

Optimización Matemática

Sesión 1

Organización de la asignatura

1. FUNDAMENTOS DE OPTIMIZACIÓN.
2. PROGRAMACIÓN CUADRÁTICA.
3. PROGRAMACIÓN CONVEXA.
4. OPTIMIZACIÓN SIN RESTRICCIONES: APLICACIONES.
5. OPTIMIZACIÓN CON RESTRICCIONES DE IGUALDAD: APLICACIONES.
6. OPTIMIZACIÓN CON RESTRICCIONES DE DESIGUALDAD: APLICACIONES.
7. PROGRAMACIÓN LINEAL. MÉTODO DEL SIMPLEX.
8. INTRODUCCIÓN ECUACIONES DIFERENCIALES.

Bibliografía

- R. BARBOLLA, E. CERDÁ y P. SANZ. Optimización: Cuestiones, ejercicios y aplicaciones a la Economía. Prentice Hall, 2001.
- HAEUSSLER, E.F., PAUL, R.S., WOOD, R.J. Matemáticas para administración y economía 12ª Edición. Pearson/ Prentice Hall, 2008

Método de evaluación

- Participación en clase de teoría y realización de pequeñas cuestiones teóricas 5 %

- Ejercicios propuestos en clases prácticas 15 %
- Trabajo sobre caso práctico 15 %
- Examen Parcial 15 %
- Examen final (cuestiones teóricas y problemas cortos) 25 %
- Examen final (problemas aplicados) 25 %

Motivación de la asignatura

1. Consideremos tres factorías de papel cada una de las cuales vierte al río arsénico y mercurio. Si en la factoría 1, el coste de procesado de una tonelada de papel por valor de 1450 euros reduce la cantidad de arsénico a 0.1 por tonelada y la de mercurio a 0.32 por tonelada. Para la factoría 2 los números son de 1210 euros y 0.15 por tonelada de arsénico y 0.2 por tonelada de mercurio. Para la factoría 3, 2200 euros y 0.3 por tonelada de arsénico y 0.3 por tonelada de mercurio. La regulación estipula qque el arsénico necesita reducirse por debajo de las 200 toneladas y el mercurio por debajo de las 300 toneladas. Plantea el problema que conteste a la siguiente pregunta. *¿Cómo podemos cumplir la norma a un coste mínimo?*
2. Para una factoría que fabrica camiones y automóviles, cada vehículo pasa por un módulo de ensamblado y otro de pintura. Si el módulo de pintura se dedica exclusivamente a la fabricación de automóviles, se pueden terminar 500 automóviles diarios. Por otro lado, si se dedica exclusivamente al pintado de camiones, 400 de ellos pueden ser terminados. Para el ensamblado los números son de 80 automóviles y 65 camiones diarios. De cada automóvil se obtiene un beneficio de 1300 euros y de cada camión 1500 euros. Plantea el problema que conteste a la siguiente pregunta. *¿Cuántos camiones y coches debería producir la factoría para maximizar el beneficio?*

Fundamentos de Optimización

Contenidos

Funciones de varias variables. Derivadas parciales. Aplicaciones e interpretación. Gradiente. Condición necesaria de primer orden de extremo local.

Motivación

Cuestión: ¿Podemos producir en cada una de las factorías de papel las cantidades necesarias para cumplir exactamente la regulación?

Solución: Necesitamos resolver:

$$0,10x_1 + 0,15x_2 + 0,30x_3 = 200$$

$$0,22x_1 + 0,20x_2 + 0,30x_3 = 300.$$

es decir

$$\begin{pmatrix} 0,10 & 0,15 & 0,30 \\ 0,22 & 0,20 & 0,30 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 200 \\ 300 \end{pmatrix}$$

que de forma simplificada podemos escribir

$$A\mathbf{x} = \mathbf{b}.$$

El sistema tiene infinitas soluciones ya que hay más incógnitas que ecuaciones. La pregunta es: **¿Como podemos calcular de forma justificada un valor para cada una de las incógnitas?**

