Extremos de funciones de varias variables

R. Álvarez-Nodarse
Universidad de Sevilla

¿Cuándo una función $f(x)$ de una variable tiene extremo?

Definición

Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ con A abierto o cerrado.

1. *Si $f(x) \leq f(a)$ ($f(x) \geq f(a)$), $\forall x \in A$, $x \neq a$, decimos que f alcanza en el punto a el máximo (mínimo)* **absoluto** *en A.*
2. *Si existe un abierto $B \subset A$ t.q. $\forall x \in B$, $x \neq a$, $f(x) \leq f(a)$ ($f(x) \geq f(a)$) decimos que f alcanza en a un máximo (mínimo)* **relativo**.

Si las desigualdades sean estrictas los extremos son estrictos.

¿Cuándo una función $f(x)$ de una variable tiene extremo?

Definición

Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ con A abierto o cerrado.

1. *Si $f(x) \leq f(a)$ ($f(x) \geq f(a)$), $\forall x \in A$, $x \neq a$, decimos que f alcanza en el punto a el máximo (mínimo)* **absoluto** *en A.*
2. *Si existe un abierto $B \subset A$ t.q. $\forall x \in B$, $x \neq a$, $f(x) \leq f(a)$ ($f(x) \geq f(a)$) decimos que f alcanza en a un máximo (mínimo)* **relativo***.*

Si las desigualdades sean estrictas los extremos son estrictos.

De lo anterior se deduce que todo extremo absoluto es un extremo relativo si este se encuentra en el interior de A.

En general, los extremos absolutos no tienen porque ser extremos relativos (si el extremo absoluto se alcanza el $x = a$ con $a \in \partial A$ no tiene por qué existir ninguna $B(a, \delta) \subset A$) ...

¿Cuándo una función $f(x)$ de una variable tiene extremo?

Definición

Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ con A abierto o cerrado.

1. *Si $f(x) \leq f(a)$ $(f(x) \geq f(a))$, $\forall x \in A$, $x \neq a$, decimos que f alcanza en el punto a el máximo (mínimo)* **absoluto** *en A.*
2. *Si existe un abierto $B \subset A$ t.q. $\forall x \in B$, $x \neq a$, $f(x) \leq f(a)$ $(f(x) \geq f(a))$ decimos que f alcanza en a un máximo (mínimo)* **relativo**.

Si las desigualdades sean estrictas los extremos son estrictos.

De lo anterior se deduce que todo extremo absoluto es un extremo relativo si este se encuentra en el interior de A.

En general, los extremos absolutos no tienen porque ser extremos relativos (si el extremo absoluto se alcanza el $x = a$ con $a \in \partial A$ no tiene por qué existir ninguna $B(a, \delta) \subset A$) ...ni los extremos relativos tienen por que ser absolutos (el extremo absoluto puede alcanzarse en ∂A).

Condición necesaria de extremo relativo

Teorema

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$, *A abierto y supongamos que* f *tiene en ,* $a \in A$ *un extremo relativo. Entonces, si* $\exists \frac{\partial f}{\partial x_k}$, $k = 1, \dots, n$ *éstas son iguales a cero en a, i.e.,* $\frac{\partial f}{\partial x_k}(a) = 0$, $k = 1, \dots, n$. *Si además* f *es diferenciable en* $a \Rightarrow Df(a) = 0$.

Condición necesaria de extremo relativo

Teorema

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$, A abierto y supongamos que f tiene en , $a \in A$ un extremo relativo. Entonces, si $\exists \frac{\partial f}{\partial x_k}$, $k = 1, \dots, n$ éstas son iguales a cero en a, i.e., $\frac{\partial f}{\partial x_k}(a) = 0$, $k = 1, \dots, n$. Si además f es diferenciable en $a \Rightarrow Df(a) = 0$.

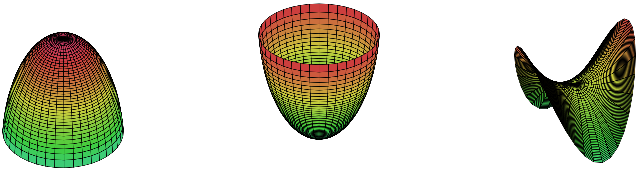
Demostración: Sea $\phi_k(x_k) := f(a_1, \dots, a_{k-1}, x_k, a_{k+1}, \dots, x_n)$. ϕ_k es diferenciable y tiene un extremo en $x_k = a_k$. Aplicamos el lema de Fermat a ϕ_k y obtenemos resultado. $\square$

Condición necesaria de extremo relativo

Teorema

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$, A abierto y supongamos que f tiene en , $a \in A$ un extremo relativo. Entonces, si $\exists \frac{\partial f}{\partial x_k}$, $k = 1, \ldots, n$ éstas son iguales a cero en a, i.e., $\frac{\partial f}{\partial x_k}(a) = 0$, $k = 1, \ldots, n$. Si además f es diferenciable en $a \Rightarrow Df(a) = 0$.

