3 Funciones inversas e implícitas

3.1 Espacios completos

Definiciones 89. *Un espacio métrico* (X, d) *es* **completo** *si toda sucesión* $\{x_n\} \subset X$ *que sea de Cauchy para la distancia* d *es convergente para dicha distancia.*
Si un espacio vectorial normado $(\mathbb{V}, \|\cdot\|)$ *es completo para la distancia* $d(x, y) = \|x - y\|$, *decimos que es un* **espacio de Banach.**

Ejemplo 1. La recta real $\mathbb{R}$ con la distancia usual es completa.

Ejemplo 2. Todo espacio métrico compacto es completo.

Ejemplo 3. Los espacios normados $(\mathbb{R}^n, \|\cdot\|_\infty)$ son completos, luego son espacios de Banach. Veamos la demostración para $n = 2$, siendo igual para n mayor. Si $\big\{(x_n, y_n)\big\} \subset \mathbb{R}^2$ es una sucesión de Cauchy para la norma $\|\cdot\|_\infty$, entonces $\{x_n\}$ e $\{y_n\}$ son sucesiones numéricas de Cauchy, luego tienen límites:

$$\{x_n\} \to x_0 \quad , \quad \{y_n\} \to y_0 \; ,$$

y fácilmente se comprueba que $\big\{(x_n, y_n)\big\} \to (x_0, y_0)$ en la norma $\|\cdot\|_\infty$.

Ejemplo 4. Si $\|\cdot\|$ es cualquier norma en $\mathbb{R}^n$, entonces $(\mathbb{R}^n, \|\cdot\|)$ es completo.
En efecto, sea $\{v_k\}_{k=1}^\infty \subset \mathbb{R}^n$ una sucesión de Cauchy para la norma $\|\cdot\|$. Como esta norma es equivalente a la $\|\cdot\|_\infty$, resulta que $\{v_k\}$ también es de Cauchy para la norma $\|\cdot\|_\infty$. Luego existe $v_0 \in \mathbb{R}^n$ con $\{v_k\} \to v_0$ respecto de la norma $\|\cdot\|_\infty$. De nuevo por la equivalencia de normas, también es $\{v_k\} \to v_0$ respecto de la norma $\|\cdot\|$.

Ejemplo 5. Si (X, d) es un espacio métrico completo y $E \subseteq X$ es un subconjunto cerrado, entonces E es completo con la distancia inducida $d|_{E\times E}$.
En particular, *todo* cerrado $E \subseteq \mathbb{R}^n$ es completo para la distancia inducida en E por cualquier norma que tengamos en $\mathbb{R}^n$.

Ejemplos de lo contrario: cualquier subconjunto $X \subset \mathbb{R}^n$ que no sea cerrado es incompleto con la distancia inducida por una norma cualquiera en $\mathbb{R}^n$.

En esos ejemplos vemos que el concepto de espacio completo es estrictamente más amplio que el de espacio compacto: todo espacio métrico compacto es completo, pero hay muchos espacios completos que no son compactos.
Otra observación es que hay muchos espacios completos que no son espacios vectoriales. El concepto de espacio completo es, pues, mucho más amplio que el de espacio de Banach.

En un espacio completo tenemos la ventaja de poder decir si una sucesión tiene límite, sin necesidad de conocer explícitamente ese límite.

3.2 Aplicaciones contractivas y punto fijo

Provisionalmente, diremos que una aplicación $f : (X_1, d_1) \to (X_2, d_2)$, entre espacios métricos, es "Lipschitz-pequeña" si es una aplicación Lipschitz con constante de Lipschitz menor que 1, es decir, si existe ε tal que

$$\varepsilon < 1 \qquad \text{y} \qquad d_2\big(f(x)\,,\, f(x')\big) \le \varepsilon\, d_1(x, x') \quad \text{para cualesquiera } x, x' \in X_1 \; . \tag{32}$$

Definición 90. *Una* **aplicación contractiva** *es una aplicación* f *que es Lipschitz-pequeña y cuyos dominio y codominio coinciden, distancias incluidas:* $f : (X, d) \to (X, d)$.

Teorema 91. (Teorema de la aplicación contractiva). *Si* $f : (X, d) \to (X, d)$ *es contractiva y* (X, d) *es completo, entonces* f *tiene un único* **punto fijo,** *es decir que existe un único* $p \in X$ *tal que* $f(p) = p$.

Demostración.
Existencia. Empecemos con cualquier punto $x_0 \in X$ y consideremos la sucesión $\{x_n\}_{n=1}^{\infty}$ que se define de la manera siguiente:

$$x_1 \;=\; f(x_0) \quad , \quad x_2 \;=\; f(x_1) \quad , \quad x_3 \;=\; f(x_2) \quad , \ldots$$

es decir $x_{n+1} = f(x_n)$ para $n = 1, 2, 3, \ldots$ Esto equivale a:

$$x_1 \;=\; f(x_0) \quad , \quad x_2 = f \circ f\,(x_0) = f^2(x_0) \quad , \quad x_3 = f \circ f \circ f\,(x_0) = f^3(x_0) \quad , \ldots$$

y en general:

$$x_n \;=\; \underbrace{f \circ \cdots \circ f}_{n \text{ factores}}(x_0) \;=\; f^n(x_0)\,.$$

Observa lo importante que es, para poder definir las f^n, que el conjunto X de salida de f sea igual a su conjunto de llegada. Es decir, que el dominio y el codominio de f coincidan.

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

$$d(x_1, x_2) \;=\; d\big(\, f(x_0)\,,\, f(x_1)\,\big) \;\leq\; \varepsilon\, d(x_0, x_1)\,,$$

y de igual modo:

$$d(x_2, x_3) \;\leq\; \varepsilon\, d(x_1, x_2) \;\leq\; \varepsilon^2\, d(x_0, x_1) \quad , \quad d(x_3, x_4) \;\leq\; \varepsilon\, d(x_2, x_3) \;\leq\; \varepsilon^3\, d(x_0, x_1) \quad , \ldots$$

y en general:

$$d(x_n, x_{n+1}) \;\leq\; \varepsilon^n\, d(x_0, x_1) \quad , \quad \text{para todo } \; n \geq 1\,.$$

Observa lo importante que es, para obtener estas desigualdades, que en la desigualdad general

$$d\big(\, f(x)\,,\, f(y)\,\big) \leq \varepsilon\, d(x, y)\,,$$

la "*d*" en el término de la izquierda sea la misma que en el término de la derecha. Es decir, que la función distancia d en el espacio métrico de salida de f sea la misma que en el espacio de llegada.

Estimemos la distancia entre dos términos de la sucesión no necesariamente consecutivos:

$$\begin{aligned} d(x_n, x_{n+s}) \;&\leq\; d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{n+s-1}, x_{n+s}) \;\leq \\ &\leq\; \big(\varepsilon^n + \varepsilon^{n+1} + \cdots + \varepsilon^{n+s-1}\big)\, d(x_0, x_1) \;\leq\; d(x_0, x_1) \sum_{j=n}^{\infty} \varepsilon^j \;=\; d(x_0, x_1)\, \frac{\varepsilon^n}{1 - \varepsilon}\,. \end{aligned}$$

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

La aplicación f, siendo Lipschitz, es continua. Entonces de $\{x_n\} \to p$ se deduce

$$\big\{\, f(x_1)\,,\, f(x_2)\,,\, f(x_3)\,,\, \ldots\, \big\} \;\to\; f(p)\,,$$

que es lo mismo que:

$$\big\{\, x_2\,,\, x_3\,,\, x_4\,,\, \ldots\, \big\} \;\to\; f(p)\,.$$

La *cola* $\{\, x_2, x_3, x_4, \ldots\}$ de $\{x_n\}$ converge tanto a p como a $f(p)$. Por la unicidad del límite, deducimos que $p = f(p)$ y así p es un punto fijo para f.

