

Funciones y Optimización en Varias Variables

Viquez

Universidad de Costa Rica

viquezejin@gmail.com

Verano 2022

Outline

- 1 Preliminares Matemáticos
 - Topología en $\mathbb{R}^n$
 - Teorema de la Función Implícita/Inversa
 - Matrices Semidefinidas y Definidas
 - Convexidad-Concavidad
 - Cuasiconvexidad-Cuasiconcavidad

Outline

- 1 Preliminares Matemáticos
 - Topología en $\mathbb{R}^n$
 - Teorema de la Función Implícita/Inversa
 - Matrices Semidefinidas y Definidas
 - Convexidad-Concavidad
 - Cuasiconvexidad-Cuasiconcavidad
- 2 Optimización
 - Motivación
 - Definiciones
 - Optimización Sin Restricciones

Outline

1 Preliminares Matemáticos

- Topología en $\mathbb{R}^n$
- Teorema de la Función Implícita/Inversa
- Matrices Semidefinidas y Definidas
- Convexidad-Concavidad
- Cuasiconvexidad-Cuasiconcavidad

2 Optimización

- Motivación
- Definiciones
- Optimización Sin Restricciones

3 Optimización Con Restricciones

- Restricciones de Igualdad: Lagrange
- Restricciones de Igualdad/Desigualdad: Karush-Kuhn-Tucker

Outline

1 Preliminares Matemáticos

- Topología en $\mathbb{R}^n$
- Teorema de la Función Implícita/Inversa
- Matrices Semidefinidas y Definidas
- Convexidad-Concavidad
- Cuasiconvexidad-Cuasiconcavidad

2 Optimización

- Motivación
- Definiciones
- Optimización Sin Restricciones

3 Optimización Con Restricciones

- Restricciones de Igualdad: Lagrange
- Restricciones de Igualdad/Desigualdad: Karush-Kuhn-Tucker

Puntos Interiores y Puntos de Frontera

Clasificación de puntos

- Se define como “bola con centro $\mathbf{a}$ ” al conjunto

$$B_r(\mathbf{a}) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{a}\| < r\}.$$

- Un punto $\mathbf{a}$ se dice un punto interior de un conjunto $S \subset \mathbb{R}^n$ si existe una bola con centro en $\mathbf{a}$ completamente contenido en S .
- Un punto $\mathbf{a}$ se dice un punto frontera de un conjunto $S \subset \mathbb{R}^n$ si toda bola centrada en $\mathbf{a}$ interseca puntos no pertenecientes a S .

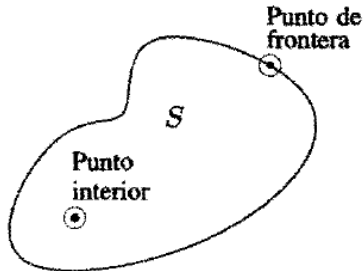
Puntos Interiores y Puntos de Frontera

Clasificación de puntos

- Se define como “bola con centro a ” al conjunto

$$B_r(a) = \{x \in \mathbb{R}^n \mid \|x - a\| < r\}.$$

- Un punto a se dice un punto interior de un conjunto $S \subset \mathbb{R}^n$ si existe una bola con centro en a completamente contenido en S .
- Un punto a se dice un punto frontera de un conjunto $S \subset \mathbb{R}^n$ si toda bola centrada en a interseca puntos no pertenecientes a S .



Abiertos y Cerrados

Clasificación de conjuntos

- Un conjunto $O \subset \mathbb{R}^n$ se dice abierto si está compuesto únicamente de puntos interiores.
- Un conjunto $C \subset \mathbb{R}^n$ se dice cerrado si todo punto frontera de S es elemento de S .

Abiertos y Cerrados

Clasificación de conjuntos

- Un conjunto $O \subset \mathbb{R}^n$ se dice abierto si está compuesto únicamente de puntos interiores.
- Un conjunto $C \subset \mathbb{R}^n$ se dice cerrado si todo punto frontera de S es elemento de S .

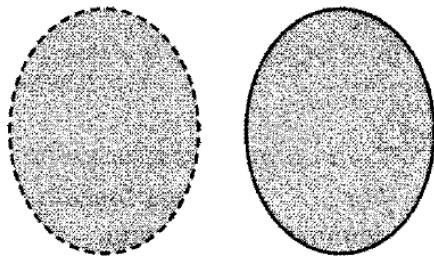


Figura: Ejemplos de conjunto abierto y cerrado

Conjuntos Compactos

Clasificación de conjuntos

- Un conjunto $A \subset \mathbb{R}^n$ se dice acotado si existe una bola que lo contenga.
- Un conjunto $C \subset \mathbb{R}^n$ se dice compacto si es cerrado y acotado.

Conjuntos Compactos

Clasificación de conjuntos

- Un conjunto $A \subset \mathbb{R}^n$ se dice acotado si existe una bola que lo contenga.
- Un conjunto $C \subset \mathbb{R}^n$ se dice compacto si es cerrado y acotado.

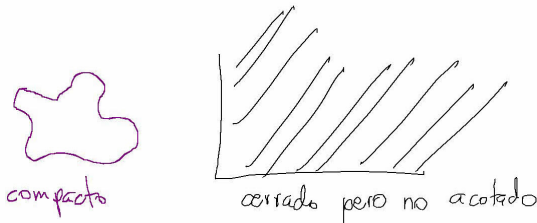


Figura: Ejemplos de conjunto compacto y no compacto

Matriz Jacobiana

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ es diferenciable, su derivada es $f' : \mathbb{R} \rightarrow \mathbb{R}$.

Matriz Jacobiana

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ es diferenciable, su derivada es $f' : \mathbb{R} \rightarrow \mathbb{R}$.

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable, su gradiente es $\nabla f = \begin{pmatrix} \partial_{x_1} f \\ \vdots \\ \partial_{x_n} f \end{pmatrix}$.

Matriz Jacobiana

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ es diferenciable, su derivada es $f' : \mathbb{R} \rightarrow \mathbb{R}$.

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable, su gradiente es $\nabla f = \begin{pmatrix} \partial_{x_1} f \\ \vdots \\ \partial_{x_n} f \end{pmatrix}$.

Jacobiano

Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función vectorial en varias variables diferenciable dada por $F(\mathbf{x}) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$. El *Jacobiano* (o matriz Jacobiana) de f , también denotado $\frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)}$, es

$$J_F(\mathbf{x}) := \begin{pmatrix} (\nabla f_1(\mathbf{x}))^T \\ \vdots \\ (\nabla f_m(\mathbf{x}))^T \end{pmatrix} = \begin{pmatrix} \partial_{x_1} f_1(\mathbf{x}) & \cdots & \partial_{x_n} f_1(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m(\mathbf{x}) & \cdots & \partial_{x_n} f_m(\mathbf{x}) \end{pmatrix}$$

Matriz Jacobiana

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ es diferenciable, su derivada es $f' : \mathbb{R} \rightarrow \mathbb{R}$.

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable, su gradiente es $\nabla f = \begin{pmatrix} \partial_{x_1} f \\ \vdots \\ \partial_{x_n} f \end{pmatrix}$.

Jacobiano

Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una función vectorial en varias variables diferenciable dada por $F(\mathbf{x}) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$. El *Jacobiano* (o matriz Jacobiana) de f , también denotado $\frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)}$, es

$$J_F(\mathbf{x}) := \begin{pmatrix} (\nabla f_1(\mathbf{x}))^T \\ \vdots \\ (\nabla f_m(\mathbf{x}))^T \end{pmatrix} = \begin{pmatrix} \partial_{x_1} f_1(\mathbf{x}) & \cdots & \partial_{x_n} f_1(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m(\mathbf{x}) & \cdots & \partial_{x_n} f_m(\mathbf{x}) \end{pmatrix}$$

Así como la derivada en la dirección $\mathbf{u}$ está dada por $\nabla f \cdot \mathbf{u}$, el vector derivada direccional en la dirección $\mathbf{u}$ está dada por la multiplicación matricial $J_F \cdot \mathbf{u}$. Note que si $m = 1$ entonces $J_F = (\nabla F)^T$.

Regla de la Cadena

Regla de la Cadena Multidimensional

Considere dos funciones vectoriales $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $G : \mathbb{R}^m \rightarrow \mathbb{R}^k$ diferenciables tales que la función vectorial $H : \mathbb{R}^n \rightarrow \mathbb{R}^k$ dada por $H(\mathbf{x}) = G \circ F(\mathbf{x}) := G(F(\mathbf{x}))$ está bien definida. Entonces H es diferenciable y $J_H(\mathbf{x}) = J_{G \circ F}(\mathbf{x}) = J_G(F(\mathbf{x}))J_F(\mathbf{x})$.

Regla de la Cadena

Regla de la Cadena Multidimensional

Considere dos funciones vectoriales $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $G : \mathbb{R}^m \rightarrow \mathbb{R}^k$ diferenciables tales que la función vectorial $H : \mathbb{R}^n \rightarrow \mathbb{R}^k$ dada por $H(\mathbf{x}) = G \circ F(\mathbf{x}) := G(F(\mathbf{x}))$ está bien definida. Entonces H es diferenciable y $J_H(\mathbf{x}) = J_{G \circ F}(\mathbf{x}) = J_G(F(\mathbf{x}))J_F(\mathbf{x})$.

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ se tiene que $J_f = (\nabla f)^T = f'$, entonces si $n = m = k = 1$ se recupera la regla de la cadena $h'(x) = (g \circ f(x))' = g'(f(x)) \cdot f'(x)$.

Regla de la Cadena

Regla de la Cadena Multidimensional

Considere dos funciones vectoriales $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $G : \mathbb{R}^m \rightarrow \mathbb{R}^k$ diferenciables tales que la función vectorial $H : \mathbb{R}^n \rightarrow \mathbb{R}^k$ dada por $H(\mathbf{x}) = G \circ F(\mathbf{x}) := G(F(\mathbf{x}))$ está bien definida. Entonces H es diferenciable y $J_H(\mathbf{x}) = J_{G \circ F}(\mathbf{x}) = J_G(F(\mathbf{x}))J_F(\mathbf{x})$.

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ se tiene que $J_f = (\nabla f)^T = f'$, entonces si $n = m = k = 1$ se recupera la regla de la cadena $h'(x) = (g \circ f(x))' = g'(f(x)) \cdot f'(x)$.

Si $n, m > 1$ y $k = 1$, entonces $(\nabla h(\mathbf{x}))^T = \nabla g(F(\mathbf{x}))^T \cdot J_F(\mathbf{x})$.

Regla de la Cadena

Regla de la Cadena Multidimensional

Considere dos funciones vectoriales $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $G : \mathbb{R}^m \rightarrow \mathbb{R}^k$ diferenciables tales que la función vectorial $H : \mathbb{R}^n \rightarrow \mathbb{R}^k$ dada por $H(\mathbf{x}) = G \circ F(\mathbf{x}) := G(F(\mathbf{x}))$ está bien definida. Entonces H es diferenciable y $J_H(\mathbf{x}) = J_{G \circ F}(\mathbf{x}) = J_G(F(\mathbf{x}))J_F(\mathbf{x})$.

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ se tiene que $J_f = (\nabla f)^T = f'$, entonces si $n = m = k = 1$ se recupera la regla de la cadena $h'(x) = (g \circ f(x))' = g'(f(x)) \cdot f'(x)$.

Si $n, m > 1$ y $k = 1$, entonces $(\nabla h(\mathbf{x}))^T = \nabla g(F(\mathbf{x}))^T \cdot J_F(\mathbf{x})$.

Si $n > 1$ y $m, k = 1$, entonces $(\nabla h(\mathbf{x}))^T = g'(f(\mathbf{x})) \cdot (\nabla f(\mathbf{x}))^T$.

Regla de la Cadena

Regla de la Cadena Multidimensional

Considere dos funciones vectoriales $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $G : \mathbb{R}^m \rightarrow \mathbb{R}^k$ diferenciables tales que la función vectorial $H : \mathbb{R}^n \rightarrow \mathbb{R}^k$ dada por $H(\mathbf{x}) = G \circ F(\mathbf{x}) := G(F(\mathbf{x}))$ está bien definida. Entonces H es diferenciable y $J_H(\mathbf{x}) = J_{G \circ F}(\mathbf{x}) = J_G(F(\mathbf{x}))J_F(\mathbf{x})$.

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ se tiene que $J_f = (\nabla f)^T = f'$, entonces si $n = m = k = 1$ se recupera la regla de la cadena $h'(x) = (g \circ f(x))' = g'(f(x)) \cdot f'(x)$.

Si $n, m > 1$ y $k = 1$, entonces $(\nabla h(\mathbf{x}))^T = \nabla g(F(\mathbf{x}))^T \cdot J_F(\mathbf{x})$.

Si $n > 1$ y $m, k = 1$, entonces $(\nabla h(\mathbf{x}))^T = g'(f(\mathbf{x})) \cdot (\nabla f(\mathbf{x}))^T$.

En general, si se tienen N funciones vectoriales $F_i : \mathbb{R}^{n_{i-1}} \rightarrow \mathbb{R}^{n_i}$, $i = 1, \dots, N$, tal que $H = F_N \circ F_{N-1} \circ \dots \circ F_2 \circ F_1$ está bien definida, entonces

$$J_H(\mathbf{x}) = J_{F_N}(F_{N-1} \circ \dots \circ F_2 \circ F_1(\mathbf{x})) \cdots J_{F_2}(F_1(\mathbf{x}))J_{F_1}(\mathbf{x}).$$

Ejemplo 1: $n_0 = 2, n_1 = 3, n_2 = 2, n_3 = 1$

$F_1(\theta, r) = (r \cos(\theta), r \sin(\theta), \tan(\theta))$, $F_2(y_1, y_2, y_3) = (e^{y_3}, \ln(y_1 y_2))$,
 $F_3(z_1, z_2) = z_1 + z_2$ y $H(\theta, r) = e^{\tan(\theta)} + \ln(r^2 \cos(\theta) \sin(\theta))$. Verifique
que $H = F_3 \circ F_2 \circ F_1$ y calcule J_H por definición y con regla de la cadena.

Ejemplo 1: $n_0 = 2, n_1 = 3, n_2 = 2, n_3 = 1$

$F_1(\theta, r) = (r \cos(\theta), r \sin(\theta), \tan(\theta))$, $F_2(y_1, y_2, y_3) = (e^{y_3}, \ln(y_1 y_2))$,
 $F_3(z_1, z_2) = z_1 + z_2$ y $H(\theta, r) = e^{\tan(\theta)} + \ln(r^2 \cos(\theta) \sin(\theta))$. Verifique
que $H = F_3 \circ F_2 \circ F_1$ y calcule J_H por definición y con regla de la cadena.

Solución: Sustituyendo $y_1 = r \cos(\theta)$, $y_2 = r \sin(\theta)$ y $y_3 = \tan(\theta)$ se
obtiene $F_2 \circ F_1(\theta, r) = (e^{\tan(\theta)}, \ln(r^2 \cos(\theta) \sin(\theta))) \Rightarrow H = F_3 \circ F_2 \circ F_1$.

Ejemplo 1: $n_0 = 2, n_1 = 3, n_2 = 2, n_3 = 1$

$F_1(\theta, r) = (r \cos(\theta), r \sin(\theta), \tan(\theta))$, $F_2(y_1, y_2, y_3) = (e^{y_3}, \ln(y_1 y_2))$,
 $F_3(z_1, z_2) = z_1 + z_2$ y $H(\theta, r) = e^{\tan(\theta)} + \ln(r^2 \cos(\theta) \sin(\theta))$. Verifique
 que $H = F_3 \circ F_2 \circ F_1$ y calcule J_H por definición y con regla de la cadena.

Solución: Sustituyendo $y_1 = r \cos(\theta)$, $y_2 = r \sin(\theta)$ y $y_3 = \tan(\theta)$ se
 obtiene $F_2 \circ F_1(\theta, r) = (e^{\tan(\theta)}, \ln(r^2 \cos(\theta) \sin(\theta))) \Rightarrow H = F_3 \circ F_2 \circ F_1$.

Derivando directamente se tiene:

$$J_H(\theta, r) = (\nabla H)^T = \left(\frac{\partial H}{\partial \theta}, \frac{\partial H}{\partial r} \right) = \left(e^{\tan(\theta)} \sec^2(\theta) + \frac{\cos^2(\theta) - \sin^2(\theta)}{\cos(\theta) \sin(\theta)}, \frac{2}{r} \right)$$

Ejemplo 1: $n_0 = 2, n_1 = 3, n_2 = 2, n_3 = 1$

$F_1(\theta, r) = (r \cos(\theta), r \sin(\theta), \tan(\theta))$, $F_2(y_1, y_2, y_3) = (e^{y_3}, \ln(y_1 y_2))$,
 $F_3(z_1, z_2) = z_1 + z_2$ y $H(\theta, r) = e^{\tan(\theta)} + \ln(r^2 \cos(\theta) \sin(\theta))$. Verifique
 que $H = F_3 \circ F_2 \circ F_1$ y calcule J_H por definición y con regla de la cadena.

Solución: Sustituyendo $y_1 = r \cos(\theta)$, $y_2 = r \sin(\theta)$ y $y_3 = \tan(\theta)$ se
 obtiene $F_2 \circ F_1(\theta, r) = (e^{\tan(\theta)}, \ln(r^2 \cos(\theta) \sin(\theta))) \Rightarrow H = F_3 \circ F_2 \circ F_1$.

Derivando directamente se tiene:

$$J_H(\theta, r) = (\nabla H)^T = \left(\frac{\partial H}{\partial \theta}, \frac{\partial H}{\partial r} \right) = \left(e^{\tan(\theta)} \sec^2(\theta) + \frac{\cos^2(\theta) - \sin^2(\theta)}{\cos(\theta) \sin(\theta)}, \frac{2}{r} \right)$$

Por regla de la cadena tenemos: $J_{F_3}(F_2 \circ F_1) = (1, 1)$;

$$J_{F_1} = \begin{pmatrix} -r \sin(\theta) & \cos(\theta) \\ r \cos(\theta) & \sin(\theta) \\ \sec^2(\theta) & 0 \end{pmatrix}; \quad J_{F_2}(F_1) = \begin{pmatrix} 0 & 0 & e^{\tan(\theta)} \\ \frac{1}{r \cos(\theta)} & \frac{1}{r \sin(\theta)} & 0 \end{pmatrix}.$$

$$\Rightarrow J_H = J_{F_3}(F_2 \circ F_1) J_{F_2}(F_1) J_{F_1} = \left(e^{\tan(\theta)} \sec^2(\theta) + \frac{\cos^2(\theta) - \sin^2(\theta)}{\cos(\theta) \sin(\theta)}, \frac{2}{r} \right).$$

Ejemplo 2: $n_0 = 1, n_1 = 3, n_2 = 1$

Si $H(t) = F_2 \circ F_1(t) = e^{t^3} \cos(4t) \sin(3t^2)$ con $F_1(t) = (2t, t^2, t^3)$ y $F_2(x, y, z) = e^z \cos(2x) \sin(3y)$, calcule $J_H(t)$ por ambas vías.

Ejemplo 2: $n_0 = 1, n_1 = 3, n_2 = 1$

Si $H(t) = F_2 \circ F_1(t) = e^{t^3} \cos(4t) \sin(3t^2)$ con $F_1(t) = (2t, t^2, t^3)$ y $F_2(x, y, z) = e^z \cos(2x) \sin(3y)$, calcule $J_H(t)$ por ambas vías.

Solución: Derivando directamente se tiene: $(J_H(t) = H'(t))$

$$H'(t) = -4e^{t^3} \sin(4t) \sin(3t^2) + 6te^{t^3} \cos(4t) \cos(3t^2) + 3t^2 e^{t^3} \cos(4t) \sin(3t^2).$$

Ejemplo 2: $n_0 = 1, n_1 = 3, n_2 = 1$

Si $H(t) = F_2 \circ F_1(t) = e^{t^3} \cos(4t) \sin(3t^2)$ con $F_1(t) = (2t, t^2, t^3)$ y $F_2(x, y, z) = e^z \cos(2x) \sin(3y)$, calcule $J_H(t)$ por ambas vías.

Solución: Derivando directamente se tiene: $(J_H(t) = H'(t))$

$$H'(t) = -4e^{t^3} \sin(4t) \sin(3t^2) + 6te^{t^3} \cos(4t) \cos(3t^2) + 3t^2 e^{t^3} \cos(4t) \sin(3t^2).$$

Por regla de la cadena tenemos: $J_{F_1}(t) = \begin{pmatrix} 2 \\ 2t \\ 3t^2 \end{pmatrix};$

$$J_{F_2}(x, y, z) = \begin{pmatrix} -2e^z \sin(2x) \sin(3y) & 3e^z \cos(2x) \cos(3y) & e^z \cos(2x) \sin(3y) \end{pmatrix}$$

y

$$J_{F_2}(F_1(t)) = \begin{pmatrix} -2e^{t^3} \sin(4t) \sin(3t^2) & 3e^{t^3} \cos(4t) \cos(3t^2) & e^{t^3} \cos(4t) \sin(3t^2) \end{pmatrix}$$

Ejemplo 2: $n_0 = 1, n_1 = 3, n_2 = 1$

Si $H(t) = F_2 \circ F_1(t) = e^{t^3} \cos(4t) \sin(3t^2)$ con $F_1(t) = (2t, t^2, t^3)$ y $F_2(x, y, z) = e^z \cos(2x) \sin(3y)$, calcule $J_H(t)$ por ambas vías.

Solución: Derivando directamente se tiene: $(J_H(t) = H'(t))$

$$H'(t) = -4e^{t^3} \sin(4t) \sin(3t^2) + 6te^{t^3} \cos(4t) \cos(3t^2) + 3t^2 e^{t^3} \cos(4t) \sin(3t^2).$$

Por regla de la cadena tenemos: $J_{F_1}(t) = \begin{pmatrix} 2 \\ 2t \\ 3t^2 \end{pmatrix};$

$$J_{F_2}(x, y, z) = \begin{pmatrix} -2e^z \sin(2x) \sin(3y) & 3e^z \cos(2x) \cos(3y) & e^z \cos(2x) \sin(3y) \end{pmatrix}$$

y

$$J_{F_2}(F_1(t)) = \begin{pmatrix} -2e^{t^3} \sin(4t) \sin(3t^2) & 3e^{t^3} \cos(4t) \cos(3t^2) & e^{t^3} \cos(4t) \sin(3t^2) \end{pmatrix}$$

$$\Rightarrow J_H(t) = H'(t) = J_{F_2}(F_1(t)) J_{F_1}(t)$$

$$= -4e^{t^3} \sin(4t) \sin(3t^2) + 6te^{t^3} \cos(4t) \cos(3t^2) + 3t^2 e^{t^3} \cos(4t) \sin(3t^2).$$

Ejemplo 3: $n_0 = 2, n_1 = 2, n_2 = 1$

Si $H(x, y) = F_2 \circ F_1(x, y) = (2x + y)^2(x - 2y) + (2x + y)(x - 2y)^2$ con $F_1(x, y) = (2x + y, x - 2y)$ y $F_2(u, v) = u^2v + uv^2$, calcule $J_H(x, y)$.

Ejemplo 3: $n_0 = 2, n_1 = 2, n_2 = 1$

Si $H(x, y) = F_2 \circ F_1(x, y) = (2x + y)^2(x - 2y) + (2x + y)(x - 2y)^2$ con $F_1(x, y) = (2x + y, x - 2y)$ y $F_2(u, v) = u^2v + uv^2$, calcule $J_H(x, y)$.

Solución: Derivando $(J_H(x, y) = (\nabla H(x, y))^T = (\partial_x H(x, y), \partial_y H(x, y))$

$$\nabla H = \begin{pmatrix} 4(2x + y)(x - 2y) + (2x + y)^2 + 2(x - 2y)^2 + 2(2x + y)(x - 2y) \\ 2(2x + y)(x - 2y) - 2(2x + y)^2 + (x - 2y)^2 - 4(2x + y)(x - 2y) \end{pmatrix}.$$

Ejemplo 3: $n_0 = 2, n_1 = 2, n_2 = 1$

Si $H(x, y) = F_2 \circ F_1(x, y) = (2x + y)^2(x - 2y) + (2x + y)(x - 2y)^2$ con $F_1(x, y) = (2x + y, x - 2y)$ y $F_2(u, v) = u^2v + uv^2$, calcule $J_H(x, y)$.

Solución: Derivando $(J_H(x, y) = (\nabla H(x, y))^T = (\partial_x H(x, y), \partial_y H(x, y)))$

$$\nabla H = \begin{pmatrix} 4(2x + y)(x - 2y) + (2x + y)^2 + 2(x - 2y)^2 + 2(2x + y)(x - 2y) \\ 2(2x + y)(x - 2y) - 2(2x + y)^2 + (x - 2y)^2 - 4(2x + y)(x - 2y) \end{pmatrix}.$$

R. de la Cadena: $J_{F_1}(x, y) = \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}$; $J_{F_2}(u, v) = (2uv + v^2, u^2 + 2uv)$.

$$J_{F_2}(F_1) = (2(2x + y)(x - 2y) + (x - 2y)^2 \quad (2x + y)^2 + 2(2x + y)(x - 2y)).$$

Ejemplo 3: $n_0 = 2, n_1 = 2, n_2 = 1$

Si $H(x, y) = F_2 \circ F_1(x, y) = (2x + y)^2(x - 2y) + (2x + y)(x - 2y)^2$ con $F_1(x, y) = (2x + y, x - 2y)$ y $F_2(u, v) = u^2v + uv^2$, calcule $J_H(x, y)$.

Solución: Derivando $(J_H(x, y) = (\nabla H(x, y))^T = (\partial_x H(x, y), \partial_y H(x, y)))$

$$\nabla H = \begin{pmatrix} 4(2x + y)(x - 2y) + (2x + y)^2 + 2(x - 2y)^2 + 2(2x + y)(x - 2y) \\ 2(2x + y)(x - 2y) - 2(2x + y)^2 + (x - 2y)^2 - 4(2x + y)(x - 2y) \end{pmatrix}.$$

$$\text{R. de la Cadena: } J_{F_1}(x, y) = \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}; J_{F_2}(u, v) = (2uv + v^2, u^2 + 2uv).$$

$$J_{F_2}(F_1) = (2(2x + y)(x - 2y) + (x - 2y)^2 \quad (2x + y)^2 + 2(2x + y)(x - 2y)).$$

$$\Rightarrow J_H(x, y) = (\nabla H(x, y))^T = J_{F_2}(F_1(x, y))J_{F_1}(x, y)$$

$$= \begin{pmatrix} 4(2x + y)(x - 2y) + (2x + y)^2 + 2(x - 2y)^2 + 2(2x + y)(x - 2y) \\ 2(2x + y)(x - 2y) - 2(2x + y)^2 + (x - 2y)^2 - 4(2x + y)(x - 2y) \end{pmatrix}^T.$$

Ejemplo 4: $n_0 = 1, n_1 = 2, n_2 = 2$

Si $H(x) = F_2 \circ F_1(x) = (\cos^2(x) \sin(x), \cos(x) \sin^2(x))$ con $F_1(x) = (\cos(x), \sin(x))$ y $F_2(y, z) = (y^2 z, y z^2)$, calcule $J_H(x)$.

Ejemplo 4: $n_0 = 1, n_1 = 2, n_2 = 2$

Si $H(x) = F_2 \circ F_1(x) = (\cos^2(x) \sin(x), \cos(x) \sin^2(x))$ con $F_1(x) = (\cos(x), \sin(x))$ y $F_2(y, z) = (y^2 z, y z^2)$, calcule $J_H(x)$.

Solución: Derivando directamente se obtiene

$$J_H(x) = \begin{pmatrix} H'_1(x) \\ H'_2(x) \end{pmatrix} = \begin{pmatrix} -2 \cos(x) \sin^2(x) + \cos^3(x) \\ -\sin^3(x) + 2 \cos^2(x) \sin(x) \end{pmatrix}.$$

Ejemplo 4: $n_0 = 1, n_1 = 2, n_2 = 2$

Si $H(x) = F_2 \circ F_1(x) = (\cos^2(x) \sin(x), \cos(x) \sin^2(x))$ con $F_1(x) = (\cos(x), \sin(x))$ y $F_2(y, z) = (y^2 z, y z^2)$, calcule $J_H(x)$.

Solución: Derivando directamente se obtiene

$$J_H(x) = \begin{pmatrix} H'_1(x) \\ H'_2(x) \end{pmatrix} = \begin{pmatrix} -2 \cos(x) \sin^2(x) + \cos^3(x) \\ -\sin^3(x) + 2 \cos^2(x) \sin(x) \end{pmatrix}.$$

Regla de la Cadena: $J_{F_1}(x) = \begin{pmatrix} -\sin(x) \\ \cos(x) \end{pmatrix}$ y $J_{F_2}(y, z) = \begin{pmatrix} 2yz & y^2 \\ z^2 & 2yz \end{pmatrix}$.

$$J_{F_2}(F_1(x)) = \begin{pmatrix} 2 \cos(x) \sin(x) & \cos^2(x) \\ \sin^2(x) & 2 \cos(x) \sin(x) \end{pmatrix}.$$

Ejemplo 4: $n_0 = 1, n_1 = 2, n_2 = 2$

Si $H(x) = F_2 \circ F_1(x) = (\cos^2(x) \sin(x), \cos(x) \sin^2(x))$ con $F_1(x) = (\cos(x), \sin(x))$ y $F_2(y, z) = (y^2 z, y z^2)$, calcule $J_H(x)$.

Solución: Derivando directamente se obtiene

$$J_H(x) = \begin{pmatrix} H'_1(x) \\ H'_2(x) \end{pmatrix} = \begin{pmatrix} -2 \cos(x) \sin^2(x) + \cos^3(x) \\ -\sin^3(x) + 2 \cos^2(x) \sin(x) \end{pmatrix}.$$

Regla de la Cadena: $J_{F_1}(x) = \begin{pmatrix} -\sin(x) \\ \cos(x) \end{pmatrix}$ y $J_{F_2}(y, z) = \begin{pmatrix} 2yz & y^2 \\ z^2 & 2yz \end{pmatrix}$.

$$J_{F_2}(F_1(x)) = \begin{pmatrix} 2 \cos(x) \sin(x) & \cos^2(x) \\ \sin^2(x) & 2 \cos(x) \sin(x) \end{pmatrix}.$$

$$\begin{aligned} \Rightarrow J_H(x) &= J_{F_2}(F_1(x)) J_{F_1}(x) = \begin{pmatrix} 2 \cos(x) \sin(x) & \cos^2(x) \\ \sin^2(x) & 2 \cos(x) \sin(x) \end{pmatrix} \begin{pmatrix} -\sin(x) \\ \cos(x) \end{pmatrix} \\ &= \begin{pmatrix} -2 \cos(x) \sin^2(x) + \cos^3(x) \\ -\sin^3(x) + 2 \cos^2(x) \sin(x) \end{pmatrix}. \end{aligned}$$

Ejemplo 5: $n_0 = 2, n_1 = 1, n_2 = 1$

Si $H(x, y) = F_2 \circ F_1(x, y) = (\sin(x) \cos(y))^3 + 2(\sin(x) \cos(y))^2$ con $F_1(x, y) = \sin(x) \cos(y)$ y $F_2(z) = z^3 + 2z^2$, calcule $J_H(x, y)$.

Ejemplo 5: $n_0 = 2, n_1 = 1, n_2 = 1$

Si $H(x, y) = F_2 \circ F_1(x, y) = (\sin(x) \cos(y))^3 + 2(\sin(x) \cos(y))^2$ con $F_1(x, y) = \sin(x) \cos(y)$ y $F_2(z) = z^3 + 2z^2$, calcule $J_H(x, y)$.

Solución: Derivando $(J_H(x, y) = (\nabla H(x, y))^T = (\partial_x H(x, y), \partial_y H(x, y))$

$$\nabla H(x, y) = \begin{pmatrix} 3 \sin^2(x) \cos(x) \cos^3(y) + 4 \sin(x) \cos(x) \cos^2(y) \\ -3 \sin^3(x) \cos(y) \sin(y) - 4 \sin^2(x) \cos(y) \sin(y) \end{pmatrix}.$$

Ejemplo 5: $n_0 = 2, n_1 = 1, n_2 = 1$

Si $H(x, y) = F_2 \circ F_1(x, y) = (\sin(x) \cos(y))^3 + 2(\sin(x) \cos(y))^2$ con $F_1(x, y) = \sin(x) \cos(y)$ y $F_2(z) = z^3 + 2z^2$, calcule $J_H(x, y)$.

Solución: Derivando $(J_H(x, y) = (\nabla H(x, y))^T = (\partial_x H(x, y), \partial_y H(x, y)))$

$$\nabla H(x, y) = \begin{pmatrix} 3 \sin^2(x) \cos(x) \cos^3(y) + 4 \sin(x) \cos(x) \cos^2(y) \\ -3 \sin^3(x) \cos(y) \sin(y) - 4 \sin^2(x) \cos(y) \sin(y) \end{pmatrix}.$$

R. Cadena: $J_{F_1} = (\cos(x) \cos(y), -\sin(x) \sin(y))$; $J_{F_2}(z) = 3z^2 + 4z$.

$$J_{F_2}(F_1(x, y)) = F_2'(F_1(x, y)) = 3(\sin(x) \cos(y))^2 + 4 \sin(x) \cos(y).$$

Ejemplo 5: $n_0 = 2, n_1 = 1, n_2 = 1$

Si $H(x, y) = F_2 \circ F_1(x, y) = (\sin(x) \cos(y))^3 + 2(\sin(x) \cos(y))^2$ con $F_1(x, y) = \sin(x) \cos(y)$ y $F_2(z) = z^3 + 2z^2$, calcule $J_H(x, y)$.

Solución: Derivando $(J_H(x, y) = (\nabla H(x, y))^T = (\partial_x H(x, y), \partial_y H(x, y)))$

$$\nabla H(x, y) = \begin{pmatrix} 3 \sin^2(x) \cos(x) \cos^3(y) + 4 \sin(x) \cos(x) \cos^2(y) \\ -3 \sin^3(x) \cos(y) \sin(y) - 4 \sin^2(x) \cos(y) \sin(y) \end{pmatrix}.$$

R. Cadena: $J_{F_1} = (\cos(x) \cos(y), -\sin(x) \sin(y))$; $J_{F_2}(z) = 3z^2 + 4z$.

$$J_{F_2}(F_1(x, y)) = F_2'(F_1(x, y)) = 3(\sin(x) \cos(y))^2 + 4 \sin(x) \cos(y).$$

$$\begin{aligned} \Rightarrow J_H(x, y) &= (\nabla H(x, y))^T = J_{F_2}(F_1(x, y)) J_{F_1}(x, y) \\ &= \begin{pmatrix} 3 \sin^2(x) \cos(x) \cos^3(y) + 4 \sin(x) \cos(x) \cos^2(y) \\ -3 \sin^3(x) \cos(y) \sin(y) - 4 \sin^2(x) \cos(y) \sin(y) \end{pmatrix}^T. \end{aligned}$$

Teorema de la Función Implícita: Motivación

Si se tiene una función vectorial $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ tal que $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, con $\mathbf{x} \in \mathbb{R}^n$ y $\mathbf{y} \in \mathbb{R}^m$,

¿Es posible definir una función $\mathbf{y} = f(\mathbf{x})$ tal que $F(\mathbf{x}, f(\mathbf{x})) = \mathbf{0}$?

Teorema de la Función Implícita: Motivación

Si se tiene una función vectorial $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ tal que $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, con $\mathbf{x} \in \mathbb{R}^n$ y $\mathbf{y} \in \mathbb{R}^m$,

¿Es posible definir una función $\mathbf{y} = f(\mathbf{x})$ tal que $F(\mathbf{x}, f(\mathbf{x})) = \mathbf{0}$?

¿Es $f(\mathbf{x})$ diferenciable y cuál sería su derivada?

Teorema de la Función Implícita: Motivación

Si se tiene una función vectorial $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ tal que $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, con $\mathbf{x} \in \mathbb{R}^n$ y $\mathbf{y} \in \mathbb{R}^m$,

¿Es posible definir una función $\mathbf{y} = f(\mathbf{x})$ tal que $F(\mathbf{x}, f(\mathbf{x})) = \mathbf{0}$?

¿Es $f(\mathbf{x})$ diferenciable y cuál sería su derivada?

Ejemplo: Sea $F(x, y) = 2x - 3y$ y tome $y = f(x) = \frac{2}{3}x$ entonces

$$F(x, f(x)) = 2x - 3 \cdot \frac{2}{3}x = 0 \quad \text{y} \quad f'(x) = \frac{2}{3} = -\frac{\partial_x F(x, f(x))}{\partial_y F(x, f(x))}.$$

Teorema de la Función Implícita: Motivación

Si se tiene una función vectorial $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ tal que $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, con $\mathbf{x} \in \mathbb{R}^n$ y $\mathbf{y} \in \mathbb{R}^m$,

¿Es posible definir una función $\mathbf{y} = f(\mathbf{x})$ tal que $F(\mathbf{x}, f(\mathbf{x})) = \mathbf{0}$?

¿Es $f(\mathbf{x})$ diferenciable y cuál sería su derivada?

Ejemplo: Sea $F(x, y) = 2x - 3y$ y tome $y = f(x) = \frac{2}{3}x$ entonces

$$F(x, f(x)) = 2x - 3 \cdot \frac{2}{3}x = 0 \quad \text{y} \quad f'(x) = \frac{2}{3} = -\frac{\partial_x F(x, f(x))}{\partial_y F(x, f(x))}.$$

Contraejemplo: Esto no siempre se puede hacer! Considere

$F(x, y) = (x^2 + y^2) - (y + x)(y - x)$. Como $F(x, y) = 2x^2$ (al simplificar), entonces $\nexists y = f(x)$ tal que $F(x, f(x)) = 2x^2 = 0$ para $x \neq 0$.

Teorema de la Función Implícita: Motivación

Si se tiene una función vectorial $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ tal que $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, con $\mathbf{x} \in \mathbb{R}^n$ y $\mathbf{y} \in \mathbb{R}^m$,

¿Es posible definir una función $\mathbf{y} = f(\mathbf{x})$ tal que $F(\mathbf{x}, f(\mathbf{x})) = \mathbf{0}$?

¿Es $f(\mathbf{x})$ diferenciable y cuál sería su derivada?

Ejemplo: Sea $F(x, y) = 2x - 3y$ y tome $y = f(x) = \frac{2}{3}x$ entonces

$$F(x, f(x)) = 2x - 3 \cdot \frac{2}{3}x = 0 \quad \text{y} \quad f'(x) = \frac{2}{3} = -\frac{\partial_x F(x, f(x))}{\partial_y F(x, f(x))}.$$

Contraejemplo: Esto no siempre se puede hacer! Considere

$F(x, y) = (x^2 + y^2) - (y + x)(y - x)$. Como $F(x, y) = 2x^2$ (al simplificar), entonces $\nexists y = f(x)$ tal que $F(x, f(x)) = 2x^2 = 0$ para $x \neq 0$.

¿Qué condiciones me indicarían si existe $\mathbf{y} = f(\mathbf{x})$ o no?

Teorema de la Función Implícita: Motivación

Si se tiene una función vectorial $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ tal que $F(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, con $\mathbf{x} \in \mathbb{R}^n$ y $\mathbf{y} \in \mathbb{R}^m$,

¿Es posible definir una función $\mathbf{y} = f(\mathbf{x})$ tal que $F(\mathbf{x}, f(\mathbf{x})) = \mathbf{0}$?

¿Es $f(\mathbf{x})$ diferenciable y cuál sería su derivada?

Ejemplo: Sea $F(x, y) = 2x - 3y$ y tome $y = f(x) = \frac{2}{3}x$ entonces

$$F(x, f(x)) = 2x - 3 \cdot \frac{2}{3}x = 0 \quad \text{y} \quad f'(x) = \frac{2}{3} = -\frac{\partial_x F(x, f(x))}{\partial_y F(x, f(x))}.$$

Contraejemplo: Esto no siempre se puede hacer! Considere

$F(x, y) = (x^2 + y^2) - (y + x)(y - x)$. Como $F(x, y) = 2x^2$ (al simplificar), entonces $\nexists y = f(x)$ tal que $F(x, f(x)) = 2x^2 = 0$ para $x \neq 0$.

¿Qué condiciones me indicarían si existe $\mathbf{y} = f(\mathbf{x})$ o no?

Note que en el primer ejemplo se tiene $\partial_y F(x, y) = -3 \neq 0$ mientras que en el segundo $\partial_y F(x, y) = 0$. ¿Esta condición sería suficiente o necesaria?

Teorema de la Función Implícita

Teorema

Sea $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ una función vectorial diferenciable en un vecindario $B_r(\mathbf{x}_0, \mathbf{y}_0)$, tal que $F(\mathbf{x}_0, \mathbf{y}_0) = (F_1(\mathbf{x}_0, \mathbf{y}_0), \dots, F_m(\mathbf{x}_0, \mathbf{y}_0)) = \mathbf{0}$, con $\mathbf{x}_0 \in \mathbb{R}^n$ y $\mathbf{y}_0 \in \mathbb{R}^m$, y

$$|J_{F,\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)| := \det \begin{vmatrix} \frac{\partial F_1}{\partial y_1}(\mathbf{x}_0, \mathbf{y}_0) & \cdots & \frac{\partial F_1}{\partial y_m}(\mathbf{x}_0, \mathbf{y}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1}(\mathbf{x}_0, \mathbf{y}_0) & \cdots & \frac{\partial F_m}{\partial y_m}(\mathbf{x}_0, \mathbf{y}_0) \end{vmatrix} \neq 0.$$

Entonces existe una única función vectorial diferenciable $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ t.q.

- $f(\mathbf{x}_0) = \mathbf{y}_0$;
- $F(\mathbf{x}, f(\mathbf{x})) = \mathbf{0}$, $\forall \mathbf{x} \in B_\epsilon(\mathbf{x}_0)$;
- $J_f(\mathbf{x}) = -J_{F,\mathbf{y}}^{-1}(\mathbf{x}, f(\mathbf{x})) J_{F,\mathbf{x}}(\mathbf{x}, f(\mathbf{x}))$.

Teorema de la Función Implícita

Teorema

Sea $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ una función vectorial diferenciable en un vecindario $B_r(\mathbf{x}_0, \mathbf{y}_0)$, tal que $F(\mathbf{x}_0, \mathbf{y}_0) = (F_1(\mathbf{x}_0, \mathbf{y}_0), \dots, F_m(\mathbf{x}_0, \mathbf{y}_0)) = \mathbf{0}$, con $\mathbf{x}_0 \in \mathbb{R}^n$ y $\mathbf{y}_0 \in \mathbb{R}^m$, y

$$|J_{F,\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)| := \det \begin{vmatrix} \frac{\partial F_1}{\partial y_1}(\mathbf{x}_0, \mathbf{y}_0) & \cdots & \frac{\partial F_1}{\partial y_m}(\mathbf{x}_0, \mathbf{y}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1}(\mathbf{x}_0, \mathbf{y}_0) & \cdots & \frac{\partial F_m}{\partial y_m}(\mathbf{x}_0, \mathbf{y}_0) \end{vmatrix} \neq 0.$$

Entonces existe una única función vectorial diferenciable $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ t.q.

- $f(\mathbf{x}_0) = \mathbf{y}_0$;
- $F(\mathbf{x}, f(\mathbf{x})) = 0, \forall \mathbf{x} \in B_\epsilon(\mathbf{x}_0)$;
- $J_f(\mathbf{x}) = -J_{F,\mathbf{y}}^{-1}(\mathbf{x}, f(\mathbf{x})) J_{F,\mathbf{x}}(\mathbf{x}, f(\mathbf{x}))$.

Si $|J_{F,\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)| \neq 0$ existe f , pero si $|J_{F,\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)| = 0$ no sabemos.

Teorema de la Función Implícita

Teorema

Sea $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ una función vectorial diferenciable en un vecindario $B_r(\mathbf{x}_0, \mathbf{y}_0)$, tal que $F(\mathbf{x}_0, \mathbf{y}_0) = (F_1(\mathbf{x}_0, \mathbf{y}_0), \dots, F_m(\mathbf{x}_0, \mathbf{y}_0)) = \mathbf{0}$, con $\mathbf{x}_0 \in \mathbb{R}^n$ y $\mathbf{y}_0 \in \mathbb{R}^m$, y

$$|J_{F,\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)| := \det \begin{vmatrix} \frac{\partial F_1}{\partial y_1}(\mathbf{x}_0, \mathbf{y}_0) & \cdots & \frac{\partial F_1}{\partial y_m}(\mathbf{x}_0, \mathbf{y}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1}(\mathbf{x}_0, \mathbf{y}_0) & \cdots & \frac{\partial F_m}{\partial y_m}(\mathbf{x}_0, \mathbf{y}_0) \end{vmatrix} \neq 0.$$

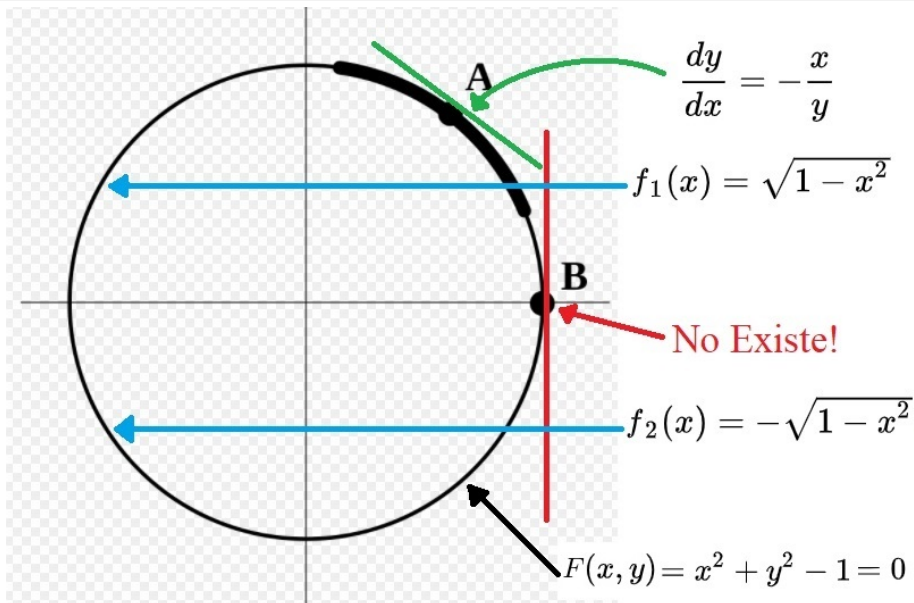
Entonces existe una única función vectorial diferenciable $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ t.q.

- $f(\mathbf{x}_0) = \mathbf{y}_0$;
- $F(\mathbf{x}, f(\mathbf{x})) = 0, \forall \mathbf{x} \in B_\epsilon(\mathbf{x}_0)$;
- $J_f(\mathbf{x}) = -J_{F,\mathbf{y}}^{-1}(\mathbf{x}, f(\mathbf{x})) J_{F,\mathbf{x}}(\mathbf{x}, f(\mathbf{x}))$.

Si $|J_{F,\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)| \neq 0$ existe f , pero si $|J_{F,\mathbf{y}}(\mathbf{x}_0, \mathbf{y}_0)| = 0$ no sabemos.

Ejemplo: Sea $F(x, y) = x - y^3$ y $y = f(x) = \sqrt[3]{x} \Rightarrow F(x, f(x)) = 0$ y $f'(x) = \frac{1}{3}x^{-\frac{2}{3}} = -\frac{\partial_x F(x, f(x))}{\partial_y F(x, f(x))}$ pero $|J_{F,\mathbf{y}}(0, 0)| = \frac{\partial F}{\partial y}(0, 0) = -3y^2|_{(0,0)} = 0$.

Ejemplo: Círculo



Ejemplo 1: $n = 2$, $m = 1$

Sea $F(x, y, z) = x^2 + y^2 + z^2 - 1$ (esfera de radio 1). Indique si existe una función $z = f(x, y)$ y determine su derivada implícita y explícitamente.

Ejemplo 1: $n = 2$, $m = 1$

Sea $F(x, y, z) = x^2 + y^2 + z^2 - 1$ (esfera de radio 1). Indique si existe una función $z = f(x, y)$ y determine su derivada implícita y explícitamente.

Solución: Calculando el Jacobiano parcial se tiene que

$$|J_{F,z}(x_0, y_0, z_0)| = 2z_0 \neq 0 \quad \text{si} \quad z_0 \neq 0.$$

Ejemplo 1: $n = 2$, $m = 1$

Sea $F(x, y, z) = x^2 + y^2 + z^2 - 1$ (esfera de radio 1). Indique si existe una función $z = f(x, y)$ y determine su derivada implícita y explícitamente.

Solución: Calculando el Jacobiano parcial se tiene que

$$|J_{F,z}(x_0, y_0, z_0)| = 2z_0 \neq 0 \quad \text{si} \quad z_0 \neq 0.$$

Por el teorema sabemos que si $z_0 \neq 0$ entonces existe $z = f(x, y)$ tal que

$F(x, y, f(x, y)) = 0$. Además, como $J_{F,(x,y)}(x, y, z) = (2x, 2y)$ y

$J_{F,z}^{-1}(x, y, z) = \frac{1}{J_{F,z}(x, y, z)} = \frac{1}{2z}$, se concluye que el Jacobiano de f es

$$J_f(x, y) = -\frac{J_{F,(x,y)}(x, y, f(x, y))}{J_{F,z}(x, y, f(x, y))} = -\frac{(2x, 2y)}{2f(x, y)} = \left(\frac{-x}{f(x, y)}, \frac{-y}{f(x, y)} \right).$$

Ejemplo 1: $n = 2$, $m = 1$

Sea $F(x, y, z) = x^2 + y^2 + z^2 - 1$ (esfera de radio 1). Indique si existe una función $z = f(x, y)$ y determine su derivada implícita y explícitamente.

Solución: Calculando el Jacobiano parcial se tiene que

$$|J_{F,z}(x_0, y_0, z_0)| = 2z_0 \neq 0 \quad \text{si} \quad z_0 \neq 0.$$

Por el teorema sabemos que si $z_0 \neq 0$ entonces existe $z = f(x, y)$ tal que

$F(x, y, f(x, y)) = 0$. Además, como $J_{F,(x,y)}(x, y, z) = (2x, 2y)$ y

$J_{F,z}^{-1}(x, y, z) = \frac{1}{J_{F,z}(x,y,z)} = \frac{1}{2z}$, se concluye que el Jacobiano de f es

$$J_f(x, y) = -\frac{J_{F,(x,y)}(x, y, f(x, y))}{J_{F,z}(x, y, f(x, y))} = -\frac{(2x, 2y)}{2f(x, y)} = \left(\frac{-x}{f(x, y)}, \frac{-y}{f(x, y)} \right).$$

Note que $z = f(x, y) = \sqrt{1 - x^2 - y^2}$ es tal que $F((x, y, f(x, y))) = 0$ y si $z \neq 0$ (equivalentemente $x^2 + y^2 < 1$), entonces

$$J_f(x, y) = \left(\frac{-x}{\sqrt{1 - x^2 - y^2}}, \frac{-y}{\sqrt{1 - x^2 - y^2}} \right) = \left(\frac{-x}{f(x, y)}, \frac{-y}{f(x, y)} \right),$$

verificando el teorema.

Ejemplo 2: $n = 2, m = 2$

Sea $F(x, y, z, w) = (x^2 + y^2 + z^2 + w^2 - 2, x^2 - y^2 + z^2 - w^2)$. Indique si existe una función $(z, w) = f(x, y)$ y determine su derivada.

Ejemplo 2: $n = 2, m = 2$

Sea $F(x, y, z, w) = (x^2 + y^2 + z^2 + w^2 - 2, x^2 - y^2 + z^2 - w^2)$. Indique si existe una función $(z, w) = f(x, y)$ y determine su derivada.

Solución: Calculando el Jacobiano parcial se tiene que

$$|J_{F,(z,w)}(x_0, y_0, z_0, w_0)| = \det \begin{vmatrix} 2z_0 & 2w_0 \\ 2z_0 & -2w_0 \end{vmatrix} = -8z_0w_0 \neq 0 \quad \text{si} \quad z_0, w_0 \neq 0.$$

Ejemplo 2: $n = 2, m = 2$

Sea $F(x, y, z, w) = (x^2 + y^2 + z^2 + w^2 - 2, x^2 - y^2 + z^2 - w^2)$. Indique si existe una función $(z, w) = f(x, y)$ y determine su derivada.

Solución: Calculando el Jacobiano parcial se tiene que

$$|J_{F,(z,w)}(x_0, y_0, z_0, w_0)| = \det \begin{vmatrix} 2z_0 & 2w_0 \\ 2z_0 & -2w_0 \end{vmatrix} = -8z_0w_0 \neq 0 \quad \text{si} \quad z_0, w_0 \neq 0.$$

Si $z_0, w_0 \neq 0$ entonces existe $(z, w) = f(x, y) = (f_1(x, y), f_2(x, y))$. Como

$$J_{F,(x,y)}(x, y, z, w) = \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix} \text{ y } J_{F,z}^{-1}(x, y, z, w) = \begin{pmatrix} \frac{1}{4z} & \frac{1}{4z} \\ \frac{1}{4w} & \frac{-1}{4w} \end{pmatrix},$$

entonces

$$J_f(x, y) = - \begin{pmatrix} \frac{1}{4f_1(x,y)} & \frac{1}{4f_1(x,y)} \\ \frac{1}{4f_2(x,y)} & \frac{-1}{4f_2(x,y)} \end{pmatrix} \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix} = \begin{pmatrix} \frac{-x}{f_1(x,y)} & 0 \\ 0 & \frac{-y}{f_2(x,y)} \end{pmatrix}.$$

Ejemplo 2: $n = 2, m = 2$

Sea $F(x, y, z, w) = (x^2 + y^2 + z^2 + w^2 - 2, x^2 - y^2 + z^2 - w^2)$. Indique si existe una función $(z, w) = f(x, y)$ y determine su derivada.