Observar que tenemos dos posiciones en el plano

$$\begin{pmatrix} 200 \\ 300 \end{pmatrix} \text{ y } \begin{pmatrix} 0,10x_1 + 0,15x_2 + 0,30x_3 \\ 0,22x_1 + 0,20x_2 + 0,30x_3 \end{pmatrix}$$

Problema: Calcular los valores x_1, x_2 y x_3 que hacen que la distancia entre los puntos del plano

$$\begin{pmatrix} 200 \\ 300 \end{pmatrix} \text{ y } \begin{pmatrix} 0,10x_1 + 0,15x_2 + 0,30x_3 \\ 0,22x_1 + 0,20x_2 + 0,30x_3 \end{pmatrix}$$

sea la mínima posible?

Recordemos que

▪ $\text{distancia}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$

- En el plano

$$\|(u_1, u_2) - (v_1, v_2)\| = \|(u_1 - v_1, u_2 - v_2)\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2}$$

- En el espacio

$$\begin{aligned} \|(u_1, u_2, u_3) - (v_1, v_2, v_3)\| &= \|(u_1 - v_1, u_2 - v_2, u_3 - v_3)\| \\ &= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2} \end{aligned}$$

Para evitar las raices cuadradas se emplea el cuadrado de la distancia:

▪ $\text{distancia}(\mathbf{u}, \mathbf{v})^2 = \|\mathbf{u} - \mathbf{v}\|^2.$

- En el plano

$$\|(u_1, u_2) - (v_1, v_2)\|^2 = \|(u_1 - v_1, u_2 - v_2)\|^2 = (u_1 - v_1)^2 + (u_2 - v_2)^2$$

- En el espacio

$$\begin{aligned}\|(u_1, u_2, u_3) - (v_1, v_2, v_3)\|^2 &= \|(u_1 - v_1, u_2 - v_2, u_3 - v_3)\|^2 \\ &= (u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2\end{aligned}$$

Entonces

$$\begin{aligned}\text{distancia} &\left(\begin{pmatrix} 0,10x_1 + 0,15x_2 + 0,30x_3 \\ 0,22x_1 + 0,20x_2 + 0,30x_3 \end{pmatrix}, \begin{pmatrix} 200 \\ 300 \end{pmatrix} \right)^2 \\ &= \left\| \begin{pmatrix} 0,10x_1 + 0,15x_2 + 0,30x_3 \\ 0,22x_1 + 0,20x_2 + 0,30x_3 \end{pmatrix} - \begin{pmatrix} 200 \\ 300 \end{pmatrix} \right\|^2 \\ &= \left\| \begin{pmatrix} 0,10x_1 + 0,15x_2 + 0,30x_3 - 200 \\ 0,22x_1 + 0,20x_2 + 0,30x_3 - 300 \end{pmatrix} \right\|^2 \\ &= (0,10x_1 + 0,15x_2 + 0,30x_3 - 200)^2 + (0,22x_1 + 0,20x_2 + 0,30x_3 - 300)^2\end{aligned}$$

Tenemos pues una función

$$f(x_1, x_2, x_3) = (0,10x_1 + 0,15x_2 + 0,30x_3 - 200)^2 + (0,22x_1 + 0,20x_2 + 0,30x_3 - 300)^2$$

que depende de tres variables x_1, x_2 y x_3 de la que queremos calcular

$$\min_{(x_1, x_2, x_3)} (0,10x_1 + 0,15x_2 + 0,30x_3 - 200)^2 + (0,22x_1 + 0,20x_2 + 0,30x_3 - 300)^2$$

en forma general

$$\min_{(x_1, x_2, x_3)} f(x_1, x_2, x_3).$$

- Vamos a considerar funciones de dos o tres variables: $f(x_1, x_2)$ o $f(x_1, x_2, x_3)$.
- **Objetivo:** Dar condiciones para resolver

$$\min_{(x_1, x_2)} f(x_1, x_2) \text{ o } \min_{(x_1, x_2, x_3)} f(x_1, x_2, x_3).$$

