Cálculo III
Funciones de $\mathbb{R}^n$ en $\mathbb{R}^m$ II: aplicaciones
Funciones inversas y extremos de funciones de varias variables
Módulo 3, Presentación 8

Raimund Bürger

21 de abril de 2025

3.3. Funciones inversas

El siguiente teorema representa una de las aplicaciones más importantes del Teorema Principal de Funciones Implícitas (Teorema 3.2).

Teorema 3.3 (Teorema Principal de las Funciones Inversas)
Sea $f : \mathbb{R}^n \to \mathbb{R}^n$. Sea $X \subset D(f)$ un conjunto abierto, y $f \in C^1(X)$. En el punto x^0 la matriz Jacobiana de f sea regular, es decir,

$$\det \left(\frac{\mathrm{d}f}{\mathrm{d}x}(x^0) \right) \neq 0.$$

1. En este caso existen vecindades abiertas $U(x^0)$ de x^0 y $V(y^0)$ de $y^0 = f(x^0)$ tales que $U(x^0)$ es mapeado de manera biyectiva a $V(y^0)$, es decir sobre $V(y^0)$ existe la aplicación inversa f^{-1}.
2. Se tiene $f^{-1} \in C^1(V(y^0))$, y para $y = f(x) \in V(y^0)$, se tiene la siguiente relación entre $\mathrm{d}f/\mathrm{d}x$ y $\mathrm{d}f^{-1}/\mathrm{d}y$:

$$\frac{\mathrm{d}f^{-1}}{\mathrm{d}y}(y) = \left(\frac{\mathrm{d}f}{\mathrm{d}y}(x) \right)^{-1}.$$

3.3. Funciones inversas

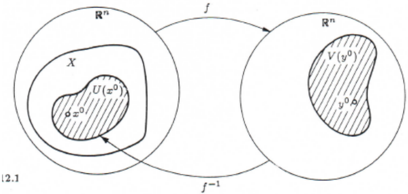


Figura: Ilustración del Teorema 3.3.

Demostración del Teorema 3.3

1. Escribimos la aplicación dada

$$y_1 = f_1(x_1, \ldots, x_n), \ldots, y_n = f_n(x_1, \ldots, x_n)$$

como un sistema de ecuaciones en $\mathbb{R}^{2n}$, donde ponemos $y_1 = x_{n+1}, \ldots, y_n = x_{2n}$ y $y_1^0 = x_{n+1}^0, \ldots, y_n^0 = x_{2n}^0$:

$$\begin{aligned} F_1(x_1, \ldots, x_{2n}) = f_1(x_1, \ldots, x_n) - x_{n+1} &= 0, \\ &\vdots \\ F_n(x_1, \ldots, x_{2n}) = f_n(x_1, \ldots, x_n) - x_{2n} &= 0. \end{aligned} \tag{3.7}$$

Demostración del Teorema 3.3 (continuación)

2. En el punto $(x_1^0, \ldots, x_{2n}^0)$ el sistema de ecuaciones está satisfecho; además, en una vecindad de este punto $F = (F_1, \ldots, F_n)$ pertenece a C^1, y

$$\begin{vmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial F_n}{\partial x_1} & \cdots & \dfrac{\partial F_n}{\partial x_n} \end{vmatrix} (x_1^0, \ldots, x_{2n}^0) = \begin{vmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{vmatrix} (x_1^0, \ldots, x_n^0) \neq 0.$$

3. Esto significa que localmente podemos despejar $x_1, \ldots, x_n$, es decir existen un $\delta > 0$ y funciones

$$x_1 = f_1^{-1}(x_{n+1}, \ldots, x_{2n}), \ldots, x_n = f_n^{-1}(x_{n+1}, \ldots, x_{2n})$$

definidas, continuas y C^1 sobre $|x_\nu - x_\nu^0| < \delta$ $(\nu = n+1, \ldots, 2n)$, las cuales allí satisfacen (3.7).

3.3. Funciones inversas

Demostración del Teorema 3.3 (continuación)

4. Si definimos $y_\nu = x_{n+\nu}$ y $y_\nu^0 = x_{n+\nu}^0$ para $\nu = 1, \dots, n$ esto significa que para $(y_1, \dots, y_n)$ tales que $|y_\nu - y_\nu^0| < \delta$ $(\nu = 1, \dots, n)$ hemos encontrado funciones tales que

$$\begin{aligned} y_1 &= f_1\bigl(f_1^{-1}(y_1, \dots, y_n), \dots, f_n^{-1}(y_1, \dots, y_n)\bigr), \\ &\vdots \\ y_n &= f_n\bigl(f_1^{-1}(y_1, \dots, y_n), \dots, f_n^{-1}(y_1, \dots, y_n)\bigr). \end{aligned}$$

La derivación parcial nos entrega

$$\frac{\partial y_\mu}{\partial y_\nu} = \sum_{i=1}^{n} \frac{\partial f_\mu}{\partial x_i} \frac{\partial f_i^{-1}}{\partial y_\nu} = \begin{cases} 1 & \text{si } \mu = \nu, \\ 0 & \text{si } \mu \neq \nu. \end{cases}$$

Esto demuestra el enunciado (2) según la definición del producto matricial. ■

3.3. Funciones inversas

Ejemplo 3.6 Consideremos la aplicación $f : \mathbb{R}^2 \to \mathbb{R}^2$ definida por

$$y_1 = f_1(x_1, x_2) = x_1^2, \quad y_2 = f_2(x_1, x_2) = x_1 + x_2.$$

$$\text{Aquí} \quad \left|\frac{\mathrm{d}f}{\mathrm{d}x}\right| = \begin{vmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 \end{vmatrix} = \begin{vmatrix} 2x_1 & 0 \\ 1 & 1 \end{vmatrix} = 2x_1.$$