Solución: Calculando el Jacobiano parcial se tiene que

$$|J_{F,(z,w)}(x_0, y_0, z_0, w_0)| = \det \begin{vmatrix} 2z_0 & 2w_0 \\ 2z_0 & -2w_0 \end{vmatrix} = -8z_0w_0 \neq 0 \quad \text{si} \quad z_0, w_0 \neq 0.$$

Si $z_0, w_0 \neq 0$ entonces existe $(z, w) = f(x, y) = (f_1(x, y), f_2(x, y))$. Como

$$J_{F,(x,y)}(x, y, z, w) = \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix} \text{ y } J_{F,z}^{-1}(x, y, z, w) = \begin{pmatrix} \frac{1}{4z} & \frac{1}{4z} \\ \frac{1}{4w} & \frac{-1}{4w} \end{pmatrix},$$

entonces

$$J_f(x, y) = - \begin{pmatrix} \frac{1}{4f_1(x,y)} & \frac{1}{4f_1(x,y)} \\ \frac{1}{4f_2(x,y)} & \frac{-1}{4f_2(x,y)} \end{pmatrix} \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix} = \begin{pmatrix} \frac{-x}{f_1(x,y)} & 0 \\ 0 & \frac{-y}{f_2(x,y)} \end{pmatrix}.$$

Si $z = f_1(x, y) = \sqrt{1 - x^2}$ y $w = f_2(x, y) = \sqrt{1 - y^2}$, se cumple que $F((x, y, f(x, y))) = 0$ y si $z, w \neq 0$ (equivalentemente $x, y < 1$), entonces

$$J_f(x, y) = \begin{pmatrix} \frac{-x}{\sqrt{1-x^2}} & 0 \\ 0 & \frac{-y}{\sqrt{1-y^2}} \end{pmatrix} = \begin{pmatrix} \frac{-x}{f_1(x,y)} & 0 \\ 0 & \frac{-y}{f_2(x,y)} \end{pmatrix}.$$

Teorema de la Función Inversa

Si se tiene una función vectorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $\mathbf{y} = F(\mathbf{x})$,

¿Es posible definir una función $F^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $F(F^{-1}(\mathbf{y})) = \mathbf{y}$?

Teorema de la Función Inversa

Si se tiene una función vectorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $\mathbf{y} = F(\mathbf{x})$,

¿Es posible definir una función $F^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $F(F^{-1}(\mathbf{y})) = \mathbf{y}$?

¿Es $F^{-1}(\mathbf{x})$ diferenciable y cuál sería su derivada?

Teorema de la Función Inversa

Si se tiene una función vectorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $\mathbf{y} = F(\mathbf{x})$,

¿Es posible definir una función $F^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $F(F^{-1}(\mathbf{y})) = \mathbf{y}$?

¿Es $F^{-1}(\mathbf{x})$ diferenciable y cuál sería su derivada?

Ejemplo: Sea $y = f(x) = \sqrt[3]{x-1}$ y tome $x = f^{-1}(y) = y^3 + 1$ entonces $f(f^{-1}(y)) = y$ y $\partial_y f^{-1}(y) = 3y^2 = \frac{1}{\frac{1}{3}([y^3+1]-1)^{-\frac{2}{3}}} = \frac{1}{f'(f^{-1}(y))} = J_f^{-1}(f^{-1}(y))$.

Teorema de la Función Inversa

Si se tiene una función vectorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $\mathbf{y} = F(\mathbf{x})$,

¿Es posible definir una función $F^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $F(F^{-1}(\mathbf{y})) = \mathbf{y}$?

¿Es $F^{-1}(\mathbf{x})$ diferenciable y cuál sería su derivada?

Ejemplo: Sea $y = f(x) = \sqrt[3]{x-1}$ y tome $x = f^{-1}(y) = y^3 + 1$ entonces $f(f^{-1}(y)) = y$ y $\partial_y f^{-1}(y) = 3y^2 = \frac{1}{\frac{1}{3}([y^3+1]-1)^{-\frac{2}{3}}} = \frac{1}{f'(f^{-1}(y))} = J_f^{-1}(f^{-1}(y))$.

Teorema

Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función vectorial diferenciable en un vecindario $B_r(\mathbf{x}_0)$, tal que $J_F(\mathbf{x})$ es continua (i.e. $\frac{f_j}{\partial x_i}$ son continuas), con $\mathbf{x} \in B_r(\mathbf{x}_0)$, y $\det|J_F(\mathbf{x}_0)| \neq 0$. Entonces existe un vecindario $B_\epsilon(\mathbf{y}_0)$ tal que existe una única función vectorial diferenciable $F^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ con $F(F^{-1}(\mathbf{y})) = \mathbf{y}$, $\forall \mathbf{y} \in B_\epsilon(\mathbf{y}_0)$, y

$$J_{F^{-1}}(\mathbf{y}) = J_F^{-1}(F^{-1}(\mathbf{y})).$$

Teorema de la Función Inversa

Si se tiene una función vectorial $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $\mathbf{y} = F(\mathbf{x})$,

¿Es posible definir una función $F^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tal que $F(F^{-1}(\mathbf{y})) = \mathbf{y}$?

¿Es $F^{-1}(\mathbf{x})$ diferenciable y cuál sería su derivada?

Ejemplo: Sea $y = f(x) = \sqrt[3]{x-1}$ y tome $x = f^{-1}(y) = y^3 + 1$ entonces $f(f^{-1}(y)) = y$ y $\partial_y f^{-1}(y) = 3y^2 = \frac{1}{\frac{1}{3}([y^3+1]-1)^{-\frac{2}{3}}} = \frac{1}{f'(f^{-1}(y))} = J_f^{-1}(f^{-1}(y))$.

Teorema

Sea $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ una función vectorial diferenciable en un vecindario $B_r(\mathbf{x}_0)$, tal que $J_F(\mathbf{x})$ es continua (i.e. $\frac{f_j}{\partial x_i}$ son continuas), con $\mathbf{x} \in B_r(\mathbf{x}_0)$, y $\det|J_F(\mathbf{x}_0)| \neq 0$. Entonces existe un vecindario $B_\epsilon(\mathbf{y}_0)$ tal que existe una única función vectorial diferenciable $F^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ con $F(F^{-1}(\mathbf{y})) = \mathbf{y}$, $\forall \mathbf{y} \in B_\epsilon(\mathbf{y}_0)$, y

$$J_{F^{-1}}(\mathbf{y}) = J_F^{-1}(F^{-1}(\mathbf{y})).$$

Nota: No siempre es posible encontrar la función inversa de manera explícita, aunque sabríamos que sí existe.

Ejemplo 1

Considere $(x, y) = F(r, \theta) = (r \cos(\theta), r \sin(\theta))$. Determine si existe F^{-1} y establezca su derivada directamente y con el teorema.

Ejemplo 1

Considere $(x, y) = F(r, \theta) = (r \cos(\theta), r \sin(\theta))$. Determine si existe F^{-1} y establezca su derivada directamente y con el teorema.

Solución: Note que $r = \sqrt{x^2 + y^2}$ y $\theta = \tan^{-1} \left(\frac{y}{x} \right)$ son tales que

$$F \left(\sqrt{x^2 + y^2}, \tan^{-1} \left(\frac{y}{x} \right) \right) = (x, y), \quad \left[\cos \left(\tan^{-1} \left(\frac{y}{x} \right) \right) = \frac{x}{\sqrt{x^2 + y^2}} \right], \text{ i.e.}$$

$$F^{-1}(x, y) = \left(\sqrt{x^2 + y^2}, \tan^{-1} \left(\frac{y}{x} \right) \right) \text{ y } J_{F^{-1}}(x, y) = \begin{pmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \\ \frac{-y}{x^2 + y^2} & \frac{x}{x^2 + y^2} \end{pmatrix}.$$

Ejemplo 1

Considere $(x, y) = F(r, \theta) = (r \cos(\theta), r \sin(\theta))$. Determine si existe F^{-1} y establezca su derivada directamente y con el teorema.

Solución: Note que $r = \sqrt{x^2 + y^2}$ y $\theta = \tan^{-1} \left(\frac{y}{x} \right)$ son tales que

$$F \left(\sqrt{x^2 + y^2}, \tan^{-1} \left(\frac{y}{x} \right) \right) = (x, y), \quad \left[\cos \left(\tan^{-1} \left(\frac{y}{x} \right) \right) = \frac{x}{\sqrt{x^2 + y^2}} \right], \text{ i.e.}$$

$$F^{-1}(x, y) = \left(\sqrt{x^2 + y^2}, \tan^{-1} \left(\frac{y}{x} \right) \right) \text{ y } J_{F^{-1}}(x, y) = \begin{pmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \\ \frac{-y}{x^2 + y^2} & \frac{x}{x^2 + y^2} \end{pmatrix}.$$

Calculando el Jacobiano se tiene que

$$|J_F(r_0, \theta_0)| = \det \begin{vmatrix} \cos(\theta_0) & -r_0 \sin(\theta_0) \\ \sin(\theta_0) & r_0 \cos(\theta_0) \end{vmatrix} = r_0 \neq 0 \quad \text{si} \quad r_0 \neq 0.$$

Ejemplo 1

Considere $(x, y) = F(r, \theta) = (r \cos(\theta), r \sin(\theta))$. Determine si existe F^{-1} y establezca su derivada directamente y con el teorema.

Solución: Note que $r = \sqrt{x^2 + y^2}$ y $\theta = \tan^{-1}(\frac{y}{x})$ son tales que

$$F\left(\sqrt{x^2 + y^2}, \tan^{-1}\left(\frac{y}{x}\right)\right) = (x, y), \quad \left[\cos\left(\tan^{-1}\left(\frac{y}{x}\right)\right) = \frac{x}{\sqrt{x^2 + y^2}}\right], \text{ i.e.}$$

$$F^{-1}(x, y) = \left(\sqrt{x^2 + y^2}, \tan^{-1}\left(\frac{y}{x}\right)\right) \text{ y } J_{F^{-1}}(x, y) = \begin{pmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \\ \frac{-y}{x^2 + y^2} & \frac{x}{x^2 + y^2} \end{pmatrix}.$$

Calculando el Jacobiano se tiene que

$$|J_F(r_0, \theta_0)| = \det \begin{vmatrix} \cos(\theta_0) & -r_0 \sin(\theta_0) \\ \sin(\theta_0) & r_0 \cos(\theta_0) \end{vmatrix} = r_0 \neq 0 \quad \text{si} \quad r_0 \neq 0.$$

Si $r_0 \neq 0$ existe $F^{-1}(x, y)$ y como $J_F^{-1}(r, \theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\frac{\sin(\theta)}{r} & \frac{\cos(\theta)}{r} \end{pmatrix}$,

$$\Rightarrow J_{F^{-1}}(x, y) = J_F^{-1}(F^{-1}(x, y)) = \begin{pmatrix} \frac{x}{\sqrt{x^2 + y^2}} & \frac{y}{\sqrt{x^2 + y^2}} \\ \frac{-y}{x^2 + y^2} & \frac{x}{x^2 + y^2} \end{pmatrix}.$$

Ejemplo 2

Considere $(z, w) = F(x, y) = \left(\frac{x^2+y^2}{x}, e^x \sin(y)\right)$. Determine si existe $F^{-1}(z, w)$ y establezca su derivada.

Ejemplo 2

Considere $(z, w) = F(x, y) = \left(\frac{x^2+y^2}{x}, e^x \sin(y)\right)$. Determine si existe $F^{-1}(z, w)$ y establezca su derivada.

Solución: Calculando el Jacobiano se tiene

$$\begin{aligned} |J_F(x_0, y_0)| &= \det \begin{vmatrix} 1 - \frac{y_0^2}{x_0^2} & \frac{2y_0}{x_0} \\ e^{x_0} \sin(y_0) & e^{x_0} \cos(y_0) \end{vmatrix} \\ &= e^{x_0} \left(\left(1 - \frac{y_0^2}{x_0^2}\right) \cos(y_0) - \frac{2y_0}{x_0} \sin(y_0) \right) \neq 0 \end{aligned}$$

Ejemplo 2

Considere $(z, w) = F(x, y) = \left(\frac{x^2+y^2}{x}, e^x \sin(y)\right)$. Determine si existe $F^{-1}(z, w)$ y establezca su derivada.

Solución: Calculando el Jacobiano se tiene

$$\begin{aligned} |J_F(x_0, y_0)| &= \det \begin{vmatrix} 1 - \frac{y_0^2}{x_0^2} & \frac{2y_0}{x_0} \\ e^{x_0} \sin(y_0) & e^{x_0} \cos(y_0) \end{vmatrix} \\ &= e^{x_0} \left(\left(1 - \frac{y_0^2}{x_0^2}\right) \cos(y_0) - \frac{2y_0}{x_0} \sin(y_0) \right) \neq 0 \end{aligned}$$

Si $\left(\frac{x_0}{y_0} - \frac{y_0}{x_0}\right) \neq 2 \tan(y_0)$ entonces existe $F^{-1}(z, w)$ y

$$\begin{aligned} J_{F^{-1}}(z, w) &= J_F^{-1}(x(z, w), y(z, w)) \\ &= \begin{pmatrix} \frac{x^2 \cos(y)}{(x^2 - y^2) \cos(y) - 2xy \sin(y)} & -\frac{2e^{-x} xy}{(x^2 - y^2) \cos(y) - 2xy \sin(y)} \\ \frac{x^2 \sin(y)}{(y^2 - x^2) \cos(y) + 2xy \sin(y)} & \frac{e^{-x} (x^2 - y^2)}{(x^2 - y^2) \cos(y) - 2xy \sin(y)} \end{pmatrix} \end{aligned}$$

Ejemplo 2

Considere $(z, w) = F(x, y) = \left(\frac{x^2+y^2}{x}, e^x \sin(y)\right)$. Determine si existe $F^{-1}(z, w)$ y establezca su derivada.

Solución: Calculando el Jacobiano se tiene

$$\begin{aligned} |J_F(x_0, y_0)| &= \det \begin{vmatrix} 1 - \frac{y_0^2}{x_0^2} & \frac{2y_0}{x_0} \\ e^{x_0} \sin(y_0) & e^{x_0} \cos(y_0) \end{vmatrix} \\ &= e^{x_0} \left(\left(1 - \frac{y_0^2}{x_0^2}\right) \cos(y_0) - \frac{2y_0}{x_0} \sin(y_0) \right) \neq 0 \end{aligned}$$

Si $\left(\frac{x_0}{y_0} - \frac{y_0}{x_0}\right) \neq 2 \tan(y_0)$ entonces existe $F^{-1}(z, w)$ y

$$\begin{aligned} J_{F^{-1}}(z, w) &= J_F^{-1}(x(z, w), y(z, w)) \\ &= \begin{pmatrix} \frac{x^2 \cos(y)}{(x^2 - y^2) \cos(y) - 2xy \sin(y)} & -\frac{2e^{-x} xy}{(x^2 - y^2) \cos(y) - 2xy \sin(y)} \\ \frac{x^2 \sin(y)}{(y^2 - x^2) \cos(y) + 2xy \sin(y)} & \frac{e^{-x} (x^2 - y^2)}{(x^2 - y^2) \cos(y) - 2xy \sin(y)} \end{pmatrix} \end{aligned}$$

Note que es posible encontrar $F^{-1}(z, w)$ explícitamente.

Menores Principales Arbitrarios (M.P.A.)

Definición

Sea $A = (a_{i,j})_{1 \leq i,j \leq n}$ una matriz cuadrada de tamaño $n \times n$. Se llaman *menores principales arbitrarios* de orden r (y se denotan Δ_r) a los determinantes de las matrices que se obtienen suprimiendo $n - r$ filas y $n - r$ columnas con los mismos índices.

Menores Principales Arbitrarios (M.P.A.)

Definición

Sea $A = (a_{i,j})_{1 \leq i,j \leq n}$ una matriz cuadrada de tamaño $n \times n$. Se llaman *menores principales arbitrarios* de orden r (y se denotan Δ_r) a los determinantes de las matrices que se obtienen suprimiendo $n - r$ filas y $n - r$ columnas con los mismos índices.

Ejemplo: Sea A una matriz cuadrada 3×3 .

- **Menores Principales Arbitrarios de Orden 1:**

$$\Delta_1 = a_{1,1}; \quad \Delta_1 = a_{2,2}; \quad \Delta_1 = a_{3,3}.$$

Menores Principales Arbitrarios (M.P.A.)

Definición

Sea $A = (a_{i,j})_{1 \leq i,j \leq n}$ una matriz cuadrada de tamaño $n \times n$. Se llaman *menores principales arbitrarios* de orden r (y se denotan Δ_r) a los determinantes de las matrices que se obtienen suprimiendo $n - r$ filas y $n - r$ columnas con los mismos índices.

Ejemplo: Sea A una matriz cuadrada 3×3 .

- **Menores Principales Arbitrarios de Orden 1:**

$$\Delta_1 = a_{1,1}; \quad \Delta_1 = a_{2,2}; \quad \Delta_1 = a_{3,3}.$$

- **Menores Principales Arbitrarios de Orden 2:**

$$\Delta_2 = \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix}; \quad \Delta_2 = \begin{vmatrix} a_{1,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{vmatrix}; \quad \Delta_2 = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix}.$$

Menores Principales Arbitrarios (M.P.A.)

Definición

Sea $A = (a_{i,j})_{1 \leq i,j \leq n}$ una matriz cuadrada de tamaño $n \times n$. Se llaman *menores principales arbitrarios* de orden r (y se denotan Δ_r) a los determinantes de las matrices que se obtienen suprimiendo $n - r$ filas y $n - r$ columnas con los mismos índices.

Ejemplo: Sea A una matriz cuadrada 3×3 .

- **Menores Principales Arbitrarios de Orden 1:**

$$\Delta_1 = a_{1,1}; \quad \Delta_1 = a_{2,2}; \quad \Delta_1 = a_{3,3}.$$

- **Menores Principales Arbitrarios de Orden 2:**

$$\Delta_2 = \begin{vmatrix} a_{2,2} & a_{2,3} \\ a_{3,2} & a_{3,3} \end{vmatrix}; \quad \Delta_2 = \begin{vmatrix} a_{1,1} & a_{1,3} \\ a_{3,1} & a_{3,3} \end{vmatrix}; \quad \Delta_2 = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix}.$$

- **Menores Principales Arbitrarios de Orden 3:**

$$\Delta_3 = \det A.$$

Menores Principales Dominantes (M.P.D.)

Definición

Sea $A = (a_{i,j})_{1 \leq i,j \leq n}$ una matriz $n \times n$. Se llaman *menores principales dominantes* de orden r (y se denotan D_r) a los siguientes determinantes:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,r} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r,1} & a_{r,2} & \cdots & a_{r,r} \end{vmatrix}$$

Menores Principales Dominantes (M.P.D.)

Definición

Sea $A = (a_{i,j})_{1 \leq i,j \leq n}$ una matriz $n \times n$. Se llaman *menores principales dominantes* de orden r (y se denotan D_r) a los siguientes determinantes:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,r} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r,1} & a_{r,2} & \cdots & a_{r,r} \end{vmatrix}$$

Ejemplo: Sea A una matriz cuadrada 3×3 .

- **Menor Principal Dominante de Orden 1:** $D_1 = a_{1,1}$

Menores Principales Dominantes (M.P.D.)

Definición

Sea $A = (a_{i,j})_{1 \leq i,j \leq n}$ una matriz $n \times n$. Se llaman *menores principales dominantes* de orden r (y se denotan D_r) a los siguientes determinantes:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,r} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r,1} & a_{r,2} & \cdots & a_{r,r} \end{vmatrix}$$

Ejemplo: Sea A una matriz cuadrada 3×3 .

- **Menor Principal Dominante de Orden 1:** $D_1 = a_{1,1}$
- **Menor Principal Dominante de Orden 2:**

$$D_2 = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

Menores Principales Dominantes (M.P.D.)

Definición

Sea $A = (a_{i,j})_{1 \leq i,j \leq n}$ una matriz $n \times n$. Se llaman *menores principales dominantes* de orden r (y se denotan D_r) a los siguientes determinantes:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,r} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r,1} & a_{r,2} & \cdots & a_{r,r} \end{vmatrix}$$

Ejemplo: Sea A una matriz cuadrada 3×3 .

- **Menor Principal Dominante de Orden 1:** $D_1 = a_{1,1}$
- **Menor Principal Dominante de Orden 2:**

$$D_2 = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

- **Menores Principales de Orden 3:** $D_3 = \det A$.

Matrices Semidefinidas y Definidas

Definición

Se dice que una matriz H es semidefinida positiva (resp. negativa) si se cumple que para todo vector $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$,

$$\mathbf{x}^T H \mathbf{x} \geq 0 \text{ (resp. } \leq 0 \text{).} \quad (1)$$

Si la desigualdad se cumple estrictamente, entonces decimos que H es definida positiva (resp. negativa).

Matrices Semidefinidas y Definidas

Definición

Se dice que una matriz H es semidefinida positiva (resp. negativa) si se cumple que para todo vector $\mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$,

$$\mathbf{x}^T H \mathbf{x} \geq 0 \text{ (resp. } \leq 0 \text{).} \quad (1)$$

Si la desigualdad se cumple estrictamente, entonces decimos que H es definida positiva (resp. negativa).

Ejemplo: Sea

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 9 \end{pmatrix}.$$

Note que para todo $\mathbf{x} = (x_1, x_2, x_3) \neq (0, 0, 0)$ se cumple que

$$\mathbf{x}^T H \mathbf{x} = x_1^2 + 4x_2^2 + 9x_3^2 > 0,$$

ergo, H es definida positiva.

Equivalencia

Semidefinida Positiva/Negativa

Las siguientes condiciones son equivalentes:

- (i) H es semidefinida positiva (semidefinida negativa).
- (ii) Los valores propios de H cumplen que $\lambda_i \geq 0$ ($\lambda_i \leq 0$).
- (iii) Los M.P.A. cumplen que $\Delta_r \geq 0$ ($(-1)^r \Delta_r \geq 0$).

Equivalencia

Semidefinida Positiva/Negativa

Las siguientes condiciones son equivalentes:

- (i) H es semidefinida positiva (semidefinida negativa).
- (ii) Los valores propios de H cumplen que $\lambda_i \geq 0$ ($\lambda_i \leq 0$).
- (iii) Los M.P.A. cumplen que $\Delta_r \geq 0$ ($(-1)^r \Delta_r \geq 0$).

Definida Positiva/Negativa

Las siguientes propiedades son equivalentes:

- (a) H es definida positiva (definida negativa)
- (b) Los valores propios de H cumplen que $\lambda_i > 0$ ($\lambda_i < 0$).
- (c) Los M.P.A. cumplen que $\Delta_r > 0$ ($(-1)^r \Delta_r > 0$).
- (d) Los M.P.D. cumplen que $D_r > 0$ ($(-1)^r D_r > 0$).

Equivalencia

Semidefinida Positiva/Negativa

Las siguientes condiciones son equivalentes:

- (i) H es semidefinida positiva (semidefinida negativa).
- (ii) Los valores propios de H cumplen que $\lambda_i \geq 0$ ($\lambda_i \leq 0$).
- (iii) Los M.P.A. cumplen que $\Delta_r \geq 0$ ($(-1)^r \Delta_r \geq 0$).

Definida Positiva/Negativa

Las siguientes propiedades son equivalentes:

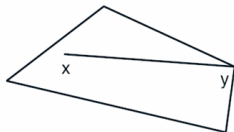
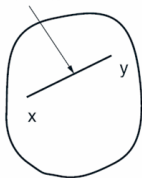
- (a) H es definida positiva (definida negativa)
- (b) Los valores propios de H cumplen que $\lambda_i > 0$ ($\lambda_i < 0$).
- (c) Los M.P.A. cumplen que $\Delta_r > 0$ ($(-1)^r \Delta_r > 0$).
- (d) Los M.P.D. cumplen que $D_r > 0$ ($(-1)^r D_r > 0$).

NOTA: Si $D_r \geq 0$ para todo r **no** implica que H sea semidefinida positiva. Lo mismo para semidefinida negativa.

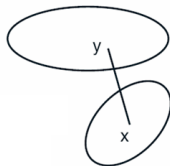
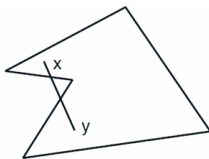
Conjuntos Convexos

Un conjunto $C \subset \mathbb{R}^n$ se dice convexo si para cualesquiera dos puntos en C , el *segmento de línea* que pasa por los puntos también está contenido en C .

$$\alpha x + (1 - \alpha)y, \quad 0 < \alpha < 1$$



Conjuntos Convexos



Conjuntos No Convexos

Para $x, y \in C$, $\lambda x + (1 - \lambda)y \in C$ para todo λ entre 0 y 1.

Operaciones que Preservan Convexidad de Conjuntos

- Multiplicación por un escalar.
- Imagen de transformaciones lineales (translación, rotación).
- Intersección de un número arbitrario de conjuntos.

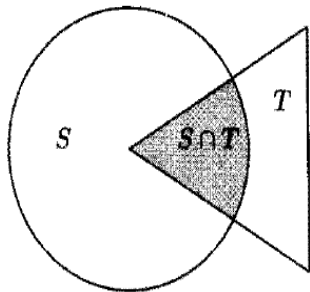


Figura: Intersección de dos conjuntos convexos

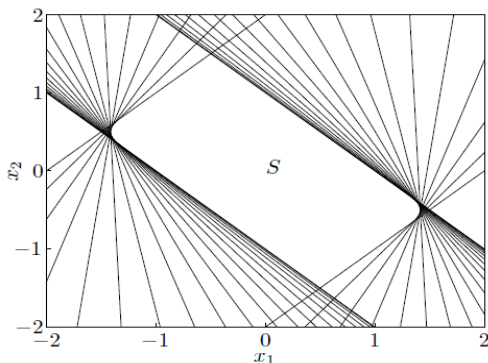


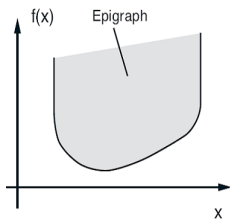
Figura: Intersección de varios conjuntos convexos

Funciones Convexas

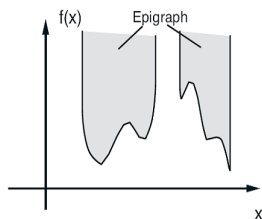
Sea C un conjunto convexo en $\mathbb{R}^n$. Una función $f : C \rightarrow \overline{\mathbb{R}} = [-\infty, \infty]$ se dice *convexa* si y solo si su epigrafo

$$\text{epi } f = \{(\mathbf{x}, r) \mid r \geq f(\mathbf{x})\}$$

es *convexo*, visto como un subconjunto de $\mathbb{R}^n \times \mathbb{R}$.



Función Convexa



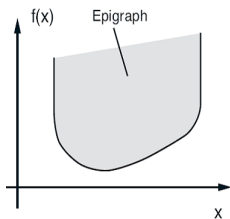
Función No Convexa

Funciones Convexas

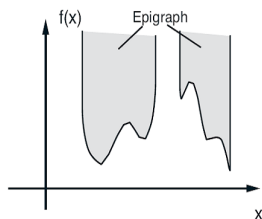
Sea C un conjunto convexo en $\mathbb{R}^n$. Una función $f : C \rightarrow \overline{\mathbb{R}} = [-\infty, \infty]$ se dice *convexa* si y solo si su epigrafo

$$\text{epi } f = \{(\mathbf{x}, r) \mid r \geq f(\mathbf{x})\}$$

es *convexo*, visto como un subconjunto de $\mathbb{R}^n \times \mathbb{R}$.



Función Convexa



Función No Convexa

Una función g es cóncava si $-g$ es convexa. De forma análoga, se habla del hipografo de una función cóncava.

Desigualdad de Jensen

Una función $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$, con S convexo, es *convexa* (resp. *cóncava*) si y solo si satisface la desigualdad

$$f(\lambda_1 \mathbf{x}_1 + \dots + \lambda_m \mathbf{x}_m) \leq \text{ (resp. } \geq \text{)} \lambda_1 f(\mathbf{x}_1) + \dots + \lambda_m f(\mathbf{x}_m)$$

para $\lambda_1 \geq 0, \dots, \lambda_m \geq 0, \lambda_1 + \dots + \lambda_m = 1, \mathbf{x}_1, \dots, \mathbf{x}_m \in S$. Si la desigualdad es estricta, f es estrictamente convexa (resp. cóncava).

Desigualdad de Jensen

Una función $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$, con S convexo, es *convexa* (resp. *cóncava*) si y solo si satisface la desigualdad

$$f(\lambda_1 \mathbf{x}_1 + \dots + \lambda_m \mathbf{x}_m) \leq \text{ (resp. } \geq \text{)} \lambda_1 f(\mathbf{x}_1) + \dots + \lambda_m f(\mathbf{x}_m)$$

para $\lambda_1 \geq 0, \dots, \lambda_m \geq 0, \lambda_1 + \dots + \lambda_m = 1, \mathbf{x}_1, \dots, \mathbf{x}_m \in S$. Si la desigualdad es estricta, f es estrictamente convexa (resp. cóncava).

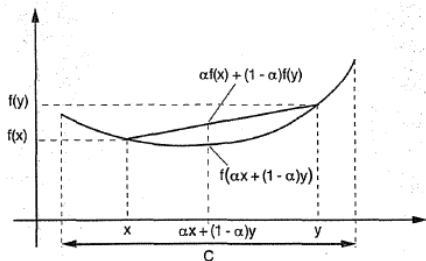


Figura: Función convexa

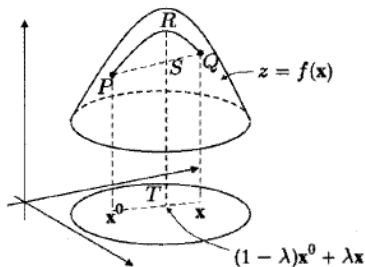


Figura: Función cóncava

Ejemplos de funciones convexas y cóncavas

En $\mathbb{R}$,

Funciones convexas

- Exponencial $f(x) = e^{ax}$, para todo a
- Potencias $f(x) = x^a$, $x > 0$, para todo $a > 1$ ó $a < 0$
- Potencias con valor absoluto $f(x) = |x|^p$, para $p > 1$
- Entropía negativa $f(x) = x \ln x$, $x > 0$

Funciones cóncavas

- Potencias $f(x) = x^a$, $x > 0$, para todo $0 < a < 1$
- Logaritmo $f(x) = \ln x$, $x > 0$

Ejemplos de funciones convexas y cóncavas

En $\mathbb{R}$,

Funciones convexas

- Exponencial $f(x) = e^{ax}$, para todo a
- Potencias $f(x) = x^a$, $x > 0$, para todo $a > 1$ ó $a < 0$
- Potencias con valor absoluto $f(x) = |x|^p$, para $p > 1$
- Entropía negativa $f(x) = x \ln x$, $x > 0$

Funciones cóncavas

- Potencias $f(x) = x^a$, $x > 0$, para todo $0 < a < 1$
- Logaritmo $f(x) = \ln x$, $x > 0$

En $\mathbb{R}^n$,

Funciones convexas

- Norma euclídea $f(\mathbf{x}) = \|\mathbf{x}\|$
- Máximo $f(\mathbf{x}) = \max_n \{x_1, \dots, x_n\}$

Funciones cóncavas

- Media geométrica $f(\mathbf{x}) = (x_1 \cdot x_2 \cdots x_n)^{1/n}$, $x_1, x_2, \dots, x_n > 0$

Ejemplo:

Consideremos la función lineal

$$f(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} + b = a_1x_1 + \cdots + a_nx_n + b$$

donde $\mathbf{a} = (a_1, \dots, a_n)$ es un vector constante y b es una constante.

Probar que f es, a la vez, cóncava y convexa.

Ejemplo:

Consideremos la función lineal

$$f(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} + b = a_1x_1 + \cdots + a_nx_n + b$$

donde $\mathbf{a} = (a_1, \dots, a_n)$ es un vector constante y b es una constante. Probar que f es, a la vez, cóncava y convexa.

Solución: Tome dos puntos $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$ y $\lambda \in (0, 1)$ arbitrarios, y probemos la desigualdad de Jensen.

$$\begin{aligned} f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) &= \mathbf{a} \cdot (\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) + \overbrace{b}^{\lambda b + (1 - \lambda)b} \\ &= \lambda(\mathbf{a} \cdot \mathbf{x}_1) + (1 - \lambda)(\mathbf{a} \cdot \mathbf{x}_2) + \lambda b + (1 - \lambda)b \\ &= \lambda(\mathbf{a} \cdot \mathbf{x}_1 + b) + (1 - \lambda)(\mathbf{a} \cdot \mathbf{x}_2 + b) \\ &= \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2). \end{aligned}$$

Ejemplo:

Consideremos la función lineal

$$f(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} + b = a_1x_1 + \cdots + a_nx_n + b$$

donde $\mathbf{a} = (a_1, \dots, a_n)$ es un vector constante y b es una constante. Probar que f es, a la vez, cóncava y convexa.

Solución: Tome dos puntos $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$ y $\lambda \in (0, 1)$ arbitrarios, y probemos la desigualdad de Jensen.

$$\begin{aligned} f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) &= \mathbf{a} \cdot (\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) + \overbrace{b}^{= \lambda b + (1 - \lambda)b} \\ &= \lambda(\mathbf{a} \cdot \mathbf{x}_1) + (1 - \lambda)(\mathbf{a} \cdot \mathbf{x}_2) + \lambda b + (1 - \lambda)b \\ &= \lambda(\mathbf{a} \cdot \mathbf{x}_1 + b) + (1 - \lambda)(\mathbf{a} \cdot \mathbf{x}_2 + b) \\ &= \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2). \end{aligned}$$

- Como $f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2)$, f es convexa.

Ejemplo:

Consideremos la función lineal

$$f(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} + b = a_1x_1 + \cdots + a_nx_n + b$$

donde $\mathbf{a} = (a_1, \dots, a_n)$ es un vector constante y b es una constante. Probar que f es, a la vez, cóncava y convexa.

Solución: Tome dos puntos $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$ y $\lambda \in (0, 1)$ arbitrarios, y probemos la desigualdad de Jensen.

$$\begin{aligned} f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) &= \mathbf{a} \cdot (\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) + \overbrace{b}^{= \lambda b + (1 - \lambda)b} \\ &= \lambda(\mathbf{a} \cdot \mathbf{x}_1) + (1 - \lambda)(\mathbf{a} \cdot \mathbf{x}_2) + \lambda b + (1 - \lambda)b \\ &= \lambda(\mathbf{a} \cdot \mathbf{x}_1 + b) + (1 - \lambda)(\mathbf{a} \cdot \mathbf{x}_2 + b) \\ &= \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2). \end{aligned}$$

- Como $f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2)$, f es convexa.
- Como $f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \geq \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2)$, f es cóncava.

Ejemplo:

Consideremos la función lineal

$$f(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} + b = a_1x_1 + \cdots + a_nx_n + b$$

donde $\mathbf{a} = (a_1, \dots, a_n)$ es un vector constante y b es una constante. Probar que f es, a la vez, cóncava y convexa.

Solución: Tome dos puntos $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^2$ y $\lambda \in (0, 1)$ arbitrarios, y probemos la desigualdad de Jensen.

$$\begin{aligned} f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) &= \mathbf{a} \cdot (\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) + \overbrace{b}^{\lambda b + (1 - \lambda)b} \\ &= \lambda(\mathbf{a} \cdot \mathbf{x}_1) + (1 - \lambda)(\mathbf{a} \cdot \mathbf{x}_2) + \lambda b + (1 - \lambda)b \\ &= \lambda(\mathbf{a} \cdot \mathbf{x}_1 + b) + (1 - \lambda)(\mathbf{a} \cdot \mathbf{x}_2 + b) \\ &= \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2). \end{aligned}$$

- Como $f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2)$, f es convexa.
- Como $f(\lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2) \geq \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2)$, f es cóncava.
- Note que la desigualdad estricta **no** se cumple, por lo tanto esta función **no** es estrictamente convexa ni estrictamente cóncava.

Propiedades

Teorema

Sean f_1, f_2, g_1 y g_2 funciones definidas sobre un conjunto convexo $S \subset \mathbb{R}^n$. Asuma que f_i son convexas y g_i son cóncavas. Entonces:

- (a) Si $a, b \geq 0$, entonces $af_1 + bf_2$ es convexa y $ag_1 + bg_2$ es cóncava.
- (b) Si F es convexa y creciente, entonces $U(\mathbf{x}) = F(f_i(\mathbf{x}))$ es convexa.
- (c) Si G es cóncava y creciente, entonces $V(\mathbf{x}) = G(g_i(\mathbf{x}))$ es cóncava.
- (d) $H := \max\{f_1, f_2\}$ es convexa y $h := \min\{g_1, g_2\}$ es cóncava.

En (b) y (c) se puede omitir que sea creciente si f_i y g_i son lineales, resp.

Propiedades

Teorema

Sean f_1, f_2, g_1 y g_2 funciones definidas sobre un conjunto convexo $S \subset \mathbb{R}^n$. Asuma que f_i son convexas y g_i son cóncavas. Entonces:

- (a) Si $a, b \geq 0$, entonces $af_1 + bf_2$ es convexa y $ag_1 + bg_2$ es cóncava.
- (b) Si F es convexa y creciente, entonces $U(\mathbf{x}) = F(f_i(\mathbf{x}))$ es convexa.
- (c) Si G es cóncava y creciente, entonces $V(\mathbf{x}) = G(g_i(\mathbf{x}))$ es cóncava.
- (d) $H := \max\{f_1, f_2\}$ es convexa y $h := \min\{g_1, g_2\}$ es cóncava.

En (b) y (c) se puede omitir que sea creciente si f_i y g_i son lineales, resp.

Ejemplo:

Demuestre que $h(x, y, z) = (x + 2y + 3z)^2$ es convexa en $[0, \infty)^3$.

Propiedades

Teorema

Sean f_1, f_2, g_1 y g_2 funciones definidas sobre un conjunto convexo $S \subset \mathbb{R}^n$. Asuma que f_i son convexas y g_i son cóncavas. Entonces:

- (a) Si $a, b \geq 0$, entonces $af_1 + bf_2$ es convexa y $ag_1 + bg_2$ es cóncava.
- (b) Si F es convexa y creciente, entonces $U(\mathbf{x}) = F(f_i(\mathbf{x}))$ es convexa.
- (c) Si G es cóncava y creciente, entonces $V(\mathbf{x}) = G(g_i(\mathbf{x}))$ es cóncava.
- (d) $H := \max\{f_1, f_2\}$ es convexa y $h := \min\{g_1, g_2\}$ es cóncava.

En (b) y (c) se puede omitir que sea creciente si f_i y g_i son lineales, resp.

Ejemplo:

Demuestre que $h(x, y, z) = (x + 2y + 3z)^2$ es convexa en $[0, \infty)^3$.

Solución: Recuerde que $f(x, y, z) = x + 2y + 3z$ es cóncava y convexa. Además, como $F(t) = t^2$ es convexa y creciente en $[0, \infty)$, entonces por (b) se sabe que $h(x, y, z) = F(f(x, y, z))$ es convexa en $[0, \infty)^3$.

Convexidad-Concavidad Vs Hessiana

Matriz Hessiana

Sea f una función dos veces continuamente diferenciable. La *matriz Hessiana* de f es

$$H_f(\mathbf{x}) := J_{\nabla f}(\mathbf{x}) = \begin{pmatrix} \partial_{1,1}f(\mathbf{x}) & \partial_{1,2}f(\mathbf{x}) & \cdots & \partial_{1,n}f(\mathbf{x}) \\ \partial_{2,1}f(\mathbf{x}) & \partial_{2,2}f(\mathbf{x}) & \cdots & \partial_{2,n}f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{n,1}f(\mathbf{x}) & \partial_{n,2}f(\mathbf{x}) & \cdots & \partial_{n,n}f(\mathbf{x}) \end{pmatrix}$$

Convexidad-Concavidad Vs Hessiana

Matriz Hessiana

Sea f una función dos veces continuamente diferenciable. La *matriz Hessiana* de f es

$$H_f(\mathbf{x}) := J_{\nabla f}(\mathbf{x}) = \begin{pmatrix} \partial_{1,1}f(\mathbf{x}) & \partial_{1,2}f(\mathbf{x}) & \cdots & \partial_{1,n}f(\mathbf{x}) \\ \partial_{2,1}f(\mathbf{x}) & \partial_{2,2}f(\mathbf{x}) & \cdots & \partial_{2,n}f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{n,1}f(\mathbf{x}) & \partial_{n,2}f(\mathbf{x}) & \cdots & \partial_{n,n}f(\mathbf{x}) \end{pmatrix}$$

Teorema

Sea f una función dos veces continuamente diferenciable, definida en un conjunto abierto convexo S de R^n . Entonces

- (a) f es cóncava si y solo si $H_f(\mathbf{x})$ es semidefinida negativa $\forall \mathbf{x} \in S$.
- (b) f es convexa si y solo si $H_f(\mathbf{x})$ es semidefinida positiva $\forall \mathbf{x} \in S$.
- (c) f es estrictamente cóncava si $H_f(\mathbf{x})$ es definida negativa $\forall \mathbf{x} \in S$.
- (d) f es estrictamente convexa si $H_f(\mathbf{x})$ es definida positiva $\forall \mathbf{x} \in S$.

Condición Suficiente para Convexidad/Concavidad en $\mathbb{R}^n$

Condición de Segundo Orden para Convexidad/Concavidad en $\mathbb{R}^n$

Sea f una función dos veces continuamente diferenciable, definida en un conjunto abierto convexo S de $\mathbb{R}^n$. Entonces

- f es convexa si y solo si $\Delta_r(\mathbf{x}) \geq 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.
- f es cóncava si y solo si $(-1)^r \Delta_r(\mathbf{x}) \geq 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.
- f es estrict. convexa si $D_r(\mathbf{x}) > 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.
- f es estrict. cóncava si $(-1)^r D_r(\mathbf{x}) > 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.

Condición Suficiente para Convexidad/Concavidad en $\mathbb{R}^n$

Condición de Segundo Orden para Convexidad/Concavidad en $\mathbb{R}^n$

Sea f una función dos veces continuamente diferenciable, definida en un conjunto abierto convexo S de $\mathbb{R}^n$. Entonces

- f es convexa si y solo si $\Delta_r(\mathbf{x}) \geq 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.
- f es cóncava si y solo si $(-1)^r \Delta_r(\mathbf{x}) \geq 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.
- f es estrict. convexa si $D_r(\mathbf{x}) > 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.
- f es estrict. cóncava si $(-1)^r D_r(\mathbf{x}) > 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.

Condición de Segundo Orden para Convexidad/Concavidad en $\mathbb{R}^2$

Sea $f(x, y)$ una función dos veces continuamente diferenciable, definida en un conjunto abierto convexo S de $\mathbb{R}^2$. Entonces, para todo $\mathbf{x} \in S$,

- f es convexa (cóncava) $\Leftrightarrow \partial_{xx}f(x, y) \geq (\leq) 0$, $\partial_{yy}f(x, y) \geq (\leq) 0$, y

$$\det H_f(x, y) = \partial_{xx}f(x, y)\partial_{yy}f(x, y) - (\partial_{xy}f(x, y))^2 \geq 0.$$
- f es estrictamente convexa (cóncava) si $\partial_{xx}f(x, y) > 0 (< 0)$, y

$$\det H_f(x, y) = \partial_{xx}f(x, y)\partial_{yy}f(x, y) - (\partial_{xy}f(x, y))^2 > 0.$$

Ejemplo:

Sea $f(x, y) = -2x^2 + (2a + 4)xy - 2y^2 + 4ay$. Averiguar para qué valores del parámetro a la función es cóncava, convexa, o ninguna de las dos.

Ejemplo:

Sea $f(x, y) = -2x^2 + (2a + 4)xy - 2y^2 + 4ay$. Averiguar para qué valores del parámetro a la función es cóncava, convexa, o ninguna de las dos.

Solución: Procedemos a calcular la matriz Hessiana de f .

$$H_f(x, y) = \begin{pmatrix} -4 & (2a + 4) \\ (2a + 4) & -4 \end{pmatrix}$$

Ejemplo:

Sea $f(x, y) = -2x^2 + (2a + 4)xy - 2y^2 + 4ay$. Averiguar para qué valores del parámetro a la función es cóncava, convexa, o ninguna de las dos.

Solución: Procedemos a calcular la matriz Hessiana de f .

$$H_f(x, y) = \begin{pmatrix} -4 & (2a + 4) \\ (2a + 4) & -4 \end{pmatrix}$$

Entonces,

$$\begin{aligned} D_2(x, y) = \Delta_2(x, y) &= \det H_f(x, y) = (-4) \cdot (-4) - (2a + 4)^2 \\ &= -2a(2a + 8) \begin{cases} < 0 & \text{si } a \in (-\infty, -4) \cup (0, \infty) \\ \geq 0 & \text{si } a \in [-4, 0] \end{cases}. \end{aligned}$$

Ejemplo:

Sea $f(x, y) = -2x^2 + (2a + 4)xy - 2y^2 + 4ay$. Averiguar para qué valores del parámetro a la función es cóncava, convexa, o ninguna de las dos.

Solución: Procedemos a calcular la matriz Hessiana de f .

$$H_f(x, y) = \begin{pmatrix} -4 & (2a + 4) \\ (2a + 4) & -4 \end{pmatrix}$$

Entonces,

$$\begin{aligned} D_2(x, y) = \Delta_2(x, y) = \det H_f(x, y) &= (-4) \cdot (-4) - (2a + 4)^2 \\ &= -2a(2a + 8) \begin{cases} < 0 & \text{si } a \in (-\infty, -4) \cup (0, \infty) \\ \geq 0 & \text{si } a \in [-4, 0] \end{cases}. \end{aligned}$$

Resumen:

- $a \in (-\infty, -4) \cup (0, \infty)$: Sabemos que no es convexa ni cóncava.
- $a \in [-4, 0]$: Como $\partial_{xx}f(x, y) = \partial_{yy}f(x, y) = -4 \leq 0$, por la condición de segundo orden sabemos que f es cóncava .

Convexa Vs Estrictamente Convexa

Si los menores principales dominantes D_k de una matriz H son positivos, entonces todos los menores arbitrarios Δ_k también son positivos (Meyer, Carl. Matrix Analysis and Applied Linear Algebra, SIAM, 2001, pág 559). Por eso basta revisar que los menores principales dominantes de la matriz Hessiana sean positivos para que la función sea estrictamente convexa.

Convexa Vs Estrictamente Convexa

Si los menores principales dominantes D_k de una matriz H son positivos, entonces todos los menores arbitrarios Δ_k también son positivos (Meyer, Carl. Matrix Analysis and Applied Linear Algebra, SIAM, 2001, pág 559). Por eso basta revisar que los menores principales dominantes de la matriz Hessiana sean positivos para que la función sea estrictamente convexa.

No funciona si la función no es estrictamente convexa.

Convexa Vs Estrictamente Convexa

Si los menores principales dominantes D_k de una matriz H son positivos, entonces todos los menores arbitrarios Δ_k también son positivos (Meyer, Carl. Matrix Analysis and Applied Linear Algebra, SIAM, 2001, pág 559). Por eso basta revisar que los menores principales dominantes de la matriz Hessiana sean positivos para que la función sea estrictamente convexa.

No funciona si la función no es estrictamente convexa.

Ejemplo: Determinar si $f(x_1, x_2, x_3) = 3x_1^2 - 2x_1x_2 - 4x_1x_3 + x_2^2 + 2x_3^2$ es convexa/cóncava, estrictamente convexa/cóncava o ninguna.

Convexa Vs Estrictamente Convexa

Si los menores principales dominantes D_k de una matriz H son positivos, entonces todos los menores arbitrarios Δ_k también son positivos (Meyer, Carl. Matrix Analysis and Applied Linear Algebra, SIAM, 2001, pág 559). Por eso basta revisar que los menores principales dominantes de la matriz Hessiana sean positivos para que la función sea estrictamente convexa.

No funciona si la función no es estrictamente convexa.

Ejemplo: Determinar si $f(x_1, x_2, x_3) = 3x_1^2 - 2x_1x_2 - 4x_1x_3 + x_2^2 + 2x_3^2$ es convexa/cóncava, estrictamente convexa/cóncava o ninguna.

Solución: Note que

$$H_f(\mathbf{x}) = \begin{pmatrix} 6 & -2 & -4 \\ -2 & 2 & 0 \\ -4 & 0 & 4 \end{pmatrix}$$

Convexa Vs Estrictamente Convexa

Si los menores principales dominantes D_k de una matriz H son positivos, entonces todos los menores arbitrarios Δ_k también son positivos (Meyer, Carl. Matrix Analysis and Applied Linear Algebra, SIAM, 2001, pág 559). Por eso basta revisar que los menores principales dominantes de la matriz Hessiana sean positivos para que la función sea estrictamente convexa.

No funciona si la función no es estrictamente convexa.

Ejemplo: Determinar si $f(x_1, x_2, x_3) = 3x_1^2 - 2x_1x_2 - 4x_1x_3 + x_2^2 + 2x_3^2$ es convexa/cóncava, estrictamente convexa/cóncava o ninguna.

Solución: Note que

$$H_f(\mathbf{x}) = \begin{pmatrix} 6 & -2 & -4 \\ -2 & 2 & 0 \\ -4 & 0 & 4 \end{pmatrix}$$

Vía M.P.D.: $D_1 = 6 > 0$; $D_2 = 8 > 0$; y $D_3 = 0$. No se puede concluir.

Convexa Vs Estrictamente Convexa

Si los menores principales dominantes D_k de una matriz H son positivos, entonces todos los menores arbitrarios Δ_k también son positivos (Meyer, Carl. Matrix Analysis and Applied Linear Algebra, SIAM, 2001, pág 559). Por eso basta revisar que los menores principales dominantes de la matriz Hessiana sean positivos para que la función sea estrictamente convexa.

No funciona si la función no es estrictamente convexa.

Ejemplo: Determinar si $f(x_1, x_2, x_3) = 3x_1^2 - 2x_1x_2 - 4x_1x_3 + x_2^2 + 2x_3^2$ es convexa/cóncava, estrictamente convexa/cóncava o ninguna.

Solución: Note que

$$H_f(\mathbf{x}) = \begin{pmatrix} 6 & -2 & -4 \\ -2 & 2 & 0 \\ -4 & 0 & 4 \end{pmatrix}$$

Vía M.P.D.: $D_1 = 6 > 0$; $D_2 = 8 > 0$; y $D_3 = 0$. No se puede concluir.

Vía M.P.A.: $\Delta_1 = \{6, 2, 4\} > 0$; $\Delta_2 = \{8, 8, 8\} > 0$; y $\Delta_3 = 0$. Entonces $H_f(\mathbf{x})$ es semidefinida positiva, lo que a su vez implica que f es convexa.

PRECAUCIÓN

f puede ser estrictamente convexa sin que H_f sea definida positiva, por eso la última parte del teorema dice “si” en lugar de “si y solo si”.

PRECAUCIÓN

f puede ser estrictamente convexa sin que H_f sea definida positiva, por eso la última parte del teorema dice “sí” en lugar de “sí y solo sí”.

Ejemplo: Tome $f(x, y) = x^4 + y^4$.

- f es Estrictamente Convexa: t^4 es estrictamente convexa en $\mathbb{R}$, ergo $(\lambda t_1 + (1 - \lambda)t_2)^4 < \lambda t_1^4 + (1 - \lambda)t_2^4$. Entonces,

$$\begin{aligned} f(\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2)) &= (\lambda x_1 + (1 - \lambda)x_2)^4 + (\lambda y_1 + (1 - \lambda)y_2)^4 \\ &< (\lambda x_1^4 + (1 - \lambda)x_2^4) + (\lambda y_1^4 + (1 - \lambda)y_2^4) \\ &= \lambda(x_1^4 + y_1^4) + (1 - \lambda)(x_2^4 + y_2^4) \\ &= \lambda f(x_1, y_1) + (1 - \lambda)f(x_2, y_2). \end{aligned}$$

PRECAUCIÓN

f puede ser estrictamente convexa sin que H_f sea definida positiva, por eso la última parte del teorema dice “si” en lugar de “si y solo si”.

Ejemplo: Tome $f(x, y) = x^4 + y^4$.

- f es Estrictamente Convexa: t^4 es estrictamente convexa en $\mathbb{R}$, ergo $(\lambda t_1 + (1 - \lambda)t_2)^4 < \lambda t_1^4 + (1 - \lambda)t_2^4$. Entonces,

$$\begin{aligned} f(\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2)) &= (\lambda x_1 + (1 - \lambda)x_2)^4 + (\lambda y_1 + (1 - \lambda)y_2)^4 \\ &< (\lambda x_1^4 + (1 - \lambda)x_2^4) + (\lambda y_1^4 + (1 - \lambda)y_2^4) \\ &= \lambda(x_1^4 + y_1^4) + (1 - \lambda)(x_2^4 + y_2^4) \\ &= \lambda f(x_1, y_1) + (1 - \lambda)f(x_2, y_2). \end{aligned}$$

- H_f No es Definida Positiva:

$$H_f(x, y) = \begin{pmatrix} 12x^2 & 0 \\ 0 & 12y^2 \end{pmatrix}$$

Entonces $H(0, 0)$ es la matrix nula 2×2 , por tanto

$\Delta_2(0, 0) = D_2(0, 0) = 0$ y $H(0, 0)$ **no** es definida positiva.

