Teorema de la función implícita y Teorema de la función Inversa

Como ya lo habíamos comentado anteriormente este documento es un pequeño resumen del capítulo 5 del libro *Cálculo diferencial de varias variables* del Profesor Javier Páez Cárdenas y tiene como objetivo, además de concluir el temario del curso, enunciar y mostrar cómo se utilizan los principales resultados de este curso, es decir, el Teorema de la función implícita y el Teorema de la función Inversa. Es por esta razón que no incluiremos las demostraciones, pero los invitamos a consultarlas en el libro y si les surgen dudas con mucho gusto los apoyamos.

1. La derivada de funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$

Comenzamos recordando la definición de derivada de funciones de $\mathbb{R}^n$ en $\mathbb{R}$.

Definición 1 *Sean $U \subseteq \mathbb{R}^n$ un conjunto abierto, $\overline{x}_0 \in U$ y $f : U \longrightarrow \mathbb{R}$ una función. Diremos que f es derivable en $\overline{x}_0$ si existe una función lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}$ tal que*

$$\lim_{x \to \overline{x}_0} \frac{f(\overline{x}) - [L(\overline{x} - \overline{x}_0) + f(\overline{x}_0)]}{\|\overline{x} - \overline{x}_0\|} = 0$$

En este caso llamaremos a la función L la derivada de f en $\overline{x}_0$ y la denotaremos por $Df(\overline{x}_0)$.

La definición de derivada de funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$ está basada en la Definición 1 y contiene la misma idea de encontrar la mejor aproximación lineal a f "cerca" de un punto $\overline{x}_0$.

Definición 2 *Sean $U \subseteq \mathbb{R}^n$ un conjunto abierto, $\overline{x}_0 \in U$ y $f : U \to \mathbb{R}^m$ una función. Decimos que f es derivable en $\overline{x}_0$ si existe una función lineal $L : \mathbb{R}^n \to \mathbb{R}^m$ tal que*

$$\lim_{\overline{x} \to \overline{x}_0} \frac{f(\overline{x}) - [L(\overline{x} - \overline{x}_0) + f(\overline{x}_0)]}{\|\overline{x} - \overline{x}_0\|} = \overline{0}.$$

Ahora, recuerde que si $U \subset \mathbb{R}^n$ y $f : U \longrightarrow \mathbb{R}^m$, entonces podemos pensar a f en términos de sus funciones coordenadas, es decir, $f = (f_1, ..., f_n) : U \longrightarrow \mathbb{R}^n$, donde $f_i : U \longrightarrow \mathbb{R}$ son las llamadas funciones coordenadas y $f(\overline{x}) = (f_1(\overline{x}), ..., f_n(\overline{x}))$ para cada $\overline{x} \in U$. Como cada función coordenada es una función de $\mathbb{R}^n$ en $\mathbb{R}$, podemos utilizar toda la teoría que hemos desarrollado para estas funciones.

Proposición 3 *Sea $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$ y $f = (f_1, ..., f_n) : U \longrightarrow \mathbb{R}$ una función. Se tiene que f es derivable en $\overline{x}_0$ si y sólo si f_i es derivable en $\overline{x}_0$ para cada $i \in \{1, ..., n\}$.*

Como sabemos que, para funciones de $\mathbb{R}^n$ en $\mathbb{R}$, la función lineal L de la Definión 1 es única, entonces, por la proposición anterior, tenemos el siguiente resultado.

Proposición 4 *Sean $U \subseteq \mathbb{R}^n, \overline{x}_0 \in U$ y $f : U \longrightarrow \mathbb{R}^m$ una función derivable en $\overline{x}_0$. Entonces existe una única función lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ que satisface la Definición 2.*

Dada la proposición anterior, llamaremos a la función L de la Definición 2 la derivada de f en $\overline{x}_0$ y la denotaremos por $Df(\overline{x}_0)$.

Antes de continuar debemos recordar que toda función lineal $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ se puede representar por una matriz de $m \times n$ con entradas reales, que depende de las bases de $\mathbb{R}^n$ y $\mathbb{R}^m$, como sigue:

Sean $\{\overline{e}_1, ..., \overline{e}_n\}$ y $\{\overline{\varepsilon}_1, ..., \overline{\varepsilon}_m\}$ dos bases ortonormales de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente. Sea $\overline{x} \in \mathbb{R}^n$, entonces

$$\begin{aligned} \overline{x} &= (x_1, ..., x_n) \\ &= x_1\overline{e}_1 + \cdots + x_n\overline{x}_n, \end{aligned}$$

luego

$$\begin{aligned} L(\overline{x}) &= L(x_1\overline{e}_1 + \cdots + x_n\overline{x}_n) \\ &= x_1 L(\overline{e}_1) + \cdots + x_n L(\overline{x}_n). \end{aligned}$$