Consideremos un punto $x^0 = (x_1^0, x_2^0) \in \mathbb{R}^2$, entonces

$$\left|\frac{\mathrm{d}f}{\mathrm{d}x}(x^0)\right| \neq 0 \Leftrightarrow x_1^0 \neq 0.$$

Si $x_1^0 > 0$, la aplicación inversa f_{R}^{-1} de f es definida por

$$x_1 = \sqrt{y_1}, \quad x_2 = y_2 - \sqrt{y_1}.$$

El dominio de f_{R}^{-1} es el semiplano derecho; la imagen igualmente es el semiplano derecho. Para $x_1^0 < 0$, obtenemos f_{L}^{-1} definida por

$$x_1 = -\sqrt{y_1}, \quad x_2 = y_2 + \sqrt{y_1}.$$

El dominio de f_{L}^{-1} es el semiplano derecho; la imagen ahora es el semiplano izquierdo.

3.3. Funciones inversas

Ejemplo 3.7 Sea la función $f := (u, v) : \mathbb{R}^2 \to \mathbb{R}^2$ dada por

$$u = u(x, y) = x \cos y, \quad v = v(x, y) = x \sin y,$$

sobre $D(f) = \{(x, y) \in \mathbb{R}^2 \,|\, x > 0\}$. analizar la invertibilidad de f notamos que

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{bmatrix} \cos y & -x \sin y \\ \sin y & x \cos y \end{bmatrix}$$

con el determinante

$$\det \frac{\partial(u, v)}{\partial(x, y)} = x \cos^2 y + x \sin^2 y = x.$$

Según el Teorema 3.3, f es invertible en una vecindad de cada punto (x_0, y_0) con $x_0 \neq 0$. Pero f no es invertible globalmente, dado que

$$\begin{aligned} u(x_0, y_0 + 2\pi) &= x_0 \cos(y_0 + 2\pi) = x_0 \cos y_0 = u(x_0, y_0), \\ v(x_0, y_0 + 2\pi) &= x_0 \sin(y_0 + 2\pi) = x_0 \sin y_0 = v(x_0, y_0), \end{aligned}$$

así que la imagen de (x_0, y_0) y $(x_0, y_0 + 2\pi)$ siempre es la misma.

3.3. Funciones inversas

Ejemplo 3.8 Se consideran las funciones

$$\begin{aligned} f_1(x,y,u,v) &= x + y + (u+v)(u-v), \\ f_2(x,y,u,v) &= 2x - y + (u^2-1)\cos v + 1. \end{aligned} \tag{3.8}$$

Sea $P = (x_0 = 0, y_0 = 0, u_0 = 0, v_0 = 0)$.

a) Analizar si en una vecindad de P, existen funciones $g_1 = g_1(u,v)$ y $g_2 = g_2(u,v)$ tales que las variables $x = g_1(u,v)$ e $y = g_2(u,v)$ puedan ser despejadas de

$$f_1(x,y,u,v) = 0, \quad f_2(x,y,u,v) = 0. \tag{3.9}$$

Si posible, calcular las derivadas parciales

$$\frac{\partial g_1}{\partial u}(0,0), \quad \frac{\partial g_1}{\partial v}(0,0), \quad \frac{\partial g_2}{\partial u}(0,0), \quad \frac{\partial g_2}{\partial v}(0,0).$$

b) Analizar si en una vecindad de P, existen funciones $h_1 = h_1(x,y)$ y $h_2 = h_2(x,y)$ tales que las variables $u = h_1(x,y)$ y $v = h_2(x,y)$ puedan ser despejadas de (3.9). Si posible, calcular las derivadas parciales $(\partial h_1/\partial x)(0,0)$, etc.

3.3. Funciones inversas

Ejemplo 3.8 (continuación)

a) Aquí calculamos

$$\det \begin{bmatrix} \partial f_1/\partial x & \partial f_1/\partial y \\ \partial f_2/\partial x & \partial f_2/\partial y \end{bmatrix} = \begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} = -3 \neq 0,$$

entonces, según el T.F.I., las funciones deseadas g_1 y g_2 existen. Por derivación implícita obtenemos

$$0 = \frac{\partial g_1}{\partial u} + \frac{\partial g_2}{\partial u} + 2u, \qquad 0 = 2\frac{\partial g_1}{\partial u} - \frac{\partial g_2}{\partial u} + 2u \cos v,$$
$$0 = \frac{\partial g_1}{\partial v} + \frac{\partial g_2}{\partial v} - 2v, \qquad 0 = 2\frac{\partial g_1}{\partial v} - \frac{\partial g_2}{\partial v} - (u^2 - 1) \sin v,$$

es decir en P_0 las derivadas parciales son soluciones de

$$\begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{pmatrix} \frac{\partial g_1}{\partial u}(0,0) \\ \frac{\partial g_2}{\partial u}(0,0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{pmatrix} \frac{\partial g_1}{\partial v}(0,0) \\ \frac{\partial g_2}{\partial v}(0,0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$$\Rightarrow \frac{\partial g_1}{\partial u}(0,0) = \frac{\partial g_1}{\partial v}(0,0) = \frac{\partial g_2}{\partial u}(0,0) = \frac{\partial g_2}{\partial v}(0,0) = 0.$$