El crecimiento viene determinado por la diferencia

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2, x_3 + \Delta x_3) - f(x_1, x_2, x_3)$$

para valores $\Delta x_1 \approx 0$, $\Delta x_2 \approx 0$, y $\Delta x_3 \approx 0$. El valor mínimo (x_1^*, x_2^*, x_3^*) tiene que cumplir

$$f(x_1^* + \Delta x_1, x_2^* + \Delta x_2, x_3^* + \Delta x_3) \geq f(x_1^*, x_2^*, x_3^*)$$

para cualquier valor $\Delta x_1 \approx 0$, $\Delta x_2 \approx 0$, y $\Delta x_3 \approx 0$, incluido 0. En particular, se tiene

$$f(x_1^* + \Delta x_1, x_2^*, x_3^*) \geq f(x_1^*, x_2^*, x_3^*) \quad (1)$$

$$f(x_1^*, x_2^* + \Delta x_2, x_3^*) \geq f(x_1^*, x_2^*, x_3^*) \quad (2)$$

y

$$f(x_1^*, x_2^*, x_3^* + \Delta x_3) \geq f(x_1^*, x_2^*, x_3^*) \quad (3)$$

Podemos entonces considerar las diferencias

$$f(x_1 + \Delta x_1, x_2, x_3) - f(x_1, x_2, x_3) = f_1(x_1 + \Delta x_1) - f_1(x_1) \quad (4)$$

considerando x_2 y x_3 valores fijos,

$$f(x_1, x_2 + \Delta x_2, x_3) - f(x_1, x_2, x_3) = f_2(x_2 + \Delta x_2) - f_2(x_2) \quad (5)$$

considerando x_1 y x_2 valores fijos, y finalmente

$$f(x_1, x_2, x_3 + \Delta x_3) - f(x_1, x_2, x_3) = f_3(x_3 + \Delta x_3) - f_3(x_3) \quad (6)$$

considerando x_1 y x_2 valores fijos. Entonces

$$\frac{f(x_1 + \Delta x_1, x_2, x_3) - f(x_1, x_2, x_3)}{\Delta x_1} = \frac{f_1(x_1 + \Delta x_1) - f_1(x_1)}{\Delta x_1} \xrightarrow{\Delta x_1 \rightarrow 0} f'_1(x_1) \quad (7)$$

$$\frac{f(x_1, x_2 + \Delta x_2, x_3) - f(x_1, x_2, x_3)}{\Delta x_2} = \frac{f_2(x_2 + \Delta x_2) - f_2(x_2)}{\Delta x_2} \xrightarrow{\Delta x_2 \rightarrow 0} f'_2(x_2) \quad (8)$$

y

$$\frac{f(x_1, x_2, x_3 + \Delta x_3) - f(x_1, x_2, x_3)}{\Delta x_3} = \frac{f_3(x_3 + \Delta x_3) - f_3(x_3)}{\Delta x_3} \xrightarrow{\Delta x_3 \rightarrow 0} f'_3(x_3) \quad (9)$$

Veamos un ejemplo con la función

$$f(x_1, x_2, x_3) = (0,10x_1 + 0,15x_2 + 0,30x_3 - 200)^2 + (0,22x_1 + 0,20x_2 + 0,30x_3 - 300)^2.$$

Si la denotamos por

$$f_2(x_2) := (0,10x_1 + 0,15x_2 + 0,30x_3 - 200)^2 + (0,22x_1 + 0,20x_2 + 0,30x_3 - 300)^2,$$

para declarar que solo depende de la variable x_2 . Derivando con respecto la variable, este caso x_2 , obtenemos

$$f'_2(x_2) := 2(0,10x_1 + 0,15x_2 + 0,30x_3 - 200)0,15 + 2(0,22x_1 + 0,20x_2 + 0,30x_3 - 300)0,20.$$

Como realmente no es una función nueva se emplea la notación

$$\frac{\partial f}{\partial x_2} := 2(0,10x_1 + 0,15x_2 + 0,30x_3 - 200)0,15 + 2(0,22x_1 + 0,20x_2 + 0,30x_3 - 300)0,20.$$

- A $\frac{\partial f}{\partial x_2}$ se le llama *derivada parcial de la función f con respecto a la variable x_2*

Formalmente no es ni más ni menos que calcular

$$\frac{\partial f}{\partial x_2}(x_1, x_2, x_3) := \lim_{\Delta x_2 \rightarrow 0} \frac{f(x_1, x_2 + \Delta x_2, x_3) - f(x_1, x_2, x_3)}{\Delta x_2},$$

siempre y cuando el límite exista.