Unicidad. Sea $x \in X$ cualquier punto fijo, es decir $f(x) = x$. Razonamos así:

$$d(p, x) \;=\; d\big(\, f(p)\,,\, f(x)\,\big) \;\leq\; \varepsilon\, d(p, x)\,,$$

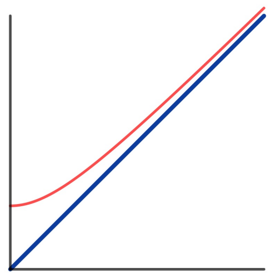
de donde $(1 - \varepsilon)\, d(p, x) \leq 0$ y, como es $1 - \varepsilon > 0$, se deduce $d(p, x) \leq 0$. Pero siempre es $d(p, x) \geq 0$, luego $d(p, x) = 0$ y, por lo tanto, es $p = x$. Esto prueba que p es el único punto fijo que tiene f. □

Es interesante que esta demostración es *constructiva:* el punto fijo se aproxima, a velocidad exponencial, por los puntos $x_n = f^n(x_0)$.
El teorema funciona en cualquier espacio métrico completo. Aquí lo vamos a aplicar al caso de una bola cerrada $X = \overline{B}(x_0, r) \subset \mathbb{R}^n$. En tal caso 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 en dicha bola.

Terminamos este apartado con un ejemplo en el que la constante de Lipschitz mínima es 1. El intervalo $X = [0, +\infty)$ es un cerrado de $\mathbb{R}$ y por lo tanto un espacio métrico completo con la distancia estándar. Definimos $f : X \to X$ por $f(x) = \sqrt{1 + x^2}$.
En lugar de la condición (32), esta función cumple la desigualdad más débil siguiente:

$$x \neq y \implies \big| f(x) - f(y) \big| < |x - y| \, ,$$

pero esta f no es Lipschitz-pequeña: no podemos bajar de la constante de Lipschitz 1 a una menor. Pues bien: resulta que f no tiene punto fijo, ya que la ecuación $x = \sqrt{1 + x^2}$ equivale a $x^2 = 1 + x^2$ y no tiene solución.
El grafo $\big\{ \big(x \, , \, f(x) \big) : x \in X \big\}$, es un "cuarto de hipérbola" y se acerca asintóticamente al grafo $\big\{ (\, x \, , \, x \,) : x \in X \big\}$ de la función identidad $\mathrm{id}_X : X \to X$, pero nunca llega a tocarlo.



3.3 Cómo obtener una constante de Lipschitz

El siguiente resultado es útil para obtener una constante de Lipschitz a partir de una cota de las derivadas parciales primeras. Nos va a permitir asegurar que ciertas funciones son Lipschitz-pequeñas.

Proposición 92. *Sea $U \subset \mathbb{R}^n$ un abierto en el que tenemos definida una función $f : U \to \mathbb{R}^m$ de clase $\mathcal{C}^1$. Sean $\| \cdot \|$ y $\| \cdot \|'$ normas cualesquiera en $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente. Supongamos que $E \subseteq U$ es un conjunto convexo y que existe una constante K tal que*

$$\big\| (Df)_x \big\| \leq K \quad , \quad \text{para todo } x \in E \, ,$$

donde $\big\| (Df)_x \big\|$ es la norma de $(Df)_x$ como operador $(\mathbb{R}^n, \| \cdot \|) \to (\mathbb{R}^m, \| \cdot \|')$. Entonces tenemos:

$$\| f(x) - f(y) \|' \leq K \, \| x - y \| \quad , \quad \text{para cualesquiera } x, y \in E \, ,$$

es decir que K es una constante de Lipschitz para $f : E \to \mathbb{R}^m$.

Demostración. Sean $x, y \in E$ y pongamos $h = y - x$. Al ser E convexo, todo el segmento rectilíneo $[x, y] = \big\{ x + th \; : \; t \in [0, 1] \big\}$ está contenido en E y tenemos la siguiente igualdad:

$$f(y) - f(x) = \int_0^1 \frac{d}{dt} f(x + th) \, dt = \int_0^1 \big(Df \big)_{x + th} h \, dt \, .$$

Es fácil probar la siguiente desigualdad, poniendo las dos integrales como límites de *sumas de Riemann*:

$$\left\| \int_{t_0}^{t_1} v(t) \, dt \right\|' \leq \int_{t_0}^{t_1} \| v(t) \|' \, dt \quad , \quad \text{para toda } v(t) : [t_0, t_1] \to \mathbb{R}^m \text{ continua.}$$

Por lo tanto:

$$\begin{aligned}\|f(y)-f(x)\|' &= \left\| \int_0^1 (Df)_{x+th}\, h\, dt, \right\|' \le \int_0^1 \left\| (Df)_{x+th}\, h \right\|' dt \le \\ &\le \int_0^1 \left\|(Df)_{x+th}\right\| \|h\|\, dt \le \|h\| \int_0^1 K\, dt = K\,\|y-x\| \,.\end{aligned}$$

□

3.4 Aplicaciones coercivas

Definición 93. *Una función entre espacios métricos* $f:(X,d_X)\to(Y,d_Y)$ *es* **coerciva** *si existe una constante* $\lambda>0$ *tal que* $d_Y\big(f(x'),f(x)\big)\ge \lambda\, d_X(x',x)$ *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:

$$x\neq x' \implies d_X(x,x')>0 \implies d_Y\big(f(x),f(x')\big) \ge \lambda\, d_X(x,x') > 0 \implies f(x)\neq f(x'),$$

pero, por ejemplo, la función $f:[0,1]\to[0,1]$ dada por $f(x)=x^2$ es una función inyectiva pero no coerciva.
Supongamos que $f:X\to Y$ 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"

$$f_{\rm im}: X\to f(X) \quad , \quad x\longmapsto f(x)\,,$$

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"

$$f_{\rm im}^{-1}: \big(\, f(X)\,,\, d_{f(X)}\big) \longrightarrow (X,d_X) \quad , \quad f(x) \longmapsto x\,,$$

es de Lipschitz con constante de Lipschitz $1/\lambda$.
La función $f:[0,1]\to[0,1]$ dada por $f(x)=x^2$ no es coerciva porque su inversa $y\longmapsto\sqrt{y}$ no es Lipschitz en $[0,1]$.

Dado un subconjunto $E\subseteq\mathbb{R}^n$, una *perturbación Lipschitz-pequeña de la identidad* es una función $f:E\to\mathbb{R}^n$ dada por $f(x)\equiv x+\varphi(x)$ donde $\varphi(x)$ es Lipschitz-pequeña. Tales funciones $f(x)$ son todas coercivas, porque la constante de coercividad 1 de la identidad "no se estropea del todo" al sumar la función Lipschitz-pequeña. Veamos el enunciado cuantitativo.

Lema 94. *Sea* $\|\cdot\|$ *una norma cualquiera en* $\mathbb{R}^n$ *y sea* $d(x,y)=\|x-y\|$.
Dado $E\subseteq\mathbb{R}^n$, *si* $\varphi(x):\big(E,d|_{E\times E}\big)\to(\mathbb{R}^n,d)$ *es Lipschitz con constante de Lipschitz* $\varepsilon<1$, *entonces* $f(x)\equiv x+\varphi(x)$ *es coerciva con constante de coercividad* $1-\varepsilon$.

Demostración. La siguiente propiedad de las normas es consecuencia fácil de la desigualdad triangular:

$$\|v+w\| \ge \|v\|-\|w\|\,,$$

y nos permite hacer el siguiente cálculo, para cualesquiera $x,x'\in E$:

$$\begin{aligned}\|\, f(x)-f(x')\,\| &= \|\, x+\varphi(x)-x'-\varphi(x')\,\| = \|\, x-x'+\varphi(x)-\varphi(x')\,\| \ge \\ &\ge \|x-x'\|-\big\|\varphi(x)-\varphi(x')\big\| \ge \|x-x'\|-\varepsilon\,\|x-x'\| = (1-\varepsilon)\,\|x-x'\|\,.\end{aligned}$$

□

Corolario 95. *En las hipótesis del lema 94, la función*

$$f_{\rm im}: E \longrightarrow f(E) \quad , \quad x \longmapsto f(x)\,,$$

es biyectiva y su inversa es Lipschitz con constante de Lipschitz $\dfrac{1}{1-\varepsilon}$.