Ahora, como cada $L(\overline{e}_i) \in \mathbb{R}^m$, entonces

$$L(\overline{e}_i) = a_{1i}\overline{\varepsilon}_1 + \cdots + a_{mi}\overline{\varepsilon}_m$$

para cada $i \in \{1, ..., n\}$. Así,

$$\begin{aligned} L(\overline{x}) &= x_1 L\left(\overline{e}_1\right) + \cdots + x_n L\left(\overline{e}_n\right) \\ &= x_1\left(a_{11}\overline{\varepsilon}_1 + \cdots + a_{m1}\overline{\varepsilon}_m\right) + \\ &\qquad\qquad \vdots \\ &+ x_n\left(a_{1n}\overline{\varepsilon}_1 + \cdots + a_{mn}\overline{\varepsilon}_m\right) \\ &= (x_1, \ldots, x_n) \cdot (a_{11}, \ldots, a_{1n})\, \overline{\varepsilon}_1 + \\ &\qquad\qquad \vdots \\ &+ (x_1, \ldots, x_n) \cdot (a_{m1}, \ldots, a_{mn})\, \overline{\varepsilon}_m. \end{aligned}$$

Así, en términos de matrices, tenemos que

$$L(\hat{x}) = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

De esta manera, tenemos que la matriz

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

representa a la función L en las bases $\{\overline{e}_1, ..., \overline{e}_n\}$ y $\{\overline{\varepsilon}_1, ..., \overline{\varepsilon}_m\}$. Ahora, note que la matriz

$$\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix}$$

representa a la función coordenada $L_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, para cada $i \in \{1, ..., m\}$, en la base $\{\overline{\varepsilon}_1, ..., \overline{\varepsilon}_m\}$. Así, si $f = (f_1, ..., f_m) : U \longrightarrow \mathbb{R}$ es una función derivable en $\overline{x}_0 \in U$, tenemos que

$$\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_i}{\partial x_1}(\overline{x}_0) & \cdots & \dfrac{\partial f_i}{\partial x_n}(\overline{x}_0) \end{bmatrix}$$

y de aquí que la derivada de f en $\overline{x}_0$ está representada por la matriz

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(\overline{x}_0) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\overline{x}_0) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(\overline{x}_0) & \cdots & \dfrac{\partial f_m}{\partial x_n}(\overline{x}_0) \end{bmatrix}.$$

A esta matriz se le conoce como **matriz jacobiana** y con cierto abuso de notación escribiremos que

$$Df(\overline{x}_0) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(\overline{x}_0) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\overline{x}_0) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(\overline{x}_0) & \cdots & \dfrac{\partial f_m}{\partial x_n}(\overline{x}_0) \end{bmatrix}.$$

Ejemplo 5 *Considere la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ dada por*

$$f(x,y) = (r\cos(x)\operatorname{sen}(y), r\operatorname{sen}(x)\operatorname{sen}(y), r\cos(y)),$$

donde r es un número real positivo fijo. Demuestra que f es derivable en cada punto $(x,y) \in \mathbb{R}^2$ y calcule la matriz jacobiana

Solución. Note que las funciones coordenadas $f_1, f_2, f_3 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ están dadas por

$$f_1(x,y) = r\cos(x)\operatorname{sen}(y), \qquad f_2(x,y) = r\operatorname{sen}(x)\operatorname{sen}(y) \qquad \text{y} \qquad f_3(x,y) = r\cos(y)$$

y que cada una de estas es derivable en cualquier punto de $\mathbb{R}^2$. Así, por la Proposición 3, f es derivable en cualquier punto de $\mathbb{R}^2$. Luego

$$\begin{aligned} Df(x,y) &= \begin{bmatrix} \dfrac{\partial f_1}{\partial x}(x,y) & \dfrac{\partial f_1}{\partial y}(x,y) \\ \dfrac{\partial f_2}{\partial x}(x,y) & \dfrac{\partial f_2}{\partial y}(x,y) \\ \dfrac{\partial f_3}{\partial x}(x,y) & \dfrac{\partial f_3}{\partial y}(x,y) \end{bmatrix} \\ &= \begin{bmatrix} -r\operatorname{sen}(x)\operatorname{sen}(y) & r\cos(x)\cos(y) \\ r\cos(x)\operatorname{sen}(y) & r\operatorname{sen}(x)\cos(y) \\ 0 & -r\operatorname{sen}(y) \end{bmatrix} \end{aligned}$$

■

Las demostraciones de las siguientes proposiciones son consecuencia de la Proposición 3.