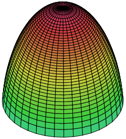
Demostración: Sea $\phi_k(x_k) := f(a_1, \ldots, a_{k-1}, x_k, a_{k+1}, \ldots, x_n)$. ϕ_k es diferenciable y tiene un extremo en $x_k = a_k$. Aplicamos el lema de Fermat a ϕ_k y obtenemos resultado. $\square$

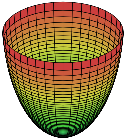


$f : \mathbb{R}^2 \mapsto \mathbb{R}$ con un máximo local, un mínimo local y un punto silla

Ejemplo: Sea la función $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}$, $A : \{(x,y) | x^2 + y^2 < 1\}$, $f(x,y) = \sqrt{1 - x^2 - y^2}$. Es *obvio* que tiene un máximo en $(0,0)$. Un cálculo directo muestra que $\frac{\partial f}{\partial x}(0,0) = 0$, $\frac{\partial f}{\partial y}(0,0) = 0$.

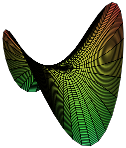
Lo mismo ocurre para la función $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}$, $A : \{(x,y) | x^2 + y^2 < 1\}$, $f(x,y) = -\sqrt{1 - x^2 - y^2}$ que tiene un mínimo local en $(0,0)$.





$f : \mathbb{R}^2 \mapsto \mathbb{R}$ con un máximo local y un mínimo local

Ejemplo: Sea la función $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}$, $f(x,y) = x^2 - y^2$. Es obvio que $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$. Sin embargo, en cualquier entorno de $(0,0)$ que escojamos f toma valores tanto positivos como negativos. En este caso el punto $(0,0)$ se denomina *punto silla* de f.



$f : \mathbb{R}^2 \mapsto \mathbb{R}$ con un punto silla

Para ello tenemos un teorema similar al del caso de una variable. Antes de enunciarlo conviene recordar que la segunda diferencial de una función de varias variables $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$, $f \in C^{(2)}(A)$ es la forma bilineal simétrica, que escribiremos convenientemente de la forma

$$d^2 f(a) := D^2 f(a)(x) = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \frac{\partial^2 f(a)}{\partial x_{i_1} \partial x_{i_2}} x_{i_1} x_{i_2} = x^T H_f(a) x,$$

donde

$$H_f(a) = \begin{pmatrix} \dfrac{\partial^2 f(a)}{\partial^2 x_1} & \cdots & \dfrac{\partial^2 f(a)}{\partial x_n \partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(a)}{\partial x_1 \partial x_n} & \cdots & \dfrac{\partial^2 f(a)}{\partial^2 x_n} \end{pmatrix}.$$

Condición suficiente de extremo

Teorema

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ *dos veces diferenciable en* $a \in A$, *A abierto, y sea* $x = a$ *un punto crítico de* f, *i.e.,* $Df(a) = 0$. *Entonces*

1. *Si la segunda diferencial* $D^2 f(a)(x)$ *es definida positiva en* a, *entonces* f *tiene un mínimo relativo en* a.
2. *Si la segunda diferencial* $D^2 f(a)(x)$ *es definida negativa, entonces* f *tiene un máximo relativo en* a.
3. *Si la segunda diferencial* $D^2 f(a)(x)$ *es indefinida, i.e., si existen* $x, y \in \mathbb{R}^n$ *tales que* $D^2 f(a)(x) > 0 > D^2 f(a)(y)$, *entonces* f *tiene un punto de silla en* a.

Condición suficiente de extremo

$$f(x)-f(a)=\frac{1}{2}\|h\|^2\left(\overbrace{\sum_{i,j=1}^{n}\frac{\partial^2 f(a)}{\partial x_i\partial x_j}\frac{h_i}{\|h\|}\frac{h_j}{\|h\|}}^{Q(x,a)}+o(1)\right)$$

$$\text{signo}(f(x)-f(a)) = \text{signo}(Q(x,a)+o(1))$$

$Q(x,a)$ está definida sobre un cerrado y acotado (esfera hueca de radio 1) entonces alcanza máx y mín absolutos: $m \le Q(x,a) \le M$.

Condición suficiente de extremo

$$f(x)-f(a)=\frac{1}{2}\|h\|^2\left(\overbrace{\sum_{i,j=1}^{n}\frac{\partial^2 f(a)}{\partial x_i\partial x_j}\frac{h_i}{\|h\|}\frac{h_j}{\|h\|}}^{Q(x,a)}+o(1)\right)$$

$$\text{signo}(f(x)-f(a))=\text{signo}(Q(x,a)+o(1))$$

$Q(x,a)$ está definida sobre un cerrado y acotado (esfera hueca de radio 1) entonces alcanza máx y mín absolutos: $m\leq Q(x,a)\leq M$.

Si $D^2f(a)(x)$ es definida positiva en a entonces $m>0$. Pero $o(1)\to 0$ luego en un entorno de a suf. pequeño $o(1)>-m/2$ luego $Q(x,a)+o(1)>0$ $\qquad f(x)-f(a)\geq 0\ \Rightarrow$ mín

Condición suficiente de extremo

$$f(x)-f(a)=\frac{1}{2}\|h\|^2\left(\overbrace{\sum_{i,j=1}^{n}\frac{\partial^2 f(a)}{\partial x_i\partial x_j}\frac{h_i}{\|h\|}\frac{h_j}{\|h\|}}^{Q(x,a)}+o(1)\right)$$

$$\text{signo}(f(x)-f(a))=\text{signo}(Q(x,a)+o(1))$$

$Q(x,a)$ está definida sobre un cerrado y acotado (esfera hueca de radio 1) entonces alcanza máx y mín absolutos: $m\leq Q(x,a)\leq M$.