Ejemplo 3.8 (continuación)

b) Aquí calculamos para $u = v = 0$

$$\det \begin{bmatrix} \partial f_1/\partial u & \partial f_1/\partial v \\ \partial f_2/\partial u & \partial f_2/\partial v \end{bmatrix} = \begin{vmatrix} 2u & -2v \\ 2u \cos v & (1-u^2) \sin v \end{vmatrix} = 0,$$

entonces el Teorema de las Funciones Implícitas no puede ser aplicad. Efectivamente, las variables u y v no pueden ser despejadas de manera única en ninguna vecindad de P. Esto es debido a que las funciones f_1 y f_2 son simétricas con respecto a u y v:

$$f_i(x, y, u, v) = f_i(x, y, \pm u, \pm v), \quad i = 1, 2.$$

Como $(u, v) = 0$ es una solución de (3.9) solamente para $(x = 0, y = 0)$, vemos que para cada $(x, y) \neq (0, 0)$, la solución de (3.8) entrega $(u, v) \neq (0, 0)$: considerando todas las combinaciones de signos en $(\pm u, \pm v)$, existe por lo menos una segunda solución.

3.3. Funciones inversas

Teorema 3.4 Sea $G \subset \mathbb{R}^n$ una región. La aplicación $f : \mathbb{R}^n \to \mathbb{R}^n$ satisfaga $f \in C^1(G)$, y sea

$$\det \frac{\mathrm{d}f}{\mathrm{d}x}(x) \neq 0 \quad \text{para todo } x \in G. \tag{3.10}$$

Entonces también $f(G)$ es una región. (Ver Teoremas 1.18 y 2.18)

Demostración. Según el Teorema 2.18, el conjunto G es abierto y conexo. Como f es continua sobre G, el Teorema 1.18 implica que $f(G)$ es conexo. Sean $y^0 \in f(G)$ y $x^0 \in G$ escogidos tales que $f(x^0) = y^0$. Entonces, en virtud de

$$\det \frac{\mathrm{d}f}{\mathrm{d}x}(x^0) \neq 0,$$

según el Teorema 3.3 existe una vecindad $V(y^0)$ de y^0 con $V(y^0) \subset f(G)$. Puesto que $y^0 \in f(G)$ es arbitrario, $f(G)$ es abierto y según el Teorema 2.18 una región.

3.4. Extremos de funciones de varias variables

Teorema 3.5 Sea la función $f : \mathbb{R}^n \to \mathbb{R}$ parcialmente diferenciable con respecto a cada variable en un punto interior x^0 de $D(f)$. Si f posee un extremo relativo en x^0, entonces

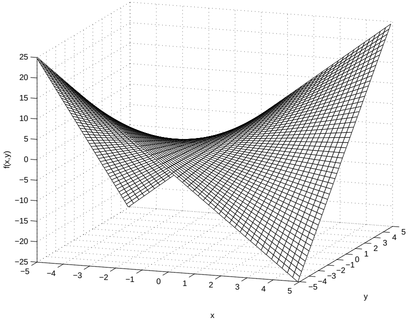
$$\nabla f(x^0) = \vec{0}.$$

Demostración. El enunciado sigue inmediatamente del hecho que f posee en x^0 un extremo en cada dirección de coordenada $\vec{e}_i$ como función de la variable escalar x_i, lo que implica $f_{x_i}(x^0) = 0$ para $i = 1, \ldots, n$. ■

La condición indicada en el Teorema 3.5, $\nabla f(x^0) = \vec{0}$, es una condición necesaria, pero no suficiente para la existencia de un extremo relativo.

3.4. Extremos de funciones de varias variables

Consideremos la función $f : \mathbb{R}^2 \to \mathbb{R}$ con $D(f) = \mathbb{R}^2$ y $f(x, y) = xy$:



Aquí $f_x(x, y) = y$ y $f_x(x, y) = x$, por lo tanto $\nabla f(x, y) = \vec{0}$ si y sólo si $(x_0, y_0) = (0, 0)$. Sin embargo, la función f no posee un extremo relativo en $(0, 0)$ dado que en cada vecindad de $(0, 0)$ existen valores de f positivo y negativos.

Aquellos puntos x^0 de $f : \mathbb{R}^n \to \mathbb{R}$ donde $\nabla f(x^0) = \vec{0}$ pero que no son extremos relativos se llaman puntos de silla.

3.4. Extremos de funciones de varias variables

Para la formulación de condiciones suficientes para la existencia de extremos locales de funciones $f : \mathbb{R}^n \to \mathbb{R}$ necesitamos estudiar formas cuadráticas.

Definición 3.3 Sea $A = (a_{ik})_{i,k=1,\dots,n}$ una matriz simétrica.

1. El polinomio $Q_A : \mathbb{R}^n \to \mathbb{R}$ definido por
$$Q_A(x) = \sum_{i,k=1}^{n} a_{ik} x_i x_k, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n$$
se llama la forma cuadrática asociada con A.
2. La matriz A o la función Q_A se llama semidefinida positiva si $Q_A(x) \geqslant 0$ para todo $x \in \mathbb{R}^n$ y semidefinida negativa si $Q_A(x) \leqslant 0$ para todo $x \in \mathbb{R}^n$.
3. La matriz A o la función Q_A se llama definida positiva si $Q_A(x) > 0$ para todo $x \in \mathbb{R}^n$, $x \neq (0, \dots, 0)$, y definida negativa si $Q_A(x) < 0$ para todo $x \in \mathbb{R}^n$, $x \neq (0, \dots, 0)$.
4. La matriz A o la función Q_A se llama indefinida si no es semidefinida ni positiva ni negativa.