Funciones Cuasiconvexas y Cuasicóncavas

Definición

Sea f una función definida sobre un conjunto convexo S en $\mathbb{R}^n$ con valores en $\mathbb{R}$. Se dice que f es cuasiconvexa (resp. cuasicóncava) si para todo $\mathbf{x}^*, \mathbf{x} \in S$ y $\lambda \in [0, 1]$ se tiene

$$f(\mathbf{x}) \leq (\geq) f(\mathbf{x}^*) \implies f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{x}^*) \leq (\geq) f(\mathbf{x}^*).$$

Se dice que es estrictamente cuasiconvexa (cuasicóncava) si $\lambda \in (0, 1)$,

$$f(\mathbf{x}) \leq (\geq) f(\mathbf{x}^*), \mathbf{x} \neq \mathbf{x}^* \implies f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{x}^*) < (>) f(\mathbf{x}^*).$$

Funciones Cuasiconvexas y Cuasicóncavas

Definición

Sea f una función definida sobre un conjunto convexo S en $\mathbb{R}^n$ con valores en $\mathbb{R}$. Se dice que f es cuasiconvexa (resp. cuasicóncava) si para todo $\mathbf{x}^*, \mathbf{x} \in S$ y $\lambda \in [0, 1]$ se tiene

$$f(\mathbf{x}) \leq (\geq) f(\mathbf{x}^*) \implies f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{x}^*) \leq (\geq) f(\mathbf{x}^*).$$

Se dice que es estrictamente cuasiconvexa (cuasicóncava) si $\lambda \in (0, 1)$,

$$f(\mathbf{x}) \leq (\geq) f(\mathbf{x}^*), \mathbf{x} \neq \mathbf{x}^* \implies f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{x}^*) < (>) f(\mathbf{x}^*).$$

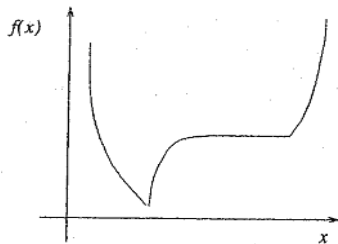
Definición Alternativa

Sea f una función definida sobre un conjunto convexo S en $\mathbb{R}^n$ con valores en $\mathbb{R}$. Se dice que f es cuasiconvexa (resp. cuasicóncava) si el conjunto bajonivel (resp. sobrenivel), definido por

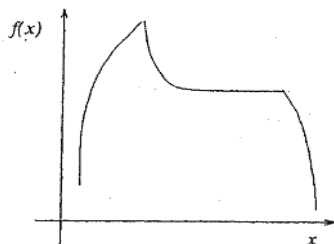
$$P_a = \{\mathbf{x} \in S \mid f(\mathbf{x}) \leq a\} \quad (\text{resp. } P_a = \{\mathbf{x} \in S \mid f(\mathbf{x}) \geq a\})$$

es convexo (como conjunto en $\mathbb{R}^n$) para todo número real a .

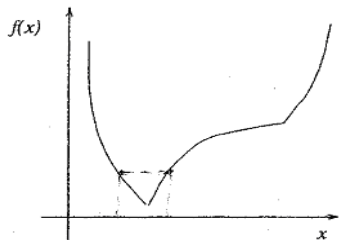
Ejemplos Gráficos



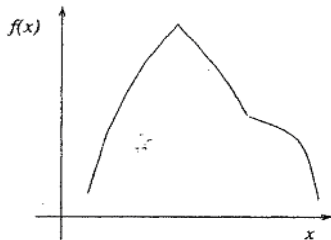
(a) Cuasiconvexa



(b) Cuasicóncava



(a) Estrictamente Cuasiconvexa



(b) Estrictamente Cuasicóncava

Observaciones

- (i) Las funciones cuasiconvexas y cuasicóncavas aparecen de forma natural en economía, en funciones de utilidad.

Observaciones

- (i) Las funciones cuasiconvexas y cuasicóncavas aparecen de forma natural en economía, en funciones de utilidad.
- (ii) Las funciones cuasiconvexas es un subconjunto más general que las convexas cuyas propiedades son de interés en optimización.

Observaciones

- (i) Las funciones cuasiconvexas y cuasicóncavas aparecen de forma natural en economía, en funciones de utilidad.
- (ii) Las funciones cuasiconvexas es un subconjunto más general que las convexas cuyas propiedades son de interés en optimización.
- (iii) Si f es cóncava, entonces f es cuasicóncava.

Observaciones

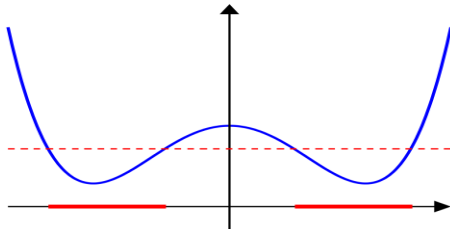
- (i) Las funciones cuasiconvexas y cuasicóncavas aparecen de forma natural en economía, en funciones de utilidad.
- (ii) Las funciones cuasiconvexas es un subconjunto más general que las convexas cuyas propiedades son de interés en optimización.
- (iii) Si f es cóncava, entonces f es cuasicóncava.
- (vi) Si f es convexa, entonces f es cuasiconvexa.

Observaciones

- (i) Las funciones cuasiconvexas y cuasicóncavas aparecen de forma natural en economía, en funciones de utilidad.
- (ii) Las funciones cuasiconvexas es un subconjunto más general que las convexas cuyas propiedades son de interés en optimización.
- (iii) Si f es cóncava, entonces f es cuasicóncava.
- (vi) Si f es convexa, entonces f es cuasiconvexa.
- (v) A la primera definición se le conoce como la versión para cuasicóncavas/cuasiconvexas de la desigualdad de Jensen.

Observaciones

- (i) Las funciones cuasiconvexas y cuasicóncavas aparecen de forma natural en economía, en funciones de utilidad.
- (ii) Las funciones cuasiconvexas es un subconjunto más general que las convexas cuyas propiedades son de interés en optimización.
- (iii) Si f es cóncava, entonces f es cuasicóncava.
- (vi) Si f es convexa, entonces f es cuasiconvexa.
- (v) A la primera definición se le conoce como la versión para cuasicóncavas/cuasiconvexas de la desigualdad de Jensen.
- (iv) Un ejemplo de función no cuasicóncava ni cuasiconvexa es la siguiente:



Propiedades

Teorema

Sea f una función definida sobre un conjunto convexo $S \subset \mathbb{R}^n$. Entonces:

- (a) Si f es cuasicóncava (cuasiconvexa) y F es estrictamente creciente, entonces $U(x) = F(f(x))$ es cuasicóncava (cuasiconvexa).
- (b) Si f es cuasicóncava (cuasiconvexa) y F es estrictamente decreciente, entonces $U(x) = F(f(x))$ es cuasiconvexa (cuasicóncava).

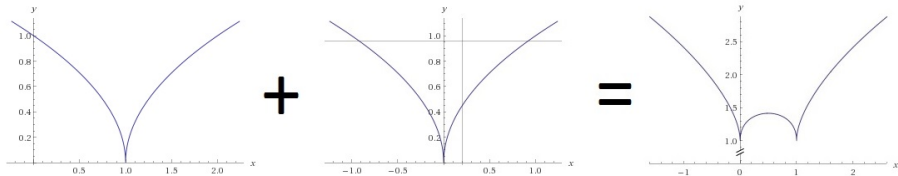
Propiedades

Teorema

Sea f una función definida sobre un conjunto convexo $S \subset \mathbb{R}^n$. Entonces:

- (a) Si f es cuasicóncava (cuasiconvexa) y F es estrictamente creciente, entonces $U(x) = F(f(x))$ es cuasicóncava (cuasiconvexa).
- (b) Si f es cuasicóncava (cuasiconvexa) y F es estrictamente decreciente, entonces $U(x) = F(f(x))$ es cuasiconvexa (cuasicóncava).

NOTA: La suma de funciones cuasicóncavas (cuasiconvexas) **no** es necesariamente cuasicóncavas (cuasiconvexas). Por ejemplo, considere las funciones cuasiconvexas $f_1(x) = \sqrt{|x-1|}$ y $f_2(x) = \sqrt{|x|}$ definidas en $\mathbb{R}$. Note que $f(x) = f_1(x) + f_2(x)$ **no** es cuasiconvexa.



Hessiana Orlada y Menores Principales

Matriz Hessiana Orlada

Sea f una función dos veces continuamente diferenciable. La *matriz Hessiana orlada* de f es

$$H_f^O(\mathbf{x}) := \begin{pmatrix} 0 & (\nabla f(\mathbf{x}))^T \\ \nabla f(\mathbf{x}) & H_f(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} 0 & \partial_1 f(\mathbf{x}) & \cdots & \partial_n f(\mathbf{x}) \\ \partial_1 f(\mathbf{x}) & \partial_{1,1} f(\mathbf{x}) & \cdots & \partial_{1,n} f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_n f(\mathbf{x}) & \partial_{n,1} f(\mathbf{x}) & \cdots & \partial_{n,n} f(\mathbf{x}) \end{pmatrix}$$

Hessiana Orlada y Menores Principales

Matriz Hessiana Orlada

Sea f una función dos veces continuamente diferenciable. La *matriz Hessiana orlada* de f es

$$H_f^O(\mathbf{x}) := \begin{pmatrix} 0 & (\nabla f(\mathbf{x}))^T \\ \nabla f(\mathbf{x}) & H_f(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} 0 & \partial_1 f(\mathbf{x}) & \cdots & \partial_n f(\mathbf{x}) \\ \partial_1 f(\mathbf{x}) & \partial_{1,1} f(\mathbf{x}) & \cdots & \partial_{1,n} f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_n f(\mathbf{x}) & \partial_{n,1} f(\mathbf{x}) & \cdots & \partial_{n,n} f(\mathbf{x}) \end{pmatrix}$$

Como $\mathbf{z}^T H_f^O(\mathbf{x}) \mathbf{z} = 0$, con $\mathbf{z}^T = (1, 0, \dots, 0) \neq \mathbf{0}$, $H_f^O(\mathbf{x})$ es indefinida.

Hessiana Orlada y Menores Principales

Matriz Hessiana Orlada

Sea f una función dos veces continuamente diferenciable. La *matriz Hessiana orlada* de f es

$$H_f^O(\mathbf{x}) := \begin{pmatrix} 0 & (\nabla f(\mathbf{x}))^T \\ \nabla f(\mathbf{x}) & H_f(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} 0 & \partial_1 f(\mathbf{x}) & \cdots & \partial_n f(\mathbf{x}) \\ \partial_1 f(\mathbf{x}) & \partial_{1,1} f(\mathbf{x}) & \cdots & \partial_{1,n} f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_n f(\mathbf{x}) & \partial_{n,1} f(\mathbf{x}) & \cdots & \partial_{n,n} f(\mathbf{x}) \end{pmatrix}$$

Como $\mathbf{z}^T H_f^O(\mathbf{x}) \mathbf{z} = 0$, con $\mathbf{z}^T = (1, 0, \dots, 0) \neq \mathbf{0}$, $H_f^O(\mathbf{x})$ es indefinida.

Menores Principales Dominantes de H_f^O

Sea $H_f^O(\mathbf{x})$ la matriz $(n+1) \times (n+1)$ Hessiana orlada. Se llaman *menores principales dominantes* de orden r (y se denotan D_r^O) a los determinantes:

$$D_r^O(\mathbf{x}) = \begin{vmatrix} 0 & \partial_1 f(\mathbf{x}) & \partial_2 f(\mathbf{x}) & \cdots & \partial_r f(\mathbf{x}) \\ \partial_1 f(\mathbf{x}) & \partial_{1,1} f(\mathbf{x}) & \partial_{1,2} f(\mathbf{x}) & \cdots & \partial_{1,r} f(\mathbf{x}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \partial_r f(\mathbf{x}) & \partial_{r,1} f(\mathbf{x}) & \partial_{r,2} f(\mathbf{x}) & \cdots & \partial_{r,r} f(\mathbf{x}) \end{vmatrix}$$

Condición Suficiente para Cuasiconvexidad/Cuasiconcavidad

Condición de Segundo Orden para Cuasiconvexidad/Cuasiconcavidad en $\mathbb{R}^n$

Sea f una función dos veces continuamente diferenciable, definida en un conjunto abierto convexo S de $\mathbb{R}^n$. Entonces

- f es cuasiconvexa si $D_r^O(\mathbf{x}) > 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.
- f es cuasicóncava si $(-1)^r D_r^O(\mathbf{x}) > 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.

Condición Suficiente para Cuasiconvexidad/Cuasiconcavidad

Condición de Segundo Orden para Cuasiconvexidad/Cuasiconcavidad en $\mathbb{R}^n$

Sea f una función dos veces continuamente diferenciable, definida en un conjunto abierto convexo S de $\mathbb{R}^n$. Entonces

- f es cuasiconvexa si $D_r^O(\mathbf{x}) > 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.
- f es cuasicóncava si $(-1)^r D_r^O(\mathbf{x}) > 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.

Ejemplo: Verificar si $f(x, y) = ye^x$ es cuasicóncava. De no serlo, encuentre el dominio máximo en donde es cuasicóncava.

Condición Suficiente para Cuasiconvexidad/Cuasiconcavidad

Condición de Segundo Orden para Cuasiconvexidad/Cuasiconcavidad en $\mathbb{R}^n$

Sea f una función dos veces continuamente diferenciable, definida en un conjunto abierto convexo S de $\mathbb{R}^n$. Entonces

- f es cuasiconvexa si $D_r^O(\mathbf{x}) > 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.
- f es cuasicóncava si $(-1)^r D_r^O(\mathbf{x}) > 0$, $\forall \mathbf{x} \in S$ y $r = 1, \dots, n$.

Ejemplo: Verificar si $f(x, y) = ye^x$ es cuasicóncava. De no serlo, encuentre el dominio máximo en donde es cuasicóncava.

Solución: Observe que

$$H_f^O(x, y) = \begin{pmatrix} 0 & ye^x & e^x \\ ye^x & ye^x & e^x \\ e^x & e^x & 0 \end{pmatrix} \Rightarrow D_1^O(x, y) = \begin{vmatrix} 0 & ye^x \\ ye^x & ye^x \end{vmatrix} = -y^2 e^{2x}.$$

$(-1)^1 D_1^O(x, y) = y^2 e^{2x} > 0$ si $y \neq 0$, y $(-1)^2 D_2^O(x, y) = ye^{3x} > 0$ si $y > 0$. Concluimos que f es cuasicóncava en $\{(x, y) \in \mathbb{R}^2 / y > 0\}$.

Cobb-Douglas

Sea $f(\mathbf{x}) = Ax_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}$, con $A > 0$ en $\{\mathbf{x} \in \mathbb{R}^n / x_1, x_2, \dots, x_n > 0\}$. Muestre que

- f es cuasicóncava si $\alpha_i > 0$ para todo $i = 1, \dots, n$.
- f es estrictamente cóncava si se restringe a $\alpha_1 + \cdots + \alpha_n < 1$.
- Qué pasa con f si $\alpha_1 + \cdots + \alpha_n \geq 1$?

Cobb-Douglas

Sea $f(\mathbf{x}) = Ax_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}$, con $A > 0$ en $\{\mathbf{x} \in \mathbb{R}^n / x_1, x_2, \dots, x_n > 0\}$. Muestre que

- f es cuasicóncava si $\alpha_i > 0$ para todo $i = 1, \dots, n$.
- f es estrictamente cóncava si se restringe a $\alpha_1 + \cdots + \alpha_n < 1$.
- Qué pasa con f si $\alpha_1 + \cdots + \alpha_n \geq 1$?

Solución: Cuasiconcavidad

Note que $\ln x_i$ es una función cóncava en $\mathbb{R}^n$. Entonces $\alpha_i \ln x_i$ es cóncava si $\alpha_i > 0$.

Cobb-Douglas

Sea $f(\mathbf{x}) = Ax_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}$, con $A > 0$ en $\{\mathbf{x} \in \mathbb{R}^n \mid x_1, x_2, \dots, x_n > 0\}$. Muestre que

- f es cuasicóncava si $\alpha_i > 0$ para todo $i = 1, \dots, n$.
- f es estrictamente cóncava si se restringe a $\alpha_1 + \cdots + \alpha_n < 1$.
- Qué pasa con f si $\alpha_1 + \cdots + \alpha_n \geq 1$?

Solución: Cuasiconcavidad

Note que $\ln x_i$ es una función cóncava en $\mathbb{R}^n$. Entonces $\alpha_i \ln x_i$ es cóncava si $\alpha_i > 0$.

$g(\mathbf{x}) := \ln A + \alpha_1 \ln x_1 + \cdots + \alpha_n \ln x_n$ es suma de funciones cóncavas, por lo que g es una función cóncava.

Cobb-Douglas

Sea $f(\mathbf{x}) = Ax_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}$, con $A > 0$ en $\{\mathbf{x} \in \mathbb{R}^n \mid x_1, x_2, \dots, x_n > 0\}$. Muestre que

- f es cuasicóncava si $\alpha_i > 0$ para todo $i = 1, \dots, n$.
- f es estrictamente cóncava si se restringe a $\alpha_1 + \cdots + \alpha_n < 1$.
- Qué pasa con f si $\alpha_1 + \cdots + \alpha_n \geq 1$?

Solución: Cuasiconcavidad

Note que $\ln x_i$ es una función cóncava en $\mathbb{R}^n$. Entonces $\alpha_i \ln x_i$ es cóncava si $\alpha_i > 0$.

$g(\mathbf{x}) := \ln A + \alpha_1 \ln x_1 + \cdots + \alpha_n \ln x_n$ es suma de funciones cóncavas, por lo que g es una función cóncava.

Como toda función cóncava es cuasicóncava, concluimos que g es cuasicóncava.

Cobb-Douglas

Sea $f(\mathbf{x}) = Ax_1^{\alpha_1}x_2^{\alpha_2}\cdots x_n^{\alpha_n}$, con $A > 0$ en $\{\mathbf{x} \in \mathbb{R}^n \mid x_1, x_2, \dots, x_n > 0\}$. Muestre que

- f es cuasicóncava si $\alpha_i > 0$ para todo $i = 1, \dots, n$.
- f es estrictamente cóncava si se restringe a $\alpha_1 + \cdots + \alpha_n < 1$.
- Qué pasa con f si $\alpha_1 + \cdots + \alpha_n \geq 1$?

Solución: Cuasiconcavidad

Note que $\ln x_i$ es una función cóncava en $\mathbb{R}^n$. Entonces $\alpha_i \ln x_i$ es cóncava si $\alpha_i > 0$.

$g(\mathbf{x}) := \ln A + \alpha_1 \ln x_1 + \cdots + \alpha_n \ln x_n$ es suma de funciones cóncavas, por lo que g es una función cóncava.

Como toda función cóncava es cuasicóncava, concluimos que g es cuasicóncava.

Luego, como e^z es una función estrictamente creciente, concluimos que $f(\mathbf{x}) = e^{g(\mathbf{x})}$ es cuasicóncava.

Cobb-Douglas: Solución - Continuación

Utilicemos ahora un criterio de segundo orden. Note que

$$H_f^O(x, y) = \begin{pmatrix} 0 & \frac{\alpha_1}{x_1} f(\mathbf{x}) & \frac{\alpha_2}{x_2} f(\mathbf{x}) & \cdots & \frac{\alpha_n}{x_n} f(\mathbf{x}) \\ \frac{\alpha_1}{x_1} f(\mathbf{x}) & \frac{\alpha_1(\alpha_1-1)}{x_1^2} f(\mathbf{x}) & \frac{\alpha_1\alpha_2}{x_1x_2} f(\mathbf{x}) & \cdots & \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) \\ \frac{\alpha_2}{x_2} f(\mathbf{x}) & \frac{\alpha_1\alpha_2}{x_1x_2} f(\mathbf{x}) & \frac{\alpha_2(\alpha_2-1)}{x_2^2} f(\mathbf{x}) & \cdots & \frac{\alpha_2\alpha_n}{x_2x_n} f(\mathbf{x}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_n}{x_n} f(\mathbf{x}) & \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) & \frac{\alpha_2\alpha_n}{x_2x_n} f(\mathbf{x}) & \cdots & \frac{\alpha_n(\alpha_n-1)}{x_n^2} f(\mathbf{x}) \end{pmatrix}$$

Cobb-Douglas: Solución - Continuación

Utilicemos ahora un criterio de segundo orden. Note que

$$H_f^O(x, y) = \begin{pmatrix} 0 & \frac{\alpha_1}{x_1} f(\mathbf{x}) & \frac{\alpha_2}{x_2} f(\mathbf{x}) & \cdots & \frac{\alpha_n}{x_n} f(\mathbf{x}) \\ \frac{\alpha_1}{x_1} f(\mathbf{x}) & \frac{\alpha_1(\alpha_1-1)}{x_1^2} f(\mathbf{x}) & \frac{\alpha_1\alpha_2}{x_1x_2} f(\mathbf{x}) & \cdots & \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) \\ \frac{\alpha_2}{x_2} f(\mathbf{x}) & \frac{\alpha_1\alpha_2}{x_1x_2} f(\mathbf{x}) & \frac{\alpha_2(\alpha_2-1)}{x_2^2} f(\mathbf{x}) & \cdots & \frac{\alpha_2\alpha_n}{x_2x_n} f(\mathbf{x}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_n}{x_n} f(\mathbf{x}) & \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) & \frac{\alpha_2\alpha_n}{x_2x_n} f(\mathbf{x}) & \cdots & \frac{\alpha_n(\alpha_n-1)}{x_n^2} f(\mathbf{x}) \end{pmatrix}$$

De acuerdo a esto, y aplicando reducción de matrices, obtenemos que

$$(-1)^r D_r^O = \frac{\alpha_1\alpha_2\cdots\alpha_r}{(x_1x_2\cdots x_r)^2} \left(\sum_{i=1}^r \alpha_i \right) f(\mathbf{x})^{r+1} > 0 \iff \alpha_i > 0, \forall i.$$

Cobb-Douglas: Solución - Continuación

Utilicemos ahora un criterio de segundo orden. Note que

$$H_f^O(x, y) = \begin{pmatrix} 0 & \frac{\alpha_1}{x_1} f(\mathbf{x}) & \frac{\alpha_2}{x_2} f(\mathbf{x}) & \cdots & \frac{\alpha_n}{x_n} f(\mathbf{x}) \\ \frac{\alpha_1}{x_1} f(\mathbf{x}) & \frac{\alpha_1(\alpha_1-1)}{x_1^2} f(\mathbf{x}) & \frac{\alpha_1\alpha_2}{x_1x_2} f(\mathbf{x}) & \cdots & \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) \\ \frac{\alpha_2}{x_2} f(\mathbf{x}) & \frac{\alpha_1\alpha_2}{x_1x_2} f(\mathbf{x}) & \frac{\alpha_2(\alpha_2-1)}{x_2^2} f(\mathbf{x}) & \cdots & \frac{\alpha_2\alpha_n}{x_2x_n} f(\mathbf{x}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_n}{x_n} f(\mathbf{x}) & \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) & \frac{\alpha_2\alpha_n}{x_2x_n} f(\mathbf{x}) & \cdots & \frac{\alpha_n(\alpha_n-1)}{x_n^2} f(\mathbf{x}) \end{pmatrix}$$

De acuerdo a esto, y aplicando reducción de matrices, obtenemos que

$$(-1)^r D_r^O = \frac{\alpha_1\alpha_2\cdots\alpha_r}{(x_1x_2\cdots x_r)^2} \left(\sum_{i=1}^r \alpha_i \right) f(\mathbf{x})^{r+1} > 0 \iff \alpha_i > 0, \forall i.$$

Por las condiciones de segundo orden, concluimos que si $\alpha_i > 0$ para todo $i = 1, \dots, n$, entonces $f(\mathbf{x})$ es cuasicóncava.

Cobb-Douglas: Solución - Continuación

Note que no se puede utilizar el mismo criterio de composición con función creciente para probar que f es cóncava, ya que e^z **no** es cóncava.

Cobb-Douglas: Solución - Continuación

Note que no se puede utilizar el mismo criterio de composición con función creciente para probar que f es cóncava, ya que e^z **no** es cóncava.

Por otra parte,

$$H_f(x, y) = \begin{pmatrix} \frac{\alpha_1(\alpha_1-1)}{x_1^2} f(\mathbf{x}) & \frac{\alpha_1\alpha_2}{x_1x_2} f(\mathbf{x}) & \cdots & \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) & \frac{\alpha_2\alpha_n}{x_2x_n} f(\mathbf{x}) & \cdots & \frac{\alpha_n(\alpha_n-1)}{x_n^2} f(\mathbf{x}) \end{pmatrix}$$

Cobb-Douglas: Solución - Continuación

Note que no se puede utilizar el mismo criterio de composición con función creciente para probar que f es cóncava, ya que e^z **no** es cóncava.

Por otra parte,

$$H_f(x, y) = \begin{pmatrix} \frac{\alpha_1(\alpha_1-1)}{x_1^2} f(\mathbf{x}) & \frac{\alpha_1\alpha_2}{x_1x_2} f(\mathbf{x}) & \cdots & \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) & \frac{\alpha_2\alpha_n}{x_2x_n} f(\mathbf{x}) & \cdots & \frac{\alpha_n(\alpha_n-1)}{x_n^2} f(\mathbf{x}) \end{pmatrix}$$

De acuerdo a esto, y aplicando reducción de matrices, obtenemos que

$$(-1)^r D_r = \frac{\alpha_1\alpha_2\cdots\alpha_r}{(x_1x_2\cdots x_r)^2} \left(1 - \sum_{i=1}^r \alpha_i\right) f(\mathbf{x})^r > 0 \iff \sum_{i=1}^r \alpha_i < 1.$$

Concluimos que $f(\mathbf{x})$ es estrictamente cóncava si $\sum_{i=1}^n \alpha_i < 1$.

Cobb-Douglas: Solución - Continuación

Note que no se puede utilizar el mismo criterio de composición con función creciente para probar que f es cóncava, ya que e^z **no** es cóncava.

Por otra parte,

$$H_f(x, y) = \begin{pmatrix} \frac{\alpha_1(\alpha_1-1)}{x_1^2} f(\mathbf{x}) & \frac{\alpha_1\alpha_2}{x_1x_2} f(\mathbf{x}) & \cdots & \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) & \frac{\alpha_2\alpha_n}{x_2x_n} f(\mathbf{x}) & \cdots & \frac{\alpha_n(\alpha_n-1)}{x_n^2} f(\mathbf{x}) \end{pmatrix}$$

De acuerdo a esto, y aplicando reducción de matrices, obtenemos que

$$(-1)^r D_r = \frac{\alpha_1\alpha_2\cdots\alpha_r}{(x_1x_2\cdots x_r)^2} \left(1 - \sum_{i=1}^r \alpha_i\right) f(\mathbf{x})^r > 0 \iff \sum_{i=1}^r \alpha_i < 1.$$

Concluimos que $f(\mathbf{x})$ es estrictamente cóncava si $\sum_{i=1}^n \alpha_i < 1$.

Si $\sum_{i=1}^n \alpha_i = 1$, el criterio no es concluyente. Como f es homogénea de grado 1 y cuasicóncava, es posible probar que f debe ser cóncava.

Cobb-Douglas: Solución - Continuación

Note que no se puede utilizar el mismo criterio de composición con función creciente para probar que f es cóncava, ya que e^z **no** es cóncava.

Por otra parte,

$$H_f(x, y) = \begin{pmatrix} \frac{\alpha_1(\alpha_1-1)}{x_1^2} f(\mathbf{x}) & \frac{\alpha_1\alpha_2}{x_1x_2} f(\mathbf{x}) & \cdots & \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_1\alpha_n}{x_1x_n} f(\mathbf{x}) & \frac{\alpha_2\alpha_n}{x_2x_n} f(\mathbf{x}) & \cdots & \frac{\alpha_n(\alpha_n-1)}{x_n^2} f(\mathbf{x}) \end{pmatrix}$$

De acuerdo a esto, y aplicando reducción de matrices, obtenemos que

$$(-1)^r D_r = \frac{\alpha_1\alpha_2\cdots\alpha_r}{(x_1x_2\cdots x_r)^2} \left(1 - \sum_{i=1}^r \alpha_i\right) f(\mathbf{x})^r > 0 \iff \sum_{i=1}^r \alpha_i < 1.$$

Concluimos que $f(\mathbf{x})$ es estrictamente cóncava si $\sum_{i=1}^n \alpha_i < 1$.

Si $\sum_{i=1}^n \alpha_i = 1$, el criterio no es concluyente. Como f es homogénea de grado 1 y cuasicóncava, es posible probar que f debe ser cóncava.

Finalmente, si $\sum_{i=1}^n \alpha_i > 1$, tenemos que $f(x, x, \dots, x) = Ax^{\sum_{i=1}^n \alpha_i}$ (para $x_1 = \dots = x_n = x$) es convexa (en $\mathbb{R}$), i.e., no es cóncava.

Outline

1 Preliminares Matemáticos

- Topología en $\mathbb{R}^n$
- Teorema de la Función Implícita/Inversa
- Matrices Semidefinidas y Definidas
- Convexidad-Concavidad
- Cuasiconvexidad-Cuasiconcavidad

2 Optimización

- Motivación
- Definiciones
- Optimización Sin Restricciones

3 Optimización Con Restricciones

- Restricciones de Igualdad: Lagrange
- Restricciones de Igualdad/Desigualdad: Karush-Kuhn-Tucker

¿Por Qué Optimización en Varias Variables?

Una enorme mayoría de problemas en Economía y Estadística consisten en encontrar el máximo o el mínimo de una función que depende de *varias* entradas.

¿Por Qué Optimización en Varias Variables?

Una enorme mayoría de problemas en Economía y Estadística consisten en encontrar el máximo o el mínimo de una función que depende de *varias* entradas.

- **Minimizar Riesgo:** Escogencia del portafolio de inversión de una compañía.

¿Por Qué Optimización en Varias Variables?

Una enorme mayoría de problemas en Economía y Estadística consisten en encontrar el máximo o el mínimo de una función que depende de *varias* entradas.

- **Minimizar Riesgo:** Escogencia del portafolio de inversión de una compañía.
- **Maximizar Utilidad:** Procesos de decisión de consumidores con múltiples alternativas.

¿Por Qué Optimización en Varias Variables?

Una enorme mayoría de problemas en Economía y Estadística consisten en encontrar el máximo o el mínimo de una función que depende de *varias* entradas.

- **Minimizar Riesgo:** Escogencia del portafolio de inversión de una compañía.
- **Maximizar Utilidad:** Procesos de decisión de consumidores con múltiples alternativas.
- **Minimizar Costos:** Diseño de presupuestos de empresas.

¿Por Qué Optimización en Varias Variables?

Una enorme mayoría de problemas en Economía y Estadística consisten en encontrar el máximo o el mínimo de una función que depende de *varias* entradas.

- **Minimizar Riesgo:** Escogencia del portafolio de inversión de una compañía.
- **Maximizar Utilidad:** Procesos de decisión de consumidores con múltiples alternativas.
- **Minimizar Costos:** Diseño de presupuestos de empresas.
- **Estimación de Parámetros:** Maximización de probabilidad de ocurrencia.

¿Por Qué Optimización en Varias Variables?

Una enorme mayoría de problemas en Economía y Estadística consisten en encontrar el máximo o el mínimo de una función que depende de *varias* entradas.

- **Minimizar Riesgo:** Escogencia del portafolio de inversión de una compañía.
- **Maximizar Utilidad:** Procesos de decisión de consumidores con múltiples alternativas.
- **Minimizar Costos:** Diseño de presupuestos de empresas.
- **Estimación de Parámetros:** Maximización de probabilidad de ocurrencia.
- **Regresión Lineal y No-Lineal:** Encontrar la función de mejor ajuste a través de la minimización de una función de “error”.

¿Por Qué Optimización Con Restricciones?

El Economista y el Estadístico se encuentran a menudo con problemas donde se debe respetar restricciones.

¿Por Qué Optimización Con Restricciones?

El Economista y el Estadístico se encuentran a menudo con problemas donde se debe respetar restricciones.

- Precios y unidades de bienes no pueden tomar valores negativos.

¿Por Qué Optimización Con Restricciones?

El Economista y el Estadístico se encuentran a menudo con problemas donde se debe respetar restricciones.

- Precios y unidades de bienes no pueden tomar valores negativos.
- Recursos limitados.

¿Por Qué Optimización Con Restricciones?

El Economista y el Estadístico se encuentran a menudo con problemas donde se debe respetar restricciones.

- Precios y unidades de bienes no pueden tomar valores negativos.
- Recursos limitados.
- Restricciones presupuestarias.

¿Por Qué Optimización Con Restricciones?

El Economista y el Estadístico se encuentran a menudo con problemas donde se debe respetar restricciones.

- Precios y unidades de bienes no pueden tomar valores negativos.
- Recursos limitados.
- Restricciones presupuestarias.
- Reglas de mercado impuestas por leyes o tratados.

¿Por Qué Optimización Con Restricciones?

El Economista y el Estadístico se encuentran a menudo con problemas donde se debe respetar restricciones.

- Precios y unidades de bienes no pueden tomar valores negativos.
- Recursos limitados.
- Restricciones presupuestarias.
- Reglas de mercado impuestas por leyes o tratados.
- Densidades no pueden ser negativas (método de máxima verosimilitud).

¿Por Qué Optimización Con Restricciones?

El Economista y el Estadístico se encuentran a menudo con problemas donde se debe respetar restricciones.

- Precios y unidades de bienes no pueden tomar valores negativos.
- Recursos limitados.
- Restricciones presupuestarias.
- Reglas de mercado impuestas por leyes o tratados.
- Densidades no pueden ser negativas (método de máxima verosimilitud).
- Regresión bajo hipótesis nula.

Definiciones de Óptimos

Óptimo Global

Sea f una función con valores reales y cuyo dominio es un subconjunto $S \subset \mathbb{R}^n$. Se dice que $\mathbf{c} = (c_1, c_2, \dots, c_n)$ es un “óptimo global” si f alcanza un máximo/mínimo global en $\mathbf{c} \in S$, es decir,

$$f(\mathbf{x}) \leq f(\mathbf{c}), \quad \text{para todo } \mathbf{x} \in S \quad \text{Máximo Global}$$

$$f(\mathbf{x}) \geq f(\mathbf{c}), \quad \text{para todo } \mathbf{x} \in S \quad \text{Mínimo Global}$$

Definiciones de Óptimos

Óptimo Global

Sea f una función con valores reales y cuyo dominio es un subconjunto $S \subset \mathbb{R}^n$. Se dice que $\mathbf{c} = (c_1, c_2, \dots, c_n)$ es un “óptimo global” si f alcanza un máximo/mínimo global en $\mathbf{c} \in S$, es decir,

$$f(\mathbf{x}) \leq f(\mathbf{c}), \quad \text{para todo } \mathbf{x} \in S \quad \text{Máximo Global}$$

$$f(\mathbf{x}) \geq f(\mathbf{c}), \quad \text{para todo } \mathbf{x} \in S \quad \text{Mínimo Global}$$

Óptimo Local

Sea $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Se dice que $\mathbf{c} = (c_1, c_2, \dots, c_n)$ es un “óptimo local” si f alcanza un máximo/mínimo local en $\mathbf{c} \in S$, es decir, si existe una bola $B_r(\mathbf{c})$, con r positivo tal que

$$f(\mathbf{x}) \leq f(\mathbf{c}), \quad \text{para todo } \mathbf{x} \in S \cap B_r(\mathbf{c}) \quad \text{Máximo Local}$$

$$f(\mathbf{x}) \geq f(\mathbf{c}), \quad \text{para todo } \mathbf{x} \in S \cap B_r(\mathbf{c}) \quad \text{Mínimo Local.}$$

(basta que la desigualdad se cumpla solo para valores $\mathbf{x}$ cercanos a $\mathbf{c}$.)

Propiedades

Teorema

- **Principio de Dualidad:** El punto c es un máximo de f si y solo si el punto c es un mínimo de $-f$.
- Maximizar (minimizar) una función equivale a maximizar (minimizar) una transformación estrictamente creciente de esa función.

Propiedades

Teorema

- **Principio de Dualidad:** El punto c es un máximo de f si y solo si el punto c es un mínimo de $-f$.
- Maximizar (minimizar) una función equivale a maximizar (minimizar) una transformación estrictamente creciente de esa función.

Ejemplo: Maximizar

$$f(x, y, z) = e^{-x^2 - (y+1)^2 - (z-2)^4 - 100}$$

es equivalente a maximizar

$$g(x, y, z) = \ln f(x, y, z) + 100 = -x^2 - (y+1)^2 - (z-2)^4,$$

pues la función $h(t) = \ln(t) + 100$ es una función estrictamente creciente.

Propiedades

Teorema

- **Principio de Dualidad:** El punto c es un máximo de f si y solo si el punto c es un mínimo de $-f$.
- Maximizar (minimizar) una función equivale a maximizar (minimizar) una transformación estrictamente creciente de esa función.

Ejemplo: Maximizar

$$f(x, y, z) = e^{-x^2 - (y+1)^2 - (z-2)^4 - 100}$$

es equivalente a maximizar

$$g(x, y, z) = \ln f(x, y, z) + 100 = -x^2 - (y+1)^2 - (z-2)^4,$$

pues la función $h(t) = \ln(t) + 100$ es una función estrictamente creciente.

Por el principio de dualidad, es equivalente a mínimizar

$$h(x, y, z) = -g(x, y, z) = x^2 + (y+1)^2 + (z-2)^4.$$

Ejemplo Gráfico de Optimización en Dos Variables

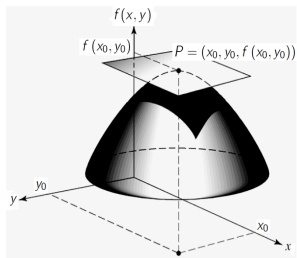


Figura: Función que alcanza un máximo en (x^*, y^*)

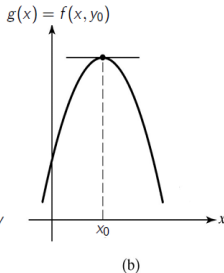
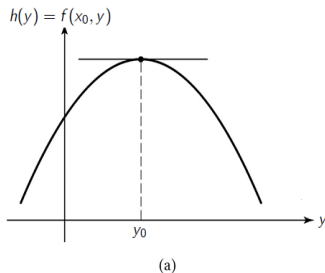


Figura: Cortes transversales de la función

Condición Necesaria para Máximo-Mínimo en $\mathbb{R}^n$

Si f es diferenciable, ocuparíamos que el plano tangencial sea completamente horizontal, es decir, se requiere que la derivada direccional ($D_{\mathbf{u}}f(\mathbf{x}^*) = \nabla f(\mathbf{x}^*) \cdot \mathbf{u}$) en cualquier dirección $\mathbf{u}$ sea 0 ($\Rightarrow \nabla f(\mathbf{x}^*) = \mathbf{0}$).

Condición Necesaria para Máximo-Mínimo en $\mathbb{R}^n$

Si f es diferenciable, ocuparíamos que el plano tangencial sea completamente horizontal, es decir, se requiere que la derivada direccional ($D_{\mathbf{u}}f(\mathbf{x}^*) = \nabla f(\mathbf{x}^*) \cdot \mathbf{u}$) en cualquier dirección $\mathbf{u}$ sea 0 ($\Rightarrow \nabla f(\mathbf{x}^*) = \mathbf{0}$).

Punto Estacionario

El punto $\mathbf{x}^*$ se dice “*punto estacionario*” de $f(\mathbf{x})$ si satisface $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Condición Necesaria para Máximo-Mínimo en $\mathbb{R}^n$

Si f es diferenciable, ocuparíamos que el plano tangencial sea completamente horizontal, es decir, se requiere que la derivada direccional ($D_{\mathbf{u}}f(\mathbf{x}^*) = \nabla f(\mathbf{x}^*) \cdot \mathbf{u}$) en cualquier dirección $\mathbf{u}$ sea 0 ($\Rightarrow \nabla f(\mathbf{x}^*) = \mathbf{0}$).

Punto Estacionario

El punto $\mathbf{x}^*$ se dice “*punto estacionario*” de $f(\mathbf{x})$ si satisface $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Condición de Primer Orden para Máximo-Mínimo en $\mathbb{R}^n$

Sea $f(\mathbf{x})$ una función diferenciable en $\mathbb{R}^n$. Si en el punto $\mathbf{x}^*$ la función f alcanza su máximo o mínimo, entonces necesariamente $\mathbf{x}^*$ debe ser un punto estacionario de f .

Condición Necesaria para Máximo-Mínimo en $\mathbb{R}^n$

Si f es diferenciable, ocuparíamos que el plano tangencial sea completamente horizontal, es decir, se requiere que la derivada direccional ($D_{\mathbf{u}}f(\mathbf{x}^*) = \nabla f(\mathbf{x}^*) \cdot \mathbf{u}$) en cualquier dirección $\mathbf{u}$ sea 0 ($\Rightarrow \nabla f(\mathbf{x}^*) = \mathbf{0}$).

Punto Estacionario

El punto $\mathbf{x}^*$ se dice “*punto estacionario*” de $f(\mathbf{x})$ si satisface $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Condición de Primer Orden para Máximo-Mínimo en $\mathbb{R}^n$

Sea $f(\mathbf{x})$ una función diferenciable en $\mathbb{R}^n$. Si en el punto $\mathbf{x}^*$ la función f alcanza su máximo o mínimo, entonces necesariamente $\mathbf{x}^*$ debe ser un punto estacionario de f .

Ejemplo: En el caso de $\mathbb{R}^2$, el punto (x^*, y^*) se dice “*punto estacionario*” de $f(x, y)$ si satisface

$$\begin{cases} \partial_x f(x^*, y^*) = 0 \\ \partial_y f(x^*, y^*) = 0 \end{cases} \quad . \quad (2)$$

Ejemplo: Maximización de Beneficios

Sea $Y = F(K, L)$ una función de producción, con

- K : Capital
- L : Trabajo ("Load")

Ejemplo: Maximización de Beneficios

Sea $Y = F(K, L)$ una función de producción, con

- K : Capital
- L : Trabajo ("Load")

Definamos

- p : Precio por unidad de output.
- r : Costo por unidad de capital.
- w : Precio (tasa salarial) por unidad de trabajo.

Ejemplo: Maximización de Beneficios

Sea $Y = F(K, L)$ una función de producción, con

- K : Capital
- L : Trabajo ("Load")

Definamos

- p : Precio por unidad de output.
- r : Costo por unidad de capital.
- w : Precio (tasa salarial) por unidad de trabajo.

Función Objetivo: Beneficio (aquella que interesa hacer óptima)

$$\pi(K, L) = pF(K, L) - rK - wL$$

Ejemplo: Maximización de Beneficios

Sea $Y = F(K, L)$ una función de producción, con

- K : Capital
- L : Trabajo ("Load")

Definamos

- p : Precio por unidad de output.
- r : Costo por unidad de capital.
- w : Precio (tasa salarial) por unidad de trabajo.

Función Objetivo: Beneficio (aquella que interesa hacer óptima)

$$\pi(K, L) = pF(K, L) - rK - wL$$

Deseamos encontrar los valores (K^*, L^*) que maximizan los beneficios.

Ejemplo: Solución

Condiciones de primer orden

$$p\partial_K F(K^*, L^*) = r \qquad p\partial_L F(K^*, L^*) = w$$

Ejemplo: Solución

Condiciones de primer orden

$$p\partial_K F(K^*, L^*) = r \qquad p\partial_L F(K^*, L^*) = w$$

Interpretación

- $\partial_K F(K^*, L^*)$: crecimiento de la producción al incrementar el capital.
- $p\partial_K F(K^*, L^*)$: aumento en los ingresos por una unidad extra de capital.
- r : aumento en el costo por un aumento unitario de capital.
- $\partial_L F(K^*, L^*)$: crecimiento de la producción al incrementar el trabajo.
- $p\partial_L F(K^*, L^*)$: aumento en los ingresos por una unidad extra de trabajo.
- w : aumento en el costo por un aumento unitario de trabajo.

Propiedad del óptimo: el ingreso extra generado por un aumento en una unidad de capital o trabajo se compensa con el costo generado por dicho aumento (costo marginal del factor = ingreso marginal del factor).

Máximo-Mínimo Global sobre Compactos

Teorema de los Valores Extremos (Weierstrass)

Sea f una función continua definida sobre un conjunto compacto $S \subset \mathbb{R}^n$. Entonces f alcanza su máximo y su mínimo globales en S .

Máximo-Mínimo Global sobre Compactos

Teorema de los Valores Extremos (Weierstrass)

Sea f una función continua definida sobre un conjunto compacto $S \subset \mathbb{R}^n$. Entonces f alcanza su máximo y su mínimo globales en $\overline{S}$.

Observaciones:

- El teorema garantiza la existencia de los puntos extremos.
- Los puntos extremos podrían ser tanto interiores como frontera.

Máximo-Mínimo Global sobre Compactos

Teorema de los Valores Extremos (Weierstrass)

Sea f una función **continua** definida sobre un conjunto **compacto** $S \subset \mathbb{R}^n$. Entonces f alcanza su máximo y su mínimo globales en $\overline{S}$.

Observaciones:

- El teorema garantiza la existencia de los puntos extremos.
- Los puntos extremos podrían ser tanto interiores como frontera.

Algoritmo de optimización de una función diferenciable f definida en un compacto S :

- Encuentre el conjunto de puntos estacionarios
$$\{(x, y) \in S \mid \nabla f(x, y) = 0\}.$$
- Hallar los puntos extremos en la frontera de S .
- Calcular los valores de f para todos los puntos hallados en los dos ítems anteriores.
- El mayor (menor) de los valores obtenidos es el valor máximo (mínimo) global de f en S .

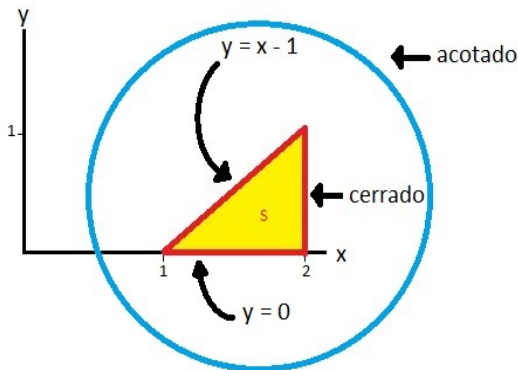
Ejemplo:

Sea $S := \{(x, y) \mid 1 \leq x \leq 2, 0 \leq y \leq x - 1\}$ y $h(x, y) = x^2 y(x - y - 1)$ definida sobre S . Pruebe que h tiene óptimos globales en S y hállelos.

Ejemplo:

Sea $S := \{(x, y) \mid 1 \leq x \leq 2, 0 \leq y \leq x - 1\}$ y $h(x, y) = x^2 y(x - y - 1)$ definida sobre S . Pruebe que h tiene óptimos globales en S y hállelos.

Solución: Note que S es un compacto como lo muestra la siguiente figura:



Por el teorema de Weierstrass, h alcanza el mínimo y el máximo en S .

Ejemplo:

Puntos Estacionarios: Como $\nabla h^T = (xy(3x - 2y - 2), x^2(x - 2y - 1))$ tenemos los siguientes puntos estacionarios:

- Si $x = 0$, $\nabla h = 0$. $(0, y)$, $y \in \mathbb{R}$, son puntos estacionarios.
- Si $y = 0$ entonces $x - 1 = 0$. $(1, 0)$ es un punto estacionario.
- Si $x \neq 0$ y $y \neq 0$, entonces

$$\begin{cases} 3x - 2y - 2 = 0 \\ x - 2y - 1 = 0 \end{cases} \Rightarrow \left(\begin{array}{cc|c} 3 & -2 & 2 \\ 1 & -2 & 1 \end{array} \right) \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{4} \end{pmatrix}$$

De todos estos puntos, el único que pertenece a S es $(1, 0)$, y $h(1, 0) = 0$.

Ejemplo:

Puntos Estacionarios: Como $\nabla h^T = (xy(3x - 2y - 2), x^2(x - 2y - 1))$ tenemos los siguientes puntos estacionarios:

- Si $x = 0$, $\nabla h = 0$. $(0, y)$, $y \in \mathbb{R}$, son puntos estacionarios.
- Si $y = 0$ entonces $x - 1 = 0$. $(1, 0)$ es un punto estacionario.
- Si $x \neq 0$ y $y \neq 0$, entonces

$$\begin{cases} 3x - 2y - 2 = 0 \\ x - 2y - 1 = 0 \end{cases} \Rightarrow \left(\begin{array}{cc|c} 3 & -2 & 2 \\ 1 & -2 & 1 \end{array} \right) \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{4} \end{pmatrix}$$

De todos estos puntos, el único que pertenece a S es $(1, 0)$, y $h(1, 0) = 0$.

Óptimos Línea $y = 0$, $1 \leq x \leq 2$: $f(x) := h(x, 0) = 0$ (constante).

Ejemplo:

Puntos Estacionarios: Como $\nabla h^T = (xy(3x - 2y - 2), x^2(x - 2y - 1))$ tenemos los siguientes puntos estacionarios:

- Si $x = 0$, $\nabla h = 0$. $(0, y)$, $y \in \mathbb{R}$, son puntos estacionarios.
- Si $y = 0$ entonces $x - 1 = 0$. $(1, 0)$ es un punto estacionario.
- Si $x \neq 0$ y $y \neq 0$, entonces

$$\begin{cases} 3x - 2y - 2 = 0 \\ x - 2y - 1 = 0 \end{cases} \Rightarrow \left(\begin{array}{cc|c} 3 & -2 & 2 \\ 1 & -2 & 1 \end{array} \right) \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{4} \end{pmatrix}$$

De todos estos puntos, el único que pertenece a S es $(1, 0)$, y $h(1, 0) = 0$.

Óptimos Línea $y = 0$, $1 \leq x \leq 2$: $f(x) := h(x, 0) = 0$ (constante).

Óptimos Línea $x = 2$, $0 \leq y \leq 1$: $g(y) := h(2, y) = 4y(1 - y)$.

$h'(y) = 4 - 8y = 0$, entonces $y = \frac{1}{2}$. Claramente $(2, \frac{1}{2}) \in S$ y $h(2, \frac{1}{2}) = 1$.

Ejemplo:

Puntos Estacionarios: Como $\nabla h^T = (xy(3x - 2y - 2), x^2(x - 2y - 1))$ tenemos los siguientes puntos estacionarios:

- Si $x = 0$, $\nabla h = 0$. $(0, y)$, $y \in \mathbb{R}$, son puntos estacionarios.
- Si $y = 0$ entonces $x - 1 = 0$. $(1, 0)$ es un punto estacionario.
- Si $x \neq 0$ y $y \neq 0$, entonces

$$\begin{cases} 3x - 2y - 2 = 0 \\ x - 2y - 1 = 0 \end{cases} \Rightarrow \left(\begin{array}{cc|c} 3 & -2 & 2 \\ 1 & -2 & 1 \end{array} \right) \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{4} \end{pmatrix}$$

De todos estos puntos, el único que pertenece a S es $(1, 0)$, y $h(1, 0) = 0$.

Óptimos Línea $y = 0, 1 \leq x \leq 2$: $f(x) := h(x, 0) = 0$ (constante).

Óptimos Línea $x = 2, 0 \leq y \leq 1$: $g(y) := h(2, y) = 4y(1 - y)$.

$h'(y) = 4 - 8y = 0$, entonces $y = \frac{1}{2}$. Claramente $(2, \frac{1}{2}) \in S$ y $h(2, \frac{1}{2}) = 1$.

Óptimos Línea $y = x - 1, 1 \leq x \leq 2$: $f(x) := h(x, x - 1) = 0$ (cte).

Ejemplo:

Puntos Estacionarios: Como $\nabla h^T = (xy(3x - 2y - 2), x^2(x - 2y - 1))$ tenemos los siguientes puntos estacionarios:

- Si $x = 0$, $\nabla h = 0$. $(0, y)$, $y \in \mathbb{R}$, son puntos estacionarios.
- Si $y = 0$ entonces $x - 1 = 0$. $(1, 0)$ es un punto estacionario.
- Si $x \neq 0$ y $y \neq 0$, entonces

$$\begin{cases} 3x - 2y - 2 = 0 \\ x - 2y - 1 = 0 \end{cases} \Rightarrow \left(\begin{array}{cc|c} 3 & -2 & 2 \\ 1 & -2 & 1 \end{array} \right) \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{4} \end{pmatrix}$$

De todos estos puntos, el único que pertenece a S es $(1, 0)$, y $h(1, 0) = 0$.

Óptimos Línea $y = 0$, $1 \leq x \leq 2$: $f(x) := h(x, 0) = 0$ (constante).

Óptimos Línea $x = 2$, $0 \leq y \leq 1$: $g(y) := h(2, y) = 4y(1 - y)$.

$h'(y) = 4 - 8y = 0$, entonces $y = \frac{1}{2}$. Claramente $(2, \frac{1}{2}) \in S$ y $h(2, \frac{1}{2}) = 1$.

Óptimos Línea $y = x - 1$, $1 \leq x \leq 2$: $f(x) := h(x, x - 1) = 0$ (cte).

Óptimo Vértices $(1, 0)$, $(2, 0)$, $(2, 1)$: $h(1, 0) = h(2, 0) = h(2, 1) = 0$.

