FACULTAD DE MATEMÁTICAS

DEPARTAMENTO DE ANÁLISIS MATEMÁTICO

El teorema de la función implícita

Trabajo realizado por

José Carlos Castaño Muñoz

Tutor: Francisco José Freniche Ibáñez

Catedrático de la Universidad de Sevilla

Curso 2017 - 2018

Summary

In this work we study several proofs of the classical implicit function theorem, which gives sufficient conditions for a vector equation to define a function. The first proof is the most elementary. It is obtained by induction on the number of dependent variables. The second proof is also based on calculus, by looking directly at the linear approximation of the mapping. Finally, Banach fixed point theorem is used in the more abstract functional analytic third proof.

In the second part we show an application of the implicit function theorem. We give four different definitions of a smooth surface and prove all of them are equivalent using the range theorem, an important consequence of the implicit function theorem. We finish the work showing an alternative proof of the implicit function theorem using a classical existence theorem of solutions in ordinary differential equations.

Índice general

1 Introducción

Comenzamos formulándonos la siguiente pregunta: ¿qué es una función?

El concepto de función ocupa un lugar central en Matemáticas y su evolución histórica ha dado lugar a importantes desarrollos. Desde las tablas babilónicas de cuadrados, cubos e inversos de naturales, pasando por los cálculos trigonométricos en la matemática de la antigua Grecia, por la introducción realizada por Descartes del álgebra en la geometría, hasta llegar a la primera definición de Euler, la cual decía: "*Una función de una cantidad variable es una expresión analítica compuesta de cualquier forma de la cantidad variable y números o cantidades constantes*".

El concepto que se maneja en la actualidad está muy cerca de esta otra definición de Euler: "*Si algunas cantidades dependen de otras de modo que si éstas cambian, las primeras experimentan cambios, entonces las primeras cantidades se dicen funciones de las últimas. Esta definición se aplica bastante ampliamente e incluye todas las formas en las que una cantidad puede ser determinada por otra. Si, por lo tanto, x representa una cantidad variable, entonces todas las cantidades que dependen de x de cualquier manera, o están determinadas por x, se llaman funciones de x*"

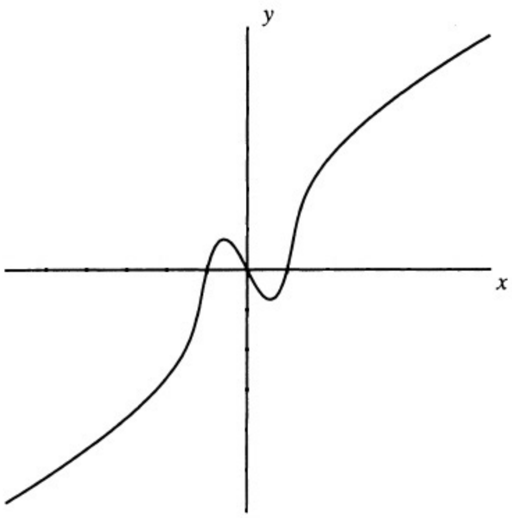
Informalmente, tendemos a pensar que una función $f : A \to B$ es una regla que asigna a cada $x \in A$, de manera única, un $y \in B$, que escribimos $y = f(x)$, y que su gráfica es el conjunto $\{(x, y) \in A \times B : y = f(x)\}$, aunque queda sin precisar qué se entiende por “regla”.

Por lo tanto, necesitamos de una definición formal, la cual formulamos en términos de teoría de conjuntos:

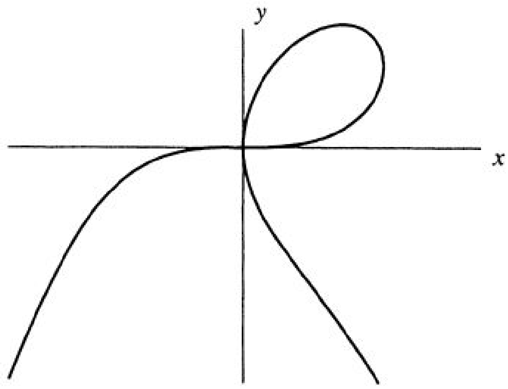
Definición. Una función $f : A \to B$ es un subconjunto f del producto cartesiano $A \times B$ tal que para todo $x \in A$ existe un único $y \in B$ tal que $(x, y) \in f$. Se escribe $y = f(x)$.

Es decir, ¡se define la función como su gráfica!

En las dos figuras siguientes podemos observar dos lugares geométricos, uno que es una función, el de la izquierda, y otro que no lo es, el de la derecha, debido a que para algunos x no tendríamos un único y:



$y^5 + 16y - 32x^3 + 32x = 0$



$y^3 + x^2y^2 - xy + x^4 = 0$

Se debe observar que en el primer ejemplo, aunque efectivamente se trata de una función, no se puede despejar y en términos de x: la función no puede expresarse explícitamente y hablamos de función definida implícitamente.

Aunque como hemos dicho, el segundo ejemplo no es o no define una función, salvo en el origen sí que podemos definir localmente una función en un entorno de cada punto que cumpla la ecuación: se dice que hay definida una función implícita local.

El teorema de la función implícita es, junto con el de la función inversa uno de los paradigmas más importantes y antiguos de la matemática moderna. Proporciona condiciones suficientes que debe cumplir una ecuación o conjunto de ecuaciones para que, localmente en el entorno de cada punto, algunas variables estén definidas como función implícita del resto de las variables.

Supongamos que tenemos un sistema de ecuaciones, en el cual hay más variables que ecuaciones ($n > m$):

$$\left.\begin{aligned} F_1(x_1, \ldots, x_n) &= c_1 \\ F_2(x_1, \ldots, x_n) &= c_2 \\ &\vdots \\ F_m(x_1, \ldots, x_n) &= c_m \end{aligned}\right\} (*)$$

Trabajando en un entorno de un punto $(p_1, \cdots, p_n)$ que verifica todas las ecuaciones, aproximamos linealmente el sistema sustituyendo las funciones $F_1, \cdots, F_m$ por sus diferenciales en ese punto:

$$\frac{\partial F_1}{\partial x_1}(p)(x_1 - p_1) + \cdots + \frac{\partial F_1}{\partial x_n}(p)(x_n - p_n) = 0$$

$$\frac{\partial F_2}{\partial x_1}(p)(x_1 - p_1) + \cdots + \frac{\partial F_2}{\partial x_n}(p)(x_n - p_n) = 0$$
$$\vdots$$
$$\frac{\partial F_m}{\partial x_1}(p)(x_1 - p_1) + \cdots + \frac{\partial F_m}{\partial x_n}(p)(x_n - p_n) = 0$$

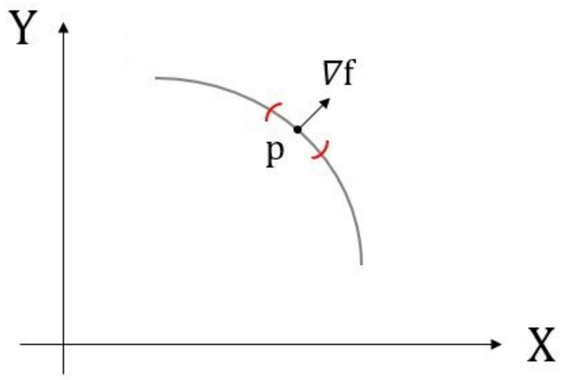
A este sistema lineal con m ecuaciones y n incógnitas se le puede aplicar el teorema de Rouché-Frobenius. Observamos que la matriz de coeficientes es la matriz jacobiana de la función vectorial $(F_1, \ldots, F_n)$ por lo que la condición para la existencia de soluciones se expresa en términos de no anulación de ciertos menores de esta matriz.

Teorema (Función implícita informal). Sean las funciones de nuestro sistema de ecuaciones $(*)$ diferenciables con continuidad. Si $(p_1, \cdots, p_n)$ verifica todas las ecuaciones y si, reemplazando $F_1, \cdots, F_m$ por sus aproximaciones lineales, un conjunto particular de m variables se puede expresar como funciones de las otras $n - m$ variables, entonces esas mismas m variables se pueden definir como funciones implícitas de las otras $n - m$ en un entorno del punto $(p_1, \cdots, p_n)$. Además, las funciones implícitas resultantes son diferenciables con continuidad.

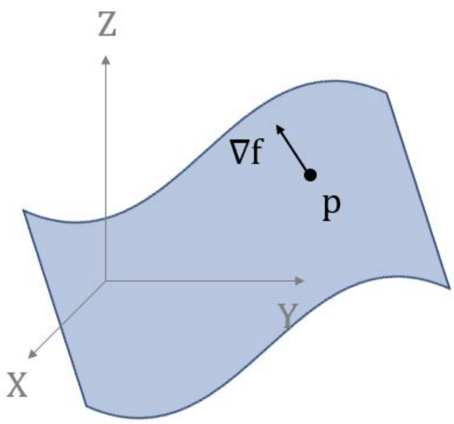
Para ilustrar geométricamente la condición de nuestro teorema, describimos a continuación tres casos sencillos que pueden visualizarse de manera intuitiva.

(a) El lugar geométrico de puntos del plano que cumplen la ecuación $f(x, y) = 0$ es generalmente una curva. Nos planteamos cuándo se define localmente y como función de x, es decir, si existe $y = y(x)$ tal que $f(x, y(x)) = 0$.

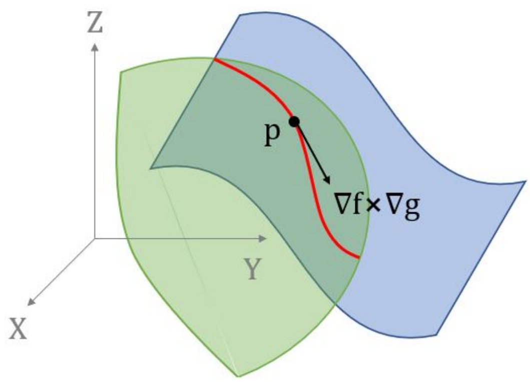
La respuesta según el teorema anterior es que basta que $\dfrac{\partial f}{\partial y}(p) \neq 0$. Geométricamente, esta condición equivale a que el vector tangente a la curva en p no sea vertical, o lo que es lo mismo, el vector normal $\nabla f(p)$ no sea horizontal.



(b) Una ecuación $f(x, y, z) = 0$ define generalmente una superficie en el espacio. Nos plantemos cuándo podremos definir localmente z como función de x e y, $z = z(x, y)$. La respuesta del teorema informal es que es suficiente con que $\dfrac{\partial f}{\partial z}(p) \neq 0$. Es decir, si el vector normal $\nabla f(p)$ a la superficie en el punto estudiado no es horizontal.



(c) Por último, si consideramos dos ecuaciones $f(x, y, z) = 0$ y $g(x, y, z) = 0$, que definirían una curva en el espacio, nos planteamos cuándo podremos definir y y z como funciones de x, es decir, $y = y(x)$ y $z = z(x)$. La condición suficiente es que el jacobiano $\dfrac{\partial(f,g)}{\partial(y,z)}(p)$ no sea nulo. Esta es la primera componente del vector $\nabla f(p) \times \nabla g(p)$, que es tangente a la curva intersección de las dos superficies. Es decir, el vector tangente a la curva no debe ser ortogonal al eje de las x.



En nuestro trabajo estudiamos con detalle el teorema de la función implícita y algunas de sus aplicaciones. En la primera parte de la memoria se presentan tres demostraciones diferentes del teorema y en la segunda parte se muestran dos aplicaciones: una a las superficies y otra a las ecuaciones diferenciales.

La primera prueba, la más elemental, se obtiene por inducción sobre el número de variables dependientes. Es, esencialmente, la realizada por Dini en 1870. El teorema se prueba primero para una variable dependiente, con una ecuación y cualquier número de variables independientes. El paso de la inducción se logra tomando una ecuación y una variable dependiente a la cual le podemos aplicar el caso base, tratando al resto de variables como independientes. La función resultante, definida implícitamente, la sustituiremos en las otras ecuaciones, reduciendo así el número de ecuaciones y variables dependientes a una.

La segunda prueba, que es la que tradicionalmente se estudia en el segundo curso del Grado de Matemáticas en Sevilla, consiste en demostrar primero el teorema de la función inversa y aprovechar que el de la función implícita es consecuencia no muy complicada de aquél. La técnica es puramente de Cálculo Vectorial, siguiendo la idea expuesta en esta introducción de aproximar linealmente la función por su diferencial. Se incluye también una demostración de que, recíprocamente, el teorema de la función inversa puede deducirse del de la implícita de manera muy sencilla.