Funciones de Utilidad

Una *función de utilidad* es una función real que mide la *satisfacción* o *utilidad* obtenida por un consumidor cuando disfruta vía consumo de cierta cantidad de bienes.

Matemáticamente puede demostrarse que si es posible modelizar la conducta de un consumidor perfectamente racional mediante funciones de utilidad convexa, entonces esta conducta puede resumirse mediante una curva de demanda decreciente. Más sencillamente, si existe una función de utilidad para el consumidor racional y se dan unos supuestos matemáticamente razonables entonces existe una *curva de demanda*.

Dada una economía en que un consumidor puede adquirir n mercancías diferentes (las cuales se suponen infinitamente divisibles o altamente divisibles), la función de utilidad se define como:

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad (Q_1, \dots, Q_n) \mapsto U = f(Q_1, \dots, Q_n)$$

Donde: $Q_i \geq 0$ se interpreta como la cantidad disponible del bien económico i -ésimo. $U \geq 0$ se interpreta como la utilidad total de una cierta combinación de bienes. Recordar que

$$\mathbb{R}^n = \{(x_1, \dots, x_n) : x_i \in \mathbb{R}\}.$$

Algunas propiedades usualmente requeridas son que:

(1) **Diferenciabilidad**, usualmente se supone que la función anterior es no sólo continua sino también diferenciable.

(2) **Monotonicidad**

$$f(Q_1, \dots, Q_i, \dots, Q_n) \leq f(Q_1, \dots, Q'_i, \dots, Q_n)$$

si $Q_i \leq Q'_i$. Si la función es monótona creciente entonces todas las derivadas parciales serán positivas o cero.

(3) **Convexidad**, si la función es convexa esto implicará cualquier mínimo relativo será un mínimo absoluto.

La condición (1) es mera conveniencia matemática, la condición (2) es importante debe ser satisfecha por toda función de utilidad, mientras que la condición (3) tiene que ver con el principio de utilidad marginal decreciente.

Derivadas parciales, diferencial y gradiente

- Una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es diferenciable si podemos calcular sus derivadas parciales con respecto cada una de sus variables, y estas son funciones continuas.

Si por ejemplo tenemos $f(x_1, x_2) = 2x_1 + 3x_1x_2^2 + 1$

$$\begin{aligned}\frac{\partial f}{\partial x_1}(x_1, x_2) &= 2 + 3x_2^2, \\ \frac{\partial f}{\partial x_2}(x_1, x_2) &= 6x_1x_2.\end{aligned}$$

Si tomamos $x_1 = 2$ y $x_2 = 3$ se tiene $f(2, 3) = 2 \cdot 2 + 3 \cdot 2 \cdot 3^2 = 58$.

- Decimos entonces que el valor de la función $2x_1 + 3x_1x_2^2 + 1$ en $(2, 3)$ es 58.

Además

$$\begin{aligned}\frac{\partial f}{\partial x_1}(2, 3) &= 2 + 3 \cdot 3^2 = 29, \\ \frac{\partial f}{\partial x_2}(2, 3) &= 6 \cdot 2 \cdot 3 = 36.\end{aligned}$$

- Decimos entonces que el valor de la derivada parcial de la función $2x_1 + 3x_1x_2^2 + 1$ con respecto la variable x_1 en $(2, 3)$ es 29 y la derivada parcial de la función $2x_1 + 3x_1x_2^2 + 1$ con respecto la variable x_2 en $(2, 3)$ es 36.

¿Cuál es la definición equivalente de derivada $f'(x)$ para una función de al menos dos variables $f(x_1, x_2)$?