Proposición 6 *Sean $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$ y $f : U \to \mathbb{R}^m$ una función. Si f es derivable en $\overline{x}_0$, entonces f es continua en $\overline{x}_0$.*

Proposición 7 *Sean $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$ y $f, g : U \to \mathbb{R}^m$ dos funciones. Si f y g son derivables en $\overline{x}_0$, entonces:*

(1) $f + g$ es derivable en $\overline{x}_0$ y

$$D(f+g)(\overline{x}_0) = Df(\overline{x}_0) + Dg(\overline{x}_0)$$

(2) Si $\lambda \in \mathbb{R}$, λf es derivable en $\overline{x}_0$ y

$$D(\lambda f)(\overline{x}_0) = \lambda Df(\overline{x}_0)$$

(3) $f \cdot g$ es derivable en $\overline{x}_0$ y

$$D(f \cdot g)(\overline{x}_0) = \begin{bmatrix} g_1(\overline{x}_0) & \cdots & g_m(\overline{x}_0) \end{bmatrix} Df(\overline{x}_0) + \begin{bmatrix} f_1(\overline{x}_0) & \cdots & f_m(\overline{x}_0) \end{bmatrix} Dg(\overline{x}_0)$$

La siguiente proposición es la versión más general de la regla de la cadena de este curso.

Proposición 8 (Regla de la Cadena) *Sean $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$, $V \subseteq \mathbb{R}^m$ y $f : U \longrightarrow \mathbb{R}^m$, $g : V \longrightarrow \mathbb{R}^k$ dos funciones tales que $f(U) \subseteq V$. Si f es derivable en $\overline{x}_0$ y g es derivable en $\overline{y}_0 = f(\overline{x}_0)$, entonces $g \circ f$ es derivable en $\overline{x}_0$ y*

$$D(g \circ f)(\overline{x}_0) = Dg(f(\overline{x}_0)) \circ Df(\overline{x}_0)$$

y en términos de matrices (con su respectivo abuso de notación)

$$D(g \circ f)(\hat{x}_0) = Dg(f(\hat{x}_0))\, Df(\hat{x}_0).$$

Las proposiciones anteriores nos proporcionan propiedades entre funciones derivables y la siguiente nos proporciona una condición suficiente para que una función sea derivable en un punto y, una vez más, su demostración es consecuencia de la Proposición 3.

Proposición 9 *Sean $U \subseteq \mathbb{R}^n$, $\overline{x}_0 \in U$, $r > 0$ tal que $B_r(\overline{x}_0) \subseteq U$ y $f = (f_1, ..., f_n) : U \longrightarrow \mathbb{R}^m$ una función. Si $\dfrac{\partial f_j}{\partial x_i}(\overline{x})$ existe para toda $\overline{x} \in B_r(\overline{x}_0)$ y $\dfrac{\partial f_j}{\partial x_i}$ es continua en $\overline{x}_0$, para cada $i \in \{1, ..., n\}$ y para cada $j \in \{1, ..., m\}$, entonces f es derivable en $\overline{x}_0$.*

Igual que en funciones de $\mathbb{R}^n$ en $\mathbb{R}$ tenemos el concepto de funciones de clase C^k.

Definición 10 *Sean $U \subseteq \mathbb{R}^n$ y $f = (f_1, ..., f_m) : U \longrightarrow \mathbb{R}^m$ una función. Diremos que f es de clase C^k en U, denotado por $f \in C^k(U)$, si $f_j \in C^k(U)$ para cada $j \in \{1, ..., m\}$. Es decir, si existen todas las derivadas parciales de orden k de f_j en cada punto de U y además estas derivadas parciales son continuas en U, para cada $j \in \{1, ..., m\}$.*

El siguiente resultado es consecuencia de la teoría para funciones de $\mathbb{R}^n$ en $\mathbb{R}$ y de la Proposición 3.

Proposición 11 *Sean $U \subseteq \mathbb{R}^n$ y $f : U \longrightarrow \mathbb{R}^m$ una función. Si $f \in C^1(U)$, entonces f es derivable en cada $\overline{x} \in U$.*

2. Superficies

Recuerde que las funciones de $\mathbb{R}$ en $\mathbb{R}^n$ se utilizaron para definir el concepto de curva, de manera análoga utilizamos las funciones de $\mathbb{R}^2$ en $\mathbb{R}^3$ para definir el concepto de superficie.

Definición 12 *Sea $S \subseteq \mathbb{R}^3$. Diremos que S es una superficie si existen $U \subseteq \mathbb{R}^2$ conjunto abierto, $A \subseteq U$ y $\sigma : U \longrightarrow \mathbb{R}^3$ derivable en cada punto de U tales que $\sigma(A) = S$. En este caso diremos que σ es una parametrización de S.*

Ejemplo 13 *Sean $a, b, c \in \mathbb{R}^+$, considere el conjunto*

$$\mathscr{E} = \left\{ (u, v, w) \in \mathbb{R}^3 \mid \frac{u^2}{a^2} + \frac{v^2}{b^2} + \frac{w^2}{c^2} = 1 \right\}.$$

Muestre que el conjunto E es una superficie.