La tercera y última de las pruebas que incluimos es la más abstracta y utiliza nociones de Análisis Funcional. Se basa en el método de aproximaciones sucesivas (teorema del punto fijo de Banach o de la aplicación contractiva) construyendo una sucesión de funciones y mostrando analíticamente cómo su límite cumple las ecuaciones implícitas y que, por tanto, es la función implícita que buscábamos. Se incluyen algunas ilustraciones para ejemplos en $\mathbb{R}^2$ y $\mathbb{R}^3$.

En la segunda parte daremos una aplicación, concretamente, se darán de forma precisa cuatro definiciones diferentes de superficie suave y demostramos su equivalencia. Nos apoyaremos en el teorema de la función implícita y en el teorema del rango, que vemos como consecuencia del de la inversa. El teorema del rango es una variante del teorema de la función implícita adaptado a situaciones donde la matriz jacobiana de la función vectorial tenga rango constante, pero no completo.

La memoria termina mostrando una demostración alternativa del teorema de la función implícita a partir del teorema de existencia de ecuaciones diferenciales de Cauchy-Peano.

Concluimos la introducción con unas breves notas históricas acerca de cómo fue surgiendo este teorema, viendo las necesidades y las complicaciones que llevaron a su formulación y demostración.

En los primeros trabajos sobre álgebra, cerca del año 850, tanto los problemas como las soluciones eran ejemplos numéricos, es decir, la noción de función no tenía sentido en este contexto. No fue hasta alrededor de 1600 cuando Vieta usó las letras para representar tanto incógnitas como coeficientes. Estos métodos de representación fueron tomados por Descartes, el cual los combinó con su sistema de coordenadas, lo que produjo un avance fundamental hacia el concepto de función como lo conocemos hoy día.

Desde el principio, muchas de las funciones se definían implícitamente, como la de la cuádrica general, dada por:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

Antes de 1800 no se sentía la necesidad de probar la existencia de funciones implícitas. De hecho Euler decía: "*... Por ejemplo, consideramos la función Z de z definida por la ecuación $Z^5 = az^2Z^3 - bz^4Z^2 + cz^3Z - 1$. A pesar de que esta ecuación no puede resolverse, sigue siendo cierto que Z es igual a una expresión compuesta de la variable z y constantes, y por esta razón Z será función de z*". Luego el enfoque de las funciones implícitas (Newton, Leibniz,...) hasta entonces era mostrar cómo se comportaban las funciones, y no comprobar la existencia de las mismas.

En 1770, Lagrange demostró el que puede considerarse el primer teorema de la función implícita, pero lo enunció como un teorema de funciones inversas, resultado que se conoce hoy día como teorema de inversión de Lagrange para series de potencias.

Se le atribuye el teorema de la función implícita a Cauchy, citando su *"Memoria de Turín"* como fuente del teorema.

No fué hasta más tarde, en el siglo XIX, cuando las diferencias entre el análisis complejo y real fueron clarificadas. Como consecuencia, la forma para variables reales del teorema no se enunció ni demostró hasta el trabajo de Dini, que presentó por primera vez en la Universidad de Pisa en el año académico 1876-1877.

2 Demostraciones

2.1. Prueba inductiva

Trataremos en este capítulo nuestra primera demostración, la formulada por Dini en la década de 1870, que fue la primera demostración formal conocida de nuestro teorema.

El teorema se prueba primero para una variable dependiente, con una ecuación y cualquier número de variables independientes, en donde, bajo suposición de monotonía de la función f, aplicando el teorema de Bolzano y demostrando la existencia de derivadas parciales de nuestras ecuaciones definidas implícitamente obtendremos el resultado requerido.

El paso de la inducción lo conseguiremos tomando una ecuación y una variable dependiente a la cual le podemos aplicar el teorema demostrado anteriormente, tratando al resto de variables como independientes. La función resultante, definida implícitamente, la sustituiremos en las otras ecuaciones, reduciendo así el número de ecuaciones y variables dependientes a una.

Comenzaremos recordando el concepto de función diferenciable.

Definición (Función diferenciable). Una función $f : \mathcal{U} \subset \mathbb{R}^m \to \mathbb{R}^n$ se dirá diferenciable en $x_0 \in \mathbb{R}^m$ si, siendo $\mathcal{U}$ abierto, existe una transformación lineal $T : \mathbb{R}^m \to \mathbb{R}^n$ que cumpla:

$f(x_0 + h) = f(x_0) + T(h) + \theta(h)$, donde θ cumple que:

$$\lim_{h\to 0} \frac{\|\theta(h)\|}{\|h\|} = 0$$

Teorema (Teorema de la función implícita para una variable dependiente). Sea $\mathcal{U} \subseteq \mathbb{R}^n$ abierto, $f : \mathcal{U} \to \mathbb{R} \in \mathcal{C}^1(\mathcal{U})$, y sea $p = (a, b) \in \mathbb{R}^{n-1} \times \mathbb{R}$, con $p \in \mathcal{U}$, tal que:

$$f(p) = 0 \qquad \text{y} \qquad \frac{\partial f}{\partial x_n}(p) \neq 0$$

Entonces, existe un conjunto abierto $\mathcal{V} \in \mathbb{R}^{n-1}$ con $a \in \mathcal{V}$ y existen c_1, c_2, con $c_1 < b < c_2$, tales que $\forall x \in \mathcal{V}$ existe un único $y \in (c_1, c_2)$ con $f(x, y) = 0$. Esto permite definir una función $g : \mathcal{V} \to \mathbb{R} \in \mathcal{C}^1(\mathcal{V})$, tal que:

(1) $b = g(a)$

(2) $f(x, g(x)) = 0, \forall x \in \mathcal{V}$

Demostración:

Sin pérdida de generalidad, podemos suponer $\dfrac{\partial f}{\partial x_n}(p) > 0$.

Por la continuidad de $\dfrac{\partial f}{\partial x_n}$, y pasando a un entorno más pequeño $\mathcal{U}$ de p si fuera necesario (aunque sin cambiar la notación), podemos asumir que:

$$\frac{\partial f}{\partial x_n}(q) > 0, \forall q \in \mathcal{U}$$

Como $f(a,\cdot)$ es una función creciente en un entorno de b, podemos encontrar c_1 y c_2, con $c_1 < b < c_2$ tales que:

$$f(a,c_1) < 0 < f(a,c_2)$$

Usando la continuidad de f podemos encontrar un entorno $\mathcal{V}$ de a tal que:

$$\mathcal{V} \times [c_1,c_2] \subseteq \mathcal{U} \quad \text{y} \quad f(x,c_1) < 0 < f(x,c_2), \quad \forall x \in \mathcal{V}.$$

Como se puede ver en la Figura 1, podemos apreciar los valores tomados por c_1 y c_2 y los valores positivos y negativos de f para todo punto $x \in \mathcal{V}$.

Por el teorema de Bolzano, tenemos que $\forall x \in \mathcal{V}$, $\exists y$ tal que $c_1 < y < c_2$, y $f(x,y) = 0$ (Figura 2), y como $\dfrac{\partial f}{\partial x_n}(q) > 0$ tenemos que y es único. Por tanto, está definida la función $g : \mathcal{V} \to \mathbb{R}$ tal que a cada x le asocia a y, donde tenemos que la unicidad de y implica la continuidad de g:

> Tomamos la sucesión $x_n \to x$, entonces $g(x_n) \to g(x)$, pues si no se cumpliera esto, por compacidad, existiría un n_k con $g(x_{n_k}) \to y' \neq y$. Pero $f(x_{n_k}, g(x_{n_k})) = 0 \ \forall k$, luego por continuidad tenemos que $f(x,y') = 0$, por tanto $y' = y$ debido a la unicidad de y.

Para completar la prueba falta ver que existe $\dfrac{\partial g}{\partial x_j}$ y que su valor es:

$$\frac{\partial g}{\partial x_j} = -\frac{\frac{\partial f}{\partial x_j}}{\frac{\partial f}{\partial x_n}}, \forall j = 1,2,...,n-1, \text{ pues así se tendría que } g \in \mathcal{C}^1(\mathcal{V}).$$

Fijados $x \in \mathcal{V}$ e $y = g(x)$, tenemos, por la diferenciabilidad de f que:

$$f(x+se_j, y+t) - f(x,y) = s\frac{\partial f}{\partial x_j}(x,y) + t\frac{\partial f}{\partial x_n}(x,y) + \epsilon\sqrt{s^2+t^2}$$

donde e_j es el vector básido $j-$ésimo y donde $\epsilon = \epsilon(s,t) \to 0$ cuando $(s,t) \to 0$.

Tomando $t = g(x+se_j) - g(x)$, obtenemos que:

$$\underbrace{f(x+se_j, g(x)+g(s+e_j)-g(x)}_{=0} - \underbrace{f(x,g(x))}_{=0} =$$

$$= s\frac{\partial f}{\partial x_j}(x,y) + t\frac{\partial f}{\partial x_n}(x,y) + \epsilon\sqrt{s^2+t^2}, \text{ luego:}$$

$$|t|\cdot\left|\frac{\partial f}{\partial x_n}(x,y)\right| \leq |s|\cdot\left|\frac{\partial f}{\partial x_j}(x,y)\right| + |\epsilon|\cdot|s|+|\epsilon|\cdot|t|$$

donde para un valor de $|s|$ suficientemente pequeño:

$$\left|\frac{g(x+se_j)-g(x)}{s} + \frac{\frac{\partial f}{\partial x_j}(x,y)}{\frac{\partial f}{\partial x_n}(x,y)}\right| = \left|\frac{t}{s} + \frac{\frac{\partial f}{\partial x_j}(x,y)}{\frac{\partial f}{\partial x_n}(x,y)}\right| \leq \left|\frac{\epsilon}{s}\cdot\frac{\sqrt{s^2+t^2}}{\frac{\partial f}{\partial x_n}(x,y)}\right|, \text{ donde:}$$

$$\sqrt{s^2+t^2} \leq 1 + \left|\frac{t}{s}\right|$$

$$\left|t\cdot\frac{\partial f}{\partial x_n}(x,y)\right| \leq |s|\cdot\left|\frac{\partial f}{\partial x_j}(x,y)\right| + |\epsilon|\cdot|s|+|\epsilon|\cdot|t|$$

donde si tomamos límite cuando $|s|\to 0$, entonces por la continuidad de g, tenemos que $|t|\to 0$, luego:

$$|\epsilon(s,t)| \leq \frac{1}{2}\cdot\left|\frac{\partial f}{\partial x_n}(x,y)\right|, \text{ con lo que tenemos que:}$$

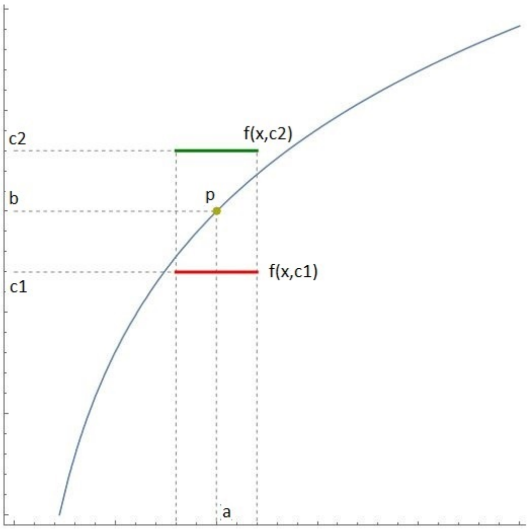
$$\frac{1}{2}\cdot|t|\cdot\left|\frac{\partial f}{\partial x_n}(x,y)\right| \leq |s|\cdot\left(\left|\frac{\partial f}{\partial x_j}\right| + \left|\frac{\partial f}{\partial x_n}\right|\right)$$

Luego obtenemos que $\frac{\partial g}{\partial x_j}$ existe, y viene dada por:

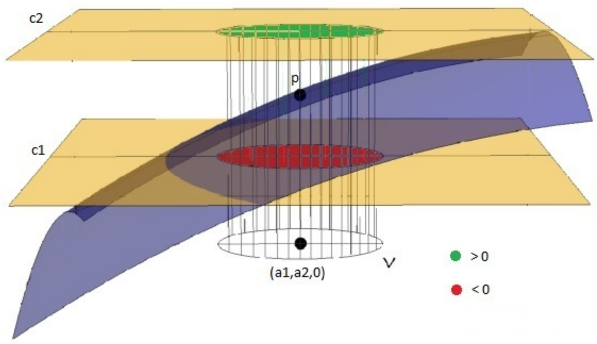
$$\frac{\partial g}{\partial x_j}(x) = -\frac{\frac{\partial f}{\partial x_j}(x,g(x))}{\frac{\partial f}{\partial x_n}(x,g(x))}, \text{ donde vemos que } g \in \mathcal{C}^1(\mathcal{V})$$

■

Figura 1

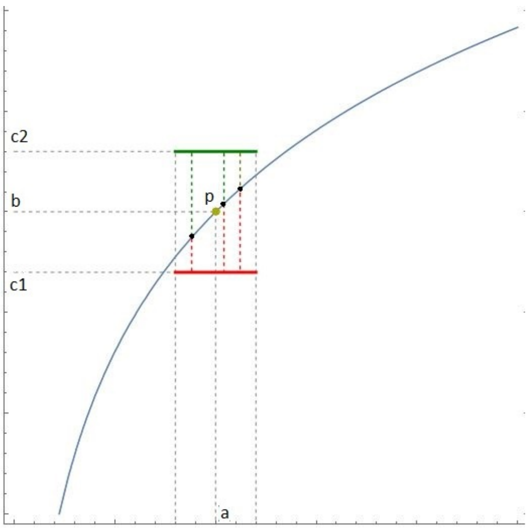


Gráfica para n=2

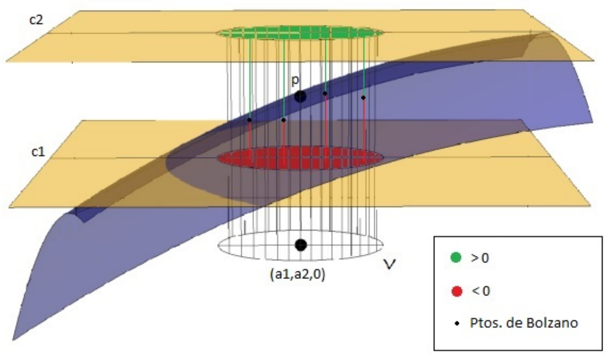


Gráfica para n=3

Figura 2



Gráfica para n=2



Gráfica para n=3

A continuación, y como será habitual en todas las secciones de nuestro trabajo, daremos unos conceptos y una notación necesaria para el seguimiento del teorema general.