Ejemplo:

Puntos Estacionarios: Como $\nabla h^T = (xy(3x - 2y - 2), x^2(x - 2y - 1))$ tenemos los siguientes puntos estacionarios:

- Si $x = 0$, $\nabla h = 0$. $(0, y)$, $y \in \mathbb{R}$, son puntos estacionarios.
- Si $y = 0$ entonces $x - 1 = 0$. $(1, 0)$ es un punto estacionario.
- Si $x \neq 0$ y $y \neq 0$, entonces

$$\begin{cases} 3x - 2y - 2 = 0 \\ x - 2y - 1 = 0 \end{cases} \Rightarrow \left(\begin{array}{cc|c} 3 & -2 & 2 \\ 1 & -2 & 1 \end{array} \right) \Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{4} \end{pmatrix}$$

De todos estos puntos, el único que pertenece a S es $(1, 0)$, y $h(1, 0) = 0$.

Óptimos Línea $y = 0$, $1 \leq x \leq 2$: $f(x) := h(x, 0) = 0$ (constante).

Óptimos Línea $x = 2$, $0 \leq y \leq 1$: $g(y) := h(2, y) = 4y(1 - y)$.

$h'(y) = 4 - 8y = 0$, entonces $y = \frac{1}{2}$. Claramente $(2, \frac{1}{2}) \in S$ y $h(2, \frac{1}{2}) = 1$.

Óptimos Línea $y = x - 1$, $1 \leq x \leq 2$: $f(x) := h(x, x - 1) = 0$ (cte).

Óptimo Vértices $(1, 0)$, $(2, 0)$, $(2, 1)$: $h(1, 0) = h(2, 0) = h(2, 1) = 0$.

Conclusión: El máximo global de h es 1, alcanzado en $(2, \frac{1}{2})$, y el mínimo global es 0, alcanzado en más de un punto.

PRECAUCIÓN

La condición de primer orden es **necesaria** y no suficiente, porque es posible que un punto estacionario no sea máximo ni mínimo.

PRECAUCIÓN

La condición de primer orden es **necesaria** y no suficiente, porque es posible que un punto estacionario no sea máximo ni mínimo.

Definición

Si c es un punto estacionario de f que no es óptimo (f no alcanza un máximo ni un mínimo en c), entonces c se llama “punto de silla”.

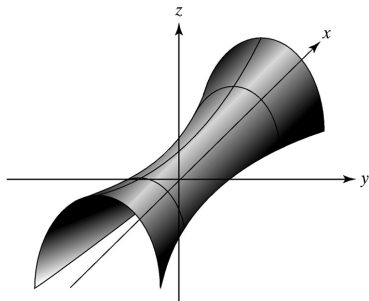


Figura: Ejemplo de punto de silla

Ejemplo

Verifiquemos que $(0, 0)$ es un punto de silla de

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

Ejemplo

Verifiquemos que $(0,0)$ es un punto de silla de

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

Punto Estacionario: Calculando el gradiente

$$\nabla f(x, y) = \begin{pmatrix} 2x \\ -2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Claramente, $(0,0)$ es un punto estacionario de f .

Ejemplo

Verifiquemos que $(0,0)$ es un punto de silla de

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

Punto Estacionario: Calculando el gradiente

$$\nabla f(x, y) = \begin{pmatrix} 2x \\ -2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Claramente, $(0,0)$ es un punto estacionario de f .

No es Mínimo: Como $g(y) := f(0, y) = -y^2$ es una cuadrática cóncava, que alcanza un máximo en $y = 0$, entonces el punto $(0,0)$ es un máximo en la trayectoria $(0, y)$, por lo que no puede ser un mínimo local ni global.

Ejemplo

Verifiquemos que $(0,0)$ es un punto de silla de

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

Punto Estacionario: Calculando el gradiente

$$\nabla f(x, y) = \begin{pmatrix} 2x \\ -2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Claramente, $(0,0)$ es un punto estacionario de f .

No es Mínimo: Como $g(y) := f(0, y) = -y^2$ es una cuadrática cóncava, que alcanza un máximo en $y = 0$, entonces el punto $(0,0)$ es un máximo en la trayectoria $(0, y)$, por lo que no puede ser un mínimo local ni global.

No es Máximo: Como $h(x) := f(x, 0) = x^2$ es una cuadrática convexa, la cual alcanza un mínimo en $x = 0$, entonces el punto $(0,0)$ es un mínimo en la trayectoria $(x, 0)$, por lo que no puede ser un máximo local ni global.

Ejemplo

Verifiquemos que $(0,0)$ es un punto de silla de

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

Punto Estacionario: Calculando el gradiente

$$\nabla f(x, y) = \begin{pmatrix} 2x \\ -2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Claramente, $(0,0)$ es un punto estacionario de f .

No es Mínimo: Como $g(y) := f(0, y) = -y^2$ es una cuadrática cóncava, que alcanza un máximo en $y = 0$, entonces el punto $(0,0)$ es un máximo en la trayectoria $(0, y)$, por lo que no puede ser un mínimo local ni global.

No es Máximo: Como $h(x) := f(x, 0) = x^2$ es una cuadrática convexa, la cual alcanza un mínimo en $x = 0$, entonces el punto $(0,0)$ es un mínimo en la trayectoria $(x, 0)$, por lo que no puede ser un máximo local ni global.

Conclusión: El punto $(0,0)$ es estacionario y para cualquier bola $B_\epsilon(0,0)$ existen puntos $|x| < \epsilon$ y $|y| < \epsilon$, con $f(0, y) \leq f(0,0) \leq f(x, 0)$, es decir, $(0,0)$ no es ni un máximo ni un mínimo, es un punto de silla.

Condición Suficiente para Máximos-Mínimos en $\mathbb{R}^n$

Condición de Segundo Orden para Máximo-Mínimo Global en $\mathbb{R}^n$

Sea $f(\mathbf{x}) \in \mathcal{C}^2(S)$, con S **convexo**, y $\mathbf{x}^*$ un punto estacionario. $f(\mathbf{x}^*)$ es:

(a) un máximo global si f es cóncava, i.e., $\forall \mathbf{x} \in S$ se cumple

$$(-1)^r \Delta_r(\mathbf{x}) \geq 0, \quad r = 1, \dots, n.$$

(b) un mínimo global si f es convexa, i.e., $\forall \mathbf{x} \in S$ se cumple

$$\Delta_r(\mathbf{x}) \geq 0, \quad r = 1, \dots, n.$$

Condición Suficiente para Máximos-Mínimos en $\mathbb{R}^n$

Condición de Segundo Orden para Máximo-Mínimo Global en $\mathbb{R}^n$

Sea $f(\mathbf{x}) \in \mathcal{C}^2(S)$, con S **convexo**, y $\mathbf{x}^*$ un punto estacionario. $f(\mathbf{x}^*)$ es:

- (a) un máximo global si f es cóncava, i.e., $\forall \mathbf{x} \in S$ se cumple

$$(-1)^r \Delta_r(\mathbf{x}) \geq 0, \quad r = 1, \dots, n.$$

- (b) un mínimo global si f es convexa, i.e., $\forall \mathbf{x} \in S$ se cumple

$$\Delta_r(\mathbf{x}) \geq 0, \quad r = 1, \dots, n.$$

Condición de Segundo Orden para Máximo-Mínimo Local en $\mathbb{R}^n$

Sea $f(\mathbf{x}) \in \mathcal{C}^2(S)$, y sea $\mathbf{x}^*$ un punto estacionario de f en S .

- (a) $f(\mathbf{x}^*)$ es un máximo local si f es estrictamente cóncava, i.e.,

$$(-1)^r D_r(\mathbf{x}^*) > 0, \quad r = 1, \dots, n.$$

- (b) $f(\mathbf{x}^*)$ es un mínimo local si f es estrictamente convexa, i.e.,

$$D_r(\mathbf{x}^*) > 0, \quad r = 1, \dots, n.$$

- (c) $\mathbf{x}^*$ es un punto de silla de f si $D_n(\mathbf{x}^*) \neq 0$ y ninguna de las condiciones anteriores ((a),(b)) se cumplen.

- (d) No se puede concluir si $D_n(\mathbf{x}^*) = 0$.

Ejemplo:

Encuentre y clasifique los puntos estacionarios de las siguientes funciones:

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

(b) $f(x_1, x_2, x_3, x_4) = 20x_2 + 48x_3 + 6x_4 + 8x_1x_2 - 4x_1^2 - 12x_3^2 - x_4^2 - 4x_2^3.$

Ejemplo:

Encuentre y clasifique los puntos estacionarios de las siguientes funciones:

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

(b) $f(x_1, x_2, x_3, x_4) = 20x_2 + 48x_3 + 6x_4 + 8x_1x_2 - 4x_1^2 - 12x_3^2 - x_4^2 - 4x_2^3$.

Solución (a): Encontrando el gradiente,

$$\nabla f(x, y, z) = \begin{pmatrix} 2x + 2y + 2z \\ 2x + 4y \\ 2x + 6z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Ejemplo:

Encuentre y clasifique los puntos estacionarios de las siguientes funciones:

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

(b) $f(x_1, x_2, x_3, x_4) = 20x_2 + 48x_3 + 6x_4 + 8x_1x_2 - 4x_1^2 - 12x_3^2 - x_4^2 - 4x_2^3$.

Solución (a): Encontrando el gradiente,

$$\nabla f(x, y, z) = \begin{pmatrix} 2x + 2y + 2z \\ 2x + 4y \\ 2x + 6z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Por otra parte, para todo $(x, y, z) \in \mathbb{R}^3$, la Hessiana es

$$H_f = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 6 \end{pmatrix} \implies D_1 = 2 > 0, \quad D_2 = \begin{vmatrix} 2 & 2 \\ 2 & 4 \end{vmatrix} = 4 > 0, \quad \text{y} \quad D_3 = 8 > 0.$$

Entonces, f es estrictamente convexa en $\mathbb{R}^3$.

Ejemplo:

Encuentre y clasifique los puntos estacionarios de las siguientes funciones:

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

(b) $f(x_1, x_2, x_3, x_4) = 20x_2 + 48x_3 + 6x_4 + 8x_1x_2 - 4x_1^2 - 12x_3^2 - x_4^2 - 4x_2^3$.

Solución (a): Encontrando el gradiente,

$$\nabla f(x, y, z) = \begin{pmatrix} 2x + 2y + 2z \\ 2x + 4y \\ 2x + 6z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Por otra parte, para todo $(x, y, z) \in \mathbb{R}^3$, la Hessiana es

$$H_f = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 4 & 0 \\ 2 & 0 & 6 \end{pmatrix} \implies D_1 = 2 > 0, \quad D_2 = \begin{vmatrix} 2 & 2 \\ 2 & 4 \end{vmatrix} = 4 > 0, \quad \text{y} \quad D_3 = 8 > 0.$$

Entonces, f es estrictamente convexa en $\mathbb{R}^3$.

En particular, f es convexa, y por ende $f(0, 0, 0) = 0$ es un mínimo global.

Ejemplo: Solución (b)

$$\nabla f(\mathbf{x}) = \begin{pmatrix} 8x_2 - 8x_1 \\ 20 + 8x_1 - 12x_2^2 \\ 48 - 24x_3 \\ 6 - 2x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow x_3 = 2, x_4 = 3, x_1 = x_2 = \begin{cases} \frac{5}{3} \\ -1 \end{cases}$$

Los puntos estacionarios son $\mathbf{x}^* = (-1, -1, 2, 3)$ y $\mathbf{x}^* = (\frac{5}{3}, \frac{5}{3}, 2, 3)$.

Ejemplo: Solución (b)

$$\nabla f(\mathbf{x}) = \begin{pmatrix} 8x_2 - 8x_1 \\ 20 + 8x_1 - 12x_2^2 \\ 48 - 24x_3 \\ 6 - 2x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow x_3 = 2, x_4 = 3, x_1 = x_2 = \begin{cases} \frac{5}{3} \\ -1 \end{cases}$$

Los puntos estacionarios son $\mathbf{x}^* = (-1, -1, 2, 3)$ y $\mathbf{x}^* = (\frac{5}{3}, \frac{5}{3}, 2, 3)$.

Por otra parte, la Hessiana es

$$H_f(\mathbf{x}) = \begin{pmatrix} -8 & 8 & 0 & 0 \\ 8 & -24x_2 & 0 & 0 \\ 0 & 0 & -24 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \Rightarrow \begin{cases} D_1(\mathbf{x}) = -8, \\ D_2(\mathbf{x}) = 64(3x_2 - 1), \\ D_3(\mathbf{x}) = -1536(3x_2 - 1), \\ D_4(\mathbf{x}) = 3072(3x_2 - 1) \end{cases}$$

Ejemplo: Solución (b)

$$\nabla f(\mathbf{x}) = \begin{pmatrix} 8x_2 - 8x_1 \\ 20 + 8x_1 - 12x_2^2 \\ 48 - 24x_3 \\ 6 - 2x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow x_3 = 2, x_4 = 3, x_1 = x_2 = \begin{cases} \frac{5}{3} \\ -1 \end{cases}$$

Los puntos estacionarios son $\mathbf{x}^* = (-1, -1, 2, 3)$ y $\mathbf{x}^* = (\frac{5}{3}, \frac{5}{3}, 2, 3)$.

Por otra parte, la Hessiana es

$$H_f(\mathbf{x}) = \begin{pmatrix} -8 & 8 & 0 & 0 \\ 8 & -24x_2 & 0 & 0 \\ 0 & 0 & -24 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \Rightarrow \begin{cases} D_1(\mathbf{x}) = -8, \\ D_2(\mathbf{x}) = 64(3x_2 - 1), \\ D_3(\mathbf{x}) = -1536(3x_2 - 1), \\ D_4(\mathbf{x}) = 3072(3x_2 - 1) \end{cases}$$

- $\mathbf{x}^* = (-1, -1, 2, 3)$: Como $D_2(\mathbf{x}^*) = (-1)^2 D_2(\mathbf{x}^*) = -256 < 0$ y $D_4(\mathbf{x}^*) = -12288 \neq 0$, se concluye que este es un punto de silla.

Ejemplo: Solución (b)

$$\nabla f(\mathbf{x}) = \begin{pmatrix} 8x_2 - 8x_1 \\ 20 + 8x_1 - 12x_2^2 \\ 48 - 24x_3 \\ 6 - 2x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow x_3 = 2, x_4 = 3, x_1 = x_2 = \begin{cases} \frac{5}{3} \\ -1 \end{cases}$$

Los puntos estacionarios son $\mathbf{x}^* = (-1, -1, 2, 3)$ y $\mathbf{x}^* = (\frac{5}{3}, \frac{5}{3}, 2, 3)$.

Por otra parte, la Hessiana es

$$H_f(\mathbf{x}) = \begin{pmatrix} -8 & 8 & 0 & 0 \\ 8 & -24x_2 & 0 & 0 \\ 0 & 0 & -24 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \Rightarrow \begin{cases} D_1(\mathbf{x}) = -8, \\ D_2(\mathbf{x}) = 64(3x_2 - 1), \\ D_3(\mathbf{x}) = -1536(3x_2 - 1), \\ D_4(\mathbf{x}) = 3072(3x_2 - 1) \end{cases}$$

- $\mathbf{x}^* = (-1, -1, 2, 3)$: Como $D_2(\mathbf{x}^*) = (-1)^2 D_2(\mathbf{x}^*) = -256 < 0$ y $D_4(\mathbf{x}^*) = -12288 \neq 0$, se concluye que este es un punto de silla.
- $\mathbf{x}^* = (\frac{5}{3}, \frac{5}{3}, 2, 3)$: $(-1)^1 D_1(\mathbf{x}^*) = 8 > 0$, $(-1)^2 D_2(\mathbf{x}^*) = 256 > 0$, $(-1)^3 D_3(\mathbf{x}^*) = 6144 > 0$, y $(-1)^4 D_4(\mathbf{x}^*) = 12288 > 0$, se concluye entonces que $f(\mathbf{x}^*) = 82,93$ es un máximo local.

Condición Suficiente para Máximo-Mínimo Global en $\mathbb{R}^2$

Condición de Segundo Orden para Máximo-Mínimo Global en $\mathbb{R}^2$

Sea $f(x, y)$ una función con derivadas parciales continuas de primer y segundo orden en un dominio **convexo** S , y sea (x^*, y^*) un punto interior de S que es punto estacionario de f . $f(x^*, y^*)$ es:

(a) un máximo global si f es cóncava, i.e., $\forall (x, y) \in S$ se cumple

- $\partial_{xx}f(x, y) \leq 0$,
- $\partial_{yy}f(x, y) \leq 0$,
- $\partial_{xx}f(x, y)\partial_{yy}f(x, y) - (\partial_{xy}f(x, y))^2 \geq 0$.

(b) un mínimo global si f es convexa, i.e., $\forall (x, y) \in S$ se cumple

- $\partial_{xx}f(x, y) \geq 0$,
- $\partial_{yy}f(x, y) \geq 0$,
- $\partial_{xx}f(x, y)\partial_{yy}f(x, y) - (\partial_{xy}f(x, y))^2 \geq 0$.

Condición Suficiente para Máximo-Mínimo Global en $\mathbb{R}^2$

Condición de Segundo Orden para Máximo-Mínimo Global en $\mathbb{R}^2$

Sea $f(x, y)$ una función con derivadas parciales continuas de primer y segundo orden en un dominio **convexo** S , y sea (x^*, y^*) un punto interior de S que es punto estacionario de f . $f(x^*, y^*)$ es:

- (a) un máximo global si f es cóncava, i.e., $\forall (x, y) \in S$ se cumple
- $\partial_{xx}f(x, y) \leq 0$,
 - $\partial_{yy}f(x, y) \leq 0$,
 - $\partial_{xx}f(x, y)\partial_{yy}f(x, y) - (\partial_{xy}f(x, y))^2 \geq 0$.
- (b) un mínimo global si f es convexa, i.e., $\forall (x, y) \in S$ se cumple
- $\partial_{xx}f(x, y) \geq 0$,
 - $\partial_{yy}f(x, y) \geq 0$,
 - $\partial_{xx}f(x, y)\partial_{yy}f(x, y) - (\partial_{xy}f(x, y))^2 \geq 0$.

Nota: Observe que se requiere que el conjunto S sea convexo y que las segundas derivadas de la función f cumplan las desigualdades en todo S .

Ejemplo 1

Los beneficios anuales (en millones de dólares) de una empresa están dados por

$$P(x, y) = -x^2 - y^2 + 22x + 18y - 102$$

donde x es la cantidad invertida en investigación (en millones de dólares) e y es el gasto publicitario (también en millones de dólares).

- 1 Hallar los beneficios cuando $x = 10$, $y = 8$ y cuando $x = 12$, $y = 10$.
- 2 Hallar los valores de x e y que podrían maximizar los beneficios.
- 3 Verifique si dichos valores de x e y maximizan (globalmente) los beneficios.
- 4Cuál es el beneficio máximo posible?

Ejemplo 1: Solución

- ① Evaluamos los puntos $(10,8)$ y $(12,10)$ en la función P .
- $P(10,8) = -10^2 - 8^2 + 22 \cdot 10 + 18 \cdot 8 - 102 = 98$
 - $P(12,10) = -12^2 - 10^2 + 22 \cdot 12 + 18 \cdot 10 - 102 = 98$

Ejemplo 1: Solución

- ① Evaluamos los puntos $(10,8)$ y $(12,10)$ en la función P .
 - $P(10,8) = -10^2 - 8^2 + 22 \cdot 10 + 18 \cdot 8 - 102 = 98$
 - $P(12,10) = -12^2 - 10^2 + 22 \cdot 12 + 18 \cdot 10 - 102 = 98$
- ② Calculamos el gradiente (y los puntos estacionarios)

$$\nabla P(x, y) = \begin{pmatrix} -2x + 22 \\ -2y + 18 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

De aquí deducimos que $(x, y) = (11, 9)$ es el único punto estacionario.

Ejemplo 1: Solución

- ① Evaluamos los puntos $(10,8)$ y $(12,10)$ en la función P .

- $P(10,8) = -10^2 - 8^2 + 22 \cdot 10 + 18 \cdot 8 - 102 = 98$
- $P(12,10) = -12^2 - 10^2 + 22 \cdot 12 + 18 \cdot 10 - 102 = 98$

- ② Calculamos el gradiente (y los puntos estacionarios)

$$\nabla P(x, y) = \begin{pmatrix} -2x + 22 \\ -2y + 18 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

De aquí deducimos que $(x, y) = (11, 9)$ es el único punto estacionario.

- ③ Note que para todo $(x, y) \in \mathbb{R}^2$ se cumple que

$$\partial_{xx}P(x, y) = -2 < 0, \quad \partial_{yy}P(x, y) = -2 < 0,$$

y

$$\partial_{xx}P(x, y)\partial_{yy}P(x, y) - (\partial_{xy}P(x, y))^2 = (-2) \cdot (-2) - 0^2 = 4 > 0.$$

Por el teorema anterior, $P(11, 9)$ es un máximo global de P .

Ejemplo 1: Solución

- ① Evaluamos los puntos $(10,8)$ y $(12,10)$ en la función P .

- $P(10,8) = -10^2 - 8^2 + 22 \cdot 10 + 18 \cdot 8 - 102 = 98$
- $P(12,10) = -12^2 - 10^2 + 22 \cdot 12 + 18 \cdot 10 - 102 = 98$

- ② Calculamos el gradiente (y los puntos estacionarios)

$$\nabla P(x, y) = \begin{pmatrix} -2x + 22 \\ -2y + 18 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

De aquí deducimos que $(x, y) = (11, 9)$ es el único punto estacionario.

- ③ Note que para todo $(x, y) \in \mathbb{R}^2$ se cumple que

$$\partial_{xx}P(x, y) = -2 < 0, \quad \partial_{yy}P(x, y) = -2 < 0,$$

y

$$\partial_{xx}P(x, y)\partial_{yy}P(x, y) - (\partial_{xy}P(x, y))^2 = (-2) \cdot (-2) - 0^2 = 4 > 0.$$

Por el teorema anterior, $P(11, 9)$ es un máximo global de P .

- ④ Finalmente, $P(11, 9) = 100$ es el beneficio máximo posible.

Ejemplo 2

Hallar el mínimo de $x^2 + y^2 + z^2$ cuando se exige que $4x + 2y - z = 5$.
(Interpretación: Punto del plano $4x + 2y - z = 5$ más cerca del origen.)

Ejemplo 2

Hallar el mínimo de $x^2 + y^2 + z^2$ cuando se exige que $4x + 2y - z = 5$.
(Interpretación: Punto del plano $4x + 2y - z = 5$ más cerca del origen.)

Solución: De la restricción tenemos que $z = 4x + 2y - 5$. Entonces

$$f(x, y) = x^2 + y^2 + (4x + 2y - 5)^2.$$

Ejemplo 2

Hallar el mínimo de $x^2 + y^2 + z^2$ cuando se exige que $4x + 2y - z = 5$.
(Interpretación: Punto del plano $4x + 2y - z = 5$ más cerca del origen.)

Solución: De la restricción tenemos que $z = 4x + 2y - 5$. Entonces

$$f(x, y) = x^2 + y^2 + (4x + 2y - 5)^2.$$

El gradiente sería $\nabla f(x, y)^T = (34x + 16y - 40, 16x + 10y - 20)$.

Resolviendo el sistema de ecuaciones

$$\begin{pmatrix} 34x + 16y \\ 16x + 10y \end{pmatrix} = \begin{pmatrix} 40 \\ 20 \end{pmatrix} \implies \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{20}{21} \\ \frac{10}{21} \end{pmatrix}$$

Ejemplo 2

Hallar el mínimo de $x^2 + y^2 + z^2$ cuando se exige que $4x + 2y - z = 5$.
(Interpretación: Punto del plano $4x + 2y - z = 5$ más cerca del origen.)

Solución: De la restricción tenemos que $z = 4x + 2y - 5$. Entonces

$$f(x, y) = x^2 + y^2 + (4x + 2y - 5)^2.$$

El gradiente sería $\nabla f(x, y)^T = (34x + 16y - 40, 16x + 10y - 20)$.

Resolviendo el sistema de ecuaciones

$$\begin{pmatrix} 34x + 16y \\ 16x + 10y \end{pmatrix} = \begin{pmatrix} 40 \\ 20 \end{pmatrix} \implies \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{20}{21} \\ \frac{10}{21} \end{pmatrix}$$

Note que $\forall (x, y) \in \mathbb{R}^2$, $\partial_{xx}f(x, y) = 34 > 0$, $\partial_{yy}f(x, y) = 10 > 0$, y
 $\partial_{xx}f(x, y)\partial_{yy}f(x, y) - (\partial_{xy}f(x, y))^2 = 34 \cdot 10 - 16^2 = 84 > 0$.

El único punto estacionario $\left(\frac{20}{21}, \frac{10}{21}\right)$ es un óptimo (mínimo) global.

Ejemplo 2

Hallar el mínimo de $x^2 + y^2 + z^2$ cuando se exige que $4x + 2y - z = 5$.
(Interpretación: Punto del plano $4x + 2y - z = 5$ más cerca del origen.)

Solución: De la restricción tenemos que $z = 4x + 2y - 5$. Entonces

$$f(x, y) = x^2 + y^2 + (4x + 2y - 5)^2.$$

El gradiente sería $\nabla f(x, y)^T = (34x + 16y - 40, 16x + 10y - 20)$.

Resolviendo el sistema de ecuaciones

$$\begin{pmatrix} 34x + 16y \\ 16x + 10y \end{pmatrix} = \begin{pmatrix} 40 \\ 20 \end{pmatrix} \implies \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{20}{21} \\ \frac{10}{21} \end{pmatrix}$$

Note que $\forall (x, y) \in \mathbb{R}^2$, $\partial_{xx}f(x, y) = 34 > 0$, $\partial_{yy}f(x, y) = 10 > 0$, y

$$\partial_{xx}f(x, y)\partial_{yy}f(x, y) - (\partial_{xy}f(x, y))^2 = 34 \cdot 10 - 16^2 = 84 > 0.$$

El único punto estacionario $\left(\frac{20}{21}, \frac{10}{21}\right)$ es un óptimo (mínimo) global.

El punto es $(x, y, z) = \left(\frac{20}{21}, \frac{10}{21}, -\frac{5}{21}\right)$ y el mínimo global es $f\left(\frac{20}{21}, \frac{10}{21}\right) = \frac{25}{21}$.

Condición Suficiente para Máximo-Mínimo Local en $\mathbb{R}^2$

Condición de Segundo Orden para Máximo-Mínimo Local en $\mathbb{R}^2$

Sea $f(x, y)$ una función con derivadas parciales continuas de primer y segundo orden en un dominio S , y sea (x^*, y^*) un punto interior de S que es punto estacionario de f .

(a) $f(x^*, y^*)$ es un máximo local si f es estrictamente cóncava, i.e.,

- $\partial_{xx}f(x^*, y^*) < 0$,
- $\partial_{xx}f(x^*, y^*)\partial_{yy}f(x^*, y^*) - (\partial_{xy}f(x^*, y^*))^2 > 0$.

(b) $f(x^*, y^*)$ es un mínimo local si f es estrictamente convexa, i.e.,

- $\partial_{xx}f(x^*, y^*) > 0$,
- $\partial_{xx}f(x^*, y^*)\partial_{yy}f(x^*, y^*) - (\partial_{xy}f(x^*, y^*))^2 > 0$.

(c) (x^*, y^*) es un punto de silla de f si

$$\partial_{xx}f(x^*, y^*)\partial_{yy}f(x^*, y^*) - (\partial_{xy}f(x^*, y^*))^2 < 0$$

(d) No se puede concluir (puede ser máx, mín, o pto de silla) si

$$\partial_{xx}f(x^*, y^*)\partial_{yy}f(x^*, y^*) - (\partial_{xy}f(x^*, y^*))^2 = 0.$$

Ejemplo

Sea la función $f(x, y) = x^2 + 2xy^2 + 2y^2$ definida para todo (x, y) .

- Calcule las derivadas parciales de primer y segundo orden de f .
- Encuentre los puntos estacionarios y clasifíquelos.

Ejemplo

Sea la función $f(x, y) = x^2 + 2xy^2 + 2y^2$ definida para todo (x, y) .

- Calcule las derivadas parciales de primer y segundo orden de f .
- Encuentre los puntos estacionarios y clasifíquelos.

Solución: El gradiente es $\nabla f(x, y) = (2x + 2y^2, 4(x + 1)y)$. Verificamos que $\partial_{xx}f(x, y) = 2$, $\partial_{yy}f(x, y) = 4(x + 1)$, y $\partial_{xy}f(x, y) = 4y$.

Ejemplo

Sea la función $f(x, y) = x^2 + 2xy^2 + 2y^2$ definida para todo (x, y) .

- Calcule las derivadas parciales de primer y segundo orden de f .
- Encuentre los puntos estacionarios y clasifíquelos.

Solución: El gradiente es $\nabla f(x, y) = (2x + 2y^2, 4(x + 1)y)$. Verificamos que $\partial_{xx}f(x, y) = 2$, $\partial_{yy}f(x, y) = 4(x + 1)$, y $\partial_{xy}f(x, y) = 4y$.

- Note que $\partial_y f(x, y) = 4(x + 1)y = 0$ solo si $x = -1$ o $y = 0$.

Ejemplo

Sea la función $f(x, y) = x^2 + 2xy^2 + 2y^2$ definida para todo (x, y) .

- Calcule las derivadas parciales de primer y segundo orden de f .
- Encuentre los puntos estacionarios y clasifíquelos.

Solución: El gradiente es $\nabla f(x, y) = (2x + 2y^2, 4(x + 1)y)$. Verificamos que $\partial_{xx}f(x, y) = 2$, $\partial_{yy}f(x, y) = 4(x + 1)$, y $\partial_{xy}f(x, y) = 4y$.

- Note que $\partial_y f(x, y) = 4(x + 1)y = 0$ solo si $x = -1$ o $y = 0$.
 - Si $x = -1$: $\partial_x f(-1, y) = -2 + 2y^2 = 0$, lo que implica que $y = \pm 1$.
Los puntos estacionarios de este caso son $(-1, -1)$ y $(-1, 1)$.

Ejemplo

Sea la función $f(x, y) = x^2 + 2xy^2 + 2y^2$ definida para todo (x, y) .

- Calcule las derivadas parciales de primer y segundo orden de f .
- Encuentre los puntos estacionarios y clasifíquelos.

Solución: El gradiente es $\nabla f(x, y) = (2x + 2y^2, 4(x + 1)y)$. Verificamos que $\partial_{xx}f(x, y) = 2$, $\partial_{yy}f(x, y) = 4(x + 1)$, y $\partial_{xy}f(x, y) = 4y$.

- Note que $\partial_y f(x, y) = 4(x + 1)y = 0$ solo si $x = -1$ o $y = 0$.
 - Si $x = -1$: $\partial_x f(-1, y) = -2 + 2y^2 = 0$, lo que implica que $y = \pm 1$.
Los puntos estacionarios de este caso son $(-1, -1)$ y $(-1, 1)$.
 - Si $y = 0$: $\partial_x f(x, 0) = 2x = 0$. $(0, 0)$ es un punto estacionario de f .

Ejemplo

Sea la función $f(x, y) = x^2 + 2xy^2 + 2y^2$ definida para todo (x, y) .

- Calcule las derivadas parciales de primer y segundo orden de f .
- Encuentre los puntos estacionarios y clasifíquelos.

Solución: El gradiente es $\nabla f(x, y) = (2x + 2y^2, 4(x + 1)y)$. Verificamos que $\partial_{xx}f(x, y) = 2$, $\partial_{yy}f(x, y) = 4(x + 1)$, y $\partial_{xy}f(x, y) = 4y$.

- Note que $\partial_y f(x, y) = 4(x + 1)y = 0$ solo si $x = -1$ o $y = 0$.
 - Si $x = -1$: $\partial_x f(-1, y) = -2 + 2y^2 = 0$, lo que implica que $y = \pm 1$.
Los puntos estacionarios de este caso son $(-1, -1)$ y $(-1, 1)$.
 - Si $y = 0$: $\partial_x f(x, 0) = 2x = 0$. $(0, 0)$ es un punto estacionario de f .

Finalmente, para clasificarlos:

- $(-1, -1)$:
 $\partial_{xx}f(-1, -1)\partial_{yy}f(-1, -1) - (\partial_{xy}f(-1, -1))^2 = -16 < 0$, por lo que es un punto de silla.

Ejemplo

Sea la función $f(x, y) = x^2 + 2xy^2 + 2y^2$ definida para todo (x, y) .

- Calcule las derivadas parciales de primer y segundo orden de f .
- Encuentre los puntos estacionarios y clasifíquelos.

Solución: El gradiente es $\nabla f(x, y) = (2x + 2y^2, 4(x + 1)y)$. Verificamos que $\partial_{xx}f(x, y) = 2$, $\partial_{yy}f(x, y) = 4(x + 1)$, y $\partial_{xy}f(x, y) = 4y$.

- Note que $\partial_y f(x, y) = 4(x + 1)y = 0$ solo si $x = -1$ o $y = 0$.
 - Si $x = -1$: $\partial_x f(-1, y) = -2 + 2y^2 = 0$, lo que implica que $y = \pm 1$.
Los puntos estacionarios de este caso son $(-1, -1)$ y $(-1, 1)$.
 - Si $y = 0$: $\partial_x f(x, 0) = 2x = 0$. $(0, 0)$ es un punto estacionario de f .

Finalmente, para clasificarlos:

- $(-1, -1)$:
 $\partial_{xx}f(-1, -1)\partial_{yy}f(-1, -1) - (\partial_{xy}f(-1, -1))^2 = -16 < 0$, por lo que es un punto de silla.
- $(-1, 1)$: $\partial_{xx}f(-1, 1)\partial_{yy}f(-1, 1) - (\partial_{xy}f(-1, 1))^2 = -16 < 0$, por lo que es un punto de silla.

Ejemplo

Sea la función $f(x, y) = x^2 + 2xy^2 + 2y^2$ definida para todo (x, y) .

- Calcule las derivadas parciales de primer y segundo orden de f .
- Encuentre los puntos estacionarios y clasifíquelos.

Solución: El gradiente es $\nabla f(x, y) = (2x + 2y^2, 4(x + 1)y)$. Verificamos que $\partial_{xx}f(x, y) = 2$, $\partial_{yy}f(x, y) = 4(x + 1)$, y $\partial_{xy}f(x, y) = 4y$.

- Note que $\partial_y f(x, y) = 4(x + 1)y = 0$ solo si $x = -1$ o $y = 0$.
 - Si $x = -1$: $\partial_x f(-1, y) = -2 + 2y^2 = 0$, lo que implica que $y = \pm 1$.
Los puntos estacionarios de este caso son $(-1, -1)$ y $(-1, 1)$.
 - Si $y = 0$: $\partial_x f(x, 0) = 2x = 0$. $(0, 0)$ es un punto estacionario de f .

Finalmente, para clasificarlos:

- **$(-1, -1)$:**
 $\partial_{xx}f(-1, -1)\partial_{yy}f(-1, -1) - (\partial_{xy}f(-1, -1))^2 = -16 < 0$, por lo que es un punto de silla.
- **$(-1, 1)$:** $\partial_{xx}f(-1, 1)\partial_{yy}f(-1, 1) - (\partial_{xy}f(-1, 1))^2 = -16 < 0$, por lo que es un punto de silla.
- **$(0, 0)$:** $\partial_{xx}f(0, 0)\partial_{yy}f(0, 0) - (\partial_{xy}f(0, 0))^2 = 8 > 0$. Como $\partial_{xx}f(0, 0) = 2 > 0$, sabemos que $f(0, 0) = 0$ es un mínimo local.

Resumen del Procedimiento

- 1 **Condiciones de Primer Orden:** Se plantea el gradiente de f y se iguala a cero, i.e.,
$$\nabla f = \mathbf{0}.$$

Se encuentran todos los puntos x^* que satisfacen esta igualdad.

Resumen del Procedimiento

- 1 **Condiciones de Primer Orden:** Se plantea el gradiente de f y se iguala a cero, i.e.,
$$\nabla f = \mathbf{0}.$$

Se encuentran todos los puntos x^* que satisfacen esta igualdad.

- 2 **Condiciones de Segundo Orden:** Primero se calcula la matriz Hessiana H_f . Aquí hay dos opciones, si el óptimo es global o local.

Resumen del Procedimiento

- ① **Condiciones de Primer Orden:** Se plantea el gradiente de f y se iguala a cero, i.e.,
- $$\nabla f = \mathbf{0}.$$

Se encuentran todos los puntos $\mathbf{x}^*$ que satisfacen esta igualdad.

- ② **Condiciones de Segundo Orden:** Primero se calcula la matriz Hessiana H_f . Aquí hay dos opciones, si el óptimo es global o local.
- **Si f es estrictamente convexa/cóncava:** Se verifican los menores principales dominantes D_r en todo el dominio, para verificar si f es estrictamente convexa/cóncava, lo que implica que es convexa/cóncava, y el óptimo encontrado es mínimo/máximo global.

Resumen del Procedimiento

- ① **Condiciones de Primer Orden:** Se plantea el gradiente de f y se iguala a cero, i.e.,
- $$\nabla f = \mathbf{0}.$$

Se encuentran todos los puntos x^* que satisfacen esta igualdad.

- ② **Condiciones de Segundo Orden:** Primero se calcula la matriz Hessiana H_f . Aquí hay dos opciones, si el óptimo es global o local.
- **Si f es estrictamente convexa/cóncava:** Se verifican los menores principales dominantes D_r en todo el dominio, para verificar si f es estrictamente convexa/cóncava, lo que implica que es convexa/cóncava, y el óptimo encontrado es mínimo/máximo global.
 - **Si f es convexa/cóncava:** Se verifican los menores principales arbitrarios Δ_r en todo el dominio, para verificar si f es convexa/cóncava, y el óptimo encontrado es mínimo/máximo global.

Resumen del Procedimiento

- ❶ **Condiciones de Primer Orden:** Se plantea el gradiente de f y se iguala a cero, i.e.,
- $$\nabla f = \mathbf{0}.$$

Se encuentran todos los puntos x^* que satisfacen esta igualdad.

- ❷ **Condiciones de Segundo Orden:** Primero se calcula la matriz Hessiana H_f . Aquí hay dos opciones, si el óptimo es global o local.
- **Si f es estrictamente convexa/cóncava:** Se verifican los menores principales dominantes D_r en todo el dominio, para verificar si f es estrictamente convexa/cóncava, lo que implica que es convexa/cóncava, y el óptimo encontrado es mínimo/máximo global.
 - **Si f es convexa/cóncava:** Se verifican los menores principales arbitrarios Δ_r en todo el dominio, para verificar si f es convexa/cóncava, y el óptimo encontrado es mínimo/máximo global.
 - **Si f no es convexa/cóncava:** Se verifican los menores principales dominantes en los puntos encontrados $D_r(x^*)$, para verificar si f es estrictamente convexa/cóncava en un vecindario de x^* , y el óptimo encontrado es mínimo/máximo local o punto de silla (o si el criterio no permite decidir).

Outline

1 Preliminares Matemáticos

- Topología en $\mathbb{R}^n$
- Teorema de la Función Implícita/Inversa
- Matrices Semidefinidas y Definidas
- Convexidad-Concavidad
- Cuasiconvexidad-Cuasiconcavidad

2 Optimización

- Motivación
- Definiciones
- Optimización Sin Restricciones

3 Optimización Con Restricciones

- Restricciones de Igualdad: Lagrange
- Restricciones de Igualdad/Desigualdad: Karush-Kuhn-Tucker

Por Qué Utilizar Lagrange?

Considere el problema de maximización de utilidad con dos bienes:

$$\text{máx } U(x, y) \quad \text{con } p_x x + p_y y = m.$$

Por Qué Utilizar Lagrange?

Considere el problema de maximización de utilidad con dos bienes:

$$\text{máx } U(x, y) \quad \text{con } p_x x + p_y y = m.$$

Podríamos reescribirlo al despejar y de la restricción, es decir, $y = \frac{m - p_x x}{p_y}$, obteniendo así un problema de optimización sin restricciones:

$$\text{máx } \tilde{U}(x) := U\left(x, \frac{m - p_x x}{p_y}\right)$$

Por Qué Utilizar Lagrange?

Considere el problema de maximización de utilidad con dos bienes:

$$\text{máx } U(x, y) \quad \text{con } p_x x + p_y y = m.$$

Podríamos reescribirlo al despejar y de la restricción, es decir, $y = \frac{m - p_x x}{p_y}$, obteniendo así un problema de optimización sin restricciones:

$$\text{máx } \tilde{U}(x) := U\left(x, \frac{m - p_x x}{p_y}\right)$$

Pero, qué pasa cuando la restricción es de la forma $x \ln(x) + ye^y = m$?

Por Qué Utilizar Lagrange?

Considere el problema de maximización de utilidad con dos bienes:

$$\text{máx } U(x, y) \quad \text{con } p_x x + p_y y = m.$$

Podríamos reescribirlo al despejar y de la restricción, es decir, $y = \frac{m - p_x x}{p_y}$, obteniendo así un problema de optimización sin restricciones:

$$\text{máx } \tilde{U}(x) := U\left(x, \frac{m - p_x x}{p_y}\right)$$

Pero, qué pasa cuando la restricción es de la forma $x \ln(x) + ye^y = m$?

No podemos (algebráicamente) despejar x o y , por lo que se torna complicado reescribir el problema a uno de optimización sin restricciones.

Por Qué Utilizar Lagrange?

Considere el problema de maximización de utilidad con dos bienes:

$$\text{máx } U(x, y) \quad \text{con } p_x x + p_y y = m.$$

Podríamos reescribirlo al despejar y de la restricción, es decir, $y = \frac{m - p_x x}{p_y}$, obteniendo así un problema de optimización sin restricciones:

$$\text{máx } \tilde{U}(x) := U\left(x, \frac{m - p_x x}{p_y}\right)$$

Pero, qué pasa cuando la restricción es de la forma $x \ln(x) + ye^y = m$?

No podemos (algebráicamente) despejar x o y , por lo que se torna complicado reescribir el problema a uno de optimización sin restricciones.

La técnica de multiplicadores de Lagrange permite resolver problemas donde la relación entre las variables es más compleja.

Definición del Problema

Considere una función objetivo f definida en $\mathbb{R}^n$, y m funciones restricción $\mathbf{g} = (g_1, \dots, g_m)$ definidas en $\mathbb{R}^n$ ($m < n$).

Definición del Problema

Considere una función objetivo f definida en $\mathbb{R}^n$, y m funciones restricción $\mathbf{g} = (g_1, \dots, g_m)$ definidas en $\mathbb{R}^n$ ($m < n$).

El problema a optimizar se puede escribir de la siguiente forma:

$$\text{máx(mín)} \ f(\mathbf{x}) \quad \text{sujeta a } \mathbf{g}(\mathbf{x}) = \mathbf{c}$$

equivalentemente,

$$\text{máx(mín)} \ f(x_1, x_2, \dots, x_n) \quad \text{sujeta a} \quad \begin{cases} g_1(x_1, x_2, \dots, x_n) = c_1 \\ \vdots \\ g_m(x_1, x_2, \dots, x_n) = c_m \end{cases}$$

Definición del Problema

Considere una función objetivo f definida en $\mathbb{R}^n$, y m funciones restricción $\mathbf{g} = (g_1, \dots, g_m)$ definidas en $\mathbb{R}^n$ ($m < n$).

El problema a optimizar se puede escribir de la siguiente forma:

$$\begin{aligned} & \text{máx(mín)} f(\mathbf{x}) \quad \text{sujeta a } \mathbf{g}(\mathbf{x}) = \mathbf{c} \\ \text{equivalentemente,} \quad & \text{máx(mín)} f(x_1, x_2, \dots, x_n) \quad \text{sujeta a } \begin{cases} g_1(x_1, x_2, \dots, x_n) = c_1 \\ \vdots \\ g_m(x_1, x_2, \dots, x_n) = c_m \end{cases} \end{aligned}$$

La función $\mathbf{G}(\mathbf{x}) := \mathbf{g}(\mathbf{x}) - \mathbf{c} = \mathbf{0}$ define implícitamente a m funciones $\mathbf{h}(x_{m+1}, \dots, x_n) = (h_1(x_{m+1}, \dots, x_n), \dots, h_m(x_{m+1}, \dots, x_n))$, en el tanto

$$\det J_{\mathbf{G}|x_1, \dots, x_m}(\mathbf{x}^*) = \det J_{\mathbf{g}|x_1, \dots, x_m}(\mathbf{x}^*) \neq 0.$$

Definición del Problema

Considere una función objetivo f definida en $\mathbb{R}^n$, y m funciones restricción $\mathbf{g} = (g_1, \dots, g_m)$ definidas en $\mathbb{R}^n$ ($m < n$).

El problema a optimizar se puede escribir de la siguiente forma:

$$\text{máx(mín)} f(\mathbf{x}) \quad \text{sujeta a } \mathbf{g}(\mathbf{x}) = \mathbf{c}$$

equivalentemente,

$$\text{máx(mín)} f(x_1, x_2, \dots, x_n) \quad \text{sujeta a } \begin{cases} g_1(x_1, x_2, \dots, x_n) = c_1 \\ \vdots \\ g_m(x_1, x_2, \dots, x_n) = c_m \end{cases}$$

La función $\mathbf{G}(\mathbf{x}) := \mathbf{g}(\mathbf{x}) - \mathbf{c} = \mathbf{0}$ define implícitamente a m funciones $\mathbf{h}(x_{m+1}, \dots, x_n) = (h_1(x_{m+1}, \dots, x_n), \dots, h_m(x_{m+1}, \dots, x_n))$, en el tanto

$$\det J_{\mathbf{G}|x_1, \dots, x_m}(\mathbf{x}^*) = \det J_{\mathbf{g}|x_1, \dots, x_m}(\mathbf{x}^*) \neq 0.$$

A lo largo de la “hiper”-superficie $\mathbf{h}(x_{m+1}, \dots, x_n)$ sabemos que se cumplen las restricciones $\mathbf{g}(\mathbf{h}(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n) = \mathbf{c}$, de modo que se puede reescribir el problema como uno sin restricciones:

$$\text{máx(mín)} \tilde{f}(x_{m+1}, \dots, x_n) := f(\mathbf{h}(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n).$$

Derivación del Método de Lagrange

Sea $\tilde{\mathbf{x}} = (x_{m+1}, \dots, x_n)$. Por el teorema de la función implícita,

$$J_{\mathbf{h}}(\tilde{\mathbf{x}}) = -J_{\mathbf{g}|x_1, \dots, x_m}^{-1}(\mathbf{h}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}) J_{\mathbf{g}|x_{m+1}, \dots, x_n}(\mathbf{h}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}). \quad (3)$$

Derivación del Método de Lagrange

Sea $\tilde{\mathbf{x}} = (x_{m+1}, \dots, x_n)$. Por el teorema de la función implícita,

$$J_{\mathbf{h}}(\tilde{\mathbf{x}}) = -J_{\mathbf{g}|x_1, \dots, x_m}^{-1}(\mathbf{h}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}) J_{\mathbf{g}|x_{m+1}, \dots, x_n}(\mathbf{h}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}). \quad (3)$$

Si $\tilde{\mathbf{x}}^* = (x_{m+1}^*, \dots, x_n^*)$ es un óptimo del problema sin restricciones, es un punto estacionario ($J_{\tilde{f}}(\tilde{\mathbf{x}}^*) = \nabla \tilde{f}^T(\tilde{\mathbf{x}}^*) = \mathbf{0}$). Por la regla de la cadena,

$$J_{\tilde{f}}(\tilde{\mathbf{x}}^*) = J_{f|x_1, \dots, x_m}(\mathbf{h}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^*) J_{\mathbf{h}}(\tilde{\mathbf{x}}^*) + J_{f|x_{m+1}, \dots, x_n}(\mathbf{h}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^*) = \mathbf{0}. \quad (4)$$

Derivación del Método de Lagrange

Sea $\tilde{\mathbf{x}} = (x_{m+1}, \dots, x_n)$. Por el teorema de la función implícita,

$$J_{\mathbf{h}}(\tilde{\mathbf{x}}) = -J_{\mathbf{g}|x_1, \dots, x_m}^{-1}(\mathbf{h}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}) J_{\mathbf{g}|x_{m+1}, \dots, x_n}(\mathbf{h}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}). \quad (3)$$

Si $\tilde{\mathbf{x}}^* = (x_{m+1}^*, \dots, x_n^*)$ es un óptimo del problema sin restricciones, es un punto estacionario ($J_{\tilde{f}}(\tilde{\mathbf{x}}^*) = \nabla \tilde{f}^T(\tilde{\mathbf{x}}^*) = \mathbf{0}$). Por la regla de la cadena,

$$J_{\tilde{f}}(\tilde{\mathbf{x}}^*) = J_{f|x_1, \dots, x_m}(\mathbf{h}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^*) J_{\mathbf{h}}(\tilde{\mathbf{x}}^*) + J_{f|x_{m+1}, \dots, x_n}(\mathbf{h}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^*) = \mathbf{0}. \quad (4)$$

Sustituyendo (3) en (4), con $\mathbf{x}^* = (\mathbf{h}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^*)$, obtenemos que

$$J_{f|x_{m+1}, \dots, x_n}(\mathbf{x}^*) = J_{f|x_1, \dots, x_m}(\mathbf{x}^*) J_{\mathbf{g}|x_1, \dots, x_m}^{-1}(\mathbf{x}^*) J_{\mathbf{g}|x_{m+1}, \dots, x_n}(\mathbf{x}^*).$$

Derivación del Método de Lagrange

Sea $\tilde{\mathbf{x}} = (x_{m+1}, \dots, x_n)$. Por el teorema de la función implícita,

$$J_{\mathbf{h}}(\tilde{\mathbf{x}}) = -J_{\mathbf{g}|x_1, \dots, x_m}^{-1}(\mathbf{h}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}) J_{\mathbf{g}|x_{m+1}, \dots, x_n}(\mathbf{h}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}). \quad (3)$$

Si $\tilde{\mathbf{x}}^* = (x_{m+1}^*, \dots, x_n^*)$ es un óptimo del problema sin restricciones, es un punto estacionario ($J_{\tilde{f}}(\tilde{\mathbf{x}}^*) = \nabla \tilde{f}^T(\tilde{\mathbf{x}}^*) = \mathbf{0}$). Por la regla de la cadena,

$$J_{\tilde{f}}(\tilde{\mathbf{x}}^*) = J_{f|x_1, \dots, x_m}(\mathbf{h}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^*) J_{\mathbf{h}}(\tilde{\mathbf{x}}^*) + J_{f|x_{m+1}, \dots, x_n}(\mathbf{h}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^*) = \mathbf{0}. \quad (4)$$

Sustituyendo (3) en (4), con $\mathbf{x}^* = (\mathbf{h}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^*)$, obtenemos que

$$J_{f|x_{m+1}, \dots, x_n}(\mathbf{x}^*) = J_{f|x_1, \dots, x_m}(\mathbf{x}^*) J_{\mathbf{g}|x_1, \dots, x_m}^{-1}(\mathbf{x}^*) J_{\mathbf{g}|x_{m+1}, \dots, x_n}(\mathbf{x}^*).$$

Tomando $(\boldsymbol{\lambda}^*)^T := J_{f|x_1, \dots, x_m}(\mathbf{x}^*) J_{\mathbf{g}|x_1, \dots, x_m}^{-1}(\mathbf{x}^*)$, obtenemos que

$$J_{f|x_{m+1}, \dots, x_n}(\mathbf{x}^*) = (\boldsymbol{\lambda}^*)^T J_{\mathbf{g}|x_{m+1}, \dots, x_n}(\mathbf{x}^*); \quad J_{f|x_1, \dots, x_m}(\mathbf{x}^*) = (\boldsymbol{\lambda}^*)^T J_{\mathbf{g}|x_1, \dots, x_m}(\mathbf{x}^*)$$

es decir, $J_f(\mathbf{x}^*) = (\boldsymbol{\lambda}^*)^T J_{\mathbf{g}}(\mathbf{x}^*) \iff \nabla f(\mathbf{x}^*) - \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*) = \mathbf{0}$.

Derivación del Método de Lagrange

Sea $\tilde{\mathbf{x}} = (x_{m+1}, \dots, x_n)$. Por el teorema de la función implícita,

$$J_{\mathbf{h}}(\tilde{\mathbf{x}}) = -J_{\mathbf{g}|x_1, \dots, x_m}^{-1}(\mathbf{h}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}) J_{\mathbf{g}|x_{m+1}, \dots, x_n}(\mathbf{h}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}). \quad (3)$$

Si $\tilde{\mathbf{x}}^* = (x_{m+1}^*, \dots, x_n^*)$ es un óptimo del problema sin restricciones, es un punto estacionario ($J_{\tilde{f}}(\tilde{\mathbf{x}}^*) = \nabla \tilde{f}^T(\tilde{\mathbf{x}}^*) = \mathbf{0}$). Por la regla de la cadena,

$$J_{\tilde{f}}(\tilde{\mathbf{x}}^*) = J_{f|x_1, \dots, x_m}(\mathbf{h}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^*) J_{\mathbf{h}}(\tilde{\mathbf{x}}^*) + J_{f|x_{m+1}, \dots, x_n}(\mathbf{h}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^*) = \mathbf{0}. \quad (4)$$

Sustituyendo (3) en (4), con $\mathbf{x}^* = (\mathbf{h}(\tilde{\mathbf{x}}^*), \tilde{\mathbf{x}}^*)$, obtenemos que

$$J_{f|x_{m+1}, \dots, x_n}(\mathbf{x}^*) = J_{f|x_1, \dots, x_m}(\mathbf{x}^*) J_{\mathbf{g}|x_1, \dots, x_m}^{-1}(\mathbf{x}^*) J_{\mathbf{g}|x_{m+1}, \dots, x_n}(\mathbf{x}^*).$$

Tomando $(\boldsymbol{\lambda}^*)^T := J_{f|x_1, \dots, x_m}(\mathbf{x}^*) J_{\mathbf{g}|x_1, \dots, x_m}^{-1}(\mathbf{x}^*)$, obtenemos que

$$J_{f|x_{m+1}, \dots, x_n}(\mathbf{x}^*) = (\boldsymbol{\lambda}^*)^T J_{\mathbf{g}|x_{m+1}, \dots, x_n}(\mathbf{x}^*); \quad J_{f|x_1, \dots, x_m}(\mathbf{x}^*) = (\boldsymbol{\lambda}^*)^T J_{\mathbf{g}|x_1, \dots, x_m}(\mathbf{x}^*)$$

es decir, $J_{\mathbf{f}}(\mathbf{x}^*) = (\boldsymbol{\lambda}^*)^T J_{\mathbf{g}}(\mathbf{x}^*) \iff \nabla f(\mathbf{x}^*) - \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*) = \mathbf{0}$.