Solución. Considere la función $\sigma : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ dada por

$$\sigma(x, y) = (a\cos(x)\operatorname{sen}(y), b\operatorname{sen}(x)\operatorname{sen}(y), c\cos(y)).$$

Por la Proposición 3, tenemos que σ es derivable en todo $\mathbb{R}^2$. Ahora, consideremos $A = [0, 2\pi] \times [0, \pi]$. Se tiene que

$$\frac{(a\cos(x)\operatorname{sen}(y))^2}{a^2} + \frac{(b\operatorname{sen}(x)\operatorname{sen}(y))^2}{b^2} + \frac{(c\cos(y))^2}{c^2} = 1,$$

para cualquier $(x, y) \in A$. Es decir, $\sigma(x, y) \in \mathscr{E}$, para cualquier $(x, y) \in A$.

Por otro lado, si $(u, v, w) \in \mathscr{E}$, entonces el punto $\left(\frac{u}{a}, \frac{v}{b}, \frac{w}{c}\right)$ es un punto en la esfera de radio 1 con centro en el origen, en $\mathbb{R}^3$, así que existe $(x, y) \in A$ de tal manera que

$$(\cos(x)\operatorname{sen}(y), \operatorname{sen}(x)\operatorname{sen}(y), \cos(y)) = \left(\frac{u}{a}, \frac{v}{b}, \frac{w}{c}\right).$$

Se sigue que $\sigma(x, y) = (u, v, w)$.
Concluimos que $\sigma(A) = \mathscr{E}$, es decir $\mathscr{E}$ es una superficie parametrizada por σ. ∎

En esta ocasión también utilizamos una parametrización de una superficie para definir el plano tangente a la superficie en un punto de esta.

Definición 14 *Sean S una superficie y $\sigma = (\sigma_1, \sigma_2) : U \longrightarrow \mathbb{R}^3$ una parametrización de S, es decir, σ es derivable en cada punto de U y existe $A \subseteq U$ tal que $\sigma(A) = S$. Si $\overline{x}_0 \in A$ y los vectores $\frac{\partial \sigma}{\partial x}(\overline{x}_0)$ y $\frac{\partial \sigma}{\partial y}(\overline{x}_0)$, definidos como sigue*

$$\frac{\partial \sigma}{\partial x}(\overline{x}_0) = \left(\frac{\partial \sigma_1}{\partial x}(\overline{x}_0), \frac{\partial \sigma_2}{\partial x}(\overline{x}_0), \frac{\partial \sigma_3}{\partial x}(\overline{x}_0) \right),$$
$$\frac{\partial \sigma}{\partial y}(\overline{x}_0) = \left(\frac{\partial \sigma_1}{\partial y}(\overline{x}_0), \frac{\partial \sigma_2}{\partial y}(\overline{x}_0), \frac{\partial \sigma_3}{\partial y}(\overline{x}_0) \right),$$

son linealmente independientes, diremos que el plano $P \subseteq \mathbb{R}^3$, definido paramétricamente como

$$P = \left\{ \sigma(\overline{x}_0) + t\frac{\partial \sigma}{\partial x}(\overline{x}_0) + s\frac{\partial \sigma}{\partial y}(\overline{x}_0) \in \mathbb{R}^3 \mid t, s \in \mathbb{R} \right\}$$

es el plano tandente a S en el punto $\sigma(\overline{x}_0)$ y que S es suave en $\sigma(\overline{x}_0)$.

Ejemplo 15 *Sean $a, b, c \in \mathbb{R}^+$, considere la superficie*

$$\mathscr{E} = \left\{ (u, v, w) \in \mathbb{R}^3 \mid \frac{u^2}{a^2} + \frac{v^2}{b^2} + \frac{w^2}{c^2} = 1 \right\}.$$

Halle el plano tangente a la superficie $\mathscr{E}$ en el punto $(0, b, 0)$.

Solución. Consideremos la parametrización

$$\sigma(x, y) = (a \cos(x) \operatorname{sen}(y), b \operatorname{sen}(x) \operatorname{sen}(y), c \cos(y))$$

de la superficie $\mathscr{E}$ vista en el Ejemplo 13. Se tiene que $\sigma\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = (0, b, 0)$, luego

$$\frac{\partial \sigma}{\partial x}(x, y) = (-a \operatorname{sen}(x) \operatorname{sen}(y), b \cos(x) \operatorname{sen}(y), 0)$$

y

$$\frac{\partial \sigma}{\partial y}(x, y) = (a \cos(x) \cos(y), b \operatorname{sen}(x) \cos(y), -c \operatorname{sen}(y)).$$