Lema. Sea A una matriz real $n \times n$. Entonces existe una matriz invertible real U tal que UA es triangular superior.

Demostración:

La matriz A puede reducirse a una matriz escalonada mediante una sucesión finita de operaciones elementales por filas. Una matriz cuadrada en forma escalonada es, necesariamente, triangular superior y cada operación elemental por filas se puede lograr a través de la multiplicación por la izquierda por una matriz invertible, luego tomando U como la matriz producto de las operaciones elementales por filas, obtenemos que UA es triangular superior.

$\blacksquare$

Notación. Supondremos que tenemos un conjunto de ecuaciones:

$f_i(x_1, ..., x_l; y_1, ..., y_n) = 0$, $i = 1, ..., n$, donde las funciones $f_1, ..., f_n \in \mathcal{C}^1$.

Supondremos que $(p; q) = (p_1, ..., p_l; q_1, ..., q_n)$ es un punto que cumple todas las ecuaciones, donde:

$$\det \begin{pmatrix} \dfrac{\partial f_1}{\partial y_1} & \cdots & \dfrac{\partial f_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial y_1} & \cdots & \dfrac{\partial f_n}{\partial y_n} \end{pmatrix} \neq 0$$

Podemos pensar en las funciones $F_i(p, \cdot)$ como la función $F : \mathbb{R}^n \to \mathbb{R}^n$ definida mediante la aplicación:

$y \to F(y) = (f_1(p; y), f_2(p; y), ..., f_n(p; y))$.

Aplicando el lema anterior, podemos conseguir una nueva función, componiendo F con una transformación lineal, de forma que obtengamos:

$y \to \tilde{F}(y) = (\tilde{f}_1(p; y), ..., \tilde{f}_n(p; y))$ tales que:

$\dfrac{\partial \tilde{f}_i}{\partial y_j}(p; q) = 0$ para $i > j$

Para simplificar, usamos la misma notación, de forma que:

$\dfrac{\partial f_i}{\partial y_j}(p; q) = 0$ para $i > j$. De esta forma, obtenemos que:

$$D = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} & \cdots & \frac{\partial f_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial y_1} & \frac{\partial f_n}{\partial y_2} & \cdots & \frac{\partial f_n}{\partial y_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \\ 0 & \frac{\partial f_2}{\partial y_2} & \cdots & \frac{\partial f_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\partial f_n}{\partial y_n} \end{pmatrix}$$

y por tanto:

$$\det(D) = \prod_{i=1}^{n} \frac{\partial f_i}{\partial y_i} \neq 0$$

Teorema (Teorema de la función implícita). Existe un entorno $\mathcal{U} \subset \mathbb{R}^l$ de p y un conjunto de funciones $g_j : \mathcal{U} \to \mathbb{R}$, $j = 1, ..., n$, con $g_j \in \mathcal{C}^1(\mathcal{U})$ tales que:

$g_j(p) = q_j$, $j = 1, ..., n$

$f_i(x; g_1(x), g_2(x), ..., g_n(x)) = 0$, $i = 1, ..., n$, $\forall x \in \mathcal{U}$

Demostración:

La prueba se realiza por inducción en n:

Para $n = 1$:

Tenemos que es cierto, debido al teorema de la función implícita para una variable dependiente

Suponemos cierto el teorema para $n-1$ y comprobamos para n:

Con la notación dada anteriormente, tenemos:

$\dfrac{\partial f_n}{\partial y_n}(p;q) \neq 0$. Llamando $y' = (y_1, ..., y_{n-1})$, $q' = (q_1, ..., q_{n-1})$

entonces el teorema de la función implícita para una variable dependiente es aplicable a la ecuación:

$f_n(x; y', y_n) = 0$ en $(p; q', q_n)$, donde tratamos a las variables $x_1, ..., x_l, y_1, ..., y_{n-1}$ como independientes y solo a y_n como dependiente.

Existe por tanto un entorno $\mathcal{V} \in \mathbb{R}^{l+n-1}$ de $(p; q')$ y una función $\varphi : \mathcal{V} \to \mathbb{R}$, con $\varphi \in \mathcal{C}^1(\mathcal{V})$ tal que:

$\varphi(p; q') = q_n$

$f_n(x; y', \varphi(x; y')) = 0, \ \ \forall (x; y') \in \mathcal{V}$

Veamos que si derivamos f_n respecto a y_j, para $1 \leq j \leq n-1$, tenemos:

$\dfrac{\partial f_n}{\partial y_j} + \dfrac{\partial f_n}{\partial y_n} \cdot \dfrac{\partial \varphi}{\partial y_j} = 0$, que evaluando esta ecuación en $x = p$ e $y' = q'$ vemos que:

$$\frac{\partial \varphi}{\partial y_j}(p; q') = 0, \ \ \forall j = 1, 2, ..., n-1$$

Ahora, para cada $i = 1, ..., n-1$ definimos h_i como:

$$h_i(x_1, ..., x_l, y_1, ..., y_{n-1}) = f_i(x; y', \varphi(x; y'))$$

Consideramos el sistema de ecuaciones:

$$h_i(x_1, ..., x_l, y_1, ..., y_{n-1}) = 0, \ \ i = 1, ..., n-1.$$

Para $j = 1, ..., n-1$, ya que $\dfrac{\partial \varphi}{\partial y_j}(p; q') = 0$, tenemos:

$$\frac{\partial h_i}{\partial y_j}(p; q') = \frac{\partial f_i}{\partial y_j}(p; q', q_n) + \frac{\partial f_i}{\partial y_n}(p; q', q_n) \cdot \frac{\partial \varphi}{\partial y_j}(p; q') = \frac{\partial f_i}{\partial y_j}(p; q)$$

luego tenemos:

$$\begin{aligned} \det(A) &= \det \begin{pmatrix} \dfrac{\partial h_1}{\partial y_1}(p; q') & \cdots & \dfrac{\partial h_1}{\partial y_{n-1}}(p; q') \\ \vdots & \ddots & \vdots \\ \dfrac{\partial h_{n-1}}{\partial y_1}(p; q') & \cdots & \dfrac{\partial h_{n-1}}{\partial y_{n-1}}(p; q') \end{pmatrix} = \\ &= \det \begin{pmatrix} \dfrac{\partial f_1}{\partial y_1}(p; q) & \cdots & \dfrac{\partial f_1}{\partial y_{n-1}}(p; q) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_{n-1}}{\partial y_1}(p; q) & \cdots & \dfrac{\partial f_{n-1}}{\partial y_{n-1}}(p; q) \end{pmatrix} = \det(B) \end{aligned}$$

Por inducción, existe un entorno $\mathcal{W}$ de p en $\mathbb{R}^l$ y funciones $g_j : \mathcal{W} \to \mathbb{R} \in \mathcal{C}^1(\mathcal{W})$ tales que:

$$g_j(p) = q_j$$

$$h_i(x; g_1(x), ..., g_{n-1}(x)) = 0, \ \ i = 1, 2, ..., n-1, \ \ \forall x \in \mathcal{W}$$

Tomando $G(x) = (x; g_1(x), ..., g_{n-1}(x))$, $\mathcal{U} = \mathcal{W} \cup G^{-1}(\mathcal{V})$, definiendo $g_n : \mathcal{U} \to \mathbb{R}$ tal que $g_n(x) = \varphi(x; g_1(x), ..., g_{n-1}(x))$, y con la definición anterior de h_i, tenemos que:

$h_i(x_1, ..., x_l, y_1, ..., y_{n-1}) = f_i(x; y', \varphi(x; y'))$, donde vemos que se cumple que:

$f_i(x; g_1(x), ..., g_n(x)) = 0, \;\; i = 1, ..., n, \;\; \forall x \in \mathcal{U}$

■

2.2. Prueba a través del teorema de la función inversa

La segunda prueba consiste en relacionar nuestro teorema con el de la función inversa. Procederemos a demostrar primero el teorema de la función inversa. Éste se demuestra en tres pasos: viendo la inyectividad de F, que F transforma abiertos en abiertos y viendo que la inversa de F restringida a un entorno lo suficientemente pequeño es de clase $\mathcal{C}^1$.

Es apropiado comenzar esta sección realizando una revisión de la matriz jacobiana, su determinante y el papel que desempeña en el cálculo.

Sean $\mathcal{U}, \mathcal{V} \subset \mathbb{R}^n$ abiertos y sea la función $G : \mathcal{U} \to \mathcal{V} \in \mathcal{C}^1$, donde:

$G(x) = (g_1(x), ..., g_n(x))$.

Si $p \in \mathcal{U}$, entonces la matriz jacobiana de G en p es:

$$J_G(p) \equiv \begin{pmatrix} \dfrac{\partial g_1}{\partial x_1}(p) & \cdots & \dfrac{\partial g_1}{\partial x_n}(p) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_n}{\partial x_1}(p) & \cdots & \dfrac{\partial g_n}{\partial x_n}(p) \end{pmatrix}$$

que es el equivalente en varias variables a la primera derivada en el caso univariante. Esta matriz nos da una buena aproximación de la variación de G, es decir, $J_G(v)$ es una buena aproximación de $v \to G(p+v) - G(p)$.
Consideramos el determinante de la matriz jacobiana de G como:

$$|J_G(p)| = \det(J_G(p)) = \begin{vmatrix} \dfrac{\partial g_1}{\partial x_1}(p) & \cdots & \dfrac{\partial g_1}{\partial x_n}(p) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_n}{\partial x_1}(p) & \cdots & \dfrac{\partial g_n}{\partial x_n}(p) \end{vmatrix}$$

El resultado básico, como veremos más adelante, es que cuando $|J_G(p)| \neq 0$, entonces G es invertible en un entorno de p. La matriz jacobiana unifica la información sobre el comportamiento de primer orden (lineal) cerca de un punto.

Para el teorema de la función implícita no consideramos funciones equidimensionales. Así, tomamos el jacobiano en las variables para las cuales queremos resolverlo.
Una notación usual en lo referente a las componentes de una función es:

$G(x) = G(x_1, ..., x_n) \equiv (g_1(x_1, ..., x_n), ..., g_m(x_1, ..., x_n))$, y una elección de m argumentos (de los n posibles) lo expresamos como:

$x_{i_1}, x_{i_2}, ..., x_{i_m}$, y expresamos el jacobiano (determinante de la matriz jacobiana) como:

$$|J_G| = \frac{\partial(g_1, ..., g_m)}{\partial(x_{i_1}, ..., x_{i_m})} = \det \begin{pmatrix} \dfrac{\partial g_1}{\partial x_{i_1}} & \cdots & \dfrac{\partial g_1}{\partial x_{i_m}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_m}{\partial x_{i_1}} & \cdots & \dfrac{\partial g_m}{\partial x_{i_m}} \end{pmatrix}$$

Con esta notación, obtenemos una formulación estándar del teorema de la función implícita y del teorema de la función inversa, como vemos a continuación.

A continuación daremos los enunciados de ambos teoremas, donde hemos modificado el enunciado del teorema de la función implícita al dado en el capítulo anterior para facilitar la demostración a partir del teorema de la inversa y viceversa, aunque ambos enunciados son completamente equivalentes, como cabía esperar.

