Matemáticas III
Tema 2
Optimización en campos escalares

Rodríguez Sánchez, F.J.
Muñoz Ruiz, M.L.
Merino Córdoba, S.

Formas cuadráticas y matrices simétricas reales

Sea A una matriz cuadrada simétrica de tamaño 2×2. La *forma cuadrática* asociada a dicha matriz es el campo escalar definido por

$$f_A(x, y) = \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix}$$

Igualmente para formas cuadráticas de $\mathbb{R}^n$, con $n > 2$.
Una forma cuadrática es siempre un polinomio de grado 2.

Ejemplo.

Si $A = \begin{pmatrix} 1 & -2 \\ -2 & -1 \end{pmatrix}$, entonces

$$f_A(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} =$$
$$= x^2 - 4xy - y^2$$

El signo de la forma cuadrática asociada permite clasificar las matrices simétricas.

1. Se dice que A es definida positiva si su forma cuadrática asociada es siempre positiva, es decir $f_A(x, y) > 0$ para todo $(x, y) \in \mathbb{R}^2$ no nulo.
2. Se dice que A es definida negativa si su forma cuadrática asociada es siempre negativa, es decir $f_A(x, y) < 0$ para todo $(x, y) \in \mathbb{R}^2$ no nulo.
3. Se dice que A es indefinida si su forma cuadrática asociada cambia de signo, es decir existen $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ tales que $f_A(x_1, y_1) > 0$ y $f_A(x_2, y_2) < 0$.
4. Existen matrices A que no son ni definidas ni indefinidas, a dichas matrices se les llama semidefinidas positivas si $f_A(x, y) \geq 0$ para todo $(x, y) \in \mathbb{R}^2$ y semidefinidas negativas si $f_A(x, y) \leq 0$ para todo $(x, y) \in \mathbb{R}^2$.

Determinación del signo de una matriz simétrica

Sea A una matriz simétrica

1. A es definida positiva si y sólo si todos sus autovalores son positivos.
2. A es definida negativa si y sólo si todos sus autovalores son negativos.
3. A es indefinida si y sólo si tiene autovalores negativos y positivos.
4. A es semidefinida positiva si los autovalores son no negativos y alguno es cero.
5. A es semidefinida negativa si los autovalores son no positivos y alguno es cero.

Criterio de Sylvester

Sea $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ y los menores principales

$$\Delta_0 = a_{11}, \quad \Delta_1 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Teorema

Sea A es una matriz simétrica de orden n. Entonces:

1. *A es definida positiva $\iff \Delta_i > 0$ para todo $i = 0, 1, \dots n$.*
2. *A es definida negativa $\iff (-1)^n \Delta_i < 0$ para todo $i = 0, 1, \dots n$.*

En cualquier otro caso la matriz es indefinida o semidefinida.

Ejemplo

1. La matriz simétrica $A = \begin{pmatrix} 1 & -1 \\ -1 & 5 \end{pmatrix}$ es definida positiva
2. La matriz $\begin{pmatrix} -3 & 1 & 1 \\ 1 & -1 & -2 \\ 1 & -2 & -5 \end{pmatrix}$ es definida negativa
3. Las siguientes matrices $A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} -1 & -1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ y $C = \begin{pmatrix} 0 & -2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ no son definidas.

Ejercicio

Clasifica las anteriores matrices simétricas A, B y C del apartado 3 calculando sus autovalores.

Polinomio de Taylor de un campo escalar de dos variables

- El polinomio de Taylor de f centrado en (x_0, y_0) de grado 1 es aquel cuya gráfica coincide con el plano tangente a f en (x_0, y_0), esto es

$$p_1(x, y) = f(x_0, y_0) + df_{(x_0,y_0)} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}.$$

  Se verifica que $\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y)-p_1(x,y)}{\|(x-x_0,y-y_0)\|} = 0$.

- El polinomio de Taylor de f centrado en (x_0, y_0) de grado 2 es

$$\begin{aligned} p_2(x, y) =& f(x_0, y_0) + df_{(x_0,y_0)} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \\ &+ \frac{1}{2!} \begin{pmatrix} x - x_0 & y - y_0 \end{pmatrix} d^2 f_{(x_0,y_0)} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \end{aligned}$$

  El error producido verifica $\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y)-p_2(x,y)}{\|(x-x_0,y-y_0)\|^2} = 0$.