Así,

$$\frac{\partial \sigma}{\partial x}\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = (-a, 0, 0) \qquad \text{y} \qquad \frac{\partial \sigma}{\partial y}\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = (0, 0, -c),$$

que claramente son linealmente independientes. Por lo tanto, el plano $P \subseteq \mathbb{R}^3$ tangente a $\mathscr{E}$ en el punto $(0, b, 0) = \sigma\left(\frac{\pi}{2}, \frac{\pi}{2}\right) = (0, b, 0)$ está definido paramétricamente como sigue

$$\begin{aligned} P &= \left\{ \sigma(\overline{x}_0) + t\frac{\partial \sigma}{\partial x}(\overline{x}_0) + s\frac{\partial \sigma}{\partial y}(\overline{x}_0) \in \mathbb{R}^3 \mid t, s \in \mathbb{R} \right\} \\ &= \left\{ (0, b, 0) + t(-a, 0, 0) + s(0, 0, -c) \in \mathbb{R}^3 \mid t, s \in \mathbb{R} \right\}. \end{aligned}$$

Se trata de un plano paralelo al plano XZ por el punto $(0, b, 0)$ (vea figura 1). ■

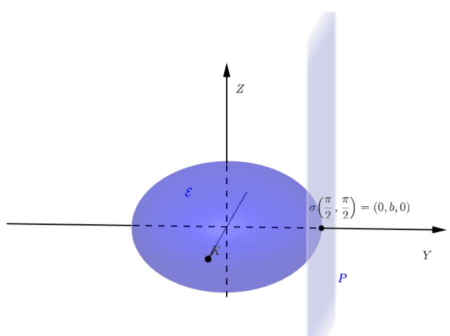


Figura 1: El plano tangente a $\mathscr{E}$ en el punto $(0, b, 0)$ es un plano paralelo al XZ por el punto $(0, b, 0)$.

3. El Teorema de la función Implícita

Consideremos el conjunto $\mathscr{C} = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 = r^2\}$, donde r es un número real positivo fijo.

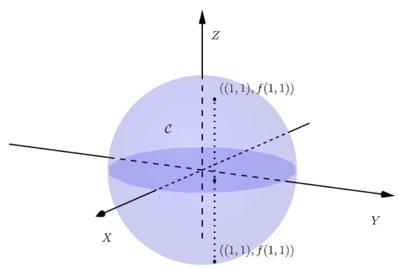


Figura 2: El conjunto $\mathscr{C}$ no se puede ver como la gráfica de una función de $\mathbb{R}^2$ en $\mathbb{R}$.

Note que este conjunto no se puede pensar como la gráfica de una función de $\mathbb{R}^2$ en $\mathbb{R}$, pues sin importar la forma en que lo pensemos, gráficamente, le asignaríamos dos puntos distintos a un punto de $\mathbb{R}^2$, con lo que no tendríamos una función (vea figura 2).
El Teorema de la Función Implícita nos proporciona condiciones para que, localmente, podamos pensar un conjunto como la gráfica de una función.

Teorema 16 (de la Función Implícita) *Sean $U \subset \mathbb{R}^m \times \mathbb{R}^k$ y $g_1, ..., g_m : U \longrightarrow \mathbb{R}$ funciones tales que $g_i \in C^k(U)$ para toda $i \in \{1, ..., m\}$. Si*

$$S = \Big\{(\overline{x}, \overline{y}) = (x_1, \ldots, x_m, y_1, \ldots, y_k) \in \mathbb{R}^m \times \mathbb{R}^k \mid g_i(\overline{x}, \overline{y}) = 0 \text{ para } i \in \{1, \ldots, m\}\Big\}$$

y $(\overline{x}_0, \overline{y}_0) \in S$ es tal que la matriz

$$\begin{bmatrix} \frac{\partial g_1}{\partial x_1}(\overline{x}_0, \overline{y}_0) & \cdots & \frac{\partial g_1}{\partial x_m}(\overline{x}_0, \overline{y}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1}(\overline{x}_0, \overline{y}_0) & \cdots & \frac{\partial g_m}{\partial x_m}(\overline{x}_0, \overline{y}_0) \end{bmatrix}$$

es invertible, entonces existen $r > 0$ y $h : B_r(\overline{y}_0) \subset \mathbb{R}^k \to \mathbb{R}^m$, con $h \in C^1(B_r(\overline{y}_0))$, tales que $h(\overline{y}_0) = \overline{x}_0$ y $(h(\overline{y}), \overline{y}) \in S$ para toda $\overline{y} \in B_r(\overline{y}_0)$.

En el caso del conjunto $\mathscr{C} = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 = r^2\}$, con el que iniciamos esta sección, usando el Teorema de la Función Implícita, tenemos los siguiente:
Sea $g : \mathbb{R} \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por

$$g(x,(y,z)) = x^2 + y^2 + z^2 - r^2.$$

Note que $g \in C^2(\mathbb{R} \times \mathbb{R}^2)$ y

$$S = \big\{(x,(y,z)) \in \mathbb{R} \times \mathbb{R}^2 \mid g(x,(y,z)) = 0\big\} = \mathscr{C}.$$

