

3 Funciones inversas e implícitas

3.1 Sucesiones de Cauchy

Definición 77. Sea (X, d) un espacio métrico. Una sucesión $\{x_n\}_{n=1}^{\infty}$ de puntos de X es una **sucesión de Cauchy** si para todo $\varepsilon > 0$ existe un $k = k(\varepsilon)$ tal que:

$$n, m \geq k \implies d(x_n, x_m) \leq \varepsilon .$$

Esta definición generaliza a cualquier espacio métrico la noción de sucesión de Cauchy de números reales (que es el caso $X = \mathbb{R}$). Recordemos que, en tal caso particular, la utilidad que tiene es que podamos decir si una sucesión de números es convergente o no *sin necesidad de saber cuál es el límite*. Nos hace el mismo servicio en cada espacio de dimensión finita.

Proposición 78. Una sucesión en \mathbb{R}^n es convergente si y sólo si es de Cauchy.

Demostración. Si $\{x_k\} \subset \mathbb{R}^n$ es de Cauchy entonces para cada $i \in \{1, \dots, n\}$ la sucesión de las i -ésimas coordenadas $x_{1i}, x_{2i}, x_{3i}, \dots$ es de Cauchy en \mathbb{R} , luego convergente a un número a_i . Se sigue que $\{x_k\}$ converge a (a_1, \dots, a_n) . El recíproco es aún más fácil de demostrar. \square

Definición 79. Un espacio métrico (X, d) es **completo** si toda sucesión de Cauchy en X tiene un límite en X .

Son completos: \mathbb{R}^n , cualquier espacio métrico compacto.

Un subconjunto $E \subseteq \mathbb{R}^n$ es completo (con la distancia inducida de \mathbb{R}^n) si y sólo si es un cerrado; esto proporciona infinidad de ejemplos de espacios completos, muchos de ellos no compactos; también proporciona infinidad de ejemplos de espacios no completos.

3.2 Aplicaciones de Lipschitz

Definición 80. Una aplicación entre espacios métricos $f : (X, d_X) \rightarrow (Y, d_Y)$ es **de Lipschitz** si existe una constante $K > 0$ tal que:

$$d_Y(f(x'), f(x)) \leq K d_X(x, x') \quad , \quad \text{para cualesquiera } x, x' \in X .$$

Las K que cumplen esta condición son las **constantes de Lipschitz** de f .

Toda aplicación Lipschitz es continua pero, por ejemplo, $f(x) \equiv \sqrt[3]{x}$ es continua y no Lipschitz. El siguiente resultado es útil para obtener una constante de Lipschitz a partir de una cota de las derivadas primeras.

Proposición 81. Sea $U \subset \mathbb{R}^n$ un abierto convexo y $f : U \rightarrow \mathbb{R}^m$ una función de clase \mathcal{C}^1 . Dadas normas cualesquiera en \mathbb{R}^n y \mathbb{R}^m , consideremos la norma de operador en $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ correspondiente. Si $\|(df)_x\| \leq K$ para todo $x \in U$, entonces f es de Lipschitz con constante K .

Demostración. Sean $x, x' \in U$ y pongamos $h = x' - x$. Al ser U convexo, todo el segmento $[x, x'] = \{x + th : t \in [0, 1]\}$ está contenido en U y tenemos la siguiente igualdad:

$$f(x') - f(x) = \int_0^1 \frac{d}{dt} f(x + th) dt = \int_0^1 (Df_{x+th})h dt .$$

Es fácil probar la desigualdad $\left\| \int_{t_0}^{t_1} v(t) dt \right\|_{\mathbb{R}^n} \leq \int_{t_0}^{t_1} \|v(t)\|_{\mathbb{R}^n} dt$ para toda función vectorial continua $v(t)$, poniendo las dos integrales como límites de *sumas de Riemann*, y se deduce:

$$\begin{aligned} \|f(x') - f(x)\|_{\mathbb{R}^m} &= \left\| \int_0^1 (Df_{x+th})h dt \right\|_{\mathbb{R}^m} \leq \int_0^1 \|(Df_{x+th})h\|_{\mathbb{R}^m} dt \leq \\ &\leq \int_0^1 \|Df_{x+th}\| \|h\|_{\mathbb{R}^n} dt \leq \|h\|_{\mathbb{R}^n} \int_0^1 K dt = K \|x' - x\|_{\mathbb{R}^n} . \end{aligned}$$

\square

3.3 Aplicaciones contractivas

Definición 82. Una *contracción* o *aplicación contractiva* es una aplicación f que cumple las dos condiciones siguientes, siendo ambas igual de importantes:

1. f es de Lipschitz y admite una constante de Lipschitz $K < 1$.
2. Dominio y codominio coinciden, distancias incluidas: $f : (X, d) \rightarrow (X, d)$.

Teorema 83. (Teorema de la aplicación contractiva). Si $f : X \rightarrow X$ es contractiva y X es completo, entonces f tiene un único **punto fijo**, es decir que existe un único $x \in X$ tal que $f(x) = x$.

Demostración.

Existencia. Empecemos con cualquier punto $a \in X$ y construyamos la sucesión $\{x_n\}_{n=1}^{\infty}$ definida por las condiciones: $x_1 = f(a)$ y $x_{n+1} = f(x_n)$, es decir $x_n = \underbrace{(f \circ \dots \circ f)}_{n \text{ factores}}(a)$.

Sea $K < 1$ una constante de Lipschitz para f . Se tiene:

$$d(x_1, x_2) = d(f(a), f(x_1)) \leq K d(a, x_1),$$

del mismo modo $d(x_2, x_3) \leq K d(x_1, x_2) \leq K^2 d(a, x_1)$. En general $d(x_n, x_{n+1}) \leq K^n d(a, x_1)$ para todo n . Estimemos la distancia entre dos términos de la sucesión no necesariamente consecutivos:

$$\begin{aligned} d(x_n, x_{n+k}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{n+k-1}, x_{n+k}) \leq \\ &\leq (K^n + K^{n+1} + \dots + K^{n+k-1}) d(a, x_1) \leq d(a, x_1) \sum_{j=n}^{\infty} K^j = d(a, x_1) \frac{K^n}{1-K}. \end{aligned}$$

Vemos que $d(x_n, x_{n+k})$ desciende a cero como $\text{cte} \cdot K^n$, luego $\{x_n\}$ es una sucesión de Cauchy y, como X es completo por hipótesis, hay un punto $x \in X$ con $x_n \rightarrow x$ cuando $n \rightarrow \infty$.

La aplicación f , siendo contractiva, es continua. Entonces de $x_n \rightarrow x$ se deduce $f(x_n) \rightarrow f(x)$ cuando $n \rightarrow \infty$. Pero esto último es $f(x) = \lim x_{n+1}$ y, como $\{x_{n+1}\}$ es una cola de $\{x_n\}$, se tiene $\lim x_{n+1} = \lim x_n = x$. Finalmente $f(x) = x$, lo que prueba la existencia de punto fijo.

Unicidad. Sea $x' \in X$ "otro" punto fijo: $f(x') = x'$. Razonamos así:

$$d(x, x') = d(f(x), f(x')) \leq K d(x, x'),$$

de donde $(1-K) d(x, x') \leq 0$ pero, como es $1-K > 0$, se deduce $d(x, x') \leq 0$, es decir $d(x, x') = 0$ y por lo tanto $x' = x$. \square

Es interesante que esta demostración es *constructiva*: el punto fijo se aproxima, a velocidad exponencial, por los puntos $x_n = f^n(a)$.

Es importante que el teorema funcione en cualquier espacio métrico, a condición de que sea completo. Aquí lo vamos a aplicar al caso de una bola cerrada $X = \overline{B}(x_0, r) \subset \mathbb{R}^n$, pero en otros contextos se lo utiliza con espacios de dimensión infinita.

Concentrándonos en el caso de dimensión finita n , el teorema nos proporciona una clase muy amplia de sistemas de n ecuaciones con n incógnitas para los que hay solución: los sistemas de la forma $f(x) = x$, donde f lleva una bola $\overline{B}(x_0, r)$ dentro de sí misma y es contractiva ahí.

Terminamos este apartado con una observación:

Fijamos una norma $\|\cdot\|$ en \mathbb{R}^n . Si h va de un dominio de \mathbb{R}^n a \mathbb{R}^n y cumple $\|h(x') - h(x)\| \leq K \|x' - x\|$, entonces para todo vector constante $\mathbf{c} \in \mathbb{R}^n$ la nueva función $F(x) \equiv \mathbf{c} + h(x)$ cumple $F(x') - F(x) = h(x') - h(x)$ para cualesquiera x, x' , luego también $\|F(x') - F(x)\| \leq K \|x' - x\|$.

3.4 Aplicaciones coercivas

Definición 84. Una función entre espacios métricos $f : (X, d_X) \rightarrow (Y, d_Y)$ es **coerciva** si existe una constante $\lambda > 0$ tal que $d_Y(f(x'), f(x)) \geq \lambda d_X(x', x)$ para cualesquiera $x, x' \in X$. Las constantes λ que cumplen esto se llaman **constantes de coercividad** de f .

La coercividad es una propiedad *estrictamente más fuerte* que la inyectividad. Toda aplicación coerciva es inyectiva:

$$x \neq x' \implies d(x, x') > 0 \implies d_Y(f(x), f(x')) \geq \lambda d(x, x') > 0 \implies f(x) \neq f(x'),$$

pero, por ejemplo, la exponencial $f(x) = e^x$ es una función inyectiva pero no coerciva.

Supongamos que $f : X \rightarrow Y$ es inyectiva y veamos qué más tiene que cumplir para ser coerciva. Una vez que es inyectiva, la función f induce una “biyección a la imagen”

$$f_{\text{im}} : X \rightarrow f(X) \quad , \quad x \mapsto f(x) \quad ,$$

y también tenemos en el conjunto imagen la distancia $d_{f(X)}$ inducida por la d_Y ; entonces f es coerciva, con constante de coercividad λ , si y sólo si la “inversa desde la imagen”

$$f_{\text{im}}^{-1} : (f(X), d_{f(X)}) \rightarrow (X, d_X) \quad , \quad f(x) \mapsto x \quad ,$$

es de Lipschitz con constante de Lipschitz $1/\lambda$. La exponencial $y = e^x$ no es coerciva porque su inversa desde la imagen $(0, +\infty) \ni y \mapsto \log y$ no es de Lipschitz.

Teorema 85. Supongamos elegida una norma en \mathbb{R}^n y sea $g : \overline{B}(\mathbf{0}, r) \rightarrow \mathbb{R}^n$ de Lipschitz con constante de Lipschitz $\varepsilon < 1$ y tal que $g(\mathbf{0}) = \mathbf{0}$. Entonces la función $f(x) \equiv x + g(x)$ tiene las dos propiedades siguientes:

(a) f es inyectiva en $\overline{B}(\mathbf{0}, r)$.

(b) $f(B(\mathbf{0}, r)) \supset B(\mathbf{0}, (1 - \varepsilon)r)$.

Demostración.

Propiedad (a). En realidad f es mejor que inyectiva: fácilmente se ve que es coerciva con constante de coercividad $1 - \varepsilon$.

Propiedad (b). Queremos ver que si y_0 es lo bastante cercano a $\mathbf{0}$ entonces el sistema $f(x) = y_0$ tiene solución $x_0 \in B(\mathbf{0}, r)$ (necesariamente única, por la inyectividad). Lo primero que hacemos es convertir ese sistema en uno de punto fijo:

$$f(x) = y_0 \iff x + g(x) = y_0 \iff x = -g(x) + y_0 \quad ,$$

es decir, queremos que la función $F_{y_0}(x) \equiv -g(x) + y_0$ tenga un punto fijo $x_0 \in B(\mathbf{0}, r)$. Ahora aplicamos la observación hecha al final del apartado 3.3: como $-g$ es de Lipschitz con constante ε , al sumarle el vector constante y_0 resulta F_{y_0} que admite la misma constante de Lipschitz. En vista del teorema 83, si F_{y_0} lleva la bola $\overline{B}(\mathbf{0}, r)$ dentro de sí misma entonces tendrá un punto fijo $x_0 \in \overline{B}(\mathbf{0}, r)$. De hecho se cumple algo un poco más fuerte:

$$\text{Para } y_0 \text{ cercano a } \mathbf{0}, \text{ se cumple } F_{y_0}(\overline{B}(\mathbf{0}, r)) \subseteq B(\mathbf{0}, r) \quad . \quad (29)$$

Entonces tendremos que el punto fijo $x_0 = F_{y_0}(x_0)$ está en la bola abierta $B(\mathbf{0}, r)$ y habremos conseguido un $x_0 \in B(\mathbf{0}, r)$ tal que $y_0 = f(x_0)$. Para probar (29) hacemos una estimación:

$$z \in \overline{B}(\mathbf{0}, r) \implies \|F_{y_0}(z)\| = \|-g(z) + y_0\| \leq \|g(\mathbf{0}) - g(z)\| + \|y_0\| \leq \varepsilon r + \|y_0\| \quad ,$$

y vemos que la desigualdad estricta $\|y_0\| < (1 - \varepsilon)r$ nos asegura que $F_{y_0}(\overline{B}(\mathbf{0}, r)) \subseteq B(\mathbf{0}, r)$. Por lo explicado, el sistema $f(x) = y_0$ tiene solución $x_0 \in B(\mathbf{0}, r)$ para todo $y_0 \in B(\mathbf{0}, (1 - \varepsilon)r)$. Es más, tenemos una fórmula para la solución:

$$x_0 = \lim_{n \rightarrow \infty} \underbrace{F_{y_0} \circ \cdots \circ F_{y_0}}_{n \text{ factores}}(\mathbf{0}) \quad .$$

□

Nota. El teorema 85 es cierto para cualquier **espacio de Banach** (espacio normado completo) en lugar de \mathbb{R}^n , sin hacer ningún cambio en la demostración.

