Funciones implícitas e inversas

por Leandro Caniglia

1. Convergencia uniforme

Hasta ahora habíamos visto sucesiones de puntos, entendiendo por "puntos" elementos de $\mathbf{R}$ o de $\mathbf{R}^n$. La palabra "punto" la utilizamos más generalmente para referirnos a un elemento de un "espacio". A su vez por *espacio* entendemos un conjunto con alguna estructura "geométrica". Estas nociones de punto y espacio son intuitivas. No hacen referencia a ninguna definción matemática. Es como cuando decimos "número"; la matemática no tiene una definción de *número* ya que no estudia esa noción en general sino en particular (números naturales, enteros, complejos, algebráicos, reales, etc.). Lo mismo ocurre con *espacio.* Hay espacios vectoriales, topológicos, métricos, pero no hay una definición matemática de espacio. En esta sección vamos a estudiar brevemente un "espacio" en donde los "puntos" son funciones.

Definiciones. Supongamos que D es un subconjunto cualquiera de $\mathbf{R}^n$. Tiene sentido considerar el conjunto de funciones de D en $\mathbf{R}^m$. Dada una función $\vec{f}$ en este conjunto, decimos que f es *acotada* si existe $M > 0$ tal que

$$\|\vec{f}(\vec{x})\| \leq M \quad \forall\, \vec{x} \in D.$$

Las funciones acotadas de D en $\mathbf{R}^m$ forman un espacio vectorial (de dimensión infinita si D es un conjunto infinito). La suma de funciones acotadas es una función acotada en virtud de la desigualdad triangular en $\mathbf{R}^m$

$$\|(\vec{f+g})(\vec{x})\| \leq \|\vec{f}(\vec{x})\| + \|\vec{g}(\vec{x})\|$$

y de la igualdad

$$\|(\lambda\vec{f})(\vec{x})\| = |\lambda| \cdot \|\vec{f}(\vec{x})\|.$$

A este espacio vectorial de funciones acotadas de D en $\mathbf{R}^m$ lo vamos a denotar $\mathcal{B}(D, \mathbf{R}^m)$. En este espacio vectorial definimos una *norma* poniendo

$$\|\vec{f}\| := \sup_{\vec{x}\in D} \|\vec{f}(\vec{x})\|.$$

El *cero* de este espacio vectorial es la función nula que a cada $\vec{x} \in D$ le asigna $\vec{0} \in \mathbf{R}^m$.

La definición anterior de "norma" cumple con las propiedades que se suelen exigir a las normas:

i) $\|\vec{f}\| \geq 0$ y $\|\vec{f}\| = 0$ si, y sólo si, $\vec{f}$ es la función nula.
ii) [Desigualdad triangular] $\|\vec{f} + g\| \leq \|\vec{f}\| + \|\vec{g}\|$.
iii) $\|\lambda\vec{f}\| = |\lambda| \cdot \|\vec{f}\|$.

Estas propiedades las puede comprobar fácilmente el lector a partir de las definiciones.

Definición. En todo espacio tiene sentido considerar sucesiones de puntos y ciertamente $\mathcal{B}(D, \mathbf{R}^m)$ no es una excepción. Diremos que una sucesión de funciones acotadas $(\vec{f}_n)_n$ es *uniformemente convergente* si existe una función (la cual resulta necesariamente acotada) $\vec{f}\colon D \to \mathbf{R}^m$ tal que

$$\forall\, \epsilon > 0,\ \exists\, n_0 \in \mathbf{N}\colon\ n \geq n_0 \Rightarrow \|\vec{f}_n - \vec{f}\| < \epsilon.$$

Si la sucesión de funciones $(\vec{f}_n)_n$ converge uniformemente, es claro que dado $\vec{x} \in D$ la sucesion de vectores $(\vec{f}_n(\vec{x}))_n$ de $\mathbf{R}^m$ también va a ser convergente (su límite será $\vec{f}(\vec{x})$). La convergencia uniforme es muchísmo más fuerte puesto que dado ϵ, el valor de δ no depende del punto $\vec{x}$ como en el caso de la segunda sucesión.

Definición. Decimos que una sucesión $(\vec{f}_n)_n$ de funciones acotadas es *de Cauchy* si para cada $\epsilon > 0$, existe $n_0 \in \mathbf{N}$ tal que

$$n, m \geq n_0 \quad \Rightarrow \quad \|\vec{f}_n - \vec{f}_m\| < \epsilon.$$

1.1 Lema. Cada sucesión de Cauchy $(\vec{y}_n)_n$ en $\mathbf{R}^m$ es convergente.

Demostración. Al proyectar los términos de la sucesión a cada una de las m coordenadas obtenemos m sucesiones de Cauchy en $\mathbf{R}$. Como en $\mathbf{R}$ las sucesiones de Cauchy son convergentes, en cada coordenada obtenemos un límite. El punto de $\mathbf{R}^m$ cuyas coordenadas sean esos puntos límites es el límte de la sucesión original ∎

1.2 Teorema. Cada sucesión de Cauchy $(\vec{f}_n)_n$ en $\mathcal{B}(D, \mathbf{R}^m)$ es uniformemente convergente.

Demostración. Para cada $\vec{x} \in D$ la sucesión $(\vec{f}(\vec{x}))_n$ es de Cauchy en $\mathbf{R}^m$. Por el lema, tiene un límite, el cual depende de $\vec{x}$, que llamaremos $\vec{g}(\vec{x})$. Esto define una función

$$\begin{aligned} \vec{g}\colon D &\longrightarrow \mathbf{R}^m \\ \vec{x} &\mapsto \lim_n \vec{f}_n(\vec{x}). \end{aligned}$$