Teorema (Teorema de la función implícita). Sea

$$F(x) = F(x_1, ..., x_n) \equiv (f_1(x_1, ..., x_n), ..., f_m(x_1, ..., x_n))$$

una función de clase $\mathcal{C}^k$, con $k \geq 1$, definida en un conjunto $\mathcal{U} \subseteq \mathbb{R}^n$ abierto, y que toma valores en $\mathbb{R}^m$. Asumimos $1 \leq m < n$. Tomamos $q = n - m$.

Sea $x^0 = (x_1^0, ..., x_n^0)$ un punto de $\mathcal{U}$ que hemos fijado, y sea $x = (x_1, ..., x_n)$ un punto cualquiera de $\mathcal{U}$. Tomemos:

$x_a = (x_1, ..., x_q)$ primeras $n - m$ coordenadas de x

$x_a^0 = (x_1^0, ..., x_q^0)$ primeras $n - m$ coordenadas de x^0

Supongamos que:

$$\frac{\partial(f_1, ..., f_m)}{\partial(x_{q+1}, ..., x_n)}(x^0) \neq 0$$

Entonces existe un entorno $\mathcal{V}$ de x^0, un conjunto $\mathcal{W} \subseteq \mathbb{R}^q$ abierto que contiene a x_a, y funciones $g_1, ..., g_m \in \mathcal{C}^k$ en $\mathcal{W}$ tales que:

$$F(x_1, ..., x_q, g_1(x_a), ..., g_m(x_a)) = 0, \ \ \forall x_a \in \mathcal{W}.$$

Además, $g_1, ..., g_m$ son las únicas funciones que satisfacen:

$$\{x \in \mathcal{V} : F(x) = 0\} = \{x \in \mathcal{V} : x_a \in \mathcal{W}, x_{q+i} = g_i(x_a), \ \ i = 1, ..., m\}$$

Teorema (Teorema de la función inversa). Sea $\mathcal{U} \subseteq \mathbb{R}^n$ abierto, y sea $F : \mathcal{U} \to \mathbb{R}^n \in \mathcal{C}^k$, con $k \geq 1$. Sea x^0 un punto fijo de $\mathcal{U}$, y asumimos que $J_F(x^0) \neq 0$. Entonces existe un entorno $\mathcal{V} \subseteq \mathcal{U}$ de x^0 tal que:

(1) La restricción $F|_{\mathcal{V}}$ es inyectiva.

(2) El conjunto $\mathcal{V} = F(\mathcal{V})$ es abierto.

(3) La inversa F^{-1} de $F|_{\mathcal{V}}$ es de clase $\mathcal{C}^k$.

Demostración:

- PASO 1: La función F es localmente inyectiva:

Fijamos un punto arbitrario $x^0 \in \mathcal{U}$. Sea h la matriz inversa de J_F, y notamos por $\|h\|$ la norma de h, considerada como un operador lineal. Se define:

$$c = \frac{1}{\|h(x^0)\|}$$

Sea L la matriz jacobiana de F en x^0, y tomamos $\tilde{F}(t) = F(t) - L$. Entonces para $s, t \in \mathcal{U}$:

$F(s) - F(t) = L(s) - L(t) + [\tilde{F}(s) - \tilde{F}(t)]$, pero:

$$\frac{|\tilde{F}(s) - \tilde{F}(t)|}{s - t} \to 0 \text{ cuando } s, t \to (0, \cdots, 0)$$

Por lo tanto, para $\epsilon > 0$: $|F(s) - F(t)| \geq |L(s) - L(t)| - \epsilon \cdot |s - t|$, siempre que s esté cerca de t. Pero claramente:

$|L(s) - L(t)| \geq c \cdot |s - t|$, con lo que obtenemos que:

$|F(s) - F(t)| \geq (c - \epsilon) \cdot |s - t|$, donde, tomando $\epsilon = \dfrac{c}{2}$ obtenemos:

$$|F(s) - F(t)| \geq \frac{c}{2} \cdot |s - t|$$

Entonces $F(s) = F(t)$ implica que $s = t$, luego vemos que la función F es localmente inyectiva en $\mathcal{U}$.

- PASO 2: El conjunto $F(\mathcal{U})$ es abierto:

Sea $\mathcal{W} = F(\mathcal{U})$. Sea x^* un punto de $\mathcal{W}$. Veamos que x^* tiene un entorno en $\mathcal{W}$. Por el paso 1, podemos tomar $t^* \in \mathcal{V}$ tal que $F(t^*) = x^*$, donde $\mathcal{V}$ es un entorno de t^* en el cual F es inyectiva. Sea B una bola abierta con centro t^*, cuya clausura está en $\mathcal{V}$, y sea S la frontera de B, es decir, $S = \partial B$.

La inyectividad implica que $x^* \notin F(S)$. Por ser F continua y S compacto, $F(S)$ es compacto. Sea:

$$\sigma^* = \frac{1}{2}\mathrm{dist}(x^*, F(S))$$

Sea $\mathcal{V}^* = B(x^*, \sigma^*)$. Ahora, fijamos un punto arbitrario $x \in \mathcal{V}^*$. Entonces, $\forall t \in S$:

$$2\sigma^* \leq |x^* - F(t)| \leq |x^* - x| + |x - F(t)|$$

Como $|x^* - x| < \sigma^*$, vemos que:

$$\sigma^* < |x - F(t)|, \ \forall t \in S$$

Para $t \in \mathcal{U}$, definimos:

$$G(t) = |x - F(t)|^2 = \sum_{j=1}^{q} (x_j - f_j(t))^2$$

donde f_j son las componentes de F. Entonces G es de clase $\mathcal{C}^k$ y debemos tener un mínimo en la bola compacta $\overline{B}$, pero:

$$\left.\begin{aligned} &G(t^*) = |x - x^*|^2 < (\sigma^*)^2 \\ &G(t) > (\sigma^*)^2 \quad \forall t \in S \end{aligned}\right\}$$

Luego el mínimo valor de G en $\overline{B}$ es menor que $(\sigma^*)^2$ y por tanto, debe alcanzarse en algún punto $\tilde{t}$ interior de B. Por tanto debe ser un punto crítico de G y las derivadas parciales de G en $\tilde{t}$ deben ser cero:

$\dfrac{\partial G(t)}{\partial t_k} = -2\sum_{j=1}^{q}(x_j - f_j(t)) \cdot \dfrac{\partial f_j(t)}{\partial t_k}$, donde, tomando:

$c_j = x_j - f_j(\tilde{t})$, obtenemos:

$0 = \sum_{j=1}^{q} c_j \cdot \dfrac{\partial f_j(\tilde{t})}{\partial t_k}$, para cada k.

Como $\det(J_F(\tilde{t})) \neq 0$, por tanto, los vectores columna:

$\left(\dfrac{\partial f_1}{\partial t_j}(\tilde{t}), \cdots, \dfrac{\partial f_q}{\partial t_j}(\tilde{t})\right)$, $j = 1, ..., q$,

son linealmente independientes. Concluimos que $c_j = 0$, $j = 1, ..., q$, luego:

$x = F(\tilde{t})$

Hemos probado que si $x \in B'$, entonces $x = F(\tilde{t})$ para algún $\tilde{t} \in B$, ó $x \in F(B)$.

Por tanto $B' \subseteq F(B) \subseteq \mathcal{W}$, luego $F(\mathcal{V})$ es abierto.

**Notación:* Para usarlo en el resto de la prueba, fijamos un entorno $\mathcal{V}$ de x^0, donde F es inyectiva, y en el cual $\det(J_F)$ nunca se anula.

- PASO 3: La función $(F|_{\mathcal{V}})^{-1}$ es de clase $\mathcal{C}^1$:

Tenemos que $(F|_{\mathcal{V}})^{-1}$ existe , por la inyectividad local demostrada en el paso 1. Sea $x^* \in \mathcal{W}$ y $t^* = (F|_{\mathcal{V}})^{-1}(x^*)$ como en el paso anterior. Sea $L^* = J_F(t^*)$. Veamos que F^{-1} es diferenciable en x^* y que $J_F^{-1}(x^*) = (L^*)^{-1}$:

Sea $\tilde{c} = \dfrac{1}{\|(L^*)^{-1}\|}$. Para algún $\epsilon > 0$, existe una bola $\tilde{B} \equiv B(t^*, r^*) \subseteq \mathcal{V}$ tal que:

$$|F(t) - F(t^*) - L^*(t - t^*)| \leq \frac{\epsilon \cdot \tilde{c} \cdot c}{2}|t - t^*| \quad \forall t \in \tilde{B} \qquad (*)$$

Aquí, la constante c es la misma que la del paso 1, donde:

$|L(s) - L(t)| \geq c \cdot |s - t|$

Por el paso 2, hay un entorno $B^* \equiv B(x^*, s)$ tal que $B^* \subseteq F(\tilde{B})$.

Sea $x \in B^*$, entonces $x = F(t)$ para algún $t \in \tilde{B}$. Como $x^* = F(t^*)$, por el paso 1, tenemos que:

$$\frac{c}{2} \cdot |t - t^*| \leq |x - x^*| \qquad (**)$$

Además, como $t = F^{-1}(x)$ y $t^* = F^{-1}(x^*)$, tenemos:

$$L^*[F^{-1}(x) - F^{-1}(x^*) - (L^*)^{-1}(x - x^*)] = -[F(t) - F(t^*) - L^*(t - t^*)]$$

Como $\tilde{c} \cdot |\omega| \leq L^*(\omega) \quad \forall \omega \in \mathbb{R}^q$, obtenemos:

$$\tilde{c}|F^{-1}(x) - F^{-1}(x^*) - (L^*)^{-1}(x - x^*)| \leq |F(t) - F(t^*) - L^*(x - x^*)|$$

Como consecuencia de $(*)$ y $(**)$, $\forall x \in B^*$, se cumple que:

$|F^{-1}(x) - F^{-1}(x^*) - (L^*)^{-1}(x - x^*)| \leq \epsilon \cdot |x - x^*|$, luego F^{-1} es diferenciable en x^* y además:

$$J_F^{-1}(x^*) = L^{-1}.$$

En resumen, la función F^{-1} es una función diferenciable, y:

$$J_F^{-1}(x) = [J_F(F^{-1}(x))] \quad \forall x \in \mathcal{W} \qquad (* * *)$$

de donde se deduce que F^{-1} es continua.

Como cada función:

$\dfrac{\partial f_i}{\partial x_j}$ es continua, la composición:

$\dfrac{\partial f_i}{\partial x_j} \circ F$ también es continua.

Además, usando $(* * *)$ y la regla de Cramer, tenemos que las derivadas parciales:

$\dfrac{\partial (F^{-1})_i}{\partial x_j}$ son continuas.

Con lo que obtenemos que la función F^{-1} es de clase $\mathcal{C}^1$.

Si ahora F es de clase $\mathcal{C}^m$, entonces cada $\dfrac{\partial f_i}{\partial x_j}$ es de clase $\mathcal{C}^{m-1}$, luego $\dfrac{\partial f_i}{\partial x_j} \circ F^{-1}$ es de clase $\mathcal{C}^{m-1}$. Obtenemos entonces que:

$\dfrac{\partial(F^{-1})_i}{\partial x_j}$ es de clase $\mathcal{C}^{m-1}$, y entonces F^{-1} es de clase $\mathcal{C}^m$.

Inductivamente, tenemos que si F es de clase $\mathcal{C}^k$, entonces también lo es F^{-1}.

Estos 3 pasos juntos completan la prueba del teorema de la función inversa.

■

Concluiremos viendo que partiendo de un teorema podemos llegar al otro y viceversa, es decir, hay una relación de equivalencia entre ambos. Para ello:

i) Partiendo del teorema de la implícita, tomando $x_a = (x_1, \cdots, x_n)$, $x_b = (x_{n+1}, \cdots, x_{2n})$ y $f_i(x_a) = x_{n+i}$, para saber si existen $h_i(x_b) = x_i$, definiremos la función $F_i(x_a, x_b) = f_i - x_{n+i}$ a la cual le aplicaremos el teorema que tenemos por hipótesis para llegar al resultado deseado.

ii) Partiendo del teorema de la inversa basta tomar la transformación $F(x) = (x_1, ..., x_q, g_1(x), ..., g_m(x))$, donde $\det(J_g)$ es no nulo, y aplicarle el teorema de la función inversa a dicha función con lo que obtendríamos que $F(x) = (x_a, 0)$ si y solo si $x = F^{-1}(x_a, 0)$.

Con esta doble implicación tendríamos resuelta nuestra segunda prueba.

Teorema (Teorema de equivalencia). Teorema de la función inversa $\Leftrightarrow$ Teorema de la funcion implícita.