- Si una función $f(x_1, x_2)$ es derivable, se dice que su derivada es una matriz fila donde en cada posición está una de las derivadas parciales.

$$Df(x_1, x_2) := \begin{bmatrix} \frac{\partial f}{\partial x_1}(x_1, x_2) & \frac{\partial f}{\partial x_2}(x_1, x_2) \end{bmatrix}$$

Para nuestro ejemplo anterior $f(x_1, x_2) = 2x_1 + 3x_1x_2^2 + 1$ tenemos que la derivada es

$$Df(x_1, x_2) := [2 + 3x_2^2 \quad 6x_1x_2]$$

Si la evaluamos en $(2, 3)$, esto es $x_1 = 2$ y $x_2 = 3$ tenemos

$$Df(2, 3) := [29 \quad 36]$$

¿Qué información nos aporta la derivada $Df(x_1, x_2)$? Recordemos que en una variable $f'(x) > 0$ nos decía que la función era creciente cerca de x .

Observemos que $Df(2, 3) := [29 \quad 36]$ es una matriz fila que podemos multiplicar por cualquier matriz columna, esto es,

$$Df(2, 3) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = [29 \quad 36] \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 29u_1 + 36u_2.$$

Esta operación coincide con el llamado *producto escalar de dos vectores*.

- El producto escalar se define como

$$\langle \mathbf{u}, \mathbf{v} \rangle = \left\langle \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right\rangle := u_1v_1 + u_2v_2.$$

A veces también se escribe como

$$\mathbf{u} \cdot \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} := u_1v_1 + u_2v_2.$$

Para nuestro caso

$$Df(2, 3) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = [29 \quad 36] \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \left\langle \begin{bmatrix} 29 \\ 36 \end{bmatrix}, \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right\rangle$$

- Al vector o matriz columna $\begin{bmatrix} 29 \\ 36 \end{bmatrix}$ que se obtiene a partir de la derivada $Df(2, 3)$ se le conoce como **vector gradiente** y se le denota por $\nabla f(2, 3)$. Formalmente,

$$\nabla f(2, 3) = Df(2, 3)^T.$$

- Dada una función diferenciable $f(x_1, x_2)$ llamaremos **vector gradiente de f en (x_1, x_2)** a

$$\nabla f(x_1, x_2) := \begin{bmatrix} \frac{\partial f}{\partial x_1}(x_1, x_2) \\ \frac{\partial f}{\partial x_2}(x_1, x_2) \end{bmatrix}$$

Una propiedad de aproximación importante es la siguiente

- Si $f(x_1, x_2)$ es diferenciable entonces se cumple

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2) - f(x_1, x_2) \approx \Delta x_1 \frac{\partial f}{\partial x_1}(x_1, x_2) + \Delta x_2 \frac{\partial f}{\partial x_2}(x_1, x_2),$$

siempre y cuando $\Delta x_1 \approx 0$ y $\Delta x_2 \approx 0$.

Recordemos una propiedad del producto escalar que dice:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta_{\mathbf{u}, \mathbf{v}} \quad (10)$$

donde $\theta_{\mathbf{u}, \mathbf{v}}$ mide el ángulo en radianes que forman los vectores $\mathbf{u}$ y $\mathbf{v}$. Observemos que si $\theta_{\mathbf{u}, \mathbf{v}} = 90^\circ (\pi/2)$, entonces $\cos \pi/2 = 0$ y en consecuencia $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. Los vectores perpendiculares tienen un producto escalar igual a cero. La propiedad más importante es la siguiente

- **El gradiente de una función diferenciable siempre apunta hacia la dirección de máximo crecimiento de la misma.**

Consideremos una función diferenciable genérica $f(x_1, x_2)$. Entonces se cumple

$$\left\langle \nabla f(x_1, x_2), \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right\rangle = u_1 \frac{\partial f}{\partial x_1}(x_1, x_2) + u_2 \frac{\partial f}{\partial x_2}(x_1, x_2).$$