Por definición la función $\vec{g}$ tiene la siguiente propiedad: para cada $\epsilon > 0$, existe $n_{\vec{x}} \in \mathbf{N}$ (dependiente de $\vec{x}$) tal que $\|\vec{f}_n(\vec{x}) - \vec{g}(\vec{x})\| < \epsilon$ si $n \geq n_{\vec{x}}$. Es suficiente probar que $\|\vec{f}_n - \vec{g}\| < 2\epsilon$ para n suficientemente grande. Tomemos $\vec{t} \in D$ arbitrario. Dado que la sucesión de funciones es de Cauchy, existe n_0 tal que $\|\vec{f}_n - \vec{f}_{n_0}\| < \epsilon$ si $n \geq n_0$. Elijamos $n_1 := \max(n_0, n_{\vec{t}})$, donde $n_{\vec{t}}$ es el entero dependiente de $\vec{t}$ tal que $\|\vec{f}_n(\vec{t}) - \vec{g}(\vec{t})\| < \epsilon$ si $n \geq n_{\vec{t}}$. Para $n \geq n_0$ tenemos

$$\begin{aligned} \|\vec{f}_n(\vec{t}) - \vec{g}(\vec{t})\| &\leq \|\vec{f}_n(\vec{t}) - \vec{f}_{n_1}(\vec{t})\| + \|\vec{f}_{n_1}(\vec{t}) - \vec{g}(\vec{t})\| \\ &< \|\vec{f} - \vec{f}_{n_1}\| + \epsilon \\ &= 2\epsilon. \end{aligned}$$

Como n_0 no depende de $\vec{t}$ y $\vec{t}$ era arbitrario, deducimos que $\|\vec{f}_n - \vec{g}\| < 2\epsilon$ si $n \gg 0$, como queríamos demostrar ∎

1.3 Teorema. Si $(\vec{f}_n)_n$ es una sucesión de funciones continuas y acotadas que converge uniformemente hacia una función $\vec{f}$, entonces $\vec{f}$ también es continua y acotada.

Demostración. Como toda sucesión convergente es de Cauchy, es claro (teorema anterior) que la función $\vec{f}$ es acotada. Por lo tanto sólo falta probar que es continua en cada punto de su dominio. Dados $\vec{x}_0 \in D$ y $\epsilon > 0$, existe n_0 tal que $\|\vec{f}_n - \vec{f}\| < \epsilon/3$ si $n \geq n_0$. Para ese n_0, la continuidad de $\vec{f}_{n_0}$ en $\vec{x}_0$ asegura la existencia de $\delta > 0$ tal que $\|\vec{f}_{n_0}(\vec{x}) - \vec{f}_{n_0}(\vec{x}_0)\| < \epsilon/3$ si $\|\vec{x} - \vec{x}_0\| < \delta$. Luego, para un tal $\vec{x}$ tenemos

$$\begin{aligned}\|\vec{f}(\vec{x}) - \vec{f}(\vec{x}_0)\| &\leq \|\vec{f}(\vec{x}) - \vec{f}_{n_0}(\vec{x})\| + \|\vec{f}_{n_0}(\vec{x}) - \vec{f}_{n_0}(\vec{x}_0)\| + \|\vec{f}_{n_0}(\vec{x}_0) - \vec{f}(\vec{x}_0)\| \\ &< \|\vec{f} - \vec{f}_{n_0}\| + \epsilon/3 + \|\vec{f} - \vec{f}_{n_0}\| \\ &= \epsilon\end{aligned}$$

como queríamos demostrar ∎

Ejercicio 1. Pruebe que el subconjunto de funciones continuas de $\mathcal{B}(D, \mathbf{R}^m)$ es un subespacio vectorial. (*Indicación*: un subespacio de un espacio vectorial V es un subconjunto S de V que tiene las siguientes tres propieadades: (1) $\vec{0} \in S$; (2) $\vec{s} + \vec{t} \in S$, simpre que $\vec{s}$ y $\vec{t}$ estén en S, y (3) $\lambda\vec{s} \in S$ para todo $\lambda \in \mathbf{R}$ y todo $\vec{s} \in S$.)

El espacio vectorial $\mathcal{B}(D, \mathbf{R}^m)$ posee como subespacio al conjunto de funciones continuas y acotadas de D en $\mathbf{R}^m$, que notaremos $\mathcal{C}^0(D, \mathbf{R}^m)$. El teorema anterior se puede interpretar diciendo que en $\mathcal{C}^0(D, \mathbf{R}^m)$ toda sucesión de Cauchy es convergente. En efecto, toda sucesión de Cauchy converge uniformemente a una función de $\mathcal{B}(D, \mathbf{R}^m)$ (anteúltimo teorema) la cual resulta continua en virtud del último teorema.

2. Aproximaciones sucesivas y puntos fijos

Los resultados de esta sección son *teoremas de existencia* en el sentido de establecer condiciones bajo las cuales es posible asegurar la existencia de una función que resuelva un problema determinado. Aquí la palabra *existencia* tiene un alcance mínimo: los teoremas nada dicen acerca de la forma en que esas funciones podrían ser calculadas o construídas, simplemente se limitan a garantizar su existencia. Este tipo de resultados hacen que la parte *computacional* de la teoría quede sin ser explorada. Aunque pueda resultar sorprendente, en matemática pura, la mera existencia de una función cumple un papel de enorme importancia y no impide que la teoría siga avanzando y obteniendo beneficios de éste tipo de resultados *no computables*.