Como $\mathbf{g}(\mathbf{x}^*) = \mathbf{c}$, $(\boldsymbol{\lambda}^*, \mathbf{x}^*)$ es un punto estacionario de la función

$$\mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}) := f(\mathbf{x}) - \boldsymbol{\lambda} \cdot (\mathbf{g}(\mathbf{x}) - \mathbf{c}) = f(x_1, \dots, x_n) - \sum_{i=1}^m \lambda_i (g_i(x_1, \dots, x_n) - c_i).$$

Interpretación Geométrica del Problema en $\mathbb{R}^2$

Como $\nabla f(x, y)$ es la dirección de máximo crecimiento en (x, y) , ocupamos que sea ortogonal (dirección de no cambio) a la trayectoria definida por $g(x, y) = c$ (i.e., $\nabla f(x^*, y^*) = \lambda \nabla g(x^*, y^*)$).

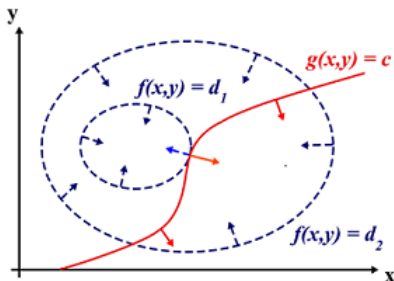
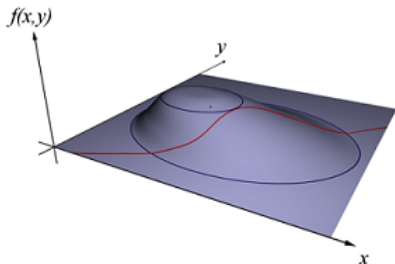
Interpretación Geométrica del Problema en $\mathbb{R}^2$

Como $\nabla f(x, y)$ es la dirección de máximo crecimiento en (x, y) , ocupamos que sea ortogonal (dirección de no cambio) a la trayectoria definida por $g(x, y) = c$ (i.e., $\nabla f(x^*, y^*) = \lambda \nabla g(x^*, y^*)$).

Buscamos donde la pendiente de la recta tangente a la curva restricción

$\left(\frac{dy}{dx} = - \left(\frac{\partial g(x, y)}{\partial y} \right)^{-1} \frac{\partial g(x, y)}{\partial x} \right)$ sea igual a la pendiente de la recta

tangente a una de las curvas de nivel de f $\left(\frac{dy}{dx} = - \left(\frac{\partial f(x, y)}{\partial y} \right)^{-1} \frac{\partial f(x, y)}{\partial x} \right)$.



Condición Necesaria para el Método de Lagrange

Note que se requirió que $\det J_{g|x_1, \dots, x_m}(\mathbf{x}^*) \neq 0$. Sin embargo, solo necesitamos que para algunas m variables se cumpla que

$$\det J_{g|x_{i_1}, \dots, x_{i_m}}(\mathbf{x}^*) \neq 0.$$

Condición Necesaria para el Método de Lagrange

Note que se requirió que $\det J_{g|x_1, \dots, x_m}(\mathbf{x}^*) \neq 0$. Sin embargo, solo necesitamos que para algunas m variables se cumpla que

$$\det J_{g|x_{i_1}, \dots, x_{i_m}}(\mathbf{x}^*) \neq 0.$$

Non-Degenerate Constraint Qualification (NDCQ)

Un punto $\mathbf{x}^*$ satisface la “Calificación de Restricción No-Degenerativa” si

$$J_g(\mathbf{x}^*) := \begin{pmatrix} \partial_{x_1} g_1(\mathbf{x}^*) & \partial_{x_2} g_1(\mathbf{x}^*) & \cdots & \partial_{x_n} g_1(\mathbf{x}^*) \\ \partial_{x_1} g_2(\mathbf{x}^*) & \partial_{x_2} g_2(\mathbf{x}^*) & \cdots & \partial_{x_n} g_2(\mathbf{x}^*) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} g_m(\mathbf{x}^*) & \partial_{x_2} g_m(\mathbf{x}^*) & \cdots & \partial_{x_n} g_m(\mathbf{x}^*) \end{pmatrix}$$

tiene rango m , y $g(\mathbf{x}^*) = \mathbf{c}$.

Condición Necesaria para el Método de Lagrange

Note que se requirió que $\det J_{g|_{x_1, \dots, x_m}}(\mathbf{x}^*) \neq 0$. Sin embargo, solo necesitamos que para algunas m variables se cumpla que

$$\det J_{g|_{x_{i_1}, \dots, x_{i_m}}}(\mathbf{x}^*) \neq 0.$$

Non-Degenerate Constraint Qualification (NDCQ)

Un punto $\mathbf{x}^*$ satisface la “Calificación de Restricción No-Degenerativa” si

$$J_g(\mathbf{x}^*) := \begin{pmatrix} \partial_{x_1} g_1(\mathbf{x}^*) & \partial_{x_2} g_1(\mathbf{x}^*) & \cdots & \partial_{x_n} g_1(\mathbf{x}^*) \\ \partial_{x_1} g_2(\mathbf{x}^*) & \partial_{x_2} g_2(\mathbf{x}^*) & \cdots & \partial_{x_n} g_2(\mathbf{x}^*) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} g_m(\mathbf{x}^*) & \partial_{x_2} g_m(\mathbf{x}^*) & \cdots & \partial_{x_n} g_m(\mathbf{x}^*) \end{pmatrix}$$

tiene rango m , y $g(\mathbf{x}^*) = c$.

En el caso de una restricción $g(x_1, \dots, x_n) = c$, se solicita que

$$\nabla g(x_1^*, \dots, x_n^*) \neq \mathbf{0}.$$

Ejemplo

Encuentre los puntos en donde las siguientes restricciones califican (NDCQ):

$$\mathbf{g}(x, y, z) = \begin{cases} g_1(x, y, z) = x^2 + y^2 + z^2 = 1, \\ g_2(x, y, z) = x + y + z = \sqrt{3}. \end{cases}$$

Ejemplo

Encuentre los puntos en donde las siguientes restricciones califican (NDCQ):

$$\mathbf{g}(x, y, z) = \begin{cases} g_1(x, y, z) = x^2 + y^2 + z^2 = 1, \\ g_2(x, y, z) = x + y + z = \sqrt{3}. \end{cases}$$

Solución: Calculemos el Jacobiano

$$J_{\mathbf{g}}(x, y, z) = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix}$$

Ejemplo

Encuentre los puntos en donde las siguientes restricciones califican (NDCQ):

$$\mathbf{g}(x, y, z) = \begin{cases} g_1(x, y, z) = x^2 + y^2 + z^2 = 1, \\ g_2(x, y, z) = x + y + z = \sqrt{3}. \end{cases}$$

Solución: Calculemos el Jacobiano

$$J_{\mathbf{g}}(x, y, z) = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix}$$

Buscamos entonces los puntos en donde las filas son linealmente dependientes,

$$\begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

con $\alpha \in \mathbb{R}$. Como $\mathbf{g}(x, y, z) = \mathbf{c}$, se debe cumplir que

$$\mathbf{g}\left(\frac{\alpha}{2}, \frac{\alpha}{2}, \frac{\alpha}{2}\right) = \begin{cases} \frac{3}{4}\alpha^2 = 1, \\ \frac{3}{2}\alpha = \sqrt{3}. \end{cases} \implies \alpha = \frac{2}{\sqrt{3}}.$$

Ejemplo

Encuentre los puntos en donde las siguientes restricciones califican (NDCQ):

$$\mathbf{g}(x, y, z) = \begin{cases} g_1(x, y, z) = x^2 + y^2 + z^2 = 1, \\ g_2(x, y, z) = x + y + z = \sqrt{3}. \end{cases}$$

Solución: Calculemos el Jacobiano

$$J_{\mathbf{g}}(x, y, z) = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix}$$

Buscamos entonces los puntos en donde las filas son linealmente dependientes,

$$\begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

con $\alpha \in \mathbb{R}$. Como $\mathbf{g}(x, y, z) = \mathbf{c}$, se debe cumplir que

$$\mathbf{g}\left(\frac{\alpha}{2}, \frac{\alpha}{2}, \frac{\alpha}{2}\right) = \begin{cases} \frac{3}{4}\alpha^2 = 1, \\ \frac{3}{2}\alpha = \sqrt{3}. \end{cases} \implies \alpha = \frac{2}{\sqrt{3}}.$$

Se califica la restricción en todo punto excepto $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$.

Contraejemplo 1: $J_g(\mathbf{x}^*)$ No Existe

Considere el problema: $\max 2x + 3y$ sujeta a $\sqrt{x} + \sqrt{y} = 5$.

Contraejemplo 1: $J_g(\mathbf{x}^*)$ No Existe

Considere el problema: $\text{máx } 2x + 3y$ sujeta a $\sqrt{x} + \sqrt{y} = 5$.

Solución: $\nabla g(x, y) = \left(\frac{1}{2\sqrt{x}}, \frac{1}{2\sqrt{y}} \right)$, el cual existe en el tanto $x \neq 0$ y $y \neq 0$. En los puntos donde $x = 0$ o $y = 0$, no se puede utilizar el teorema de la función inversa, por lo que el método de Lagrange falla.

Contraejemplo 1: $J_g(\mathbf{x}^*)$ No Existe

Considere el problema: $\text{máx } 2x + 3y$ sujeta a $\sqrt{x} + \sqrt{y} = 5$.

Solución: $\nabla g(x, y) = \left(\frac{1}{2\sqrt{x}}, \frac{1}{2\sqrt{y}} \right)$, el cual existe en el tanto $x \neq 0$ y $y \neq 0$. En los puntos donde $x = 0$ o $y = 0$, no se puede utilizar el teorema de la función inversa, por lo que el método de Lagrange falla.

El único punto estacionario de $\mathcal{L}(\lambda, x, y) = 2x + 3y - \lambda(\sqrt{x} + \sqrt{y} - 5)$ es $(\lambda^*, x^*, y^*) = (12, 9, 4) \implies P = (x^*, y^*) = (9, 4)$.

Contraejemplo 1: $J_g(x^*)$ No Existe

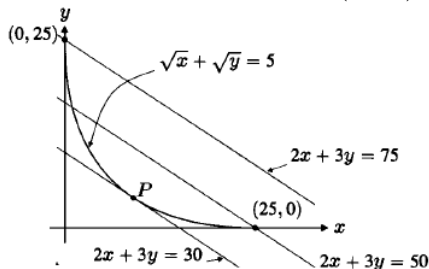
Considere el problema: máx $2x + 3y$ sujeta a $\sqrt{x} + \sqrt{y} = 5$.

Solución: $\nabla g(x, y) = \left(\frac{1}{2\sqrt{x}}, \frac{1}{2\sqrt{y}} \right)$, el cual existe en el tanto $x \neq 0$ y $y \neq 0$. En los puntos donde $x = 0$ o $y = 0$, no se puede utilizar el teorema de la función inversa, por lo que el método de Lagrange falla.

El único punto estacionario de $\mathcal{L}(\lambda, x, y) = 2x + 3y - \lambda(\sqrt{x} + \sqrt{y} - 5)$ es

$$(\lambda^*, x^*, y^*) = (12, 9, 4) \implies P = (x^*, y^*) = (9, 4).$$

Sin embargo, el óptimo se alcanza en el punto $(0, 25)$.



Contraejemplo 2: $J_g(\mathbf{x}^*)$ No Tiene Rango m

Considere el problema: $\max x$ sujeta a $x^3 + y^2 = 0$.

Contraejemplo 2: $J_g(\mathbf{x}^*)$ No Tiene Rango m

Considere el problema: $\text{máx } x$ sujeta a $x^3 + y^2 = 0$.

Solución: $g(0, 0) = 0^3 + 0^2 = 0$. Además, $\nabla g(x, y) = (3x^2, 2y) = (0, 0)$, si $x = 0$ y $y = 0$, por lo que el punto $(0, 0)$ no califica la restricción.

Contraejemplo 2: $J_g(\mathbf{x}^*)$ No Tiene Rango m

Considere el problema: $\text{máx } x$ sujeta a $x^3 + y^2 = 0$.

Solución: $g(0, 0) = 0^3 + 0^2 = 0$. Además, $\nabla g(x, y) = (3x^2, 2y) = (0, 0)$, si $x = 0$ y $y = 0$, por lo que el punto $(0, 0)$ no califica la restricción.

Por otra parte, $\mathcal{L}(\lambda, x, y) = x - \lambda(x^3 + y^2)$ no tiene puntos estacionarios, por lo que $(0, 0)$ no es un candidato con el método de Lagrange.

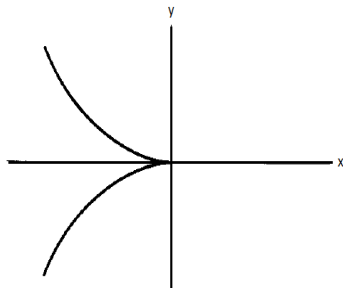
Contraejemplo 2: $J_g(\mathbf{x}^*)$ No Tiene Rango m

Considere el problema: máx x sujeta a $x^3 + y^2 = 0$.

Solución: $g(0, 0) = 0^3 + 0^2 = 0$. Además, $\nabla g(x, y) = (3x^2, 2y) = (0, 0)$, si $x = 0$ y $y = 0$, por lo que el punto $(0, 0)$ no califica la restricción.

Por otra parte, $\mathcal{L}(\lambda, x, y) = x - \lambda(x^3 + y^2)$ no tiene puntos estacionarios, por lo que $(0, 0)$ no es un candidato con el método de Lagrange.

Sin embargo, $(0, 0)$ es un máximo de $f(x, y) = x$ sujeto a $g(x, y) = 0$.



Condición Necesaria para Máximo-Mínimo

Punto Admisible / Punto Factible

Considere el problema

$$\text{máx(mín)} f(\mathbf{x}) \quad \text{sujeta a } \mathbf{g}(\mathbf{x}) = \mathbf{c}. \quad (5)$$

Un punto $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ se dice “admisible” o “factible” si cumple que

$$g_i(\mathbf{x}^*) = c_i \text{ para todo } i = 1, \dots, m.$$

Condición Necesaria para Máximo-Mínimo

Punto Admisble / Punto Factible

Considere el problema

$$\text{máx(mín)} f(\mathbf{x}) \quad \text{sujeta a } \mathbf{g}(\mathbf{x}) = \mathbf{c}. \quad (5)$$

Un punto $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ se dice “admisble” o “factible” si cumple que

$$g_i(\mathbf{x}^*) = c_i \text{ para todo } i = 1, \dots, m.$$

Teorema de Lagrange: Condiciones de Primer Orden

Sea f una función continuamente diferenciable en $\mathbb{R}^n$ y sean g_i funciones continuamente diferenciables en $\mathbb{R}^n$, para $i = 1, \dots, m$. Asuma que $\mathbf{x}^*$ es un punto admisble que resuelve el problema (5). Suponga además que $\mathbf{x}^*$ es un punto donde se califica la restricción (NDCQ). Entonces existe un único vector $(\boldsymbol{\lambda}^*)^T = (\lambda_1^*, \dots, \lambda_m^*)$ tal que $(\boldsymbol{\lambda}^*, \mathbf{x}^*)$ es un punto estacionario del Lagrangiano

$$\mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}) := f(\mathbf{x}) - \boldsymbol{\lambda} \cdot (\mathbf{g}(\mathbf{x}) - \mathbf{c}) = f(x_1, \dots, x_n) - \sum_{i=1}^m \lambda_i (g_i(x_1, \dots, x_n) - c_i),$$

esto es

$$\mathcal{L}_{x_1}(\boldsymbol{\lambda}^*, \mathbf{x}^*) = 0, \dots, \mathcal{L}_{x_n}(\boldsymbol{\lambda}^*, \mathbf{x}^*) = 0, \mathcal{L}_{\lambda_1}(\boldsymbol{\lambda}^*, \mathbf{x}^*) = 0, \dots, \mathcal{L}_{\lambda_m}(\boldsymbol{\lambda}^*, \mathbf{x}^*) = 0.$$

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n = m$$

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n = m$$

Todo punto califica la restricción (NDCQ), pues $\nabla g = (p_1, \dots, p_n) \neq \mathbf{0}$.

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n = m$$

Todo punto califica la restricción (NDCQ), pues $\nabla g = (p_1, \dots, p_n) \neq \mathbf{0}$.

Lagrangiano: $\mathcal{L}(\lambda, \mathbf{x}) = U(x_1, \dots, x_n) - \lambda(p_1x_1 + \dots + p_nx_n - m)$

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n = m$$

Todo punto califica la restricción (NDCQ), pues $\nabla g = (p_1, \dots, p_n) \neq \mathbf{0}$.

Lagrangiano: $\mathcal{L}(\lambda, \mathbf{x}) = U(x_1, \dots, x_n) - \lambda(p_1x_1 + \dots + p_nx_n - m)$

Por el teorema de Lagrange, si existe un punto $(x_1^*, \dots, x_n^*)$ que maximice la utilidad gastando m , entonces existe un valor λ^* tal que

$$\begin{cases} \mathcal{L}_{x_i}(\lambda^*, x_1^*, \dots, x_n^*) = U_{x_i}(x_1^*, \dots, x_n^*) - \lambda^*p_i = 0; & i = 1, \dots, n. \\ \mathcal{L}_\lambda(\lambda^*, x_1^*, \dots, x_n^*) = m - (p_1x_1^* + \dots + p_nx_n^*) = 0. \end{cases}$$

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n = m$$

Todo punto califica la restricción (NDCQ), pues $\nabla g = (p_1, \dots, p_n) \neq \mathbf{0}$.

Lagrangiano: $\mathcal{L}(\lambda, \mathbf{x}) = U(x_1, \dots, x_n) - \lambda(p_1x_1 + \dots + p_nx_n - m)$

Por el teorema de Lagrange, si existe un punto $(x_1^*, \dots, x_n^*)$ que maximice la utilidad gastando m , entonces existe un valor λ^* tal que

$$\begin{cases} \mathcal{L}_{x_i}(\lambda^*, x_1^*, \dots, x_n^*) = U_{x_i}(x_1^*, \dots, x_n^*) - \lambda^*p_i = 0; & i = 1, \dots, n. \\ \mathcal{L}_\lambda(\lambda^*, x_1^*, \dots, x_n^*) = m - (p_1x_1^* + \dots + p_nx_n^*) = 0. \end{cases}$$

De este modo, las condiciones de primer orden se reducen a

$$\underbrace{\frac{U_{x_i}(x_1^*, \dots, x_n^*)}{p_i} = \frac{U_{x_j}(x_1^*, \dots, x_n^*)}{p_j}}_{\text{Utilidad por Colón Gastado}} \iff \underbrace{\frac{U_{x_i}(x_1^*, \dots, x_n^*)}{U_{x_j}(x_1^*, \dots, x_n^*)}}_{\text{Tasa Marginal de Sustitución}} = \underbrace{\frac{p_i}{p_j}}_{\text{Precios Relativos}}.$$

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n = m$$

Todo punto califica la restricción (NDCQ), pues $\nabla g = (p_1, \dots, p_n) \neq \mathbf{0}$.

Lagrangiano: $\mathcal{L}(\lambda, \mathbf{x}) = U(x_1, \dots, x_n) - \lambda(p_1x_1 + \dots + p_nx_n - m)$

Por el teorema de Lagrange, si existe un punto $(x_1^*, \dots, x_n^*)$ que maximice la utilidad gastando m , entonces existe un valor λ^* tal que

$$\begin{cases} \mathcal{L}_{x_i}(\lambda^*, x_1^*, \dots, x_n^*) = U_{x_i}(x_1^*, \dots, x_n^*) - \lambda^*p_i = 0; & i = 1, \dots, n. \\ \mathcal{L}_{\lambda}(\lambda^*, x_1^*, \dots, x_n^*) = m - (p_1x_1^* + \dots + p_nx_n^*) = 0. \end{cases}$$

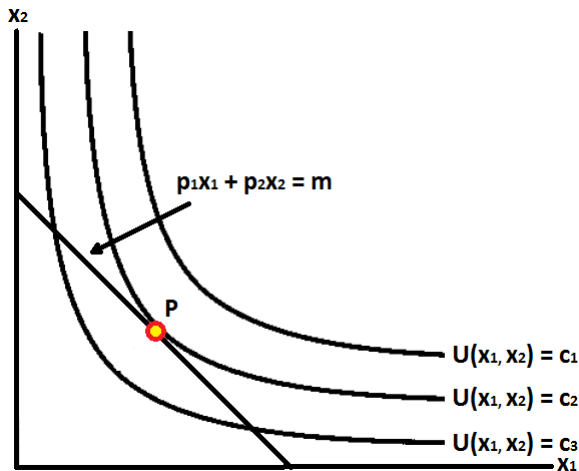
De este modo, las condiciones de primer orden se reducen a

$$\underbrace{\frac{U_{x_i}(x_1^*, \dots, x_n^*)}{p_i} = \frac{U_{x_j}(x_1^*, \dots, x_n^*)}{p_j}}_{\text{Utilidad por Colón Gastado}} \iff \underbrace{\frac{U_{x_i}(x_1^*, \dots, x_n^*)}{U_{x_j}(x_1^*, \dots, x_n^*)}}_{\text{Tasa Marginal de Sustitución}} = \underbrace{\frac{p_i}{p_j}}_{\text{Precios Relativos}}.$$

Propiedad del óptimo: La utilidad (marginal) generada por una unidad extra del bien relativa a su costo, debe ser igual para todos los bienes.

Maximización de Utilidad (Interpretación Geométrica)

Encontrar la cesta óptima de consumo implica encontrar un punto en el que la recta presupuestaria sea tangente a una curva de indiferencia.



PRECAUCIÓN

Note que el Teorema de Lagrange nos indica que el punto x^* que maximiza f sujeto a $g = c$, es junto al valor λ^* , un punto estacionario del Lagrangiano $\mathcal{L}$. Esto **NO** significa que x^* sea un óptimo de $\mathcal{L}(\lambda^*, \cdot)$.

PRECAUCIÓN

Note que el Teorema de Lagrange nos indica que el punto x^* que maximiza f sujeto a $g = c$, es junto al valor λ^* , un punto estacionario del Lagrangiano $\mathcal{L}$. Esto **NO** significa que x^* sea un óptimo de $\mathcal{L}(\lambda^*, \cdot)$.

Considere el problema:

$$\max xy \quad \text{sujeta a } x + y = 2.$$

PRECAUCIÓN

Note que el Teorema de Lagrange nos indica que el punto x^* que maximiza f sujeto a $g = c$, es junto al valor λ^* , un punto estacionario del Lagrangiano $\mathcal{L}$. Esto **NO** significa que x^* sea un óptimo de $\mathcal{L}(\lambda^*, \cdot)$.

Considere el problema:

$$\text{máx } xy \quad \text{sujeta a } x + y = 2.$$

Este se puede reescribir como máx $x(2 - x)$ (problema sin restricciones).

PRECAUCIÓN

Note que el Teorema de Lagrange nos indica que el punto x^* que maximiza f sujeto a $g = c$, es junto al valor λ^* , un punto estacionario del Lagrangiano $\mathcal{L}$. Esto **NO** significa que x^* sea un óptimo de $\mathcal{L}(\lambda^*, \cdot)$.

Considere el problema:

$$\text{máx } xy \quad \text{sujeta a } x + y = 2.$$

Este se puede reescribir como máx $x(2 - x)$ (problema sin restricciones).

Notamos que

$$f'(x) = 2 - 2x = 0 \quad \Rightarrow \quad x^* = 1 \quad \Rightarrow \quad (x^*, y^*) = (1, 1).$$

Como f es estrictamente cóncava, $f(1, 1) = 1$ es un máximo global!

PRECAUCIÓN

Note que el Teorema de Lagrange nos indica que el punto x^* que maximiza f sujeto a $g = c$, es junto al valor λ^* , un punto estacionario del Lagrangiano $\mathcal{L}$. Esto **NO** significa que x^* sea un óptimo de $\mathcal{L}(\lambda^*, \cdot)$.

Considere el problema:

$$\text{máx } xy \quad \text{sujeta a } x + y = 2.$$

Este se puede reescribir como máx $x(2 - x)$ (problema sin restricciones).

Notamos que

$$f'(x) = 2 - 2x = 0 \quad \Rightarrow \quad x^* = 1 \quad \Rightarrow \quad (x^*, y^*) = (1, 1).$$

Como f es estrictamente cóncava, $f(1, 1) = 1$ es un máximo global!

Utilizando Lagrange, vemos que $\mathcal{L}(\lambda, x, y) = xy - \lambda(x + y - 2)$, y

$$\mathcal{L}_x = y - \lambda = 0, \quad \mathcal{L}_y = x - \lambda = 0, \quad \mathcal{L}_\lambda = -(x + y) + 2 = 0,$$

concluyendo que $(\lambda^*, x^*, y^*) = (1, 1, 1)$.

PRECAUCIÓN

Note que el Teorema de Lagrange nos indica que el punto x^* que maximiza f sujeto a $g = c$, es junto al valor λ^* , un punto estacionario del Lagrangiano $\mathcal{L}$. Esto **NO** significa que x^* sea un óptimo de $\mathcal{L}(\lambda^*, \cdot)$.

Considere el problema:

$$\text{máx } xy \quad \text{sujeta a } x + y = 2.$$

Este se puede reescribir como máx $x(2 - x)$ (problema sin restricciones).

Notamos que

$$f'(x) = 2 - 2x = 0 \quad \Rightarrow \quad x^* = 1 \quad \Rightarrow \quad (x^*, y^*) = (1, 1).$$

Como f es estrictamente cóncava, $f(1, 1) = 1$ es un máximo global!

Utilizando Lagrange, vemos que $\mathcal{L}(\lambda, x, y) = xy - \lambda(x + y - 2)$, y

$$\mathcal{L}_x = y - \lambda = 0, \quad \mathcal{L}_y = x - \lambda = 0, \quad \mathcal{L}_\lambda = -(x + y) + 2 = 0,$$

concluyendo que $(\lambda^*, x^*, y^*) = (1, 1, 1)$.

Sin embargo, $(1, 1, 1)$ **NO** es un óptimo de $\mathcal{L}(1, x, y)$, ya que

$$\mathcal{L}(1, 0, 2) = 0 < \mathcal{L}(1, 1, 1) = 1 < 2 = \mathcal{L}(1, 2, 2).$$

Condición Suficiente para Máximo-Mínimo Global

Sea (λ^*, x^*) un punto estacionario de $\mathcal{L}$. Si para λ^* fijo, x^* es un óptimo **global** de $\mathcal{L}(\lambda^*, \cdot)$, entonces es un óptimo de f sujeto a $g(x) = c$.

Condición Suficiente para Máximo-Mínimo Global

Sea (λ^*, x^*) un punto estacionario de $\mathcal{L}$. Si para λ^* fijo, x^* es un óptimo **global** de $\mathcal{L}(\lambda^*, \cdot)$, entonces es un óptimo de f sujeto a $g(x) = c$.

Condición de Segundo Orden para Máximo-Mínimo Global en $\mathbb{R}^n$

Sean $f, g \in \mathcal{C}^1(S)$, con S **convexo**, y $x^* \in S$ un punto estacionario de $\mathcal{L}(\lambda^*, \cdot)$, admisible que satisface NDCQ.

- (a) Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava, entonces $f(x^*)$ es un máximo global.
- (b) Si $\mathcal{L}(\lambda^*, \cdot)$ es convexa, entonces $f(x^*)$ es un mínimo global.

Condición Suficiente para Máximo-Mínimo Global

Sea (λ^*, x^*) un punto estacionario de $\mathcal{L}$. Si para λ^* fijo, x^* es un óptimo **global** de $\mathcal{L}(\lambda^*, \cdot)$, entonces es un óptimo de f sujeto a $g(x) = c$.

Condición de Segundo Orden para Máximo-Mínimo Global en $\mathbb{R}^n$

Sean $f, g \in \mathcal{C}^1(S)$, con S **convexo**, y $x^* \in S$ un punto estacionario de $\mathcal{L}(\lambda^*, \cdot)$, admisible que satisface NDCQ.

- (a) Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava, entonces $f(x^*)$ es un máximo global.
- (b) Si $\mathcal{L}(\lambda^*, \cdot)$ es convexa, entonces $f(x^*)$ es un mínimo global.

Observaciones:

- Se exige que la función $\mathcal{L}(\lambda^*, \cdot)$ sea cóncava/convexa **en todo** S .

Condición Suficiente para Máximo-Mínimo Global

Sea (λ^*, x^*) un punto estacionario de $\mathcal{L}$. Si para λ^* fijo, x^* es un óptimo **global** de $\mathcal{L}(\lambda^*, \cdot)$, entonces es un óptimo de f sujeto a $g(x) = c$.

Condición de Segundo Orden para Máximo-Mínimo Global en $\mathbb{R}^n$

Sean $f, g \in \mathcal{C}^1(S)$, con S **convexo**, y $x^* \in S$ un punto estacionario de $\mathcal{L}(\lambda^*, \cdot)$, admisible que satisface NDCQ.

- (a) Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava, entonces $f(x^*)$ es un máximo global.
- (b) Si $\mathcal{L}(\lambda^*, \cdot)$ es convexa, entonces $f(x^*)$ es un mínimo global.

Observaciones:

- Se exige que la función $\mathcal{L}(\lambda^*, \cdot)$ sea cóncava/convexa **en todo** S .
- Si f es cóncava (convexa), g_i es convexa (cóncava) si $\lambda_i^* \geq 0$, y g_j es cóncava (convexa) si $\lambda_j^* \leq 0$, entonces $\mathcal{L}(\lambda^*, \cdot)$ sería una suma de funciones cóncavas (convexas), i.e., $f(x^*)$ sería un máximo (mínimo).

Condición Suficiente para Máximo-Mínimo Global

Sea (λ^*, x^*) un punto estacionario de $\mathcal{L}$. Si para λ^* fijo, x^* es un óptimo **global** de $\mathcal{L}(\lambda^*, \cdot)$, entonces es un óptimo de f sujeto a $g(x) = c$.

Condición de Segundo Orden para Máximo-Mínimo Global en $\mathbb{R}^n$

Sean $f, g \in C^1(S)$, con S **convexo**, y $x^* \in S$ un punto estacionario de $\mathcal{L}(\lambda^*, \cdot)$, admisible que satisface NDCQ.

- (a) Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava, entonces $f(x^*)$ es un máximo global.
- (b) Si $\mathcal{L}(\lambda^*, \cdot)$ es convexa, entonces $f(x^*)$ es un mínimo global.

Observaciones:

- Se exige que la función $\mathcal{L}(\lambda^*, \cdot)$ sea cóncava/convexa **en todo** S .
- Si f es cóncava (convexa), g_i es convexa (cóncava) si $\lambda_i^* \geq 0$, y g_j es cóncava (convexa) si $\lambda_j^* \leq 0$, entonces $\mathcal{L}(\lambda^*, \cdot)$ sería una suma de funciones cóncavas (convexas), i.e., $f(x^*)$ sería un máximo (mínimo).
- Como $a \cdot x + b$ es cóncava y convexa, si las restricciones son lineales basta que f sea cóncava/convexa para que sea un máximo/mínimo.

Ejemplo 1

Resuelva el siguiente problema:

$$\text{mín } x^2 + y^2 + z^2 \quad \text{sujeta a } x + y + z = 1.$$

Ejemplo 1

Resuelva el siguiente problema:

$$\text{mín } x^2 + y^2 + z^2 \quad \text{sujeta a } x + y + z = 1.$$

Solución: Note que la restricción es lineal. Basta con verificar que f sea convexa.

Ejemplo 1

Resuelva el siguiente problema:

$$\text{mín } x^2 + y^2 + z^2 \quad \text{sujeta a } x + y + z = 1.$$

Solución: Note que la restricción es lineal. Basta con verificar que f sea convexa.

$$H_f(x, y, z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \implies D_1 = 2 > 0, D_2 = 4 > 0, D_3 = 8 > 0,$$

por lo que f es convexa y por ende $\mathcal{L}(\lambda, \cdot)$ es convexa para todo λ .

Ejemplo 1

Resuelva el siguiente problema:

$$\text{mín } x^2 + y^2 + z^2 \quad \text{sujeta a } x + y + z = 1.$$

Solución: Note que la restricción es lineal. Basta con verificar que f sea convexa.

$$H_f(x, y, z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \implies D_1 = 2 > 0, D_2 = 4 > 0, D_3 = 8 > 0,$$

por lo que f es convexa y por ende $\mathcal{L}(\lambda, \cdot)$ es convexa para todo λ .

De acuerdo al teorema, como $\nabla g = (1, 1, 1) \neq \mathbf{0}$, el punto que satisfaga las condiciones de primer orden es el único mínimo del problema.

Ejemplo 1

Resuelva el siguiente problema:

$$\text{mín } x^2 + y^2 + z^2 \quad \text{sujeta a } x + y + z = 1.$$

Solución: Note que la restricción es lineal. Basta con verificar que f sea convexa.

$$H_f(x, y, z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \implies D_1 = 2 > 0, D_2 = 4 > 0, D_3 = 8 > 0,$$

por lo que f es convexa y por ende $\mathcal{L}(\lambda, \cdot)$ es convexa para todo λ .

De acuerdo al teorema, como $\nabla g = (1, 1, 1) \neq \mathbf{0}$, el punto que satisfaga las condiciones de primer orden es el único mínimo del problema.

$$\text{Finalmente, } \nabla \mathcal{L} = \begin{pmatrix} 1 - x - y - z \\ 2x - \lambda \\ 2y - \lambda \\ 2z - \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ x^* \\ y^* \\ z^* \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}.$$

El mínimo es $f\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = \frac{1}{3}$.

Ejemplo 1

Resuelva el siguiente problema:

$$\text{mín } x^2 + y^2 + z^2 \quad \text{sujeta a } x + y + z = 1.$$

Solución: Note que la restricción es lineal. Basta con verificar que f sea convexa.

$$H_f(x, y, z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \implies D_1 = 2 > 0, D_2 = 4 > 0, D_3 = 8 > 0,$$

por lo que f es convexa y por ende $\mathcal{L}(\lambda, \cdot)$ es convexa para todo λ .

De acuerdo al teorema, como $\nabla g = (1, 1, 1) \neq \mathbf{0}$, el punto que satisfaga las condiciones de primer orden es el único mínimo del problema.

$$\text{Finalmente, } \nabla \mathcal{L} = \begin{pmatrix} 1 - x - y - z \\ 2x - \lambda \\ 2y - \lambda \\ 2z - \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ x^* \\ y^* \\ z^* \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}.$$

El mínimo es $f\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = \frac{1}{3}$.

Cuál es la interpretación gráfica? Tiene sentido maximizar?

Ejemplo 2

Una persona posee una función de utilidad $U(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ ($\alpha > 0$), para los bienes x_1 y x_2 , con precios p_1 y p_2 , respectivamente. Si él posee m colones, cuáles cantidades x_1^* y x_2^* maximizan su felicidad?

Ejemplo 2

Una persona posee una función de utilidad $U(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ ($\alpha > 0$), para los bienes x_1 y x_2 , con precios p_1 y p_2 , respectivamente. Si él posee m colones, cuáles cantidades x_1^* y x_2^* maximizan su felicidad?

Solución: $\max x_1^\alpha x_2^{1-\alpha}$ sujeta a $p_1 x_1 + p_2 x_2 = m$.

Ejemplo 2

Una persona posee una función de utilidad $U(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ ($\alpha > 0$), para los bienes x_1 y x_2 , con precios p_1 y p_2 , respectivamente. Si él posee m colones, cuáles cantidades x_1^* y x_2^* maximizan su felicidad?

Solución: máx $x_1^\alpha x_2^{1-\alpha}$ sujeta a $p_1 x_1 + p_2 x_2 = m$.

Vimos que la condición de óptimo es

$$\frac{p_1}{p_2} = \frac{U_{x_1}(x_1, x_2)}{U_{x_2}(x_1, x_2)} = \frac{\alpha x_1^{\alpha-1} x_2^{1-\alpha}}{(1-\alpha) x_1^\alpha x_2^{-\alpha}} = \frac{\alpha}{1-\alpha} \frac{x_2}{x_1} \implies x_2 = \frac{(1-\alpha)p_1}{\alpha p_2} x_1.$$

Sustituyendo en la restricción tenemos

$$m = p_1 x_1 + p_2 \left(\frac{(1-\alpha)p_1}{\alpha p_2} x_1 \right) = \frac{p_1}{\alpha} x_1 \implies x_1^* = \frac{\alpha}{p_1} m \text{ y } x_2^* = \frac{(1-\alpha)}{p_2} m.$$

Ejemplo 2

Una persona posee una función de utilidad $U(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ ($\alpha > 0$), para los bienes x_1 y x_2 , con precios p_1 y p_2 , respectivamente. Si él posee m colones, cuáles cantidades x_1^* y x_2^* maximizan su felicidad?

Solución: máx $x_1^\alpha x_2^{1-\alpha}$ sujeta a $p_1 x_1 + p_2 x_2 = m$.

Vimos que la condición de óptimo es

$$\frac{p_1}{p_2} = \frac{U_{x_1}(x_1, x_2)}{U_{x_2}(x_1, x_2)} = \frac{\alpha x_1^{\alpha-1} x_2^{1-\alpha}}{(1-\alpha) x_1^\alpha x_2^{-\alpha}} = \frac{\alpha}{1-\alpha} \frac{x_2}{x_1} \implies x_2 = \frac{(1-\alpha)p_1}{\alpha p_2} x_1.$$

Sustituyendo en la restricción tenemos

$$m = p_1 x_1 + p_2 \left(\frac{(1-\alpha)p_1}{\alpha p_2} x_1 \right) = \frac{p_1}{\alpha} x_1 \implies x_1^* = \frac{\alpha}{p_1} m \text{ y } x_2^* = \frac{(1-\alpha)}{p_2} m.$$

$$\text{Además, } \lambda^* = \frac{U_{x_1}(x_1^*, x_2^*)}{p_1} = \frac{\alpha \left(\frac{\alpha}{p_1} m \right)^{\alpha-1} \left(\frac{(1-\alpha)}{p_2} m \right)^{1-\alpha}}{p_1} = \left(\frac{\alpha}{p_1} \right)^\alpha \left(\frac{1-\alpha}{p_2} \right)^{1-\alpha} > 0.$$

Ejemplo 2

Una persona posee una función de utilidad $U(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ ($\alpha > 0$), para los bienes x_1 y x_2 , con precios p_1 y p_2 , respectivamente. Si él posee m colones, cuáles cantidades x_1^* y x_2^* maximizan su felicidad?

Solución: máx $x_1^\alpha x_2^{1-\alpha}$ sujeta a $p_1 x_1 + p_2 x_2 = m$.

Vimos que la condición de óptimo es

$$\frac{p_1}{p_2} = \frac{U_{x_1}(x_1, x_2)}{U_{x_2}(x_1, x_2)} = \frac{\alpha x_1^{\alpha-1} x_2^{1-\alpha}}{(1-\alpha) x_1^\alpha x_2^{-\alpha}} = \frac{\alpha}{1-\alpha} \frac{x_2}{x_1} \implies x_2 = \frac{(1-\alpha)p_1}{\alpha p_2} x_1.$$

Sustituyendo en la restricción tenemos

$$m = p_1 x_1 + p_2 \left(\frac{(1-\alpha)p_1}{\alpha p_2} x_1 \right) = \frac{p_1}{\alpha} x_1 \implies x_1^* = \frac{\alpha}{p_1} m \text{ y } x_2^* = \frac{(1-\alpha)}{p_2} m.$$

$$\text{Además, } \lambda^* = \frac{U_{x_1}(x_1^*, x_2^*)}{p_1} = \frac{\alpha \left(\frac{\alpha}{p_1} m \right)^{\alpha-1} \left(\frac{(1-\alpha)}{p_2} m \right)^{1-\alpha}}{p_1} = \left(\frac{\alpha}{p_1} \right)^\alpha \left(\frac{1-\alpha}{p_2} \right)^{1-\alpha} > 0.$$

$\mathcal{L}(\lambda^*, \cdot)$ es cóncava, pues g es convexa por ser lineal (y $\lambda^* > 0$) y f es cóncava (Cobb-Douglas con $\alpha + (1-\alpha) = 1$). Entonces la utilidad máxima gastando m colones es $U(x_1^*, x_2^*) = \left(\frac{\alpha}{p_1} \right)^\alpha \left(\frac{1-\alpha}{p_2} \right)^{1-\alpha} m$.

Ejemplo 3

Una empresa posee una función de producción $F(K, L) = K^{\frac{1}{2}}L^{\frac{1}{4}}$, con K y L unidades de capital y trabajo, respectivamente. Si se desea producir Q unidades de producto, cuáles cantidades K^* y L^* minizan los costos?

Ejemplo 3

Una empresa posee una función de producción $F(K, L) = K^{\frac{1}{2}}L^{\frac{1}{4}}$, con K y L unidades de capital y trabajo, respectivamente. Si se desea producir Q unidades de producto, cuáles cantidades K^* y L^* minizan los costos?

Solución: mín $C(K, L) = rK + wL$ sujeta a $K^{\frac{1}{2}}L^{\frac{1}{4}} = Q$.

Ejemplo 3

Una empresa posee una función de producción $F(K, L) = K^{\frac{1}{2}}L^{\frac{1}{4}}$, con K y L unidades de capital y trabajo, respectivamente. Si se desea producir Q unidades de producto, cuáles cantidades K^* y L^* minizan los costos?

Solución: mín $C(K, L) = rK + wL$ sujeta a $K^{\frac{1}{2}}L^{\frac{1}{4}} = Q$.

$\nabla g = \frac{1}{4}L^{-\frac{3}{4}}K^{-\frac{1}{2}}(2L, K) \neq (0, 0)$ para todo $K, L \neq 0$. En $K = 0$ o $L = 0$ el gradiente no existe, pero $F(K, L) = 0 \neq Q$, por lo que no es admisible. Concluimos que todos los puntos admisibles satisfacen NDCQ.

Ejemplo 3

Una empresa posee una función de producción $F(K, L) = K^{\frac{1}{2}}L^{\frac{1}{4}}$, con K y L unidades de capital y trabajo, respectivamente. Si se desea producir Q unidades de producto, cuáles cantidades K^* y L^* minizan los costos?

Solución: mín $C(K, L) = rK + wL$ sujeta a $K^{\frac{1}{2}}L^{\frac{1}{4}} = Q$.

$\nabla g = \frac{1}{4}L^{-\frac{3}{4}}K^{-\frac{1}{2}}(2L, K) \neq (0, 0)$ para todo $K, L \neq 0$. En $K = 0$ o $L = 0$ el gradiente no existe, pero $F(K, L) = 0 \neq Q$, por lo que no es admisible. Concluimos que todos los puntos admisibles satisfacen NDCQ.

Planteando el Lagrangiano $\mathcal{L} = rK + wL - \lambda \left(K^{\frac{1}{2}}L^{\frac{1}{4}} - Q \right)$, tenemos que

$$\nabla \mathcal{L} = \begin{pmatrix} Q - K^{\frac{1}{2}}L^{\frac{1}{4}} \\ r - \lambda \frac{1}{2}K^{-\frac{1}{2}}L^{\frac{1}{4}} \\ w - \lambda \frac{1}{4}K^{\frac{1}{2}}L^{-\frac{3}{4}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ K^* \\ L^* \end{pmatrix} = \begin{pmatrix} \sqrt[3]{16r^2wQ} \\ \sqrt[3]{\frac{2wQ^4}{r}} \\ \sqrt[3]{\frac{r^2Q^4}{4w^2}} \end{pmatrix}.$$

Ejemplo 3

Una empresa posee una función de producción $F(K, L) = K^{\frac{1}{2}}L^{\frac{1}{4}}$, con K y L unidades de capital y trabajo, respectivamente. Si se desea producir Q unidades de producto, cuáles cantidades K^* y L^* minizan los costos?

Solución: mín $C(K, L) = rK + wL$ sujeta a $K^{\frac{1}{2}}L^{\frac{1}{4}} = Q$.

$\nabla g = \frac{1}{4}L^{-\frac{3}{4}}K^{-\frac{1}{2}}(2L, K) \neq (0, 0)$ para todo $K, L \neq 0$. En $K = 0$ o $L = 0$ el gradiente no existe, pero $F(K, L) = 0 \neq Q$, por lo que no es admisible. Concluimos que todos los puntos admisibles satisfacen NDCQ.

Planteando el Lagrangiano $\mathcal{L} = rK + wL - \lambda \left(K^{\frac{1}{2}}L^{\frac{1}{4}} - Q \right)$, tenemos que

$$\nabla \mathcal{L} = \begin{pmatrix} Q - K^{\frac{1}{2}}L^{\frac{1}{4}} \\ r - \lambda \frac{1}{2}K^{-\frac{1}{2}}L^{\frac{1}{4}} \\ w - \lambda \frac{1}{4}K^{\frac{1}{2}}L^{-\frac{3}{4}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ K^* \\ L^* \end{pmatrix} = \begin{pmatrix} \sqrt[3]{16r^2wQ} \\ \sqrt[3]{\frac{2wQ^4}{r}} \\ \sqrt[3]{\frac{r^2Q^4}{4w^2}} \end{pmatrix}.$$

f es convexa por ser lineal. g es cóncava (Cobb-Douglas con $\alpha_1 + \alpha_2 < 1$) y $\lambda^* > 0$. Entonces $\mathcal{L}(\lambda^*, \cdot)$ es convexa, i.e., $C(K^*, L^*) = 34^{-\frac{1}{3}}r^{\frac{2}{3}}w^{\frac{1}{3}}Q^{\frac{4}{3}}$ es el costo mínimo absoluto de producir Q unidades.

PRECAUCIÓN

La condición del teorema es suficiente pero no necesaria, es decir, x^* puede ser un óptimo global pero $\mathcal{L}(\lambda^*, \cdot)$ no ser cóncava ni convexa.

PRECAUCIÓN

La condición del teorema es suficiente pero no necesaria, es decir, $\mathbf{x}^*$ puede ser un óptimo global pero $\mathcal{L}(\boldsymbol{\lambda}^*, \cdot)$ no ser cóncava ni convexa.

Considere el problema: $\text{máx } 2x_1x_2 + 3x_1$ sujeta a $x_1 + 2x_2 = 83$.

PRECAUCIÓN

La condición del teorema es suficiente pero no necesaria, es decir, $\mathbf{x}^*$ puede ser un óptimo global pero $\mathcal{L}(\boldsymbol{\lambda}^*, \cdot)$ no ser cóncava ni convexa.

Considere el problema: $\text{máx } 2x_1x_2 + 3x_1$ sujeta a $x_1 + 2x_2 = 83$.

Como $\nabla g = (1, 2) \neq (0, 0)$, todo (x_1, x_2) satisface NDCQ.

PRECAUCIÓN

La condición del teorema es suficiente pero no necesaria, es decir, $\mathbf{x}^*$ puede ser un óptimo global pero $\mathcal{L}(\boldsymbol{\lambda}^*, \cdot)$ no ser cóncava ni convexa.

Considere el problema: $\text{máx } 2x_1x_2 + 3x_1$ sujeta a $x_1 + 2x_2 = 83$.

Como $\nabla g = (1, 2) \neq (0, 0)$, todo (x_1, x_2) satisface NDCQ.

Para el Lagrangiano $\mathcal{L} = 2x_1x_2 + 3x_1 - \lambda(x_1 + 2x_2 - 83)$ tenemos,

$$\nabla \mathcal{L}(\lambda, x_1, x_2) = \begin{pmatrix} 83 - (x_1 + 2x_2) \\ 2x_2 + 3 - \lambda \\ 2x_1 - 2\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} 43 \\ 43 \\ 20 \end{pmatrix}.$$

PRECAUCIÓN

La condición del teorema es suficiente pero no necesaria, es decir, $\mathbf{x}^*$ puede ser un óptimo global pero $\mathcal{L}(\boldsymbol{\lambda}^*, \cdot)$ no ser cóncava ni convexa.

Considere el problema: $\text{máx } 2x_1x_2 + 3x_1$ sujeta a $x_1 + 2x_2 = 83$.

Como $\nabla g = (1, 2) \neq (0, 0)$, todo (x_1, x_2) satisface NDCQ.

Para el Lagrangiano $\mathcal{L} = 2x_1x_2 + 3x_1 - \lambda(x_1 + 2x_2 - 83)$ tenemos,

$$\nabla \mathcal{L}(\lambda, x_1, x_2) = \begin{pmatrix} 83 - (x_1 + 2x_2) \\ 2x_2 + 3 - \lambda \\ 2x_1 - 2\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} 43 \\ 43 \\ 20 \end{pmatrix}.$$

Por otra parte,

$$H_{\mathcal{L}(\lambda^*, \cdot, \cdot)}(x_1, x_2) = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \implies (-1)^2 D_2 = -4 < 0,$$

por lo que $\mathcal{L}(\lambda^*, x_1, x_2)$ no es cóncava (ni convexa), y no podemos utilizar el teorema para concluir que $f(43, 20)$ es un máximo global.

PRECAUCIÓN

La condición del teorema es suficiente pero no necesaria, es decir, $\mathbf{x}^*$ puede ser un óptimo global pero $\mathcal{L}(\boldsymbol{\lambda}^*, \cdot)$ no ser cóncava ni convexa.

Considere el problema: $\text{máx } 2x_1x_2 + 3x_1$ sujeta a $x_1 + 2x_2 = 83$.

Como $\nabla g = (1, 2) \neq (0, 0)$, todo (x_1, x_2) satisface NDCQ.

Para el Lagrangiano $\mathcal{L} = 2x_1x_2 + 3x_1 - \lambda(x_1 + 2x_2 - 83)$ tenemos,

$$\nabla \mathcal{L}(\lambda, x_1, x_2) = \begin{pmatrix} 83 - (x_1 + 2x_2) \\ 2x_2 + 3 - \lambda \\ 2x_1 - 2\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} 43 \\ 43 \\ 20 \end{pmatrix}.$$

Por otra parte,

$$H_{\mathcal{L}(\lambda^*, \cdot, \cdot)}(x_1, x_2) = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \implies (-1)^2 D_2 = -4 < 0,$$

por lo que $\mathcal{L}(\lambda^*, x_1, x_2)$ no es cóncava (ni convexa), y no podemos utilizar el teorema para concluir que $f(43, 20)$ es un máximo global.

El problema equivalente (sin restricciones) es $\text{máx } x_1(86 - x_1)$. Como $\tilde{f}(x_1)$ es cuadrática, alcanza su máximo global en $x_1 = 43$, es decir, $f(43, 20)$ es un máximo global (aunque $\mathcal{L}$ no sea cóncava).

Hessiana Orlada Generalizada

Matriz Hessiana Orlada Generalizada

Sea f una función dos veces continuamente diferenciable en $\mathbb{R}^n$ y sean g_i funciones continuamente diferenciables en $\mathbb{R}^n$, para $i = 1, \dots, m$.

La *matriz Hessiana orlada generalizada* de f con las restricciones $\mathbf{g}$ es

$$H_{f,\mathbf{g}}^O(\boldsymbol{\lambda}, \mathbf{x}) := H_{\mathcal{L}}(\boldsymbol{\lambda}, \mathbf{x}) \quad (\text{Hessiana de la función Lagrangiana})$$

Hessiana Orlada Generalizada

Matriz Hessiana Orlada Generalizada

Sea f una función dos veces continuamente diferenciable en $\mathbb{R}^n$ y sean g_i funciones continuamente diferenciables en $\mathbb{R}^n$, para $i = 1, \dots, m$.

La *matriz Hessiana orlada generalizada* de f con las restricciones $\mathbf{g}$ es

$$H_{f,\mathbf{g}}^O(\boldsymbol{\lambda}, \mathbf{x}) := H_{\mathcal{L}}(\boldsymbol{\lambda}, \mathbf{x}) \quad (\text{Hessiana de la función Lagrangiana})$$

Observaciones:

- $H_{f,\mathbf{g}}^O$ es una matriz de tamaño $(n + m) \times (n + m)$.

Hessiana Orlada Generalizada

Matriz Hessiana Orlada Generalizada

Sea f una función dos veces continuamente diferenciable en $\mathbb{R}^n$ y sean g_i funciones continuamente diferenciables en $\mathbb{R}^n$, para $i = 1, \dots, m$.

La *matriz Hessiana orlada generalizada* de f con las restricciones $\mathbf{g}$ es

$$H_{f,\mathbf{g}}^O(\boldsymbol{\lambda}, \mathbf{x}) := H_{\mathcal{L}}(\boldsymbol{\lambda}, \mathbf{x}) \quad (\text{Hessiana de la función Lagrangiana})$$

Observaciones:

- $H_{f,\mathbf{g}}^O$ es una matriz de tamaño $(n + m) \times (n + m)$.
- Como $\mathcal{L}_{\lambda_i \lambda_j} = 0$, para todo $i, j = 1, \dots, m$, entonces $H_{f,\mathbf{g}}^O$ tiene una submatriz $\mathbf{0}_{m \times m}$ de ceros en la esquina superior izquierda.