Teorema 96. *Supongamos elegida una norma en* $\mathbb{R}^n$ *y sea* $g : \overline{B}(\mathbf{0}, r) \to \mathbb{R}^n$ *de Lipschitz con constante de Lipschitz* $\varepsilon < 1$ *y tal que* $g(\mathbf{0}) = \mathbf{0}$*. Entonces la función* $f : \overline{B}(\mathbf{0}, r) \to \mathbb{R}^n$ *dada por* $f(x) \equiv x + g(x)$*, además de ser coerciva con constante de coercividad* $1 - \varepsilon$*, satisface lo siguiente:*

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

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

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

es decir, queremos que la función $F_{y_0}(x) \stackrel{\text{def}}{=} -g(x) + y_0$ tenga un punto fijo $x \in B(\mathbf{0}, r)$. Es trivial ver que F_{y_0} es Lipschitz, con la misma constante de Lipschitz ε que la función $g(x)$. En vista del teorema 91, 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:

$$\text{Para } y_0 \text{ cercano a } \mathbf{0}\,, \quad \text{se cumple} \quad F_{y_0}\big(\overline{B}(\mathbf{0}, r)\big) \;\subseteq\; B(\mathbf{0}, r)\,, \tag{33}$$

y al final el punto fijo $x_0 = F_{y_0}(x_0)$ estará en la bola abierta $B(\mathbf{0}, r)$ y habremos conseguido un $x_0 \in B(\mathbf{0}, r)$ tal que $f(x_0) = y_0$.
Demostramos una versión cuantitativa de (33) mediante la siguiente estimación:

$$\begin{aligned} z \in \overline{B}(\mathbf{0}, r) \implies \|F_{y_0}(z)\| \;&=\; \|-g(z) + y_0\| \;=\; \|g(\mathbf{0}) - g(z) + y_0\| \;\le \\ &\le\; \|g(\mathbf{0}) - g(z)\| + \|y_0\| \;\le\; \varepsilon\, r + \|y_0\|\,, \end{aligned}$$

luego para $\|y_0\| < (1-\varepsilon)r$ está asegurado que $F_{y_0}\big(\overline{B}(\mathbf{0}, r)\big) \subseteq B(\mathbf{0}, r)$.

Por lo explicado, el sistema $f(x) = y_0$ tiene solución $x_0 \in B(\mathbf{0}, r)$ para todo $y_0 \in B\big(\mathbf{0}\,,\,(1-\varepsilon)r\big)$.
Es más, tenemos una fórmula para la solución: $x_0 \;=\; \lim\limits_{n\to\infty} \underbrace{F_{y_0} \circ \cdots \circ F_{y_0}}_{n \text{ factores}}(\mathbf{0})$. $\square$

Este teorema nos proporciona dos abiertos conteniendo a $\mathbf{0}$:

$$V \;=\; B\big(\mathbf{0}\,,\,(1-\varepsilon)\,r\big) \quad,\quad U \;=\; \big\{\, x \in B(\mathbf{0}, r) \;:\; f(x) \in V \,\big\}\,,$$

tales que $f : U \to V$ es biyectiva. Además la inversa $f^{-1} : V \to U$ es Lipschitz con constante de Lipschitz $\dfrac{1}{1-\varepsilon}$.

3.5 Función inversa

Vamos a dar una condición suficiente (no necesaria) para que una función de clase $\mathcal{C}^1$ sea biyectiva entre un entorno de un punto x_0 y un entorno del punto imagen $f(x_0)$.

3.5.1 Teorema de la función inversa

Enunciamos ya el teorema.

Teorema 97. (Teorema de la función inversa). *Sean* $W \subseteq \mathbb{R}^n$ *un abierto y* $f : W \to \mathbb{R}^n$ *función de clase* $\mathcal{C}^1$*. Si en un punto* $x_0 \in W$ *la matriz jacobiana* $(Df)_{x_0}$ *es invertible, 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 este caso la inversa* $f^{-1} : V \to U$ *es diferenciable en* y_0*.*

Varios comentarios antes de pasar a la demostración.

(1) 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.

(2) La condición que da el teorema es suficiente, pero no necesaria, para la existencia de una inversa. Pero sí es necesaria para que dicha inversa (si existe) sea diferenciable en y_0.
Supongamos que la inversa existe y es diferenciable en $y_0 = f(x_0)$, entonces es correcto aplicar la regla de la cadena a las identidades:

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

y deducir las siguientes igualdades:

$$I_n = \big(D(f^{-1})\big)_{y_0} (Df)_{x_0} = (Df)_{x_0} \big(D(f^{-1})\big)_{y_0} \, .$$

La conclusión es que, para que f^{-1} sea diferenciable en y_0, **es necesario que $(Df)_{x_0}$ sea invertible,** y en tal caso

$$\boxed{D(f^{-1})_{y_0} = (Df)_{x_0}^{-1}} \tag{34}$$

De este modo, la regla de la cadena predice un único valor posible para la jacobiana de f^{-1} en $y_0 = f(x_0)$, caso de que f^{-1} exista *y sea diferenciable en* y_0.
Si f^{-1} existe cerca de y_0 y $(Df)_{x_0}$ no es invertible, *entonces* f^{-1} *definitivamente no es diferenciable en* y_0. Por ejemplo $f(x) \equiv x^3$ tiene $f'(0) = 0$ y tiene inversa $f^{-1}(y) \equiv \sqrt[3]{y}$, pero f^{-1} no es diferenciable en $y = f(0) = 0$.

(3) De las dos funciones f y f^{-1}, una puede ser elemental y la otra no. Es lo que ocurre, por ejemplo, con $f(x) \equiv x + e^x$. Esta función, que es biyectiva $\mathbb{R} \to \mathbb{R}$ y con derivada siempre positiva, es elemental mientras que su inversa no lo es. El teorema da una condición suficiente para que haya una inversa, pero no proporciona ninguna fórmula elemental para dicha inversa.

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

$$f(x) = \begin{cases} x + x^2 \operatorname{sen} \dfrac{1}{x^2} & \text{si} \quad x \neq 0 \\[2ex] 0 & \text{si} \quad x = 0 \end{cases}$$

es $\mathcal{C}^\infty$ en $\mathbb{R} \setminus \{0\}$ y derivable en $x = 0$ con $f'(0) = 1$, 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 97.
Caso particular: $x_0 = y_0 = \mathbf{0}$ y $(Df)_{\mathbf{0}} = I_n$.
Escribimos f como una perturbación de la identidad:

$$f(x) \equiv x + g(x) \quad , \quad \text{con } g(x) \stackrel{\text{def}}{=} f(x) - x \, ,$$

entonces:

$$g \in \mathcal{C}^1(U_0) \quad , \quad g(\mathbf{0}) = \mathbf{0} \quad , \quad (Dg)_{\mathbf{0}} = 0_{n \times n} \, ,$$

en particular:

$$g(x) = \mathrm{o}(\|x\|) \, . \tag{35}$$

Fijamos una norma cualquiera $\| \cdot \|$ en $\mathbb{R}^n$ y escribimos $\big\| (Dg)_x \big\|$ para la norma de $(Dg)_x$ como operador $\big(\mathbb{R}^n, \| \cdot \|\big) \to \big(\mathbb{R}^n, \| \cdot \|\big)$. Fijamos también un ε con $0 < \varepsilon < 1$.
La función $x \longmapsto \big\| (Dg)_x \big\|$ es continua y nula en $x = \mathbf{0}$, luego existe un $r > 0$ tal que

$$\big\| (Dg)_x \big\| \leq \varepsilon \quad , \quad \text{para } x \in \overline{B}(\mathbf{0}, r) \, . \tag{36}$$

Esta desigualdad, junto con el hecho de que la bola $\overline{B}(\mathbf{0}, r)$ es un conjunto convexo, nos permite aplicar la proposición 92 y concluir que la función $g : \overline{B}(\mathbf{0}, r) \to \mathbb{R}^n$ es Lipschitz con constante de Lipschitz ε, o sea que es Lipschitz-pequeña. Ahora aplicamos el teorema 96 y concluimos que $f : \overline{B}(\mathbf{0}, r) \to \mathbb{R}^n$ es coerciva, con constante de coercividad $1-\varepsilon$, y además

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

Concluimos que $f : U \to V$ es biyectiva, siendo U, V los siguientes abiertos:

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

ambos entornos de $\mathbf{0}$.