2.1 Proposición. Supongamos que $B := B(\vec{0}, \alpha)$ y $B' := B(\vec{0}, \beta)$ son dos bolas abiertas en $\mathbf{R}^n$ y $\mathbf{R}^m$ respectivamente. Supongamos dada una función continua

$$\begin{aligned}\vec{v}\colon B \times B' &\longrightarrow \mathbf{R}^m \\ (\vec{x}, \vec{y}) &\mapsto v(\vec{x}, \vec{y})\end{aligned}$$

tal que

i) $\|\vec{v}(\vec{x}, \vec{y}_1) - \vec{v}(\vec{x}, \vec{y}_2)\| \leq k \cdot \|\vec{y}_1 - \vec{y}_2\|$ para todo $\vec{x} \in B$ donde k es una constante $0 \leq k < 1$ y

ii) $\|\vec{v}(\vec{x},\vec{0})\| < \beta(1-k)$ para todo $\vec{x} \in B$.

Entonces, en estas condiciones existe una única aplicación $\vec{y}\colon B \to B'$ con la siguiente propiedad

$$\vec{y}(\vec{x}) = \vec{v}(\vec{x},\vec{y}(\vec{x})) \qquad (\forall\, \vec{x} \in B).$$

Además, la función $\vec{y}$ es continua en B.

Demostración. Fijemos $\vec{x} \in B$. Para ese $\vec{x}$ construimos inductivamente una sucesión $(\vec{y}_i)_i$ de la siguiente manera: $\vec{y}_0 := \vec{0}$, $\vec{y}_1 := \vec{v}(\vec{x},\vec{y}_0)$, $\vec{y}_2 := \vec{v}(\vec{x},\vec{y}_1)$, etc. En general $\vec{y}_i := \vec{v}(\vec{x},\vec{y}_{i-1})$. Para ver que esta sucesión está bien definida tenemos que probar que todos los $\vec{y}_i$ pertenecen a B' (en caso contrario no tendría sentido hablar $\vec{v}(\vec{x},\vec{y}_i)$). Esto equivale a ver que $\|\vec{y}_i\| < \beta$. Ahora, es claro que esta propiedad es cierta cuando $i = 0, 1$ puesto que $\vec{y}_0 = \vec{0} \in B'$ y $\vec{y}_1 = \vec{v}(\vec{x},\vec{0})$ que tiene norma menor que β en virtud de (ii) y del hecho de que $k < 1$. Por otro lado, si $i > 0$ tenemos

$$\|\vec{y}_i - \vec{y}_{i-1}\| = \|\vec{v}(\vec{x},\vec{y}_{i-1}) - \vec{v}(\vec{x},\vec{y}_{i-2})\| \leq k \cdot \|\vec{y}_{i-1} - \vec{y}_{i-2}\|.$$

Por la misma razón

$$\|\|\vec{y}_i - \vec{y}_{i-1}\| \leq k^2 \cdot \|\vec{y}_{i-2} - \vec{y}_{i-3}\| \leq k^3 \cdot \|\vec{y}_{i-3} - \vec{y}_{i-4}\| \leq \cdots$$

y así sucesivamente llegamos a

$$\|\vec{y}_i - \vec{y}_{i-1}\| \leq k^{i-1}\|\vec{y}_1\|. \tag{1}$$

y para $1 \leq j \leq i$

$$\begin{aligned}
\|\vec{y}_i - \vec{y}_{i-j}\| &\leq \|\vec{y}_i - \vec{y}_{i-1}\| + \|\vec{y}_{i-1} - \vec{y}_{i-2}\| + \cdots + \|\vec{y}_{i-j+1} - \vec{y}_{i-j}\| \\
&\leq (k^{i-1} + k^{i-2} + \cdots + k^{i-j})\|\vec{y}_1\| \\
&= \frac{k^{i-j} - k^i}{1-k} \cdot \|\vec{y}_1\| \\
&= k^{i-j} \cdot \frac{1-k^j}{1-k} \cdot \|\vec{y}_1\|
\end{aligned} \quad . \tag{2}$$

Luego, para $j = i$

$$\|y_i\| \leq \frac{1-k^i}{1-k} \cdot \|\vec{y}_1\| \leq \beta(1-k^i) < \beta. \tag{3}$$

Para resaltar ahora el hecho de que la sucesión $(\vec{y}_i)_i$ depende de $\vec{x}$, ponemos $\vec{y}_i = \vec{y}_i(\vec{x})$. Esto define una función $\vec{y}_i\colon B \to B$ para cada $i \geq 0$. Además, estas funciones verifican

$$\vec{y}_i(\vec{x}) = \vec{v}(\vec{x},\vec{y}_{i-1}(\vec{x})). \tag{4}$$

Como $\vec{y}_0$ es la función constante nula, deducimos inductivamente de la igualdad anterior que todas las $\vec{y}_i$ son continuas. Además, de (2) se deduce inmediatamente que estas funciones forman una sucesión de Cauchy. Por los teoremas de convergencia que vimos, el límite de la sucesión es una función continua $\vec{y}: B \to B'$. Haciendo tender $i \to \infty$ en (4) queda

$$\vec{y}(\vec{x}) = \vec{v}(\vec{x}, \vec{y}(\vec{x}))$$

lo que prueba la existencia de la función $\vec{y}$. Falta demostrar la unicidad. Pero si $\vec{g}$ es otra función con la misma propiedad, es decir

$$\vec{g}(\vec{x}) = \vec{v}(\vec{x}, \vec{g}(\vec{x})),$$

resulta de las hipótesis que para cada $\vec{x} \in B$ es

$$\|\vec{y}(\vec{x}) - \vec{g}(\vec{x})\| \leq k \cdot \|\vec{y}(\vec{x}) - \vec{g}(\vec{x})\| < \|\vec{y}(\vec{x}) - \vec{g}(\vec{x})\|$$