3.4. Extremos de funciones de varias variables

Teorema 3.6 Sea $A = (a_{ik}) \in \mathbb{R}^{n\times n}$ una matriz simétrica; y sea $-A = (-a_{ik})$.

1. La matriz A es semidefinida positiva si y sólo si $-A$ es semidefinida negativa.
2. La matriz A es definida positiva si y sólo si $-A$ es definida negativa.

Demostración Las afirmaciones son una consecuencia inmediata de

$$Q_A(x) = \sum_{i,k=1}^{n} a_{ik}x_ix_k = -\sum_{i,k=1}^{n}(-a_{ik})x_ix_k = -Q_{(-A)}(x). \quad \blacksquare$$

3.4. Extremos de funciones de varias variables

Teorema 3.7 La matriz

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix} \quad \text{es}$$

- semidefinida positiva si y sólo si $ad - b^2 \geqslant 0$, $a \geqslant 0$ y $d \geqslant 0$,
- semidefinida negativa si y sólo si $ad - b^2 \geqslant 0$, $a \leqslant 0$ y $d \leqslant 0$,
- definida positiva si y sólo si $ad - b^2 > 0$ y $a > 0$,
- definida negativa si y sólo si $ad - b^2 > 0$ y $a < 0$,
- indefinida si y sólo si $ad - b^2 < 0$.

Demostración Tarea. Se recomienda utilizar que para $(x, y) \in \mathbb{R}^2$ el polinomio $Q_A(x, y) = ax^2 + 2bxy + dy^2$ satisface

$$aQ_A(x, y) = (ax + by)^2 + (ad - b^2)y^2. \quad \blacksquare$$

3.4. Extremos de funciones de varias variables

En general es muy dificil decidir si una matriz es definida; los métodos especializados son tópico del álgebra lineal. Aquí mencionamos el siguiente criterio útil (sin demostración).

Teorema 3.8 Se considera la matriz simétrica $A = (a_{ik})$ con las submatrices principales

$$A_\nu = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1\nu} \\ a_{21} & a_{22} & \dots & a_{2\nu} \\ \vdots & \vdots & & \vdots \\ a_{\nu 1} & a_{\nu 2} & \dots & a_{\nu\nu} \end{bmatrix}.$$

Sea $\det A_\nu$ el determinante de A_ν. Entonces

1. A es definida positiva si y sólo si
$$\det A_\nu > 0 \quad \text{para } \nu = 1, \dots, n,$$
2. A es definida negativa si y sólo si
$$(-1)^\nu \det A_\nu > 0 \quad \text{para } \nu = 1, \dots, n.$$

3.4. Extremos de funciones de varias variables

(Ejemplo ad-hoc)

Teorema 3.9 Para un $\varepsilon > 0$ sean

$$U'_\varepsilon := \left\{ x \in \mathbb{R}^n \,\middle|\, 0 < d(x,0) = \sqrt{x_1^2 + \cdots + x_n^2} < \varepsilon \right\},$$
$$S_\varepsilon := \left\{ x \in \mathbb{R}^n \,\middle|\, d(x,0) = \sqrt{x_1^2 + \cdots + x_n^2} = \varepsilon \right\}.$$

Si sobre U'_ε o sobre S_ε se tiene que

- $Q_A(x) \geqslant 0$, entonces A es semidefinida positiva,
- $Q_A(x) \leqslant 0$, entonces A es semidefinida negativa,
- $Q_A(x) > 0$, entonces A es definida positiva,
- $Q_A(x) < 0$, entonces A es definida negativa.

Demostración Demostraremos solamente los primeros dos enunciados; la demostración de los dos demás enunciados es análoga.

3.4. Extremos de funciones de varias variables

Demostración del Teorema 3.9 (continuación)

1. Sea $Q_A \geqslant 0$ sobre U'_ε. Entonces para $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ con $x \neq (0, \ldots, 0)$ existe un $R > 0$ tal que
$$y = \left(\frac{x_1}{R}, \ldots, \frac{x_n}{R}\right) \in U'_\varepsilon.$$
Ahora $Q_A(y) \geqslant 0$, y por lo tanto $Q_A(x) = R^2 Q_A(y) \geqslant 0$.
2. Sobre S_ε sea $Q_A \geqslant 0$. Entonces para $(0, \ldots, 0) \neq x \in \mathbb{R}^n$,
$$y = \left(\frac{\varepsilon x_1}{d(x,0)}, \ldots, \frac{\varepsilon x_n}{d(x,0)}\right) \in S_\varepsilon.$$
Por lo tanto $Q_A(y) \geqslant 0$, y
$$Q_A(x) = \frac{d^2(x,0)}{\varepsilon^2} Q_A(y) \geqslant 0.$$
3. Para $Q_A \leqslant 0$ el enunciado sigue del Teorema 3.6.

3.4. Extremos de funciones de varias variables

Teorema 3.10 Sea $f : \mathbb{R}^n \to \mathbb{R}$. Sea $x^0 \in D(f)$, y para una vecindad $U(x^0)$ de x^0 sea $f \in C^2(U(x^0))$. Además, sea $\nabla f(x^0) = \vec{0}$, y consideremos la matriz

$$A(x^0) = \left(f_{x_i x_k}(x^0)\right)_{i,k=1,\ldots,n}.$$

Entonces:

1. Si $A(x^0)$ es definida positiva, entonces f posee un mínimo local en x^0. Si $A(x^0)$ es definida negativa, entonces f posee un máximo local en x^0.
2. Si f posee un mínimo relativo en x^0, entonces $A(x^0)$ es semidefinida positiva. Si f posee un máximo relativo en x^0, entonces $A(x^0)$ es semidefinida negativa.
3. Si $A(x^0)$ es indefinida, entonces f no posee en x^0 ni un mínimo relativo ni un máximo relativo.