Demostración:

$\boxed{\Leftarrow}$: Sean $x_a = (x_1, \cdots, x_n)$ y $x_b = (x_{n+1}, \cdots, x_{2n})$

Tenemos que $f_i(x_a) = x_{n+i}$, $i = 1, \cdots, n$, y queremos saber si existen funciones g_i tales que:

$g_i(x_b) = x_i$, $i = 1, \cdots, n$.

Definamos $F_i(x_a, x_b) = f_i - x_{n+i}$, $i = 1, \cdots, n$, una familia de funciones de clase $\mathcal{C}^1$, con $J_F(x_a, x_b) \neq 0$ a las cuales les vamos a aplicar el teorema de la función implícita, ya que cumplen las condiciones necesarias, con lo que obtenemos que existen n funciones $g_1, \cdots, g_n$ de clase $\mathcal{C}^1$ y un entorno $\mathcal{V}$ abierto de x_b tales que:

$F_i(g_i(x_b), x_{n+i}) = 0$, lo que implica que:

$f_i(g_i(x_b)) - x_{n+i} = 0 \Rightarrow f_i(g_i(x_b)) = x_{n+i}$

por tanto tenemos que $g_i \circ f_i = Id$, lo que implica que una es la inversa de la otra. Si llamamos $G = (g_1, \cdots, g_n)$, tenemos que tomando $\mathcal{W} = G^{-1}(\mathcal{V})$ obtenemos el resultado deseado.

$\boxed{\Rightarrow}$: Nuestra función F es, al menos, de clase $\mathcal{C}^1$, por tanto el jacobiano:

$\dfrac{\partial(f_1, ..., f_m)}{\partial(x_{q+1}, ..., x_n)}$ es continuo, luego hay un entorno $\mathcal{U}$ de x^0, en el cual, el jacobiano no se anula.

Consideremos la transformación $G : \mathcal{U} \to \mathbb{R}^n$, dada por:

$G(x) = (x_1, ..., x_q, f_1(x), ..., f_m(x))$. Entonces, $G \in \mathcal{C}^k$ al igual que F. La matriz jacobiana es:

$$J_F = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_q} & \dfrac{\partial f_1}{\partial x_{q+1}} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \dfrac{\partial f_m}{\partial x_2} & \cdots & \dfrac{\partial f_m}{\partial x_q} & \dfrac{\partial f_m}{\partial x_{q+1}} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Es obvio que el determinante de esta matriz es solo el determinante de $m \times m$ del bloque inferior derecho, luego vemos que $|J_G(x)| \neq 0, \ \forall x \in \mathcal{U}$.

Por el teorema de la función inversa concluimos que hay un entorno $\mathcal{V}$ de x^0 tal que $G(\mathcal{V})$ es un conjunto abierto y la restricción $G|_{\mathcal{V}}$ tiene una inversa G^{-1} de clase $\mathcal{C}^k$.

Ahora escribimos $(x_a, 0) = (x_1, ..., x_q, 0, ..., 0)$. Tomamos:

$R = \{x_a : (x_a, 0) \in G(\mathcal{V})\}$. Ya que $G(\mathcal{V})$ es un subconjunto abierto de $\mathbb{R}^n$, R también es abierto. Ahora $\forall x_a \in R$ tomamos:

$g_i(x_a) = f_{q+i}(x_a, 0)$, para $i = 1, ..., m$.

Para $x \in \mathcal{V}$, $G(x) = 0$ si y solo si $x_a \in R$ y $G(x) = (x_a, 0)$. Ya que $G|_{\mathcal{V}}$ y G^{-1} son inversas, $G(x) = (x_a, 0)$ si y solo si $x = G^{-1}(x_a, 0)$.

$\blacksquare$

2.3. Prueba mediante el método de aproximaciones sucesivas

Esta tercera y última de las pruebas que incluimos es la más abstracta y utiliza nociones de Análisis Funcional. Se basa en el método de aproximaciones sucesivas (el cual se basa en tomar un punto inicial e ir mejorando nuestro resultado a través de las iteraciones), en el cual construiremos una sucesión de funciones y mostraremos analíticamente cómo su límite (que sabemos que existe debido al teorema de convergencia global de este método) cumple las ecuaciones implícitas y que, por tanto, es la función implícita que buscábamos.

Se incluirán tanto ilustraciones del teorema del punto fijo para aclarar su demostración como ejemplos en $\mathbb{R}^2$ y $\mathbb{R}^3$ del teorema de la función implícita para ver gráficamente como nuestra sucesión alcanza su límite en la función implícita.

Comencemos, como es habitual, con unas definiciones y notaciones previas para la mejor comprensión del capítulo.

Notación. En este apartado vamos a usar las letras F para contracciones del espacio métrico X en si mismo, es decir, $F : X \to X$, y usaremos la letra ρ para indicar la métrica de dicho espacio.

Definición. Sea X un espacio métrico completo con métrica ρ. Una función $F : X \to X$ se llama una contracción si existe una constante $0 < c < 1$ tal que:

$$\rho(F(x), F(y)) \leq c \cdot \rho(x, y) \quad \forall x, y \in X$$

El hecho de que c sea menor que 1 nos dice que la imagen de un conjunto bajo F es contraída, los puntos en $F(X)$ están más juntos que los puntos de sus contraimágenes en X. El teorema básico sobre funciones contractivas es el siguiente:

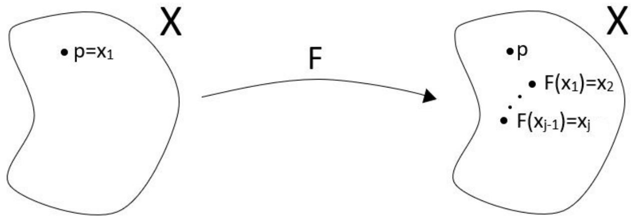
Teorema (Teorema del punto fijo de Banach). Sea $F : X \to X$ una contracción del espacio métrico completo X. Entonces F tiene un único punto fijo, es decir, hay un único punto $x_0 \in X$ tal que $F(x_0) = x_0$.

Demostración:

Sea $p \in X$ un punto cualquiera. Definimos una sucesión inductivamente como:

$$x_1 = p$$
$$x_2 = F(x_1)$$
$$\vdots$$
$$x_j = F(x_{j-1})$$

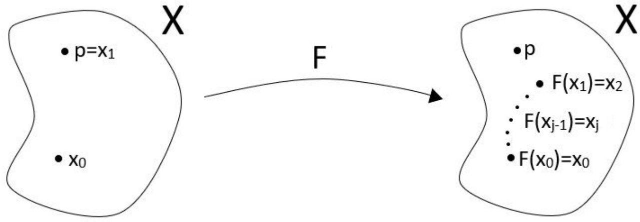
la cual representamos en la siguiente figura:



Afirmamos que $\{x_j\}$ es una sucesión de Cauchy en X. Supongamos que esta afirmación ha sido probada.

Entonces, como X es completo, existe el límite x_0 de la sucesión. Además:

$$F(x_0) = F\left(\lim_{j\to\infty} x_j\right) \overset{cont.}{=} \lim_{j\to\infty} F(x_j) = \lim_{j\to\infty} x_{j+1} = x_0$$



Luego x_0 es efectivamente un punto fijo de la función F. Si $\tilde{x}_0$ fuera otro punto fijo, entonces tendríamos:

$$\rho(x_0, \tilde{x}_0) = \rho(F(x_0), F(\tilde{x}_0)) \leq c \cdot \rho(x_0, \tilde{x}_0)$$

Ya que $0 < c < 1$, las únicas posibilidades son $\rho(x_0, \tilde{x}_0) = 0$ ó $x_0 = \tilde{x}_0$. Esto prueba la existencia y unicidad de x_0.

Falta demostrar la afirmación anterior. Calculamos para $j \geq 1$:

$$\begin{aligned} \rho(x_j, x_{j-1}) &= \rho(F(x_{j-1}), F(x_{j-2})) \leq c \cdot \rho(x_{j-1}, x_{j-2}) = \\ &= c \cdot \rho(F(x_{j-2}), F(x_{j-3})) \leq c^2 \cdot \rho(x_{j-2}, x_{j-3}) \leq \cdots \leq \\ &\leq c^{j-1} \cdot \rho(x_1, x_0) \end{aligned}$$

Como resultado:

$$\begin{aligned} \rho(x_{j+k}, x_j) &\leq \rho(x_{j+k}, x_{j+k-1}) + \rho(x_{j+k-1}, x_{j+k-2}) + \cdots + \rho(x_{j+1}, x_j) \leq \\ &\leq [c^{j+k-1} + c^{j+k-2} + \cdots + c^j] \cdot \rho(x_1, x_0) \leq c^j \cdot \frac{1}{1-c} \cdot \rho(x_1, x_0) \end{aligned}$$

En particular, si $\epsilon > 0$ y si j es lo suficientemente grande, independientemente del valor de $k \geq 1$:

$$\rho(x_{j+k}, x_j) < \epsilon$$

Así, la sucesión $\{x_j\}$ es de Cauchy, y la afirmación quedaría probada.

∎

Teorema (Teorema de la función implícita). Sean $\mathcal{U} \subseteq \mathbb{R}^n$ abierto, $\mathcal{V} \subseteq \mathbb{R}^m$ abierto, y sea $G : \mathcal{U} \times \mathcal{V} \to \mathbb{R}^m$ de clase $\mathcal{C}^1$ tal que $\exists (a,b) \in \mathcal{U} \times \mathcal{V}$ con $G(a,b) = 0$. Sea T_0 la matriz jacobiana de $G(a,\cdot)$ en b invertible. Entonces $\exists\ \mathcal{U}_0 \subset \mathcal{U}$ entorno abierto de a, $\mathcal{V}_0 \subset \mathcal{V}$ entorno abierto de b, tales que $\forall x_0 \in \mathcal{U}_0\ \exists! y \in \mathcal{V}_0$ con $G(x,y) = 0$ y existe una aplicación $u : \mathcal{U}_0 \to \mathcal{V}_0$ de clase $\mathcal{C}^1$ tal que:

(1) $u(a) = b$, $(x, u(x)) \in \mathcal{U} \times \mathcal{V}$

(2) $G(x, u(x)) = 0\ \forall x \in \mathcal{U}$

Demostración:

Definimos la siguiente sucesión de funciones de forma recursiva como:

$$\begin{cases} u_0(x) = y_0 \\ u_n(x) = g(x, u_{n-1}(x)) \ \in \mathcal{C}(\mathcal{U}_0, \mathbb{R}^m) \end{cases}$$

donde $g(x,y) = y - T_0^{-1}(G(x,y))$, la cual cumple que $g(x,y) = y \Leftrightarrow G(x,y) = 0$. Consideramos la siguiente aplicación:

$$H : v \in \mathcal{C}^0(\mathcal{U}_0, \mathbb{R}^m) \to [x \to g(x, v(x))] \in \mathcal{C}^0(\mathcal{U}_0, \mathbb{R}^m)$$

donde un punto fijo de esta aplicación verifica que:

$$Hv = v \Leftrightarrow g(x, v(x)) = v(x) \Leftrightarrow G(x, v(x)) = 0$$

Vamos a ver que H es contractiva si $\mathcal{U}_0$ es lo suficientemente pequeño, es decir:

$\|Hv_1 - Hv_2\|_{\infty,\mathcal{U}_0} \leq c \cdot \|v_1 - v_2\|_{\infty,\mathcal{U}_0}$, con $0 < c < 1$, por tanto:

$$\begin{aligned} \|Hv_1 - Hv_2\|_{\infty,\mathcal{U}_0} &= \sup_{x\in\mathcal{U}_0} |Hv_1(x) - Hv_2(x)|_{\mathbb{R}^m} = \\ &= \sup_{x\in\mathcal{U}_0} |g(x, v_1(x)) - g(x, v_2(x))| \leq (*) \\ &\leq \sup_{x\in\mathcal{U}_0} \frac{1}{2} \cdot |v_1(x) - v_2(x)|_{\mathbb{R}^m} = \\ &= \frac{1}{2} \cdot \|v_1 - v_2\|_{\infty,\mathcal{U}_0} \end{aligned}$$

donde en $(*)$ hemos usado que:

$$d_2g(a,b) = Id_{\mathbb{R}^m} - T_0^{-1}(d_2G(a,b)) = 0$$

y por el teorema de los incrementos finitos, siendo $\mathcal{U}_0$ y $\mathcal{V}_0$ tales que:

$$\|d_2g(x,y)\| \leq \frac{1}{2},\ \forall x \in \mathcal{U}_0,\ \forall y \in \mathcal{V}_0$$

tenemos que existe la función de punto fijo de la transformación.