Ejemplo: Para $f(x, y) = \frac{y+1}{x-1}$, los polinomios de Taylor de grado 1 y 2 centrados en $(0, 0)$.

La diferencial es $df = \begin{pmatrix} -\frac{y+1}{(x-1)^2} & \frac{1}{x-1} \end{pmatrix}$, luego $df_{(0,0)} = \begin{pmatrix} -1 & -1 \end{pmatrix}$, por tanto

$$p_1(x, y) = f(0, 0) + df_{(0,0)} \begin{pmatrix} x - 0 \\ y - 0 \end{pmatrix} = -1 + \begin{pmatrix} -1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} =$$
$$= -1 - x - y \text{ (plano tangente)}$$

A partir de df obtenemos la matriz hessiana $d^2 f = \begin{pmatrix} \frac{2(y+1)}{(x-1)^3} & -\frac{1}{(x-1)^2} \\ -\frac{1}{(x-1)^2} & 0 \end{pmatrix}$,

en el origen es $d^2 f_{(0,0)} = \begin{pmatrix} -2 & -1 \\ -1 & 0 \end{pmatrix}$, por tanto

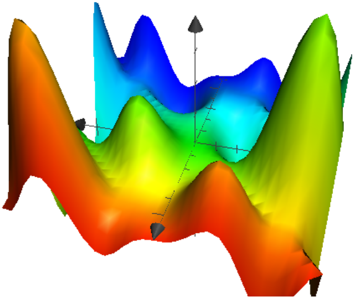
$$p_2(x, y) = -1 + \begin{pmatrix} -1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2!} \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} -2 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} =$$
$$= -1 - x - y - xy - x^2$$

Extremos relativos

Se dice que f alcanza en un punto interior $(x_0, y_0) \in U$ un máximo (resp. mínimo) relativo si existe un entorno de (x_0, y_0) de manera que para todo (x, y) en dicho entorno se verifica

$$f(x, y) \leq f(x_0, y_0) \text{ (resp. } f(x, y) \geq f(x_0, y_0)).$$

Los máximos y mínimos relativos de f se denominan extremos relativos



Condición necesaria de la primera derivada.

Si $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ es un campo escalar de clase C^1 en el abierto U y f alcanza un extremo relativo en $(x_0, y_0) \in U$, entonces $df_{(x_0,y_0)} = (0, 0)$.

Definición

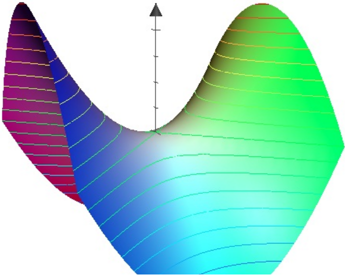
Los puntos $(x_0, y_0) \in U$ que satisfacen $df_{(x_0,y_0)} = (0, 0)$ se denominan puntos críticos de f. Es decir, los puntos críticos son las soluciones del sistema

$$df = (0, 0)$$

Definición

Un punto (x_0, y_0) del interior de $U \subseteq \mathbb{R}^2$ que **no** *es un punto crítico de f se dice que es un punto regular de f. A la imagen de un punto regular, $f(x_0, y_0)$ se le llama valor regular.*

Los puntos críticos de f que no son extremos relativos se denominan puntos de silla.



Condición suficiente de la segunda derivada.

Sea $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$ campo escalar de clase C^3 en un abierto U y $(x_0, y_0) \in U$ un punto crítico de f.

1. Si $d^2 f_{(x_0,y_0)}$ es definida positiva entonces f alcanza un mínimo relativo en (x_0, y_0).
2. Si $d^2 f_{(x_0,y_0)}$ es definida negativa entonces f alcanza un máximo relativo en (x_0, y_0).
3. Si $d^2 f_{(x_0,y_0)}$ es indefinida entonces f tiene un punto de silla en (x_0, y_0).
4. Si $d^2 f_{(x_0,y_0)}$ es semidefinida (positiva o negativa) no podemos decidir.