Si $D^2f(a)(x)$ es definida positiva en a entonces $m>0$. Pero $o(1)\to 0$ luego en un entorno de a suf. pequeño $o(1)>-m/2$ luego $Q(x,a)+o(1)>0$ $\qquad f(x)-f(a)\geq 0 \Rightarrow$ mín

Si $D^2f(a)(x)$ es definida negativa en a entonces $M<0$. Pero $o(1)\to 0$ luego en un entorno de a suf. pequeño $o(1)<|M|/2$ luego $Q(x,a)+o(1)<0$ $\qquad f(x)-f(a)\leq 0 \Rightarrow$ máx

Condición suficiente de extremo

$$f(x) - f(a) = \frac{1}{2}\|h\|^2 \left(\overbrace{\sum_{i,j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i \partial x_j} \frac{h_i}{\|h\|} \frac{h_j}{\|h\|}}^{Q(x,a)} + o(1) \right)$$

$$\text{signo}(f(x) - f(a)) = \text{signo}(Q(x,a) + o(1))$$

Sean a_m y a_M los puntos donde Q alcanza su mínimo y máximo absolutos sobre S. Como Q es indefinida $m < 0 < M$.

Condición suficiente de extremo

$$f(x) - f(a) = \frac{1}{2}\|h\|^2 \left(\overbrace{\sum_{i,j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i \partial x_j} \frac{h_i}{\|h\|} \frac{h_j}{\|h\|}}^{Q(x,a)} + o(1) \right)$$

$$\text{signo}(f(x) - f(a)) = \text{signo}(Q(x,a) + o(1))$$

Sean a_m y a_M los puntos donde Q alcanza su mínimo y máximo absolutos sobre S. Como Q es indefinida $m < 0 < M$.

Si elegimos $h = x - a = ta_m$ con $t \ll 1$ tendremos $f(x) - f(a) = \frac{t^2}{2}(m + o(1))$ y para ciertos valores de x en un entorno de a signo$(f(x) - f(a)) = $ signo$(m) < 0$.

Condición suficiente de extremo

$$f(x) - f(a) = \frac{1}{2}\|h\|^2 \left(\overbrace{\sum_{i,j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i \partial x_j} \frac{h_i}{\|h\|} \frac{h_j}{\|h\|}}^{Q(x,a)} + o(1) \right)$$

$$\text{signo}(f(x) - f(a)) = \text{signo}(Q(x,a) + o(1))$$

Sean a_m y a_M los puntos donde Q alcanza su mínimo y máximo absolutos sobre S. Como Q es indefinida $m < 0 < M$.

Si elegimos $h = x - a = ta_m$ con $t \ll 1$ tendremos $f(x) - f(a) = \frac{t^2}{2}(m + o(1))$ y para ciertos valores de x en un entorno de a signo$(f(x) - f(a)) =$ signo$(m) < 0$.

Repitiendo el razonamiento para $h = x - a = ta_M$ con $t \ll 1$ obtendremos que para ciertos x en un entorno de a $f(x) - f(a) > 0$, luego en a no puede haber ningún extremo.

Condición suficiente de extremo

Teorema

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ dos veces diferenciable en $a \in A$, A abierto, y sea $x = a$ un punto crítico de f, i.e., $Df(a) = 0$. Entonces

1. *Si la segunda diferencial $D^2 f(a)(x)$ es definida positiva en a, entonces f tiene un mínimo relativo en a.*
2. *Si la segunda diferencial $D^2 f(a)(x)$ es definida negativa, entonces f tiene un máximo relativo en a.*
3. *Si la segunda diferencial $D^2 f(a)(x)$ es indefinida, i.e., si existen $x, y \in \mathbb{R}^n$ tales que $D^2 f(a)(x) > 0 > D^2 f(a)(y)$, entonces f tiene un punto de silla en a.*

¿Cómo saber el signo de $D^2 f(a)(x)$?

Sea $B(x,y)$ una aplicación bilineal simétrica y sea $B = [b_{i,j}]_{i,j=1,n}$ su matriz. Las siguientes condiciones son equivalentes:

1. B es definida positiva.
2. Todos los autovalores de B son positivos.
3. Los menores principales Δ_k de B son positivos, i.e. $\Delta_k > 0$, $k = 1, 2, \dots, n$ donde

$$\Delta_k := \det \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,k} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,k} \\ \\ \vdots & \ddots & \vdots & \\ b_{k,1} & b_{k,2} & \cdots & b_{1,k} \end{pmatrix}, \quad \forall k = 1, 2, \dots n.$$

Análogamente se tiene para las formas biliniales definidas negativas las siguientes condiciones equivalentes:

1. B es definida negativa.
2. Todos los autovalores de B son negantivos.
3. Los menores principales Δ_k de B son tales que $(-1)^k \Delta_k > 0$, $k = 1, 2, \dots, n$.