Observa que el abierto V es una bola abierta para la norma $\|\cdot\|$.

También sabemos que la inversa $f^{-1}(y) : V \to U$ es Lipschitz con constante de Lipschitz $\dfrac{1}{1-\varepsilon}$.
Falta ver que $f^{-1}(y)$ es diferenciable en $y = \mathbf{0}$. Recordando que $f(x) = x + g(x)$, para cada punto $y \in V$ tenemos:

$$y = f\big(f^{-1}(y)\big) = f^{-1}(y) + g\left(f^{-1}(y)\right) ,$$

luego, por la fórmula (35):

$$f^{-1}(y) = y - g\left(f^{-1}(y)\right) = y + \mathrm{o}\big(\|f^{-1}(y)\|\big) , \tag{37}$$

pero $f^{-1}(y)$ es lipschitz, luego $\|f^{-1}(y)\| = \|f^{-1}(y) - f^{-1}(\mathbf{0})\| = \mathrm{O}(\|y - \mathbf{0}\|) = \mathrm{O}(\|y\|)$, lo cual convierte (37) en:

$$f^{-1}(y) = y + \mathrm{o}(\|y\|) ,$$

que significa que $f^{-1}(y)$ es diferenciable en $y = \mathbf{0}$ con $\big(d(f^{-1})\big)_{\mathbf{0}} = I_n$.

Caso general: x_0 e $y_0 = f(x_0)$ cualesquiera, $(Df)_{x^0} = A$ una matriz invertible $n \times n$.
Definimos $\widetilde{f} = A^{-1}\left(f(x + x_0) - y_0\right)$, que también es $\mathcal{C}^1$ con:

$$\widetilde{f}(\mathbf{0}) = \mathbf{0} \quad , \quad (D\widetilde{f})_{\mathbf{0}} = I_n .$$

Además, recuperamos la función f a partir de $\widetilde{f}$ mediante la fórmula:

$$f(x) = y_0 + A\,\widetilde{f}(x - x_0) . \tag{38}$$

Aplicamos a $\widetilde{f}$ el caso particular y obtenemos abiertos U_0, V_0, ambos entornos de $\mathbf{0}$, tales que $\widetilde{f} : U_0 \to V_0$ es biyectiva, la inversa $\widetilde{f}^{-1} : V_0 \to U_0$ es diferenciable en $\mathbf{0}$ y $\big(D(\widetilde{f}^{-1})\big)_{\mathbf{0}} = I_n$.
Entonces f, dada por la fórmula (38), es inyectiva en $x_0 + B(\mathbf{0}, r) = B(x_0, r)$ y lleva biyectivamente el entorno $U = x_0 + U_0 \subseteq B(x_0, r)$ de x_0 al entorno $V = y_0 + A\,V_0$ de y_0.

Observa que V es la imagen de una bola V_0 por un automorfismo afín de $\mathbb{R}^n$. En el caso particular en que $\|\cdot\|$ sea una norma euclídea, el abierto V es el interior de un *elipsoide sólido* centrado en y_0. En cualquier caso, el abierto V es *convexo.*

La fórmula (38) nos permite hacer la siguiente manipulación:

$$\begin{aligned} y &= f(x) = y_0 + A\,\widetilde{f}(x - x_0) , \\ y - y_0 &= A\,\widetilde{f}(x - x_0) , \\ x - x_0 &= A^{-1}\,\widetilde{f}^{-1}(y - y_0) , \\ f^{-1}(y) &= x = x_0 + A^{-1}\,\widetilde{f}^{-1}(y - y_0) , \end{aligned}$$

de donde se deduce que $f^{-1}(y)$ es diferenciable en $y = y_0$, con

$$\big(D(f^{-1})\big)_{y_0} = A^{-1}\left(D(\widetilde{f}^{-1})\right)_{\mathbf{0}} = A^{-1}\,I_n = A^{-1} ,$$

que es el valor que habíamos predicho en la fórmula (34) para la jacobiana de la función inversa. □

Si la norma utilizada $\|\cdot\|$ es euclídea, entonces V es un elipsoide abierto y en él existe una única inversa f^{-1} que tome todos sus valores dentro de la bola $B(x_0, r)$.

3.5.2 Funciones regulares e inversas locales

Definición 98. *Dado un abierto $U_0 \subseteq \mathbb{R}^n$, una función $f : U_0 \to \mathbb{R}^m$ es* **regular en $x^0 \in U_0$** *si es de clase $\mathcal{C}^1$ en algún entorno de x^0 y la diferencial $(df)_{x^0} : \mathbb{R}^n \to \mathbb{R}^m$ tiene el máximo rango que puede tener, es decir:*

$$\text{rango}\big((df)_{x_0}\big) \;=\; \min(n, m)\;.$$

Decimos que f es **regular en U_0** *si cumple esto en todo punto $x_0 \in U_0$.*

En este apartado nos centramos en el caso $n = m$. Entonces f es regular en U_0 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 $B = B(x, r)$ y un entorno V de $f(x)$ tales que $f|_B$ es inyectiva y V está totalmente cubierto por la imagen $f(B)$.

Teorema-definición 99. *Una* **aplicación abierta** *es una $h : (X, d) \to (Y, d')$ tal que para todo abierto U de (X, d) la imagen $h(U)$ es un abierto de (Y, d').*
Toda función $f \in \mathcal{C}^1$, de un abierto $U_0 \subseteq \mathbb{R}^n$ a $\mathbb{R}^n$, es abierta en cada abierto en el que es regular. Llamamos **inversas locales de f** *a las funciones $(f|_U)^{-1} : f(U) \to U$, donde $U \subseteq U_0$ es cualquier abierto en el que f es regular e inyectiva.*

Lo único que hay que demostrar es que, si f es regular en un abierto $U \subseteq U_0$, entonces $f(U)$ es un abierto. Dado $y_0 \in f(U)$, elegimos un $x_0 \in U$ con $y_0 = f(x_0)$. Existen una bola $B(x_0, r) \subseteq U$ y un elipsoide abierto V con $y_0 \in V \subseteq f\big(B(x_0, r)\big) \subseteq f(U)$, luego y_0 es interior a $f(U)$. Como y_0 era un punto cualquiera de $f(U)$, vemos que $f(U)$ es abierto.