3.5 Teorema de la función inversa: espacios normados

Teorema 86. (Teorema de la función inversa). *Sea $(\mathbb{V}, \|\cdot\|)$ un espacio de Banach. Sean $U_0 \subseteq \mathbb{V}$ un abierto y $f : U_0 \rightarrow \mathbb{V}$ función de clase \mathcal{C}^1 . Si en un punto $x^0 \in U_0$ la diferencial $L = (df)_{x^0}$ es un elemento invertible de $\mathcal{L}(\mathbb{V}, \mathbb{V})$, es decir que L es lineal acotada, biyectiva y con inversa L^{-1} también acotada, entonces existen abiertos $U \ni x^0$ y $V \ni y^0 = f(x^0)$ tales que f es biyectiva de U a V . Además, en ese caso la inversa $f^{-1} : V \rightarrow U$ es diferenciable en y^0 y su diferencial en y^0 es la que predice la regla de la cadena.*

Varios comentarios antes de pasar a la demostración.

(1) La condición $f \in \mathcal{C}^1$ significa que f es diferenciable en todo punto de U_0 y que la aplicación

$$U_0 \longrightarrow \mathcal{L}(\mathbb{V}, \mathbb{V}) \quad , \quad x \longmapsto (df)_x \quad ,$$

es continua cuando en $\mathcal{L}(\mathbb{V}, \mathbb{V})$ ponemos la norma de operador asociada a la de \mathbb{V} .

(2) En el caso $\mathbb{V} = \mathbb{R}^n$, las dimensiones de salida y de llegada son iguales. Esto es imprescindible para que la matriz jacobiana sea *cuadrada*, condición sin la cual no puede ser invertible.

(3) Si ya hemos probado que la inversa f^{-1} existe y es diferenciable en $y^0 = f(x^0)$, entonces es correcto aplicar la regla de la cadena a las identidades

$$x \equiv (f^{-1} \circ f)(x) \quad , \quad y \equiv (f \circ f^{-1})(y) \quad ,$$

y deducir las siguientes igualdades:

$$\text{id}_{\mathbb{V}} = (d(\text{id}_{\mathbb{V}}))_{x^0} = (d(f^{-1}))_{y^0} \circ (df)_{x^0} \quad , \quad \text{id}_{\mathbb{V}} = (d(\text{id}_{\mathbb{V}}))_{y^0} = (df)_{x^0} \circ (d(f^{-1}))_{y^0} \quad ,$$

o sea, para que f^{-1} sea diferenciable en y^0 es necesario que $(df)_{x^0}$ sea invertible, y en tal caso

$$\boxed{(d(f^{-1}))_{y^0} = [(df)_{x^0}]^{-1}} \quad (30)$$

De este modo la regla de la cadena predice un único valor posible para la diferencial de f^{-1} en $y^0 = f(x^0)$, caso de que f^{-1} exista y sea diferenciable en y^0 .

(4) Es de señalar que a veces f^{-1} puede existir cerca de y^0 aunque $(df)_{x^0}$ no sea invertible, pero en tal caso f^{-1} definitivamente no es diferenciable en y^0 . Por ejemplo $f(x) \equiv x^3$ tiene $f'(0) = 0$ y la inversa $f^{-1}(y) \equiv \sqrt[3]{y}$ existe pero no es diferenciable en $y = f(0) = 0$.

(5) De las dos funciones f y f^{-1} una puede ser elemental y la otra no. Por ejemplo $f(x) \equiv x + e^x$, que es biyectiva $\mathbb{R} \rightarrow \mathbb{R}$ y con derivada siempre positiva, es elemental pero su inversa no lo es.

(6) La hipótesis “ f de clase \mathcal{C}^1 ” del teorema no se puede debilitar a “ f diferenciable en todo punto”. Por ejemplo, la función

$$f(x) = \begin{cases} x + x^2 \text{sen} \frac{1}{x^4} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

es derivable en todo punto de \mathbb{R} y tiene $f'(0) = 1$, pero cambia infinitas veces de creciente a decreciente en cualquier entorno de $x = 0$. Esto le impide ser inyectiva en tales entornos.

Demostración del teorema 86.

Caso cero: $x^0 = y^0 = \mathbf{0}$ y $(df)_{\mathbf{0}} = \text{id}_{\mathbb{V}}$.

Definimos $g(x) = f(x) - x$, con lo cual g es \mathcal{C}^1 y se cumple la identidad $f(x) = x + g(x)$, que nos lleva al teorema 85. Además $(dg)_{\mathbf{0}} = 0$, que equivale a:

$$g(x) = o(\|x\|) \quad . \quad (31)$$

Toda norma es de Lipschitz, con constante 1, respecto de sí misma. En particular la norma de operador $\|\cdot\| : (\mathcal{L}(\mathbb{V}, \mathbb{V}), \|\cdot\|) \rightarrow \mathbb{R}$ es de Lipschitz y la función compuesta $\|(dg)_x\|$ es función continua de x y nula en $x = \mathbf{0}$. Entonces existen $r > 0$ y $0 < \varepsilon < 1$ tales que:

$$x \in \overline{B}(\mathbf{0}, r) \implies \|(dg)_x\| \leq \varepsilon \quad .$$

Como las bolas son convexas, la siguiente estimación:

$$\|g(x') - g(x)\| \leq \varepsilon \|x' - x\| \quad , \quad \text{para cualesquiera } x, x' \in \overline{B(\mathbf{0}, r)} ,$$

está garantizada en el caso $\mathbb{V} = \mathbb{R}^n$. También para \mathbb{V} general porque, aunque no lo vamos a probar, la proposición 81 del apartado 3.2 es válida en todos los espacios de Banach.

El teorema 85 se aplica directamente y nos dice que f es inyectiva en $\overline{B(\mathbf{0}, r)}$ y que la imagen inyectiva $f(B(\mathbf{0}, r))$ contiene a la bola abierta $B(\mathbf{0}, (1 - \varepsilon)r)$. Definimos los abiertos:

$$V = B(\mathbf{0}, (1 - \varepsilon)r) \quad , \quad U = \{x \in B(\mathbf{0}, r) : f(x) \in V\} = (f|_{B(\mathbf{0}, r)})^{-1}(V) .$$

y es obvio, por la construcción de U y V , que f lleva U biyectivamente a V y que ambos abiertos son entornos de $\mathbf{0}$. Tenemos, pues, una inversa $f^{-1} : V \rightarrow U$.

Como $f(x) = x + g(x)$ es coerciva con constante de coercividad $1 - \varepsilon$ en $B(x^0, r)$, la inversa f^{-1} es de Lipschitz con constante de Lipschitz $\frac{1}{1 - \varepsilon}$ y en particular $\|x\| \leq \frac{1}{1 - \varepsilon} \|y\|$ para $y \in V$ y $x = f^{-1}(y)$. Veamos que f^{-1} es diferenciable en $y = \mathbf{0}$ con $(d(f^{-1}))_{\mathbf{0}} = \text{id}_{\mathbb{V}}$. En primer lugar:

$$\begin{aligned} y \neq \mathbf{0} \implies 0 &\leq \frac{\|f^{-1}(y) - f^{-1}(\mathbf{0}) - \text{id}_{\mathbb{V}}(y)\|}{\|y\|} = \frac{\|f^{-1}(y) - \mathbf{0} - y\|}{\|y\|} = \frac{\|x - f(x)\|}{\|y\|} = \\ &= \frac{\| -g(x)\|}{\|x\|} \frac{\|x\|}{\|y\|} \leq \frac{\|g(x)\|}{\|x\|} \frac{1}{1 - \varepsilon} . \end{aligned}$$

En segundo lugar, cuando $y \rightarrow \mathbf{0}$ se tiene $x = f^{-1}(y) \rightarrow \mathbf{0}$ y $\frac{\|g(x)\|}{\|x\|} \rightarrow 0$ por (31), luego

$$\lim_{y \rightarrow \mathbf{0}} \frac{\|f^{-1}(y) - f^{-1}(\mathbf{0}) - \text{id}_{\mathbb{V}}(y)\|}{\|y\|} = 0 ,$$

y, por definición, la inversa f^{-1} es diferenciable en $y = \mathbf{0}$ con $(d(f^{-1}))_{\mathbf{0}} = \text{id}_{\mathbb{V}}$.

Caso general: x^0 e $y^0 = f(x^0)$ puntos cualesquiera, $L = (df)_{x^0}$ lineal acotada y biyectiva, con la inversa L^{-1} también acotada.

Para cualquier vector constante $\mathbf{c} \in \mathbb{R}^n$ definimos la **traslación** $T_{\mathbf{c}}(x) = x + \mathbf{c}$, aplicación afín cuya diferencial en todo punto es $\text{id}_{\mathbb{V}}$. Utilizamos esto para modificar f ligeramente, de manera que volvamos al caso cero. En primer lugar, la función

$$h \longmapsto \bar{f}(h) \equiv f(x^0 + h) - y^0 ,$$

es \mathcal{C}^1 y cumple que $\bar{f}(\mathbf{0}) = \mathbf{0}$. De $\bar{f} \equiv T_{-y^0} \circ f \circ T_{x^0}$ sale que $(d\bar{f})_{\mathbf{0}} = \text{id}_{\mathbb{V}} \circ L \circ \text{id}_{\mathbb{V}} = L$, luego \bar{f} todavía no está en la hipótesis del caso cero. Lo solucionamos definiendo $\tilde{f}(h) \equiv L^{-1} \circ \bar{f}(h)$, con lo cual $\tilde{f}(\mathbf{0}) = \mathbf{0}$ y $(d\tilde{f})_{\mathbf{0}} = \text{id}_{\mathbb{V}}$. Aplicando el caso cero a \tilde{f} , encontramos $r > 0$ y $0 < \varepsilon < 1$ tales que \tilde{f} es inyectiva en $B(\mathbf{0}, r)$ y la imagen inyectiva $\tilde{f}(B(\mathbf{0}, r))$ contiene a $B(\mathbf{0}, (1 - \varepsilon)r)$. Si ahora recuperamos f a partir de \tilde{f} : $f(x) \equiv y^0 + L \circ \tilde{f}(x - x^0)$, vemos de inmediato que f es inyectiva en $B(x^0, r)$ y que:

$$f(B(x^0, r)) \supset y^0 + L \left(B(\mathbf{0}, (1 - \varepsilon)r) \right) .$$

El conjunto $V_0 = L \left(B(\mathbf{0}, (1 - \varepsilon)r) \right)$ es la *preimagen* por L^{-1} de $B(\mathbf{0}, (1 - \varepsilon)r)$. Como suponemos que L^{-1} es acotada (es decir, continua), se sigue que V_0 es un abierto y, de hecho, es un entorno de $\mathbf{0}$ porque $\mathbf{0} \in B(\mathbf{0}, (1 - \varepsilon)r) \implies V_0 \ni L(\mathbf{0}) = \mathbf{0}$.

Por lo tanto $V \stackrel{\text{def}}{=} y_0 + V_0$ es un entorno de y_0 que, según acabamos de demostrar, está contenido en la imagen inyectiva $f(B(x^0, r))$. Definiendo ahora el siguiente abierto:

$$U = \{x \in B(x^0, r) : f(x) \in V\} = (f|_{B(x^0, r)})^{-1}(V) ,$$

tenemos que $x^0 \in U$ y que f es biyectiva de U a V .

Si la función inversa $f^{-1} : V \rightarrow U$ es diferenciable en y^0 , ya hemos comentado que su diferencial en y^0 sólo puede ser L^{-1} , debido a la regla de la cadena. Para demostrar tal diferenciabilidad, expresemos f^{-1} en términos de $\tilde{f}^{-1}(h)$ que, por el caso cero, sabemos es diferenciable en $h = \mathbf{0}$. A partir de $y = f(x) = y^0 + L \circ \tilde{f}(x - x^0)$, una sencilla manipulación nos da:

$$f^{-1}(y) = x = x^0 + \tilde{f}^{-1}(L^{-1}(y - y^0)),$$

identidad que puede expresarse de la manera siguiente:

$$f^{-1} \equiv T_{x^0} \circ \tilde{f}^{-1} \circ L^{-1} \circ T_{-y^0}.$$

Como además $(L^{-1} \circ T_{-y^0})(y^0) = \mathbf{0}$, punto en el que \tilde{f}^{-1} es diferenciable con diferencial $\text{id}_{\mathbb{V}}$, la regla de la cadena nos dice que f^{-1} es diferenciable en y^0 con la siguiente diferencial:

$$(d(f^{-1}))_{y^0} = \text{id}_{\mathbb{V}} \circ (d(\tilde{f}^{-1}))_{\mathbf{0}} \circ L^{-1} \circ \text{id}_{\mathbb{V}} = \text{id}_{\mathbb{V}} \circ \text{id}_{\mathbb{V}} \cdot L^{-1} \circ \text{id}_{\mathbb{V}} = L^{-1},$$

o sea f^{-1} es diferenciable en y^0 con $(d(f^{-1}))_{y^0} = [(df)_{x^0}]^{-1}$, tal como predice la regla de la cadena. \square

Bien entendido que en esta demostración ha quedado pendiente un detalle: cuando \mathbb{V} es un espacio de Banach de dimensión infinita, se necesita probar el análogo para \mathbb{V} de la proposición 81 del apartado 3.2.