Corolario (Condición suficiente de extremo)

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ *dos veces diferenciable en* $a \in A$, *A abierto, y sea* $x = a$ *un punto crítico de* f, *i.e.,* $Df(a) = 0$ *y sea*

$$\Delta_k := \det \begin{pmatrix} \dfrac{\partial^2 f(a)}{\partial^2 x_1} & \cdots & \dfrac{\partial^2 f(a)}{\partial x_k \partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(a)}{\partial x_1 \partial x_k} & \cdots & \dfrac{\partial^2 f(a)}{\partial^2 x_k} \end{pmatrix}.$$

1. *Si todos los menores principales* $\Delta_k > 0$, $k = 1, 2, \ldots, n$, *entonces* f *tiene un mínimo relativo en* a.
2. *Si todos los menores principales son tales que* $(-1)^k \Delta_k > 0$, $k = 1, 2, \ldots, n$, *entonces* f *tiene un máximo relativo en* a.

Corolario (Condición suficiente de extremo)

Sea $f : A \subset \mathbb{R}^n \to \mathbb{R}$ dos veces diferenciable en $a \in A$, A abierto, y sea $x = a$ un punto crítico de f, i.e., $Df(a) = 0$ y sea

$$\Delta_k := \det \begin{pmatrix} \dfrac{\partial^2 f(a)}{\partial^2 x_1} & \cdots & \dfrac{\partial^2 f(a)}{\partial x_k \partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(a)}{\partial x_1 \partial x_k} & \cdots & \dfrac{\partial^2 f(a)}{\partial^2 x_k} \end{pmatrix}.$$

1. *Si todos los menores principales $\Delta_k > 0$, $k = 1, 2, \ldots, n$, entonces f tiene un mínimo relativo en a.*
2. *Si todos los menores principales son tales que $(-1)^k \Delta_k > 0$, $k = 1, 2, \ldots, n$, entonces f tiene un máximo relativo en a.*

En el caso especial de dos variables se puede ir más allá:

Corolario

Sea $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ *dos veces diferenciable en* $a \in A$, A *abierto,* $Df(a) = 0$.

1. *Si* $\dfrac{\partial^2 f(a)}{\partial^2 x_1} > 0$ *y* $\det \begin{pmatrix} \dfrac{\partial^2 f(a)}{\partial^2 x_1} & \dfrac{\partial^2 f(a)}{\partial x_2 \partial x_1} \\ \dfrac{\partial^2 f(a)}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f(a)}{\partial^2 x_2} \end{pmatrix} > 0$, f *tiene un mínimo relativo en* a.
2. *Si* $\dfrac{\partial^2 f(a)}{\partial^2 x_1} < 0$ *y* $\det \begin{pmatrix} \dfrac{\partial^2 f(a)}{\partial^2 x_1} & \dfrac{\partial^2 f(a)}{\partial x_2 \partial x_1} \\ \dfrac{\partial^2 f(a)}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f(a)}{\partial^2 x_2} \end{pmatrix} > 0$, f *tiene un máximo relativo en* a.
3. *Si* $\det \begin{pmatrix} \dfrac{\partial^2 f(a)}{\partial^2 x_1} & \dfrac{\partial^2 f(a)}{\partial x_2 \partial x_1} \\ \dfrac{\partial^2 f(a)}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f(a)}{\partial^2 x_2} \end{pmatrix} < 0$, f *tiene un punto silla en* a.
4. *Si el* $\det H_f(a) = 0$, *nada puede decirse.*

Ejemplos

Calcular los extremos de $f(x,y) = x^4 + y^4 + 6x^2y^2 + 8x^3$.

Ejemplos

Calcular los extremos de $f(x, y) = x^4 + y^4 + 6x^2y^2 + 8x^3$.

$(0, 0)$ $(-6, 0)$

Ejemplos

Calcular los extremos de $f(x,y) = x^4 + y^4 + 6x^2y^2 + 8x^3$.

$(0,0)$ $(-6,0)$

Calcular los extremos de $f(x,y,z) = x^4 + y^4 - 2(x^2 + z^2) + y^2$

Ejemplos

Calcular los extremos de $f(x,y) = x^4 + y^4 + 6x^2y^2 + 8x^3$.

$(0,0)$ $(-6,0)$

Calcular los extremos de $f(x,y,z) = x^4 + y^4 - 2(x^2 + z^2) + y^2$

$(0,0,0)$ $(0,0,\pm 1)$ $(\pm 1,0,0)$ $(\pm 1,0,1)$ $(-1,0,\pm 1)$

Ejemplos

Calcular los extremos de $f(x,y) = x^4 + y^4 + 6x^2y^2 + 8x^3$.

$(0,0)$ $(-6,0)$

Calcular los extremos de $f(x,y,z) = x^4 + y^4 - 2(x^2 + z^2) + y^2$

$(0,0,0)$ $(0,0,\pm 1)$ $(\pm 1,0,0)$ $(\pm 1,0,1)$ $(-1,0,\pm 1)$

Calcular los extremos de $f(x,y) = xy\log(x^2 + y^2)$.