Hessiana Orlada Generalizada

Matriz Hessiana Orlada Generalizada

Sea f una función dos veces continuamente diferenciable en $\mathbb{R}^n$ y sean g_i funciones continuamente diferenciables en $\mathbb{R}^n$, para $i = 1, \dots, m$.

La *matriz Hessiana orlada generalizada* de f con las restricciones $\mathbf{g}$ es

$$H_{f,\mathbf{g}}^O(\boldsymbol{\lambda}, \mathbf{x}) := H_{\mathcal{L}}(\boldsymbol{\lambda}, \mathbf{x}) \quad (\text{Hessiana de la función Lagrangiana})$$

Observaciones:

- $H_{f,\mathbf{g}}^O$ es una matriz de tamaño $(n + m) \times (n + m)$.
- Como $\mathcal{L}_{\lambda_i \lambda_j} = 0$, para todo $i, j = 1, \dots, m$, entonces $H_{f,\mathbf{g}}^O$ tiene una submatriz $\mathbf{0}_{m \times m}$ de ceros en la esquina superior izquierda.
- Como $\mathcal{L}_{\lambda_i x_j} = -\partial_{x_j} g_i$, entonces $-J_{\mathbf{g}}$ se encuentra en la esquina superior derecha y $-J_{\mathbf{g}}^T$ en la esquina inferior izquierda.

Hessiana Orlada Generalizada

Matriz Hessiana Orlada Generalizada

Sea f una función dos veces continuamente diferenciable en $\mathbb{R}^n$ y sean g_i funciones continuamente diferenciables en $\mathbb{R}^n$, para $i = 1, \dots, m$.

La *matriz Hessiana orlada generalizada* de f con las restricciones $\mathbf{g}$ es

$$H_{f,\mathbf{g}}^O(\boldsymbol{\lambda}, \mathbf{x}) := H_{\mathcal{L}}(\boldsymbol{\lambda}, \mathbf{x}) \quad (\text{Hessiana de la función Lagrangiana})$$

Observaciones:

- $H_{f,\mathbf{g}}^O$ es una matriz de tamaño $(n + m) \times (n + m)$.
- Como $\mathcal{L}_{\lambda_i \lambda_j} = 0$, para todo $i, j = 1, \dots, m$, entonces $H_{f,\mathbf{g}}^O$ tiene una submatriz $\mathbf{0}_{m \times m}$ de ceros en la esquina superior izquierda.
- Como $\mathcal{L}_{\lambda_i x_j} = -\partial_{x_j} g_i$, entonces $-J_{\mathbf{g}}$ se encuentra en la esquina superior derecha y $-J_{\mathbf{g}}^T$ en la esquina inferior izquierda.
- Como $\mathcal{L}_{x_i x_j} = \partial_{x_i x_j} f - \sum_{k=1}^m \lambda_k \partial_{x_i x_j} g_k$, entonces $H_f - \sum_{k=1}^m \lambda_k H_{g_k}$.

Hessiana Orlada Generalizada

Matriz Hessiana Orlada Generalizada

Sea f una función dos veces continuamente diferenciable en $\mathbb{R}^n$ y sean g_i funciones continuamente diferenciables en $\mathbb{R}^n$, para $i = 1, \dots, m$.

La *matriz Hessiana orlada generalizada* de f con las restricciones $\mathbf{g}$ es

$$H_{f,\mathbf{g}}^O(\boldsymbol{\lambda}, \mathbf{x}) := H_{\mathcal{L}}(\boldsymbol{\lambda}, \mathbf{x}) \quad (\text{Hessiana de la función Lagrangiana})$$

Observaciones:

- $H_{f,\mathbf{g}}^O$ es una matriz de tamaño $(n + m) \times (n + m)$.
- Como $\mathcal{L}_{\lambda_i \lambda_j} = 0$, para todo $i, j = 1, \dots, m$, entonces $H_{f,\mathbf{g}}^O$ tiene una submatriz $\mathbf{0}_{m \times m}$ de ceros en la esquina superior izquierda.
- Como $\mathcal{L}_{\lambda_i x_j} = -\partial_{x_j} g_i$, entonces $-J_{\mathbf{g}}$ se encuentra en la esquina superior derecha y $-J_{\mathbf{g}}^T$ en la esquina inferior izquierda.
- Como $\mathcal{L}_{x_i x_j} = \partial_{x_i x_j} f - \sum_{k=1}^m \lambda_k \partial_{x_i x_j} g_k$, entonces $H_f - \sum_{k=1}^m \lambda_k H_{g_k}$.

En conclusión,

$$H_{f,\mathbf{g}}^O(\boldsymbol{\lambda}, \mathbf{x}) = \begin{pmatrix} \mathbf{0}_{m \times m} & -J_{\mathbf{g}}(\mathbf{x}) \\ -J_{\mathbf{g}}^T(\mathbf{x}) & H_f(\mathbf{x}) - \sum_{k=1}^m \lambda_k H_{g_k}(\mathbf{x}) \end{pmatrix}$$

Condición Suficiente para Máximo-Mínimo Local

Menores Principales Dominantes de $H_{f,g}^O$

Los *menores principales dominantes* de orden r son los determinantes:

$$D_r^O(\boldsymbol{\lambda}, \mathbf{x}) = \begin{vmatrix} 0 & \cdots & 0 & -\partial_{x_1} g_1(\mathbf{x}) & \cdots & -\partial_{x_r} g_1(\mathbf{x}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & -\partial_{x_1} g_m(\mathbf{x}) & \cdots & -\partial_{x_r} g_m(\mathbf{x}) \\ -\partial_{x_1} g_1(\mathbf{x}) & \cdots & -\partial_{x_1} g_m(\mathbf{x}) & \partial_{x_1 x_1}^2 \mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}) & \cdots & \partial_{x_1 x_r}^2 \mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\partial_{x_r} g_1(\mathbf{x}) & \cdots & -\partial_{x_r} g_m(\mathbf{x}) & \partial_{x_1 x_r}^2 \mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}) & \cdots & \partial_{x_r x_r}^2 \mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}) \end{vmatrix}$$

Condición Suficiente para Máximo-Mínimo Local

Menores Principales Dominantes de $H_{f,g}^O$

Los *menores principales dominantes* de orden r son los determinantes:

$$D_r^O(\lambda, x) = \begin{vmatrix} 0 & \cdots & 0 & -\partial_{x_1} g_1(x) & \cdots & -\partial_{x_r} g_1(x) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & -\partial_{x_1} g_m(x) & \cdots & -\partial_{x_r} g_m(x) \\ -\partial_{x_1} g_1(x) & \cdots & -\partial_{x_1} g_m(x) & \partial_{x_1 x_1}^2 \mathcal{L}(\lambda, x) & \cdots & \partial_{x_1 x_r}^2 \mathcal{L}(\lambda, x) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ -\partial_{x_r} g_1(x) & \cdots & -\partial_{x_r} g_m(x) & \partial_{x_1 x_r}^2 \mathcal{L}(\lambda, x) & \cdots & \partial_{x_r x_r}^2 \mathcal{L}(\lambda, x) \end{vmatrix}$$

Condición de Segundo Orden para Máximo-Mínimo Local en $\mathbb{R}^n$

Sean $f, g \in \mathcal{C}^2(S)$. Si $x^* \in S$ cumple las condiciones de primer orden,

- (a) $f(x^*)$ es un máximo local si $(-1)^r D_r^O(\lambda^*, x^*) > 0$, $r = m+1, \dots, n$.
- (b) $f(x^*)$ es un mínimo local si $(-1)^m D_r^O(\lambda^*, x^*) > 0$, $r = m+1, \dots, n$.
- (c) x^* es un punto de silla si $D_n^O(\lambda^*, x^*) \neq 0$ y ((a),(b)) no se cumplen.
- (d) No se puede concluir si $D_n^O(\lambda^*, x^*) = 0$.

Ejemplo 1: $n = 2$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2$ sujeta a $x^2 + xy + y^2 = 3$.

Ejemplo 1: $n = 2$ y $m = 1$

Resuelva el problema: máx(mín) $x^2 + y^2$ sujeta a $x^2 + xy + y^2 = 3$.

Solución: $\nabla g = (2x + y, 2y + x) = (0, 0)$ solo si $x = 0 = y$, pero $g(0, 0) = 0 \neq 3$, i.e., no es admisible. Todo x admisible satisface NDCQ.

Ejemplo 1: $n = 2$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2$ sujeta a $x^2 + xy + y^2 = 3$.

Solución: $\nabla g = (2x + y, 2y + x) = (0, 0)$ solo si $x = 0 = y$, pero $g(0, 0) = 0 \neq 3$, i.e., no es admisible. Todo x admisible satisface NDCQ.

Para el Lagrangiano $\mathcal{L} = x^2 + y^2 - \lambda(x^2 + xy + y^2 - 3)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 3 - (x^2 + xy + y^2) \\ 2x - \lambda(2x + y) \\ 2y - \lambda(x + 2y) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ x^* \\ y^* \end{pmatrix} = \left\{ \begin{pmatrix} \frac{2}{3} \\ \pm 1 \\ \pm 1 \end{pmatrix} ; \begin{pmatrix} 2 \\ \pm \sqrt{3} \\ \mp \sqrt{3} \end{pmatrix} \right\}.$$

Ejemplo 1: $n = 2$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2$ sujeta a $x^2 + xy + y^2 = 3$.

Solución: $\nabla g = (2x + y, 2y + x) = (0, 0)$ solo si $x = 0 = y$, pero $g(0, 0) = 0 \neq 3$, i.e., no es admisible. Todo x admisible satisface NDCQ.

Para el Lagrangiano $\mathcal{L} = x^2 + y^2 - \lambda(x^2 + xy + y^2 - 3)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 3 - (x^2 + xy + y^2) \\ 2x - \lambda(2x + y) \\ 2y - \lambda(x + 2y) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ x^* \\ y^* \end{pmatrix} = \left\{ \begin{pmatrix} \frac{2}{3} \\ \pm 1 \\ \pm 1 \end{pmatrix} ; \begin{pmatrix} 2 \\ \pm \sqrt{3} \\ \mp \sqrt{3} \end{pmatrix} \right\}.$$

Por otra parte,

$$H_{f,g}^O = \begin{pmatrix} 0 & -(2x + y) & -(2y + x) \\ -(2x + y) & 2 - 2\lambda & -\lambda \\ -(2y + x) & -\lambda & 2 - 2\lambda \end{pmatrix},$$

$$\Rightarrow D_2^O = -2(\lambda(2x + y)(x + 2y) + (1 - \lambda)((2x + y)^2 + (x + 2y)^2)).$$

Ejemplo 1: $n = 2$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2$ sujeta a $x^2 + xy + y^2 = 3$.

Solución: $\nabla g = (2x + y, 2y + x) = (0, 0)$ solo si $x = 0 = y$, pero $g(0, 0) = 0 \neq 3$, i.e., no es admisible. Todo x admisible satisface NDCQ.

Para el Lagrangiano $\mathcal{L} = x^2 + y^2 - \lambda(x^2 + xy + y^2 - 3)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 3 - (x^2 + xy + y^2) \\ 2x - \lambda(2x + y) \\ 2y - \lambda(x + 2y) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ x^* \\ y^* \end{pmatrix} = \left\{ \begin{pmatrix} \frac{2}{3} \\ \pm 1 \\ \pm 1 \end{pmatrix}; \begin{pmatrix} 2 \\ \pm \sqrt{3} \\ \mp \sqrt{3} \end{pmatrix} \right\}.$$

Por otra parte,

$$H_{f,g}^O = \begin{pmatrix} 0 & -(2x + y) & -(2y + x) \\ -(2x + y) & 2 - 2\lambda & -\lambda \\ -(2y + x) & -\lambda & 2 - 2\lambda \end{pmatrix},$$

$$\implies D_2^O = -2(\lambda(2x + y)(x + 2y) + (1 - \lambda)((2x + y)^2 + (x + 2y)^2)).$$

- $(2, \pm\sqrt{3}, \mp\sqrt{3}) : (-1)^2 D_2^O = 216 > 0$, son máximos locales.

Ejemplo 1: $n = 2$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2$ sujeta a $x^2 + xy + y^2 = 3$.

Solución: $\nabla g = (2x + y, 2y + x) = (0, 0)$ solo si $x = 0 = y$, pero $g(0, 0) = 0 \neq 3$, i.e., no es admisible. Todo x admisible satisface NDCQ.

Para el Lagrangiano $\mathcal{L} = x^2 + y^2 - \lambda(x^2 + xy + y^2 - 3)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 3 - (x^2 + xy + y^2) \\ 2x - \lambda(2x + y) \\ 2y - \lambda(x + 2y) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ x^* \\ y^* \end{pmatrix} = \left\{ \begin{pmatrix} \frac{2}{3} \\ \pm 1 \\ \pm 1 \end{pmatrix}; \begin{pmatrix} 2 \\ \pm \sqrt{3} \\ \mp \sqrt{3} \end{pmatrix} \right\}.$$

Por otra parte,

$$H_{f,g}^O = \begin{pmatrix} 0 & -(2x + y) & -(2y + x) \\ -(2x + y) & 2 - 2\lambda & -\lambda \\ -(2y + x) & -\lambda & 2 - 2\lambda \end{pmatrix},$$

$$\Rightarrow D_2^O = -2(\lambda(2x + y)(x + 2y) + (1 - \lambda)((2x + y)^2 + (x + 2y)^2)).$$

- $(2, \pm\sqrt{3}, \mp\sqrt{3}) : (-1)^2 D_2^O = 216 > 0$, son máximos locales.
- $(\frac{2}{3}, \pm 1, \pm 1) : (-1)^1 D_2^O = 24 > 0$, por lo que son mínimos locales.

Ejemplo 1: $n = 2$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2$ sujeta a $x^2 + xy + y^2 = 3$.

Solución: $\nabla g = (2x + y, 2y + x) = (0, 0)$ solo si $x = 0 = y$, pero $g(0, 0) = 0 \neq 3$, i.e., no es admisible. Todo x admisible satisface NDCQ.

Para el Lagrangiano $\mathcal{L} = x^2 + y^2 - \lambda(x^2 + xy + y^2 - 3)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 3 - (x^2 + xy + y^2) \\ 2x - \lambda(2x + y) \\ 2y - \lambda(x + 2y) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \begin{pmatrix} \lambda^* \\ x^* \\ y^* \end{pmatrix} = \left\{ \begin{pmatrix} \frac{2}{3} \\ \pm 1 \\ \pm 1 \end{pmatrix}; \begin{pmatrix} 2 \\ \pm\sqrt{3} \\ \mp\sqrt{3} \end{pmatrix} \right\}.$$

Por otra parte,

$$H_{f,g}^O = \begin{pmatrix} 0 & -(2x + y) & -(2y + x) \\ -(2x + y) & 2 - 2\lambda & -\lambda \\ -(2y + x) & -\lambda & 2 - 2\lambda \end{pmatrix},$$

$$\implies D_2^O = -2(\lambda(2x + y)(x + 2y) + (1 - \lambda)((2x + y)^2 + (x + 2y)^2)).$$

- $(2, \pm\sqrt{3}, \mp\sqrt{3}) : (-1)^2 D_2^O = 216 > 0$, son máximos locales.
- $(\frac{2}{3}, \pm 1, \pm 1) : (-1)^1 D_2^O = 24 > 0$, por lo que son mínimos locales.

NOTA: Como $r = m + 1 = 1 + 1 = 2 = n$, solo tenemos que calcular D_2^O .

Ejemplo 2: $n = 3$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2 + z^2$ sujeta a $z - xy = 2$.

Ejemplo 2: $n = 3$ y $m = 1$

Resuelva el problema: máx(mín) $x^2 + y^2 + z^2$ sujeta a $z - xy = 2$.

Solución: $\nabla g = (-y, -x, 1) \neq (0, 0, 0)$, i.e., todos califican la restricción.

Ejemplo 2: $n = 3$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2 + z^2$ sujeta a $z - xy = 2$.

Solución: $\nabla g = (-y, -x, 1) \neq (0, 0, 0)$, i.e., todos califican la restricción.

Para el Lagrangiano $\mathcal{L} = x^2 + y^2 + z^2 - \lambda(z - xy - 2)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 2 - (z - xy) \\ 2x + y\lambda \\ 2y + x\lambda \\ 2z - \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \lambda^* \\ x^* \\ y^* \\ z^* \end{pmatrix} = \left\{ \begin{pmatrix} 2 \\ \mp 1 \\ \pm 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

Ejemplo 2: $n = 3$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2 + z^2$ sujeta a $z - xy = 2$.

Solución: $\nabla g = (-y, -x, 1) \neq (0, 0, 0)$, i.e., todos califican la restricción.

Para el Lagrangiano $\mathcal{L} = x^2 + y^2 + z^2 - \lambda(z - xy - 2)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 2 - (z - xy) \\ 2x + y\lambda \\ 2y + x\lambda \\ 2z - \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \lambda^* \\ x^* \\ y^* \\ z^* \end{pmatrix} = \left\{ \begin{pmatrix} 2 \\ \mp 1 \\ \pm 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

Entonces,

$$H_{f,g}^O = \begin{pmatrix} 0 & y & x & -1 \\ y & 2 & \lambda & 0 \\ x & \lambda & 2 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix} \Rightarrow \begin{aligned} D_2^O &= -2(x^2 - xy\lambda + y^2) \\ D_3^O &= -4(x^2 - xy\lambda + y^2 + 1) + \lambda^2 \end{aligned}$$

Ejemplo 2: $n = 3$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2 + z^2$ sujeta a $z - xy = 2$.

Solución: $\nabla g = (-y, -x, 1) \neq (0, 0, 0)$, i.e., todos califican la restricción.

Para el Lagrangiano $\mathcal{L} = x^2 + y^2 + z^2 - \lambda(z - xy - 2)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 2 - (z - xy) \\ 2x + y\lambda \\ 2y + x\lambda \\ 2z - \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \lambda^* \\ x^* \\ y^* \\ z^* \end{pmatrix} = \left\{ \begin{pmatrix} 2 \\ \mp 1 \\ \pm 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

Entonces, $H_{f,g}^O = \begin{pmatrix} 0 & y & x & -1 \\ y & 2 & \lambda & 0 \\ x & \lambda & 2 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix} \Rightarrow \begin{aligned} D_2^O &= -2(x^2 - xy\lambda + y^2) \\ D_3^O &= -4(x^2 - xy\lambda + y^2 + 1) + \lambda^2 \end{aligned}$

• $(2, \mp 1, \pm 1, 1)$: Como $(-1)^1 D_2^O = 8 > 0$ y $(-1)^1 D_3^O = 16 > 0$, entonces ambos puntos son mínimos locales.

Ejemplo 2: $n = 3$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2 + z^2$ sujeta a $z - xy = 2$.

Solución: $\nabla g = (-y, -x, 1) \neq (0, 0, 0)$, i.e., todos califican la restricción.

Para el Lagrangiano $\mathcal{L} = x^2 + y^2 + z^2 - \lambda(z - xy - 2)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 2 - (z - xy) \\ 2x + y\lambda \\ 2y + x\lambda \\ 2z - \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \lambda^* \\ x^* \\ y^* \\ z^* \end{pmatrix} = \left\{ \begin{pmatrix} 2 \\ \mp 1 \\ \pm 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

Entonces, $H_{f,g}^O = \begin{pmatrix} 0 & y & x & -1 \\ y & 2 & \lambda & 0 \\ x & \lambda & 2 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix} \Rightarrow \begin{aligned} D_2^O &= -2(x^2 - xy\lambda + y^2) \\ D_3^O &= -4(x^2 - xy\lambda + y^2 + 1) + \lambda^2 \end{aligned}$

- $(2, \mp 1, \pm 1, 1)$: Como $(-1)^1 D_2^O = 8 > 0$ y $(-1)^1 D_3^O = 16 > 0$, entonces ambos puntos son mínimos locales.
- $(4, 0, 0, 2)$: Punto de silla pues $(-1)^1 D_3^O = (-1)^3 D_3^O = -12 < 0$.

Ejemplo 2: $n = 3$ y $m = 1$

Resuelva el problema: $\text{máx}(\text{mín}) \ x^2 + y^2 + z^2$ sujeta a $z - xy = 2$.

Solución: $\nabla g = (-y, -x, 1) \neq (0, 0, 0)$, i.e., todos califican la restricción.

Para el Lagrangiano $\mathcal{L} = x^2 + y^2 + z^2 - \lambda(z - xy - 2)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 2 - (z - xy) \\ 2x + y\lambda \\ 2y + x\lambda \\ 2z - \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \lambda^* \\ x^* \\ y^* \\ z^* \end{pmatrix} = \left\{ \begin{pmatrix} 2 \\ \mp 1 \\ \pm 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

Entonces,

$$H_{f,g}^O = \begin{pmatrix} 0 & y & x & -1 \\ y & 2 & \lambda & 0 \\ x & \lambda & 2 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix} \Rightarrow \begin{aligned} D_2^O &= -2(x^2 - xy\lambda + y^2) \\ D_3^O &= -4(x^2 - xy\lambda + y^2 + 1) + \lambda^2 \end{aligned}$$

- $(2, \mp 1, \pm 1, 1)$: Como $(-1)^1 D_2^O = 8 > 0$ y $(-1)^1 D_3^O = 16 > 0$, entonces ambos puntos son mínimos locales.
- $(4, 0, 0, 2)$: Punto de silla pues $(-1)^1 D_3^O = (-1)^3 D_3^O = -12 < 0$.

NOTA: Como $r = m + 1 = 1 + 1 = 2 < 3 = n$, debemos calcular D_2^O, D_3^O .

Ejemplo 3: $n = 3$ y $m = 2$

Resuelva el problema: $\text{máx}(\text{mín}) \ yz + xz$ sujeta a $y^2 + z^2 = 1, \ xz = 3$.

Ejemplo 3: $n = 3$ y $m = 2$

Resuelva el problema: máx(mín) $yz + xz$ sujeta a $y^2 + z^2 = 1$, $xz = 3$.

Solución:

$$J_g = \begin{pmatrix} 0 & 2y & 2z \\ z & 0 & x \end{pmatrix} \implies \begin{pmatrix} 0 \\ 2y \\ 2z \end{pmatrix} = \alpha \begin{pmatrix} z \\ 0 \\ x \end{pmatrix} \implies z = 0.$$

$g_2(x, y, 0) = 0 \neq 3$, i.e., no es admisible. Todos satisfacen NDCQ.

Ejemplo 3: $n = 3$ y $m = 2$

Resuelva el problema: $\text{máx}(\text{mín}) \ yz + xz$ sujeta a $y^2 + z^2 = 1$, $xz = 3$.

Solución:

$$J_g = \begin{pmatrix} 0 & 2y & 2z \\ z & 0 & x \end{pmatrix} \implies \begin{pmatrix} 0 \\ 2y \\ 2z \end{pmatrix} = \alpha \begin{pmatrix} z \\ 0 \\ x \end{pmatrix} \implies z = 0.$$

$g_2(x, y, 0) = 0 \neq 3$, i.e., no es admisible. Todos satisfacen NDCQ.

Para el Lagrangiano $\mathcal{L} = yz + xz - \lambda_1(y^2 + z^2 - 1) - \lambda_2(xz - 3)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 1 - (y^2 + z^2) \\ 3 - xz \\ z - z\lambda_2 \\ z - 2y\lambda_1 \\ x + y - 2z\lambda_1 - x\lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{matrix} (1) \\ (2) \\ (3) \\ (4) \\ (5) \end{matrix}$$

Ejemplo 3: $n = 3$ y $m = 2$

Resuelva el problema: $\text{máx}(\text{mín}) \ yz + xz$ sujeta a $y^2 + z^2 = 1, \ xz = 3$.

Solución:

$$J_g = \begin{pmatrix} 0 & 2y & 2z \\ z & 0 & x \end{pmatrix} \implies \begin{pmatrix} 0 \\ 2y \\ 2z \end{pmatrix} = \alpha \begin{pmatrix} z \\ 0 \\ x \end{pmatrix} \implies z = 0.$$

$g_2(x, y, 0) = 0 \neq 3$, i.e., no es admisible. Todos satisfacen NDCQ.

Para el Lagrangiano $\mathcal{L} = yz + xz - \lambda_1(y^2 + z^2 - 1) - \lambda_2(xz - 3)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 1 - (y^2 + z^2) \\ 3 - xz \\ z - z\lambda_2 \\ z - 2y\lambda_1 \\ x + y - 2z\lambda_1 - x\lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{matrix} (1) \\ (2) \\ (3) \\ (4) \\ (5) \end{matrix}$$

- De (3) $z = 0$ o $\lambda_2 = 1$. Como $(x, y, 0)$ no es admisible, $z \neq 0$ y $\lambda_2 = 1$.

Ejemplo 3: $n = 3$ y $m = 2$

Resuelva el problema: $\text{máx}(\text{mín}) \ yz + xz$ sujeta a $y^2 + z^2 = 1, \ xz = 3$.

Solución:

$$J_g = \begin{pmatrix} 0 & 2y & 2z \\ z & 0 & x \end{pmatrix} \implies \begin{pmatrix} 0 \\ 2y \\ 2z \end{pmatrix} = \alpha \begin{pmatrix} z \\ 0 \\ x \end{pmatrix} \implies z = 0.$$

$g_2(x, y, 0) = 0 \neq 3$, i.e., no es admisible. Todos satisfacen NDCQ.

Para el Lagrangiano $\mathcal{L} = yz + xz - \lambda_1(y^2 + z^2 - 1) - \lambda_2(xz - 3)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 1 - (y^2 + z^2) \\ 3 - xz \\ z - z\lambda_2 \\ z - 2y\lambda_1 \\ x + y - 2z\lambda_1 - x\lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{matrix} (1) \\ (2) \\ (3) \\ (4) \\ (5) \end{matrix}$$

- De (3) $z = 0$ o $\lambda_2 = 1$. Como $(x, y, 0)$ no es admisible, $z \neq 0$ y $\lambda_2 = 1$.
- Con $\lambda_2 = 1$ en (5) tenemos $y - 2z\lambda_1 = 0$. Multiplicando ésta por y y (4) por z e igualando, $z^2 = 2yz\lambda_1 = y^2$, es decir, $z = -y$ o $z = y$.

Ejemplo 3: $n = 3$ y $m = 2$

Resuelva el problema: $\text{máx}(\text{mín}) \ yz + xz$ sujeta a $y^2 + z^2 = 1$, $xz = 3$.

Solución:

$$J_g = \begin{pmatrix} 0 & 2y & 2z \\ z & 0 & x \end{pmatrix} \implies \begin{pmatrix} 0 \\ 2y \\ 2z \end{pmatrix} = \alpha \begin{pmatrix} z \\ 0 \\ x \end{pmatrix} \implies z = 0.$$

$g_2(x, y, 0) = 0 \neq 3$, i.e., no es admisible. Todos satisfacen NDCQ.

Para el Lagrangiano $\mathcal{L} = yz + xz - \lambda_1(y^2 + z^2 - 1) - \lambda_2(xz - 3)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 1 - (y^2 + z^2) \\ 3 - xz \\ z - z\lambda_2 \\ z - 2y\lambda_1 \\ x + y - 2z\lambda_1 - x\lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{matrix} (1) \\ (2) \\ (3) \\ (4) \\ (5) \end{matrix}$$

- De (3) $z = 0$ o $\lambda_2 = 1$. Como $(x, y, 0)$ no es admisible, $z \neq 0$ y $\lambda_2 = 1$.
- Con $\lambda_2 = 1$ en (5) tenemos $y - 2z\lambda_1 = 0$. Multiplicando ésta por y y (4) por z e igualando, $z^2 = 2yz\lambda_1 = y^2$, es decir, $z = -y$ o $z = y$.
- Con $0 \neq z = y$ en (4) tenemos $\lambda_1 = \frac{1}{2}$. Si $z = -y$, $\lambda_1 = -\frac{1}{2}$.

Ejemplo 3: $n = 3$ y $m = 2$

Resuelva el problema: $\text{máx}(\text{mín}) \ yz + xz$ sujeta a $y^2 + z^2 = 1$, $xz = 3$.

Solución:

$$J_g = \begin{pmatrix} 0 & 2y & 2z \\ z & 0 & x \end{pmatrix} \implies \begin{pmatrix} 0 \\ 2y \\ 2z \end{pmatrix} = \alpha \begin{pmatrix} z \\ 0 \\ x \end{pmatrix} \implies z = 0.$$

$g_2(x, y, 0) = 0 \neq 3$, i.e., no es admisible. Todos satisfacen NDCQ.

Para el Lagrangiano $\mathcal{L} = yz + xz - \lambda_1(y^2 + z^2 - 1) - \lambda_2(xz - 3)$ tenemos,

$$\nabla \mathcal{L} = \begin{pmatrix} 1 - (y^2 + z^2) \\ 3 - xz \\ z - z\lambda_2 \\ z - 2y\lambda_1 \\ x + y - 2z\lambda_1 - x\lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{matrix} (1) \\ (2) \\ (3) \\ (4) \\ (5) \end{matrix}$$

- De (3) $z = 0$ o $\lambda_2 = 1$. Como $(x, y, 0)$ no es admisible, $z \neq 0$ y $\lambda_2 = 1$.
- Con $\lambda_2 = 1$ en (5) tenemos $y - 2z\lambda_1 = 0$. Multiplicando ésta por y y (4) por z e igualando, $z^2 = 2yz\lambda_1 = y^2$, es decir, $z = -y$ o $z = y$.
- Con $0 \neq z = y$ en (4) tenemos $\lambda_1 = \frac{1}{2}$. Si $z = -y$, $\lambda_1 = -\frac{1}{2}$.
- $z = \pm y$ en (1) implica $z = \pm \frac{1}{\sqrt{2}}$. Usando esto en (2), $x = \frac{3}{z} = \pm 3\sqrt{2}$.

Ejemplo 3: $n = 3$ y $m = 2$ (Continuación)

Los puntos estacionarios son

$$\left(\frac{1}{2}, 1, \pm 3\sqrt{2}, \pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right); \quad \left(-\frac{1}{2}, 1, \pm 3\sqrt{2}, \mp \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right)$$

Ejemplo 3: $n = 3$ y $m = 2$ (Continuación)

Los puntos estacionarios son

$$\left(\frac{1}{2}, 1, \pm 3\sqrt{2}, \pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right); \quad \left(-\frac{1}{2}, 1, \pm 3\sqrt{2}, \mp \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right)$$

La Hessiana sería:

$$H_{f,g}^O = \begin{pmatrix} 0 & 0 & 0 & -2y & -2z \\ 0 & 0 & -z & 0 & -x \\ 0 & -z & 0 & 0 & (1 - \lambda_2) \\ -2y & 0 & 0 & -2\lambda_1 & 1 \\ -2z & -x & (1 - \lambda_2) & 1 & -2\lambda_1 \end{pmatrix},$$

$$\Rightarrow D_3^O = -8(yz^3 + y^2z^2\lambda_1 + z^4\lambda_1 + xy^2z(1 - \lambda_2)).$$

Ejemplo 3: $n = 3$ y $m = 2$ (Continuación)

Los puntos estacionarios son

$$\left(\frac{1}{2}, 1, \pm 3\sqrt{2}, \pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right); \quad \left(-\frac{1}{2}, 1, \pm 3\sqrt{2}, \mp \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right)$$

La Hessiana sería:

$$H_{f,g}^O = \begin{pmatrix} 0 & 0 & 0 & -2y & -2z \\ 0 & 0 & -z & 0 & -x \\ 0 & -z & 0 & 0 & (1 - \lambda_2) \\ -2y & 0 & 0 & -2\lambda_1 & 1 \\ -2z & -x & (1 - \lambda_2) & 1 & -2\lambda_1 \end{pmatrix},$$

$$\Rightarrow D_3^O = -8(yz^3 + y^2z^2\lambda_1 + z^4\lambda_1 + xy^2z(1 - \lambda_2)).$$

- $\left(\frac{1}{2}, 1, \pm 3\sqrt{2}, \pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right) : (-1)^3 D_3^O = 4 > 0$, son máximos locales.

Ejemplo 3: $n = 3$ y $m = 2$ (Continuación)

Los puntos estacionarios son

$$\left(\frac{1}{2}, 1, \pm 3\sqrt{2}, \pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right); \quad \left(-\frac{1}{2}, 1, \pm 3\sqrt{2}, \mp \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right)$$

La Hessiana sería:

$$H_{f,g}^O = \begin{pmatrix} 0 & 0 & 0 & -2y & -2z \\ 0 & 0 & -z & 0 & -x \\ 0 & -z & 0 & 0 & (1 - \lambda_2) \\ -2y & 0 & 0 & -2\lambda_1 & 1 \\ -2z & -x & (1 - \lambda_2) & 1 & -2\lambda_1 \end{pmatrix},$$

$$\Rightarrow D_3^O = -8(yz^3 + y^2z^2\lambda_1 + z^4\lambda_1 + xy^2z(1 - \lambda_2)).$$

- $\left(\frac{1}{2}, 1, \pm 3\sqrt{2}, \pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right) : (-1)^3 D_3^O = 4 > 0$, son máximos locales.
- $\left(-\frac{1}{2}, 1, \pm 3\sqrt{2}, \mp \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right) : (-1)^2 D_3^O = 4 > 0$, son mínimos locales.

Ejemplo 3: $n = 3$ y $m = 2$ (Continuación)

Los puntos estacionarios son

$$\left(\frac{1}{2}, 1, \pm 3\sqrt{2}, \pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right); \quad \left(-\frac{1}{2}, 1, \pm 3\sqrt{2}, \mp \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right)$$

La Hessiana sería:

$$H_{f,g}^O = \begin{pmatrix} 0 & 0 & 0 & -2y & -2z \\ 0 & 0 & -z & 0 & -x \\ 0 & -z & 0 & 0 & (1 - \lambda_2) \\ -2y & 0 & 0 & -2\lambda_1 & 1 \\ -2z & -x & (1 - \lambda_2) & 1 & -2\lambda_1 \end{pmatrix},$$

$$\Rightarrow D_3^O = -8(yz^3 + y^2z^2\lambda_1 + z^4\lambda_1 + xy^2z(1 - \lambda_2)).$$

- $\left(\frac{1}{2}, 1, \pm 3\sqrt{2}, \pm \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right) : (-1)^3 D_3^O = 4 > 0$, son máximos locales.
- $\left(-\frac{1}{2}, 1, \pm 3\sqrt{2}, \mp \frac{1}{\sqrt{2}}, \pm \frac{1}{\sqrt{2}}\right) : (-1)^2 D_3^O = 4 > 0$, son mínimos locales.

NOTA: Como $r = m + 1 = 2 + 1 = 3 = n$, solo tenemos que calcular D_3^O .

Definición e Interpretación de λ

Una vez obtenido el óptimo (λ^*, x^*) para el problema

$$\text{máx(mín)} \ f(x) \quad \text{sujeto a } g(x) = c,$$

tenemos que $\lambda^*(c)$ y $x^*(c)$ son funciones de la restricción c .

Definición e Interpretación de λ

Una vez obtenido el óptimo (λ^*, x^*) para el problema

$$\text{máx(mín)} \ f(x) \quad \text{sujeto a } g(x) = c,$$

tenemos que $\lambda^*(c)$ y $x^*(c)$ son funciones de la restricción c .

De la restricción $g(x^*(c)) = c$ notamos (por regla de la cadena) que

$$I_{m \times m} = J_{g|c}(c) = J_{g|x}(x^*(c))J_x(c).$$

Definición e Interpretación de λ

Una vez obtenido el óptimo (λ^*, x^*) para el problema

$$\text{máx(mín)} f(x) \quad \text{sujeto a } g(x) = c,$$

tenemos que $\lambda^*(c)$ y $x^*(c)$ son funciones de la restricción c .

De la restricción $g(x^*(c)) = c$ notamos (por regla de la cadena) que

$$I_{m \times m} = J_{g|c}(c) = J_{g|x}(x^*(c)) J_x(c).$$

Por la condición de primer orden $(0 = \nabla \mathcal{L} = \nabla f - \sum_{i=1}^m \lambda_i \nabla g_i)$,

$$J_{f|x}(x^*(c)) = (\lambda^*(c))^T J_{g|x}(x^*(c)).$$

Definición e Interpretación de λ

Una vez obtenido el óptimo (λ^*, x^*) para el problema

$$\text{máx(mín)} f(x) \quad \text{sujeto a } g(x) = c,$$

tenemos que $\lambda^*(c)$ y $x^*(c)$ son funciones de la restricción c .

De la restricción $g(x^*(c)) = c$ notamos (por regla de la cadena) que

$$I_{m \times m} = J_{g|c}(c) = J_{g|x}(x^*(c)) J_x(c).$$

Por la condición de primer orden $(0 = \nabla \mathcal{L} = \nabla f - \sum_{i=1}^m \lambda_i \nabla g_i)$,

$$J_{f|x}(x^*(c)) = (\lambda^*(c))^T J_{g|x}(x^*(c)).$$

Finalmente, por la regla de la cadena,

$$J_{f|c}(c) = J_{f|x}(x^*(c)) J_x(c) = (\lambda^*(c))^T \overbrace{J_{g|x}(x^*(c)) J_x(c)}^{= I_{m \times m}} = (\lambda^*(c))^T.$$

Definición e Interpretación de λ

Una vez obtenido el óptimo (λ^*, x^*) para el problema

$$\text{máx(mín)} f(x) \quad \text{sujeito a } g(x) = c,$$

tenemos que $\lambda^*(c)$ y $x^*(c)$ son funciones de la restricción c .

De la restricción $g(x^*(c)) = c$ notamos (por regla de la cadena) que

$$I_{m \times m} = J_{g|c}(c) = J_{g|x}(x^*(c)) J_x(c).$$

Por la condición de primer orden $(0 = \nabla \mathcal{L} = \nabla f - \sum_{i=1}^m \lambda_i \nabla g_i)$,

$$J_{f|x}(x^*(c)) = (\lambda^*(c))^T J_{g|x}(x^*(c)).$$

Finalmente, por la regla de la cadena,

$$J_{f|c}(c) = J_{f|x}(x^*(c)) J_x(c) = (\lambda^*(c))^T \overbrace{J_{g|x}(x^*(c)) J_x(c)}^{= I_{m \times m}} = (\lambda^*(c))^T.$$

Multiplicadores de Lagrange

A los valores λ_i del vector λ se le llaman “*multiplicadores de Lagrange*”.
Éstos representan el cambio en el óptimo por un cambio en la restricción, es decir, $\nabla_c f(x^(c)) = \lambda^*(c)$, o equivalentemente, $\frac{\partial f(x^*(\cdot))}{\partial c_i}(c) = \lambda_i(c)$.*

Interpretación Económica de λ

Si se incrementa la restricción en la dirección $\Delta \mathbf{c}$, cuánto cambiaría el máximo/mínimo de la función objetivo? Vimos que su cambio sería

$$D_{\Delta \mathbf{c}} f(\mathbf{x}^*(\mathbf{c})) = \nabla_{\mathbf{c}} f(\mathbf{x}^*(\mathbf{c})) \cdot \Delta \mathbf{c} = \lambda^*(\mathbf{c}) \cdot \Delta \mathbf{c}.$$

Interpretación Económica de λ

Si se incrementa la restricción en la dirección $\Delta \mathbf{c}$, cuánto cambiaría el máximo/mínimo de la función objetivo? Vimos que su cambio sería

$$D_{\Delta \mathbf{c}} f(\mathbf{x}^*(\mathbf{c})) = \nabla_{\mathbf{c}} f(\mathbf{x}^*(\mathbf{c})) \cdot \Delta \mathbf{c} = \boldsymbol{\lambda}^*(\mathbf{c}) \cdot \Delta \mathbf{c}.$$

Aproximadamente tenemos que

$$f(\mathbf{x}^*(\mathbf{c} + \Delta \mathbf{c})) - f(\mathbf{x}^*(\mathbf{c})) \approx \boldsymbol{\lambda}^*(\mathbf{c}) \cdot \Delta \mathbf{c} = \sum_{i=1}^m \lambda_i^*(\mathbf{c}) \Delta c_i.$$

En particular, si $\Delta \mathbf{c} = (0, \dots, 0, 1, 0, \dots, 0)$,

$$f(\mathbf{x}^*(\mathbf{c} + \Delta \mathbf{c})) - f(\mathbf{x}^*(\mathbf{c})) \approx \lambda_i^*(\mathbf{c}).$$

Interpretación Económica de λ

Si se incrementa la restricción en la dirección $\Delta \mathbf{c}$, cuánto cambiaría el máximo/mínimo de la función objetivo? Vimos que su cambio sería

$$D_{\Delta \mathbf{c}} f(\mathbf{x}^*(\mathbf{c})) = \nabla_{\mathbf{c}} f(\mathbf{x}^*(\mathbf{c})) \cdot \Delta \mathbf{c} = \boldsymbol{\lambda}^*(\mathbf{c}) \cdot \Delta \mathbf{c}.$$

Aproximadamente tenemos que

$$f(\mathbf{x}^*(\mathbf{c} + \Delta \mathbf{c})) - f(\mathbf{x}^*(\mathbf{c})) \approx \boldsymbol{\lambda}^*(\mathbf{c}) \cdot \Delta \mathbf{c} = \sum_{i=1}^m \lambda_i^*(\mathbf{c}) \Delta c_i.$$

En particular, si $\Delta \mathbf{c} = (0, \dots, 0, 1, 0, \dots, 0)$,

$$f(\mathbf{x}^*(\mathbf{c} + \Delta \mathbf{c})) - f(\mathbf{x}^*(\mathbf{c})) \approx \lambda_i^*(\mathbf{c}).$$

Los precios convierten un bien en términos de otro (galones de leche por dolar, manzanas por naranja, etc), por lo que λ_i convierte el aumentar la restricción por un incremento/disminución en la función objetivo, es decir, el optimizante está dispuesto a “pagar” λ_i unidades de la función objetivo para aumentar la restricción c_i en 1 unidad.

Interpretación Económica de λ

Si se incrementa la restricción en la dirección $\Delta \mathbf{c}$, cuánto cambiaría el máximo/mínimo de la función objetivo? Vimos que su cambio sería

$$D_{\Delta \mathbf{c}} f(\mathbf{x}^*(\mathbf{c})) = \nabla_{\mathbf{c}} f(\mathbf{x}^*(\mathbf{c})) \cdot \Delta \mathbf{c} = \boldsymbol{\lambda}^*(\mathbf{c}) \cdot \Delta \mathbf{c}.$$

Aproximadamente tenemos que

$$f(\mathbf{x}^*(\mathbf{c} + \Delta \mathbf{c})) - f(\mathbf{x}^*(\mathbf{c})) \approx \boldsymbol{\lambda}^*(\mathbf{c}) \cdot \Delta \mathbf{c} = \sum_{i=1}^m \lambda_i^*(\mathbf{c}) \Delta c_i.$$

En particular, si $\Delta \mathbf{c} = (0, \dots, 0, 1, 0, \dots, 0)$,

$$f(\mathbf{x}^*(\mathbf{c} + \Delta \mathbf{c})) - f(\mathbf{x}^*(\mathbf{c})) \approx \lambda_i^*(\mathbf{c}).$$

Los precios convierten un bien en términos de otro (galones de leche por dolar, manzanas por naranja, etc), por lo que λ_i convierte el aumentar la restricción por un incremento/disminución en la función objetivo, es decir, el optimizante está dispuesto a “pagar” λ_i unidades de la función objetivo para aumentar la restricción c_i en 1 unidad.

Precio Sombra (Costo Marginal de la Restricción)

El valor $\lambda_i^*(\mathbf{c})$ es el “precio sombra” del parámetro c_i .

Ejemplo 1: Maximización de Utilidad

Recordemos el ejemplo en que buscamos maximizar la utilidad

$U(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ con restricción “presupuestaria” $p_1 x_1 + p_2 x_2 = m$.

Ejemplo 1: Maximización de Utilidad

Recordemos el ejemplo en que buscamos maximizar la utilidad

$U(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ con restricción “presupuestaria” $p_1 x_1 + p_2 x_2 = m$.

Vimos que

$$x_1^*(m) = \frac{\alpha}{p_1} m, \quad x_2^*(m) = \frac{(1-\alpha)}{p_2} m, \quad \lambda^*(m) = \left(\frac{\alpha}{p_1}\right)^\alpha \left(\frac{1-\alpha}{p_2}\right)^{1-\alpha} \text{ y}$$
$$U(x_1^*(m), x_2^*(m)) = \left(\frac{\alpha}{p_1}\right)^\alpha \left(\frac{1-\alpha}{p_2}\right)^{1-\alpha} m.$$

Ejemplo 1: Maximización de Utilidad

Recordemos el ejemplo en que buscamos maximizar la utilidad

$U(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ con restricción “presupuestaria” $p_1 x_1 + p_2 x_2 = m$.

Vimos que

$$x_1^*(m) = \frac{\alpha}{p_1} m, \quad x_2^*(m) = \frac{(1-\alpha)}{p_2} m, \quad \lambda^*(m) = \left(\frac{\alpha}{p_1}\right)^\alpha \left(\frac{1-\alpha}{p_2}\right)^{1-\alpha} \text{ y}$$
$$U(x_1^*(m), x_2^*(m)) = \left(\frac{\alpha}{p_1}\right)^\alpha \left(\frac{1-\alpha}{p_2}\right)^{1-\alpha} m.$$

Se verifica que

$$\frac{\partial U(x_1^*(\cdot), x_2^*(\cdot))}{\partial m}(m) = \left(\frac{\alpha}{p_1}\right)^\alpha \left(\frac{1-\alpha}{p_2}\right)^{1-\alpha} = \lambda^*(m),$$

es decir, el aumento en la utilidad provocada por un incremento de 1 colón en el presupuesto sería aproximadamente $\lambda^*(m)$.

Ejemplo 1: Maximización de Utilidad

Recordemos el ejemplo en que buscamos maximizar la utilidad

$U(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ con restricción “presupuestaria” $p_1 x_1 + p_2 x_2 = m$.

Vimos que

$$x_1^*(m) = \frac{\alpha}{p_1} m, \quad x_2^*(m) = \frac{(1-\alpha)}{p_2} m, \quad \lambda^*(m) = \left(\frac{\alpha}{p_1}\right)^\alpha \left(\frac{1-\alpha}{p_2}\right)^{1-\alpha} \text{ y}$$

$$U(x_1^*(m), x_2^*(m)) = \left(\frac{\alpha}{p_1}\right)^\alpha \left(\frac{1-\alpha}{p_2}\right)^{1-\alpha} m.$$

Se verifica que

$$\frac{\partial U(x_1^*(\cdot), x_2^*(\cdot))}{\partial m}(m) = \left(\frac{\alpha}{p_1}\right)^\alpha \left(\frac{1-\alpha}{p_2}\right)^{1-\alpha} = \lambda^*(m),$$

es decir, el aumento en la utilidad provocada por un incremento de 1 colón en el presupuesto sería aproximadamente $\lambda^*(m)$.

Precio Sombra $\lambda^*(m)$: Se “compran” $\lambda^*(m)$ unidades de “felicidad” con el aumento de un colón en la riqueza del consumidor.

Ejemplo 2: Minimización de Costos

Recordemos el ejemplo en que buscamos minimizar los costos

$C(K, L) = rK + wL$ con la restricción de “producción” $K^{\frac{1}{2}}L^{\frac{1}{4}} = Q$.

Ejemplo 2: Minimización de Costos

Recordemos el ejemplo en que buscamos minimizar los costos

$C(K, L) = rK + wL$ con la restricción de “producción” $K^{\frac{1}{2}}L^{\frac{1}{4}} = Q$.

Vimos que

$$K^*(Q) = \sqrt[3]{\frac{2wQ^4}{r}}, L^*(Q) = \sqrt[3]{\frac{r^2Q^4}{4w^2}}, \lambda^*(Q) = \sqrt[3]{16r^2wQ} \text{ y} \\ C(K^*(Q), L^*(Q)) = 34^{-\frac{1}{3}} r^{\frac{2}{3}} w^{\frac{1}{3}} Q^{\frac{4}{3}}.$$

Ejemplo 2: Minimización de Costos

Recordemos el ejemplo en que buscamos minimizar los costos

$C(K, L) = rK + wL$ con la restricción de “producción” $K^{\frac{1}{2}}L^{\frac{1}{4}} = Q$.

Vimos que

$$K^*(Q) = \sqrt[3]{\frac{2wQ^4}{r}}, L^*(Q) = \sqrt[3]{\frac{r^2Q^4}{4w^2}}, \lambda^*(Q) = \sqrt[3]{16r^2wQ} \text{ y} \\ C(K^*(Q), L^*(Q)) = 34^{-\frac{1}{3}} r^{\frac{2}{3}} w^{\frac{1}{3}} Q^{\frac{4}{3}}.$$

Se verifica que

$$\frac{\partial C(K^*(\cdot), L^*(\cdot))}{\partial Q}(Q) = \sqrt[3]{16r^2wQ} = \lambda^*(Q),$$

es decir, el aumento en los costos provocado por un incremento de 1 unidad en la producción sería aproximadamente $\lambda^*(Q)$.

Ejemplo 2: Minimización de Costos

Recordemos el ejemplo en que buscamos minimizar los costos

$C(K, L) = rK + wL$ con la restricción de “producción” $K^{\frac{1}{2}}L^{\frac{1}{4}} = Q$.

Vimos que

$$K^*(Q) = \sqrt[3]{\frac{2wQ^4}{r}}, \quad L^*(Q) = \sqrt[3]{\frac{r^2Q^4}{4w^2}}, \quad \lambda^*(Q) = \sqrt[3]{16r^2wQ} \text{ y}$$

$$C(K^*(Q), L^*(Q)) = 34^{-\frac{1}{3}} r^{\frac{2}{3}} w^{\frac{1}{3}} Q^{\frac{4}{3}}.$$

Se verifica que

$$\frac{\partial C(K^*(\cdot), L^*(\cdot))}{\partial Q}(Q) = \sqrt[3]{16r^2wQ} = \lambda^*(Q),$$

es decir, el aumento en los costos provocado por un incremento de 1 unidad en la producción sería aproximadamente $\lambda^*(Q)$.

Precio Sombra $\lambda^*(Q)$: El “precio” a pagar por aumentar la producción en una unidad es $\lambda^*(Q)$ colones (aumento en los costos).

Resumen del Procedimiento

① Condiciones de Primer Orden:

- ① **NDCQ:** Se calcula J_g y se encuentran los puntos $\tilde{x}$ admisibles en donde este Jacobiano no existe o no tiene rango m .

Resumen del Procedimiento

① Condiciones de Primer Orden:

- ① **NDCQ:** Se calcula J_g y se encuentran los puntos $\tilde{x}$ admisibles en donde este Jacobiano no existe o no tiene rango m .
- ② **Lagrangiano:** Se plantea la función $\mathcal{L}(\lambda, x) = f(x) - \sum_{k=1}^m \lambda_k g_k(x)$, y se encuentran los puntos (λ^*, x^*) en donde $\nabla \mathcal{L}(\lambda^*, x^*) = 0$.

Resumen del Procedimiento

① Condiciones de Primer Orden:

- ① **NDCQ:** Se calcula J_g y se encuentran los puntos $\tilde{x}$ admisibles en donde este Jacobiano no existe o no tiene rango m .
- ② **Lagrangiano:** Se plantea la función $\mathcal{L}(\lambda, x) = f(x) - \sum_{k=1}^m \lambda_k g_k(x)$, y se encuentran los puntos (λ^*, x^*) en donde $\nabla \mathcal{L}(\lambda^*, x^*) = 0$.

② Condiciones de Segundo Orden: 2 opciones: máximo global o local.

Resumen del Procedimiento

1 Condiciones de Primer Orden:

- 1 **NDCQ:** Se calcula J_g y se encuentran los puntos $\tilde{x}$ admisibles en donde este Jacobiano no existe o no tiene rango m .
- 2 **Lagrangiano:** Se plantea la función $\mathcal{L}(\lambda, x) = f(x) - \sum_{k=1}^m \lambda_k g_k(x)$, y se encuentran los puntos (λ^*, x^*) en donde $\nabla \mathcal{L}(\lambda^*, x^*) = 0$.

2 Condiciones de Segundo Orden: 2 opciones: máximo global o local.

- Si $\mathcal{L}(\lambda^*, \cdot)$ es **convexa/cóncava**: Se pueden verificar los menores principales arbitrarios de $H_f(x) - \sum_{k=1}^m \lambda_k^* H_{g_k}(x)$ o verificar que la correspondiente convexidad/concavidad de las funciones f, g (tomando en cuenta el signo de λ_i), para verificar si $\mathcal{L}(\lambda^*, \cdot)$ es convexa/cóncava en todo el dominio, y el óptimo encontrado es mínimo/máximo global.

Resumen del Procedimiento

① Condiciones de Primer Orden:

- ① **NDCQ:** Se calcula J_g y se encuentran los puntos $\tilde{x}$ admisibles en donde este Jacobiano no existe o no tiene rango m .
- ② **Lagrangiano:** Se plantea la función $\mathcal{L}(\lambda, x) = f(x) - \sum_{k=1}^m \lambda_k g_k(x)$, y se encuentran los puntos (λ^*, x^*) en donde $\nabla \mathcal{L}(\lambda^*, x^*) = 0$.

② Condiciones de Segundo Orden: 2 opciones: máximo global o local.

- **Si $\mathcal{L}(\lambda^*, \cdot)$ es convexa/cóncava:** Se pueden verificar los menores principales arbitrarios de $H_f(x) - \sum_{k=1}^m \lambda_k^* H_{g_k}(x)$ o verificar que la correspondiente convexidad/concavidad de las funciones f, g (tomando en cuenta el signo de λ_i), para verificar si $\mathcal{L}(\lambda^*, \cdot)$ es convexa/cóncava en todo el dominio, y el óptimo encontrado es mínimo/máximo global.
- **Si $\mathcal{L}(\lambda^*, \cdot)$ no es convexa/cóncava:** Primero se calcula la matriz Hessiana $H_{f,g}^O$. Se verifican los menores principales dominantes en los puntos encontrados $D_r^O(\lambda^*, x^*)$, para verificar si el óptimo encontrado es mínimo/máximo local o punto de silla (o si el criterio no decide).