La demostración, por otra parte, está completa cuando $\mathbb{V} = \mathbb{R}^n$. Es de destacar que es válida para *cualquier* norma que se utilice en \mathbb{R}^n . Ahora concretamos al caso de norma euclídea estándar $\|\cdot\|_2$ y vamos a escribir las fórmulas con la matriz jacobiana $A = Df_{x_0}$ en representación de la diferencial de f en x_0 . Entonces el abierto V_0 , que hemos definido durante la demostración, admite la descripción $V_0 = A \cdot B_{\text{estándar}}(\mathbf{0}, (1 - \varepsilon)r)$ y es un **elipsoide** abierto centrado en $\mathbf{0}$. Luego el abierto $V = y^0 + V_0$ es un elipsoide abierto centrado en y^0 . Así, la demostración que hemos hecho aquí proporciona un elipsoide abierto V en el que existe la inversa f^{-1} , la cual además toma todos sus valores dentro de la bola $B_{\text{estándar}}(x_0, r)$.

3.6 Inversas locales

Definición 87. Dado un abierto $U_0 \subseteq \mathbb{R}^n$, una función $f : U_0 \rightarrow \mathbb{R}^m$ es **regular en** $x^0 \in U_0$ si es de clase \mathcal{C}^1 en algún entorno de x^0 y la diferencial $(df)_{x^0} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tiene el máximo rango que puede tener, que es $\min(n, m)$. Decimos que f es **regular** si cumple esto en todo punto de U_0 .

En este apartado nos centramos en el caso $n = m$. Entonces f es regular si es \mathcal{C}^1 y todas sus jacobianas son matrices invertibles. En esa situación el teorema de la función inversa proporciona, para cada x del dominio de f , una bola $B(x, r)$, en la que f es inyectiva, y un elipsoide V centrado en $f(x)$ y totalmente cubierto por la imagen $f(B(x, r))$.

Definición 88. Sean X, Y espacios métricos. Una aplicación $f : X \rightarrow Y$ es **abierta** si para todo abierto $U \subseteq X$ la imagen directa $f(U)$ es un abierto de Y .

Una aplicación abierta puede no ser inyectiva, véase el apartado 3.8.

Proposición 89. Sean $U_0 \subseteq \mathbb{R}^n$ un abierto y $f : U_0 \rightarrow \mathbb{R}^n$ función de clase \mathcal{C}^1 . El conjunto

$$E = \{x^0 \in U_0 : f \text{ es regular en } x^0\},$$

es un abierto y $f|_E$ es abierta.

Demostración. Sea $x^0 \in E$. Existe un entorno $U_1 \subseteq U_0$ de x^0 en el cual f es de clase \mathcal{C}^1 . Como la función $U_1 \ni x \mapsto \det Df_x$ es continua, el conjunto $U_2 = \{x \in U_1 : \det Df_x \neq 0\}$ es un entorno de x^0 contenido en E . Esto indica que x_0 es interior a E y, como x^0 era cualquier punto de E , el conjunto E es un abierto (puede ser vacío).

Tomamos un subconjunto abierto $U \subseteq E$, la imagen directa $f(U)$ y un punto $y^0 \in f(U)$. Elegimos una preimagen, o sea un $a \in U$ tal que $y^0 = f(a)$. Como $a \in E$, es $\det Df_a \neq 0$ y el teorema de las funciones inversas proporciona un abierto $U^a \subseteq U$ y un elipsoide abierto V , centrado en y^0 , tales que f es biyectiva de U^a a V . Entonces

$$y^0 = f(a) \in V = f(U^a) \subseteq f(U),$$

o sea $y^0 \in V \subseteq f(U)$, luego el punto y^0 es *interior* a $f(U)$. Al ser todo punto $y^0 \in f(U)$ interior, la imagen $f(U)$ es un abierto. \square

Definición 90. Sean $U_0 \subseteq \mathbb{R}^n$ abierto y $f : U_0 \rightarrow \mathbb{R}^n$. Las **inversas locales de f** son las funciones $(f|_U)^{-1} : f(U) \rightarrow U$, donde $U \subseteq U_0$ es cualquier abierto tal que $f(U)$ es abierto y $f|_U$ es inyectiva.

Proposición 91. Sean $U_0 \subseteq \mathbb{R}^n$ un abierto y $f : U_0 \rightarrow \mathbb{R}^n$. Si f es regular, entonces las inversas locales de f heredan el grado de suavidad de f : si f es \mathcal{C}^1 entonces ellas son \mathcal{C}^1 , si f es \mathcal{C}^k entonces ellas son \mathcal{C}^k , si f es \mathcal{C}^∞ entonces ellas son \mathcal{C}^∞ .

Demostración. Sea $U \subseteq U_0$ cualquier abierto en el que f es regular e inyectiva. Por la proposición 89 sabemos que $V = f(U)$ es un abierto. Además f es biyectiva de U a V . Denotemos por $g : V \rightarrow U$ la correspondiente inversa local de f . Sabemos, por el teorema 86, que g es diferenciable en todo punto de V . A fortiori g es continua.

La fórmula (30) del apartado 3.5 da lugar a: $(Dg)_{f(x)} = [Df_x]^{-1}$ para todo $x \in U$, que también puede ponerse en la siguiente forma:

$$Dg_y = [(Df)_{g(y)}]^{-1}, \quad \text{para todo } y \in V. \quad (32)$$

Sea ahora $\text{inv} : \text{GL}(\mathbb{R}, n) \rightarrow \text{GL}(\mathbb{R}, n)$ la función dada por $A \mapsto A^{-1}$. En el apartado 2.7 del capítulo 2 vimos que $\text{inv} \in \mathcal{C}^\infty$. La fórmula (32) es lo mismo que la siguiente identidad entre funciones con valores matrices, es decir funciones $V \rightarrow M_{n \times n}(\mathbb{R})$:

$$Dg \equiv \text{inv} \circ Df \circ g. \quad (33)$$

Si f es \mathcal{C}^1 entonces Df es continua y los tres factores en el lado derecho de (33) son continuos, con lo cual la función Dg es continua y g es \mathcal{C}^1 .

Sea ahora f de clase \mathcal{C}^2 . En particular es \mathcal{C}^1 y ya hemos visto que g es también \mathcal{C}^1 . Pero entonces los tres factores en el lado derecho de (33) son \mathcal{C}^1 , luego la función Dg es \mathcal{C}^1 y por lo tanto g es \mathcal{C}^2 .

Este proceso continúa indefinidamente y prueba, para todo k , que si f es \mathcal{C}^k entonces g es \mathcal{C}^k . Si f es \mathcal{C}^∞ entonces es \mathcal{C}^k para todo k , con lo cual g es \mathcal{C}^k para todo k y por lo tanto es \mathcal{C}^∞ . \square

Atención. Insistimos en que la condición $\det Df \neq 0$ es imprescindible para que la suavidad de f la hereden sus inversas locales. Recordemos que $f(x) = x^3$ es \mathcal{C}^∞ pero $f^{-1}(y) = \sqrt[3]{y}$ ni siquiera es derivable en $y = 0 = f(0)$, por culpa de la anulación de $f'(0)$.

3.7 Teorema de la función inversa: espacio euclídeo

La demostración del teorema de las funciones inversas del apartado 3.5 proporciona números r, ε tales que f es inyectiva en $B(x^0, r)$ y en $V = y^0 + A \cdot B(\mathbf{0}, (1 - \varepsilon)r)$ está definida la inversa local de f con valores en $B(x^0, r)$. Para ello se definen $\tilde{f}(h) = A^{-1}(f(x^0 + h) - y^0)$ y $g(h) = \tilde{f}(h) - h$; entonces r, ε se escogen de modo que sea $\|h\| \leq r \implies \|Dg_h\| \leq \varepsilon$, que es equivalente a:

$$x \in \overline{B}(x^0, r) \implies \|A^{-1}(Df_x - A)\| \leq \varepsilon.$$

Si utilizamos la norma $\|\cdot\|_2$ para vectores de \mathbb{R}^n entonces $V = y^0 + A \cdot B(\mathbf{0}, (1 - \varepsilon)r)$ es un elipsoide, pero el cálculo de la norma de la matriz variable $A^{-1}(Df - A)$ es muy difícil porque involucra cálculo de autovalores. Si utilizamos la norma $\|\cdot\|_1$ o la $\|\cdot\|_\infty$ para vectores de \mathbb{R}^n el cálculo de $\|A^{-1}(Df - A)\|$ se hace llevadero, pero entonces el entorno V es un *poliedro* algunas de cuyas esquinas pueden ser muy afiladas: ganamos por un lado y perdemos por otro.

En este apartado damos un método alternativo para obtener una bola euclídea $B(x^0, r)$, centrada en x^0 y en la que f es inyectiva, y otra centrada en y^0 en la que está definida la inversa local. Además tendremos una descripción satisfactoria de toda la imagen $f(B(x^0, r))$.

Definición 92. Sea A una matriz real $n \times n$. Decimos que A es **coerciva** si existe una constante $\lambda > 0$ tal que:

$$\text{para todo } v \in \mathbb{R}^n \quad v^t A v \geq \lambda \|v\|_2^2. \quad (34)$$

Las constantes λ que cumplen (34) se llaman **constantes de coercividad de A** .

Este concepto tiene interés en el contexto actual debido al siguiente resultado.

Lema 93. Sean dados un abierto $U_0 \subseteq \mathbb{R}^n$, una función $f : U_0 \rightarrow \mathbb{R}^n$ de clase \mathcal{C}^1 y un subconjunto convexo $B \subset U_0$. Si existen una matriz ortogonal $P \in O(n)$ y una constante $\lambda > 0$ tales que para todo $z \in B$ la matriz $(Df)_z P$ es coerciva con constante de coercividad λ , entonces $f|_B$ es coerciva con esa misma constante de coercividad.

Demostración. Supongamos primero que $P = I_n$, es decir:

$$\text{para todo } z \in B \quad \text{y todo } v \in \mathbb{R}^n, \quad v^t Df_z v \geq \lambda \|v\|_2^2. \quad (35)$$

Dados $x, x' \in B$ con $x \neq x'$, el segmento $[x, x']$ está contenido en B y por lo tanto en U_0 , luego tenemos definida para $t \in [0, 1]$ la siguiente función escalar:

$$\varphi(t) = (x' - x) \cdot f(x + t(x' - x)) = (x' - x)^t f(x + t(x' - x)),$$

que es diferenciable con $\varphi'(t) = (x' - x)^t Df_{x+t(x'-x)}(x' - x)$. Existe un $\theta \in (0, 1)$ tal que $\varphi(1) - \varphi(0) = (1 - 0)\varphi'(\theta) = \varphi'(\theta)$, igualdad que se desarrolla en la siguiente:

$$(x' - x) \cdot (f(x') - f(x)) = (x' - x)^t Df_z (x' - x) \quad , \quad \text{con } z = x + \theta(x' - x) \in E,$$

que junto con (35) nos da:

$$(x' - x) \cdot (f(x') - f(x)) \geq \lambda \|x' - x\|_2^2.$$

Utilizando ahora la desigualdad de Cauchy-Schwarz, llegamos a:

$$\|x' - x\|_2 \|f(x') - f(x)\|_2 \geq (x' - x) \cdot (f(x') - f(x)) \geq \lambda \|x' - x\|_2^2,$$

y dividiendo por $\|x' - x\|_2$ deducimos $\|f(x') - f(x)\|_2 \geq \lambda \|x' - x\|_2$. Como esta última desigualdad se cumple de manera trivial en el caso $x = x'$, la función $f|_B$ es coerciva con λ como constante de coercividad.

Para el caso de P general consideramos el abierto $\tilde{U} = P^t \cdot U_0$, que contiene al convexo $\tilde{B} = P^t \cdot B$, y la función:

$$\tilde{f} : \tilde{U} \rightarrow \mathbb{R}^n \quad , \quad \tilde{f}(x) = f(Px),$$

que es diferenciable con $D\tilde{f}_{\tilde{z}} = (Df_{P\tilde{z}})P$ para todo $\tilde{z} \in \tilde{U}$. Para todo $\tilde{z} \in \tilde{B}$ se tiene $P\tilde{z} \in B$ y, por la hipótesis del lema, la matriz $D\tilde{f}_{\tilde{z}}$ es coerciva con constante de coercividad λ . Entonces, por lo ya demostrado, sabemos que $\tilde{f}|_{\tilde{B}}$ es coerciva con constante de coercividad λ . Dados $x, x' \in B$ tenemos:

$$\|\tilde{f}(P^t x) - \tilde{f}(P^t x')\|_2 \geq \lambda \|P^t x - P^t x'\|_2 = \lambda \|P^t(x' - x)\|_2 = \lambda \|x' - x\|_2.$$

Pero $\tilde{f}(P^t x) = f(x)$ y $\tilde{f}(P^t x') = f(x')$, luego $\|f(x') - f(x)\|_2 \geq \lambda \|x' - x\|_2$ como se quería demostrar. \square

Nota. Si A es coerciva entonces $Av = \mathbf{0} \implies v = \mathbf{0}$, luego una tal A es forzosamente invertible. Se sigue que una función f en las hipótesis del lema 93 es regular en cada punto $z \in B$.