con lo cual $\vec{y}(\vec{x}) = \vec{g}(\vec{x})$ para todo $\vec{x} \in B$ ∎

2.2 Corolario (*Teorema del Punto Fijo***).** Supongamos que $B := B(\vec{y}_0, \beta)$ es una bola abierta en $\mathbf{R}^n$ y que $\vec{v}: B \to \mathbf{R}^n$ es una función que verifica las siguientes dos condiciones

i) $\|\vec{v}(\vec{y}_1) - \vec{v}(\vec{y}_2)\| \leq k \cdot \|\vec{y}_1 - \vec{y}_2\|$ cualesquiera sean $\vec{y}_1, \vec{y}_2 \in B$, donde k es una constante $0 \leq k < 1$ y

ii) $\|\vec{v}(\vec{y}_0) - \vec{y}_0\| < \beta(1-k)$.

Entonces, en estas condiciones, existe un único punto $\vec{z} \in B$ tal que $\vec{v}(\vec{z}) = \vec{z}$.

Demostración. La primera condición asegura que $\vec{v}$ es continua. Por lo tanto se puede aplicar el teorema a la función

$$\begin{aligned} B \times B &\longrightarrow \mathbf{R}^n \\ (\vec{x}, \vec{y}) &\mapsto \vec{v}(\vec{y} + \vec{y}_0) - \vec{y}_0 \end{aligned}$$

que no depende de $\vec{x}$, y deducir la existencia de una función $\vec{z}$ tal que

$$\vec{z}(\vec{x}) = \vec{v}(\vec{z}(\vec{x}) + \vec{y}_0) - \vec{y}_0$$

de donde

$$\vec{v}(\vec{z}(\vec{x}) + \vec{y}_0) = \vec{z}(\vec{x}) + \vec{y}_0$$

lo que muestra que los puntos $\vec{z}(\vec{x}) + \vec{y}_0$ son puntos fijos de $\vec{v}$. La unicidad (que a su vez equivale a que $\vec{z}$ sea una función constante) se deduce de la condición (i) ∎

3. Teorema de la función implícita

Como todas las nociones que hemos venido estuidando hasta ahora, el contenido geométrico del Teorema de la Función Implícita es claro y sencillo. Imagine el gráfico de una función f de $\mathbf{R}^2$ en $\mathbf{R}$. Su aspecto es el de una superficie en $\mathbf{R}^3$. Piense ahora en la intersección de esa superficie con el plano $z = 0$. El dibujo que obtiene sobre ese plano (o cualquier otro plano horizontal) es el de una curva. Esa curva yace en el plano xy. Sus puntos verifican la ecuación $f(x, y) = 0$. Párese ahora en un punto de esa curva, digamos (a, b). En un entorno de (a, b) tenemos una descripción de la curva: ésta comprende a todos los puntos (x, y) tales que $f(x, y) = 0$. Sin embargo, esta ecuación es *implícita* puesto que no muestra claramente cómo despejar y en función de x ni x en función de y. Una ecuación *explícita* seriá de la forma $y = g(x)$ o $x = h(y)$. Cualquiera de estas dos ecuaciones sería preferible a la implícita puesto que siempre una ecuación explícita puede transformase en una implícita de manera trivial poniendo $f(x, y) = y-g(x)$ (o $x-h(y)$, según corresponda). La pregunta es entonces: ¿cuándo es posible encontrar una función que describa explícitamente y en función de x o x en función de y en un entorno de (a, b)? Esta es una pregunta que atiende únicamente a la existencia de una función que resuelva el problema. La pregunta se desentiende del problema siguiente que consiste en encontrar una tal función en el supuesto caso de que no haya un impedimento teórico inapelable. Es decir, la pregunta es filosófica, no computacional y sin embargo tiene una respuesta totalmente computacional, a saber: cuando alguna de las derivadas parciales $\partial f/\partial x$ o $\partial f/\partial y$ sea diferente de 0 en (a, b). Esta sección está dedicada a probar este teorema. Por supuesto, su enunciado y demostración los daremos con un poco más de generalidad para funciones de varias variables.

Notaciones. Si $\vec{f}\colon A \to \mathbf{R}^m$ es una función diferenciable en un abierto A de $\mathbf{R}^n \times \mathbf{R}^m$ y $(\vec{a}, \vec{b}) \in A$ es un punto de A, existe un entorno U de $\vec{a}$ tal que los puntos de la forma $(\vec{x}, \vec{b})$ con $\vec{x} \in U$ pertenecen a A. Queda así definida la función

$$\begin{aligned} \vec{f}_1\colon U &\longrightarrow \mathbf{R}^m \\ \vec{x} &\mapsto \vec{f}(\vec{x}, \vec{b}) \end{aligned}$$

que fija en $\vec{b}$ las m últimas coordenadas y deja variar las n primeras. A la diferencial de $\vec{f}_1$ en el punto $\vec{a}$ la notaremos $D_1\vec{f}(\vec{a}, \vec{b})$. Esta diferencial es una transformación lineal de $\mathbf{R}^n$ en $\mathbf{R}^m$. Análogamente, si fijamos la segunda variable obtenemos la función

$$\begin{aligned} \vec{f}_2\colon V &\longrightarrow \mathbf{R}^m \\ \vec{y} &\mapsto \vec{f}(\vec{a}, \vec{y}) \end{aligned}$$

definida en un entorno V de $\vec{b}$ cuyos puntos $\vec{y}$ satisfacen $(\vec{a}, \vec{y}) \in A$. A su diferencial en el punto $\vec{b}$ la notaremos $D_2\vec{f}(\vec{a}, \vec{b})$.