Resumen del Procedimiento

1 Condiciones de Primer Orden:

- 1 **NDCQ:** Se calcula J_g y se encuentran los puntos $\tilde{x}$ admisibles en donde este Jacobiano no existe o no tiene rango m .
- 2 **Lagrangiano:** Se plantea la función $\mathcal{L}(\lambda, x) = f(x) - \sum_{k=1}^m \lambda_k g_k(x)$, y se encuentran los puntos (λ^*, x^*) en donde $\nabla \mathcal{L}(\lambda^*, x^*) = 0$.

2 Condiciones de Segundo Orden: 2 opciones: máximo global o local.

- Si $\mathcal{L}(\lambda^*, \cdot)$ es **convexa/cóncava**: Se pueden verificar los menores principales arbitrarios de $H_f(x) - \sum_{k=1}^m \lambda_k^* H_{g_k}(x)$ o verificar que la correspondiente convexidad/concavidad de las funciones f, g (tomando en cuenta el signo de λ_i), para verificar si $\mathcal{L}(\lambda^*, \cdot)$ es convexa/cóncava en todo el dominio, y el óptimo encontrado es mínimo/máximo global.
- Si $\mathcal{L}(\lambda^*, \cdot)$ **no es convexa/cóncava**: Primero se calcula la matriz Hessiana $H_{f,g}^O$. Se verifican los menores principales dominantes en los puntos encontrados $D_r^O(\lambda^*, x^*)$, para verificar si el óptimo encontrado es mínimo/máximo local o punto de silla (o si el criterio no decide).

3 Óptimos: Si $f(\tilde{x})$ es mayor/menor que los máximos/mínimos de la segunda parte, entonces se agrega $\tilde{x}$ a la canasta de óptimos hallados.

Por Qué Utilizar Karush-Kuhn-Tucker (KKT)?

Considere el problema de maximización de utilidad con dos bienes:

$$\max U(x, y) \quad \text{con } p_x x + p_y y \leq m.$$

Por Qué Utilizar Karush-Kuhn-Tucker (KKT)?

Considere el problema de maximización de utilidad con dos bienes:

$$\max U(x, y) \quad \text{con } p_x x + p_y y \leq m.$$

Aquí tenemos dos escenarios:

- **Consumo Máximo:** Si nos encontramos en una región en donde más consumo implica más “felicidad”, nuestra utilidad máxima se dará gastando todo el dinero, es decir, resolviendo el problema:

$$\max U(x, y) \quad \text{con } p_x x + p_y y = m$$

Por Qué Utilizar Karush-Kuhn-Tucker (KKT)?

Considere el problema de maximización de utilidad con dos bienes:

$$\max U(x, y) \quad \text{con } p_x x + p_y y \leq m.$$

Aquí tenemos dos escenarios:

- **Consumo Máximo:** Si nos encontramos en una región en donde más consumo implica más “felicidad”, nuestra utilidad máxima se dará gastando todo el dinero, es decir, resolviendo el problema:

$$\max U(x, y) \quad \text{con } p_x x + p_y y = m$$

- **Punto de Saciedad:** Si alcanzamos nuestro punto de saciedad, es decir, consumir por encima de ese punto reduce nuestra “felicidad”, entonces el máximo se alcanzará sin gastar todo el dinero, es decir, resolviendo el problema sin restricciones:

$$\max U(x, y)$$

Por Qué Utilizar Karush-Kuhn-Tucker (KKT)?

Considere el problema de maximización de utilidad con dos bienes:

$$\max U(x, y) \quad \text{con } p_x x + p_y y \leq m.$$

Aquí tenemos dos escenarios:

- **Consumo Máximo:** Si nos encontramos en una región en donde más consumo implica más “felicidad”, nuestra utilidad máxima se dará gastando todo el dinero, es decir, resolviendo el problema:

$$\max U(x, y) \quad \text{con } p_x x + p_y y = m$$

- **Punto de Saciedad:** Si alcanzamos nuestro punto de saciedad, es decir, consumir por encima de ese punto reduce nuestra “felicidad”, entonces el máximo se alcanzará sin gastar todo el dinero, es decir, resolviendo el problema sin restricciones:

$$\max U(x, y)$$

Vemos que debemos resolver dos problemas, uno de Lagrange y uno sin restricciones. La técnica KKT permite resolver este tipo de problemas de una sola vez sin tener que plantearlos por separado.

Definición del Problema KKT

Tome una función objetivo f , $m + k$ funciones restricción, (m de igualdad) $\mathbf{g} = (g_1, \dots, g_m)$, (k de desigualdad) $\mathbf{h} = (h_1, \dots, h_k)$, en $\mathbb{R}^n$ ($m < n$).

Definición del Problema KKT

Tome una función objetivo f , $m + k$ funciones restricción, (m de igualdad) $\mathbf{g} = (g_1, \dots, g_m)$, (k de desigualdad) $\mathbf{h} = (h_1, \dots, h_k)$, en $\mathbb{R}^n$ ($m < n$).

El problema a optimizar se puede escribir de la siguiente forma:

$$\text{máx } f(\mathbf{x}) \quad \text{sujeta a} \quad \begin{cases} \mathbf{g}(\mathbf{x}) = \mathbf{c} \\ \mathbf{h}(\mathbf{x}) \leq \mathbf{b} \end{cases}$$

equivalentemente,

$$\text{máx } f(x_1, x_2, \dots, x_n) \quad \text{sujeta a} \quad \begin{cases} g_1(x_1, x_2, \dots, x_n) = c_1 \\ \vdots \\ g_m(x_1, x_2, \dots, x_n) = c_m \\ h_1(x_1, x_2, \dots, x_n) \leq b_1 \\ \vdots \\ h_k(x_1, x_2, \dots, x_n) \leq b_k \end{cases}$$

Definición del Problema KKT

Tome una función objetivo f , $m + k$ funciones restricción, (m de igualdad) $\mathbf{g} = (g_1, \dots, g_m)$, (k de desigualdad) $\mathbf{h} = (h_1, \dots, h_k)$, en $\mathbb{R}^n$ ($m < n$).

El problema a optimizar se puede escribir de la siguiente forma:

$$\text{máx } f(\mathbf{x}) \quad \text{sujeta a} \quad \begin{cases} \mathbf{g}(\mathbf{x}) = \mathbf{c} \\ \mathbf{h}(\mathbf{x}) \leq \mathbf{b} \end{cases}$$

equivalentemente,

$$\text{máx } f(x_1, x_2, \dots, x_n) \quad \text{sujeta a} \quad \begin{cases} g_1(x_1, x_2, \dots, x_n) = c_1 \\ \vdots \\ g_m(x_1, x_2, \dots, x_n) = c_m \\ h_1(x_1, x_2, \dots, x_n) \leq b_1 \\ \vdots \\ h_k(x_1, x_2, \dots, x_n) \leq b_k \end{cases}$$

Solo se considera el caso de maximización con desigualdad de menor o igual. Si se desea minimizar f basta con maximizar $-f$, y si la restricción es $\tilde{h}_i(\mathbf{x}) \geq \tilde{b}_i$, entonces tome $h_i = -\tilde{h}_i$ y $b_i = -\tilde{b}_i$, i.e., $h_i(\mathbf{x}) \leq b_i$.

Restricción Activa/Saturada Vs Inactiva/No-Saturada

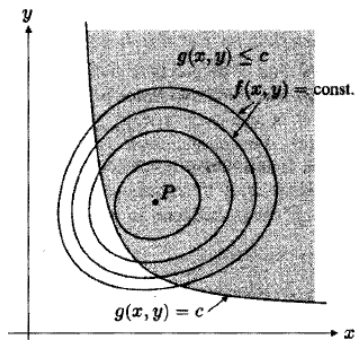
Definición

Sea $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ un óptimo del problema KKT. Si $h_i(\mathbf{x}^*) < b_i$, se dice que la restricción $h_j(\mathbf{x}^*) \leq b_j$ está “inactiva” o “no-saturada” en $\mathbf{x}^*$. Por el contrario, si $h_i(\mathbf{x}^*) = b_i$, se dice que la restricción $h_j(\mathbf{x}^*) \leq b_j$ está “activa” o “saturada” en $\mathbf{x}^*$.

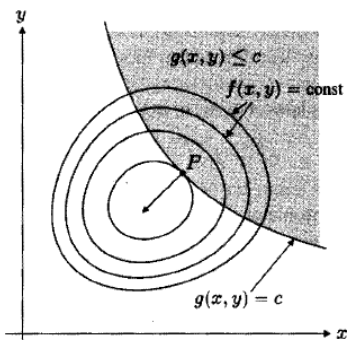
Restricción Activa/Saturada Vs Inactiva/No-Saturada

Definición

Sea $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ un óptimo del problema KKT. Si $h_i(\mathbf{x}^*) < b_i$, se dice que la restricción $h_j(\mathbf{x}^*) \leq b_j$ está “inactiva” o “no-saturada” en $\mathbf{x}^*$. Por el contrario, si $h_i(\mathbf{x}^*) = b_i$, se dice que la restricción $h_j(\mathbf{x}^*) \leq b_j$ está “activa” o “saturada” en $\mathbf{x}^*$.



**Máximo ocurre en el interior
del conjunto restricción**



**Máximo ocurre en la frontera
del conjunto restricción**

Derivación del Método KKT

Sea $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ un óptimo del problema KKT. Suponga que $h_{i_1}, \dots, h_{i_l}$ están saturadas ($l \leq k$) y las demás están inactivas.

Derivación del Método KKT

Sea $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ un óptimo del problema KKT. Suponga que $h_{i_1}, \dots, h_{i_l}$ están saturadas ($l \leq k$) y las demás están inactivas.

Esto implica que $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ es un óptimo del problema Lagrangiano

$$\text{máx } f(\mathbf{x}) \quad \text{sujeta a } \tilde{\mathbf{g}}(\mathbf{x}) = \tilde{\mathbf{c}},$$

donde $\tilde{\mathbf{g}} = (g_1, \dots, g_m, h_{i_1}, \dots, h_{i_l})$ y $\tilde{\mathbf{c}} = (c_1, \dots, c_m, b_{i_1}, \dots, b_{i_l})$.

Derivación del Método KKT

Sea $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ un óptimo del problema KKT. Suponga que $h_{i_1}, \dots, h_{i_l}$ están saturadas ($l \leq k$) y las demás están inactivas.

Esto implica que $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ es un óptimo del problema Lagrangiano

$$\text{máx } f(\mathbf{x}) \quad \text{sujeta a } \tilde{\mathbf{g}}(\mathbf{x}) = \tilde{\mathbf{c}},$$

donde $\tilde{\mathbf{g}} = (g_1, \dots, g_m, h_{i_1}, \dots, h_{i_l})$ y $\tilde{\mathbf{c}} = (c_1, \dots, c_m, b_{i_1}, \dots, b_{i_l})$.

Si $\mathbf{x}^*$ califica la restricción NDCQ para $\tilde{\mathbf{g}}$, entonces la condición de primer orden nos confirma que existe $\tilde{\boldsymbol{\lambda}}^* = (\lambda_1^*, \dots, \lambda_m^*, \tilde{\mu}_{i_1}^*, \dots, \tilde{\mu}_{i_l}^*)$, para

$$\tilde{\mathcal{L}}(\boldsymbol{\lambda}, \mathbf{x}) = f(\mathbf{x}) - \sum_{j=1}^m \lambda_j g_j(\mathbf{x}) - \sum_{r=1}^l \tilde{\mu}_{i_r} h_{i_r}(\mathbf{x}), \text{ t.q. } \nabla \tilde{\mathcal{L}}(\tilde{\boldsymbol{\lambda}}^*, \mathbf{x}^*) = \mathbf{0}.$$

Derivación del Método KKT

Sea $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ un óptimo del problema KKT. Suponga que $h_{i_1}, \dots, h_{i_l}$ están saturadas ($l \leq k$) y las demás están inactivas.

Esto implica que $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ es un óptimo del problema Lagrangiano

$$\text{máx } f(\mathbf{x}) \quad \text{sujeta a } \tilde{\mathbf{g}}(\mathbf{x}) = \tilde{\mathbf{c}},$$

donde $\tilde{\mathbf{g}} = (g_1, \dots, g_m, h_{i_1}, \dots, h_{i_l})$ y $\tilde{\mathbf{c}} = (c_1, \dots, c_m, b_{i_1}, \dots, b_{i_l})$.

Si $\mathbf{x}^*$ califica la restricción NDCQ para $\tilde{\mathbf{g}}$, entonces la condición de primer orden nos confirma que existe $\tilde{\boldsymbol{\lambda}}^* = (\lambda_1^*, \dots, \lambda_m^*, \tilde{\mu}_{i_1}^*, \dots, \tilde{\mu}_{i_l}^*)$, para

$$\tilde{\mathcal{L}}(\boldsymbol{\lambda}, \mathbf{x}) = f(\mathbf{x}) - \sum_{j=1}^m \lambda_j g_j(\mathbf{x}) - \sum_{r=1}^l \tilde{\mu}_{i_r} h_{i_r}(\mathbf{x}), \text{ t.q. } \nabla \tilde{\mathcal{L}}(\tilde{\boldsymbol{\lambda}}^*, \mathbf{x}^*) = \mathbf{0}.$$

Defina $\boldsymbol{\lambda}^* = (\lambda_1^*, \dots, \lambda_m^*, \mu_1^*, \dots, \mu_k^*)$, donde $\mu_i^* = \tilde{\mu}_{i_r}^*$ si $i = i_r$, y $\mu_i^* = 0$ si $i \neq i_r$. Note que para $\mathcal{L} = f - \sum_{j=1}^m \lambda_j g_j - \sum_{i=1}^k \mu_i h_i$, se cumpliría

$$\begin{aligned} \nabla \mathcal{L}(\boldsymbol{\lambda}^*, \mathbf{x}^*) &= \nabla f(\mathbf{x}^*) - \sum_{j=1}^m \lambda_j^* \nabla g_j(\mathbf{x}^*) - \sum_{i=1}^k \mu_i^* \nabla h_i(\mathbf{x}^*) \\ &= \nabla f(\mathbf{x}^*) - \sum_{j=1}^m \lambda_j^* \nabla g_j(\mathbf{x}^*) - \sum_{r=1}^l \mu_{i_r}^* \nabla h_{i_r}(\mathbf{x}^*) = \nabla \tilde{\mathcal{L}}(\tilde{\boldsymbol{\lambda}}^*, \mathbf{x}^*) = \mathbf{0}, \end{aligned}$$

es decir, $\mathbf{x}^*$ es un punto estacionario de $\mathcal{L}(\boldsymbol{\lambda}^*, \cdot)$.

Derivación del Método KKT

Sea $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ un óptimo del problema KKT. Suponga que $h_{i_1}, \dots, h_{i_l}$ están saturadas ($l \leq k$) y las demás están inactivas.

Esto implica que $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ es un óptimo del problema Lagrangiano

$$\text{máx } f(\mathbf{x}) \quad \text{sujeta a } \tilde{\mathbf{g}}(\mathbf{x}) = \tilde{\mathbf{c}},$$

donde $\tilde{\mathbf{g}} = (g_1, \dots, g_m, h_{i_1}, \dots, h_{i_l})$ y $\tilde{\mathbf{c}} = (c_1, \dots, c_m, b_{i_1}, \dots, b_{i_l})$.

Si $\mathbf{x}^*$ califica la restricción NDCQ para $\tilde{\mathbf{g}}$, entonces la condición de primer orden nos confirma que existe $\tilde{\boldsymbol{\lambda}}^* = (\lambda_1^*, \dots, \lambda_m^*, \tilde{\mu}_{i_1}^*, \dots, \tilde{\mu}_{i_l}^*)$, para

$$\tilde{\mathcal{L}}(\boldsymbol{\lambda}, \mathbf{x}) = f(\mathbf{x}) - \sum_{j=1}^m \lambda_j g_j(\mathbf{x}) - \sum_{r=1}^l \tilde{\mu}_{i_r} h_{i_r}(\mathbf{x}), \text{ t.q. } \nabla \tilde{\mathcal{L}}(\tilde{\boldsymbol{\lambda}}^*, \mathbf{x}^*) = \mathbf{0}.$$

Defina $\boldsymbol{\lambda}^* = (\lambda_1^*, \dots, \lambda_m^*, \mu_1^*, \dots, \mu_k^*)$, donde $\mu_i^* = \tilde{\mu}_{i_r}^*$ si $i = i_r$, y $\mu_i^* = 0$ si $i \neq i_r$. Note que para $\mathcal{L} = f - \sum_{j=1}^m \lambda_j g_j - \sum_{i=1}^k \mu_i h_i$, se cumpliría

$$\begin{aligned} \nabla \mathcal{L}(\boldsymbol{\lambda}^*, \mathbf{x}^*) &= \nabla f(\mathbf{x}^*) - \sum_{j=1}^m \lambda_j^* \nabla g_j(\mathbf{x}^*) - \sum_{i=1}^k \mu_i^* \nabla h_i(\mathbf{x}^*) \\ &= \nabla f(\mathbf{x}^*) - \sum_{j=1}^m \lambda_j^* \nabla g_j(\mathbf{x}^*) - \sum_{r=1}^l \mu_{i_r}^* \nabla h_{i_r}(\mathbf{x}^*) = \nabla \tilde{\mathcal{L}}(\tilde{\boldsymbol{\lambda}}^*, \mathbf{x}^*) = \mathbf{0}, \end{aligned}$$

es decir, $\mathbf{x}^*$ es un punto estacionario de $\mathcal{L}(\boldsymbol{\lambda}^*, \cdot)$. Finalmente,

$$\mu_i^*(\mathbf{c}, \mathbf{b}) = \partial_{b_i} f(\mathbf{x}^*(\mathbf{c}, \mathbf{b})) \geq 0, \text{ pues } \{\mathbf{x} / \mathbf{h}(\mathbf{x}) \leq \mathbf{b}\} \subset \{\mathbf{x} / \mathbf{h}(\mathbf{x}) \leq \mathbf{b} + \epsilon\}.$$

Condición Necesaria para el Método de KKT

Note que se requirió que x^* califique la restricción NDCQ para las restricciones $\tilde{\mathbf{g}} = (g_1, \dots, g_m, h_{i_1}, \dots, h_{i_l})$, es decir, las restricciones activas en x^* deben tener gradientes linealmente independientes.

Condición Necesaria para el Método de KKT

Note que se requirió que $\mathbf{x}^*$ califique la restricción NDCQ para las restricciones $\tilde{\mathbf{g}} = (g_1, \dots, g_m, h_{i_1}, \dots, h_{i_l})$, es decir, las restricciones activas en $\mathbf{x}^*$ deben tener gradientes linealmente independientes.

Linear Independence Constraint Qualification (LICQ)

Un punto $\mathbf{x}^*$ satisface la “Calificación de Restricción Linealmente Independiente” si

$$J_{\tilde{\mathbf{g}}}(\mathbf{x}^*) := \begin{pmatrix} \partial_{x_1} g_1(\mathbf{x}^*) & \partial_{x_2} g_1(\mathbf{x}^*) & \cdots & \partial_{x_n} g_1(\mathbf{x}^*) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} g_m(\mathbf{x}^*) & \partial_{x_2} g_m(\mathbf{x}^*) & \cdots & \partial_{x_n} g_m(\mathbf{x}^*) \\ \partial_{x_1} h_{i_1}(\mathbf{x}^*) & \partial_{x_2} h_{i_1}(\mathbf{x}^*) & \cdots & \partial_{x_n} h_{i_1}(\mathbf{x}^*) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} h_{i_l}(\mathbf{x}^*) & \partial_{x_2} h_{i_l}(\mathbf{x}^*) & \cdots & \partial_{x_n} h_{i_l}(\mathbf{x}^*) \end{pmatrix}$$

tiene rango $m + l$, y $\tilde{\mathbf{g}}(\mathbf{x}^*) = \tilde{\mathbf{c}} = (c_1, \dots, c_m, b_{i_1}, \dots, b_{i_l})$.

Condición Necesaria para el Método de KKT

Note que se requirió que $\mathbf{x}^*$ califique la restricción NDCQ para las restricciones $\tilde{\mathbf{g}} = (g_1, \dots, g_m, h_{i_1}, \dots, h_{i_l})$, es decir, las restricciones activas en $\mathbf{x}^*$ deben tener gradientes linealmente independientes.

Linear Independence Constraint Qualification (LICQ)

Un punto $\mathbf{x}^*$ satisface la “Calificación de Restricción Linealmente Independiente” si

$$J_{\tilde{\mathbf{g}}}(\mathbf{x}^*) := \begin{pmatrix} \partial_{x_1} g_1(\mathbf{x}^*) & \partial_{x_2} g_1(\mathbf{x}^*) & \cdots & \partial_{x_n} g_1(\mathbf{x}^*) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} g_m(\mathbf{x}^*) & \partial_{x_2} g_m(\mathbf{x}^*) & \cdots & \partial_{x_n} g_m(\mathbf{x}^*) \\ \partial_{x_1} h_{i_1}(\mathbf{x}^*) & \partial_{x_2} h_{i_1}(\mathbf{x}^*) & \cdots & \partial_{x_n} h_{i_1}(\mathbf{x}^*) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_1} h_{i_l}(\mathbf{x}^*) & \partial_{x_2} h_{i_l}(\mathbf{x}^*) & \cdots & \partial_{x_n} h_{i_l}(\mathbf{x}^*) \end{pmatrix}$$

tiene rango $m + l$, y $\tilde{\mathbf{g}}(\mathbf{x}^*) = \tilde{\mathbf{c}} = (c_1, \dots, c_m, b_{i_1}, \dots, b_{i_l})$.

En el caso de una restricción activa, $g(x_1, \dots, x_n) = c$, se solicita que

$$\nabla g(x_1^*, \dots, x_n^*) \neq \mathbf{0}.$$

Ejemplo

Encuentre los puntos en donde las siguientes restricciones califican (LICQ):

$$\mathbf{h}(x, y, z) = \begin{cases} h_1(x, y, z) = x^2 + y^2 + z^2 \leq 1, \\ h_2(x, y, z) = x + y + z \leq 1. \end{cases}$$

Ejemplo

Encuentre los puntos en donde las siguientes restricciones califican (LICQ):

$$\mathbf{h}(x, y, z) = \begin{cases} h_1(x, y, z) = x^2 + y^2 + z^2 \leq 1, \\ h_2(x, y, z) = x + y + z \leq 1. \end{cases}$$

Solución: Si ambas están activas:

$$J_{\mathbf{h}}(x, y, z) = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

con $\alpha \in \mathbb{R}$. Como $\mathbf{h}(x, y, z) = \mathbf{c}$, se debe cumplir que

$$\mathbf{h}\left(\frac{\alpha}{2}, \frac{\alpha}{2}, \frac{\alpha}{2}\right) = \begin{cases} \frac{3}{4}\alpha^2 = 1, \\ \frac{3}{2}\alpha = 1. \end{cases} \implies \text{No hay puntos que no califiquen.}$$

Ejemplo

Encuentre los puntos en donde las siguientes restricciones califican (LICQ):

$$\mathbf{h}(x, y, z) = \begin{cases} h_1(x, y, z) = x^2 + y^2 + z^2 \leq 1, \\ h_2(x, y, z) = x + y + z \leq 1. \end{cases}$$

Solución: Si ambas están activas:

$$J_{\mathbf{h}}(x, y, z) = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

con $\alpha \in \mathbb{R}$. Como $\mathbf{h}(x, y, z) = \mathbf{c}$, se debe cumplir que

$$\mathbf{h}\left(\frac{\alpha}{2}, \frac{\alpha}{2}, \frac{\alpha}{2}\right) = \begin{cases} \frac{3}{4}\alpha^2 = 1, \\ \frac{3}{2}\alpha = 1. \end{cases} \implies \text{No hay puntos que no califiquen.}$$

Si solo h_1 está activa: $\nabla h_1(x, y, z) = (2x, 2y, 2z) = (0, 0, 0)$ si $x = y = z = 0$. Pero $h_1(0, 0, 0) = 0 \neq 1$, i.e., todos los puntos califican.

Ejemplo

Encuentre los puntos en donde las siguientes restricciones califican (LICQ):

$$\mathbf{h}(x, y, z) = \begin{cases} h_1(x, y, z) = x^2 + y^2 + z^2 \leq 1, \\ h_2(x, y, z) = x + y + z \leq 1. \end{cases}$$

Solución: Si ambas están activas:

$$J_{\mathbf{h}}(x, y, z) = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

con $\alpha \in \mathbb{R}$. Como $\mathbf{h}(x, y, z) = \mathbf{c}$, se debe cumplir que

$$\mathbf{h}\left(\frac{\alpha}{2}, \frac{\alpha}{2}, \frac{\alpha}{2}\right) = \begin{cases} \frac{3}{4}\alpha^2 = 1, \\ \frac{3}{2}\alpha = 1. \end{cases} \implies \text{No hay puntos que no califiquen.}$$

Si solo h_1 está activa: $\nabla h_1(x, y, z) = (2x, 2y, 2z) = (0, 0, 0)$ si $x = y = z = 0$. Pero $h_1(0, 0, 0) = 0 \neq 1$, i.e., todos los puntos califican.

Si solo h_2 está activa: $\nabla h_2(x, y, z) = (1, 1, 1) \neq (0, 0, 0)$, por lo que no hay puntos que no califiquen.

Ejemplo

Encuentre los puntos en donde las siguientes restricciones califican (LICQ):

$$\mathbf{h}(x, y, z) = \begin{cases} h_1(x, y, z) = x^2 + y^2 + z^2 \leq 1, \\ h_2(x, y, z) = x + y + z \leq 1. \end{cases}$$

Solución: Si ambas están activas:

$$J_{\mathbf{h}}(x, y, z) = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

con $\alpha \in \mathbb{R}$. Como $\mathbf{h}(x, y, z) = \mathbf{c}$, se debe cumplir que

$$\mathbf{h}\left(\frac{\alpha}{2}, \frac{\alpha}{2}, \frac{\alpha}{2}\right) = \begin{cases} \frac{3}{4}\alpha^2 = 1, \\ \frac{3}{2}\alpha = 1. \end{cases} \implies \text{No hay puntos que no califiquen.}$$

Si solo h_1 está activa: $\nabla h_1(x, y, z) = (2x, 2y, 2z) = (0, 0, 0)$ si $x = y = z = 0$. Pero $h_1(0, 0, 0) = 0 \neq 1$, i.e., todos los puntos califican.

Si solo h_2 está activa: $\nabla h_2(x, y, z) = (1, 1, 1) \neq (0, 0, 0)$, por lo que no hay puntos que no califiquen.

Conclusión: Todos los puntos califican la restricción LICQ.

Contraejemplo: $J_g(\mathbf{x}^*)$ No Tiene Rango m

Considere el problema: $\max xy$ sujeta a $(x + y - 2)^2 \leq 0$.

Contraejemplo: $J_g(\mathbf{x}^*)$ No Tiene Rango m

Considere el problema: $\max xy$ sujeta a $(x + y - 2)^2 \leq 0$.

Solución: $\nabla g(x, y) = (2(x + y - 2), 2(x + y - 2)) = (0, 0)$ si $y = 2 - x$, y $g(x, 2 - x) = (x + (2 - x) - 2)^2 = 0$, por lo que los puntos $(x, 2 - x)$ no califican la restricción LICQ, en particular el punto $(1, 1)$.

Contraejemplo: $J_g(\mathbf{x}^*)$ No Tiene Rango m

Considere el problema: $\max xy$ sujeta a $(x + y - 2)^2 \leq 0$.

Solución: $\nabla g(x, y) = (2(x + y - 2), 2(x + y - 2)) = (0, 0)$ si $y = 2 - x$, y $g(x, 2 - x) = (x + (2 - x) - 2)^2 = 0$, por lo que los puntos $(x, 2 - x)$ no califican la restricción LICQ, en particular el punto $(1, 1)$.

Por otra parte, $0 \leq (x + y - 2)^2 \leq 0$, es decir, $(x + y - 2)^2 = 0$. Entonces, el problema se puede reescribir como un problema de Lagrange (la restricción siempre está activa): $\max xy$ sujeta a $(x + y - 2)^2 = 0$.

Contraejemplo: $J_g(\mathbf{x}^*)$ No Tiene Rango m

Considere el problema: $\max xy$ sujeta a $(x + y - 2)^2 \leq 0$.

Solución: $\nabla g(x, y) = (2(x + y - 2), 2(x + y - 2)) = (0, 0)$ si $y = 2 - x$, y $g(x, 2 - x) = (x + (2 - x) - 2)^2 = 0$, por lo que los puntos $(x, 2 - x)$ no califican la restricción LICQ, en particular el punto $(1, 1)$.

Por otra parte, $0 \leq (x + y - 2)^2 \leq 0$, es decir, $(x + y - 2)^2 = 0$. Entonces, el problema se puede reescribir como un problema de Lagrange (la restricción siempre está activa): $\max xy$ sujeta a $(x + y - 2)^2 = 0$.

El Lagrangiano es $\mathcal{L}(\lambda, x, y) = xy - \lambda(x + y - 2)^2$, y

$$\nabla \mathcal{L} = \begin{pmatrix} (x + y - 2)^2 \\ y - 2\lambda(x + y - 2) \\ x - 2\lambda(x + y - 2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies y = 2 - x \text{ pero } x = y = 0.$$

$\mathcal{L}$ no tiene puntos estacionarios, i.e., $(1, 1)$ no es un candidato de KKT.

Contraejemplo: $J_g(x^*)$ No Tiene Rango m

Considere el problema: $\max xy$ sujeta a $(x + y - 2)^2 \leq 0$.

Solución: $\nabla g(x, y) = (2(x + y - 2), 2(x + y - 2)) = (0, 0)$ si $y = 2 - x$, y $g(x, 2 - x) = (x + (2 - x) - 2)^2 = 0$, por lo que los puntos $(x, 2 - x)$ no califican la restricción LICQ, en particular el punto $(1, 1)$.

Por otra parte, $0 \leq (x + y - 2)^2 \leq 0$, es decir, $(x + y - 2)^2 = 0$. Entonces, el problema se puede reescribir como un problema de Lagrange (la restricción siempre está activa): $\max xy$ sujeta a $(x + y - 2)^2 = 0$.

El Lagrangiano es $\mathcal{L}(\lambda, x, y) = xy - \lambda(x + y - 2)^2$, y

$$\nabla \mathcal{L} = \begin{pmatrix} (x + y - 2)^2 \\ y - 2\lambda(x + y - 2) \\ x - 2\lambda(x + y - 2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies y = 2 - x \text{ pero } x = y = 0.$$

$\mathcal{L}$ no tiene puntos estacionarios, i.e., $(1, 1)$ no es un candidato de KKT.

Sin embargo, $(x + y - 2)^2 = 0$ si y solo si $x + y = 2$. Planteando el problema como $\max x(2 - x)$ (sin restricciones), tenemos que alcanza su máximo en $x = 1$, i.e., $f(1, 1) = 1$ es un máximo sujeto a $g(x, y) \leq 0$.

Condición Necesaria para Máximo

Punto Admisible / Punto Factible

Considere el problema KKT. Un punto $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ se dice que es “admisible” o “factible” si cumple que $g_i(\mathbf{x}^*) = c_i$ y $h_j(\mathbf{x}^*) \leq b_j$, $\forall i, j$.

Condición Necesaria para Máximo

Punto Admisible / Punto Factible

Considere el problema KKT. Un punto $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ se dice que es “admisible” o “factible” si cumple que $g_i(\mathbf{x}^*) = c_i$ y $h_j(\mathbf{x}^*) \leq b_j$, $\forall i, j$.

Teorema KKT: Condiciones de Primer Orden

Sea f una función continuamente diferenciable en $\mathbb{R}^n$ y sean g_i, h_j funciones continuamente diferenciables en $\mathbb{R}^n$, para $i = 1, \dots, m$ y $j = 1, \dots, k$. Asuma que $\mathbf{x}^*$ es un punto admisible que resuelve el problema KKT. Suponga además que $\mathbf{x}^*$ es un punto donde se califica la restricción (LICQ). Entonces existe un único vector $(\boldsymbol{\lambda}^*)^T = (\lambda_1^*, \dots, \lambda_m^*, \mu_1^*, \dots, \mu_k^*)$, con $\mu_j^* \geq 0$, tal que $\mathbf{x}^*$ es un punto estacionario del Lagrangiano

$\mathcal{L}(\boldsymbol{\lambda}^*, \mathbf{x}) := f(\mathbf{x}) - \sum_{i=1}^m \lambda_i^* (g_i(\mathbf{x}) - c_i) - \sum_{j=1}^k \mu_j^* (h_j(\mathbf{x}) - b_j)$,
esto es

- $\nabla f(\mathbf{x}^*) - \sum_{i=1}^m \lambda_i^* \nabla g_i(\mathbf{x}^*) - \sum_{j=1}^k \mu_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0}$,
- $g_i(\mathbf{x}^*) = c_i$,
- $\mu_i^* = 0$ si $h_i(\mathbf{x}^*) < b_i$ (**Condición de Holgura**).

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n \leq m$$

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n \leq m$$

Todo punto califica la restricción (LICQ), pues $\nabla h = (p_1, \dots, p_n) \neq \mathbf{0}$.

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n \leq m$$

Todo punto califica la restricción (LICQ), pues $\nabla h = (p_1, \dots, p_n) \neq \mathbf{0}$.

Lagrangiano: $\mathcal{L}(\lambda, \mathbf{x}) = U(x_1, \dots, x_n) - \lambda(p_1x_1 + \dots + p_nx_n - m)$

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n \leq m$$

Todo punto califica la restricción (LICQ), pues $\nabla h = (p_1, \dots, p_n) \neq \mathbf{0}$.

Lagrangiano: $\mathcal{L}(\lambda, \mathbf{x}) = U(x_1, \dots, x_n) - \lambda(p_1x_1 + \dots + p_nx_n - m)$

Por el teorema KKT, si existe un punto $(x_1^*, \dots, x_n^*)$ que maximice la utilidad gastando a lo más m , entonces existe un valor $\lambda^* \geq 0$ tal que

$$\begin{cases} \mathcal{L}_{x_i}(\lambda^*, x_1^*, \dots, x_n^*) = U_{x_i}(x_1^*, \dots, x_n^*) - \lambda^*p_i = 0; & i = 1, \dots, n. \\ \lambda^* = 0 \text{ si } p_1x_1^* + \dots + p_nx_n^* < m. \end{cases}$$

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n \leq m$$

Todo punto califica la restricción (LICQ), pues $\nabla h = (p_1, \dots, p_n) \neq \mathbf{0}$.

Lagrangiano: $\mathcal{L}(\lambda, \mathbf{x}) = U(x_1, \dots, x_n) - \lambda(p_1x_1 + \dots + p_nx_n - m)$

Por el teorema KKT, si existe un punto $(x_1^*, \dots, x_n^*)$ que maximice la utilidad gastando a lo más m , entonces existe un valor $\lambda^* \geq 0$ tal que

$$\begin{cases} \mathcal{L}_{x_i}(\lambda^*, x_1^*, \dots, x_n^*) = U_{x_i}(x_1^*, \dots, x_n^*) - \lambda^*p_i = 0; & i = 1, \dots, n. \\ \lambda^* = 0 \text{ si } p_1x_1^* + \dots + p_nx_n^* < m. \end{cases}$$

Si se gasta todo el dinero ($p_1x_1^* + \dots + p_nx_n^* = m$), recuperamos la condición del óptimo: tasa marginal de sustitución = precios relativos.

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n \leq m$$

Todo punto califica la restricción (LICQ), pues $\nabla h = (p_1, \dots, p_n) \neq \mathbf{0}$.

Lagrangiano: $\mathcal{L}(\lambda, \mathbf{x}) = U(x_1, \dots, x_n) - \lambda(p_1x_1 + \dots + p_nx_n - m)$

Por el teorema KKT, si existe un punto $(x_1^*, \dots, x_n^*)$ que maximice la utilidad gastando a lo más m , entonces existe un valor $\lambda^* \geq 0$ tal que

$$\begin{cases} \mathcal{L}_{x_i}(\lambda^*, x_1^*, \dots, x_n^*) = U_{x_i}(x_1^*, \dots, x_n^*) - \lambda^*p_i = 0; & i = 1, \dots, n. \\ \lambda^* = 0 \text{ si } p_1x_1^* + \dots + p_nx_n^* < m. \end{cases}$$

Si se gasta todo el dinero ($p_1x_1^* + \dots + p_nx_n^* = m$), recuperamos la condición del óptimo: tasa marginal de sustitución = precios relativos.

Si no se gasta todo ($p_1x_1^* + \dots + p_nx_n^* < m$), entonces $\nabla U(\mathbf{x}^*) = \mathbf{0}$, pues $\lambda^* = 0$, (condición de óptimo global sin restricciones - si U es cóncava).

Ejemplo: Maximización de Utilidad

Considere n bienes $x_1, \dots, x_n$ y una función de utilidad U . El problema es

$$\text{máx } U(x_1, \dots, x_n) \quad \text{sujeta a } p_1x_1 + \dots + p_nx_n \leq m$$

Todo punto califica la restricción (LICQ), pues $\nabla h = (p_1, \dots, p_n) \neq \mathbf{0}$.

Lagrangiano: $\mathcal{L}(\lambda, \mathbf{x}) = U(x_1, \dots, x_n) - \lambda(p_1x_1 + \dots + p_nx_n - m)$

Por el teorema KKT, si existe un punto $(x_1^*, \dots, x_n^*)$ que maximice la utilidad gastando a lo más m , entonces existe un valor $\lambda^* \geq 0$ tal que

$$\begin{cases} \mathcal{L}_{x_i}(\lambda^*, x_1^*, \dots, x_n^*) = U_{x_i}(x_1^*, \dots, x_n^*) - \lambda^*p_i = 0; & i = 1, \dots, n. \\ \lambda^* = 0 \text{ si } p_1x_1^* + \dots + p_nx_n^* < m. \end{cases}$$

Si se gasta todo el dinero ($p_1x_1^* + \dots + p_nx_n^* = m$), recuperamos la condición del óptimo: tasa marginal de sustitución = precios relativos.

Si no se gasta todo ($p_1x_1^* + \dots + p_nx_n^* < m$), entonces $\nabla U(\mathbf{x}^*) = \mathbf{0}$, pues $\lambda^* = 0$, (condición de óptimo global sin restricciones - si U es cóncava).

Propiedad del óptimo: Se consume hasta el punto de saciedad si el dinero alcanza, sino se distribuirá el consumo hasta que la utilidad generada por una unidad extra del bien relativa a su costo sea igual para todos.

Condición Suficiente para Máximo Global

Condición de Segundo Orden para Máximo Global

Sean $f, g, h \in \mathcal{C}^1(S)$, con S **convexo**, y $x^* \in S$ un punto estacionario de $\mathcal{L}(\lambda^*, \cdot)$, admisible que satisface LICQ y la condición de holgura.

Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava, entonces $f(x^*)$ es un máximo global.

Condición Suficiente para Máximo Global

Condición de Segundo Orden para Máximo Global

Sean $f, g, h \in \mathcal{C}^1(S)$, con S **convexo**, y $x^* \in S$ un punto estacionario de $\mathcal{L}(\lambda^*, \cdot)$, admisible que satisface LICQ y la condición de holgura.

Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava, entonces $f(x^*)$ es un máximo global.

Observaciones:

- Se exige que la función $\mathcal{L}(\lambda^*, \cdot)$ sea cóncava/convexa **en todo** S .

Condición Suficiente para Máximo Global

Condición de Segundo Orden para Máximo Global

Sean $f, g, h \in \mathcal{C}^1(S)$, con S **convexo**, y $x^* \in S$ un punto estacionario de $\mathcal{L}(\lambda^*, \cdot)$, admisible que satisface LICQ y la condición de holgura.

Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava, entonces $f(x^*)$ es un máximo global.

Observaciones:

- Se exige que la función $\mathcal{L}(\lambda^*, \cdot)$ sea cóncava/convexa **en todo** S .
- Si f es cóncava, g_i es convexa si $\lambda_i^* \geq 0$, y g_j es cóncava si $\lambda_j^* \leq 0$, y h_r es convexa (pues $\mu_r \geq 0$), entonces $\mathcal{L}(\lambda^*, \cdot)$ sería una suma de funciones cóncavas, es decir, $f(x^*)$ sería un máximo global.

Condición Suficiente para Máximo Global

Condición de Segundo Orden para Máximo Global

Sean $f, g, h \in \mathcal{C}^1(S)$, con S **convexo**, y $x^* \in S$ un punto estacionario de $\mathcal{L}(\lambda^*, \cdot)$, admisible que satisface LICQ y la condición de holgura.

Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava, entonces $f(x^*)$ es un máximo global.

Observaciones:

- Se exige que la función $\mathcal{L}(\lambda^*, \cdot)$ sea cóncava/convexa **en todo** S .
- Si f es cóncava, g_i es convexa si $\lambda_i^* \geq 0$, y g_j es cóncava si $\lambda_j^* \leq 0$, y h_r es convexa (pues $\mu_r \geq 0$), entonces $\mathcal{L}(\lambda^*, \cdot)$ sería una suma de funciones cóncavas, es decir, $f(x^*)$ sería un máximo global.
- Si las restricciones son lineales basta que f sea cóncava.

Condición Suficiente para Máximo Global

Condición de Segundo Orden para Máximo Global

Sean $f, g, h \in \mathcal{C}^1(S)$, con S **convexo**, y $x^* \in S$ un punto estacionario de $\mathcal{L}(\lambda^*, \cdot)$, admisible que satisface LICQ y la condición de holgura.

Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava, entonces $f(x^*)$ es un máximo global.

Observaciones:

- Se exige que la función $\mathcal{L}(\lambda^*, \cdot)$ sea cóncava/convexa **en todo** S .
- Si f es cóncava, g_i es convexa si $\lambda_i^* \geq 0$, y g_j es cóncava si $\lambda_j^* \leq 0$, y h_r es convexa (pues $\mu_r \geq 0$), entonces $\mathcal{L}(\lambda^*, \cdot)$ sería una suma de funciones cóncavas, es decir, $f(x^*)$ sería un máximo global.
- Si las restricciones son lineales basta que f sea cóncava.

Teorema de Arrow-Enthoven

Sean $f, g, h \in \mathcal{C}^1(S)$, con S **convexo**, y $x^* \in S$ un punto estacionario de $\mathcal{L}(\lambda^*, \cdot)$, admisible que satisface LICQ y la condición de holgura.

Si f es cuasicóncava, $\lambda_i^* g_i$ y $\mu_j^* h_j$ son cuasiconvexas, y $\nabla f(x^*) \neq 0$, entonces $f(x^*)$ es un máximo global.

Ejemplos

① Resuelva $\min x^2 + y^2$ sujeta a $x + 2y \geq 5$.

② Considere el problema
$$\max \ln x_1 + x_2 + x_3 \quad \text{sujeta a} \quad \begin{cases} x_1 + x_2 + x_3 \leq 1 \\ x_1 \geq 1 \\ x_1^2 + x_2^2 \leq 2 \end{cases} .$$

Muestre que hay infinitos puntos que resuelven el problema planteado.Cuál es la solución?

③ Resuelva el problema
$$\max \ln(x + y + 3) \quad \text{sujeta a} \quad \begin{cases} x^2 + y^2 + z^2 \leq 1 \\ x + y + z \leq 1 \end{cases} .$$

④ Aplique el teorema de Arrow-Enthoven al problema

$$\max xy \quad \text{sujeta a} \quad \begin{cases} x + y \leq 6 \\ x \geq 0 \\ y \geq 0 \end{cases} .$$

Ejemplos: Solución (mín $x^2 + y^2$ s.a. $x + 2y \geq 5$)

Reescribimos el problema (principio de dualidad) como

$$\text{máx } -x^2 - y^2 \text{ sujeta a } -x - 2y \leq -5.$$

Ejemplos: Solución (mín $x^2 + y^2$ s.a. $x + 2y \geq 5$)

Reescribimos el problema (principio de dualidad) como

$$\text{máx } -x^2 - y^2 \text{ sujeta a } -x - 2y \leq -5.$$

Como $\nabla h = (-1, -2) \neq (0, 0)$, i.e., todos los puntos califican LICQ.

Ejemplos: Solución (mín $x^2 + y^2$ s.a. $x + 2y \geq 5$)

Reescribimos el problema (principio de dualidad) como

$$\text{máx } -x^2 - y^2 \text{ sujeta a } -x - 2y \leq -5.$$

Como $\nabla h = (-1, -2) \neq (0, 0)$, i.e., todos los puntos califican LICQ.

Planteamos el Lagrangiano:

$$\mathcal{L}(\mu, x, y) = -x^2 - y^2 - \mu(-x - 2y - (-5)) = -x^2 - y^2 + \mu(x + 2y - 5).$$

Ejemplos: Solución (mín $x^2 + y^2$ s.a. $x + 2y \geq 5$)

Reescribimos el problema (principio de dualidad) como

$$\text{máx } -x^2 - y^2 \text{ sujeta a } -x - 2y \leq -5.$$

Como $\nabla h = (-1, -2) \neq (0, 0)$, i.e., todos los puntos califican LICQ.

Planteamos el Lagrangiano:

$$\mathcal{L}(\mu, x, y) = -x^2 - y^2 - \mu(-x - 2y - (-5)) = -x^2 - y^2 + \mu(x + 2y - 5).$$

Buscamos las condiciones de primer orden:

$$\nabla_{x,y} \mathcal{L}(\mu, \cdot, \cdot) = \begin{pmatrix} -2x + \mu \\ -2y + \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies (\mu^*, x^*, y^*) = \left(\frac{10}{3}, \frac{5}{3}, \frac{5}{3} \right).$$

Si $-x - 2y < -5$ entonces $\mu = 0$, lo que implica $x = 0 = y$, pero $h(0, 0) = 0 > -5$. Entonces $(0, 0)$ no es factible.

Ejemplos: Solución (mín $x^2 + y^2$ s.a. $x + 2y \geq 5$)

Reescribimos el problema (principio de dualidad) como

$$\text{máx } -x^2 - y^2 \text{ sujeta a } -x - 2y \leq -5.$$

Como $\nabla h = (-1, -2) \neq (0, 0)$, i.e., todos los puntos califican LICQ.

Planteamos el Lagrangiano:

$$\mathcal{L}(\mu, x, y) = -x^2 - y^2 - \mu(-x - 2y - (-5)) = -x^2 - y^2 + \mu(x + 2y - 5).$$

Buscamos las condiciones de primer orden:

$$\nabla_{x,y} \mathcal{L}(\mu, \cdot, \cdot) = \begin{pmatrix} -2x + \mu \\ -2y + \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies (\mu^*, x^*, y^*) = \left(\frac{10}{3}, \frac{5}{3}, \frac{5}{3} \right).$$

Si $-x - 2y < -5$ entonces $\mu = 0$, lo que implica $x = 0 = y$, pero $h(0, 0) = 0 > -5$. Entonces $(0, 0)$ no es factible.

Como $h = -x - 2y$ es lineal y $f = -x^2 - y^2$ es cóncava, entonces

$$f\left(\frac{5}{3}, \frac{5}{3}\right) = -\frac{25}{9} \text{ es el máximo global s.a. } -x - 2y \leq -5.$$

Ejemplos: Solución (mín $x^2 + y^2$ s.a. $x + 2y \geq 5$)

Reescribimos el problema (principio de dualidad) como

$$\text{máx } -x^2 - y^2 \text{ sujeta a } -x - 2y \leq -5.$$

Como $\nabla h = (-1, -2) \neq (0, 0)$, i.e., todos los puntos califican LICQ.

Planteamos el Lagrangiano:

$$\mathcal{L}(\mu, x, y) = -x^2 - y^2 - \mu(-x - 2y - (-5)) = -x^2 - y^2 + \mu(x + 2y - 5).$$

Buscamos las condiciones de primer orden:

$$\nabla_{x,y} \mathcal{L}(\mu, \cdot, \cdot) = \begin{pmatrix} -2x + \mu \\ -2y + \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies (\mu^*, x^*, y^*) = \left(\frac{10}{3}, \frac{5}{3}, \frac{5}{3} \right).$$

Si $-x - 2y < -5$ entonces $\mu = 0$, lo que implica $x = 0 = y$, pero $h(0, 0) = 0 > -5$. Entonces $(0, 0)$ no es factible.

Como $h = -x - 2y$ es lineal y $f = -x^2 - y^2$ es cóncava, entonces

$$f\left(\frac{5}{3}, \frac{5}{3}\right) = -\frac{25}{9} \text{ es el máximo global s.a. } -x - 2y \leq -5.$$

Concluimos que $-f\left(\frac{5}{3}, \frac{5}{3}\right) = \frac{25}{9}$ es el mínimo global sobre $x + 2y \geq 5$.

Ejemplos: Solución Problema 2

Calificación de la Restricción (LICQ):

- Si todas están activas ($x_1 + x_2 + x_3 = 1$, $x_1 = 1$ y $x_1^2 + x_2^2 = 2$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}; \quad \det J_{\tilde{g}} = -2x_2 = 0 \text{ si } x_2 = 0.$$

Por h_2 sabemos que $x_1 = 1$, pero $h_3(1, 0, x_3) = 1^2 + 0^2 = 1 \neq 2$, i.e., el punto no es admisible. Si todas están activas siempre califica la restricción

Ejemplos: Solución Problema 2

Calificación de la Restricción (LICQ):

- Si todas están activas ($x_1 + x_2 + x_3 = 1$, $x_1 = 1$ y $x_1^2 + x_2^2 = 2$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}; \quad \det J_{\tilde{g}} = -2x_2 = 0 \text{ si } x_2 = 0.$$

Por h_2 sabemos que $x_1 = 1$, pero $h_3(1, 0, x_3) = 1^2 + 0^2 = 1 \neq 2$, i.e., el punto no es admisible. Si todas están activas siempre califica la restricción

- Si h_1 y h_2 están activas ($x_1 + x_2 + x_3 = 1$ y $x_1 = 1$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

Todos los puntos califican la restricción.

Ejemplos: Solución Problema 2

Calificación de la Restricción (LICQ):

- Si todas están activas ($x_1 + x_2 + x_3 = 1$, $x_1 = 1$ y $x_1^2 + x_2^2 = 2$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}; \quad \det J_{\tilde{g}} = -2x_2 = 0 \text{ si } x_2 = 0.$$

Por h_2 sabemos que $x_1 = 1$, pero $h_3(1, 0, x_3) = 1^2 + 0^2 = 1 \neq 2$, i.e., el punto no es admisible. Si todas están activas siempre califica la restricción

- Si h_1 y h_2 están activas ($x_1 + x_2 + x_3 = 1$ y $x_1 = 1$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

Todos los puntos califican la restricción.

- Si h_1 y h_3 están activas ($x_1 + x_2 + x_3 = 1$ y $x_1^2 + x_2^2 = 2$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 & 1 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

Todos los puntos califican la restricción.

Ejemplos: Solución Problema 2 (Continuación)

Calificación de la Restricción (LICQ):

- Si h_2 y h_3 están activas ($x_1 = 1$ y $x_1^2 + x_2^2 = 2$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 0 & 0 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}; \quad \text{no tiene rango 2 si } x_2 = 0.$$

Pero como $x_1 = 1$, $h_3(1, 0, x_3) = 1 \neq 2$, i.e., no es admisible.

Ejemplos: Solución Problema 2 (Continuación)

Calificación de la Restricción (LICQ):

- Si h_2 y h_3 están activas ($x_1 = 1$ y $x_1^2 + x_2^2 = 2$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 0 & 0 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}; \quad \text{no tiene rango 2 si } x_2 = 0.$$

Pero como $x_1 = 1$, $h_3(1, 0, x_3) = 1 \neq 2$, i.e., no es admisible.

- Si solo h_1 está activa ($x_1 + x_2 + x_3 = 1$):

$$\nabla h_1 = (1, 1, 1) \neq (0, 0, 0).$$

Ejemplos: Solución Problema 2 (Continuación)

Calificación de la Restricción (LICQ):

- Si h_2 y h_3 están activas ($x_1 = 1$ y $x_1^2 + x_2^2 = 2$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 0 & 0 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}; \quad \text{no tiene rango 2 si } x_2 = 0.$$

Pero como $x_1 = 1$, $h_3(1, 0, x_3) = 1 \neq 2$, i.e., no es admisible.

- Si solo h_1 está activa ($x_1 + x_2 + x_3 = 1$):

$$\nabla h_1 = (1, 1, 1) \neq (0, 0, 0).$$

- Si solo h_2 está activa ($x_1 = 1$):

$$\nabla h_2 = (1, 0, 0) \neq (0, 0, 0).$$

Ejemplos: Solución Problema 2 (Continuación)

Calificación de la Restricción (LICQ):

- Si h_2 y h_3 están activas ($x_1 = 1$ y $x_1^2 + x_2^2 = 2$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 0 & 0 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}; \quad \text{no tiene rango 2 si } x_2 = 0.$$

Pero como $x_1 = 1$, $h_3(1, 0, x_3) = 1 \neq 2$, i.e., no es admisible.

- Si solo h_1 está activa ($x_1 + x_2 + x_3 = 1$):

$$\nabla h_1 = (1, 1, 1) \neq (0, 0, 0).$$

- Si solo h_2 está activa ($x_1 = 1$):

$$\nabla h_2 = (1, 0, 0) \neq (0, 0, 0).$$

- Si solo h_3 está activa ($x_1^2 + x_2^2 = 2$):

$$\nabla h_3 = (2x_1, 2x_2, 0) = (0, 0, 0),$$

solo si $x_1 = x_2 = 0$, pero $h_3(0, 0, x_3) = 0 \neq 2$, i.e., no es admisible.