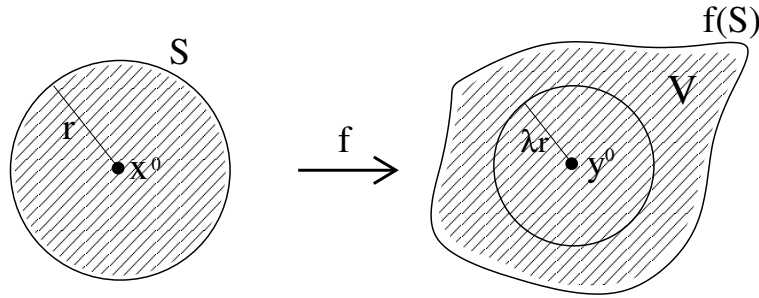
Teorema 94. Sean $U_0 \subseteq \mathbb{R}^n$ abierto y $f : U_0 \rightarrow \mathbb{R}^n$ función \mathcal{C}^1 . Dado un punto $x^0 \in U_0$, supongamos que existen una matriz ortogonal $P \in O(n)$, una bola $\overline{B}(x^0, r) \subset U_0$ y una constante $\lambda > 0$ tales que:

$$\text{para todo } x \in \overline{B}(x^0, r) \text{ y todo } v \in \mathbb{R}^n, \quad v^t (Df_x) P v \geq \lambda \|v\|_2^2. \quad (36)$$

Sean, además $y^0 = f(x^0)$ el punto imagen y $S = \{x : \|x - x^0\|_2 = r\}$ la esfera euclídea de centro x^0 y radio r . Entonces:

$$B(y^0, \lambda r) \subseteq f(B(x^0, r)) = V,$$

donde V es la componente conexa por caminos del abierto $\mathbb{R}^n \setminus f(S)$ que contiene al punto y^0 .



Demostración.

Paso preliminar. El conjunto $f(S)$ es compacto y por lo tanto cerrado, luego $\mathbb{R}^n \setminus f(S)$ es un abierto y sus componentes conexas por caminos son abiertos.

Aplicamos el lema 93 con $B = \overline{B}(x^0, r)$ y deducimos que $f|_B$ es coerciva con constante de coercividad λ . En particular, f es inyectiva en B .

La bola $B = \overline{B}(x^0, r)$ es unión *disjunta* de $B(x^0, r)$ y S . Al ser f inyectiva en B , las imágenes $f(B(x^0, r))$ y $f(S)$ son disjuntas. De manera equivalente:

$$f(B(x^0, r)) \subset \mathbb{R}^n \setminus f(S).$$

Como a bola $B(x^0, r)$ es conexa por caminos, resulta que $f(B(x^0, r))$ es un subconjunto conexo por caminos de $\mathbb{R}^n \setminus f(S)$ conteniendo a $y^0 = f(x^0)$. Por lo tanto, se tiene $y^0 \in f(B(x^0, r)) \subseteq V$.

Paso 1. La imagen $f(B(x^0, r))$ es un abierto de \mathbb{R}^n .

Como estamos haciendo una prueba nueva del teorema de la función inversa, no podemos hacer uso de la proposición 89, de la que hemos dada una demostración basada en dicho teorema.

Definición 95. La **distancia a un conjunto no vacío** E es la función $\text{dist}(\cdot, E)$ que se define de la manera siguiente:

$$\text{dist}(p, E) = \inf \{ \|p - q\|_2 : q \in E \}. \quad (37)$$

Es fácil ver que esta función es de Lipschitz con constante de Lipschitz 1, luego continua.

Sea $x \in B(x^0, r)$. Como $f(S)$ es compacto, el ínfimo que define $\text{dist}(f(x), f(S))$ es un mínimo y es positivo, porque $f(x) \notin f(S)$. Esto implica que el siguiente conjunto contiene al punto $f(x)$:

$$U^{f(x)} = \{ y \in \mathbb{R}^n : \|y - f(x)\|_2 < \text{dist}(y, f(S)) \}.$$

Como es evidente que $U^{f(x)}$ es un abierto, resulta que es un entorno de $f(x)$. Afirmamos que se cumple:

$$U^{f(x)} \subseteq f(B(x^0, r)), \quad (38)$$

con lo cual el punto $f(x)$ va a ser interior a $f(B(x^0, r))$. Esto demostrará que $f(B(x^0, r))$ es abierto, porque $f(x)$ es un punto arbitrario suyo.

Sea $y \in U^{f(x)}$. La fórmula (38), que estamos intentando demostrar, dice que una solución x' de la ecuación $y = f(x')$ se halla en la bola $B(x^0, r)$. Busquemos un x' que sea una **solución aproximada**, de manera que el “error” $\|y - f(x')\|_2$ sea lo menor posible. Con ese propósito, definimos la función:

$$\varphi_y : \overline{B}(x^0, r) \longrightarrow \mathbb{R} \quad , \quad \varphi_y(x') = \|y - f(x')\|_2 ,$$

que es continua y alcanza su mínimo en algún punto $a \in \overline{B}(x^0, r)$. Por definición de $U^{f(x)}$, el punto y está más cerca de $f(x)$ que de cualquier punto $f(x'')$ con $x'' \in S$, es decir $\varphi_y(x) < \varphi_y(x'')$ para todo $x'' \in S$, y por lo tanto el mínimo de φ_y *no se alcanza en la esfera S* . Aunque se necesitaba la compacidad de $\overline{B}(x^0, r)$ para asegurar que φ_y alcanza un mínimo, ha resultado que dicho mínimo *se alcanza en el interior*, o sea en la bola abierta:

$$a \in \overline{B}(x^0, r) \setminus S = B(x^0, r) .$$

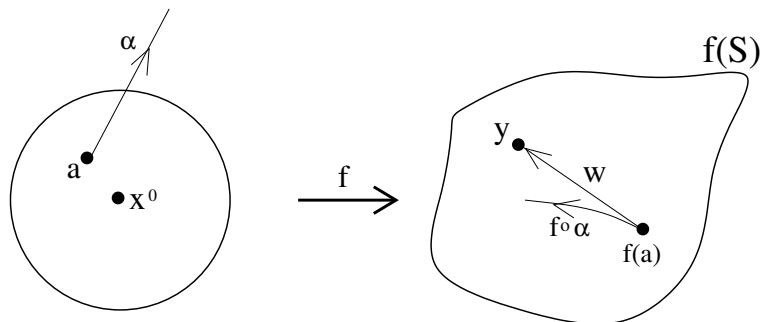
Ahora probaremos que dicho valor mínimo es nulo, es decir que $\|y - f(a)\|_2 = 0$ y la solución aproximada a es una solución de verdad: $f(a) = y$. Lo vamos a probar por reducción al absurdo: suponemos que $\|y - f(a)\|_2 > 0$ y deducimos una contradicción.

Suponemos, pues, que el vector $w = y - f(a)$ es distinto del $\mathbf{0}$. La matriz Df_a , siendo coerciva, es invertible. Existe, pues, un vector v tal que $Df_a v = w$ y definimos el camino $\alpha(t) = a + tv$. Consideremos la función $\varphi_y \circ \alpha(t) = \varphi_y(a + tv)$ y calculemos su derivada. Para ello recordemos que $g(\cdot) = \|\cdot\|_2$ es diferenciable en $\mathbb{R}^n \setminus \{\mathbf{0}\}$, con las siguientes derivadas:

$$\text{para cualesquiera } z, w \in \mathbb{R}^n \text{ con } z \neq \mathbf{0} \quad , \quad D_w g(z) = \frac{z \cdot w}{\|z\|_2} .$$

Entonces:

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} \varphi_y \circ \alpha(t) &= \left. \frac{d}{dt} \right|_{t=0} g(y - f \circ \alpha(t)) = \frac{[y - f(\alpha(0))] \cdot (-(f \circ \alpha)'(0))}{\|y - f(\alpha(0))\|_2} = \\ &= \frac{(y - f(a)) \cdot (-Df_a v)}{\|y - f(a)\|_2} = \frac{w \cdot (-w)}{\|w\|_2} = -\|w\|_2 < 0 , \end{aligned}$$



Gracias a que a está en la bola abierta, para t pequeño se cumple $\alpha(t) \in B(x^0, r)$ (esto podría ser falso si a hubiese estado en la esfera S) y además se tiene:

$$\varphi_y(\alpha(t)) < \varphi_y(\alpha(0)) = \varphi_y(a) .$$

que contradice que $\varphi_y(a)$ sea el mínimo de φ_y en $\overline{B}(x^0, r)$. Esta contradicción muestra que tiene que ser $w = \mathbf{0}$, es decir $f(a) = y$, lo que demuestra (38) y que $f(B(x^0, r))$ es un abierto.

Paso 2. La imagen $f(B(x^0, r))$ es un cerrado relativo de V .

Como $V \subset \mathbb{R}^n \setminus f(S)$, es $f(S) \cap V = \emptyset$ y tenemos las igualdades:

$$f(\overline{B}(x^0, r)) \cap V = [f(B(x^0, r)) \cap V] \cup (f(S) \cap V) = f(B(x^0, r)) \cup \emptyset = f(B(x^0, r))$$

que exhiben $f(B(x^0, r))$ como la intersección de V con el compacto $f(\overline{B}(x^0, r))$, luego como un cerrado relativo de V .

Paso 3. Se completa la demostración.

El conjunto $f(B(x^0, r))$ es no vacío y es a la vez abierto relativo y cerrado relativo de V . Como V es conexo por caminos por definición, tiene que ser $f(B(x^0, r)) = V$.

Como f es coerciva en $\overline{B}(x^0, r)$ con constante de coercividad λ , en particular:

$$x'' \in S \implies \|f(x'') - y^0\|_2 = \|f(x'') - f(x^0)\|_2 \geq \lambda \|x'' - x^0\|_2 = \lambda r,$$

por lo tanto $f(S) \cap B(y^0, \lambda r) = \emptyset$, que equivale a $B(y^0, \lambda r) \subset \mathbb{R}^n \setminus f(S)$. Al ser $B(y^0, \lambda r)$ un subconjunto conexo por caminos de $\mathbb{R}^n \setminus f(S)$ y conteniendo a y^0 , tiene ser parte de la componente conexa por caminos que contiene a y^0 , es decir

$$B(y^0, \lambda r) \subseteq V = f(B(x^0, r)).$$

□

Para validar el teorema 94 como una versión del teorema de la función inversa, falta probar que dada una matriz invertible A siempre hay matrices ortogonales P tales que AP es coerciva. Antes de demostrar eso en general veamos, con varios ejemplos, un procedimiento práctico para encontrar P y λ , primero para una matriz constante y luego para la matriz variable Df .

Toda matriz cuadrada A se descompone $A = M + N$, con $M = (1/2)(A + A^t)$ la **parte simétrica** y $N = (1/2)(A - A^t)$ la **parte antisimétrica**. Además $v^t N v = 0$ para todo $v \in \mathbb{R}^n$, luego se tiene la identidad $v^t A v \equiv v^t M v$ y la desigualdad (34) dice que la parte simétrica de A es definida positiva con autovalores acotados inferiormente por λ . Por lo tanto, el valor óptimo (el más grande) de λ que cumple (34) es el mínimo de los autovalores de la parte simétrica de A . En la búsqueda de λ y P que cumplan (36), vamos a seguir tres ideas:

1. Nos vamos a conformar con un valor no óptimo de λ , pero que se pueda hallar sin calcular autovalores.
2. Una condición suficiente para que A sea coerciva es que las entradas de su diagonal sean positivas y, en cierto modo, mayores que las otras entradas: que en la forma cuadrática $Q(v) = v^t A v$ los cuadrados “dominen” a los términos mixtos.
3. Si $\{e_{i_1}, \dots, e_{i_n}\}$ es una permutación de la base estándar, y si $P = [\pm e_{i_1} \mid \pm e_{i_2} \mid \dots \mid \pm e_{i_n}]$, entonces la operación $A \mapsto AP$ consiste en aplicar la misma permutación a las columnas de A y multiplicar algunas por -1 .

Ejemplo 1. Vamos a usar las dos primeras ideas con $A = \begin{bmatrix} 1 & 1 \\ 3 & 5 \end{bmatrix}$. Calculamos $v^t A v$, confrontando los cuadrados con el término mixto:

$$[x \ y] A \begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 5y^2 + 4xy \geq x^2 + 5y^2 - |4xy|.$$

Se saca más partido a la desigualdad de Cauchy-Schwarz $2|ab| \leq a^2 + b^2$ intercalando $c \frac{1}{c}$ como sigue:

$$2|ab| = 2 \left| (ca) \frac{b}{c} \right| \geq c^2 a^2 + \frac{b^2}{c^2},$$

y haciendo una elección inteligente de c (este truco es útil también con la desigualdad de Young). Para poder ver qué valor de c nos conviene en nuestro caso, calculamos:

$$x^2 + 5y^2 + 4xy \geq x^2 + 5y^2 - 2 \left| (cx) \frac{2y}{c} \right| \geq (1 - c^2) x^2 + \left(5 - \frac{4}{c^2} \right) y^2,$$

y vemos que hace falta $1 - c^2 > 0$, por lo cual probamos con $c = 0'9$:

$$x^2 + 5y^2 + 4xy \geq (1 - 0'9^2) x^2 + \left(5 - \frac{4}{0'9^2} \right) y^2 = 0'19 x^2 + 0'0617 y^2,$$

y el resultado tiene coeficientes positivos (si bien uno de ellos es bastante pequeño). Aceptamos la estimación $x^2 + 5y^2 + 4xy \geq 0'061(x^2 + y^2)$, es decir $v^t Av \geq 0'061\|v\|_2^2$. Esto proporciona la constante de coercividad $\lambda = 0'061$.

La parte simétrica de A es $\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$, con autovalor mínimo $3 - 2\sqrt{2} = 0'1716$ que, por supuesto, es más grande. Pero nos conformamos con el valor $0'061$ hallado sin autovalores.

