominador no se anula.
En el caso en el que la f es de clase C^p procedemos por inducción; una vez sabemos que es de clase C^1, como Df es de clase C^{p-1} y la inversa C^∞, resultará que Df^{-1} es de clase C^1, y así sucesivamente hasta tener que la diferencial es de clase C^{p-1} y por tanto f de clase C^p. □

El teorema de la función inversa me está diciendo que por el hecho de ser C^1 y tener la diferencial rango máximo, localmente tengo un difeomorfismo entre un abierto que contiene a mi punto y la imagen; el ser un difeomorfismo me dice en particular que f es localmente inyectiva.
Conviene tener muy presente que inyectividad global más ser difeomorfismo local implica que es un difeomorfismo con la imagen: para ello hay que ver que tengo una inversa global bien definida de clase C^1 (eso es fácil, dado un punto considero su antiimagen y a partir de él tengo unos entornos del punto y la imagen donde todo funciona bien) y además tiene la virtud de transformar abiertos en abiertos (es lo que se dice una aplicación abierta).
Un corolario casi inmediato del teorema de la función inversa es el teorema de la función implícita:

Teorema 2. *Sea $A \subset \mathbb{R}^n \times \mathbb{R}^m$ un abierto y sea $F : A \to \mathbb{R}^m$ una función de clase C^p. Sea $(x_0, y_0) \in A$ de forma que $F(x_0, y_0) = 0$. Consideremos el menor de la jacobiana formado al derivar respecto a las m últimas variables, y supongamos que en (x_0, y_0)*

no es cero. Entonces, existe un entorno abierto $U \subset \mathbb{R}^n$ *de* x_0 *y un entorno* V *de* y_0 *en* $\mathbb{R}^m$ *de forma que existe* $f : U \to V$ *de forma que* $F(x, f(x)) = 0$ *para todo* $x \in U$*. Además* f *es de clase* C^p*.*

La demostración se basa simplemente en definir $G(x, y) = (x, F(x, y))$. La matriz jacobiana tiene por hipótesis determinante no nulo y todo es C^1. Así que por el teorema de la función inversa resulta muy sencillo concluir.