3.4. Extremos de funciones de varias variables

Ejemplo 3.9 Consideremos $f : \mathbb{R}^3 \to \mathbb{R}$ con $D(f) = \mathbb{R}^3$ y

$$f(x, y, z) = 35 - 6x + 2z + x^2 - 2xy + 2y^2 + 2yz + 3z^2$$

$$\Rightarrow \qquad f_x = -6 + 2x - 2y, \quad f_y = -2x + 4y + 2z, \quad f_z = 2 + 2y + 6z.$$

Aquí $(x_0, y_0, z_0) = (8, 5, -2)$ es el único punto donde $\nabla f = \vec{0}$. Allí

$$f_{xx} = 2, \quad f_{xy} = -2, \quad f_{yy} = 4, \quad f_{xz} = 0, \quad f_{yz} = 2, \quad f_{zz} = 6$$

$$\Rightarrow \qquad A(x_0, y_0, z_0) = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 4 & 2 \\ 0 & 2 & 6 \end{bmatrix}.$$

Los determinantes de las submatrices principales son

$$|2| = 2, \quad \begin{vmatrix} 2 & -2 \\ -2 & 4 \end{vmatrix} = 4, \quad \begin{vmatrix} 2 & -2 & 0 \\ -2 & 4 & 2 \\ 0 & 2 & 6 \end{vmatrix} = 16.$$

Todos estos determinantes son positivos. Concluimos que según el Teorema 3.8 la matriz $A(x_0, y_0, z_0)$ es definida positiva, por lo tanto la función f posee en $(x_0, y_0, z_0) = (8, 5, -2)$ un mínimo local.

Demostración del Teorema 3.10 Según el Teorema 2.10, $A(x^0)$ es simétrica. Según el Teorema de Taylor existe un $\delta > 0$ tal que para todo $\vec{h} = \{h_1, \ldots, h_n\}$ con $\|\vec{h}\| < \delta$ y un $\vartheta = \vartheta(\vec{h}) \in (0,1)$,

$$\begin{aligned} f(x^0 + \vec{h}) - f(x^0) &= (\vec{h} \cdot \nabla f)(x^0) + \frac{1}{2!}(\vec{h} \cdot \nabla f)^2 \left(x^0 + \vartheta(\vec{h})\vec{h}\right) \\ &= \frac{1}{2} \sum_{i,k=1}^{n} f_{x_i x_k}\left(x^0 + \vartheta(\vec{h})\vec{h}\right) h_i h_k. \end{aligned} \tag{3.11}$$

1. Sea $A(x^0)$ definida positiva. Entonces la forma cuadrática

$$\sum_{i,k=1}^{n} f_{x_i x_k}(x^0) h_i h_k,$$

como función continua, asume un mínimo absoluto $m > 0$ sobre el conjunto compacto $S_1 = \{\vec{h} \mid \|\vec{h}\| = 1\}$, es decir

$$\sum_{i,k=1}^{n} f_{x_i x_k}(x^0) h_i h_k \geqslant m > 0 \quad \text{sobre } S_1.$$

3.4. Extremos de funciones de varias variables

Demostración del Teorema 3.10 (continuación)

1. Como todas $f_{x_i x_k}(x)$ son continuas en x^0, existe $\delta' \in (0, \delta)$ tal que para todo x con $d(x, x^0) < \delta'$ y todo $\vec{h} \in S_1$,

$$\sum_{i,k=1}^{n} f_{x_i x_k}(x) h_i h_k > 0.$$

Según el Teorema 3.9, entonces la forma cuadrática

$$\sum_{i,k=1}^{n} f_{x_i x_k}(x) h_i h_k$$

es definida positiva para todo x con $d(x, x^0) < \delta'$. Como consecuencia de (3.11), aplicado para todo $\vec{h}$ con $\|\vec{h}\| < \delta'$, obtenemos que

$$f(x^0 + \vec{h}) - f(x^0) > 0,$$

es decir f posee un mínimo relativo en x^0.

3.4. Extremos de funciones de varias variables

Demostración del Teorema 3.10 (continuación)

1. Si $A(x^0)$ es definida negativa, entonces $-A(x^0)$ es definida positiva, por lo tanto $-f$ posee un mínimo local en x^0, es decir, f posee un máximo local.
2. Si f posee un mínimo relativo en x^0, entonces existe un $\delta'' \in (0, \delta)$ tal que

$$\sum_{i,k=1}^{n} f_{x_i x_k}\big(x^0 + \vartheta(\vec{h})\vec{h}\big) h_i h_k \geqslant 0$$

para todo $\vec{h}$ con $\|\vec{h}\| < \delta''$. Fijemos un tal $\vec{h}$, entonces para $\lambda \in (0, 1)$ arbitrario,

$$\sum_{i,k=1}^{n} f_{x_i x_k}\big(x^0 + \vartheta(\lambda\vec{h})\lambda\vec{h}\big)(\lambda h_i)(\lambda h_k) \geqslant 0.$$

3.4. Extremos de funciones de varias variables

Demostración del Teorema 3.10 (continuación)

2. Después de la división por λ^2 y tomando $\lambda \to 0$ obtenemos

$$\sum_{i,k=1}^{n} f_{x_i x_k}(x^0) h_i h_k \geqslant 0;$$

   esto es válido para todo $\vec{h}$ con $\|\vec{h}\| < \delta''$. Según el Teorema 3.9, la matriz $A(x^0)$ es semidefinida positiva. Por otro lado, si f posee un máximo relativo en x^0, entonces $-f$ posee un mínimo relativo, por lo tanto $-A(x^0)$ es semidefinida positiva y $A(x^0)$ es semidefinida negativa.