Ejemplo 2. Vamos a usar la tercera idea con la matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$. Las entradas de la diagonal son pequeñas y las otras grandes. Intercambiamos las columnas multiplicando por $P = [\mathbf{e}_2 | \mathbf{e}_1]$:

$$AP = A[\mathbf{e}_2 | \mathbf{e}_1] = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix},$$

y estimamos:

$$v^t(AP)v = 2x^2 + 3y^2 + 2xy \geq 2x^2 + 3y^2 - 2|xy| \geq (2-1)x^2 + (3-1)y^2 = x^2 + 2y^2.$$

Llegamos a $v^t(AP)v \geq \|v\|_2^2$ y aceptamos la constante de coercividad $\lambda = 1$ para AP .

La matriz A , antes de multiplicar por P , no es coerciva porque su parte simétrica $\begin{bmatrix} 1 & 2'5 \\ 2'5 & 1 \end{bmatrix}$ es indefinida. Luego en este ejemplo es inevitable utilizar una matriz $P \neq I_2$.

Ejemplo 3. Para $A = \begin{bmatrix} 1 & 1 \\ -5 & 2 \end{bmatrix}$ probamos con $P = [\mathbf{e}_2 | -\mathbf{e}_1]$:

$$AP = \begin{bmatrix} 1 & 1 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 2 & 5 \end{bmatrix},$$

y estimamos:

$$v^t(AP)v = x^2 + 5y^2 + xy \geq \left(1 - \frac{1}{2}\right)x^2 + \left(5 - \frac{1}{2}\right)y^2 = 0'5x^2 + 4'5y^2,$$

y $\lambda = 0'5$ es una constante de coercividad para AP .

Ejemplo 4. Veamos ahora un ejemplo donde A es una jacobiana, o sea una matriz variable:

$$f \begin{pmatrix} x \\ y \end{pmatrix} \equiv \begin{pmatrix} 5 \operatorname{sen} x + e^{-y} \\ x + 3y + y^2 \end{pmatrix}, \quad x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad y^0 = f(x^0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Probemos con $r = 1/2$ y $P = I_2$. Calculamos y estimamos:

$$\begin{aligned} v^t(Df)v &= 5(\cos x)v_1^2 + (1 - e^{-y})v_1v_2 + (3 + 2y)v_2^2 \geq \\ &\geq \left[5\cos x - \frac{|1 - e^{-y}|}{2}\right]v_1^2 + \left[3 + 2y - \frac{|1 - e^{-y}|}{2}\right]v_2^2. \end{aligned}$$

Si $\|(x, y)\| \leq 1/2$ entonces $|x|, |y| \leq 1/2$. Ahora bien:

$$\begin{aligned} \min_{|x| \leq 1/2} 5 \cos x &= 5 \cos \frac{1}{2} = 4'38791 \dots, \\ \min_{|y| \leq 1/2} \frac{-|1 - e^{-y}|}{2} &= \frac{1 - e^{1/2}}{2} = -0'324 \dots, \\ \min_{|y| \leq 1/2} (3 + 2y) &= 3 - 1 = 2, \end{aligned}$$

luego:

$$\|(x, y)\| \leq 1/2 \implies v^t Df_{(x,y)} v \geq (4'38 - 0'33)v_1^2 + (2 - 0'33)v_2^2 = 4'05v_1^2 + 1'67v_2^2.$$

Tomando el menor de los dos coeficientes:

$$\|(x, y)\| \leq 1/2 \implies v^t Df_{(x,y)} v \geq 1'67 \|v\|_2^2.$$

Para $(x, y) \in B((0, 0), 1/2)$ la matriz $Df_{(x,y)}$ es coerciva con constante de coercividad 1'67. Dada la circunferencia $S = \{(x, y) : x^2 + y^2 = (1/2)^2\}$, el teorema 94 nos dice que f es inyectiva en $U = B((0, 0), 1/2)$ y que el dominio de definición de la inversa local $(f|_U)^{-1}$ (la componente conexa V de $\mathbb{R}^2 \setminus f(S)$ con $(1, 0) \in V$) contiene a la bola euclídea $B((1, 0), (1'67)(1/2)) = B((1, 0), 0'835)$.

Terminamos este apartado demostrando que para toda A invertible hay una matriz ortogonal P tal que AP es coerciva.

Teorema 96. (Descomposición polar). *Si A es una matriz real invertible entonces existen una matriz simétrica definida positiva S y una matriz ortogonal Q tales que $A = SQ$.*

Tomando $P = Q^t$ tenemos que $AP = S$ es simétrica definida positiva, por lo tanto coerciva.

Demostración. Sea $\{u_1, \dots, u_n\}$ una base ortonormal de autovectores de $A^t A$, con autovalores respectivos $\lambda_1, \dots, \lambda_n$. Tenemos $A^t A u_j = \lambda_j u_j$ para $j = 1, \dots, n$. Los vectores imagen $w_j = A u_j$ resultan ser ortogonales dos a dos:

$$i \neq j \implies (A u_i) \cdot (A u_j) = (A u_i)^t (A u_j) = u_i^t A^t A u_j = u_i^t (\lambda_j u_j) = \lambda_j (u_i \cdot u_j) = 0.$$

Nota. Esto es un poco milagroso: la transformación $v \mapsto Av$ no conserva el ángulo recto en general, pero hay ciertos vectores ortonormales u_1, \dots, u_n cuyas imágenes sí son ortogonales. De hecho los w_j son los semiejes principales del elipsoide $A \cdot B(\mathbf{0}, 1)$, imagen de la bola euclídea unidad por la transformación $v \mapsto Av$.

El cálculo análogo con $i = j$ nos da que $w_j \cdot w_j = \lambda_j$, por lo tanto $\|w_j\|_2 = \sqrt{\lambda_j}$, $j = 1, \dots, n$. Los vectores ortonormales $v_j = w_j / \|w_j\|_2 = w_j / \sqrt{\lambda_j}$ son tales que:

$$A u_j = \sqrt{\lambda_j} v_j \quad , \quad j = 1, \dots, n,$$

lo que matricialmente se expresa así:

$$A [u_1 | u_2 | \dots | u_n] = [v_1 | v_2 | \dots | v_n] \begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_n} \end{bmatrix},$$

es decir $AQ_2 = Q_1 D$ donde $Q_1, Q_2 \in O(n)$ y D es una matriz diagonal definida positiva. Despejamos:

$$A = Q_1 D Q_2^t = (Q_1 D Q_1^t) (Q_1 Q_2^t) = S Q,$$

donde $S = Q_1 D Q_1^t$ es simétrica definida positiva y $Q = Q_1 Q_2^t$ es ortogonal. \square

3.8 Consideraciones globales

Por el teorema de la función inversa, si f es regular en x^0 entonces es *localmente inyectiva*, es decir inyectiva en un entorno U de x^0 , y también tiene una suerte de *suprayectividad local*: la imagen $f(U)$ recubre todo un entorno V de $y^0 = f(x^0)$, es decir que $f(x^0)$ es un punto *interior* a $f(U)$. El concepto de *aplicación abierta* le da un significado preciso a la suprayectividad local. Es natural preguntarse cuánta inyectividad global y cuánta suprayectividad presenta una función regular, siendo parte de la cuestión el precisar el significado de esas dos propiedades. La respuesta a las dos preguntas resulta ser mucho más sencilla en dimensión 1 que en las otras dimensiones.

Inyectividad en dimensión=1. Sea $I \subseteq \mathbb{R}$ un **intervalo** y $f : I \rightarrow \mathbb{R}$ regular. Esto significa que f' no se anula nunca en I y, como es continua, tiene que ser positiva en todo I o negativa

en todo I . Luego f ha de ser estrictamente creciente o estrictamente decreciente. Esto hace que f sea inyectiva en todo el intervalo I , no importa cuán largo sea éste.

Suprayectividad en dimensión=1. Una función regular $f : I \rightarrow \mathbb{R}$ puede no ser suprayectiva. Un ejemplo es $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x/\sqrt{1+x^2}$, regular en todo \mathbb{R} pero con imagen $J = f(\mathbb{R}) = (-1, 1)$ acotada. La función está lejos de ser suprayectiva porque no tiene ningún valor $y \leq -1$ ni tampoco $y \geq 1$.

En vista de eso, nos hacemos la siguiente pregunta: dada $f : I \rightarrow \mathbb{R}$, función de clase \mathcal{C}^1 pero no necesariamente regular ¿cuáles son los intervalos más largos en los que está definida una inversa local diferenciable?

Fijamos un y^0 que admite una preimagen $x^0 \in I$ en la que f es regular, es decir $f(x^0) = y^0$ y $f'(x^0) \neq 0$. Definamos a como el mínimo valor tal que $f'(x) \neq 0$ para $x \in (a, x^0)$ (se admite aquí la posibilidad $a = -\infty$). De igual modo definamos b como el máximo valor tal que $f'(x) \neq 0$ para $x \in (x^0, b)$ (se admite la posibilidad $b = +\infty$). Entonces $(a, b) \subseteq I$ es un intervalo abierto con las siguientes propiedades:

- (1) contiene a x^0 ,
- (2) f es regular en él,
- (3) cualquier otro intervalo con las dos primeras propiedades es subconjunto suyo,

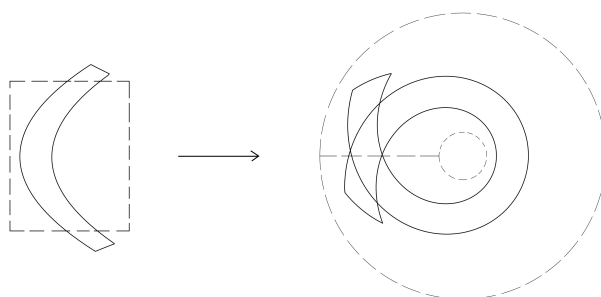
en este sentido es el **máximo** intervalo con las dos primeras propiedades. La imagen $J = f((a, b))$ es un intervalo abierto conteniendo a y^0 y es el máximo dominio, conteniendo a y^0 , de una inversa local diferenciable $f^{-1}(y)$ que lleve $y^0 \mapsto x^0$. Dicho de otra manera, tenemos $f^{-1}(y) : J \rightarrow (a, b)$ que es **“la más amplia inversa local diferenciable que lleva $y^0 \mapsto x^0$ ”**. Cualquier otra inversa local diferenciable que lleve $y^0 \mapsto x^0$ es una restricción de ésta.

Inyectividad en dimensión ≥ 2 . Una función regular f de un abierto de \mathbb{R}^n a \mathbb{R}^n puede no ser globalmente inyectiva porque **“tiene sitio para dar la vuelta”**. Es lo que hace, por ejemplo, la *función cambio a polares*:

$$\text{CP} : (0, +\infty) \times \mathbb{R} \longrightarrow \mathbb{R}^2 \quad , \quad \text{CP}(r, \theta) \equiv (r \cos \theta, r \sin \theta) ,$$

que es regular porque tiene $\det D(\text{CP}) = r$, positivo en todo punto de su dominio.

La siguiente figura muestra cómo la imagen por CP de una región en forma de letra C (y de altura mayor que 2π) “da la vuelta y se solapa consigo misma”

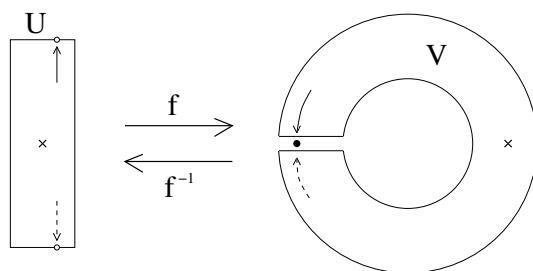


La función CP repite valores en cuanto se suma $\pm 2\pi$ a la variable θ , de modo que todo punto de $\mathbb{R}^2 \setminus \{(0, 0)\} = f(\mathbb{R}^2)$ tiene infinitas preimágenes por CP.

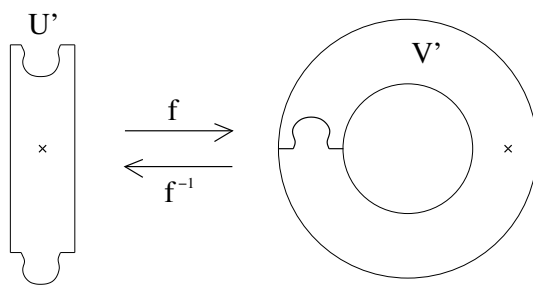
Obsevación. La función CP es abierta, por la proposición 89 del apartado 3.6. Es, pues, un ejemplo de aplicación abierta (de dimensión 2 a dimensión 2) que no es inyectiva.

Suprayectividad en dimensión ≥ 2 . Una función regular de un abierto de \mathbb{R}^n a \mathbb{R}^n no tiene por qué ser suprayectiva. Un ejemplo es $f(x, y) = (x, y)/\sqrt{1+x^2+y^2}$, regular con imagen acotada $f(\mathbb{R}^2) = B(\mathbf{0}, 1)$. En vista de esto, nos hacemos la misma pregunta que para dimensión 1: dado un abierto $U_0 \subseteq \mathbb{R}^n$ y $f : U_0 \rightarrow \mathbb{R}^n$ de clase \mathcal{C}^1 pero no necesariamente regular ¿cuáles son los dominios más grandes en los que está definida una inversa local diferenciable?