Pasemos ahora a un problema muy relacionado con el anterior. Imaginemos que queremos encontrar los extremos de una función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ donde las variables no son todas independientes sino que han de satisfacer una serie de condiciones de *ligadura* $\Phi_k(x_1, \dots, x_n) = 0$, $k = 1, \dots, m$, $m < n$.

Ejemplo 1: Encontrar en mínimo de $f(x, y) = xy$ si $x^2 + y^2 = 1$.

Ejemplo 2: Hallar la distancia mínima del origen de coordenadas a la cónica de ecuación $x^2 + 3xy + y^2 = 4$.

Pasemos ahora a un problema muy relacionado con el anterior. Imaginemos que queremos encontrar los extremos de una función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ donde las variables no son todas independientes sino que han de satisfacer una serie de condiciones de *ligadura* $\Phi_k(x_1, \ldots, x_n) = 0$, $k = 1, \ldots, m$, $m < n$.

Ejemplo 1: Encontrar en mínimo de $f(x, y) = xy$ si $x^2 + y^2 = 1$.

Ejemplo 2: Hallar la distancia mínima del origen de coordenadas a la cónica de ecuación $x^2 + 3xy + y^2 = 4$.

¿Cómo proceder?

Un ejemplo para centrar ideas

Queremos encontrar el máximo y/o mínimo absolutos de $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}$, si sus variables satisfacen la ecuación $\Phi(x, y) = 0$.

Una forma de resolver el problema es como sigue:

Queremos encontrar el máximo y/o mínimo absolutos de $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}$, si sus variables satisfacen la ecuación $\Phi(x, y) = 0$.

Una forma de resolver el problema es como sigue:

1) resolvemos la ecuación $\Phi(x, y) = 0$ respecto a una variable, digamos $y = g(x)$, y sustituimos la función resultante en nuestra f.

Un ejemplo para centrar ideas

Queremos encontrar el máximo y/o mínimo absolutos de $f : A \subset \mathbb{R}^2 \mapsto \mathbb{R}$, si sus variables satisfacen la ecuación $\Phi(x, y) = 0$.

Una forma de resolver el problema es como sigue:

1) resolvemos la ecuación $\Phi(x, y) = 0$ respecto a una variable, digamos $y = g(x)$, y sustituimos la función resultante en nuestra f.

2) obtenemos una función de una variable $F(x) = f(x, g(x))$ a la que podemos calcularle los extremos al ser x una variable *libre*.

Por el teorema de la función implícita bastaría que $\Phi'_y(x, y) \neq 0$ en A para tener garantizado que exista la función $y = g(x)$.

Veamos como funciona en uno de los ejemplos que mencionamos antes:

Ejemplo 2: Hallar la distancia mínima del origen de coordenadas a la cónica de ecuación $x^2 + 3xy + y^2 = 4$.

Veamos como funciona en uno de los ejemplos que mencionamos antes:

Ejemplo 2: Hallar la distancia mínima del origen de coordenadas a la cónica de ecuación $x^2 + 3xy + y^2 = 4$.

¿Es "bueno" este método?

Veamos como funciona en uno de los ejemplos que mencionamos antes:

Ejemplo 2: Hallar la distancia mínima del origen de coordenadas a la cónica de ecuación $x^2 + 3xy + y^2 = 4$.

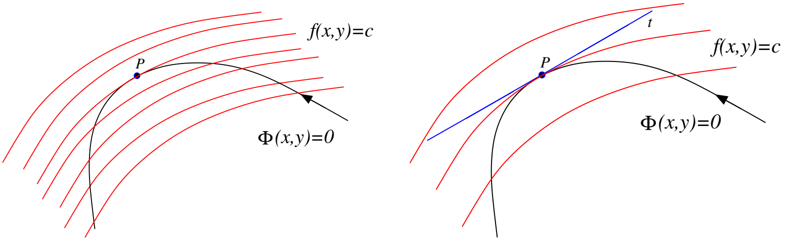
¿Es "bueno" este método? **NO**

Un ejemplo para centrar ideas

Veamos como funciona en uno de los ejemplos que mencionamos antes:

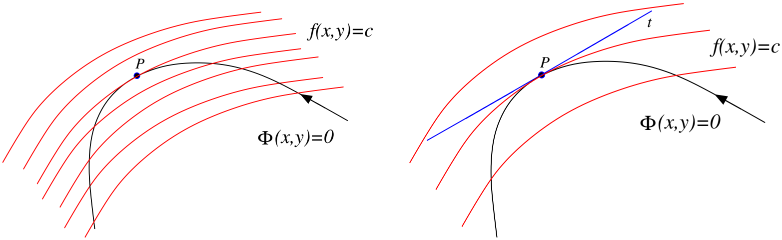
Ejemplo 2: Hallar la distancia mínima del origen de coordenadas a la cónica de ecuación $x^2 + 3xy + y^2 = 4$.

¿Es "bueno" este método? **NO** $g(x)$??? y ¿por qué no $x = h(y)$?



En rojo las curvas de nivel de la función f, i.e., $f(x, y) = c$. Recorramos la curva $\Phi(x, y) = 0$ en contra de las manecillas del reloj y supongamos que f tiene un extremo a lo largo de la curva Φ en P. ¿Qué vemos?