Para concluir, vemos que, como $G(x, u(x)) = 0$, obtenemos que:

$$\left.\begin{array}{c}
\dfrac{\partial G}{\partial x_1} + \dfrac{\partial G}{\partial x_{n+1}} \cdot \dfrac{\partial u_1}{\partial x_1} + \cdots + \dfrac{\partial G}{\partial x_{n+m}} \cdot \dfrac{\partial u_m}{\partial x_1} = 0 \\[2ex]
\dfrac{\partial G}{\partial x_2} + \dfrac{\partial G}{\partial x_{n+1}} \cdot \dfrac{\partial u_1}{\partial x_2} + \cdots + \dfrac{\partial G}{\partial x_{n+m}} \cdot \dfrac{\partial u_m}{\partial x_2} = 0 \\[2ex]
\vdots \\[2ex]
\dfrac{\partial G}{\partial x_n} + \dfrac{\partial G}{\partial x_{n+1}} \cdot \dfrac{\partial u_1}{\partial x_n} + \cdots + \dfrac{\partial G}{\partial x_{n+m}} \cdot \dfrac{\partial u_m}{\partial x_n} = 0
\end{array}\right\}$$

con lo que obtenemos que, por la regla de Cramer, las parciales: $\dfrac{\partial u}{\partial x_1}, \cdots, \dfrac{\partial u}{\partial x_n}$ pueden expresarse como funciones continuas, con lo que tenemos que $u \in \mathcal{C}^1$.

Faltaría comprobar que dichas parciales de u existen, para lo cual necesitamos el siguiente lema:

Lema:

Sea f un homeomorfismo y f diferenciable en a. Entonces se cumple que:

$$g = f^{-1} \text{ es diferenciable en } b = f(a) \Leftrightarrow f'(a) \in \text{Isomorfismos}$$

Demostración del lema:

$\boxed{\Rightarrow}$: Directo mediante la regla de la cadena.

$\boxed{\Leftarrow}$: Sea $y = f(x)$ con x cercano a a, entonces:

$f(x) - f(a) = y - b = f'(a) \cdot (x - a) + \|x - a\| \cdot \varphi(x)$, donde:

$$\lim_{x \to a} \varphi(x) = 0, \text{ luego:}$$

$$f^{-1}(y) - f^{-1}(b) = x - a = f'^{-1}(a)(y - b) - \|x - a\| \cdot f'^{-1}(a)(\varphi(x))$$

Como tenemos que:

$\|f'^{-1}(a)(y - b)\| \geq \|x - a\|(1 - \psi(x)) \Rightarrow$

$$\Rightarrow \|x - a\| \leq \|y - b\| \cdot \frac{\|f'^{-1}(a)\|}{1 - \psi(x)} \text{ (si } \psi(x) \to 0) \Rightarrow$$

$$\Rightarrow \|x - a\| \cdot \frac{\psi(x)}{\|y - b\|} \to 0 \text{ si } y \to b$$

■

Una vez tenemos el lema comenzamos a probar que existen las parciales de u:

Tenemos que como $G(x, u(x)) = 0$ entonces:

$\frac{\partial G}{\partial y}(x, y)$ es un difeomorfismo de $\mathbb{R}^m$ en el entorno de (a, b), luego:

$F(x, u(x)) = (x, G(x, u(x)))$ es inyectiva y $dF(x, y)$ es un isomorfismo en un entorno de (a, b). Por tanto:

F es un difeomorfismo de $\mathcal{C}^1$, luego F^{-1} es un difeomorfismo de $\mathcal{C}^1$ también, y por el lema anterior quedaría probada la existencia de las parciales de u.

■

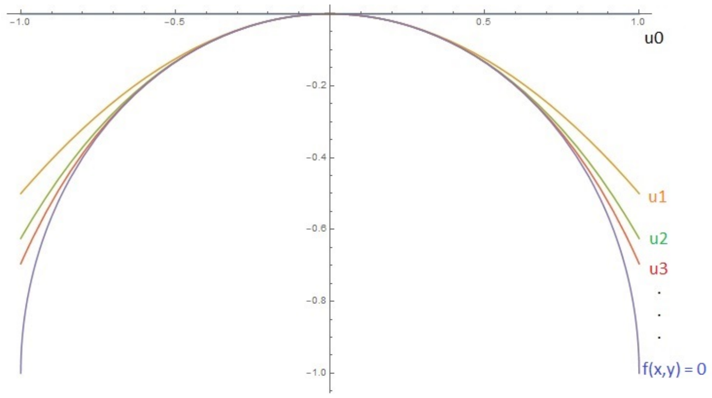
En las siguientes representaciones podemos apreciar el método de las aproximaciones sucesivas.

Se puede apreciar como las funciones van convergiendo a nuestra solución para el caso bidimensional, en el cual tomamos la ecuación:

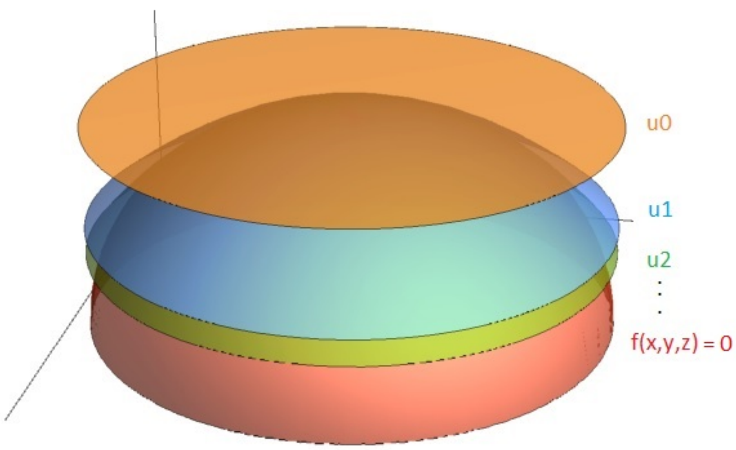
$$f(x, y) = x^2 + (y + 1)^2 - 1$$

y en el caso tridimensional, en el cual tomamos la siguiente función:

$$f(x, y, z) = x^2 + y^2 + (z + 1)^2 - 1:$$



Gráfica en $\mathbb{R}^2$



Gráfica en $\mathbb{R}^3$

3 Aplicaciones

En el estudio del análisis geométrico, a menudo podemos estar interesados en considerar una superficie en el espacio euclídeo.

Todos podemos creer que entendemos intuitivamente el término de “superficie suave en el espacio euclídeo”, pero debemos dar una definición formal y precisa de ello.

Vamos a considerar las siguientes definiciones informales de superficie suave:

- La superficie puede, localmente, enderezarse suavemente.
- La superficie es localmente el conjunto de soluciones de un sistema de ecuaciones suaves.
- La superficie es localmente la gráfica de una función suave.
- La superficie se puede parametrizar suavemente de forma local.

En esta sección enunciaremos de forma precisa estas definiciones y veremos que son equivalentes. Para ello necesitaremos un resultado previo, el cual vemos como consecuencia del teorema de la función inversa, cuyo nombre es el teorema del rango.

3.1. Teorema del rango

El teorema del rango es una consecuencia del teorema de la función implícita, adaptado a las situaciones en las cuales la matriz jacobiana de la función estudiada es de rango constante, pero no rango completo.

Consideremos, por ejemplo, la función $f : \mathbb{R}^3 \to \mathbb{R}^3$ dada por:

$$f(x, y, z) = (x^3 + y + z, y^3 + y, x^3 - y^3 + z) = (F_1, F_2, F_3)$$

Entonces la matriz jacobiana de F es:

$$J_f = \begin{pmatrix} 3x^2 & 1 & 1 \\ 0 & 3y^2 + 1 & 0 \\ 3x^2 & -3y^2 & 1 \end{pmatrix}$$

Ya que las primeras y tercera columnas de J_f son dependientes, el rango de J_f es siempre menor o igual a 2. Por otra parte, las segunda y tercera filas de J_f son independientes, pues

$$\begin{vmatrix} 1 & 1 \\ 3y^2 + 1 & 0 \end{vmatrix} = -(3y^2 + 1) \neq 0$$

Así, J_f tiene rango 2 para cualquier punto.

El teorema del rango nos dirá que, salvo difeomorfismos, f es como la proyección $(x, y, z) \to (0, y, z)$. Lo comprobaremos:

Sea $g(x, y, z) = (x, y^3 + y, x^3 - y^3 + z)$ y calculamos su inversa:

$g^{-1}(x, y, z) = (x', y', z')$ donde:

$$\begin{cases} x' = x \\ y' = u(y') \\ z' = u(y')^3 - x'^3 + z' \end{cases}$$

Por tanto $g^{-1}(x, y, z) = (x, u(y), u(y)^3 - x^3 + z)$, luego:

$f \circ g^{-1}(x, y, z) = (x^3 + u(y) + u(y)^3 - x^3 + z, y, z) = (y + z, y, z)$

Tomamos la función h de la siguiente forma:

$h(x, y, z) = (x - y - z, y, z)$

Se cumple que $h \circ f \circ g^{-1}(x, y, z) = h(y + z, y, z) = (0, y, z)$

Y esto es lo que obtenemos aplicando el teorema.

*Nota: Para poner $(x, y, 0)$ en lugar de $(0, y, z)$ como aparecerá en el enunciado del teorema, solo tenemos que realizar una transformación ortogonal que cambie el orden de las variables, la cual iría dentro del isomorfismo en cuestión.

Realizaremos a continuación un par de observaciones al ejemplo desarrollado anteriormente:

(1) Sea $b = (30, 10, 20) = f(3, 2, 1)$, luego:

$f^{-1}(b) = \{(x, y, z) : f(x, y, z) = (30, 10, 20)\}$, es decir:

$$\left.\begin{array}{c} x^3 + y + z = 30 \\ y^3 + y = 10 \\ x^3 - y^3 + z = 20 \end{array}\right\} \Leftrightarrow \{y = 2; z = 28 - x^3\}$$

Lo que nos da una curva en el plano vertical $y = 2$, luego el punto $(3, 2, 1)$ es un elemento de la variedad:

$f^{-1}(b) = \{(x, 2, z) : z = 28 - x^3\}$

(2) Inversamente, tenemos que:

$f(\mathbb{R}^3) = \{(u, v, w) : u = x^3 + y + z, v = y^3 + y, w = x^3 - y^3 + z\} = \{(u, v, w) : u - v = w\}$, pues si $u - v = w$, donde $y^3 + y = v$ y $x^3 + z = u - y$, entonces:

$$\begin{cases} u = x^3 + z + y \\ w = u - v = (x^3 + z + y) - (y^3 + y) = x^3 - y^3 + z \end{cases}$$

Lo que nos devuelve un plano en el espacio tridimensional.

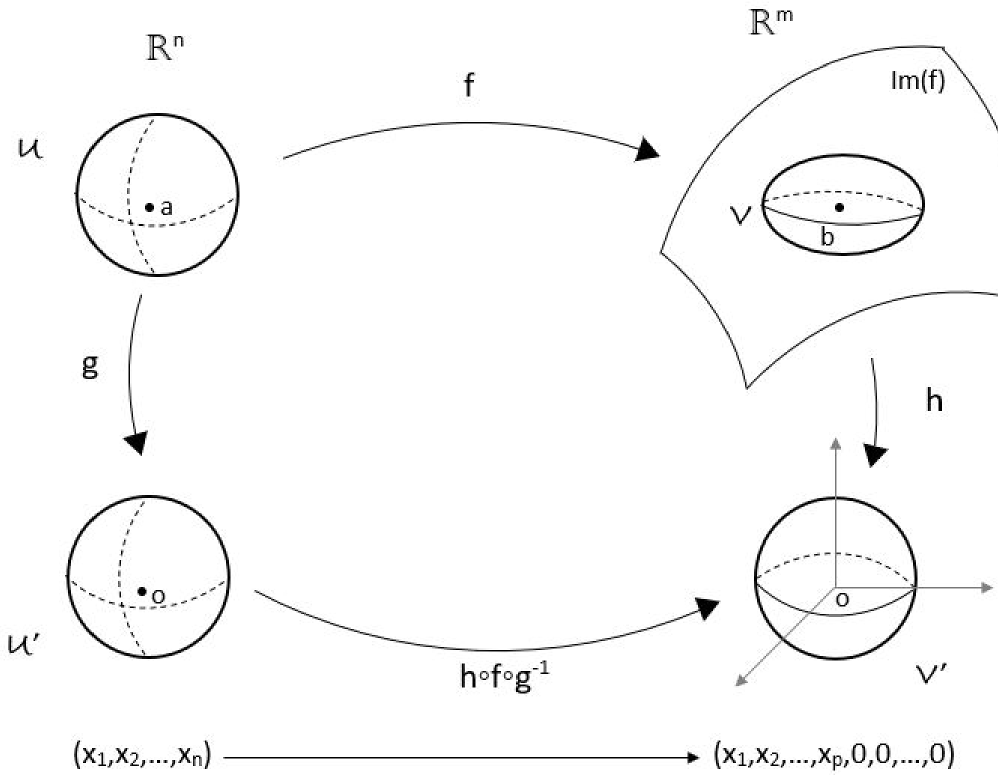
Pasemos entonces, después de ver lo que hace nuestro teorema, a dar un enunciado formal y una prueba de éste:

Teorema (Teorema del rango). Sea $D \subseteq \mathbb{R}^n$ abierto y $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ con $f \in \mathcal{C}^1(D)$. Supongamos que J_f tiene rango p $\forall x \in D$ y para $a \in D$ pongamos $b = f(a) \in \mathbb{R}^m$. Entonces:

(1) Existen conjuntos abiertos $\mathcal{U}$ de a, $\mathcal{U}'$ de $0 \in \mathbb{R}^n$ y un difeomorfismo $g : \mathcal{U} \to \mathcal{U}'$ de clase $\mathcal{C}^1$.