Nota.- Los conceptos y resultados anteriores pueden extenderse de forma análoga a campos escalares de tres o más variables.

Ejemplo. Calcular y clasificar los puntos críticos del campo escalar $f(x,y) = (x-y)\, e^{xy}$

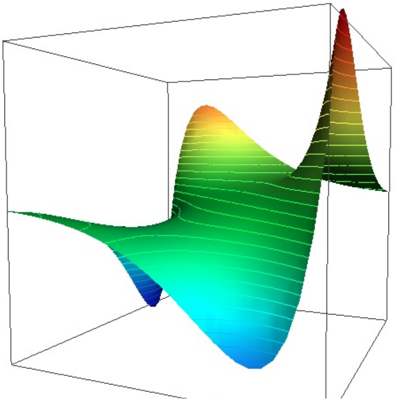
$$df = \left(-y^2 + xy + 1, -x\,y + x^2 - 1\right) e^{xy} = (0,0)$$

$$\text{Puntos críticos: } P_0 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \text{ y } P_1 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

$$d^2 f = \begin{pmatrix} -y\,\left(y^2 - x\,y - 2\right)\, e^{x\,y} & (x-y)\,(x\,y+2)\, e^{x\,y} \\ (x-y)\,(x\,y+2)\, e^{x\,y} & -x\,\left(x\,y - x^2 + 2\right)\, e^{x\,y} \end{pmatrix}$$

- En $P_0 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $H_0 = \begin{pmatrix} \frac{1}{\sqrt{2}\,\sqrt{e}} & -\frac{3}{\sqrt{2}\,\sqrt{e}} \\ -\frac{3}{\sqrt{2}\,\sqrt{e}} & \frac{1}{\sqrt{2}\,\sqrt{e}} \end{pmatrix}$ indefinida. **Punto de silla**.
- En $P_1 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$, $H_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}\,\sqrt{e}} & \frac{3}{\sqrt{2}\,\sqrt{e}} \\ \frac{3}{\sqrt{2}\,\sqrt{e}} & -\frac{1}{\sqrt{2}\,\sqrt{e}} \end{pmatrix}$ indefinida. **Punto de silla**.

En la figura se puede apreciar la gráfica de esta función cerca de los puntos de silla.



Extremos absolutos

Sea el campo escalar $f\colon U \subseteq \mathbb{R}^2 \to \mathbb{R}$. Se dice que f alcanza un máximo (resp. mínimo) absoluto para U en $(x_0, y_0) \in U$ si para todo $(x, y) \in U$ se verifica que $f(x, y) \leq f(x_0, y_0)$ (resp. $f(x, y) \geq f(x_0, y_0)$).
Los máximos y mínimos absolutos de f se denominan, conjuntamente, extremos absolutos de f en U.

Relación entre extremos relativos y absolutos.

Si f alcanza un máximo (mínimo) absoluto para U en un punto (x_0, y_0) interior a U entonces f alcanza un máximo (mínimo) relativo en (x_0, y_0).

Teorema de Weierstrass para campos escalares

Dominios acotados en varias variables.

Un conjunto $U \subseteq \mathbb{R}^2$ se dice acotado si está contenido en algún entorno del cero.

Teorema

Si $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ *es un campo escalar continuo en un conjunto* U *cerrado y acotado entonces existen el máximo absoluto y el mínimo absoluto de* f *en* U*.*

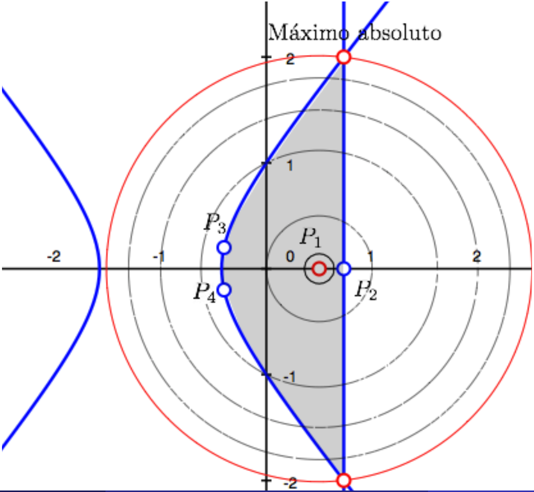
Búsqueda de los extremos absolutos de un campo escalar