Proposición 100. *Si f es regular e inyectiva en un abierto U, entonces la correspondiente inversa local $g = (f|_U)^{-1}$ hereda el grado de suavidad de f: si f es $\mathcal{C}^1$, entonces g es $\mathcal{C}^1$; si f es $\mathcal{C}^k$, entonces g es $\mathcal{C}^k$; si f es $\mathcal{C}^\infty$, entonces g es $\mathcal{C}^\infty$.*

Demostración. Sea $U \subseteq U_0$ cualquier abierto en el que f es regular e inyectiva. Sabemos que $V = f(U)$ es un abierto. Además f es biyectiva de U a V. Denotemos por $g : V \to U$ la correspondiente inversa local de f. Aplicando el teorema 97 a cada punto $x_0 \in U$, se deduce que g es localmente Lipschitz y diferenciable en todo punto de V.
La fórmula (34) del apartado 3.5.1 da lugar a:

$$(Dg)_{f(x)} \;=\; \big[(Df)_x\big]^{-1} \qquad \text{para todo } x \in U\;,$$

que también puede ponerse en la siguiente forma:

$$Dg_y \;=\; \big[(Df)_{g(y)}\big]^{-1} \quad , \quad \text{para todo } y \in V\;. \tag{39}$$

Sea ahora $\text{inv} : \text{GL}(\mathbb{R}, n) \to \text{GL}(\mathbb{R}, n)$ la función dada por $A \mapsto A^{-1}$. En el apartado 2.2.2 del capítulo 2 vimos que $\text{inv} \in \mathcal{C}^\infty$. La fórmula (39) es lo mismo que la siguiente identidad entre funciones con valores matrices, es decir funciones $V \to M_{n\times n}(\mathbb{R})$:

$$Dg \;\equiv\; \text{inv} \circ Df \circ g\;. \tag{40}$$

Si f es $\mathcal{C}^1$, entonces Df es continua y los tres factores en el lado derecho de (40) 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 (40) 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 $\mathcal{C}^\infty$, en particular es $\mathcal{C}^k$ para todo k. Entonces g es $\mathcal{C}^k$ para todo k, luego es $\mathcal{C}^\infty$. □

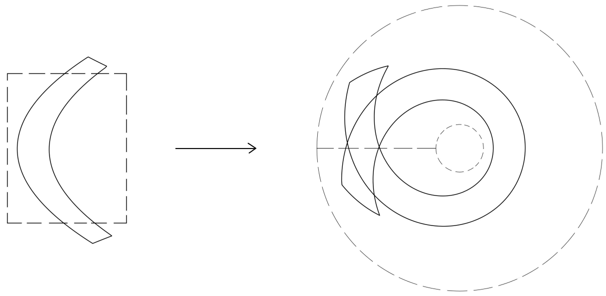
Atención. Insistimos en que la condición $\det Df \neq 0$ es imprescindible para que la suavidad de f la hereden sus inversas locales. Recordemos que $f(x) = x^3$ es $\mathcal{C}^\infty$ pero $f^{-1}(y) = \sqrt[3]{y}$ ni siquiera es derivable en $y = 0 = f(0)$, por culpa de la anulación de $f'(0)$.

Una función regular es **localmente inyectiva:** inyectiva en abiertos suficientemente pequeños. ¿Es inyectiva en todo su dominio?
Para funciones de una variable $f(x)$, definidas en un *intervalo* $I \subseteq \mathbb{R}$, la regularidad implica inyectividad en todo I porque la función derivada $f'(x)$ se mantiene positiva en todo I o negativa en todo I. La función f es estrictamente creciente en el primer caso y estrictamente decreciente en el segundo.

Funciones regulares de dos o más variables **tienen sitio para dar la vuelta** y pueden no ser globalmente inyectivas. Una función que "da la vuelta" es, por ejemplo, el *cambio a polares:*

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

que es regular porque tiene $\det D(\mathrm{CP}) \equiv r$, positivo en todo punto de su dominio. La siguiente figura muestra una región en forma de letra C (y de altura mayor que 2π) y cómo su imagen por CP "da la vuelta y se solapa consigo misma"



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

Observación. La función CP es abierta, por ser regular. Es, pues, un ejemplo de aplicación abierta (de dimensión 2 a dimensión 2) que no es inyectiva.

Una función regular de un abierto de $\mathbb{R}^n$ a $\mathbb{R}^n$ **no tiene por qué ser suprayectiva,** ni siquiera en el caso $n = 1$. Como primer ejemplo $f(x) = e^x$ es regular en todo $\mathbb{R}$, pero su imagen $f(\mathbb{R}) = (0, +\infty)$ no es todo $\mathbb{R}$. Un segundo ejemplo es $h(x,y) = (x,y)/\sqrt{1 + x^2 + y^2}$, regular en todo el plano pero con imagen acotada $h(\mathbb{R}^2) = B(\mathbf{0}, 1)$.

3.5.3 Cómo especificar una inversa local

Dado un abierto U, en el que f es inyectiva, la inversa local $(f|_U)^{-1}$ puede definirse, de manera unívoca, como **la que está definida en $\mathbf{f(U)}$ y toma sus valores en $\mathbf{U}$.**
Por ejemplo $g : (-1,1) \to \mathbb{R}$, dada por $g(y) \equiv \mathrm{arcsen}\, y$, es la inversa local de $f(x) \equiv \mathrm{sen}\, x$ definida para $y \in (-1,1) = f\big((\, -\pi/2 \, , \, \pi/2 \,)\big) = f\big((\, 3\pi/2 \, , \, 5\pi/2 \,)\big)$ y tomando todos sus valores en $U = (\, -\pi/2 \, , \, \pi/2 \,)$, o sea $g \equiv \big(f|_U\big)^{-1}$. En cambio $h : (-1,1) \to \mathbb{R}$, dada por $h(y) = 2\pi + \mathrm{arcsen}\, y$, es la inversa local de f definida también para $y \in (-1,1)$ pero tomando sus valores en $U' = \big(\, (3/2)\pi \, , \, (5/2)\pi \,\big)$, o sea $h \equiv \big(f|_{U'}\big)^{-1}$.

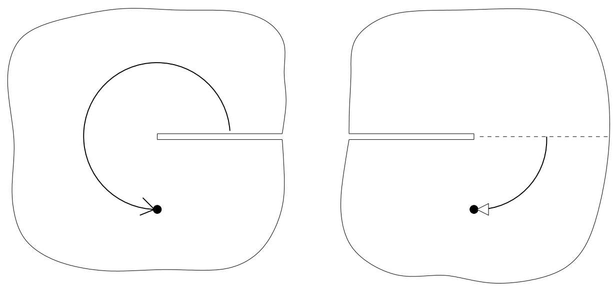
Si conocemos una bola $B = B(x_0, r) \supset U$ tal que f es inyectiva en B y $f(B) \supset V$, también queda determinada **la única inversa local g de f definida en V y tomando sus valores dentro de B.** Ahora U queda determinado a posteriori, como imagen directa $U = g(V) \subseteq B$.

La siguiente proposición 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 $g(y)$ de $f(x)$ definida para $y \in V$ y con $g(y_0) = x_0$ **o bien es única o bien no existe.**

Proposición 101. (Unicidad de la prolongación). *Sean 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.* □

Así, para especificar una inversa local g basta dar un dominio conexo por caminos V, especificar el valor de g en un punto concreto $y_0 \in V$ y ver si existe o no esa inversa local definida en V. Por ejemplo, la función $g : (-1,1) \to \mathbb{R}$, dada por $g(y) \equiv \operatorname{arcsen} y$, puede definirse como la inversa local de $f(x) \equiv \operatorname{sen} x$ definida para $y \in (-1,1)$ y con el valor particular $g(0) = 0$.
Del mismo modo la función $h : (-1,1) \to \mathbb{R}$, dada por $h(y) = 2\pi + \operatorname{arcsen} y$, es la inversa local de f definida también para $y \in (-1,1)$ pero con $h(0) = 2\pi$.

Atención: Se necesita especificar el dominio V tanto como la condición $g(y_0) = x_0$ para determinar (si existe) la inversa local g. Por ejemplo, la función CP tiene infinitas inversas locales distintas que satisfacen la condición $g(0,1) = (1, \pi/2)$. En particular, tiene una definida en $V_1 = \mathbb{R}^2 \setminus \big([0,+\infty) \times \{0\}$ y otra definida en $V_2 = \mathbb{R}^2 \setminus \big((-\infty, 0] \times \{0\}\big)$. Llamando g_1 y g_2, respectivamente, a esas dos inversas locales de CP, se tiene $g_1(-1,0) = (1, 3\pi/2)$ pero $g_2(-1,0) = (1, -\pi/2)$. O sea que g_1 y g_2 valen igual en $(0,1)$ pero valen distinto en el punto $(-1,0) \in V_1 \cap V_2$ común a sus dominios.