3. Si $A(x^0)$ es indefinida podemos concluir de (1) y (2) que f no puede tener ni un máximo relativo ni un mínimo relativo. ■

3.4. Extremos de funciones de varias variables

Teorema 3.11 Se considera la función $f : \mathbb{R}^2 \to \mathbb{R}$. Sea $(x_0, y_0) \in D(f)$, y en para una vecindad $U(x_0, y_0)$ de (x_0, y_0) sea $f \in C^2(U(x_0, y_0))$. Sea $\nabla f(x_0, y_0) = \vec{0}$, y el discriminante $\delta(x_0, y_0)$ definido por

$$\delta(x_0, y_0) := f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - \big(f_{xy}(x_0, y_0)\big)^2.$$

1. Si $\delta(x_0, y_0) > 0$, entonces f posee un extremo en (x_0, y_0). Se trata de un máximo relativo si $f_{xx}(x_0, y_0) < 0$ y de un mínimo relativo si $f_{xx}(x_0, y_0) > 0$.
2. Si $\delta(x_0, y_0) < 0$, entonces f no posee ningún extremo en (x_0, y_0).

Demostración del Teorema 3.11

Consideremos la matriz simétrica

$$A = \begin{bmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{bmatrix}.$$

Utilizando el Teorema 3.7 obtenemos lo siguiente.

1. Si $\delta(x_0, y_0) > 0$ y $f_{xx}(x_0, y_0) < 0$, entonces A es definida negativa, y según el Teorema 3.10 la función f posee un máximo local en (x_0, y_0).
2. Si $\delta(x_0, y_0) > 0$ y $f_{xx}(x_0, y_0) > 0$, entonces A es definida positiva, y según el Teorema 3.10 la función f posee un mínimo local en (x_0, y_0).
3. Si $\delta(x_0, y_0) < 0$, entonces A es indefinida, y según el Teorema 3.10 la función f no posee ni un máximo local ni un mínimo local en (x_0, y_0). $\blacksquare$

3.4. Extremos de funciones de varias variables

Ejemplo 3.10 Consideremos la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x,y) = xy + x - y + 1, \quad D(f) = \mathbb{R}^2.$$

Aquí obtenemos las derivadas parciales

$$\begin{aligned} &f_x(x,y) = y + 1, \quad f_y(x,y) = x - 1, \\ &f_{xy}(x,y) = 1, \quad f_{xx}(x,y) = f_{yy}(x,y) = 0. \end{aligned}$$

Los únicos candidatos a ser extremos de f son aquellos puntos (x_0, y_0) donde $f_x(x_0,y_0) = f_y(x_0,y_0) = 0$; en este caso el único punto con esta propiedad es $(1,-1)$.

$$\delta(1,-1) = f_{xx}(1,-1)f_{yy}(1,-1) - \big(f_{xy}(1,-1)\big)^2 = -1 < 0,$$

entonces según el Teorema 3.11, f no posee ningún extremo.

Ejemplo 3.11 Consideremos la función $f : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$f(x,y) = x^3 - 12xy + 8y^3, \quad D(f) = \mathbb{R}^2.$$

Aquí obtenemos las derivadas parciales

$$\begin{aligned} &f_x(x,y) = 3x^2 - 12y, \quad f_y(x,y) = -12x + 24y^2, \\ &f_{xy}(x,y) = -12, \quad f_{xx}(x,y) = 6x, \quad f_{yy}(x,y) = 48y. \end{aligned}$$

Los puntos extremos (x_0, y_0) de f deben satisfacer

$$\begin{aligned} 3x_0^2 - 12y_0 &= 0, \\ -12x_0 + 24y_0^2 &= 0. \end{aligned}$$

La primera ecuación implica que $y_0 = \frac{1}{4}x_0^2$. Insertando esto en la segunda ecuación obtenemos

$$-12x_0 + \frac{24}{16}x_0^4 = 0 \iff x_0^4 - 8x_0 = 0.$$

Esta ecuación tiene solamente las soluciones $x_0^{(1)} = 0$ y $x_0^{(2)} = 2$, es decir los únicos puntos que hay que examinar son $(0,0)$ y $(2,1)$.

3.4. Extremos de funciones de varias variables

Ejemplo 3.11 (continuación)

$$\delta(0,0) = f_{xx}(0,0)f_{yy}(0,0) - \left(f_{xy}(0,0)\right)^2 = -144 < 0,$$

por lo tanto f no posee un extremo en $(0,0)$, mientras que

$$\delta(2,1) = f_{xx}(2,1)f_{yy}(2,1) - \left(f_{xy}(2,1)\right)^2 = 432 > 0,$$

es decir, f posee un extremo local en $(2,1)$; puesto que $f_{xx}(2,1) = 12 > 0$, se trata de un mínimo local.

3.5. Extremos con restricciones

El Teorema de las Funciones Implícitas nos permite el tratmiento de extremos de funciones sujetos a restricciones.