Ejercicio 1. Pruebe que con las notaciones recién introducidas vale la siguiente relación

$$D\vec{f}(\vec{a}, \vec{b})(\vec{s}, \vec{t}) = D_1\vec{f}(\vec{a}, \vec{b})(\vec{s}) + D_2\vec{f}(\vec{a}, \vec{b})(\vec{t}).$$

Endomorfismos y automorfismos. Una transformación lineal $L\colon \mathbf{R}^n \to \mathbf{R}^n$ cuyo dominio y codominio coinciden se denomina un *endomorfismo*. La matriz asociada a un endomorphismo es cuadrada de $n \times n$. La transformación lineal es biyectiva si, y sólo si, su matriz A tiene inversa A^{-1}. En ese caso se dice que la transformación lineal L es un *automorfismo* de $\mathbf{R}^n$.

3.1 Teorema. Supongamos que $\vec{f}: A \to \mathbf{R}^m$ es una función diferenciable y con todas sus derivadas parciales continuas en un abierto $A \subseteq \mathbf{R}^n \times \mathbf{R}^m$ (el mismo m que el del codominio). Supongamos además que $(\vec{a}, \vec{b}) \in A$ es un punto del dominio de $\vec{f}$ tal que $\vec{f}(\vec{a}, \vec{b}) = \vec{0}$ y que la diferencial $D_2\vec{f}(\vec{a}, \vec{b})$ (notación anterior) es un automorfismo de $\mathbf{R}^m$. En estas condiciones, existe un entorno U de $\vec{a}$ y una única aplicación continua $\vec{u}: U \to \mathbf{R}^m$ tal que

$$\vec{u}(\vec{a}) = \vec{b} \quad \text{y} \quad \forall\, \vec{x} \in U:\ \vec{f}(\vec{x}, \vec{u}(\vec{x})) = \vec{0}.$$

Además, $\vec{u}$ es diferenciable en U, todas sus derivadas parciales son continuas y su diferencial en cada punto de U está dado por

$$D\vec{u}(\vec{x}) = -(D_2\vec{f}(\vec{x}, \vec{u}(\vec{x})))^{-1} \circ (D_1\vec{f}(\vec{x}, \vec{u}(\vec{x}))).$$

Demostración. Llamemos $L: \mathbf{R}^m \to \mathbf{R}^m$ a la diferencial $D_2\vec{f}(\vec{a}, \vec{b})$. Por hipótesis L es un automorfismo de $\mathbf{R}^m$ y por lo tanto tiene una inversa L^{-1}. Si reemplazamos $\vec{f}$ por $L^{-1} \circ \vec{f}$ podemos suponer, sin pérdida de generalidad que $L = D_2\vec{f}(\vec{a}, \vec{b})$ es la identidad. También, reemplazando $\vec{f}$ por la función $(\vec{x}, \vec{y}) \mapsto \vec{f}(\vec{a} + \vec{x}, \vec{b} + \vec{y}))$ podemos suponer que $(\vec{a}, \vec{b})$ es $(\vec{0}, \vec{0}) \in \mathbf{R}^n \times \mathbf{R}^m$. Consideremos la función

$$\vec{v}(\vec{s}, \vec{t}) = \vec{t} - \vec{f}(\vec{s}, \vec{t}).$$

La función $\vec{v}$ está definida en un entorno de $(\vec{0}, \vec{0})$ en $\mathbf{R}^n \times \mathbf{R}^m$. Hacemos ahora la siguiente

Afirmación. La función $\vec{v}$ verifica las hipótesis de la prop. 2.1.

Demostración de la afirmación. De acuerdo a las definiciones tenemos

$$\vec{v}(\vec{s}, \vec{t}_1) - \vec{v}(\vec{s}, \vec{t}_2) = \vec{t}_1 - \vec{t}_2 - (\vec{f}(\vec{s}, \vec{t}_1) - \vec{f}(\vec{s}, \vec{t}_2))$$

Por otro lado, las condiciones de diferenciabilidad de $\vec{f}$ aseguran que dado $\epsilon > 0$ y arbitrario existen dos bolas abiertas $B := B(\vec{0}, \alpha)$ y $B' := B(\vec{0}, \beta)$ tales que

$$\begin{aligned} \|\vec{f}(\vec{s}, \vec{t}_1) - \vec{f}(\vec{s}, \vec{t}_2) - (\vec{t}_1 - \vec{t}_2)\| &\le \|\vec{f}(\vec{s}, \vec{t}_1) - \vec{f}(\vec{s}, \vec{t}_2) - D_2\vec{f}(\vec{s}, \vec{t}_1)(\vec{t}_1 - \vec{t}_2)\| \\ &\quad + \|D_2\vec{f}(\vec{s}, \vec{t}_2)(\vec{t}_1 - \vec{t}_2) - (\vec{t}_1 - \vec{t}_2)\| \\ &< \epsilon \cdot \|\vec{t}_2 - \vec{t}_1\| \end{aligned}$$

puesto que la matriz de $D_2\vec{f}(\vec{x}, \vec{y})$ está formada por derivadas parciales de $\vec{f}$ (con respecto a las últimas m variables) y tales derivadas son continuas por hipótesis. Por lo tanto si tomamos ϵ tal que $0 < \epsilon \le 1/2$, tendremos

$$\|\vec{v}(\vec{s}, \vec{t}_1) - \vec{v}(\vec{s}, \vec{t}_2)\| < \epsilon \cdot \|\vec{t}_2 - \vec{t}_1\| \le \frac{1}{2} \cdot \|\vec{t}_2 - \vec{t}_1\|$$

que asegura el cumplimiento del punto (i) en la prop. 2.1. Para ver que se satisface el punto (ii) de la misma prop. calculemos

$$\vec{v}(\vec{s},\vec{0}) = -\vec{f}(\vec{s},\vec{0}).$$

Como $\vec{f}(\vec{0},\vec{0}) = \vec{0}$ y $\vec{f}$ es continua, podemos achicar aún más α para que sea

$$\|\vec{f}(\vec{s},\vec{0})\| < \frac{\beta}{2}$$

con lo cual logramos que se verifique la condición (ii)

$$\|\vec{v}(\vec{s},\vec{0})\| < \beta\left(1 - \frac{1}{2}\right).$$

Esto pone fin a la demostración de la afirmación.