En ciertos abiertos $V \subseteq f(U_0)$ puede no existir ninguna inversa local de f. Por ejemplo CP no tiene ninguna inversa local definida en una corona circular centrada en $(0,0)$, pero sí tiene inversas locales definidas en el resultado de quitarle a esa corona circular un segmento que vaya de su circunferencia interior a la exterior.
Gracias al teorema 97, entre los abiertos $V \subseteq f(U_0)$ en los que sí hay definida una inversa local se incluyen, en particular, los *suficientemente pequeños.*

3.5.4 Radios euclídeos de inyectividad y suprayectividad

En este apartado la norma es siempre la euclídea estándar y todas las bolas son de esta norma.

El teorema 97 dice que, si f es de clase $\mathcal{C}^1$ y $(Df)_a$ es invertible, entonces f es inyectiva en entornos suficientemente pequeños $U \ni a$ y las imágenes $f(U)$ contienen entornos $V \ni f(a)$. Presentamos un método elemental para *cuantificar* ese resultado:

> Hallar dos radios explícitos $r, r' > 0$, tales que f es inyectiva en la bola euclídea $B(a,r)$ y la imagen $f\big(B(a,r)\big)$ contiene la bola euclídea $B\big(f(a), r'\big)$.
> Entonces sabremos que f tiene una, y sólo una, inversa local definida en $B\big(f(a), r'\big)$ y que tome sus valores dentro de $B(a,r)$.

No vamos a demostrar el siguiente resultado, sólo utilizarlo para ese propósito.

Teorema 102. *Sean un abierto $W \subseteq \mathbb{R}^n$, un punto $a \in W$ y $f : W \to \mathbb{R}^n$ de clase $\mathcal{C}^1$. Supongamos que existen dos números $r, \lambda > 0$ y una matriz* **ortogonal** *$P \in O(n)$ tales que*

$$\text{para todo } x \in \overline{B}(a,r) \text{ y todo } v \in \mathbb{R}^n \text{ se tiene } v^t\, P\, (Df)_x\, v \geq \lambda \|v\|^2 . \tag{41}$$

Entonces f es una función coerciva en $B(a,r)$, con constante de coercividad λ, y además:

$$f\big(B(a,r)\big) \supset V = B\big(f(a) \,,\, \lambda\, r \big) .$$

□

La constante de coercividad λ, que así se obtiene para f, es a veces mayor que 1.

Definiciones 103. *La pareja* $(r, \lambda r)$ *es un* **par de radios de inyectividad y suprayectividad para f,** *porque la restricción* $f|_{B(a,r)}$ *es a la vez inyectiva y "suprayectiva a* $B(a, \lambda r)$*", con lo cual* f *tiene una inversa local* $B\big(f(a), \lambda r\big) \longrightarrow U \subseteq B(a, r)$ *que es única.*

Decimos que una matriz $B \in M_{n\times n}(\mathbb{R})$ *es* **coerciva con constante de coercividad $\boldsymbol{\lambda}$** *si se cumple*

$$v^t\, B\, v \;\geq\; \lambda\, \|v\|^2 \quad , \quad \textit{para todo } v \in \mathbb{R}^n \,. \tag{42}$$

Veamos brevemente la relación entre esto y los autovalores. La desigualdad (42) equivale a

$$v^t\, S\, v \;\geq\; \lambda\, \|v\|^2 \,,$$

siendo $S = (B + B^t)/2$ la **parte simétrica de B.** Por lo tanto (42) es posible si y sólo si S es definida positiva, en cuyo caso la más grande constante de coercividad para B es el mínimo autovalor de su parte simétrica S.
La primera idea es evitar calcular autovalores, al precio de conformarnos con valores r, λ no del todo óptimos. Esto puede ser muy conveniente cuando la dimensión n es mayor que 2, ya que entonces los autovalores no se calculan de manera elemental. Las otras dos ideas son:

1. Una condición suficiente para que B tenga una constante de coercividad es que las entradas de su diagonal sean positivas y, en cierto modo, mayores que las otras entradas: que en la forma cuadrática $Q(v) = v^t B v$ "los cuadrados dominen a los términos mixtos".

2. Si $\{\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}\}$ es un permutación de la base estándar, y si P es la matriz cuyas filas son $\pm\mathbf{e}_{i_1}^t, \pm\mathbf{e}_{i_2}^t \ldots \pm \mathbf{e}_{i_n}^t$, entonces la operación $A \mapsto PA = B$ consiste en aplicar esa misma permutación a las filas de A y multiplicar algunas por -1.

Buscaremos, pues, permutar y cambiar de signo las filas de la jacobiana, de modo que la diagonal de $B = P(Df)$ sea positiva y "domine" al resto de la matriz B.

Ejemplo 1. Vamos a usar la idea 1. con $B = \begin{bmatrix} 1 & 1 \\ 3 & 5 \end{bmatrix}$. Estimamos $v^t B v$ confrontando los cuadrados con el término mixto:

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

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

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

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 $v^t A v \geq 0'061\|v\|^2$. Esto proporciona la constante de coercividad $\lambda = 0'061$.
La parte simétrica de B 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 calcular autovalores.

Ejemplo 2. Vamos a usar la idea 2. para la matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$. Como las entradas de la diagonal son pequeñas y las otras grandes, intercambiamos las filas para poner las entradas grandes en la diagonal:

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \implies PA = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} .$$

Entonces, usando $2|xy| \leq x^2 + y^2$, estimamos:

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

Llegamos a $v^t(PA)v \geq 2x^2 + y^2 \geq \|v\|^2$ y aceptamos la constante de coercividad $\lambda = 1$ para PA. 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 ortogonal $P \neq I_2$.

Ejemplo 3. Para $A = \begin{bmatrix} 1 & 1 \\ -5 & 2 \end{bmatrix}$ probamos con $P = \left[\begin{array}{c} -\mathbf{e}_2^t \\ \hline \mathbf{e}_1^t \end{array}\right] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$:

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

y la siguiente estimación nos dice que $\lambda = 0'5$ es una constante de coercividad para PA:

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

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

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

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

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

Si $(x, y) \in \overline{B}\big(a, 1/2\big)$, es decir $\|(x, y)\| \leq 1/2$, entonces $|x|, |y| \leq 1/2$. Ahora bien:

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

luego:

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

Tomando el menor de los dos coeficientes:

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

Para $(x, y) \in \overline{B}\big(a, 1/2\big)$ la matriz $Df_{(x,y)}$ es coerciva con constante de coercividad $1'67$. Como $(1'67)(1/2) = 0'835$, el teorema 102 nos dice que f es inyectiva en $B\big(a\,,\,1/2\big)$ y tiene una inversa local:

$$B\big(b\,,\,0'835\big) \longrightarrow U \;\subseteq\; B\big(a\,,\,1/2\big) \,,$$

y además sólo hay una inversa local satisfaciendo estas condiciones. Por supuesto, esta inversa local no es una función elemental.