Una forma más elegante de proceder



En rojo las curvas de nivel de la función f, i.e., $f(x,y)=c$.
Recorramos la curva $\Phi(x,y)=0$ en contra de las manecillas del reloj y supongamos que f tiene un extremo a lo largo de la curva Φ en P. ¿Qué vemos?
Si suponemos además que tanto la curva $f(x,y)=c_P$ como $\Phi(x,y)=0$ son *suaves* entonces ambas tienen la misma recta tangente en P.

La pendiente de dicha tangente se calcula, en general, por la fórmula $-f'_x(x_0, y_0)/f'_y(x_0, y_0)$ o bien $-\Phi'_x(x_0, y_0)/\Phi'_y(x_0, y_0)$. Lo anterior nos conduce a

$$\frac{f'_x(x_0, y_0)}{f'_y(x_0, y_0)} = \frac{\Phi'_x(x_0, y_0)}{\Phi'_y(x_0, y_0)} \quad \Leftrightarrow \quad \frac{f'_x(x_0, y_0)}{\Phi'_x(x_0, y_0)} = \frac{f'_y(x_0, y_0)}{\Phi'_y(x_0, y_0)} = -\lambda.$$

Es decir, que si en P hay un extremo de f cuando nos restringimos a la curva $\Phi(x, y) = 0$, entonces ha de cumplirse las siguientes condiciones:

$$\left\{ \begin{array}{r} f'_x(x_0, y_0) + \lambda \Phi'_x(x_0, y_0) = 0, \\ f'_y(x_0, y_0) + \lambda \Phi'_y(x_0, y_0) = 0, \\ \Phi(x, y) = 0, \end{array} \right.$$

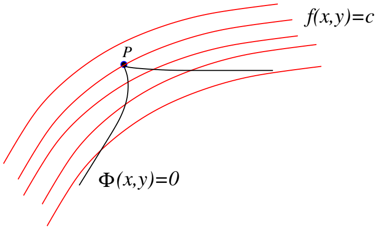
donde λ es cierta constante. O sea, P ha de ser un punto crítico de la función F de tres variables $L(x, y, \lambda) = f(x, y) + \lambda \Phi(x, y)$.

La función L se suele denominar función de Lagrange y la forma de encontrar el extremo según el sistema anterior es conocido como el método de los coeficientes indeterminados de Lagrange.

Nótese que lo anterior sólo nos da condiciones necesarias.

Si querenos una condición suficiente tenemos que calcular el segundo diferencial de f en el punto crítico y luego usar la identidad $d\Phi(x,y) = \Phi'_x(x_0,y_0)dx + \Phi'_y(x_0,y_0)dy = 0$ que relaciona los diferenciales de las dos variables. Sustituyendo esta última relación en la expresión de $d^2f(x_0,y_0)$ obtendremos una forma cuadrática (en este caso de una única variable) cuyo signo determinará el tipo de extremo.

Antes de continuar conviene observar que el método anterior falla si la curva Φ tiene picos pues puede ocurrir que el extremo se alcance justo en ese punto tal y como se muestra en la figura



Problema general

Sea $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ una función $f(x_1, x_2, \ldots, x_m, x_{m+1}, x_n)$ cuyas n variables no son indep. i.e., satisfacen las ec. de ligadura

$$\begin{cases} \Phi_1(x_1, x_2, \cdots, x_n) = 0, \\ \Phi_2(x_1, x_2, \cdots, x_n) = 0, \\ \vdots \\ \Phi_m(x_1, x_2, \cdots, x_n) = 0. \end{cases}$$

Por simplicidad supondremos que todas las ecuaciones de ligadura son independientes. Sea $a \in A$. Asumiremos que el siguiente jacobiano es no nulo en todo un entorno de a

$$\det J_\Phi := \det \begin{pmatrix} \dfrac{\partial \Phi_1(x)}{\partial x_1} & \dfrac{\partial \Phi_1(x)}{\partial x_2} & \cdots & \dfrac{\partial \Phi_1(x)}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial \Phi_m(x)}{\partial x_1} & \dfrac{\partial \Phi_m(x)}{\partial x_2} & \cdots & \dfrac{\partial \Phi_m(x)}{\partial x_m} \end{pmatrix} \neq 0$$

Bajo las condiciones anteriores se tiene

Theorem

Sea la función $f : A \subset \mathbb{R}^n \mapsto \mathbb{R}$ una función de clase $C^{(1)}(A)$ cuyas n variables satisfacen las ecuaciones de ligadura y sea $a \in A$ un extremo de f. Dicho extremo se suele denominar extremo condicionado de f por las ecuaciones de ligadura. Entonces existen m constantes λ_1, λ_2, ..., λ_m reales tales que la función $L : \mathbb{R}^{n+m} \mapsto \mathbb{R}$, que se denomina función de Lagrange,

$$L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = f(x_1, \ldots, x_n) + \lambda_1 \Phi_1(x_1, \ldots, x_n) + \cdots$$
$$+ \lambda_m \Phi_m(x_1, \ldots, x_n)$$

tiene un punto crítico en a.