Ahora podemos aplicar la Prop. 2.1 a $\vec{v}$ y deducir la existencia de una única función continua $\vec{u}\colon B \to B'$ tal que

$$\begin{aligned}\vec{u}(\vec{s}) &= \vec{v}(\vec{s},\vec{u}(\vec{s}))\\ &= \vec{u}(\vec{s}) - \vec{f}(\vec{s},\vec{u}(\vec{s}))\end{aligned}$$

es decir

$$\vec{f}(\vec{s},\vec{u}(\vec{s})) = \vec{0}.$$

Vamos a mostrar que $\vec{u}$ es diferenciable en B (aquí tal vez tengamos que achicar α un poco más). Si llamamos $S(\vec{x}) := D_1(\vec{x},\vec{u}(\vec{x}))$ y $T(\vec{x}) := D_2\vec{f}(\vec{x},\vec{u}(\vec{x}))$, obtenemos

$$D\vec{f}(\vec{x},\vec{u}(\vec{x}))(\vec{s},\vec{t}) = S(\vec{x})(\vec{s}) + T(\vec{x})(\vec{t})$$

de donde

$$\lim_{(\vec{s},\vec{t})\to(\vec{0},\vec{0})} \frac{\|\vec{f}(\vec{x}+\vec{s},\vec{u}(\vec{x})+\vec{t}) - \vec{f}(\vec{x},\vec{u}(\vec{x})) - S(\vec{x})(\vec{s}) - T(\vec{x})(\vec{t})\|}{\|(\vec{s},\vec{t})\|} = 0$$

Tomemos $\vec{x}$ y $\vec{x}+\vec{s}$ en B y pongamos $\vec{t} := \vec{u}(\vec{x}+\vec{s})-\vec{u}(\vec{x})$. Ya hemos visto que $\vec{f}(\vec{x},\vec{u}(\vec{x})) = \vec{0}$ y que $\vec{f}(\vec{x}+\vec{s},\vec{u}(\vec{x})+\vec{t}) = \vec{0}$, de modo que vale la siguiente propiedad:
Dado $\eta > 0$, existe $r > 0$ tal que si $\|\vec{s}\|, \|\vec{t}\| < r$, entonces

$$\|S(\vec{x})(\vec{s}) - T(\vec{x})(\vec{t})\| < \eta \cdot \|(\vec{s},\vec{t})\| \leq \eta \cdot (\|\vec{s}\| + \|\vec{t}\|)$$

y como $T(\vec{x})$ es un automorfismo,

$$\|T^{-1}(\vec{x}) \circ S(\vec{x})(\vec{s}) - \vec{t}\| < \eta \cdot \|T^{-1}(\vec{x})\| \cdot (\|\vec{s}\| + \|\vec{t}\|).$$

Eligiendo $\eta \leq 1/2 \cdot \|T(\vec{x})\|$ para $\vec{x} \in B$ y definiendo

$$a := 2\|T^{-1}(\vec{x}) \circ S(\vec{x})\| + 1,$$

nos queda

$$\|\vec{t}\| - \frac{a-1}{2} \cdot \|\vec{s}\| \leq \frac{1}{2} \cdot (\|\vec{s}\| + \|\vec{t}\|)$$

es decir $\|\vec{t}\| \leq a\|\vec{s}\|$. Luego

$$\|\vec{t} + (T^{-1}(\vec{x}) \circ S(\vec{x}))(\vec{s})\| \leq \eta(a+1) \cdot \|T^{-1}(\vec{x})\| \cdot \|\vec{s}\|$$

siempre que $\|\vec{s}\| \leq r$. Reemplazando $\vec{t}$ por su valor llegamos a lo que queríamos demostrar ∎

3.2 Corolario (*Teorema de la Función Inversa*). Supongamos que $\vec{f}\colon A \to \mathbf{R}^n$ es una función diferenciable, con todas sus derivadas continuas en un abierto A de $\mathbf{R}^n$ (el mismo 'n' que en el codominio). Si la transformación lineal $D\vec{f}(\vec{a})$ es un automorfismo de $\mathbf{R}^n$, entonces existe un entorno abierto U de $\vec{a}$ en $\mathbf{R}^n$ tal que la restricción $\vec{f}|_U$ de $\vec{f}$ a U establece una biyección de U en un entorno abierto $V = \vec{f}(U)$ de $\vec{b} := \vec{f}(\vec{a})$ en $\mathbf{R}^n$ cuya inversa es también diferenciable.