En muchas aplicaciones no solamente queremos estudiar una función $f : \mathbb{R}^n \to \mathbb{R}$, sino que se plantean una o varias restricciones adicionales a las cuales las variables estan sujetas. Las restricciones están dadas en la forma de un sistema de ecuaciones

$$\begin{aligned} g_1(x_1, \ldots, x_n) = g_1(x) &= 0, \\ &\vdots \\ g_m(x_1, \ldots, x_n) = g_m(x) &= 0 \end{aligned}$$

o brevemente

$$g(x) = 0, \quad g = (g_1, \ldots, g_m) : \mathbb{R}^n \to \mathbb{R}^m.$$

Se dice que f posee en x^0 un máximo local sujeto a las restricciones definidas por g si existe una vecindad $U(x^0)$ de x^0 tal que $f(x) \leqslant f(x^0)$ para todo $x \in U(x^0)$ tal que $g_1(x) = 0, \ldots, g_m(x) = 0$. Una defincón análoga es válida para un mínimo sujeto a una restricción.

3.5. Extremos con restricciones

Para la determinación de los extremos de una función diferenciable $f : \mathbb{R}^n \to \mathbb{R}$ sin restricción nos interesan solamente aquellos puntos x^0 que satisfacen $\nabla f(x^0) = \vec{0}$. Hay una caracterización similar de aquellos puntos que son candidatos a ser extremo de f sujeto a la restricción $g(x) = 0$.

Teorema 3.12 Se consideran $f : \mathbb{R}^n \to \mathbb{R}$ y $g = (g_1, \ldots, g_m) : \mathbb{R}^n \to \mathbb{R}^m$ con $m < n$. Sea $x^0 = (x_1^0, \ldots, x_n^0) \in D(f) \cap D(g)$, y sobre una vecindad $U(x^0)$ de x^0 sea $f \in C^1(U(x^0))$ y $g_\mu \in C^1(U(x^0))$ para $\mu = 1, \ldots, m$. Sea el rango de la matriz

$$\left(\frac{\partial g_\mu}{\partial x_\nu}(x^0) \right)_{\substack{\mu=1,\ldots,m \\ \nu=1,\ldots,n}}$$

igual m. Supongamos que la función f posee un extremo local en x^0 bajo la restricción $g(x) = 0$. Entonces existen constantes $\lambda_1, \ldots, \lambda_m$, llamadas multiplicadores de Lagrange, tales que

$$\nabla f(x^0) = \lambda_1 \nabla g_1(x^0) + \cdots + \lambda_m \nabla g_m(x^0).$$

3.5. Extremos con restricciones

El Teorema 3.12 entrega solamente un criterio necesario. Con su ayuda podemos solamente determinar aquellos puntos $(x_1^0, \dots, x_n^0)$ donde podría existir un extremo local de f sujeto a la restricción

$$g_1(x_1, \dots, x_n) = 0, \tag{3.18}$$
$$\vdots \tag{3.19}$$
$$g_m(x_1, \dots, x_n) = 0. \tag{3.20}$$

Para tal efecto se forman las $n + m$ ecuaciones

$$g_\mu(x_1^0, \dots, x_n^0) = 0, \quad \mu = 1, \dots, m,$$
$$\frac{\partial f}{\partial x_\nu}(x_1^0, \dots, x_n^0) = \lambda_1 \frac{\partial g_1}{\partial x_\nu}(x_1^0, \dots, x_n^0) + \dots + \lambda_m \frac{\partial g_m}{\partial x_\nu}(x_1^0, \dots, x_n^0),$$
$$\nu = 1, \dots, n,$$

la cuales permiten determinar $\lambda_1, \dots, \lambda_m$ y $x_1^0, \dots, x_n^0$.

3.5. Extremos con restricciones

Para el tratamiento de extremos con restricciones no existen criterios simples suficientes (tales como para problemas de extremos sin restricciones) para poder decidir si en un punto $(x_1^0, \dots, x_n^0)$ efectivamente se tiene un extremo local de f sujeto a la restricción (3.18). Hay que decidir esto considerando la forma particular de f o utilizando consideraciones geométricas.

Ejemplo 3.14 Queremos estudiar los extremos de la función

$$f(x, y, z) = x + y + z$$

bajo las restricciones

$$g_1(x, y, z) = x^2 + y^2 - 2 = 0,$$
$$g_2(x, y, z) = x + z - 1 = 0.$$

3.5. Extremos con restricciones

Ejemplo 3.14 (contiuación)
Según el Teorema 3.12, los puntos críticos (candidatos a extremo) son aquellos puntos (x_0, y_0, z_0) que satisfacen

$$\begin{aligned} g_1(x_0, y_0, z_0) &= 0, \\ g_2(x_0, y_0, z_0) &= 0, \\ \nabla f(x_0, y_0, z_0) &= \lambda_1 \nabla g_1(x_0, y_0, z_0) + \lambda_2 \nabla g_2(x_0, y_0, z_0). \end{aligned}$$

3.5. Extremos con restricciones

Ejemplo 3.14 (contiuación)
Aquí obtenemos las cinco ecuaciones

$$
\begin{aligned}
x_0^2 + y_0^2 - 2 &= 0, && (3.21)\\
x_0 + z_0 - 1 &= 0, && (3.22)\\
1 &= \lambda_1 \cdot 2x_0 + \lambda_2 \cdot 1, && (3.23)\\
1 &= \lambda_1 \cdot 2y_0 + \lambda_2 \cdot 0, && (3.24)\\
1 &= \lambda_1 \cdot 0 + \lambda_2 \cdot 1. && (3.25)
\end{aligned}
$$

De estas ecuaciones debemos determinar x_0, y_0, z_0, λ_1 y λ_2. De (3.25) obtenemos $\lambda_2 = 1$, por lo tanto (3.23) entrega que $2\lambda_1 x_0 = 0$ y (3.24) implica que $2\lambda_1 y_0 = 1$, es decir $\lambda_1 \neq 0$ y por lo tanto $x_0 = 0$, luego $y_0 = \sqrt{2}$ o $y_0 = -\sqrt{2}$ y $z_0 = 1$. Los puntos críticos son $(0, \sqrt{2}, 1)$ y $(0, -\sqrt{2}, 1)$. En este caso, el punto $(0, \sqrt{2}, 1)$ corresponde a un máximo y el punto $(0, -\sqrt{2}, 1)$ a un mínimo.