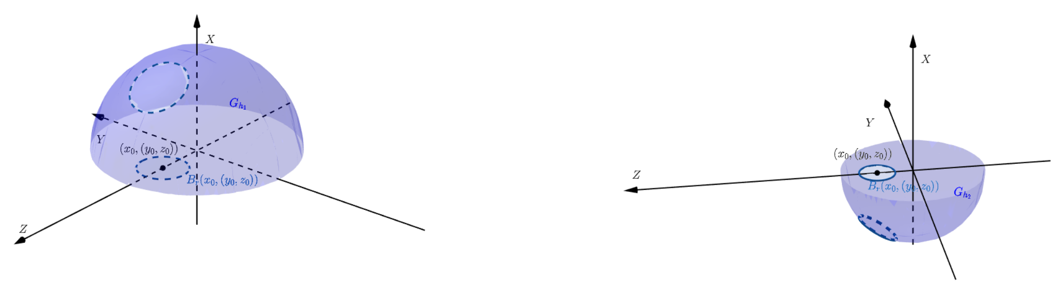
Ahora, la matriz del enunciado del teorema de la función implícita, en este caso, es una matriz de 1×1, por lo que pedir que sea invertible es equivalente a pedir que su única entrada sea distinta de cero. Luego, la única entrada de dicha matriz es

$$\frac{\partial g}{\partial x}(x,(y,z)) = 2x,$$

y esta es distinta de cero si x es distinto de cero. Se sigue, por el Teorema de la Función Implícita, que si $(x_0,(y_0,z_0)) \in S = \mathscr{C}$ es tal que $x_0 \neq 0$, entonces existen $r > 0$ y $h : B_r(y_0,z_0) \longrightarrow \mathbb{R}$, con $h \in C^1(B_r(y_0,z_0))$, tales que $h(y_0,z_0) = x_0$ y $(h(y,z),(y,z)) \in S = \mathscr{C}$ para toda $(y,z) \in B_r(y_0,z_0)$.

Algo que es fácil de notar en este ejemplo es que para un punto $(x_0,(y_0,z_0)) \in S = \mathscr{C}$ con $x_0 \neq 0$, tenemos exactamente dos funciones h_1, h_2 (vea la figura 3) que cumplen la conclusión del Teorema de la Función Implícita, a saber,

$$h_1(y,z) = \sqrt{r^2 - (y^2+z^2)} \qquad \text{y} \qquad h_2(y,z) = -\sqrt{r^2 - (y^2+z^2)}.$$



(a) Gráfica de la función h_1 (b) Gráfica de la función h_2

Figura 3: En un punto $(x_0,(y_0,z_0)) \in S = \mathscr{C}$ tal que $x_0 \neq 0$ existen dos funciones que satisfacen la conclusión del Teorema de la Función Implícita.

Veamos un par de ejemplos más donde apliquemos el Teorema 16.

Ejemplo 17 *Halle los números reales a para los cuales existen $r > 0$ y una función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ derivable en $B_r(0,0)$ tal que $f(0,0) = 0$ y*

$$af^3(x,y) + (a+x+1)f(x,y) - y^2 + x = 0$$

para todo $(x,y) \in B_r(0,0)$.

Solución. Consideremos la función $g : \mathbb{R} \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por

$$g(z,(x,y)) = az^3 + (a+x+1)z - y^2 + x$$

y sea $S = \{(z,(x,y)) \mid g(z,(x,y)) = 0\}$. Note que $g \in C^2(\mathbb{R} \times \mathbb{R}^2)$ y que $(0,(0,0)) \in S$. Luego,

$$\frac{\partial g}{\partial z}(x,y,z) = 3az^2 + a + x + 1,$$

de donde

$$\frac{\partial g}{\partial z}(0,(0,0)) = a + 1.$$

Así que si $a \neq -1$, por el Teorema de la Función Implícita, existen $r > 0$ y $f : B_r(0,0) \longrightarrow \mathbb{R}$, con $f \in C^1(B_r(0,0))$, tales que

$$f(0,0) = 0 \qquad \text{y} \qquad (f(x,y),(x,y)) \in S, \quad \text{para todo} \ \ (x,y) \in B_r(0,0),$$

es decir,

$$f(0,0) = 0 \qquad \text{y} \qquad af^3(x,y) + (a+x+1)f(x,y) - y^2 + x = 0, \quad \text{para todo} \ \ (x,y) \in B_r(0,0).$$

■

Ejemplo 18 *Considere el siguiente sistema de ecuaciones*

$$\begin{aligned} z^3x + w^2y^3 + 2xy &= 0 \\ xyzw - 1 &= 0. \end{aligned}$$

Note que la cuaterna $(x,y,z,w) = (-1,-1,1,1)$ *es una solución del sistema. Muestre que la pareja* (z,w) *se puede expresar en función de la pareja* (x,y) *en alguna bola de* $\mathbb{R}^2$ *con centro en* $(-1,-1)$.