Demostración. Pongamos $\vec{h}(\vec{x}, \vec{y}) := \vec{f}(\vec{x}) - \vec{y})$ para $\vec{x}, \vec{y} \in \mathbf{R}^n$. Como $D_1\vec{h}(\vec{a}, \vec{b}) = D\vec{f}(\vec{a})$ es un automorfismo, podemos aplicar el teorema (de la función implícita) intercambiando los roles de $\vec{x}$ e $\vec{y}$. Al hacer eso, deducimos la existencia de una función $\vec{u}\colon V \to \mathbf{R}^n$ diferenciable, con derivadas parciales continuas, definida en un entorno abierto V de $\vec{b}$ en $\mathbf{R}^n$ que envía $\vec{b}$ en $\vec{a}$ y verifica

$$\vec{f}(\vec{u}(\vec{y})) - y = \vec{h}(\vec{u}(\vec{y}), \vec{y}) = \vec{0}.$$

La función $\vec{u}$ es inyectiva (puesto que tiene una inversa a izquierda). Por lo tanto induce una biyección de V en $U := \vec{u}(V)$. Pero $U = \vec{u}(V) = \vec{f}^{-1}(V)$ es abierto, lo que termina la demostración ∎

Epimorfismos. Antes de enunciar el siguiente teorema, necesitamos recordar algunas cosas sobre transformaciones lineales. Ya vimos que una transformación lineal $L\colon \mathbf{R}^n \to \mathbf{R}^m$ tiene asociada una matriz A de m filas y n columnas. Esta asociación coloca en la $i$ésima columna de A al vector $L(\vec{e}_i)$, imagen por L del $i$ésimo vector canónico $\vec{e}_i$ de $\mathbf{R}^n$. Es bien sabido que la imagen de L es un *subespacio lineal* de $\mathbf{R}^m$. La dimensión de ese subespacio coincide con el *rango* de la matriz A, es decir el máximo número de columnas (o filas) linealmente independientes que posee A. En particular, L es sobreyectiva si, y sólo si, A tiene rango m. Las aplicaciones lineales sobreyectivas se llaman *epimorfismos*. Ejemplos de epimorphismos son las proyecciones π_i que a cada $\vec{x}$ le asignan su $i$ésima coordenada x_i. Más generalmente es posible definir una proyección sobre varias coordenadas. Pongamos por caso la función

$$\begin{aligned} \pi\colon \mathbf{R}^n &\longrightarrow \mathbf{R}^m \\ (x_1, \ldots, x_n) &\mapsto (x_{n-m+1}, \ldots, x_n) \end{aligned}$$

que a cada vector de n coordenadas le asigna el vector de sus últimas m coordenadas, desde la $(n-m+1)$ésima hasta la $n$ésima. La transformación lineal π es un epimorfismo (demuestre). De algún modo es el epimorfismo más sencillo que podemos definir de $\mathbf{R}^n$ en $\mathbf{R}^m$. Obviamente ésto sólo es posible si $n \geq m$.

3.3 Corolario (*Teorema del Rango*). Supongamos que $g\colon A \to \mathbf{R}^m$ es una función diferenciable, con derivadas parciales continuas en un abierto A de R^n, donde $n \geq m$. Si $\vec{a}$ es un punto de A tal que $D\vec{g}(\vec{a})$ es un epimorfismo, entonces existe una función inversible y diferenciable $\vec{h}\colon V \to U$ de un abierto V de $\mathbf{R}^n$ sobre un entorno abierto U de $\vec{a}$ en $\mathbf{R}^n$ con inversa diferenciable y tal que

$$\vec{g} \circ \vec{h}(x_1, \ldots, x_n) = (x_{n-p+1}, \ldots, x_n).$$

Demostración. Llamemos L a la diferencial $D\vec{g}(\vec{a})$. Supongamos primero que las últimas m columnas de la matriz de L son linealmente independientes, es decir, supongamos que los vectores $L(\vec{e}_{n-m+1}), \ldots, L(\vec{e}_n)$ forman un conjunto linealmente independiente en $\mathbf{R}^m$. Bajo esta suposición tomemos dentro de A un entorno de $\vec{a}$ de la forma $A_1 \times A_2$ con $A_1 \subseteq \mathbf{R}^{n-m}$ y $A_2 \subseteq \mathbf{R}^m$. Dentro de $A_1 \times A_2$ pensaremos a los puntos de $\mathbf{R}^n$ como pares $(\vec{x}, \vec{y})$ en donde $\vec{x}$ representa a las primeras $n - m$ coordenadas e $\vec{y}$ a las últimas m. Consideremos la función

$$\begin{aligned} \vec{G}\colon A_1 \times A_2 &\longrightarrow \mathbf{R}^n \\ (\vec{x}, \vec{y}) &\mapsto (\vec{x}, \vec{g}(\vec{x}, \vec{y})). \end{aligned}$$

La diferencial $D\vec{G}(\vec{a})$ tiene asociada una matriz de la forma

$$M = \begin{pmatrix} I_{n-m} & 0 \\ [D_1\vec{g}(\vec{a})] & [D_2\vec{g}(\vec{a})] \end{pmatrix}$$

donde I_{n-m} es la matriz identidad de $(n-m)\times(n-m)$, 0 es la matriz nula de $(n-m)\times m$ y los corchetes denotan a las matrices de las diferenciales $D_1\vec{g}(\vec{a})$ y $D_2\vec{g}(\vec{a})$ (vea las notaciones al comienzo de la sección §3). De acuerdo a nuestra suposición, $[D_2\vec{g}(\vec{a})]$ es una matriz de rango máximo y por lo tanto inversible. En consecuencia M resulta inversible. Podemos aplicar entonces el Teorema de la función inversa y deducir la existencia de una inversa de $\vec{G}$ definida en un entorno abierto V de $b := \vec{G}(\vec{a}) = (a_1, \ldots, a_{n-m}, g_1(\vec{a}), \ldots, g_m(\vec{a}))$. Por la forma de $\vec{F}$ es claro que su inversa tiene la forma