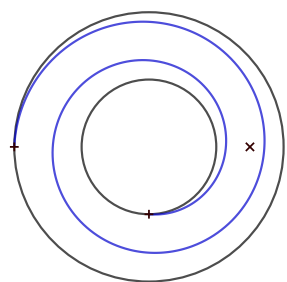
Más en concreto, planteamos esa cuestión para $f(r, \theta) = (r \cos(2\theta), r \sin(2\theta))$, $(r, \theta) \in (1, 2) \times \mathbb{R}$. El conjunto imagen de f es la corona circular $E = \{(x, y) : 1 < x^2 + y^2 < 4\}$. En esta corona tomamos el punto $y^0 = (3/2, 0)$ y fijamos la preimagen $x^0 = (3/2, 0)$. Como f es biyectiva del rectángulo abierto $U = (1, 2) \times (-\pi/2, \pi/2) \ni x^0$ al abierto $V = E \setminus (-2, -1) \times \{0\} \ni y^0$, hay una inversa local $V \rightarrow U$. La siguiente figura muestra la situación, con x^0, y^0 representados por aspás.



Se ve que es imposible extender f^{-1} a ningún conjunto mayor que V : si se la intenta extender a un punto $a \in (1, 2) \times \{0\}$ la extensión es discontinua en a , porque f^{-1} tiene límites distintos en a según nos acerquemos desde arriba (en la figura, flecha curvada continua) o desde abajo (flecha curvada a trazos). Esto significa que V , resultado de “hacerle un corte” a E a lo largo de un segmento, es un dominio **maximal** de definición de una inversa local. Pero hay infinitos otros dominios que contienen a y^0 y presentan esa misma maximalidad: los que resultan de hacerle un corte a E a lo largo de cualquier arco que vaya de la circunferencia interior a la exterior sin pasar por y^0 .



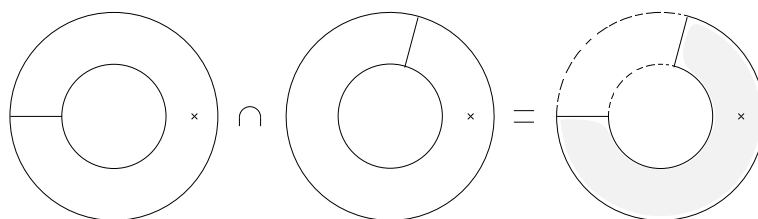
Conviene advertir, asimismo, que el arco que une las dos circunferencias puede dar muchas vueltas (lo único que se le exige es que esquive al punto y^0). En tal caso el dominio que resulta es “delgado y arrollado” y su imagen por la correspondiente inversa local puede ser muy alta (es decir, tener una proyección muy larga sobre el eje vertical, que es el de la variable θ).



Al contrario de lo que ocurría en dimensión=1, aquí no existe un dominio máximo de definición de inversas locales con $y^0 \mapsto x^0$:

los hay maximales pero distintos, de hecho una cantidad infinita de maximales.

Dados dos dominios maximales distintos $V_1, V_2 \ni y^0$, las correspondientes inversas locales son distintas pero coinciden en la componente conexa por caminos de $V_1 \cap V_2$ que contiene a y^0 . La siguiente figura muestra dicha componente conexa en línea continua y la otra componente, en la que las dos inversas locales difieren, en línea a trazos.



Este fenómeno es un caso particular del resultado que enunciaremos a continuación.

Proposición 97. Sea f regular y V un abierto **conexo por caminos** en el que están definidas dos inversas locales g_1, g_2 de f . Si $g_1(y^0) = g_2(y^0)$ para algún $y^0 \in V$, entonces $g_1 \equiv g_2$ en V .

Demostración. El conjunto $Y = \{y \in V : g_1(y) = g_2(y)\}$ es no vacío, pues al menos $y^0 \in Y$. Es, además, muy fácil ver que es un *cerrado relativo* de V .

Dado $y \in Y$, el punto $a = g_1(y) = g_2(y)$ es tal que $f(a) = y$. Existen entornos U^a de a y U^y de y tales que f es biyectiva de U^a a U^y . Entonces $U^y \cap V$ es un entorno de y en el que están definidas g_1 y g_2 . Como g_1, g_2 son ambas continuas (inversas locales de una función regular) y llevan $y \mapsto a$, existe una bola $B(y, r) \subseteq U^y \cap V$ tal que

$$g_1(B(y, r)) \subseteq U^a \quad \text{y} \quad g_2(B(y, r)) \subseteq U^a,$$

de donde para todo $z \in B(y, r)$ tenemos:

$$g_1(z) = (\text{el único } x \in U^a \text{ tal que } f(x) = z) = g_2(z),$$

es decir que $g_1(z) = g_2(z)$ para todo $z \in B(y, r)$. Esto significa que $B(y, r) \subseteq Y$, luego el punto y es interior a Y . Como y era cualquier punto de Y , resulta que Y es abierto.

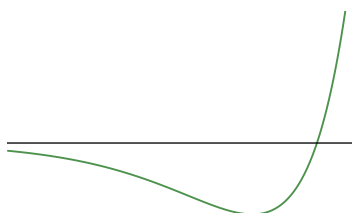
El conjunto Y es, pues, cerrado relativo y abierto relativo de V , además de no vacío. Como por hipótesis V es conexo por caminos, tenemos $Y = V$ que equivale a $g_1 \equiv g_2$ en V . \square

Para dar una inversa local lo más seguro es especificar un dominio U , en el que f sea inyectiva, y pedir que la inversa local tome sus valores en U , o sea definirla como $f|_U^{-1}$. Por ejemplo, la función elemental $\arcsen y$ es, para $-1 < y < 1$, la inversa local de $\sen x$ que llega a $(-\pi/2, \pi/2)$. La proposición 97 nos da un método alternativo de especificar una inversa local: dada f regular, dados un abierto *conexo por caminos* V y un punto $y^0 = f(x_0) \in V$, la inversa local g de f con $g(y^0) = x^0$ **o bien es única o bien no existe**. Así, para especificar una inversa local en un dominio conexo por caminos basta con especificar su valor en un punto concreto (y asegurarse de que existe).

El resto de este apartado se dedica a ejemplos de inversas locales maximales en dimensión=1.

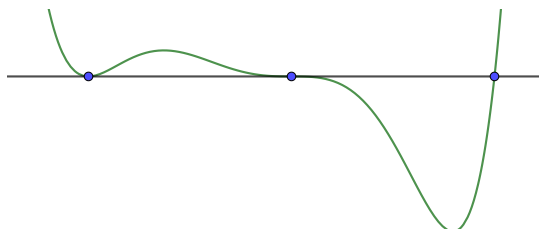
Ejemplo 1. $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^3$. Tiene dos inversas locales diferenciables maximales, una definida en $(0, +\infty)$ y la otra definida en $(-\infty, 0)$. Las demás inversas locales diferenciables son restricciones de una de esas dos. También hay una inversa global $f^{-1}(y)$, continua pero **no** diferenciable en $y = 0$.

Ejemplo 2. $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = xe^x$. Tiene un único mínimo en $x = -1$, siendo $f(-1) = -1/e$, $f'(x) < 0$ en $x < -1$ y $f'(x) > 0$ en $x > -1$. La siguiente figura muestra el grafo de f y el eje de abscisas.



Elegido un valor $y^0 \in (-1/e, 0)$, por ejemplo $y^0 = -1/4$, tiene dos preimágenes $x_1 < -1 < x_2$. La inversa local más amplia que lleva $y^0 \mapsto x_1$ está definida en $(-\frac{1}{e}, 0) = f(\{x < -1\})$; cualquier otra inversa local con $y^0 \mapsto x_1$ es restricción de ésta. La inversa local más amplia que lleva $y^0 \mapsto x_2$ está definida en $(-\frac{1}{e}, +\infty) = f(\{x > -1\})$, siendo restricción suya cualquier otra inversa local que lleve $y^0 \mapsto x_2$. No hay ninguna inversa local $f^{-1}(y)$ definida en un entorno de $y = -1/e$.

Ejemplo 3. Sea $f(x) = 7 + x^2(x - 1)^3(x - 2)$, un polinomio de grado 6. Las soluciones de $f(x) = 7$ son 0, 1, 2. La siguiente figura muestra el grafo $y = f(x)$, la recta horizontal $y = 7$ y los puntos $(1, 7)$, $(2, 7)$, $(3, 7)$.



Como $f'(0) = 0$ y $f''(0) > 0$, la función tiene un mínimo local estricto en $x = 0$. En entornos pequeños de $x = 0$ la función no es ni inyectiva ni suprayectiva a ningún entorno de $y = 7$; de hecho, la imagen por f de un entorno pequeño de $x = 0$ es $[7, 7 + \delta)$ con $\delta > 0$ pequeño. Por lo tanto **no existe** ninguna inversa local definida en un entorno de $y = 7$ y que lleve $7 \mapsto 0$.

Como $f'(1) = 0 = f''(1)$ y $f'''(1) < 0$, la función es estrictamente decreciente en un entorno U de $x = 1$, luego biyectiva de U a un entorno V de $y = 7$. Existe una inversa local $f^{-1}(y)$ definida para $y \in V$, caracterizada por llevar $7 \mapsto 1$, pero que **no es diferenciable en $y = 7$** .

Como $f'(2) > 0$ la función es regular en $x = 2$. Es biyectiva de un entorno de $x = 2$ a un entorno de $y = 7$ y la correspondiente inversa local (caracterizada por llevar $7 \mapsto 2$) es C^∞ .

Ejemplo 4. La función $f(x) = \sin x$ no es regular en todo \mathbb{R} pero lo es en el abierto E que resulta de quitarle a \mathbb{R} los ceros de la derivada $f'(x) = \cos x$, es decir:

$$E = \mathbb{R} \setminus \left(\frac{\pi}{2} + \pi \mathbb{Z} \right) = \dots \cup \left(-\frac{3\pi}{2}, -\frac{\pi}{2} \right) \cup \left(-\frac{\pi}{2}, \frac{\pi}{2} \right) \cup \left(\frac{\pi}{2}, \frac{3\pi}{2} \right) \cup \left(\frac{3\pi}{2}, \frac{5\pi}{2} \right) \cup \dots$$

En cada uno de estos intervalos la función $\sin x$ es, además de regular, biyectiva al intervalo $(-1, 1)$. Por lo tanto hay definidas en $(-1, 1)$ infinitas inversas locales diferenciables máximas:

$$f^{-1} : (-1, 1) \longrightarrow \left(\frac{(2s-1)\pi}{2}, \frac{(2s+1)\pi}{2} \right) \quad , \quad s \in \mathbb{Z} . \quad (39)$$

Estas inversas locales de $\sin x$ son funciones distintas dos a dos, porque sus intervalos de llegada son disjuntos dos a dos (además, unas son crecientes y otras decrecientes).

La función elemental $\arcsen y$ es, para $-1 < y < 1$, la única inversa local de $\sin x$ que lleva 0 a 0. Del mismo modo $\pi - \arcsen y$ puede definirse de dos maneras: como la inversa local de $\sin x$ que toma valores en $(\pi/2, 3\pi/2)$ o como la inversa local que lleva $0 \mapsto \pi$. Igual para las otras inversas locales de $\sin x$.

Ejemplo 6. Para la función $f(x) = e^x \cdot \sin(x - (\pi/4))$ calculamos:

$$\begin{aligned} f'(x) &= e^x \cdot \left[\sin\left(x - \frac{\pi}{4}\right) + \cos\left(x - \frac{\pi}{4}\right) \right] = \\ &= \sqrt{2} \cdot e^x \cdot \left[\sin\left(x - \frac{\pi}{4}\right) \cos \frac{\pi}{4} + \cos\left(x - \frac{\pi}{4}\right) \sin \frac{\pi}{4} \right] = \sqrt{2} \cdot e^x \cdot \sin x , \end{aligned}$$

luego f es regular en el siguiente abierto:

$$E = \{x \in \mathbb{R} : \sin x \neq 0\} = \mathbb{R} \setminus \pi \mathbb{Z} = \dots \cup (-2\pi, -\pi) \cup (-\pi, 0) \cup (0, \pi) \cup (\pi, 2\pi) \cup \dots$$

Al contrario que el ejemplo 5, las imágenes de esos intervalos son distintas entre sí:

$$f\left((s\pi, (s+1)\pi)\right) = \begin{cases} \left(-e^{s\pi}, e^{(s+1)\pi}\right) & \text{si } s \text{ es impar,} \\ \left(-e^{(s+1)\pi}, e^{s\pi}\right) & \text{si } s \text{ es par,} \end{cases}$$

de hecho son intervalos minúsculos cuando s es negativo y gigantes cuando s es positivo. El único punto común a todos esos intervalos imagen es $y = 0$. Cada valor $y \neq 0$ pertenece a una infinidad de esos intervalos, pero no a todos. La inversa local maximal $f^{-1}(y)$ que lleva $0 \mapsto \pi/4$ es la que toma sus valores en $(0, \pi)$ y está definida para $y \in (-e^\pi, 1)$. La inversa local maximal $f^{-1}(y)$ que lleva $0 \mapsto 5\pi/4$ es la que toma sus valores en $(\pi, 2\pi)$ y está definida para $y \in (-e^\pi, e^{2\pi})$. Al tomar valores en intervalos disjuntos $(0, \pi) \cap (\pi, 2\pi) = \emptyset$, esas dos inversas **no valen lo mismo** en ningún punto del dominio común $y \in (-e^\pi, 1) = (-e^\pi, 1) \cap (-e^\pi, e^{2\pi})$.

3.9 Funciones implícitas

Sean x, y variables escalares o vectoriales. Una función φ está dada *implícitamente* si se define su valor $y = \varphi(x)$ como solución de una ecuación o sistema de ecuaciones en las variables x, y . Si se cumplen unas condiciones de “bondad” de dicho sistema, esperamos **despejar tantas variables como ecuaciones tenga el sistema**.

Empezamos con tres ejemplos.