(2) Existen conjuntos abiertos $\mathcal{V}$ de b, $\mathcal{V}'$ de $0 \in \mathbb{R}^m$ y un difeomorfismo $h : \mathcal{V} \to \mathcal{V}'$ de clase $\mathcal{C}^1$.

(3) $h \circ f \circ g^{-1}(x_1, x_2, \cdots, x_n) = (x_1, x_2, \cdots, x_p, 0, \cdots, 0) \in \mathbb{R}^m$



Demostración:

Primero se supondrá $a = 0$ y $f(a) = b = 0$, pues el caso general se obtendrá poniendo $F(x) = f(x+a) - b$.

Debido a que $J_f(x)$ tiene rango p, $\forall x \in D$ entonces la matriz:

$$J_f(0) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \dfrac{\partial f_m}{\partial x_2} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix}$$

tiene rango p. Por lo tanto sabemos que existe una submatriz de tamaño $p \times p$ invertible y que esto no se cumple para submatrices de mayor tamaño. Además se puede suponer que tal submatriz esta colocada en la esquina de la parte superior de $J_f(0)$. De no ser así, podemos permutar las filas y columnas para que esto se pueda tener como muestra la matriz (lo que equivale a una transformación ortogonal de permutar variables):

$$\left(\begin{array}{ccc|c} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_p} & \\ \vdots & \ddots & \vdots & \text{———} \\ \dfrac{\partial f_p}{\partial x_1} & \cdots & \dfrac{\partial f_p}{\partial x_p} & \\ \hline & & & \\ & \text{———} & & \text{———} \\ & & & \end{array}\right)$$

Después de esto definimos $g : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ dada por:

$g(x_1, x_2, \cdots, x_n) = (f_1(x), f_2(x), \cdots, f_p(x), x_{p+1}, \cdots, x_n)$, en donde x es el vector $x = (x_1, x_2, \cdots, x_n)$.

Es claro que $g \in \mathcal{C}^1(D)$ y que $g(0) = 0$, luego si calculamos la matriz jacobiana de g podemos ver que es:

$$J_g(x) = \left(\begin{array}{ccc|cccc} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_p} & & & & \\ \vdots & \ddots & \vdots & & & \text{---} & \\ \dfrac{\partial f_p}{\partial x_1} & \cdots & \dfrac{\partial f_p}{\partial x_p} & & & & \\ \hline 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{array}\right)$$

la cual es invertible en $x = 0$ y por el teorema de la función inversa tenemos que existen abiertos $\mathcal{U}$ de $0, \mathcal{U}'$ de $g(0) = 0$ y un difeomorfismo $g : \mathcal{U} \to \mathcal{U}'$ de clase $\mathcal{C}^1$ en donde además $g(\mathcal{U}) = \mathcal{U}'$. Esto prueba (1).

Pongamos ahora en las coordenadas de g que:

$$(y_1, y_2, \cdots, y_n) = g(x_1, x_2, \cdots, x_n) = (f_1(x), f_2(x), \cdots, f_p(x), x_{p+1}, \cdots, x_n)$$

y observemos que g^{-1} funciona como la inversa de f en las primeras p coordenadas, pues, g y f son iguales en esas componentes. De aquí que tengamos la siguiente igualdad:

$$\begin{aligned} f \circ g^{-1}(y_1, y_2, \cdots, y_n) &= (y_1, y_2, \cdots, y_p, f_{p+1} \circ g^{-1}(y), \cdots, f_m \circ g^{-1}(y)) = \\ &= (y_1, y_2, \cdots, y_p, \psi_{p+1}(y), \cdots, \psi_m(y)) \end{aligned}$$

en donde $y = (y_1, y_2, \cdots, y_n) \in \mathcal{U}'$.

Si calculamos la matriz jacobiana de esta composición se puede ver que:

$$J_{[f\circ g^{-1}]}(y) = \left(\begin{array}{c|ccc} & 0 & \cdots & 0 \\ I_p & \vdots & \ddots & \vdots \\ & 0 & \cdots & 0 \\ \hline & \dfrac{\partial \psi_{p+1}}{\partial y_{p+1}} & \cdots & \dfrac{\partial \psi_{p+1}}{\partial y_n} \\ \text{———} & \vdots & \ddots & \vdots \\ & \dfrac{\partial \psi_m}{\partial y_{p+1}} & \cdots & \dfrac{\partial \psi_m}{\partial y_n} \end{array}\right)$$

Ahora, el rango de esta matriz debe ser p lo cual implica que todas las parciales deben ser cero, ya que si para algún $y \in \mathcal{U}'$ al menos una de ellas es distinta de cero se tendría una submatriz de tamaño $(p+1)\times(p+1)$ lo cual sería contrario a nuestra hipótesis. Por lo tanto las funciones $\psi_{p+1}\cdots\psi_m$ no dependen de las variables $y_{p+1},\cdots,y_m$ y en base a esto definimos $h:\mathbb{R}^m \to \mathbb{R}^m$ dado por:

$$h(z_1,z_2,\cdots,z_m) = (z_1,z_2,\cdots,z_p,z_{p+1}-\psi_{p+1}(z),\cdots,z_m-\psi_m(z))$$

donde $z=(z_1,z_2,\cdots,z_m)$. Así, la matriz jacobiana de h es:

$$J_h(z) = \left(\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \hline & & & & 1 & 0 & \cdots & 0 \\ & & \text{———} & & 0 & 1 & \cdots & 0 \\ & & & & \vdots & \vdots & \ddots & \vdots \\ & & & & 0 & 0 & \cdots & 1 \end{array}\right)$$

la cual es obviamente invertible, y nuevamente por el teorema de la función inversa obtenemos que existen conjuntos abiertos $\mathcal{V}$, $\mathcal{V}'$ de 0 y $h(0) = 0$ en $\mathbb{R}^m$, respectivamente, tales que la función $h : \mathcal{V} \to \mathcal{V}'$ es un difeomorfismo de clase $\mathcal{C}^1$, en donde además $h(\mathcal{V}) = \mathcal{V}'$. Esto prueba (2).

Por último, haciendo la composición de $f \circ g^{-1}$ con h obtenemos:

$$h \circ f \circ g^{-1}(y_1, y_2, \cdots, y_n) = (y_1, y_2, \cdots, y_p, \psi_{p+1} - \psi_{p+1}, \cdots, \psi_m - \psi_m) =$$
$$= (y_1, y_2, \cdots, y_p, 0, 0, \cdots, 0) \in \mathbb{R}^m, \ \forall y \in \mathcal{U}'.$$

■

3.2. Superficies

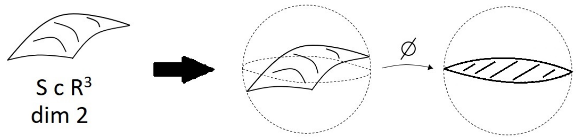
Ahora sí estamos preparados para poder probar el teorema de equivalencia de todas las definiciones de superficies.

El teorema siguiente expresa de una forma precisa cada una de estas descripciones y nos prueba su equivalencia de forma cíclica.

Teorema. Sean n, m enteros con $1 \leq n < m$, sea $k \geq 1$ otro entero ó $+\infty$ y sea S un subconjunto de $\mathbb{R}^m$. Son equivalentes:

(1) (La superficie puede, localmente, enderezarse suavemente). Para cada punto $a \in S$ existe un entorno abierto $\mathcal{V} \subset \mathbb{R}^m$ de a, un difeomorfismo $\phi : \mathcal{V} \to \mathbb{R}^m \in \mathcal{C}^k$ y un subespacio vectorial n-dimensional $L \subset \mathbb{R}^m$ tal que:

$\phi(S \cap \mathcal{V}) = L \cap \phi(\mathcal{V})$



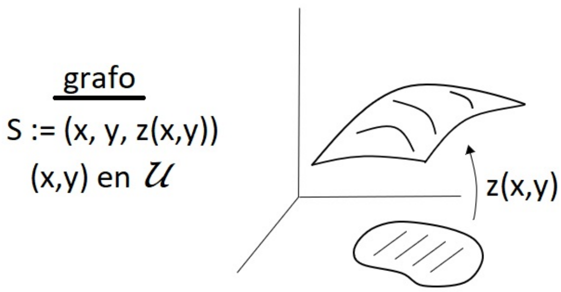
(2) (La superficie es localmente el conjunto de soluciones de un sistema de ecuaciones suaves). Para cada punto $a \in S$ existe un entorno abierto $\mathcal{V} \subset \mathbb{R}^m$ de a y una función de clase $\mathcal{C}^k$ $g : \mathcal{V} \to \mathbb{R}^{m-n}$ tal que:

$S \cap \mathcal{V} = \{x \in \mathcal{V} : g(x) = 0\}$, y:

$\text{rango}(J_g(x)) = m - n \ \forall x \in \mathcal{V}$

(3) (La superficie localmente es la gráfica de una función suave). Para cada punto $a \in S$ existe un entorno $\mathcal{V} \subset \mathbb{R}^m$ de a tal que $S \cap \mathcal{V}$ es la gráfica de una función de clase $\mathcal{C}^k$. Es decir, existe una permutación de coordenadas $\Phi : \mathbb{R}^m \to \mathbb{R}^m$, un conjunto abierto $\mathcal{U} \subseteq \mathbb{R}^n$ y una función $F : \mathcal{U} \to \mathbb{R}^{m-n}$ tal que:

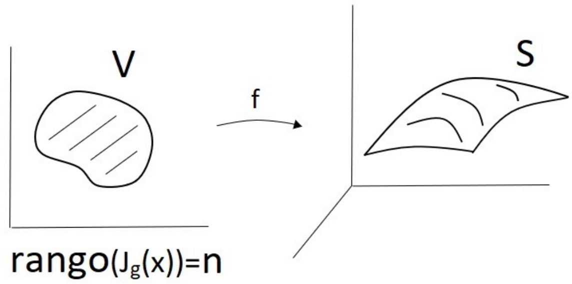
$S \cap \mathcal{V} = \Phi\left(\{(u, F(u)) : u \in \mathcal{U} \cap \mathbb{R}^n\}\right)$



(4) (Localmente, la superficie se puede parametrizar suavemente). Para cada punto $a \in S$ existe un entorno abierto $\mathcal{V}$ de a, existe un entorno abierto $\mathcal{U} \subset \mathbb{R}^n$ y una función de clase $\mathcal{C}^k$ $f : \mathcal{U} \to \mathbb{R}^m$ inyectiva, tales que:

$f(\mathcal{U}) = S \cap \mathcal{V}$, y

$\text{rango}(J_f(x)) = n \ \forall x \in \mathcal{U}$



Demostración:

(1)⇒(2) Sean $\mathcal{V}$ y ϕ como en (1). Elegimos un conjunto ortonormal de $m - n$ vectores $v_1, \cdots, v_{m-n}$, todos ortogonales a L. Definimos $g : \mathcal{V} \to \mathbb{R}^{m-n}$ como:

$$g(x) = (v_1 \cdot \phi(x), v_2 \cdot \phi(x), \cdots, v_{m-n} \cdot \phi(x)) \in \mathcal{C}^k$$

Se cumple que $\phi(S \cap \mathcal{V}) = L \cap \phi(\mathcal{V})$:

i) Si $x \in S \cap \mathcal{V}$, entonces:

$$\left.\begin{array}{r} \phi(x) \in L \\ v_k \in L^\perp \end{array}\right\} \Rightarrow g(x) = 0$$

ii) Si $x \in \mathcal{V}$, $g(x) = 0$, entonces:

$$\phi(x) \in < v_1, \cdots, v_{m-n} >^\perp = L \Rightarrow \phi(x) \in L \cap \phi(\mathcal{V}) = \phi(S \cap \mathcal{V}) \Rightarrow x \in S \cap \mathcal{V}$$

Como g es composición de:

$x \to \phi(x)$ de $\mathbb{R}^m \to \mathbb{R}^m$

$y \to (v_1 \cdot y, \cdots, v_{m-n} \cdot y)$ de $\mathbb{R}^m \to \mathbb{R}^{m-n}$

Tenemos que la primera matriz $J_\phi(x)$ es $m \times m$ invertible, luego, aplicando la regla de la cadena, obtenemos que la segunda matriz J_g tiene rango $m-n$, ya que sus filas son los vectores $v_1, \cdots, v_{m-n}$:

$$J_g = \begin{pmatrix} — & v_1 & — \\ & \vdots & \\ — & v_{m-n} & — \end{pmatrix}$$

(2)$\Rightarrow$(3) Ahora $S \cap \mathcal{V} = \{x \in \mathcal{V} : g(x) = 0\}$, donde $g : \mathcal{V} \to \mathbb{R}^{m-n}$, la cual tiene rango $m-n$ en $\mathcal{V}$.