Un comentario formal

Como el sistema que define las ligaduras tiene solución en el punto $a \in A$, y el jacobiano $\det J_\Phi$ es distinto de cero, entonces por el teorema de la función implícita el sistema que define las ligaduras es resoluble en las variables $x_1, \ldots, x_m$, es decir, en un entorno de $a \in A$ existen las funciones $x_k = g_k(x_{m+1}, \ldots, x_n)$, $k = 1, 2, \ldots, m$ tales que

$$\Phi_k(g_1(x_{m+1}, \ldots, x_n), \ldots, g_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n) = 0$$

son identidades en el entorno de $a \in A$.

Un comentario formal

Como el sistema que define las ligaduras tiene solución en el punto $a \in A$, y el jacobiano $\det J_\Phi$ es distinto de cero, entonces por el teorema de la función implícita el sistema que define las ligaduras es resoluble en las variables $x_1, \ldots, x_m$, es decir, en un entorno de $a \in A$ existen las funciones $x_k = g_k(x_{m+1}, \ldots, x_n)$, $k = 1, 2, \ldots, m$ tales que

$$\Phi_k(g_1(x_{m+1}, \ldots, x_n), \ldots, g_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n) = 0$$

son identidades en el entorno de $a \in A$.

O sea, en las condiciones dadas el problema del cálculo de un extremo condicionado se puede transformar en el de un extremo libre (sin ecuaciones de ligadura) sustituyendo las funciones x_k, $k = 1, \ldots, m$ así obtenidas en la expresión de f, i.e, encontrando los extremos *libres* de la función

$$F(x_{m+1}, \ldots, x_n) := f(g_1(x_{m+1}, \ldots, x_n), \ldots, g_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n).$$

Un comentario formal

Como el sistema que define las ligaduras tiene solución en el punto $a \in A$, y el jacobiano $\det J_\Phi$ es distinto de cero, entonces por el teorema de la función implícita el sistema que define las ligaduras es resoluble en las variables $x_1, \ldots, x_m$, es decir, en un entorno de $a \in A$ existen las funciones $x_k = g_k(x_{m+1}, \ldots, x_n)$, $k = 1, 2, \ldots, m$ tales que

$$\Phi_k(g_1(x_{m+1}, \ldots, x_n), \ldots, g_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n) = 0$$

son identidades en el entorno de $a \in A$.

O sea, en las condiciones dadas el problema del cálculo de un extremo condicionado se puede transformar en el de un extremo libre (sin ecuaciones de ligadura) sustituyendo las funciones x_k, $k = 1, \ldots, m$ así obtenidas en la expresión de f, i.e, encontrando los extremos *libres* de la función

$$F(x_{m+1}, \ldots, x_n) := f(g_1(x_{m+1}, \ldots, x_n), \ldots, g_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n).$$

Esto no es factible. ¿Por qué?

¿Cómo proceder?

Está claro que los extremos de f con las ligaduras son los mismos que los de la función de Lagrange L por tanto la idea es encontrar los puntos críticos de L a partir de sus derivadas parciales, donde ahora las constantes indeterminadas λ_k, $k = 1, \ldots, m$ se consideran variables independientes. Eso nos conduce a un sistema de $n + m$ ecuaciones, a saber

$$\frac{\partial L}{\partial x_k} = 0, \quad k = 1, \ldots, n, \quad \frac{\partial L}{\partial \lambda_i} = 0, \quad i = 1, \ldots, m.$$

Nótese que las ecuaciones $\frac{\partial L}{\partial \lambda_i} = 0$ se transforman en las ecuaciones de ligadura, que sabemos de antemano que han de cumplirse.

Este sistema nos proporciona cierta cantidad de puntos críticos. Supongamos que $a = (x_0, \lambda_0) \in \mathbb{R}^{n+m}$ es uno de dichos puntos críticos. Para saber si dicho punto crítico es un extremo hemos de calcular la segunda diferencial de L en dicho punto:

$$d^2L(a) = \sum_{i,j=1}^{n} \frac{\partial^2 L(a)}{\partial x_i \partial x_j} dx_i dx_j + \sum_{i=1}^{n} \frac{\partial^2 L(a)}{\partial^2 \lambda_i} d\lambda_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{\partial^2 L(a)}{\partial x_i \partial \lambda_j} dx_i d\lambda_j.$$

Este sistema nos proporciona cierta cantidad de puntos críticos. Supongamos que $a = (x_0, \lambda_0) \in \mathbb{R}^{n+m}$ es uno de dichos puntos críticos. Para saber si dicho punto crítico es un extremo hemos de calcular la segunda diferencial de L en dicho punto:

$$d^2L(a) = \sum_{i,j=1}^{n} \frac{\partial^2 L(a)}{\partial x_i \partial x_j} dx_i dx_j + \sum_{i=1}^{n} \frac{\partial^2 L(a)}{\partial^2 \lambda_i} d\lambda_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{\partial^2 L(a)}{\partial x_i \partial \lambda_j} dx_i d\lambda_j.$$

Dado que $\frac{\partial L}{\partial \lambda_i} = \Phi_i(x_1, \cdots, x_n) = 0$ es una identidad, entonces todas las derivadas de orden dos $\frac{\partial^2 L(a)}{\partial^2 \lambda_i} = \frac{\partial^2 L(a)}{\partial x_i \partial \lambda_j} = 0$ por lo que

$$d^2L(a) = d^2 f(a)\Big|_{\Phi_i(x_0)=0,\ i=1,\ldots,m}.$$

Lo anterior nos dice que debemos calcular la segunda diferencial en $a = (x_0, \lambda_0)$, pero teniendo en cuenta que las diferenciales de las variables dx_k, $k = 1, \ldots, n$ no son independientes.