Primer ejemplo. La ecuación $e^x - y^3 = 0$ define implícitamente la función $y = e^{x/3}$. El conjunto de las soluciones $\{(x, y) : e^x - y^3 = 0\}$ es el grafo $\{(x, e^{x/3}) : x \in \mathbb{R}\}$ de la función implícita.

Segundo ejemplo. Esperamos que el sistema

$$\left. \begin{aligned} \operatorname{sen} x_1 + \frac{1}{3} x_4^2 - e^{x_3} x_5 &= -1 \\ 6x_1 x_2 x_3^2 + x_4 &= 0 \end{aligned} \right\}, \quad (40)$$

al tener dos ecuaciones, nos permita despejar dos variables. En efecto, podemos resolver (40) despejando las dos últimas variables:

$$\begin{aligned} x_4 &= -6x_1 x_2 x_3^2 && \equiv x_4(x_1, x_2, x_3), \\ x_5 &= e^{-x_3} (1 + \operatorname{sen} x_1 + 12x_1^2 x_2^2 x_3^4) && \equiv x_5(x_1, x_2, x_3), \end{aligned}$$

que quedan expresadas como funciones de las tres primeras, las cuales a su vez quedan libres (en este ejemplo) de recorrer todo \mathbb{R}^3 . Definiendo:

$$f(x_1, x_2, x_3, x_4, x_5) = \begin{bmatrix} \operatorname{sen} x_1 + \frac{1}{3} x_4^2 - e^{x_3} x_5 \\ 6x_1 x_2 x_3^2 + x_4 \end{bmatrix}, \quad b = \begin{bmatrix} -1 \\ 0 \end{bmatrix},$$

el sistema (40) equivale a la ecuación vectorial $f(x_1, x_2, x_3, x_4, x_5) = b$ y la solución que le hemos dado es:

$$f(x_1, x_2, x_3, x_4, x_5) = b \iff (x_4, x_5) = \varphi(x_1, x_2, x_3), \quad (41)$$

donde φ es la siguiente función vectorial:

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2, \quad \varphi(x_1, x_2, x_3) = \begin{bmatrix} x_4(x_1, x_2, x_3) \\ x_5(x_1, x_2, x_3) \end{bmatrix},$$

que es la *función vectorial implícita* definida por el sistema (40). La fórmula (41) dice que la *preimagen* $f^{-1}(\{b\}) = \{x \in \mathbb{R}^5 : f(x) = b\}$ coincide con el *grafo* de φ :

$$f^{-1}(\{b\}) = \{((\bar{x}, \varphi(\bar{x})) : \bar{x} \in \mathbb{R}^3)\}.$$

Tercer ejemplo. La ecuación $xe^x + y \operatorname{sen} y - z^2 = 2$ se puede resolver despejando, por ejemplo, la tercera variable. Salen dos opciones para el valor de z :

$$z_+(x, y) = \sqrt{xe^x + y \operatorname{sen} y - 2} \quad , \quad z_-(x, y) = -\sqrt{xe^x + y \operatorname{sen} y - 2} \quad ,$$

y ahora (x, y) no es libre de recorrer todo \mathbb{R}^2 , sino sólo el subconjunto:

$$C = \{(x, y) : xe^x + y \operatorname{sen} y \geq 2\} \quad .$$

Definiendo $f(x, y, z) = xe^x + y \operatorname{sen} y + z^2$, tenemos una preimagen que es *unión de dos grafos*:

$$f^{-1}(\{2\}) = \{(x, y, z_+(x, y)) : (x, y) \in C\} \cup \{(x, y, z_-(x, y)) : (x, y) \in C\} \quad .$$

El planteamiento general es el siguiente. Escribimos:

$$\mathbb{R}^{k+m} = \mathbb{R}^k \times \mathbb{R}^m = \{(x, y) : x = (x_1, \dots, x_k) \in \mathbb{R}^k, y = (y_1, \dots, y_m) \in \mathbb{R}^m\} \quad .$$

Partimos de un abierto $E_0 \subseteq \mathbb{R}^{k+m}$ y una función $f(x, y) : E_0 \rightarrow \mathbb{R}^m$ al menos de clase \mathcal{C}^1 . Dado un punto $a \in E_0$, tomamos $b = f(a)$ y queremos describir la parte de la preimagen $f^{-1}(\{b\})$ que se encuentra cerca de a , es decir el conjunto

$$E \cap f^{-1}(\{b\}) = \{x \in E : f(x) = b\} \quad ,$$

para cierto entorno E de a en \mathbb{R}^{k+m} . El teorema de las funciones implícitas da una condición suficiente (no necesaria) para que, eligiendo E adecuadamente, la parte $E \cap f^{-1}(\{b\})$ sea el grafo de una función diferenciable $\varphi : W \rightarrow \mathbb{R}^m$ definida en un abierto $W \subseteq \mathbb{R}^k$:

$$\{(x, y) \in E : f(x, y) = b\} = \{(x, \varphi(x)) : x \in W\} \quad .$$

Dos comentarios:

(1) Insistimos en que no se intenta describir la preimagen entera $f^{-1}(\{b\})$, sino solamente la parte de la misma que está cerca de a .

(2) El sistema $f(x, y) = b$, junto con la condición $(x, y) \in E$, equivale a $y = \varphi(x)$, $x \in W$. Es decir que *m ecuaciones escalares nos van a permitir despejar m variables (y_1, \dots, y_m) en función de las otras k variables (x_1, \dots, x_k) que quedan libres de recorrer un abierto $W \subseteq \mathbb{R}^k$.*

Volviendo con el tercer ejemplo, el punto $a = (1, 0, \sqrt{e-2})$ está en $f^{-1}(\{2\})$ y tiene $z > 0$. Para el entorno $E = \mathbb{R}^2 \times (0, +\infty)$ de a , la parte $E \cap f^{-1}(\{2\})$ involucra un solo grafo:

$$E \cap f^{-1}(\{2\}) = \{(x, y, z) : xe^x + y \operatorname{sen} y > 2, y = z_+(x, y)\} \quad ,$$

y además $z_+(x, y)$ es suave en esa parte, por ser la raíz cuadrada de una función suave y positiva.

Teorema 98. (Existencia de la función implícita). Sean un abierto $E_0 \subseteq \mathbb{R}^{k+m}$ y una función $f(x, y) : E_0 \rightarrow \mathbb{R}^m$ al menos de clase \mathcal{C}^1 . Si en un punto $a = (x^0, y^0) \in E_0$ los vectores $f_{y_1}(a), \dots, f_{y_m}(a)$ son linealmente independientes (y por lo tanto una base de \mathbb{R}^m) entonces hay entornos $x^0 \in W \subseteq \mathbb{R}^k$ e $y^0 \in U \subseteq \mathbb{R}^m$, con $W \times U \subseteq E_0$, y una función $\varphi : W \rightarrow \mathbb{R}^m$ tales que:

$$(W \times U) \cap f^{-1}(\{f(a)\}) \stackrel{\text{def}}{=} \{(x, y) : x \in W, y \in U, f(x, y) = f(a)\} = \{(x, \varphi(x)) : x \in W\} \quad ,$$

o bien, expresado con sistemas de condiciones:

$$\left. \begin{array}{l} f(x, y) = f(a) \\ x \in W \\ y \in U \end{array} \right\} \iff \left\{ \begin{array}{l} y = \varphi(x) \\ x \in W \end{array} \right.$$

Demostración. Para cada $x = (x_1, \dots, x_k)$ cercano a x^0 , tenemos la función $F_x(y) = f(x, y)$. Entonces $(F_x)_{x \text{ cercano a } x^0}$ es una familia de funciones con k parámetros x_1, \dots, x_k , para la

que nos planteamos encontrar una familia de inversas locales F_x^{-1} . La matriz jacobiana DF_x es la matriz $m \times m$ cuyas columnas son las derivadas parciales f_{y_1}, \dots, f_{y_m} .

La hipótesis del teorema es que $DF_{x^0}(y^0)$ es una matriz invertible. Por el teorema 96, existe una matriz ortogonal $P \in O(h)$ tal que $DF_{x^0}(y^0)P$ es coerciva con constante de coercividad $\lambda_0 > 0$.

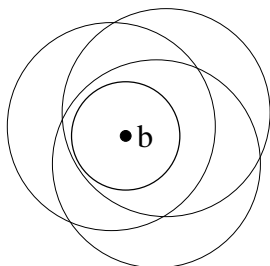
Ejercicio: tomando una constante un poco menor, por ejemplo $\lambda = \lambda_0/2$, se tiene que para A suficientemente cercana a $DF_{x^0}(y^0)$ la matriz AP es coerciva con constante de coercividad λ .

Por lo tanto existen radios $r_0, r_2 > 0$ tales que $B(x^0, r_0) \times \overline{B}(y^0, r_2) \subset E$ y si $x \in B(x^0, r_0)$ e $y \in \overline{B}(y^0, r_2)$ entonces $DF_x(y)P$ es coerciva con constante de coercividad λ .

Conseguido eso, el teorema 94 del apartado 3.7 nos dice que para todo $x \in B(x^0, r_0)$ la función $F_x(y)$ es inyectiva en $y \in \overline{B}(y^0, r_2)$ y además :

$$F_x(B(y^0, r_2)) \supseteq B(F_x(y^0), \lambda r_2) .$$

Las bolas $B(F_x(y^0), \lambda r_2)$ tienen el radio λr_2 fijo pero el centro $F_x(y^0)$ móvil. Si este centro se mueve poco entonces todas esas bolas contendrán una centrada en el punto $b = F_{x^0}(y^0) = f(a)$.



Elijamos un ε con $0 < \varepsilon < \lambda r_2$ y tomemos un radio $0 < r_1 < r_0$ tal que:

$$x \in B(x^0, r_1) \implies \varepsilon > \|F_x(y^0) - F_{x^0}(y^0)\|_2 = \|f(x, y^0) - b\|_2 ,$$

entonces un manejo fácil con la desigualdad triangular prueba que:

$$\overline{B}(b, \lambda r_2 - \varepsilon) \subseteq B(F_x(y^0), \lambda r_2) \quad \text{para todo } x \in B(x^0, r_1) .$$

Finalmente, para $x \in B(x^0, r_1)$ la función F_x es inyectiva en $B(y^0, r_2)$ y la imagen $F_x(B(y^0, r_2))$ contiene a la bola $\overline{B}(b, \lambda r_2 - \varepsilon)$.

Definiendo $W = B(x^0, r_1)$, $U = B(y^0, r_2)$ y $V = B(b, \lambda r_2 - \varepsilon)$, tenemos definida la familia de inversas locales:

$$(F_x|_U)^{-1} : V \longrightarrow U \subset \mathbb{R}^h \quad , \quad x \in W . \tag{42}$$

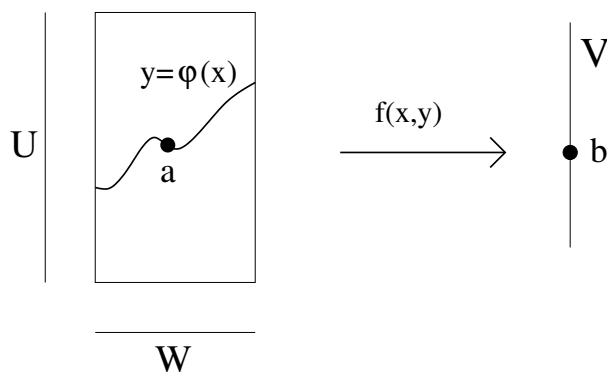
Haciendo actuar estas inversas locales sobre el punto $b \in V$, vemos que:

$$(x, y) \in (W \times U) \cap f^{-1}(\{b\}) \iff x \in W \quad , \quad y = (F_x|_U)^{-1}(b) . \tag{43}$$

Definimos, pues, la función:

$$\varphi : W \longrightarrow U \subset \mathbb{R}^h \quad , \quad \varphi(x) = (F_x|_U)^{-1}(b) , \tag{44}$$

y es obvio que $(W \times U) \cap f^{-1}(\{b\}) = \{(x, \varphi(x)) : x \in W\}$. □



Si sólo se quiere construir φ podemos tomar $\varepsilon = \lambda r_2$, pues esto basta para que las imágenes inyectivas $F_x(B(y^0, r_2))$ contengan $\{b\}$. Para que (43) y (44) tengan validez basta, pues, con elegir $r_1, r_2, \lambda > 0$ y $P \in O(h)$ tales que:

$$\left. \begin{array}{l} x \in B(x^0, r_1) \\ y \in \overline{B}(y^0, r_2) \end{array} \right\} \implies \left\{ \begin{array}{l} (DF_x)P \text{ es coerciva con constante de coercividad } \lambda \\ \|f(x, y^0) - b\|_2 < \lambda r_2 \end{array} \right.$$

Teorema 99. (Versión paramétrica del teorema de la función inversa). *En la situación del teorema anterior, si $f(x, y)$ es C^s , con $s \geq 1$, entonces las inversas locales (42) **dependen C^s del parámetro x** , es decir que la función conjunta $(x, z) \mapsto F_x^{-1}(z)$ está en $C^s(W \times V)$.*

Demostración. La función $(x, z) \mapsto F_x^{-1}(z)$ es de clase C^s si y sólo si lo es $(x, z) \mapsto (x, F_x^{-1}(z))$, que no es sino la inversa de la función $F(x, y) = (x, F_x(y))$. Establecemos la notación:

$$D_x f = [f_{x_1} | f_{x_2} | \cdots | f_{x_k}]_{m \times k}, \quad D_y f = [f_{y_1} | f_{y_2} | \cdots | f_{y_m}]_{m \times m},$$

y la jacobiana de F :

$$DF = \left[\begin{array}{c|c} I_k & 0_{k \times m} \\ \hline D_x f & D_y f \end{array} \right],$$

es triangular por cajas, luego invertible donde f_{y_1}, \dots, f_{y_m} sean linealmente independientes. Esto ocurre, en particular, cuando $(x, y) \in W \times U$. Entonces el teorema de la función inversa dice que

$$W \times V \ni (x, z) \mapsto (F|_{W \times U})^{-1}(x, z) = (x, (F_x|_U)^{-1}(z)),$$

es tan suave como F , que es C^s por serlo f . □

Corolario 100. *Si $f(x, y)$ es C^s , con $s \geq 1$, entonces la función $\varphi(x)$ definida en (44) es también C^s .*

La función vectorial implícita $\varphi(x)$ satisface la siguiente identidad (como funciones $W \rightarrow \mathbb{R}^m$):

$$f(x, \varphi(x)) \equiv b.$$

Tomando en ambos miembros de esta identidad derivadas parciales con respecto a x_1, \dots, x_k , se obtiene:

$$D_x f + (D_y f) D_x \varphi \equiv 0_{m \times k},$$

de donde es inmediato despejar $D_x \varphi$, porque $D_y f$ es invertible. Este procedimiento se conoce con el nombre de **derivación implícita**.