Para encontrar los extremos absolutos de f en U se realizan los siguientes pasos:

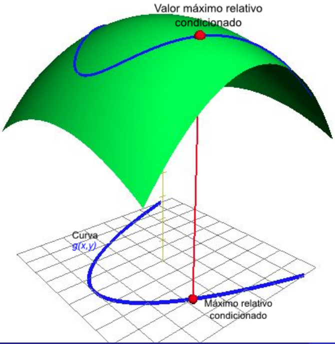
1. Garantizar la existencia de los extremos requeridos.
2. Encontrar los puntos críticos contenidos en el interior de U.
3. Encontrar los puntos candidatos a extremo absoluto contenidos en la frontera de U.
4. Evaluar f en todos los puntos seleccionados en los apartados 2 y 3. En los puntos de mayor evaluación se tiene que f alcanza el máximo absoluto en U, y en los de menor evaluación el mínimo absoluto.

Ejemplo

Comprobemos el teorema de Weierstrass para la función $f(x, y) = x^2 + y^2 - x$ en el conjunto cerrado y acotado que delimita la hipérbola $3x^2 - 2y^2 + 6x = -2$ y la recta $x = \sqrt{3} - 1$.



Sean un c. e. $f(x, y)$ y $g(x, y) = 0$ una curva definida en $U \in \mathbb{R}^2$ abierto. Se dice que f alcanza máximo (resp. mínimo) relativo condicionado en un punto $P = (x_0, y_0)$ de la curva si existe un entorno de P de manera que para todo punto (x, y) en dicho entorno tal que si $g(x, y) = 0$ se verifica que $f(x, y) \leq f(x_0, y_0)$ (resp. $f(x, y) \geq f(x_0, y_0)$). Los máximos y mínimos relativos de f condicionados a g(x,y)=0 se denominan conjuntamente extremos relativos de f condicionados a $g(x, y) = 0$.



Multiplicadores de Lagrange para una ecuación en el plano.

Definimos la función de Lagrange como $L(x,y,\lambda) = f(x,y) + \lambda g(x,y)$. Entonces los extremos relativos condicionados se obtienen entre los puntos que anulan la diferencial $dL_{(x,y,\lambda)} = (0,0,0)$, o lo que es equivalente entre las soluciones del sistema

$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) + \lambda \dfrac{\partial g}{\partial x}(x,y) = 0 \\ \dfrac{\partial f}{\partial y}(x,y) + \lambda \dfrac{\partial g}{\partial y}(x,y) = 0 \\ g(x,y) = 0 \end{cases}$$

Si (x_0, y_0, λ_0) es un punto crítico de la función de Lagrange, el signo de la forma cuadrática

$$d^2L_{(x_0,y_0,\lambda_0)} = \begin{pmatrix} f_{xx}(x_0,y_0) + \lambda_0 g_{xx}(x_0,y_0) & f_{xy}(x_0,y_0) + \lambda_0 g_{xy}(x_0,y_0) & g_x(x_0,y_0) \\ f_{yx}(x_0,y_0) + \lambda_0 g_{yx}(x_0,y_0) & f_{yy}(x_0,y_0) + \lambda_0 g_{yy}(x_0,y_0) & g_y(x_0,y_0) \\ g_x(x_0,y_0) & g_y(x_0,y_0) & 0 \end{pmatrix}$$

determina el tipo de extremo relativo (máximo o mínimo) condicionado.