3.5. Extremos con restricciones

Demostración del Teorema 3.12

1. Sin pérdida de la generalidad podemos suponer que

$$\begin{vmatrix} \partial g_1 \partial x_1(x^0) & \cdots & \partial g_1/\partial x_m(x^0) \\ \vdots & & \vdots \\ \partial g_m/x_1(x^0) & \cdots & \partial g_m/\partial x_m(x^0) \end{vmatrix} \neq 0. \tag{3.12}$$

Según el Teorema 3.2 podemos localmente despejar $x_1, \ldots, x_m$ del sistema de ecuaciones

$$g_1(x_1, \ldots, x_n) = 0, \ldots, g_m(x_1, \ldots, x_n) = 0. \tag{3.13}$$

Entonces, existen un $\delta > 0$ y m funciones

$$x_1 = \varphi_1(x_{m+1}, \ldots, x_n), \ldots, x_m = \varphi_m(x_{m+1}, \ldots, x_n)$$

que son C^1 para $|x_\nu - x_\nu^0| < \delta$ $(\nu = m+1, \ldots, n)$ con

$$\begin{aligned} g_1\big(\varphi_1(x_{m+1}, \ldots, x_n), \ldots, \varphi_m(\ldots), x_{m+1}, \ldots, x_n\big) &= 0, \\ &\vdots \\ g_m\big(\varphi_1(x_{m+1}, \ldots, x_n), \ldots, \varphi_m(\ldots), x_{m+1}, \ldots, x_n\big) &= 0. \end{aligned} \tag{3.14}$$

3.5. Extremos con restricciones

Demostración del Teorema 3.12 (continuación)

1. Insertando estas funciones en f, obtenemos una función de las variables $x_{m+1}, \dots, x_n$:

$$F(x_{m+1}, \dots, x_n) := f\big(\varphi_1(x_{m+1}, \dots, x_n), \dots, \varphi_m(x_{m+1}, \dots, x_n), \\ x_{m+1}, \dots, x_n\big).$$

Según la hipótesis, esta función posee un extremo local en $\xi^0 = (x^0_{m+1}, \dots, x^0_n)$, por lo tanto

$$\nabla F(\xi^0) = \vec{0}.$$

Utilizando la regla de la cadena obtenemos

$$\frac{\partial F}{\partial x_\nu}(\xi^0) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x^0) \frac{\partial \varphi_i}{\partial x_\nu}(\xi^0) + \frac{\partial f}{\partial x_\nu}(x^0) = 0, \\ \nu = m+1, \dots, n. \tag{3.15}$$

Demostración del Teorema 3.12 (continuación)

2. En virtud de (3.12), el sistema

$$\sum_{\mu=1}^{m} \lambda_\mu \frac{\partial g_\mu}{\partial x_i}(x^0) = \frac{\partial f}{\partial x_i}(x^0), \quad i = 1, \dots, m \tag{3.16}$$

posee una solución única $(\lambda_1, \dots, \lambda_m)$. De (3.14) obtenemos

$$\begin{gathered} \sum_{i=1}^{m} \frac{\partial g_\mu}{\partial x_i}(x^0) \frac{\partial \varphi_i}{\partial x_\nu}(\xi^0) + \frac{\partial g_\mu}{\partial x_\nu}(x^0) = 0, \\ \nu = m+1, \dots, n, \quad \mu = 1, \dots, m. \end{gathered} \tag{3.17}$$

A partir (3.15) y (3.16),

$$\begin{aligned} \frac{\partial f}{\partial x_\nu}(x^0) &= -\sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(x^0) \frac{\partial \varphi_i}{\partial x_\nu}(\xi^0) = -\sum_{i=1}^{m} \left(\sum_{\mu=1}^{m} \lambda_\mu \frac{\partial g_\mu}{\partial x_i}(x^0) \right) \frac{\partial \varphi_i}{\partial x_\nu}(\xi^0) \\ &= -\sum_{\mu=1}^{m} \lambda_\mu \left(\sum_{i=1}^{m} \frac{\partial g_\mu}{\partial x_i}(x^0) \frac{\partial \varphi_i}{\partial x_\nu}(\xi^0) \right). \end{aligned}$$

3.5. Extremos con restricciones

Demostración del Teorema 3.12 (continuación)

2. En virtud de (3.17), esto implica que

$$\frac{\partial f}{\partial x_\nu}(x^0) = \sum_{\mu=1}^{m} \lambda_\mu \frac{\partial g_\mu}{\partial x_\nu}(x^0), \quad \nu = m+1, \dots, n.$$

Resumiendo (a) y (b) obtenemos

$$\frac{\partial f}{\partial x_\nu}(x^0) = \sum_{\mu=1}^{m} \lambda_\mu \frac{\partial g_\mu}{\partial x_\nu}(x^0), \quad \nu = 1, \dots, n,$$

lo que significa que

$$\nabla f(x^0) = \sum_{\mu=1}^{m} \lambda_\mu \nabla g_\mu(x^0). \quad \blacksquare$$