Además,

$$f(x_1 + u_1, x_2 + u_2) - f(x_1, x_2) \approx u_1 \frac{\partial f}{\partial x_1}(x_1, x_2) + u_2 \frac{\partial f}{\partial x_2}(x_1, x_2) \quad (11)$$

siempre y cuando $u_1 \approx 0$ y $u_2 \approx 0$. Para elegir la dirección de máximo crecimiento, es decir, que la diferencia

$$f(x_1 + u_1, x_2 + u_2) - f(x_1, x_2)$$

en función de u_1 y u_2 sea la máxima posible, nos fijaremos exclusivamente en direcciones $\mathbf{u} := (u_1, u_2)$ de forma que $u_1^2 + u_2^2 = \varepsilon^2$, con $\varepsilon > 0$ fijado sea suficientemente pequeño. Esto implica que $\|\mathbf{u}\| = \|(u_1, u_2)\| = \sqrt{u_1^2 + u_2^2} = \varepsilon$. Entonces empleando (10) obtenemos

$$\begin{aligned} \left\langle \nabla f(x_1, x_2), \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right\rangle &= \|\nabla f(x_1, x_2)\| \|(u_1, u_2)\| \cos \theta_{\nabla f(x_1, x_2), \mathbf{u}} \\ &= \underbrace{\varepsilon \|\nabla f(x_1, x_2)\|}_{\text{no depende de } \mathbf{u}} \cos \theta_{\nabla f(x_1, x_2), \mathbf{u}} \end{aligned}$$

Usando (11) se tiene que

$$f(x_1 + u_1, x_2 + u_2) - f(x_1, x_2) \approx \underbrace{\varepsilon \|\nabla f(x_1, x_2)\|}_{\text{es positivo}} \cos \theta_{\nabla f(x_1, x_2), \mathbf{u}}, \quad (12)$$

y en consecuencia el máximo valor para (11) en función de $\mathbf{u}$, se obtiene para el valor máximo del $\cos \theta_{\nabla f(x_1, x_2), \mathbf{u}}$ como función de $\mathbf{u}$. Basta entonces recordar que

$$-1 \leq \cos \theta_{\nabla f(x_1, x_2), \mathbf{u}} \leq 1.$$

Esto nos dice que el valor máximo en (12) se obtiene para una dirección $\mathbf{u}$ para la que $\theta_{\nabla f(x_1, x_2), \mathbf{u}} = 1$, esto es cuando el ángulo entre $\nabla f(x_1, x_2)$ y $\mathbf{u}$ sea de cero grados. Más explícitamente, $\mathbf{u}$ y $\nabla f(x_1, x_2)$ tienen que ser paralelos y apuntar en la misma dirección.

Extremos locales

- Un valor (x_1, x_2) se dice que es un **mínimo local de la función** $f(x_1, x_2)$ si

$$f(x_1 + u_1, x_2 + u_2) \geq f(x_1, x_2) \quad (13)$$

para cualquier $\mathbf{u} = (u_1, u_2)$ de forma que $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2} \leq \varepsilon$ con $\varepsilon > 0$.

- Un valor (x_1, x_2) se dice que es un **máximo local de la función** $f(x_1, x_2)$ si

$$f(x_1 + u_1, x_2 + u_2) \leq f(x_1, x_2) \quad (14)$$

para cualquier $\mathbf{u} = (u_1, u_2)$ de forma que $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2} \leq \varepsilon$ con $\varepsilon > 0$.

- Un valor (x_1, x_2) se dice que es un **extremo local de la función** $f(x_1, x_2)$ si es o bien un máximo local o bien un mínimo local.

Un primer resultado es el siguiente

- Si (x_1, x_2) es extremo local de la función $f(x_1, x_2)$ entonces cumple

$$\frac{\partial f}{\partial x_1}(x_1, x_2) = 0, \quad \frac{\partial f}{\partial x_2}(x_1, x_2) = 0, \quad (15)$$

o de forma equivalente

$$\nabla f(x_1, x_2) = [0 \quad 0] = \mathbf{0}. \quad (16)$$

Está claro que si el gradiente es distinto de cero entonces la función debe de crecer en la dirección que marca el gradiente. Como consecuencia si un punto es un extremo local, es decir, la función no puede crecer o decrecer más hacia ninguna otra dirección, su gradiente ha de ser obligatoriamente igual al vector cero.

Criterio: Los puntos donde el gradiente se anula son los candidatos para ser los extremos locales (máximos o mínimos) de una función de varias variables. Ahora, bien puede existir algún punto donde el gradiente se anule y no ser un extremo local.