3.6 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, junto con condiciones sobre qué abierto recorre la variable x y en qué abierto toma valores la variable y.
Esperamos **despejar tantas variables escalares como ecuaciones tenga el sistema.**

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

Segundo ejemplo. Sea $f : \mathbb{R}^5 \to \mathbb{R}^2$ dada por

$$f(x_1,x_2,x_3,x_4,x_5) \equiv \begin{bmatrix} \operatorname{sen} x_1 + \frac{1}{3}\,x_4^2 - e^{x_3}\,x_5 \\ 6x_1x_2x_3^2 + x_4 \end{bmatrix},$$

y sea $b = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$. Esperamos que el sistema $f(x) = b$, es decir

$$\left.\begin{array}{rcr} \operatorname{sen} x_1 + \frac{1}{3}\,x_4^2 - e^{x_3}\,x_5 &=& -1 \\ 6x_1x_2x_3^2 + x_4 &=& 0 \end{array}\right\}, \tag{43}$$

que tiene *dos* ecuaciones, nos permita despejar *dos* variables. En este ejemplo, podemos resolver (43) despejando las dos últimas variables:

$$\begin{array}{rcc} x_4 &=& -6x_1x_2x_3^2\,, \\ x_5 &=& e^{-x_3}\left(1 + \operatorname{sen} x_1 + 12x_1^2x_2^2x_3^4\right), \end{array}$$

que quedan expresadas como funciones de las tres primeras variables, las cuales a su vez son libres (en este ejemplo) de recorrer todo $\mathbb{R}^3$. Hemos obtenido la siguiente equivalencia:

$$f(x_1,x_2,x_3,x_4,x_5) = b \iff (x_4,x_5) = \varphi(x_1,x_2,x_3)\,, \tag{44}$$

donde φ es la siguiente función vectorial:

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2 \quad , \quad \varphi(x_1,x_2,x_3) \equiv \begin{bmatrix} -6x_1x_2x_3^2 \\ e^{-x_3}\left(1 + \operatorname{sen} x_1 + 12x_1^2x_2^2x_3^4\right) \end{bmatrix},$$

que es una *función vectorial implícita* definida por el sistema (43). La fórmula (44) dice que la *preimagen* $f^{-1}(\{b\}) = \{x \in \mathbb{R}^5 : f(x) = b\}$ coincide con el grafo de φ:

$$f^{-1}(\{b\}) = \left\{ \left(x\,,\,\varphi(x)\right) : x = (x_1,x_2,x_3) \in \mathbb{R}^3 \right\}.$$

Tercer ejemplo. Sea ahora $f : \mathbb{R}^3 \to \mathbb{R}$ dada por $f(x,y,z) \equiv xe^x + y \operatorname{sen} y - z^2$. Podemos resolver ecuación $f(x,y,z) = 2$, es decir

$$xe^x + y \operatorname{sen} y - z^2 = 2\,,$$

despejando la tercera variable. Salen dos opciones para el valor de z:

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

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

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

La preimagen queda descrita como *unión de dos grafos:*

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

Consideremos el abierto $U = \{ (x,y) : xe^x + y \operatorname{sen} y > 2 \} \subset C$. Para definir la función $z = z_+(x,y)|_U$, además de dar la ecuación $f(x,y,z) = 2$ y la restricción $(x,y) \in U$, hay que decir *dónde ha de tomar valores la función implícita,* o sea imponer la condición $z \in (0, +\infty)$:

$$\left.\begin{array}{r} z = z_+(x,y) \\ (x,y) \in U \end{array}\right\} \iff \left\{\begin{array}{l} f(x,y,z) = 2 \\ (x,y) \in U \\ z \in (0,+\infty) \end{array}\right.$$

Del mismo modo, la condición $z \in (-\infty, 0)$ da lugar a la función $z_-(x,y)|_U$. La identidad $f\big(x \,,\, y \,,\, \text{función}\,(x,y) \big) \equiv b$ es satisfecha por *dos* funciones implícitas definidas en el abierto U, pero podemos distinguir esas dos funciones por el abierto donde toman sus valores:

1. $\varphi(x) = z_+(x,y)|_U$ es la única función definida en U que satisface $f\big(x \,,\, y \,,\, \varphi(x,y) \big) \equiv b$ y toma todos sus valores dentro de $V_1 = (0, +\infty)$.

2. $\psi(x) = z_-(x,y)|_U$ es la única función definida en U que satisface $f\big(x \,,\, y \,,\, \psi(x,y) \big) \equiv b$ y toma todos sus valores dentro de $V_2 = (-\infty, 0)$.

El planteamiento general es el siguiente. Escribimos:

$$\mathbb{R}^{m+k} = \mathbb{R}^m \times \mathbb{R}^k = \big\{ (x,y) : x = (x_1, \ldots, x_m) \in \mathbb{R}^m \,,\, y = (y_1, \ldots, y_k) \in \mathbb{R}^k \big\} \,.$$

Partimos de un abierto $W_0 \subseteq \mathbb{R}^{m+k}$ y una función $f(x,y) : W_0 \to \mathbb{R}^k$ al menos de clase $\mathcal{C}^1$. Dado un punto $a \in W_0$, tomamos $b = f(a)$ y nos preguntamos si se puede describir la parte de la preimagen $f^{-1}(\{b\})$ que se encuentra cerca de a, es decir el conjunto

$$W \cap f^{-1}(\{b\}) = \{ x \in W : f(x) = b \} \quad , \quad W \text{ un cierto entorno de } a \text{ en } \mathbb{R}^{m+k} \,,$$

como el grafo de una función diferenciable $\varphi : U \to \mathbb{R}^k$ definida en un abierto $U \subseteq \mathbb{R}^m$:

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

Cuidado: 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, donde "cerca" significa "dentro de W".

Utilizaremos para W un producto cartesiano $W = U \times V$, con U abierto de $\mathbb{R}^m$ y V abierto de $\mathbb{R}^k$, entonces el sistema $\left.\begin{array}{r} f(x,y) = b \\ (x,y) \in W \end{array}\right\}$ equivale al sistema $\left.\begin{array}{r} f(x,y) = b \\ x \in U \\ y \in V \end{array}\right\}$, con el que pretendemos definir una función $y = \varphi(x) : U \to \mathbb{R}^m$ especificando tres cosas:

1. El dominio $U \subseteq \mathbb{R}^m$ donde $\varphi(x) \equiv \big(\varphi_1(x) \,,\, \ldots \,,\, \varphi_k(x) \big)$ va a estar definida.
2. El sistema de $\mathbf{k}$ ecuaciones $f\big(x \,,\, \varphi(x) \big) = b$, del cual esperamos despejar las $\mathbf{k}$ funciones escalares $\varphi_1(x), \ldots, \varphi_k(x)$ como funciones de las m variables $(x_1, \ldots, x_m) \in U$.
3. El abierto $V \subseteq \mathbb{R}^k$, en el que permitimos que tome valores φ. Esperamos que esto delimite una única función $\varphi(x)$ solución del sistema.

El teorema de las funciones implícitas da una condición *suficiente* (no necesaria) para que las $\mathbf{k}$ ecuaciones escalares $f(x,y) = b$, junto con las condiciones $y \in V$, nos permitan determinar (con unicidad) los valores de $\mathbf{k}$ incógnitas $y = (y_1, \ldots, y_k)$ en función de las otras m variables $(x_1, \ldots, x_m)$ que, a su vez, queden libres de recorrer un cierto abierto $U \subseteq \mathbb{R}^m$.

En el tercer ejemplo, si $W_1 = U \times (0, +\infty)$ entonces la parte $W_1 \cap f^{-1}(\{2\})$ de la preimagen $f^{-1}(\{2\})$ coincide con el grafo de una función unívocamente determinada:

$$\begin{aligned} W_1 \cap f^{-1}(\{2\}) &= \big\{ (x,y,z) : xe^x + y \operatorname{sen} y - z^2 = 2 \,,\, (x,y) \in U \,,\, z > 0 \big\} = \\ &= \big\{ \big(x \,,\, y \,,\, z_+(x,y) \big) \,,\, (x,y) \in U \big\} = \text{grafo de } z_+(x,y)|_U \,, \end{aligned}$$

y además $z_+(x,y)|_U \in \mathcal{C}^\infty(U)$, por ser raíz cuadrada de una función $\mathcal{C}^\infty$ y positiva en U. Si $W_2 = U \times (-\infty, 0)$, entonces la parte $W_2 \cap f^{-1}(\{2\})$ es el grafo de $z_-(x,y)|_U$.