Solución. Considere las funciones $g_1, g_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dadas por

$$g_1\left((z,w),(x,y)\right) = z^3x + w^2y^3 + 2xy$$

y

$$g_2\left((z,w),(x,y)\right) = xyzw - 1$$

y sea

$$S = \left\{((z,w),(x,y)) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid g_1\left((z,w),(x,y)\right) = 0, \ g_2\left((z,w),(x,y)\right) = 0\right\}.$$

Note que $g_1, g_2 \in C^2(\mathbb{R}^2 \times \mathbb{R}^2)$ y que $((1,1),(-1,-1)) \in S$. Ahora,

$$\begin{bmatrix} \dfrac{\partial g_1}{\partial z}\left((z,w),(x,y)\right) & \dfrac{\partial g_1}{\partial w}\left((z,w),(x,y)\right) \\ \dfrac{\partial g_2}{\partial z}\left((z,w),(x,y)\right) & \dfrac{\partial g_2}{\partial w}\left((z,w),(x,y)\right) \end{bmatrix} = \begin{bmatrix} 3z^2x & 2wy^3 \\ xyw & xyz \end{bmatrix},$$

por lo que

$$\begin{bmatrix} \dfrac{\partial g_1}{\partial z}\left((1,1),(-1,-1)\right) & \dfrac{\partial g_1}{\partial w}\left((1,1),(-1,-1)\right) \\ \dfrac{\partial g_2}{\partial z}\left((1,1),(-1,-1)\right) & \dfrac{\partial g_2}{\partial w}\left((1,1),(-1,-1)\right) \end{bmatrix} = \begin{bmatrix} -3 & -2 \\ 1 & 1 \end{bmatrix}$$

y como

$$det \begin{bmatrix} -3 & -2 \\ 1 & 1 \end{bmatrix} = (-3)(1) - (-2)(1) = -1 \neq 0$$

se sigue que

$$\begin{bmatrix} -3 & -2 \\ 1 & 1 \end{bmatrix}$$

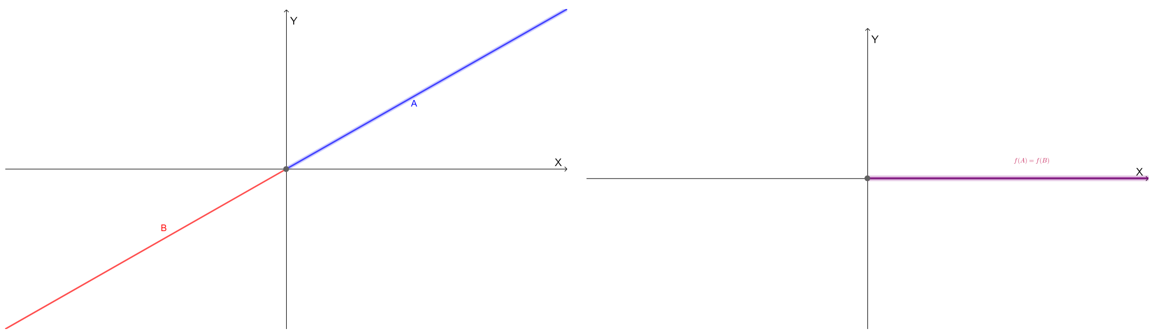
es invertible. Luego, por el Teorema de la Función Implícita, tenemos que existen $r > 0$ y $h : B_r(-1,-1) \longrightarrow \mathbb{R}^2$, con $h \in C^1(\mathbb{R}^2 \times \mathbb{R}^2)$, tales que $h(-1,-1) = (1,1)$ y $h(x,y) = (z,w)$ para todo $(x,y) \in B_r(-1,-1)$. ■

4. Teorema de la función inversa

Como ya sabemos, para que exista la función inversa de una función dada es necesario que la función sea inyectiva en su dominio, por ejemplo la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, dada por

$$f(x, y) = (x^2 - y^2, 2xy)$$

no es inyectiva en $\mathbb{R}^2$, pues $f(-x, -y) = f(x, y)$ para toda $(x, y) \in \mathbb{R}^2$, por lo que no existe la función inversa de f.



(a) $A = \{(x, x) \mid x > 0\}$ y $B = \{(x, x) \mid x < 0\}$ son subconjuntos del dominio de f.

(b) Las imágenes bajo f de los conjuntos A y B coinciden, es decir, $f(A) = f(B)$.

Figura 4: La función $f(x, y) = (x^2 - y^2, 2xy)$ no es inyectiva en su dominio.

El Teorema de la Función Inversa nos proporciona condiciones para que en un punto dado exista una bola, que tiene como centro a dicho punto, donde la función es inyectiva y por lo tanto invertible (a esto lo llamamos localmente invertible).