- A los puntos de una función $f(x_1, x_2)$ donde el gradiente se anula se les llama **puntos estacionarios de la función**.

El ejemplo de un punto estacionario que no es un extremo local es el siguiente. Considerar la función

$$f(x_1, x_2) = x_1^2 - x_2^2,$$

donde los posibles puntos estacionarios tienen que cumplir:

$$\begin{aligned}\frac{\partial f}{\partial x_1}(x_1, x_2) &= 2x_1 = 0 \\ \frac{\partial f}{\partial x_2}(x_1, x_2) &= -2x_2 = 0\end{aligned}$$

esto es,

$$\begin{aligned}2x_1 &= 0 \\ -2x_2 &= 0.\end{aligned}$$

El único punto que cumple estas condiciones es $x_1 = 0$ y $x_2 = 0$. Ahora bien si $(0, 0)$ es un máximo para f se ha de cumplir

$$f(u_1, u_2) = u_1^2 - u_2^2 \leq f(0, 0) = 0,$$

que en el caso de un mínimo para f ha de ser

$$f(u_1, u_2) = u_1^2 - u_2^2 \geq f(0, 0) = 0,$$

para valores $\mathbf{u} := (u_1, u_2)$ de forma que $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2} \leq \varepsilon$. Supongamos que $(0, 0)$ fuera un máximo, entonces se ha de cumplir

$$u_1^2 - u_2^2 \leq 0.$$

para valores $\mathbf{u} := (u_1, u_2)$ de forma que $\|\mathbf{u}\| = \sqrt{u_1^2 + u_2^2} \leq \varepsilon$. Basta considerar por ejemplo $(u_1 = 0,001, u_2 = 0)$ para darse cuenta que no se cumple

$$u_1^2 - u_2^2 = (0,001)^2 \leq 0,$$

y en consecuencia $(0, 0)$ no puede ser máximo. Algo parecido ocurre con ser un mínimo. En consecuencia, $(0, 0)$ es un punto estacionario para la función $x_1^2 - x_2^2$ que no es un extremo local de la misma.

Curvas de nivel

Cuando una función tiene un máximo o mínimo local se suele observar comportamiento que podemos ver en la Figuras y . En la Figura unimos mediante una línea los puntos que tienen la imagen igual a un valor fijado. Conforme aumentamos este valor nos acercamos al valor máximo.

Un ejemplo de este tipo de representación son las líneas de presión constante o isobaras que vemos en televisión (ver Figura)

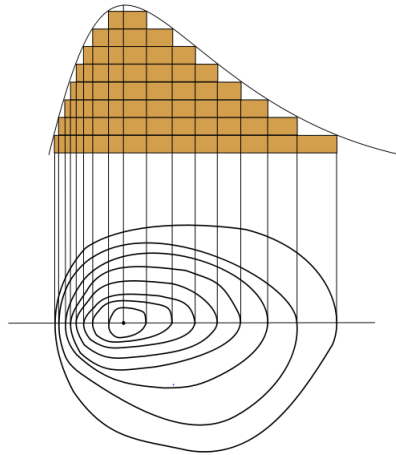


Figura 1: Las curvas de nivel cerca de un máximo local.

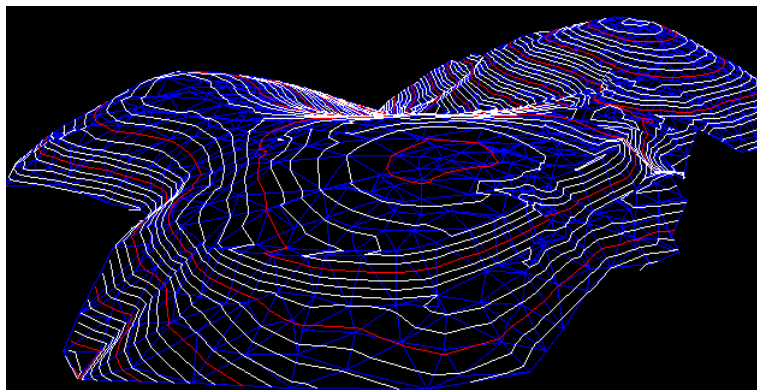


Figura 2: Las curvas de nivel cerca de un máximo local en 3D.