Teorema 104. (De las funciones implícitas). *Sean un abierto $W_0 \subseteq \mathbb{R}^{m+k}$ y una función $f(x,y): W_0 \to \mathbb{R}^k$ de clase $\mathcal{C}^s$ con $1 \leq s \leq \infty$.*
Si en un punto $\boldsymbol{a = (x^0, y^0) \in W_0}$ los vectores $\boldsymbol{f_{y_1}(a), \ldots, f_{y_k}(a)}$ son independientes *(y por lo tanto una base de $\mathbb{R}^k$), entonces hay entornos $x^0 \in U \subseteq \mathbb{R}^m$ e $y^0 \in V \subseteq \mathbb{R}^k$, con $U \times V \subseteq W_0$, y una función $\varphi: U \to V$ tales que:*

$$(U \times V) \cap f^{-1}\big(\{f(a)\}\big) \;=\; \big\{ \big(\, x \,,\, \varphi(x) \,\big) \;:\; x \in U \,\big\} \,,$$

o bien, expresado con sistemas de condiciones: $$\left.\begin{array}{r} f(x,y) = f(a) \\ x \in U \\ y \in V \end{array}\right\} \iff \left\{ \begin{array}{l} y = \varphi(x) \\ x \in U \end{array}\right.$$
Además $\varphi(x) \in \mathcal{C}^s(U)$, o sea que $\varphi(x)$ es, al menos, tan suave como la función $f(x,y)$.

Demostración. Establecemos la notación:

$$D_x f \;=\; \big[\, f_{x_1} \,|\, f_{x_2} \,|\, \cdots \,|\, f_{x_m} \,\big]_{k\times m} \quad , \quad D_y f \;=\; \big[\, f_{y_1} \,|\, f_{y_2} \,|\, \cdots \,|\, f_{y_k} \,\big]_{k\times k} \,.$$

Partimos de la hipótesis que $D_y f(a)$ es una matriz invertible $k \times k$.
Definimos una función auxiliar:

$$F: W_0 \longrightarrow \mathbb{R}^{m+k} \quad , \quad F(x,y) \;\equiv\; \big(\, x \,,\, f(x,y) \,\big) \,,$$

y es evidente que $F \in \mathcal{C}^s(W_0)$, siendo s el mismo que para f. Además:

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

es triangular por cajas, luego invertible donde $f_{y_1}, \ldots, f_{y_k}$ sean linealmente independientes. Nuestra hipótesis equivale, pues, a que $(DF)_a$ sea una matriz invertible $(m+k)\times(m+k)$.
Si $a = (x_0, y_0)$, con $x_0 \in \mathbb{R}^m$ e $y_0 \in \mathbb{R}^k$, definimos $b = f(a)$ y $c = \big(x_0, f(a)\big) = (x_0, b)$. Entonces $F(a) = c$.
Aplicamos a F el teorema de la función inversa y obtenemos un abierto $W^a \subseteq \mathbb{R}^{m+k}$ tal que $a \in W^a$, la función F es regular e inyectiva en W^a y la inversa $F^{-1}: F(W^a) \to W^a$ es de clase $\mathcal{C}^s$. Elegimos abiertos $U_0 \subseteq \mathbb{R}^m$, $V \subseteq \mathbb{R}^k$ tales que

$$x_0 \in U_0 \quad , \quad y_0 \in V \quad , \quad a \;\in\; U_0 \times V \;\subseteq\; W^a \,.$$

Entonces F *es inyectiva en* $U_0 \times V$ y $F(U_0 \times V) \subseteq F(W^a)$ es un abierto de $\mathbb{R}^{m+k}$ que contiene el punto $c = (x_0, b)$, luego

$$\text{existe un abierto } U \subseteq \mathbb{R}^m \text{ tal que } \quad x_0 \in U \subseteq U_0 \quad \text{y} \quad U \times \{b\} \subset F(U_0 \times V) \,. \tag{45}$$

Sea x_0' un punto cualquiera de U. Por la fórmula que define F, sabemos que las preimágenes de (x_0', b) por $F|_{U_0 \times V}$ son los puntos (x_0', y) donde y es (cualquier) punto de V que cumpla la ecuación $f(x_0', y) = b$. Pero por (45) sabemos que (x_0', b) está en $F(U_0 \times V)$, luego tiene una única preimagen por F en $U_0 \times V$. De la unicidad de dicha preimagen se deduce, entonces, la unicidad del elemento $y \in V$ tal que $f(x_0', y) = b$. Como x_0' era un punto arbitrario de U, es correcto definir una función $\varphi: U \to V$ de la siguiente manera:

$$\begin{aligned} U \;\ni\; x \;\longmapsto\; \varphi(x) \;&=\; \big(\text{ el único } y \in V \text{ tal que } F(x,y) = (x,b) \,\big) \;= \\ &=\; \big(\text{ el único } y \in V \text{ tal que } f(x,y) = b \,\big) \,, \end{aligned}$$

Atención: fijado $x \in U$, la ecuación $f(x,y) = b$ puede satisfacerse por otros valores de y, pero esos valores no están en V.
Así llegamos a la siguiente igualdad conjuntista:

$$(U \times V) \cap f^{-1}(\{b\}) \;=\; \big\{ \big(\, x \,,\, \varphi(x) \,\big) \;:\; x \in U \,\big\} \,,$$

en particular $\boxed{\varphi(x_0) = y_0}$ porque $(x_0, y_0) \in (U \times V) \cap f^{-1}(\{b\})$.
La **inclusión** $i_b : \mathbb{R}^m \to \mathbb{R}^{m+k}$, dada por $i_b(x) \equiv (x, b)$, y la **proyección** $\pi : \mathbb{R}^{m+k} \to \mathbb{R}^k$, dada por $\pi(x, y) \equiv y$, son claramente funciones $\mathcal{C}^\infty$. Como se verifica:

$$\varphi \;\equiv\; \pi \circ \left(F|_{U_0 \times V}\right)^{-1} \circ \left(i_b|_U\right) ,$$

deducimos que $\varphi(x)$ es, al menos, igual de suave que $\left(F|_{U_0 \times V}\right)^{-1}$. Es decir $\varphi \in \mathcal{C}^s(U)$. □

Cuatro comentarios:

(1) Si mantenemos la ecuación $f(x, y) = b$ y la condición $x \in U$, pero cambiamos el abierto V en el que $y = \varphi(x)$ ha de tomar sus valores, la función implícita puede cambiar. Es lo que hemos visto más arriba en el tercer ejemplo.

(2) Se satisface la siguiente identidad entre funciones $U \to \mathbb{R}^k$:

$$f\big(\, x \,,\, \varphi(x) \,\big) \;\equiv\; b \,,$$

y, tomando en ambos miembros de esta identidad derivadas parciales con respecto a $x_1, \ldots, x_m$, se obtiene la siguiente identidad entre funciones matriz:

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

de donde es inmediato despejar $D_x\varphi$, porque los valores de $D_y f$ son matrices invertibles. Este procedimiento se conoce con el nombre de **derivación implícita.** Puesto que llegamos a una *identidad* $D_x\varphi \equiv \cdots$, podemos volver a derivar y así ir obteniendo indentidades para las derivadas segundas, terceras, etc, de la función φ.

(3) La función implícita $\varphi(x)$ puede no ser elemental, aunque $f(x, y)$ sí sea elemental.

(4) La función φ puede ser estrictamente más suave que la f. Es lo que ocurre, por ejemplo, con $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) \equiv y + |y|y$, que es $\mathcal{C}^1$ pero no $\mathcal{C}^2$. Esta función satisface $f_y > 0$ en todo punto, y, mediante la ecuación $f(x, y) = 0$, define implícitamente $y(x) = 0$ que es una función $\mathcal{C}^\infty$ de la variable x.