Extremos relativos condicionados a una superficie en el espacio

Sean $f(x,y,z)$ un campo escalar y $g(x,y,z)=0$ una superficie definida en un abierto U de $\mathbb{R}^3$.
Sea (x_0,y_0,z_0) un punto de la superficie, esto es $g(x_0,y_0,z_0)=0$.
Se dice que f alcanza en (x_0,y_0,z_0) un máximo (resp. mínimo) relativo condicionado a $g(x,y,z)=0$ si existe un entorno de (x_0,y_0,z_0) de manera que para todo punto (x,y,z) perteneciente a dicho entorno tal que $g(x,y,z)=0$ se verifica que $f(x,y,z)\leq f(x_0,y_0,z_0)$ (resp. $f(x,y,z)\geq f(x_0,y_0,z_0)$).
Los máximos y mínimos relativos de f condicionados a $g(x,y,z)=0$ se denominan conjuntamente extremos relativos de f condicionados a la superficie $g(x,y,z)=0$.

Multiplicadores de Lagrange para una ecuación en el espacio.

Definimos la función de Lagrange como

$$L(x, y, z, \lambda) = f(x, y, z) + \lambda g(x, y, z).$$

Entonces los extremos relativos condicionados se obtienen entre los puntos que anulan la diferencial $dL_{(x,y,z,\lambda)} = (0, 0, 0, 0)$, o lo que es equivalente entre las soluciones del sistema

$$\begin{cases} f_x(x, y, z) + \lambda g_x(x, y, z) = 0 \\ f_y(x, y, z) + \lambda g_y(x, y, z) = 0 \\ f_z(x, y, z) + \lambda g_z(x, y, z) = 0 \\ g(x, y, z) = 0 \end{cases}$$

Si $(x_0, y_0, z_0, \lambda_0)$ es un punto crítico de la función de Lagrange, el signo de la forma cuadrática $d^2 L_{(x_0,y_0,z_0,\lambda_0)}$ (de dimensión 4).

Extremo relativo condicionado a una curva en el espacio

Sean un campo escalar $f(x,y,z)$ y $\left\{\begin{array}{l} g_1(x,y,z)=0 \\ g_2(x,y,z)=0 \end{array}\right.$ una curva definidos en un abierto U del espacio $\mathbb{R}^3$.

Sea (x_0,y_0,z_0) un punto de la curva, se dice que f alcanza en (x_0,y_0,z_0) un máximo (resp. mínimo) relativo condicionado a $\left\{\begin{array}{l} g_1(x,y,z)=0 \\ g_2(x,y,z)=0 \end{array}\right.$ si existe un entorno de (x_0,y_0,z_0) de manera que para todo punto (x,y,z) perteneciente a dicho entorno tal que $\left\{\begin{array}{l} g_1(x_0,y_0,z_0)=0 \\ g_2(x_0,y_0,z_0)=0 \end{array}\right.$ se verifica la desigualdad $f(x,y,z)\leq f(x_0,y_0,z_0)$ (resp. $f(x,y,z)\geq f(x_0,y_0,z_0)$).

Los máximos y mínimos relativos de f condicionados se denominan conjuntamente extremos relativos de f condicionados .

Multiplicadores de Lagrange para dos ecuaciones en el espacio.

Definimos la función de Lagrange como

$$L(x,y,z,\lambda,\mu) = f(x,y,z) + \lambda g_1(x,y,z) + \mu g_2(x,y,z).$$

Entonces los extremos relativos condicionados se obtienen entre los puntos que anulan la diferencial $dL_{(x,y,z,\lambda,\mu)} = (0,0,0,0,0)$, o lo que es equivalente entre las soluciones del sistema

$$\left\{ \begin{array}{l} f_x(x,y,z) + \lambda g_{1x}(x,y,z) + \mu g_{2x}(x,y,z) = 0 \\ f_y(x,y,z) + \lambda g_{1y}(x,y,z) + \mu g_{2y}(x,y,z) = 0 \\ f_z(x,y,z) + \lambda g_{1z}(x,y,z) + \mu g_{2z}(x,y,z) = 0 \\ g_1(x,y,z) = 0 \\ g_2(x,y,z) = 0 \end{array} \right.$$

Si $(x_0, y_0, z_0, \lambda_0, \mu_0)$ es un punto crítico de la función de Lagrange, el signo de la forma cuadrática $d^2L_{(x_0,y_0,z_0,\lambda_0,\mu_0)}$ (de dimensión 5).

OCW UMA

Rodríguez Sánchez, F.J.

Muñoz Ruiz, M.L.

Merino Córdoba, S.

2014.