3.10 Dos ejemplos de estimación del dominio

En el primer ejemplo vamos a determinar una inversa local y a la vez estimar el dominio donde está definida. En el segundo ejemplo vamos a determinar una función (vectorial) implícita y a la vez estimar el dominio donde está definida.

Primer ejemplo: construcción de una inversa local.

Consideramos el abierto $U_0 = \{(x, y) \in \mathbb{R}^2 : y > 0\} = \mathbb{R} \times (0, +\infty)$ y la función $f : U_0 \rightarrow \mathbb{R}^2$ dada por:

$$f(x, y) = \begin{bmatrix} e^y + xy^2 \\ x^3 - x \log y \end{bmatrix}.$$

Dado el punto $a = (2, 1)$ y su imagen $b = f(a) = \begin{bmatrix} e + 2 \\ 8 \end{bmatrix}$, buscamos bolas $B(a, r_1) \subset U_0$ y $B(b, r_2)$ tales que f sea *regular e inyectiva* en $B(a, r_1)$ y la imagen $f(B(a, r_1))$ contenga a $B(b, r_2)$, con lo cual quedará determinada una inversa local

$$f^{-1} : B(b, r_2) \longrightarrow B(a, r_1) \subset \mathbb{R}^2,$$

de clase C^∞ y llevando $b \mapsto a$. El propósito es hallar valores explícitos para r_1 y r_2 . Calculamos:

$$Df = \begin{bmatrix} y^2 & e^y + 2xy \\ 3x^2 - \log y & -x/y \end{bmatrix}.$$

Buscamos una matriz $P \in O(2)$ y números $r, \lambda > 0$ tales que:

$$\text{para todo } (x, y) \in B(a, r) \text{ y todo } v \in \mathbb{R}^2 \text{ es } v^t (Df_{(x,y)} P) v \geq \lambda \|v\|_2^2. \quad (45)$$

Conseguido eso, tendremos que f es regular e inyectiva en $B(a, r)$ y que $f(B(a, r)) \supseteq B(b, \lambda r)$.

En vista de que $Df_a = \begin{bmatrix} 1 & e+4 \\ 12 & -2 \end{bmatrix}$ tiene bastante más grandes (y positivas) las entradas fuera de la diagonal, elegimos $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, con lo cual $(Df)P = \begin{bmatrix} e^y + 2xy & y^2 \\ -x/y & 3x^2 - \log y \end{bmatrix}$ es el resultado de intercambiar las columnas de Df , y calculamos:

$$\begin{aligned} v^t (Df) P v &= (e^y + 2xy) v_1^2 + (3x^2 - \log y) v_2^2 + \left(y^2 - \frac{x}{y}\right) v_1 v_2 \geq \\ &\geq (e^y + 2xy) v_1^2 + (3x^2 - \log y) v_2^2 - \left| \left(y^2 - \frac{x}{y}\right) v_1 v_2 \right| \geq \\ &\geq \left(e^y + 2xy - \frac{1}{2} \left| y^2 - \frac{x}{y} \right| \right) v_1^2 + \left(3x^2 - \log y - \frac{1}{2} \left| y^2 - \frac{x}{y} \right| \right) v_2^2. \end{aligned}$$

Probemos a ver si para $r = 1/2$ existe un $\lambda > 0$ cumpliendo (45). Si $(x, y) \in B(a, 1/2)$, entonces:

$$\frac{3}{2} < x < \frac{5}{2} \quad \text{y} \quad \frac{1}{2} < y < \frac{3}{2}.$$

Para $t \neq 0$ se tiene $e^t \geq t + 1$ y $\log t \leq t - 1$. Usando eso deducimos que si $(x, y) \in B(a, 1/2)$ entonces:

$$\begin{aligned} e^y + 2xy &> e^{1/2} + (2) \left(\frac{3}{2}\right) \frac{1}{2} > \frac{3}{2} + \frac{3}{2} = 3, \\ 3x^2 - \log y &> 3 \left(\frac{3}{2}\right)^2 - \log \frac{3}{2} > \frac{27}{4} - \frac{1}{2} = \frac{25}{4} > 6, \\ \frac{1}{2} \left| y^2 - \frac{x}{y} \right| &< \frac{1}{2} \left| \left(\frac{1}{2}\right)^2 - \frac{5/2}{1/2} \right| = \frac{1}{2} \left| \frac{1}{4} - 5 \right| = \frac{19}{8}, \end{aligned}$$

y para $(x, y) \in B(a, 1/2)$ se tiene $v^t (Df) P v \geq \left(3 - \frac{19}{8}\right) \|v\|_2^2 = \frac{5}{8} \|v\|_2^2$, es decir que se cumple (45) con $r = 1/2$ y $\lambda = 5/8$. Ahora sabemos que f es regular e inyectiva en $B((2, 1), 1/2)$ y que $f(B((2, 1), 1/2)) \supseteq B((e+2, 8), r_2)$ con $r_2 = \lambda r = 5/16 > 0'3$.

Conclusión. Hay una inversa local f^{-1} de clase C^∞ , definida en $V = B((e+2, 8), 0'3)$ y determinada de manera única por la condición de que toma todos sus valores dentro de $B((2, 1), 0'5)$. Como la bola V es *conexa por caminos*, es también la única inversa local que lleva $(e+2, 8) \mapsto (2, 1)$ y está definida en todo V .

Segundo ejemplo: construcción de una función implícita

Sea de nuevo $U_0 = \{(x, y) \in \mathbb{R}^2 : y > 0\}$ y consideremos una familia de funciones $F_p : U_0 \rightarrow \mathbb{R}^2$, dependiendo de un parámetro $p \in \mathbb{R}$, dadas por la siguiente fórmula:

$$F_p(x, y) = \begin{bmatrix} p e^y + x y^2 \\ p - 1 + x^3 - x \log y \end{bmatrix}.$$

También podemos considerarla como una función de tres variables $h(p, x, y) = F_p(x, y)$, pero vamos a usar la notación $F_p(x, y)$ cuando queramos exhibir la variable p como parámetro que,

al ir cambiando de valor, provoca una deformación de la función F_p . Esta es una **deformación suave**, en el sentido de que la función conjunta $h(p, x, y)$ está en $\mathcal{C}^\infty(\mathbb{R} \times U_0)$.

En particular F_1 es la función f del primer ejemplo. De nuevo consideramos los puntos $a = (2, 1)$ y $b = F_1(a) = \begin{bmatrix} e+2 \\ 8 \end{bmatrix}$ y establecemos el sistema $h(p, x, y) = b$, es decir

$$\left. \begin{aligned} p e^y + x y^2 &= e + 2 \\ p - 1 + x^3 - x \log y &= 8 \end{aligned} \right\} \quad (46)$$

cuyo conjunto de soluciones es la preimagen $h^{-1}(\{b\}) \subset \mathbb{R}_{pxy}^3$. Es un sistema de dos ecuaciones (escalares) con tres incógnitas p, x, y . Dada la solución particular $(p, x, y) = (1, a) = (1, 2, 1)$, queremos describir las soluciones cercanas a esta solución particular despejando x, y como funciones del parámetro p . Dicho con más precisión, buscamos una bola $U' = B((1, 2), r')$ y un intervalo $W = (1 - \delta, 1 + \delta)$ tales que las soluciones (p, x, y) de (46) con $p \in W$ y $(x, y) \in U'$ sean $(p, x, y) = (p, x(p), y(p))$, donde

$$\varphi(p) = (x(p), y(p)) : W \longrightarrow U' \subset \mathbb{R}^2,$$

será una función vectorial (implícita) que va a estar en $\mathcal{C}^\infty(W)$. Dicho de otra manera, las soluciones (p, x, y) de (46) con $(p, x, y) \in W \times U'$ van a formar el grafo de φ :

$$(W \times U') \cap h^{-1}(\{b\}) = \{(p, x(p), y(p)) : p \in W\} = \{(p, \varphi(p)) : p \in W\}.$$

El hecho de que $\varphi(p)$ tome sus valores en U' la va a determinar de manera única. En particular, estará obligada a cumplir $\varphi(1) = a$, es decir que el grafo de $\varphi(p)$ pasa por $(1, a)$.

Para cualesquiera $W \subseteq \mathbb{R}$ y $U' \subseteq U_0$ se cumple la siguiente igualdad conjuntista:

$$(W \times U') \cap h^{-1}(\{b\}) = \bigcup_{p \in W} \{p\} \times (F_p|_{U'})^{-1}(\{b\}).$$

Se trata, por lo tanto, de elegir W, U' de modo que para $p \in W$ la función F_p sea regular e inyectiva en U' y la imagen $F_p(U')$ contenga a b . Para cada $p \in W$, la ecuación $F_p(x, y) = b$ tendrá en U' una única solución que será el punto $\varphi(p)$.

Calculamos, para cada p :

$$DF_p = \begin{bmatrix} y^2 & p e^y + 2xy \\ 3x^2 - \log y & -x/y \end{bmatrix} = D_{x,y} h.$$

Elegimos, igual que en el primer ejemplo, la matriz ortogonal $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ y obtenemos:

$$v^t (DF_p) P v \geq \left(p e^y + 2xy - \frac{1}{2} \left| y^2 - \frac{x}{y} \right| \right) v_1^2 + \left(3x^2 - \log y - \frac{1}{2} \left| y^2 - \frac{x}{y} \right| \right) v_2^2.$$

Para $U' = B(a, 1/2)$, es decir $r' = 1/2$, buscamos un $W = (1 - \delta, 1 + \delta)$ adecuado a U' y, en todo caso, cumpliendo $\delta \leq 1/3$; en particular $p \geq 1 - 1/3 = 2/3$. Con esas limitaciones para p y (x, y) , tenemos:

$$p e^y + 2xy > \frac{2}{3} e^{1/2} + (2) \left(\frac{3}{2} \right) \frac{1}{2} > \frac{2}{3} \cdot \frac{3}{2} + \frac{3}{2} = \frac{5}{2},$$

mientras que las otras dos estimaciones son las mismas del apartado anterior:

$$3x^2 - \log y > 6, \quad \frac{1}{2} \left| y^2 - \frac{x}{y} \right| < \frac{19}{8},$$

con lo cual:

$$v^t (DF_p) P v \geq \left(\frac{5}{2} - \frac{19}{8} \right) v_1^2 + \left(6 - \frac{19}{8} \right) v_2^2 \geq \frac{1}{8} \|v\|_2^2.$$

Concluimos que para $p \geq 2/3$ todas las funciones F_p son coercivas en $B(a, 1/2)$, con constante de coercividad $\lambda = 1/8$, y cada imagen $F_p(B(a, 1/2))$ contiene la bola $B(F_p(a), r'_2)$ siendo $r'_2 = \lambda r' = 1/16 > 0'06$.

Las bolas $B(F_p(a), 0'06)$ tienen todas el mismo radio pero centro variable. Lo que necesitamos es que todas ellas contengan al punto b , lo cual equivale a lo siguiente:

$$\begin{aligned} 0'06 &> \|F_p(a) - b\|_2 = \left\| \begin{bmatrix} p e^1 + 2 \cdot 1^2 & - (e^1 + 2 \cdot 1^2) \\ p - 1 + 2^3 - 2 \cdot \log 1 & - (0 + 2^3 - 2 \cdot \log 1) \end{bmatrix} \right\|_2 = \\ &= \left\| \begin{bmatrix} (p-1)e \\ p-1 \end{bmatrix} \right\|_2 = |p-1| \sqrt{e^2 + 1}. \end{aligned}$$

Como definimos $W = (1 - \delta, 1 + \delta)$, es decir $W = B(1, \delta)$, debemos tomar $\delta \leq 0'06/\sqrt{e^2 + 1}$. Por ejemplo $\delta = 0'02$ y $W = (0'98, 1'02)$.

Conclusión. Para $0'98 < p < 1'02$ el sistema (46) tiene dentro de $B(a, 0'5)$ una única solución $(x, y) = \varphi(p)$. Esto define una función vectorial implícita

$$(x(p), y(p)) = \varphi(p) : (0'98, 1'02) \longrightarrow B(a, 0'5) \subset \mathbb{R}^2,$$

que es de clase \mathcal{C}^∞ porque la jacobiana $DF_p = D_{x,y}h$ permanece coerciva (luego invertible) para $p \geq 2/3$ y $(x, y) \in B(a, 1/2)$.