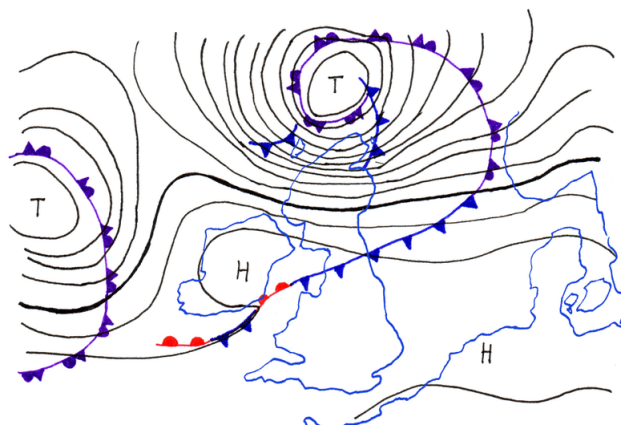


Figura 3: Las isobaras como curvas de nivel.

- Consideremos una función diferenciable $f(x_1, x_2)$ y un número real k fijado, se llama **curva de nivel k de la función f** al conjunto de puntos

$$\{(x_1, x_2) : f(x_1, x_2) = k\} \quad (17)$$

Consideremos la función $f(x_1, x_2) = x_1^2 + x_2^2$. Entonces si fijamos un número $r^2 > 0$ la curvas de nivel para r^2 viene determinada por los puntos del conjunto

$$\{(x_1, x_2) : x_1^2 + x_2^2 = r^2\}.$$

Este conjunto representa una circunferencia en el plano de radio r .

La idea genérica es que las curvas de nivel se pueden parametrizar. Por ejemplo la circunferencia de radio r^2 se describe mediante un parámetro t (podemos considerarlo el tiempo) mediante

$$t \mapsto (r \cos(2\pi t), r \sin(2\pi t)) \quad 0 \leq t \leq 1.$$

Si consideramos

$$\begin{aligned} x_1(t) &= r \cos(2\pi t), \\ x_2(t) &= r \sin(2\pi t). \end{aligned}$$

Entonces

$$x_1^2(t) + x_2^2(t) = r^2 \cos^2(2\pi t) + r^2 \sin^2(2\pi t) = r^2 \left(\underbrace{\cos^2(2\pi t) + \sin^2(2\pi t)}_{=1} \right) = r^2.$$

Como la función

$$f(x_1(t), x_2(t)) = x_1^2(t) + x_2^2(t) = r^2,$$

es constante con respecto a t , derivando con respecto a esta variable obtenemos

$$\frac{d}{dt} f(x_1(t), x_2(t)) = 2x_1(t)x_1'(t) + 2x_2(t)x_2'(t) = 0.$$

Esta última expresión la podemos escribir

$$\begin{aligned} 0 &= 2x_1(t)x_1'(t) + 2x_2(t)x_2'(t) = \left\langle \begin{bmatrix} 2x_1(t) \\ 2x_2(t) \end{bmatrix}, \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} \right\rangle \\ &= \left\langle \begin{bmatrix} \frac{\partial f}{\partial x_1}(x_1(t), x_2(t)) \\ \frac{\partial f}{\partial x_2}(x_1(t), x_2(t)) \end{bmatrix}, \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} \right\rangle \\ &= \left\langle \nabla f(x_1(t), x_2(t)), \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} \right\rangle \end{aligned}$$

Podemos enunciar entonces el resultado siguiente:

- Consideremos que existe una parametrización $t \mapsto (x_1(t), x_2(t))$ de una curva de nivel k para una función diferenciable $f(x_1, x_2)$. Entonces el vector gradiente evaluado sobre la curva $\nabla f(x_1(t), x_2(t))$ es perpendicular al vector tangente a la misma

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix},$$

esto es,

$$\left\langle \nabla f(x_1(t), x_2(t)), \begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} \right\rangle = 0. \quad (18)$$