Para ello vamos a escribir las diferenciales de $\Phi_i(x_0)$, $i = 1, \ldots, m$. Tomando diferenciales en ambos lados de las eq. de ligadura tenemos

$$d\Phi_i(x_0) = \sum_{k=1}^{n} \frac{\partial \Phi_i}{\partial x_k}(x_0) dx_k = 0, \quad k = 1, \ldots, m.$$

El sistema anterior es un sistema lineal respecto a $dx_1, \ldots, dx_m$, cuyo determinante (que es el jacobiano) es distinto de cero en un entorno del punto crítico a por lo que existen ciertas funciones lineales $g_j : \mathbb{R}^{n-m} \mapsto \mathbb{R}$, $j = 1, \ldots, m$ tales que $dx_j = g_j(dx_{m+1}, \ldots, dx_n) = A_{j,1} dx_{m+1} + \cdots + A_{j,n-m} dx_n$, $j = 1, \ldots, m$. Es decir *podemos resolverlo respecto a las diferenciales de las variables $x_1, \ldots, x_m$*.

Para ello vamos a escribir las diferenciales de $\Phi_i(x_0)$, $i = 1, \dots, m$. Tomando diferenciales en ambos lados de las eq. de ligadura tenemos

$$d\Phi_i(x_0) = \sum_{k=1}^{n} \frac{\partial \Phi_i}{\partial x_k}(x_0) dx_k = 0, \quad k = 1, \dots, m.$$

El sistema anterior es un sistema lineal respecto a dx_1, ..., dx_m, cuyo determinante (que es el jacobiano) es distinto de cero en un entorno del punto crítico a por lo que existen ciertas funciones lineales $g_j : \mathbb{R}^{n-m} \mapsto \mathbb{R}$, $j = 1, \dots, m$ tales que $dx_j = g_j(dx_{m+1}, \dots, dx_n) = A_{j,1} dx_{m+1} + \cdots + A_{j,n-m} dx_n$, $j = 1, \dots, m$. Es decir *podemos resolverlo respecto a las diferenciales de las variables x_1, ..., x_m*.

Sustituyendo los valores de las diferenciales dx_1, ..., dx_m en la expresión de la segunda diferencial obtenemos la expresión de esta en las variables independientes. Estudiando el signo de la forma cuadrática resultante tal y como se indica en el teorema de la Cond. Suf. podremos decidir si el punto x_0 es un extremo o no de f bajo las condiciones de ligadura.

Ejemplo: Hallar la distancia mínima del origen de coordenadas a la cónica de ecuación $x^2 + 3xy + y^2 = 4$.

Ejemplo: Hallar la distancia mínima del origen de coordenadas a la cónica de ecuación $x^2 + 3xy + y^2 = 4$.

Ejemplo: Encontrar los extremos de la función $f(x, y) = x^2 + y^2$ con la condición de ligadura $(x - 3)^2 + (y - 4)^2 = 10^2$.

Ejemplo: Encontrar el máximo y mínimo absolutos de la función $f(x, y) = x^2 + y^2 - x - y + 1$ en la región definida por $x^2 + y^2 \leq 1$.

Más ejemplos

Problema. Sea la curva definida por la intersección de las superficies

$$x + y + z = 12, \qquad z = x^2 + y^2$$

Encontrar:

1. Los puntos de mayor y menor altura
2. Los punto más cercanos y más lejanos al origen

Más ejemplos

Problema. Sea la curva definida por la intersección de las superficies

$$x + y + z = 12, \qquad z = x^2 + y^2$$

Encontrar:

1. Los puntos de mayor y menor altura
2. Los punto más cercanos y más lejanos al origen

Problema. Sea la curva definida por la intersección de las superficies

$$x^2 + y^2 - z^2 - xy = 1, \qquad x^2 + y^2 = 1.$$

Encontrar los punto más cercanos y más lejanos al origen

Problema. Sea la curva definida por la intersección de las superficies

$$x^2 + y^2 + z^2 = 11, \qquad x + y + z = 3.$$

Encontrar los extremos absolutos de $f(x, y, z) = z^2 + 2x + 2y + 20$ sobre la curva anterior.

Problema. Sea la curva definida por la intersección de las superficies

$$x^2 + y^2 + z^2 = 11, \qquad x + y + z = 3.$$

Encontrar los extremos absolutos de $f(x, y, z) = z^2 + 2x + 2y + 20$ sobre la curva anterior.

Problema. Calcula los puntos del trozo de paraboloide definido por

$$z = x^2 + y^2, \qquad z \leq 3$$

que son más cercanos y más lejanos del punto $(3, 3, 1)$.