Teorema 19 (de la Función Inversa) *Sea $U \subset \mathbb{R}^n$ un conjunto abierto, $\overline{x}_0 \in U$ y $f : U \longrightarrow \mathbb{R}^n$ una función tal que $f \in C^1(U)$. Si $Df(\overline{x}_0)$ es invertible, entonces existe $r > 0$ tal que:*

(1) f es inyectiva en $B_r(\overline{x}_0) \subseteq U$,

(2) $f^{-1} : f(B_r(\overline{x}_0)) \longrightarrow \mathbb{R}^n$ es continua en $f(B_r(\overline{x}_0))$,

(3) $f(B_r(\overline{x}_0)) \subseteq \mathbb{R}^n$ es un conjunto abierto,

(4) $f^{-1} \in C^1\left(f\left(B_r(\overline{x}_0)\right)\right)$ y si $\overline{y} = f(\overline{x}) \in f(B_r(\overline{x}_0))$, entonces

$$Df^{-1}(\overline{y}) = Df^{-1}(f(\overline{x})) = \left(Df(\overline{x})\right)^{-1}.$$

Veamos con un par de ejemplos cómo aplicar este teorema.

Ejemplo 20 *Considere la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, dada por*

$$f(x, y) = (x^2 - y^2, 2xy).$$

Halle los puntos $(x, y) \in \mathbb{R}^2$ donde f es localmente invertible.

Solución. Note que las funciones coordenadas $f_1, f_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ están dadas por

$$f_1(x,y) = x^2 - y^2 \qquad \text{y} \qquad f_2(x,y) = 2xy.$$

Así,

$$\begin{aligned} Df(x,y) &= \begin{bmatrix} \dfrac{\partial f_1}{\partial x}(x,y) & \dfrac{\partial f_1}{\partial y}(x,y) \\ \dfrac{\partial f_2}{\partial x}(x,y) & \dfrac{\partial f_2}{\partial y}(x,y) \end{bmatrix} \\ &= \begin{bmatrix} 2x & -2y \\ 2y & 2x \end{bmatrix}. \end{aligned}$$

Por lo tanto $det(Df(x,y)) = 4x^2 + 4y^2$ y de aquí que $Df(x,y)$ es invertible si $(x,y) \neq (0,0)$. Concluimos, por el Teorema de la Función Inversa, que f es localmente invertible en cada $(x,y) \in \mathbb{R}^2$, distinto de $(0,0)$. ■

Ejemplo 21 *Considere la función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por*

$$f(x,y) = (x^3 + 2xy + y^2, x^2 + y).$$

¿Es f inyectiva? ¿Es f localmente invertible en $(1,1)$? En caso afirmativo, estime el valor de $f^{-1}(4.1, 1.8)$.

Solución. Note que $f(-2,-2) = (4,2) = f(1,1)$ por lo que f no es inyectiva. Ahora, las funciones coordenadas $f_1, f_2 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ están dadas por

$$f_1(x,y) = x^3 + 2xy + y^2 \qquad \text{y} \qquad f_2(x,y) = x^2 + y.$$

Así,

$$\begin{aligned} Df(x,y) &= \begin{bmatrix} \dfrac{\partial f_1}{\partial x}(x,y) & \dfrac{\partial f_1}{\partial y}(x,y) \\ \dfrac{\partial f_2}{\partial x}(x,y) & \dfrac{\partial f_2}{\partial y}(x,y) \end{bmatrix} \\ &= \begin{bmatrix} 3x^2 + 2y & 2x + 2y \\ 2x & 1 \end{bmatrix}. \end{aligned}$$

Luego

$$Df(1,1) = \begin{bmatrix} 5 & 4 \\ 2 & 1 \end{bmatrix}$$

y como $det(Df(1,1)) = -3 \neq 0$, entonces $Df(1,1)$ es invertible. Así, por inciso (1) del Teorema de la Función Inversa, existe $r > 0$ tal que f es inyectiva en $B_r(\overline{x}_0)$, luego es invertible en $B_r(\overline{x}_0)$. Ahora, por el inciso (4), tenemos que $f^{-1} \in C^1(B_r(1,1))$ y para $(4,2) = f(1,1) \in f(B_r(1,1))$ que

$$Df^{-1}(4,2) = (Df(1,1))^{-1} = \begin{bmatrix} 5 & 4 \\ 2 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1/3 & 4/3 \\ 2/3 & -5/3 \end{bmatrix}.$$

Finalmente, como la derivada de una función en un punto es la mejor aproximación lineal cerca de dicho punto, tenemos que

$$f^{-1}(4.1, 1.8) \approx f^{-1}(4, 2) + Df^{-1}[(4.1, 1.8) - (4, 2)],$$

es decir,

$$\begin{aligned} f^{-1}(4.1, 1.8) &\approx (1, 1) + \begin{bmatrix} -1/3 & 4/3 \\ 2/3 & -5/3 \end{bmatrix} \begin{bmatrix} 0.1 \\ -0.2 \end{bmatrix} \\ &= (1, 1) + (-3/10, 2/5) \\ &= (0.7, 1.4). \end{aligned}$$

$\blacksquare$