$$(\vec{x}, \vec{y}) \mapsto (\vec{x}, \vec{k}(\vec{x}, \vec{y})).$$

Usando esta información llegamos a

$$(\vec{x}, \vec{y}) = \vec{G}(\vec{x}, \vec{k}(\vec{x}, \vec{y})) = (\vec{x}, \vec{g}(\vec{x}, \vec{k}(\vec{x}, \vec{y})))$$

de donde

$$\vec{g}(\vec{x}, \vec{k}(\vec{x}, \vec{y})) = \vec{y}.$$

Por lo tanto $h(\vec{x}, \vec{y}) = (\vec{x}, \vec{k}(\vec{x}, \vec{y}))$ es la función que queríamos encontrar. Falta estudiar el caso general, en el que las m columnas independientes de la matriz de L no son las últimas. Pero si $\vec{\sigma}$ es una permutación de las variables que envía esos m índices a los índices $n - m + 1, \ldots, n$, entonces $\vec{\sigma}$ es un automorfismo (lineal), la composición $\vec{g} \circ \vec{\sigma}$

es una función del tipo ya considerado. Así, si $\vec{h}$ es la función que nos sirve para $\vec{g} \circ \vec{\sigma}$, entonces $\vec{\sigma} \circ \vec{h}$ nos servirá para $\vec{g}$ ∎

3.4 Corolario (*Multiplicadores de Lagrange***).** Supongamos que $\vec{g} = (g_1, \ldots, g_m)\colon A \to \mathbf{R}^m$ es una función como en el Teorema del Rango (corolario anterior). Llamemos Z al conjunto de ceros de $\vec{g}$, es decir,

$$Z = \vec{g}^{-1}(\vec{0}) = \{\vec{z} \in A \mid \vec{g}(\vec{z}) = \vec{0}\}$$

y supongamos que $D\vec{g}(\vec{z})$ es un epimorfismo para todo $\vec{z} \in Z$. Si $f\colon A \to \mathbf{R}$ es diferenciable y $\vec{a}$ es un extremo local (máximo o mínimo) de la restricción $f|_Z$ de f a Z, entonces existen $\lambda_1, \ldots, \lambda_m$ en $\mathbf{R}$ tales que

$$Df(\vec{a}) = \sum_{i=1}^{m} \lambda_i Dg_i(\vec{a})$$

Demostración. Por el Teorema del Rango (coro. anterior) aplicado a la función $\vec{g}$ y al punto $\vec{a} \in Z$, existe una una función $\vec{h}\colon V \to U$ biyectiva, diferenciable, con inversa diferenciable, donde U es un entrono abierto de $\vec{a}$ y tal que

$$\vec{g} \circ \vec{h}(x_1, \ldots, x_n) = (x_{n-m+1}, \ldots, x_n).$$

Pongamos $\vec{\pi} := \vec{g} \circ \vec{h}$. Entonces $\vec{\pi}$ es un epimorphismo y, por ser lineal, coincide con su diferencial. Si $\vec{b} := \vec{h}^{-1}(\vec{a})$, tenemos

$$\vec{\pi} = D\vec{\pi}(\vec{b}) = D\vec{g}(\vec{a}) \circ D\vec{h}(\vec{b}).$$

Dibujemos todas las funciones en un diagrama:

$$\begin{tikzcd}
\vec{h}^{-1}(Z) \arrow[r] \arrow[d, "\iota"] & Z \cap U \arrow[r, "f|_Z"] \arrow[d, "\iota"] & \mathbf{R} \\
V \arrow[r, "\vec{h}"] \arrow[dr, "\pi"] & U \arrow[r, "f"] \arrow[d, "\vec{g}"] & \mathbf{R} \\
& \mathbf{R}^m &
\end{tikzcd}$$

donde las flechas verticales rotuladas con ι son las inclusiones y la flecha $\vec{h}^{-1}(Z) \to Z \cap U$ es la restricción de $\vec{h}$. Por hipótesis, $f|_Z$ tiene un extremo local en $\vec{a}$. En consecuencia la restricción

$$f \circ \vec{h}|_{\vec{h}^{-1}(Z)}\colon \vec{h}^{-1}(Z) \to \mathbf{R}$$

tiene un extremo local en $\vec{b} = \vec{h}^{-1}(\vec{a})$. Y como

$$\vec{h}^{-1}(Z) = V \cap \pi^{-1}(\vec{0}) = V \cap \{(x_1, \ldots, x_{n-m}, 0, \ldots, 0) \mid x_i \in \mathbf{R}\}$$

deducimos que las $n-m$ primeras derivadas parciales de $f \circ \vec{h}$ deben anularse en $\vec{b}$. Eso significa que la diferencial de $f \circ \vec{h}$ en $\vec{b}$ es de la forma

$$D(f \circ \vec{h})(\vec{b})(\vec{s}) = \sum_{i=n-m+1}^{n} \lambda_i s_i$$

donde λ_i es la derivada parcial

$$\left.\frac{\partial(f \circ \vec{h})}{\partial x_i}\right|_{\vec{b}}$$

es decir

$$D(f \circ \vec{h})(\vec{b}) = \sum_{i=n-m+1}^{n} \lambda_i \pi_i = \sum_{i=n-m+1}^{n} \lambda_i D(g_i \circ \vec{h})(\vec{b})$$

o sea

$$Df(\vec{a}) \circ D\vec{h}(\vec{b}) = \sum_{i=n-m+1}^{n} \lambda_i Dg_i(\vec{a}) \circ D\vec{h}(\vec{b})$$

y componiendo a derecha por la inversa del automorfismo $D\vec{h}(\vec{b})$ obtenemos

$$Df(\vec{a}) = \sum_{i=n-m+1}^{n} \lambda_i Dg_i(\vec{a})$$

como queríamos demostrar ∎