Como rango$(J_g(a)) = m-n$, por el teorema de la función implícita tenemos que $m-n$ variables son función de las otras n, por tanto:

$$\left.\begin{array}{c} \exists \mathcal{V}_0 \subset \mathbb{R}^{m-n}, \text{ con } (a_{l_1}, \cdots, a_{l_{m-n}}) \in \mathcal{V}_0 \\ \exists \mathcal{U} \subset \mathbb{R}^n, \text{ con } (a_{j_1}, \cdots, a_{j_n}) \in \mathcal{U} \\ \exists F : U \to \mathcal{V}_0 \in \mathcal{C}^k \end{array}\right\} \text{tales que :}$$

$\forall x \in \mathcal{U} \times \mathcal{V}_0$, donde $\{l_1, \cdots, l_{m-n}, j_1, \cdots, j_n\} = \{1, \cdots, m\}$

Tenemos que:

$g(x) = 0 \Leftrightarrow (x_{l_1}, \cdots, x_{l_{m-n}}) = F(x_{j_1}, \cdots, x_{j_n})$, por tanto:

$x \in S \cap (\mathcal{U} \times \mathcal{V}_0) \Leftrightarrow x = \Phi(x_{j_1}, \cdots, x_{j_n}, f_1(x_{j_1}, \cdots, x_{j_n}), \cdots, f_{m-n}(x_{j_1}, \cdots, x_{j_n}))$

donde Φ es una permutación de índices. Luego obtenemos que:

$S \cap (\mathcal{U} \times \mathcal{V}_0) = \Phi(\{(u, F(u)) : u \in \mathcal{U}\})$

(3)⇒(4) Tenemos que $\forall a \in S$:

$\exists \mathcal{V} \subset \mathbb{R}^m$, con $a \in \mathcal{V}$

$\exists \mathcal{U} \subset \mathbb{R}^n$, abierto

$\exists F : U \to \mathbb{R}^{m-n} \in \mathcal{C}^k$ y una permutación de índices Φ tal que:

$S \cap \mathcal{U} = \Phi(\{(u, F(u)) : u \in \mathcal{U}\})$, donde $f(\mathcal{U}) = s \cap \mathcal{U}$

Sea $f(u) = \Phi(u, F(u))$ de clase $\mathcal{C}^k$ e inyectiva. Como:

$$J_f(u) = J_\Phi(u, F(u)) = \left(\begin{array}{ccc|c} 1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & J_F(u) \\ 0 & \cdots & 1 & \end{array}\right)$$

f es de rango n.

(4)⇒(1) Si $b \in \mathcal{U}$ con $f(b) = a$, Por el teorema del rango el siguiente diagrama:

$$\begin{tikzcd} b \in \mathcal{U}' \subset \mathcal{U} \subset \mathbb{R}^n \arrow[r, "f"] \arrow[d, "g"] & a \in \mathcal{V}' \subset \mathcal{V} \subset \mathbb{R}^m \arrow[d, "h"] \\ 0 \in \mathcal{U}'' \subset \mathbb{R}^n \arrow[r, "\text{Inclusión}"] & 0 \in \mathcal{V}'' \subset \mathbb{R}^m \end{tikzcd}$$

donde g y h son difeomorfismos de clase $\mathcal{C}^k$ tales que:

$h \circ f \circ g^{-1}(x_1 \cdots, x_n) = (x_1, \cdots, x_n, 0, \cdots, 0)$.

$$\left.\begin{array}{c} \phi \text{ es } h \\ L \text{ es } < e_1, \cdots, e_n > \subset \mathbb{R}^m \\ \mathcal{V} \text{ de (1) es } \mathcal{V}' \text{ aquí} \end{array}\right\}$$

donde $< e_1, \cdots, e_n >$ es la base canónica.

Se tiene que:

$$y \in \phi(S \cap \mathcal{V}') \Leftrightarrow \exists x \in S \cap \mathcal{V}' \text{ tal que } \phi(x) = y$$

Pero $x \in S$ y $x \in \mathcal{V}'$, luego por (4) tenemos que existe un único $u \in \mathcal{U}$ tal que $f(u) = x$, por tanto:

$$y = \phi(x) = \phi(f(u)) = h(f(u)) = h \circ f \circ g^{-1}(g(u)) \in L, \text{ luego } y \in L \cap \phi(S)$$

■

3.3. Relación con las ecuaciones diferenciales

Hay una conexión entre el teorema de la función implícita y la teoría de ecuaciones diferenciales ordinarias.
Incluso desde el punto de vista histórico podemos apreciar dicha relación, debido a que la prueba iterativa del teorema de existencia y unicidad de ecuaciones diferenciales ordinarias de Picard inspiró a Goursat a dar su prueba iterativa del teorema de la función implícita (ver Goursat [3]).

A mediados del siglo XX, John Nash inició el uso de una forma sofisticada del teorema de la función implícita en ecuaciones en derivadas parciales, pero dicho tema no lo vamos a abordar en este trabajo.

En esta sección mostraremos cómo un teorema de existencia de soluciones de ecuaciones diferenciales ordinarias clásico (teorema de Cauchy-Peano) puede ser usado para probar nuestro teorema de la ecuación implícita.

Aunque en el libro de Krantz y Parks (ver Krantz [5])se afirma que una versión para espacios de Banach del teorema de la función implícita puede ser usada para probar el teorema de Cauchy-Peano, hemos encontrado ciertas dificultades en su demostración por lo que hemos optado finalmente por no inluirlo en esta memoria.

El teorema de Cauchy-Peano tiene el siguiente enunciado:

Teorema (Teorema de Cauchy-Peano). Si $F(t,x)$ con $(t,x) \in \mathbb{R}\times\mathbb{R}^N$ es continua en la región $(N+1)$dimensional $(t_0 - a, t_0 + a) \times B(x_0, r)$, entonces existe una solución $x(t)$ de:

$$\begin{cases} \dfrac{dx}{dt} = F(t,x) \\ x(t_0) = x_0 \end{cases}$$

definido en un intervalo $(t_0 - h, t_0 + h)$.

Observación. La solución de

$$\begin{cases} \dfrac{dx}{dt} = F(t,x) \\ x(t_0) = x_0 \end{cases}$$

no necesariamente es única solo si F es continua. Por ejemplo, el problema de encontrar $x(t)$ que satisfaga el problema:

$$\begin{cases} x' = x^{2/3} \\ x(0) = 0 \end{cases}$$

tiene dos soluciones $x \equiv 0$ y $x(t) = \dfrac{t}{3}^3$.
Luego para garantizar la unicidad de solución, necesitamos añadir una condición adicional a F, la cual habitualmente es que sea localmente lipschitziana como función de x, la cual añadida como condición al teorema da lugar al teorema de

existencia y unicidad de Picard.

Pasamos, pues, a dar la prueba alternativa del teorema de la función implícita, el cual lo reformularemos para una mayor facilidad de demostración, siendo así una consecuencia del teorema de Cauchy-Peano.

Teorema (Teorema de la función implícita) Supongamos $\mathcal{U} \subset \mathbb{R}^{N+1}$ y $f : \mathcal{U} \to \mathbb{R} \in \mathcal{C}^1$. Si $f(t_0, x_0) = 0$, $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^N$, y la matriz $N \times N$:

$$\left(\frac{\partial f_i}{\partial x_j}(t_0, x_0) \right)_{i,j=1,2,\cdots,N}$$

es no singular, entonces existe un intervalo abierto $(t_0 - h, t_0 + h)$ y una función de clase $\mathcal{C}^1$ $u : (t_0 - h, t_0 + h) \to \mathbb{R}^N$ tal que:

(1) $u(t_0) = x_0$

(2) $f(t, u(t)) = 0$, $\forall t \in (t_0 - \delta, t_0 + \delta)$

Demostración:

Consideremos en detalle el caso $N = 1$:
Primero elegimos $a, r > 0$, luego $(t_0 - a, t_0 + a) \times (x_0 - r, x_0 + r) \subset \mathcal{U}$, y $\dfrac{\partial f}{\partial x}(t, x)$ es no nula en $(t_0 - a, t_0 + a) \times (x_0 - r, x_0 + r)$.

Entonces definimos $F : (t_0 - a, t_0 + a) \times (x_0 - r, x_0 + r) \to \mathbb{R}$ como:

$$F(t, x) = \frac{-\dfrac{\partial f}{\partial t}(t, x)}{\dfrac{\partial f}{\partial x}(t, x)}$$

Como F es continua, podemos aplicar el teorema de Cauchy-Peano para concluir que existe una solución del problema:

$$\begin{cases} \dfrac{dx}{dt} = F(t,x) \\ x(t_0) = x_0 \end{cases}$$

definido en un intervalo $(t_0 - h, t_0 + h)$.

Sea $\varphi : (t_0 - h, t_0 + h) \to \mathbb{R}$ la solución. Se cumple:

$$\begin{cases} \varphi'(t) = \dfrac{-\dfrac{\partial f}{\partial t}(t,\varphi(t))}{\dfrac{\partial f}{\partial x}(t,\varphi(t))} \\ \varphi(t_0) = x_0 \end{cases}$$

Como $\varphi(t_0) = x_0$, tenemos que $f(t,\varphi(t)) = f(t_0,x_0) = 0$, y por la regla de la cadena tenemos que:

$$\frac{d}{dt}f(t,\varphi(t)) = \frac{\partial f}{\partial t}(t,\varphi(t)) + \frac{\partial f}{\partial x}(t,\varphi(t)) \cdot \varphi'(t) = 0$$

Luego $f(t,\varphi(t)) = 0$ es constante en $(t_0 - h, t_0 + h)$, y como en t_0 se anula, tenemos que:

$f(t,\varphi(t)) = 0$ en el intervalo.

Para el caso $N > 1$, elegimos $a, r > 0$ tales que $(t_0 - a, t_0 + a) \times B(x_0, r) \subseteq \mathcal{U}$, y por lo tanto tenemos que la matriz $N \times N$:

$$d_x f = \left(\frac{\partial f_i}{\partial x_j}\right)_{i,j=1,\cdots,N} \quad \text{es no singular en } (t_0 - a, t_0 + a) \times B(x_0, r).$$

Reemplazando la definición de $F(t,x)$ por:

$$F(t,x) = -[d_x H(t + t_0, x + x_0)]^{-1}\left(\frac{\partial H}{\partial t}(t + t_0, x + x_0)\right)$$

obtenemos que la prueba es equivalente al caso $N = 1$.

$\blacksquare$

Observación. La prueba del teorema de la función implícita dada anteriormente está limitada al caso de una variable independiente en la función definida implícitamente. El caso de una variable dependiente y varias independientes se puede obtener reemplazando la definición de $F(t, x)$ de la demostración por un sistema apropiado de ecuaciones en derivadas parciales de primer orden, y el caso general (varias variables dependientes y varias independientes) puede obtenerse usando el procedimiento de inducción de Dini visto en la primera demostración.

Bibliografía

[1] Henri Cartan, *Cálculo Diferencial*, Hermann, 1967.

[2] Jean Dieudonné, *Fundamentos de Análisis Moderno*, Reverté, 1969.

[3] Edouard Jean Baptiste Goursat, *Sur la théorie des fonctions implicites*, Bulletin de la Société Mathématique de France 31, 1903.

[4] Steven G. Krantz and Harold R. Parks, *A Primer of Real Analytic Functions*, Springer Science & Business Media, 2002.

[5] Steven G. Krantz and Harold R. Parks, *The Implicit Function Theorem. History, theory and applications*, Springer Science & Business Media, 2012.

[6] J.J. O'Connor and E.F. Robertson, *The Function Concept*, MacTutor History of Mathematics Archive, 2005.

[7] W. Rudin, *Principios de Análisis Matemático*, McGraw-Hill, 1980.