Ejemplos: Solución Problema 2 (Continuación)

Calificación de la Restricción (LICQ):

- Si h_2 y h_3 están activas ($x_1 = 1$ y $x_1^2 + x_2^2 = 2$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 0 & 0 \\ 2x_1 & 2x_2 & 0 \end{pmatrix}; \quad \text{no tiene rango 2 si } x_2 = 0.$$

Pero como $x_1 = 1$, $h_3(1, 0, x_3) = 1 \neq 2$, i.e., no es admisible.

- Si solo h_1 está activa ($x_1 + x_2 + x_3 = 1$):

$$\nabla h_1 = (1, 1, 1) \neq (0, 0, 0).$$

- Si solo h_2 está activa ($x_1 = 1$):

$$\nabla h_2 = (1, 0, 0) \neq (0, 0, 0).$$

- Si solo h_3 está activa ($x_1^2 + x_2^2 = 2$):

$$\nabla h_3 = (2x_1, 2x_2, 0) = (0, 0, 0),$$

solo si $x_1 = x_2 = 0$, pero $h_3(0, 0, x_3) = 0 \neq 2$, i.e., no es admisible.

Conclusión: Todos los puntos califican la restricción.

Ejemplos: Solución Problema 2 (Continuación)

Planteamos el Lagrangiano (cambiando la segunda restricción):

$$\mathcal{L}(\boldsymbol{\lambda}, \boldsymbol{x}) = \ln x_1 + x_2 + x_3 - \mu_1(x_1 + x_2 + x_3 - 1) - \mu_2(-x_1 - (-1)) - \mu_3(x_1^2 + x_2^2 - 2)$$

Ejemplos: Solución Problema 2 (Continuación)

Planteamos el Lagrangiano (cambiando la segunda restricción):

$$\mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}) = \ln x_1 + x_2 + x_3 - \mu_1(x_1 + x_2 + x_3 - 1) - \mu_2(-x_1 - (-1)) - \mu_3(x_1^2 + x_2^2 - 2)$$

CPO:

$$\nabla_{x_1, x_2, x_3} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot, \cdot) = \begin{pmatrix} \frac{1}{x_1} - \mu_1 + \mu_2 - 2x_1\mu_3 \\ 1 - \mu_1 - 2x_2\mu_3 \\ 1 - \mu_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* = 1.$$

Ejemplos: Solución Problema 2 (Continuación)

Planteamos el Lagrangiano (cambiando la segunda restricción):

$$\mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}) = \ln x_1 + x_2 + x_3 - \mu_1(x_1 + x_2 + x_3 - 1) - \mu_2(-x_1 - (-1)) - \mu_3(x_1^2 + x_2^2 - 2)$$

CPO:

$$\nabla_{x_1, x_2, x_3} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot, \cdot) = \begin{pmatrix} \frac{1}{x_1} - \mu_1 + \mu_2 - 2x_1\mu_3 \\ 1 - \mu_1 - 2x_2\mu_3 \\ 1 - \mu_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* = 1.$$

De la segunda sabemos que $2x_2\mu_3 = 0$, i.e., $x_2 = 0$ o $\mu_3 = 0$. Si $\mu_3 \neq 0$, entonces $x_2 = 0$ y $h_3(x_1, 0, x_3) = x_1^2 = 2$, i.e., $x_1 = \pm\sqrt{2}$. Por h_2 sabemos que $x_1 = \sqrt{2} > 1$ (único factible), por lo que $\mu_2 = 0$. De la primera vemos que $0 \leq \mu_3 = \frac{1}{2x_1^2} - \frac{\mu_1}{2x_1} - \frac{\mu_2}{2x_1} = \frac{1}{4} - \frac{1}{2\sqrt{2}} < 0$ ($\Rightarrow \Leftarrow$). Concluimos que $\mu_3 = 0$. Si $\mu_2 \neq 0$, entonces $x_1 = 1$, pero de la primera vemos que $\mu_2 = 0$. Concluimos que $\mu_2 = 0$ y de la primera tenemos que $x_1 = 1$. Por h_1 vemos que $x_2 = -x_3$, y por h_3 tenemos que $x_2^2 \leq 1$. Concluimos que

$$(\mu_1^*, \mu_2^*, \mu_3^*, x_1^*, x_2^*, x_3^*) = (1, 0, 0, 1, t, -t) \quad -1 \leq t \leq 1.$$

Ejemplos: Solución Problema 2 (Continuación)

Planteamos el Lagrangiano (cambiando la segunda restricción):

$$\mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}) = \ln x_1 + x_2 + x_3 - \mu_1(x_1 + x_2 + x_3 - 1) - \mu_2(-x_1 - (-1)) - \mu_3(x_1^2 + x_2^2 - 2)$$

CPO:

$$\nabla_{x_1, x_2, x_3} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot, \cdot) = \begin{pmatrix} \frac{1}{x_1} - \mu_1 + \mu_2 - 2x_1\mu_3 \\ 1 - \mu_1 - 2x_2\mu_3 \\ 1 - \mu_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* = 1.$$

De la segunda sabemos que $2x_2\mu_3 = 0$, i.e., $x_2 = 0$ o $\mu_3 = 0$. Si $\mu_3 \neq 0$, entonces $x_2 = 0$ y $h_3(x_1, 0, x_3) = x_1^2 = 2$, i.e., $x_1 = \pm\sqrt{2}$. Por h_2 sabemos que $x_1 = \sqrt{2} > 1$ (único factible), por lo que $\mu_2 = 0$. De la primera vemos que $0 \leq \mu_3 = \frac{1}{2x_1^2} - \frac{\mu_1}{2x_1} - \frac{\mu_2}{2x_1} = \frac{1}{4} - \frac{1}{2\sqrt{2}} < 0$ ($\Rightarrow \Leftarrow$). Concluimos que $\mu_3 = 0$. Si $\mu_2 \neq 0$, entonces $x_1 = 1$, pero de la primera vemos que $\mu_2 = 0$. Concluimos que $\mu_2 = 0$ y de la primera tenemos que $x_1 = 1$. Por h_1 vemos que $x_2 = -x_3$, y por h_3 tenemos que $x_2^2 \leq 1$. Concluimos que

$$(\mu_1^*, \mu_2^*, \mu_3^*, x_1^*, x_2^*, x_3^*) = (1, 0, 0, 1, t, -t) \quad -1 \leq t \leq 1.$$

Como $f = \ln x_1 + x_2 + x_3$ es cóncava, h_1 y h_2 son lineales, y h_3 es convexa, entonces $f(1, t, -t) = 0$ es el máximo global (infinitos óptimos).

Ejemplos: Solución Problema 3

Calificación de la Restricción (LICQ):

- Si todas están activas ($x^2 + y^2 + z^2 = 1$ y $x + y + z = 1$):

$$J_{\tilde{g}} = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix}; \quad \text{no tiene rango 2 si } x = y = z.$$

Sustituyendo en h_2 tenemos $x = y = z = \frac{1}{3}$, pero $h_1\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = \frac{1}{3} \neq 1$.

Ejemplos: Solución Problema 3

Calificación de la Restricción (LICQ):

- Si todas están activas ($x^2 + y^2 + z^2 = 1$ y $x + y + z = 1$):

$$J_{\tilde{g}} = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix}; \quad \text{no tiene rango 2 si } x = y = z.$$

Sustituyendo en h_2 tenemos $x = y = z = \frac{1}{3}$, pero $h_1\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = \frac{1}{3} \neq 1$.

- Si solo h_1 está activa ($x^2 + y^2 + z^2 = 1$):

$$\nabla h_1 = (2x, 2y, 2z) = (0, 0, 0),$$

solo si $x = y = z = 0$, pero $h_1(0, 0, 0) = 0 \neq 1$.

Ejemplos: Solución Problema 3

Calificación de la Restricción (LICQ):

- Si todas están activas ($x^2 + y^2 + z^2 = 1$ y $x + y + z = 1$):

$$J_{\tilde{g}} = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix}; \quad \text{no tiene rango 2 si } x = y = z.$$

Sustituyendo en h_2 tenemos $x = y = z = \frac{1}{3}$, pero $h_1\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = \frac{1}{3} \neq 1$.

- Si solo h_1 está activa ($x^2 + y^2 + z^2 = 1$):

$$\nabla h_1 = (2x, 2y, 2z) = (0, 0, 0),$$

solo si $x = y = z = 0$, pero $h_1(0, 0, 0) = 0 \neq 1$.

- Si solo h_2 está activa ($x + y + z = 1$):

$$\nabla h_2 = (1, 1, 1) \neq (0, 0, 0).$$

Ejemplos: Solución Problema 3

Calificación de la Restricción (LICQ):

- Si todas están activas ($x^2 + y^2 + z^2 = 1$ y $x + y + z = 1$):

$$J_{\tilde{g}} = \begin{pmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \end{pmatrix}; \quad \text{no tiene rango 2 si } x = y = z.$$

Sustituyendo en h_2 tenemos $x = y = z = \frac{1}{3}$, pero $h_1\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = \frac{1}{3} \neq 1$.

- Si solo h_1 está activa ($x^2 + y^2 + z^2 = 1$):

$$\nabla h_1 = (2x, 2y, 2z) = (0, 0, 0),$$

solo si $x = y = z = 0$, pero $h_1(0, 0, 0) = 0 \neq 1$.

- Si solo h_2 está activa ($x + y + z = 1$):

$$\nabla h_2 = (1, 1, 1) \neq (0, 0, 0).$$

Conclusión: Todos los puntos califican la restricción.

Ejemplos: Solución Problema 3 (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\boldsymbol{\lambda}, x, y, z) = \ln(x + y + 3) - \mu_1(x^2 + y^2 + z^2 - 1) - \mu_2(x + y + z - 1).$$

Ejemplos: Solución Problema 3 (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\boldsymbol{\lambda}, x, y, z) = \ln(x + y + 3) - \mu_1(x^2 + y^2 + z^2 - 1) - \mu_2(x + y + z - 1).$$

CPO:

$$\nabla_{x,y,z} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot, \cdot) = \begin{pmatrix} \frac{1}{x+y+3} - 2x\mu_1 - \mu_2 \\ \frac{1}{x+y+3} - 2y\mu_1 - \mu_2 \\ -2z\mu_1 - \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* \neq 0.$$

Ejemplos: Solución Problema 3 (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\boldsymbol{\lambda}, x, y, z) = \ln(x + y + 3) - \mu_1(x^2 + y^2 + z^2 - 1) - \mu_2(x + y + z - 1).$$

CPO:

$$\nabla_{x,y,z} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot, \cdot) = \begin{pmatrix} \frac{1}{x+y+3} - 2x\mu_1 - \mu_2 \\ \frac{1}{x+y+3} - 2y\mu_1 - \mu_2 \\ -2z\mu_1 - \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* \neq 0.$$

Si $\mu_2^* = 0$ entonces $z = 0$ (por la tercera). Restando la primera a la segunda obtenemos, $-2(y - x)\mu_1^* = 0$, i.e., $x = y$. De la primera restricción, $2x^2 = 1$, i.e., $x^* = y^* = \pm \frac{1}{\sqrt{2}}$. $\mu_1^* = \frac{-1}{\sqrt{2}(3-\sqrt{2})} < 0$ si $x, y < 0$, por lo que $x^* = y^* = \frac{1}{\sqrt{2}}$. No es admisible ($x^* + y^* + z^* = \sqrt{2} > 1$).

Ejemplos: Solución Problema 3 (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\lambda, x, y, z) = \ln(x + y + 3) - \mu_1(x^2 + y^2 + z^2 - 1) - \mu_2(x + y + z - 1).$$

CPO:

$$\nabla_{x,y,z} \mathcal{L}(\lambda, \cdot, \cdot, \cdot) = \begin{pmatrix} \frac{1}{x+y+3} - 2x\mu_1 - \mu_2 \\ \frac{1}{x+y+3} - 2y\mu_1 - \mu_2 \\ -2z\mu_1 - \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* \neq 0.$$

Si $\mu_2^* = 0$ entonces $z = 0$ (por la tercera). Restando la primera a la segunda obtenemos, $-2(y - x)\mu_1^* = 0$, i.e., $x = y$. De la primera restricción, $2x^2 = 1$, i.e., $x^* = y^* = \pm \frac{1}{\sqrt{2}}$. $\mu_1^* = \frac{-1}{\sqrt{2}(3-\sqrt{2})} < 0$ si $x, y < 0$, por lo que $x^* = y^* = \frac{1}{\sqrt{2}}$. No es admisible ($x^* + y^* + z^* = \sqrt{2} > 1$).

Si $\mu_2^* \neq 0$, entonces $2x + z = x + y + z = 1$ (condición de holgura). De la primera restricción se obtiene $1 = 2x^2 + (1 - 2x)^2 = 6x^2 - 4x + 1$, i.e., $x^* = y^* = 0$ y $z^* = 1$, o $x^* = y^* = \frac{2}{3}$ y $z^* = -\frac{1}{3}$. El primer punto no es posible óptimo porque $\mu_2 = -2\mu_1 < 0$ (de la tercera). Con el segundo punto tenemos $\mu_2 = \frac{2}{3}\mu_1$, y de la primera obtenemos $\mu_1^* = \frac{1}{13}$ ($\mu_2^* = \frac{2}{39}$).

Ejemplos: Solución Problema 3 (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\boldsymbol{\lambda}, x, y, z) = \ln(x + y + 3) - \mu_1(x^2 + y^2 + z^2 - 1) - \mu_2(x + y + z - 1).$$

CPO:

$$\nabla_{x,y,z} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot, \cdot) = \begin{pmatrix} \frac{1}{x+y+3} - 2x\mu_1 - \mu_2 \\ \frac{1}{x+y+3} - 2y\mu_1 - \mu_2 \\ -2z\mu_1 - \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* \neq 0.$$

Si $\mu_2^* = 0$ entonces $z = 0$ (por la tercera). Restando la primera a la segunda obtenemos, $-2(y - x)\mu_1^* = 0$, i.e., $x = y$. De la primera restricción, $2x^2 = 1$, i.e., $x^* = y^* = \pm \frac{1}{\sqrt{2}}$. $\mu_1^* = \frac{-1}{\sqrt{2}(3-\sqrt{2})} < 0$ si $x, y < 0$, por lo que $x^* = y^* = \frac{1}{\sqrt{2}}$. No es admisible ($x^* + y^* + z^* = \sqrt{2} > 1$).

Si $\mu_2^* \neq 0$, entonces $2x + z = x + y + z = 1$ (condición de holgura). De la primera restricción se obtiene $1 = 2x^2 + (1 - 2x)^2 = 6x^2 - 4x + 1$, i.e., $x^* = y^* = 0$ y $z^* = 1$, o $x^* = y^* = \frac{2}{3}$ y $z^* = -\frac{1}{3}$. El primer punto no es posible óptimo porque $\mu_2 = -2\mu_1 < 0$ (de la tercera). Con el segundo punto tenemos $\mu_2 = \frac{2}{3}\mu_1$, y de la primera obtenemos $\mu_1^* = \frac{1}{13}$ ($\mu_2^* = \frac{2}{39}$). f cóncava y h convexas, i.e., $f\left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right) = \ln\left(\frac{13}{3}\right)$ es el máximo global.

Ejemplos: Solución Problema 4

Calificación de la Restricción (LICQ):

- Todas no pueden estar activas ($x + y = 6$, $x = 0$ y $y = 0$).

Ejemplos: Solución Problema 4

Calificación de la Restricción (LICQ):

- Todas no pueden estar activas ($x + y = 6$, $x = 0$ y $y = 0$).
- Si solo h_1 y h_2 está activa ($x + y = 6$, $x = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

Ejemplos: Solución Problema 4

Calificación de la Restricción (LICQ):

- Todas no pueden estar activas ($x + y = 6$, $x = 0$ y $y = 0$).
- Si solo h_1 y h_2 está activa ($x + y = 6$, $x = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

- Si solo h_1 y h_3 está activa ($x + y = 6$, $y = 0$): Análogo al anterior.

Ejemplos: Solución Problema 4

Calificación de la Restricción (LICQ):

- Todas no pueden estar activas ($x + y = 6$, $x = 0$ y $y = 0$).
- Si solo h_1 y h_2 está activa ($x + y = 6$, $x = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

- Si solo h_1 y h_3 está activa ($x + y = 6$, $y = 0$): Análogo al anterior.
- Si solo h_2 y h_3 está activa ($x = 0$, $y = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

Ejemplos: Solución Problema 4

Calificación de la Restricción (LICQ):

- Todas no pueden estar activas ($x + y = 6$, $x = 0$ y $y = 0$).
- Si solo h_1 y h_2 está activa ($x + y = 6$, $x = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

- Si solo h_1 y h_3 está activa ($x + y = 6$, $y = 0$): Análogo al anterior.
- Si solo h_2 y h_3 está activa ($x = 0$, $y = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

- Si solo h_1 está activa ($x + y = 6$): $\nabla h_1 = (1, 1) \neq (0, 0)$.

Ejemplos: Solución Problema 4

Calificación de la Restricción (LICQ):

- Todas no pueden estar activas ($x + y = 6$, $x = 0$ y $y = 0$).
- Si solo h_1 y h_2 está activa ($x + y = 6$, $x = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

- Si solo h_1 y h_3 está activa ($x + y = 6$, $y = 0$): Análogo al anterior.
- Si solo h_2 y h_3 está activa ($x = 0$, $y = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

- Si solo h_1 está activa ($x + y = 6$): $\nabla h_1 = (1, 1) \neq (0, 0)$.
- Si solo h_2 está activa ($x = 0$): $\nabla h_2 = (1, 0) \neq (0, 0)$.

Ejemplos: Solución Problema 4

Calificación de la Restricción (LICQ):

- Todas no pueden estar activas ($x + y = 6$, $x = 0$ y $y = 0$).
- Si solo h_1 y h_2 está activa ($x + y = 6$, $x = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

- Si solo h_1 y h_3 está activa ($x + y = 6$, $y = 0$): Análogo al anterior.
- Si solo h_2 y h_3 está activa ($x = 0$, $y = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

- Si solo h_1 está activa ($x + y = 6$): $\nabla h_1 = (1, 1) \neq (0, 0)$.
- Si solo h_2 está activa ($x = 0$): $\nabla h_2 = (1, 0) \neq (0, 0)$.
- Si solo h_3 está activa ($y = 0$): $\nabla h_3 = (0, 1) \neq (0, 0)$.

Ejemplos: Solución Problema 4

Calificación de la Restricción (LICQ):

- Todas no pueden estar activas ($x + y = 6$, $x = 0$ y $y = 0$).
- Si solo h_1 y h_2 está activa ($x + y = 6$, $x = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

- Si solo h_1 y h_3 está activa ($x + y = 6$, $y = 0$): Análogo al anterior.
- Si solo h_2 y h_3 está activa ($x = 0$, $y = 0$):

$$J_{\tilde{g}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

- Si solo h_1 está activa ($x + y = 6$): $\nabla h_1 = (1, 1) \neq (0, 0)$.
- Si solo h_2 está activa ($x = 0$): $\nabla h_2 = (1, 0) \neq (0, 0)$.
- Si solo h_3 está activa ($y = 0$): $\nabla h_3 = (0, 1) \neq (0, 0)$.

Conclusión: Todos los puntos califican la restricción.

Ejemplos: Solución Problema 4 (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\mu, x, y) = xy - \mu_1(x + y - 6) - \mu_2(-x) - \mu_3(-y).$$

Ejemplos: Solución Problema 4 (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\mu, x, y) = xy - \mu_1(x + y - 6) - \mu_2(-x) - \mu_3(-y).$$

CPO:

$$\nabla_{x,y}\mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot) = \begin{pmatrix} y - \mu_1 + \mu_2 \\ x - \mu_1 + \mu_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Ejemplos: Solución Problema 4 (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\mu, x, y) = xy - \mu_1(x + y - 6) - \mu_2(-x) - \mu_3(-y).$$

CPO:

$$\nabla_{x,y} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot) = \begin{pmatrix} y - \mu_1 + \mu_2 \\ x - \mu_1 + \mu_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Si $\mu_1^* = 0$ entonces $0 \leq \mu_2 = -y \leq 0$ y $0 \leq \mu_3 = -x \leq 0$ (por la primera y la segunda, respectivamente). El punto es $(0, 0, 0, 0, 0)$.

Ejemplos: Solución Problema 4 (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\mu, x, y) = xy - \mu_1(x + y - 6) - \mu_2(-x) - \mu_3(-y).$$

CPO:

$$\nabla_{x,y}\mathcal{L}(\lambda, \cdot, \cdot) = \begin{pmatrix} y - \mu_1 + \mu_2 \\ x - \mu_1 + \mu_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Si $\mu_1^* = 0$ entonces $0 \leq \mu_2 = -y \leq 0$ y $0 \leq \mu_3 = -x \leq 0$ (por la primera y la segunda, respectivamente). El punto es $(0, 0, 0, 0, 0)$.

Si $\mu_1^* \neq 0$ entonces $x + y = 6$. Como las tres restricciones no pueden estar saturadas al mismo tiempo, $\mu_2^* = 0$ o $\mu_3^* = 0$ (o ambas). Si $\mu_2^* \neq 0$ entonces $x = 0$, $y = 6$ y $\mu_3^* = 0$, i.e., de la segunda $\mu_1^* = 0$ ($\Rightarrow \Leftarrow$). Idem si $\mu_3^* \neq 0$. Se concluye que $\mu_2^* = 0 = \mu_3^*$ y $y = \mu_1^* = x$ (por la 1ra y 2da). De la 1ra restricción tenemos $\mu_1^* = x = y = 3$. El punto es $(3, 0, 0, 3, 3)$.

Ejemplos: Solución Problema 4 (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\mu, x, y) = xy - \mu_1(x + y - 6) - \mu_2(-x) - \mu_3(-y).$$

CPO:

$$\nabla_{x,y} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot) = \begin{pmatrix} y - \mu_1 + \mu_2 \\ x - \mu_1 + \mu_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Si $\mu_1^* = 0$ entonces $0 \leq \mu_2 = -y \leq 0$ y $0 \leq \mu_3 = -x \leq 0$ (por la primera y la segunda, respectivamente). El punto es $(0, 0, 0, 0, 0)$.

Si $\mu_1^* \neq 0$ entonces $x + y = 6$. Como las tres restricciones no pueden estar saturadas al mismo tiempo, $\mu_2^* = 0$ o $\mu_3^* = 0$ (o ambas). Si $\mu_2^* \neq 0$ entonces $x = 0$, $y = 6$ y $\mu_3^* = 0$, i.e., de la segunda $\mu_1^* = 0$ ($\Rightarrow \Leftarrow$). Idem si $\mu_3^* \neq 0$. Se concluye que $\mu_2^* = 0 = \mu_3^*$ y $y = \mu_1^* = x$ (por la 1ra y 2da). De la 1ra restricción tenemos $\mu_1^* = x = y = 3$. El punto es $(3, 0, 0, 3, 3)$.

Como f **no** es cóncava, no se puede aplicar el teorema de máximo global. Además, $\nabla f(0, 0) = (0, 0)$, por lo que no se puede aplicar Arrow-Enthoven.

Ejemplos: Solución Problema 4 (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\mu, x, y) = xy - \mu_1(x + y - 6) - \mu_2(-x) - \mu_3(-y).$$

CPO:

$$\nabla_{x,y} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot) = \begin{pmatrix} y - \mu_1 + \mu_2 \\ x - \mu_1 + \mu_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Si $\mu_1^* = 0$ entonces $0 \leq \mu_2 = -y \leq 0$ y $0 \leq \mu_3 = -x \leq 0$ (por la primera y la segunda, respectivamente). El punto es $(0, 0, 0, 0, 0)$.

Si $\mu_1^* \neq 0$ entonces $x + y = 6$. Como las tres restricciones no pueden estar saturadas al mismo tiempo, $\mu_2^* = 0$ o $\mu_3^* = 0$ (o ambas). Si $\mu_2^* \neq 0$ entonces $x = 0$, $y = 6$ y $\mu_3^* = 0$, i.e., de la segunda $\mu_1^* = 0$ ($\Rightarrow \Leftarrow$). Idem si $\mu_3^* \neq 0$. Se concluye que $\mu_2^* = 0 = \mu_3^*$ y $y = \mu_1^* = x$ (por la 1ra y 2da). De la 1ra restricción tenemos $\mu_1^* = x = y = 3$. El punto es $(3, 0, 0, 3, 3)$.

Como f **no** es cóncava, no se puede aplicar el teorema de máximo global. Además, $\nabla f(0, 0) = (0, 0)$, por lo que no se puede aplicar Arrow-Enthoven.

Como f es cuasicóncava, h son lineales, y $\nabla f(3, 3) = (3, 3) \neq (0, 0)$, por Arrow-Enthoven se concluye que $f(3, 3) = 9$ es el máximo global.

Condición Suficiente para Máximo Local en $\mathbb{R}^n$

Matriz Hessiana Orlada Activa

Sean $f, \mathbf{g}, \mathbf{h} \in \mathcal{C}^2(S)$. La *Hessiana orlada* de f y las restricciones $(\mathbf{g}, \mathbf{h})$ en

$$(\boldsymbol{\lambda}, \mathbf{x}) \text{ es } H_{f, \tilde{\mathbf{g}}}^O(\boldsymbol{\lambda}, \mathbf{x}) := \begin{pmatrix} \mathbf{0}_{(m+l) \times (m+l)} & -J_{\tilde{\mathbf{g}}}(\mathbf{x}) \\ -J_{\tilde{\mathbf{g}}}^T(\mathbf{x}) & H_f(\mathbf{x}) - \sum_{k=1}^m \lambda_k H_{g_k}(\mathbf{x}) - \sum_{i=1}^l \mu_i H_{h_i}(\mathbf{x}) \end{pmatrix}$$

donde $\tilde{\mathbf{g}} = (g_1, \dots, g_m, h_{i_1}, \dots, h_{i_l})$ son las restricciones activas en $\mathbf{x}$.

Condición Suficiente para Máximo Local en $\mathbb{R}^n$

Matriz Hessiana Orlada Activa

Sean $f, \mathbf{g}, \mathbf{h} \in \mathcal{C}^2(S)$. La Hessiana orlada de f y las restricciones $(\mathbf{g}, \mathbf{h})$ en

$$H_{f, \tilde{\mathbf{g}}}^O(\boldsymbol{\lambda}, \mathbf{x}) \text{ es } \begin{pmatrix} \mathbf{0}_{(m+l) \times (m+l)} & -J_{\tilde{\mathbf{g}}}(\mathbf{x}) \\ -J_{\tilde{\mathbf{g}}}^T(\mathbf{x}) & H_f(\mathbf{x}) - \sum_{k=1}^m \lambda_k H_{g_k}(\mathbf{x}) - \sum_{i=1}^l \mu_i H_{h_i}(\mathbf{x}) \end{pmatrix}$$

donde $\tilde{\mathbf{g}} = (g_1, \dots, g_m, h_{i_1}, \dots, h_{i_l})$ son las restricciones activas en $\mathbf{x}$.

Menores Principales Dominantes de $H_{f, \tilde{\mathbf{g}}}^O$

Los MPD de orden r de $H_{f, \tilde{\mathbf{g}}}^O$ son los determinantes $D_r^O(\boldsymbol{\lambda}, \mathbf{x})$ dados por

$$\begin{vmatrix} \mathbf{0}_{(m+l) \times (m+l)} & -J_{\tilde{\mathbf{g}}|x_1, \dots, x_r}(\mathbf{x}) \\ -J_{\tilde{\mathbf{g}}|x_1, \dots, x_r}^T(\mathbf{x}) & H_{f|x_1, \dots, x_r}(\mathbf{x}) - \sum_{k=1}^m \lambda_k H_{g_k|x_1, \dots, x_r}(\mathbf{x}) - \sum_{i=1}^l \mu_i H_{h_i|x_1, \dots, x_r}(\mathbf{x}) \end{vmatrix}$$

Condición Suficiente para Máximo Local en $\mathbb{R}^n$

Matriz Hessiana Orlada Activa

Sean $f, g, h \in \mathcal{C}^2(S)$. La Hessiana orlada de f y las restricciones (g, h) en

$$H_{f,\tilde{g}}^O(\lambda, x) := \begin{pmatrix} 0_{(m+l) \times (m+l)} & -J_{\tilde{g}}(x) \\ -J_{\tilde{g}}^T(x) & H_f(x) - \sum_{k=1}^m \lambda_k H_{g_k}(x) - \sum_{i=1}^l \mu_i H_{h_i}(x) \end{pmatrix}$$

donde $\tilde{g} = (g_1, \dots, g_m, h_{i_1}, \dots, h_{i_l})$ son las restricciones activas en x .

Menores Principales Dominantes de $H_{f,\tilde{g}}^O$

Los MPD de orden r de $H_{f,\tilde{g}}^O$ son los determinantes $D_r^O(\lambda, x)$ dados por

$$\begin{vmatrix} 0_{(m+l) \times (m+l)} & -J_{\tilde{g}|x_1, \dots, x_r}(x) \\ -J_{\tilde{g}|x_1, \dots, x_r}^T(x) & H_{f|x_1, \dots, x_r}(x) - \sum_{k=1}^m \lambda_k H_{g_k|x_1, \dots, x_r}(x) - \sum_{i=1}^l \mu_i H_{h_i|x_1, \dots, x_r}(x) \end{vmatrix}$$

Condición de Segundo Orden para Máximo Local en $\mathbb{R}^n$

Sean $f, g, h \in \mathcal{C}^2(S)$. Si $x^* \in S$ cumple las condiciones de primer orden, $f(x^*)$ es un máximo local si $(-1)^r D_r^O(\lambda^*, x^*) > 0$, $r = m + l + 1, \dots, n$.

Ejemplo: $n = 3$, $m = 0$ y $l = ?$

Resuelva el problema

$$\text{máx } f(x, y, z) = x + \ln(1 + z) \quad \text{sujeta a } \begin{cases} x^2 + y^2 \leq 1 \\ x + y + z \leq 1 \end{cases}.$$

Cuál sería el cambio estimado en el óptimo si se utiliza $x + y + z \leq 1,02$?

Ejemplo: $n = 3$, $m = 0$ y $l = ?$

Resuelva el problema

$$\text{máx } f(x, y, z) = x + \ln(1 + z) \quad \text{sujeta a } \begin{cases} x^2 + y^2 \leq 1 \\ x + y + z \leq 1 \end{cases}.$$

Cuál sería el cambio estimado en el óptimo si se utiliza $x + y + z \leq 1,02$?

Solución: Primero calificuemos la restricción:

- Si solo h_1 está activa ($x^2 + y^2 = 1$): $\nabla h_1 = (2x, 2y, 0) = (0, 0, 0)$ si $x = y = 0$, pero $h_1(0, 0, z) = 0 \neq 1$, por lo que todos los puntos admisibles califican la restricción.

Ejemplo: $n = 3$, $m = 0$ y $l = ?$

Resuelva el problema

$$\text{máx } f(x, y, z) = x + \ln(1 + z) \quad \text{sujeta a } \begin{cases} x^2 + y^2 \leq 1 \\ x + y + z \leq 1 \end{cases}.$$

Cuál sería el cambio estimado en el óptimo si se utiliza $x + y + z \leq 1,02$?

Solución: Primero calificuemos la restricción:

- Si solo h_1 está activa ($x^2 + y^2 = 1$): $\nabla h_1 = (2x, 2y, 0) = (0, 0, 0)$ si $x = y = 0$, pero $h_1(0, 0, z) = 0 \neq 1$, por lo que todos los puntos admisibles califican la restricción.
- Si solo h_2 está activa ($x + y + z = 1$): $\nabla h_2 = (1, 1, 1) \neq (0, 0, 0)$.

Ejemplo: $n = 3$, $m = 0$ y $l = ?$

Resuelva el problema

$$\text{máx } f(x, y, z) = x + \ln(1 + z) \quad \text{sujeta a } \begin{cases} x^2 + y^2 \leq 1 \\ x + y + z \leq 1 \end{cases}.$$

Cuál sería el cambio estimado en el óptimo si se utiliza $x + y + z \leq 1,02$?

Solución: Primero calificuemos la restricción:

- Si solo h_1 está activa ($x^2 + y^2 = 1$): $\nabla h_1 = (2x, 2y, 0) = (0, 0, 0)$ si $x = y = 0$, pero $h_1(0, 0, z) = 0 \neq 1$, por lo que todos los puntos admisibles califican la restricción.
- Si solo h_2 está activa ($x + y + z = 1$): $\nabla h_2 = (1, 1, 1) \neq (0, 0, 0)$.
- Si todas están activas ($x^2 + y^2 = 1$ y $x + y + z = 1$):

$$J_{\tilde{g}} = \begin{pmatrix} 2x & 2y & 0 \\ 1 & 1 & 1 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

Ejemplo: $n = 3$, $m = 0$ y $l = ?$

Resuelva el problema

$$\text{máx } f(x, y, z) = x + \ln(1 + z) \quad \text{sujeta a } \begin{cases} x^2 + y^2 \leq 1 \\ x + y + z \leq 1 \end{cases}.$$

Cuál sería el cambio estimado en el óptimo si se utiliza $x + y + z \leq 1,02$?

Solución: Primero calificuemos la restricción:

- Si solo h_1 está activa ($x^2 + y^2 = 1$): $\nabla h_1 = (2x, 2y, 0) = (0, 0, 0)$ si $x = y = 0$, pero $h_1(0, 0, z) = 0 \neq 1$, por lo que todos los puntos admisibles califican la restricción.
- Si solo h_2 está activa ($x + y + z = 1$): $\nabla h_2 = (1, 1, 1) \neq (0, 0, 0)$.
- Si todas están activas ($x^2 + y^2 = 1$ y $x + y + z = 1$):

$$J_{\tilde{g}} = \begin{pmatrix} 2x & 2y & 0 \\ 1 & 1 & 1 \end{pmatrix}; \quad \text{tiene rango 2 siempre.}$$

Conclusión: Todos los puntos califican la restricción.

Ejemplo: $n = 3$, $m = 0$ y $l = ?$ (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\boldsymbol{\lambda}, x, y, z) = x + \ln(1 + z) - \mu_1(x^2 + y^2 - 1) - \mu_2(x + y + z - 1).$$

Ejemplo: $n = 3$, $m = 0$ y $l = ?$ (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\boldsymbol{\lambda}, x, y, z) = x + \ln(1 + z) - \mu_1(x^2 + y^2 - 1) - \mu_2(x + y + z - 1).$$

CPO:

$$\nabla_{x,y,z} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot, \cdot) = \begin{pmatrix} 1 - 2x\mu_1 - \mu_2 \\ -2y\mu_1 - \mu_2 \\ \frac{1}{1+z} - \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* \neq 0.$$

Ejemplo: $n = 3$, $m = 0$ y $l = ?$ (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\boldsymbol{\lambda}, x, y, z) = x + \ln(1 + z) - \mu_1(x^2 + y^2 - 1) - \mu_2(x + y + z - 1).$$

CPO:

$$\nabla_{x,y,z} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot, \cdot) = \begin{pmatrix} 1 - 2x\mu_1 - \mu_2 \\ -2y\mu_1 - \mu_2 \\ \frac{1}{1+z} - \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* \neq 0.$$

De la primera y la segunda tenemos $\mu_1^* = \frac{1}{2(x-y)}$ y $\mu_2^* = -\frac{y}{x-y}$. Si $x < y = 0$ entonces $\mu_1^* < 0$ ($\Rightarrow \Leftarrow$). Como $x > y$, se concluye que $y \leq 0$ para que $\mu_2^* \geq 0$. Si $y = 0$ entonces $\mu_2^* = 0$ y de la tercera $\frac{1}{1+z} = 0$ ($\Rightarrow \Leftarrow$).

Ejemplo: $n = 3$, $m = 0$ y $l = ?$ (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\boldsymbol{\lambda}, x, y, z) = x + \ln(1 + z) - \mu_1(x^2 + y^2 - 1) - \mu_2(x + y + z - 1).$$

CPO:

$$\nabla_{x,y,z} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot, \cdot) = \begin{pmatrix} 1 - 2x\mu_1 - \mu_2 \\ -2y\mu_1 - \mu_2 \\ \frac{1}{1+z} - \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* \neq 0.$$

De la primera y la segunda tenemos $\mu_1^* = \frac{1}{2(x-y)}$ y $\mu_2^* = -\frac{y}{x-y}$. Si $x < y = 0$ entonces $\mu_1^* < 0$ ($\Rightarrow \Leftarrow$). Como $x > y$, se concluye que $y \leq 0$ para que $\mu_2^* \geq 0$. Si $y = 0$ entonces $\mu_2^* = 0$ y de la tercera $\frac{1}{1+z} = 0$ ($\Rightarrow \Leftarrow$). Como $\mu_1^* \neq 0$ y $\mu_2^* \neq 0$, h_1 y h_2 deben estar saturadas, i.e., $x^2 + y^2 = 1$ y $x + y + z = 1$. Por la tercera $\frac{1}{1+z} = \mu_2^* = \frac{-y}{x-y}$ ergo $z = \frac{-x}{y}$. Utilizando la restricción saturada h_2 , i.e. $x = 1 - y - z$, se tiene que $z = \frac{-1+y+z}{y} \Rightarrow (z-1)(y-1) = 0$. Como $y < 0$, entonces $z = 1$.

Ejemplo: $n = 3$, $m = 0$ y $l = ?$ (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\lambda, x, y, z) = x + \ln(1 + z) - \mu_1(x^2 + y^2 - 1) - \mu_2(x + y + z - 1).$$

CPO:

$$\nabla_{x,y,z} \mathcal{L}(\lambda, \cdot, \cdot, \cdot) = \begin{pmatrix} 1 - 2x\mu_1 - \mu_2 \\ -2y\mu_1 - \mu_2 \\ \frac{1}{1+z} - \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* \neq 0.$$

De la primera y la segunda tenemos $\mu_1^* = \frac{1}{2(x-y)}$ y $\mu_2^* = -\frac{y}{x-y}$. Si $x < y = 0$ entonces $\mu_1^* < 0$ ($\Rightarrow \Leftarrow$). Como $x > y$, se concluye que $y \leq 0$ para que $\mu_2^* \geq 0$. Si $y = 0$ entonces $\mu_2^* = 0$ y de la tercera $\frac{1}{1+z} = 0$ ($\Rightarrow \Leftarrow$). Como $\mu_1^* \neq 0$ y $\mu_2^* \neq 0$, h_1 y h_2 deben estar saturadas, i.e., $x^2 + y^2 = 1$ y $x + y + z = 1$. Por la tercera $\frac{1}{1+z} = \mu_2^* = \frac{-y}{x-y}$ ergo $z = \frac{-x}{y}$. Utilizando la restricción saturada h_2 , i.e. $x = 1 - y - z$, se tiene que $z = \frac{-1+y+z}{y} \Rightarrow (z-1)(y-1) = 0$. Como $y < 0$, entonces $z = 1$. Si $z = 1$ entonces $x = -y > 0 \Rightarrow x = \frac{1}{\sqrt{2}}, y = \frac{-1}{\sqrt{2}}$.

Ejemplo: $n = 3$, $m = 0$ y $l = ?$ (Continuación)

Planteamos el Lagrangiano:

$$\mathcal{L}(\boldsymbol{\lambda}, x, y, z) = x + \ln(1 + z) - \mu_1(x^2 + y^2 - 1) - \mu_2(x + y + z - 1).$$

CPO:

$$\nabla_{x,y,z} \mathcal{L}(\boldsymbol{\lambda}, \cdot, \cdot, \cdot) = \begin{pmatrix} 1 - 2x\mu_1 - \mu_2 \\ -2y\mu_1 - \mu_2 \\ \frac{1}{1+z} - \mu_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies \mu_1^* \neq 0.$$

De la primera y la segunda tenemos $\mu_1^* = \frac{1}{2(x-y)}$ y $\mu_2^* = -\frac{y}{x-y}$. Si $x < y = 0$ entonces $\mu_1^* < 0$ ($\Rightarrow \Leftarrow$). Como $x > y$, se concluye que $y \leq 0$ para que $\mu_2^* \geq 0$. Si $y = 0$ entonces $\mu_2^* = 0$ y de la tercera $\frac{1}{1+z} = 0$ ($\Rightarrow \Leftarrow$). Como $\mu_1^* \neq 0$ y $\mu_2^* \neq 0$, h_1 y h_2 deben estar saturadas, i.e., $x^2 + y^2 = 1$ y $x + y + z = 1$. Por la tercera $\frac{1}{1+z} = \mu_2^* = \frac{-y}{x-y}$ ergo $z = \frac{-x}{y}$. Utilizando la restricción saturada h_2 , i.e. $x = 1 - y - z$, se tiene que $z = \frac{-1+y+z}{y} \Rightarrow (z-1)(y-1) = 0$. Como $y < 0$, entonces $z = 1$. Si $z = 1$ entonces $x = -y > 0 \Rightarrow x = \frac{1}{\sqrt{2}}, y = \frac{-1}{\sqrt{2}}$.

El único punto factible sería $\left(\frac{1}{2\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 1\right)$.

Ejemplo: $n = 3$, $m = 0$ y $l = 2$ (Continuación)

Planteamos el Hessiano orlado con dos restricciones activas:

$$H_{f,\tilde{g}}^O(\mu_1, \mu_2, x, y, z) = \begin{pmatrix} 0 & 0 & 2x & 2y & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 2x & 1 & -2\mu_1 & 0 & 0 \\ 2y & 1 & 0 & -2\mu_1 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{(1+z)^2} \end{pmatrix}$$

$$\implies D_3^O = -8w(x^2 + y^2) - \frac{4(x - y)^2}{(z + 1)^2}$$

Ejemplo: $n = 3$, $m = 0$ y $l = 2$ (Continuación)

Planteamos el Hessiano orlado con dos restricciones activas:

$$H_{f,\tilde{g}}^O(\mu_1, \mu_2, x, y, z) = \begin{pmatrix} 0 & 0 & 2x & 2y & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 2x & 1 & -2\mu_1 & 0 & 0 \\ 2y & 1 & 0 & -2\mu_1 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{(1+z)^2} \end{pmatrix}$$

$$\Rightarrow D_3^O = -8w(x^2 + y^2) - \frac{4(x - y)^2}{(z + 1)^2}$$

Condición de Segundo Orden:

$$(-1)^3 D_3^O \left(\frac{1}{2\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 1 \right) = 2(\sqrt{2} + 1) > 0 \quad \Rightarrow \quad \text{Máximo Local.}$$

Ejemplo: $n = 3$, $m = 0$ y $l = 2$ (Continuación)

Planteamos el Hessiano orlado con dos restricciones activas:

$$H_{f,\tilde{g}}^O(\mu_1, \mu_2, x, y, z) = \begin{pmatrix} 0 & 0 & 2x & 2y & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 2x & 1 & -2\mu_1 & 0 & 0 \\ 2y & 1 & 0 & -2\mu_1 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{(1+z)^2} \end{pmatrix}$$

$$\implies D_3^O = -8w(x^2 + y^2) - \frac{4(x - y)^2}{(z + 1)^2}$$

Condición de Segundo Orden:

$$(-1)^3 D_3^O \left(\frac{1}{2\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 1 \right) = 2(\sqrt{2} + 1) > 0 \quad \Rightarrow \quad \text{Máximo Local.}$$

NOTA: Como $r = m + l + 1 = 0 + 2 + 1 = 3 = n$, sólo se calcula D_3^O .

Ejemplo: $n = 3$, $m = 0$ y $l = 2$ (Continuación)

Planteamos el Hessiano orlado con dos restricciones activas:

$$H_{f,\tilde{g}}^O(\mu_1, \mu_2, x, y, z) = \begin{pmatrix} 0 & 0 & 2x & 2y & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 2x & 1 & -2\mu_1 & 0 & 0 \\ 2y & 1 & 0 & -2\mu_1 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{(1+z)^2} \end{pmatrix}$$

$$\Rightarrow D_3^O = -8w(x^2 + y^2) - \frac{4(x - y)^2}{(z + 1)^2}$$

Condición de Segundo Orden:

$$(-1)^3 D_3^O \left(\frac{1}{2\sqrt{2}}, \frac{1}{2}, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 1 \right) = 2(\sqrt{2} + 1) > 0 \Rightarrow \text{Máximo Local.}$$

NOTA: Como $r = m + l + 1 = 0 + 2 + 1 = 3 = n$, sólo se calcula D_3^O .

El cambio estimado en el óptimo si se utiliza $x + y + z \leq 1,02$ sería

$$\mu_2 \cdot \Delta c \approx \frac{1}{2} \cdot 0,2 = 0,01. \quad (\text{porque esta activa!})$$

Resumen del Procedimiento

① Condiciones de Primer Orden:

- ① **LICQ:** Calcule $J_{\tilde{g}}$ para las distintas restricciones activas y encuentre los $\tilde{x}$ admisibles en donde $J_{\tilde{g}}$ no existe o no tiene rango $(m + l)$.

Resumen del Procedimiento

1 Condiciones de Primer Orden:

- 1 **LICQ:** Calcule $J_{\tilde{g}}$ para las distintas restricciones activas y encuentre los $\tilde{x}$ admisibles en donde $J_{\tilde{g}}$ no existe o no tiene rango $(m + l)$.
- 2 **Lagrangiano:** $\mathcal{L}(\lambda, x) = f(x) - \sum_{i=1}^m \lambda_i g_i(x) - \sum_{j=1}^k \mu_j h_j(x)$.
Halle (λ^*, x^*) t.q. $\nabla_x \mathcal{L}(\lambda^*, x^*) = 0$, $g_i(x^*) = c_i$ y $\mu_j^* = 0$ si $h_j(x^*) < b_j$.

Resumen del Procedimiento

1 Condiciones de Primer Orden:

- 1 **LICQ:** Calcule $J_{\tilde{g}}$ para las distintas restricciones activas y encuentre los $\tilde{x}$ admisibles en donde $J_{\tilde{g}}$ no existe o no tiene rango $(m + l)$.
- 2 **Lagrangiano:** $\mathcal{L}(\lambda, x) = f(x) - \sum_{i=1}^m \lambda_i g_i(x) - \sum_{j=1}^k \mu_j h_j(x)$.
Halle (λ^*, x^*) t.q. $\nabla_x \mathcal{L}(\lambda^*, x^*) = 0$, $g_i(x^*) = c_i$ y $\mu_j^* = 0$ si $h_j(x^*) < b_j$.

2 Condiciones de Segundo Orden: 2 opciones: máximo global o local.

Resumen del Procedimiento

1 Condiciones de Primer Orden:

- 1 **LICQ:** Calcule $J_{\tilde{g}}$ para las distintas restricciones activas y encuentre los $\tilde{x}$ admisibles en donde $J_{\tilde{g}}$ no existe o no tiene rango $(m + l)$.
- 2 **Lagrangiano:** $\mathcal{L}(\lambda, x) = f(x) - \sum_{i=1}^m \lambda_i g_i(x) - \sum_{j=1}^k \mu_j h_j(x)$.
Halle (λ^*, x^*) t.q. $\nabla_x \mathcal{L}(\lambda^*, x^*) = 0$, $g_i(x^*) = c_i$ y $\mu_j^* = 0$ si $h_j(x^*) < b_j$.

2 Condiciones de Segundo Orden: 2 opciones: máximo global o local.

- Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava: Para $H_f - \sum_{i=1}^m \lambda_i^* H_{g_i} - \sum_{j=1}^k \mu_j^* H_{h_j}$ verifique los MPA ó pruebe la convexidad/concavidad en todo el dominio de $f, \lambda_i^* g_i, \mu_j^* h_j$, así el óptimo es un máximo global.

Resumen del Procedimiento

1 Condiciones de Primer Orden:

- 1 **LICQ:** Calcule $J_{\tilde{g}}$ para las distintas restricciones activas y encuentre los $\tilde{x}$ admisibles en donde $J_{\tilde{g}}$ no existe o no tiene rango ($m + l$).
- 2 **Lagrangiano:** $\mathcal{L}(\lambda, x) = f(x) - \sum_{i=1}^m \lambda_i g_i(x) - \sum_{j=1}^k \mu_j h_j(x)$.
Halle (λ^*, x^*) t.q. $\nabla_x \mathcal{L}(\lambda^*, x^*) = 0$, $g_i(x^*) = c_i$ y $\mu_j^* = 0$ si $h_j(x^*) < b_j$.

2 Condiciones de Segundo Orden: 2 opciones: máximo global o local.

- **Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava:** Para $H_f - \sum_{i=1}^m \lambda_i^* H_{g_i} - \sum_{j=1}^k \mu_j^* H_{h_j}$ verifique los MPA ó pruebe la convexidad/concavidad en todo el dominio de $f, \lambda_i^* g_i, \mu_j^* h_j$, así el óptimo es un máximo global.
- **Si f es cuasicóncava y $\lambda_i^* g_i, \mu_j^* h_j$ son cuasiconvexas:** Para probar la cuasiconvexidad/cuasiconcavidad de $f, \lambda_i^* g_i, \mu_j^* h_j$ verifique los MPA de $H_f^O, H_{\lambda_i^* g_i}^O$ y $H_{\mu_j^* h_j}^O$, y si $\nabla f(x^*) \neq 0$, el óptimo es global.

Resumen del Procedimiento

1 Condiciones de Primer Orden:

- 1 **LICQ:** Calcule $J_{\tilde{g}}$ para las distintas restricciones activas y encuentre los $\tilde{x}$ admisibles en donde $J_{\tilde{g}}$ no existe o no tiene rango $(m + l)$.
- 2 **Lagrangiano:** $\mathcal{L}(\lambda, x) = f(x) - \sum_{i=1}^m \lambda_i g_i(x) - \sum_{j=1}^k \mu_j h_j(x)$.
Halle (λ^*, x^*) t.q. $\nabla_x \mathcal{L}(\lambda^*, x^*) = 0$, $g_i(x^*) = c_i$ y $\mu_j^* = 0$ si $h_j(x^*) < b_j$.

2 Condiciones de Segundo Orden: 2 opciones: máximo global o local.

- Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava: Para $H_f - \sum_{i=1}^m \lambda_i^* H_{g_i} - \sum_{j=1}^k \mu_j^* H_{h_j}$ verifique los MPA ó pruebe la convexidad/concavidad en todo el dominio de $f, \lambda_i^* g_i, \mu_j^* h_j$, así el óptimo es un máximo global.
- Si f es cuasicóncava y $\lambda_i^* g_i, \mu_j^* h_j$ son cuasiconvexas: Para probar la cuasiconvexidad/cuasiconcavidad de $f, \lambda_i^* g_i, \mu_j^* h_j$ verifique los MPA de $H_f^O, H_{\lambda_i^* g_i}^O$ y $H_{\mu_j^* h_j}^O$, y si $\nabla f(x^*) \neq 0$, el óptimo es global.
- Si f no es cuasicóncava ó $\lambda_i^* g_i, \mu_j^* h_j$ no son cuasiconvexas:
Primero calcule la Hessiana $H_{f, \tilde{g}}^O$. Verifique los MPD en los puntos encontrados $D_r^O(\lambda^*, x^*)$, para saber si el óptimo es un máximo local.

Resumen del Procedimiento

1 Condiciones de Primer Orden:

- 1 **LICQ:** Calcule $J_{\tilde{g}}$ para las distintas restricciones activas y encuentre los $\tilde{x}$ admisibles en donde $J_{\tilde{g}}$ no existe o no tiene rango $(m + l)$.
- 2 **Lagrangiano:** $\mathcal{L}(\lambda, x) = f(x) - \sum_{i=1}^m \lambda_i g_i(x) - \sum_{j=1}^k \mu_j h_j(x)$.
Halle (λ^*, x^*) t.q. $\nabla_x \mathcal{L}(\lambda^*, x^*) = 0$, $g_i(x^*) = c_i$ y $\mu_j^* = 0$ si $h_j(x^*) < b_j$.

2 Condiciones de Segundo Orden: 2 opciones: máximo global o local.

- Si $\mathcal{L}(\lambda^*, \cdot)$ es cóncava: Para $H_f - \sum_{i=1}^m \lambda_i^* H_{g_i} - \sum_{j=1}^k \mu_j^* H_{h_j}$ verifique los MPA ó pruebe la convexidad/concavidad en todo el dominio de $f, \lambda_i^* g_i, \mu_j^* h_j$, así el óptimo es un máximo global.
- Si f es cuasicóncava y $\lambda_i^* g_i, \mu_j^* h_j$ son cuasiconvexas: Para probar la cuasiconvexidad/cuasiconcavidad de $f, \lambda_i^* g_i, \mu_j^* h_j$ verifique los MPA de $H_f^O, H_{\lambda_i^* g_i}^O$ y $H_{\mu_j^* h_j}^O$, y si $\nabla f(x^*) \neq 0$, el óptimo es global.
- Si f no es cuasicóncava ó $\lambda_i^* g_i, \mu_j^* h_j$ no son cuasiconvexas: Primero calcule la Hessiana $H_{f, \tilde{g}}^O$. Verifique los MPD en los puntos encontrados $D_r^O(\lambda^*, x^*)$, para saber si el óptimo es un máximo local.

- 3 **Óptimos:** Si $f(\tilde{x})$ es mayor que los máximos de la segunda parte, entonces se agrega $\tilde{x}$ a la canasta de óptimos hallados.

References



Sydsæter, K., Hammond, P., Soto Prieto, M.J. y Vicente Córdoba, J.L. (1998)

Matemáticas para el análisis económico.

Prentice Hall, Madrid.

Final de clase