Matemática aplicada a Economía

Juan Andrés Cabral

Versión 0.2 — última versión del libro (click acá)

Sobre este libro

Este libro reúne notas y resoluciones de ejercicios de clase de Matemática Aplicada I (FCE–UBA). Su objetivo es acompañar la cursada: ordenar ideas, fijar métodos y ofrecer ejemplos resueltos. Puede contener errores u omisiones; toda corrección o sugerencia es bienvenida en `jcabral@udesa.edu.ar`.

Índice

1 Curvas de nivel

Introducción

Las curvas de nivel (o líneas de contorno) de una función de dos variables $z = f(x, y)$ se obtienen fijando un valor constante $z = c$ y resolviendo la ecuación

$$f(x, y) = c$$

Cada valor de c determina una curva en el plano xy que representa los puntos en los que la función adquiere ese valor.

Regla Práctica para Graficar Curvas de Nivel

1. **Fijar el valor de z:** Escoger un número c (por ejemplo, $c = 0$, 1, -1, etc.).
2. **Obtener la ecuación en x e y:** Sustituir $z = c$ en $f(x, y) = c$ y simplificar la ecuación.
3. **Identificar la forma de la curva:** Reconocer si la ecuación representa un círculo, elipse, hipérbola, recta u otra figura.
4. **Determinar puntos clave:** Encontrar intersecciones con los ejes, vértices, asíntotas u otros puntos notables.
5. **Graficar:** Con la información anterior, traza la curva en el plano xy.

Ejemplos de Curvas de Nivel

1. Círculo: $z = x^2 + y^2$

- **Curva de nivel:** Al fijar $z = c$, se tiene

$$x^2 + y^2 = c$$

- **Condición:** $c \geq 0$, radio $r = \sqrt{c}$; si $c < 0$, no hay soluciones reales.

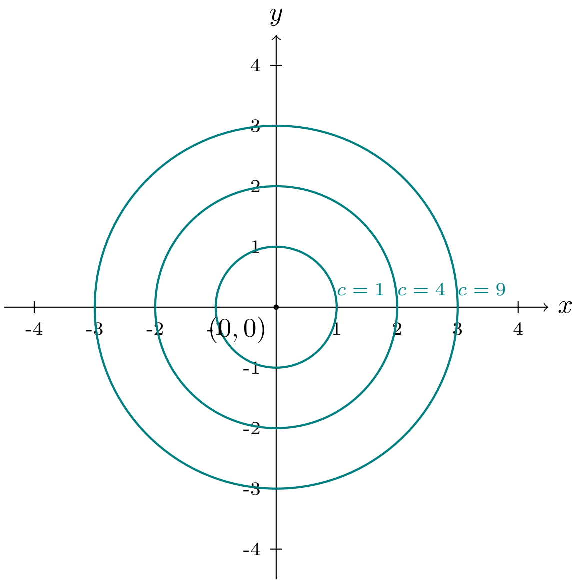


Figure 1: Curvas de nivel de $z = x^2 + y^2$: círculos para distintos valores de c.

2a. Elipse: $z = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$

- **Curva de nivel:**

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = c$$

- **Semiejes:** $a\sqrt{c}$ en x, $b\sqrt{c}$ en y; si $c = 0$, la curva se reduce al punto $(0, 0)$

2b. Elipse: $z = \dfrac{x^2}{a} + \dfrac{y^2}{b}$

- **Curva de nivel:**

$$\frac{x^2}{a} + \frac{y^2}{b} = c$$

- **Forma alternativa:**

$$\frac{x^2}{ac} + \frac{y^2}{bc} = 1$$

- **Semiejes:** $\sqrt{ac}$ en x, $\sqrt{bc}$ en y; para $c = 0$, punto $(0, 0)$

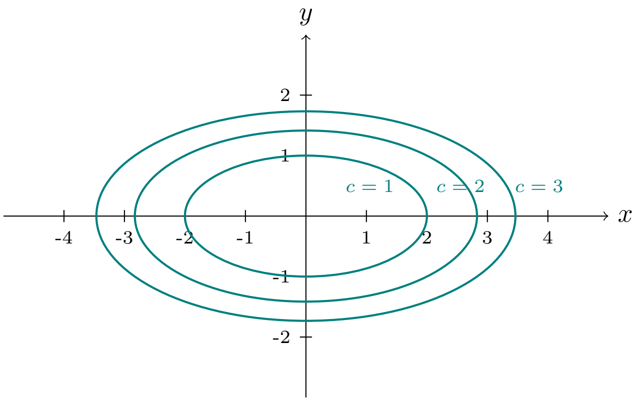


Figure 2: Curvas de nivel de $z = \frac{x^2}{4} + \frac{y^2}{1}$

3. Hipérbola: $z = axy$

- **Curva de nivel:**

$$y = \frac{c}{ax}$$

- **Casos:** Para $c = 0$, la curva es $x = 0 \cup y = 0$, es decir, los ejes coordenados.

4. Plano: $ax + by = cz$

- **Forma explícita:**

$$z = \frac{a}{c}x + \frac{b}{c}y$$

- **Curvas de nivel:**

$$ax + by = ck$$

que representan rectas paralelas para distintos valores de k.

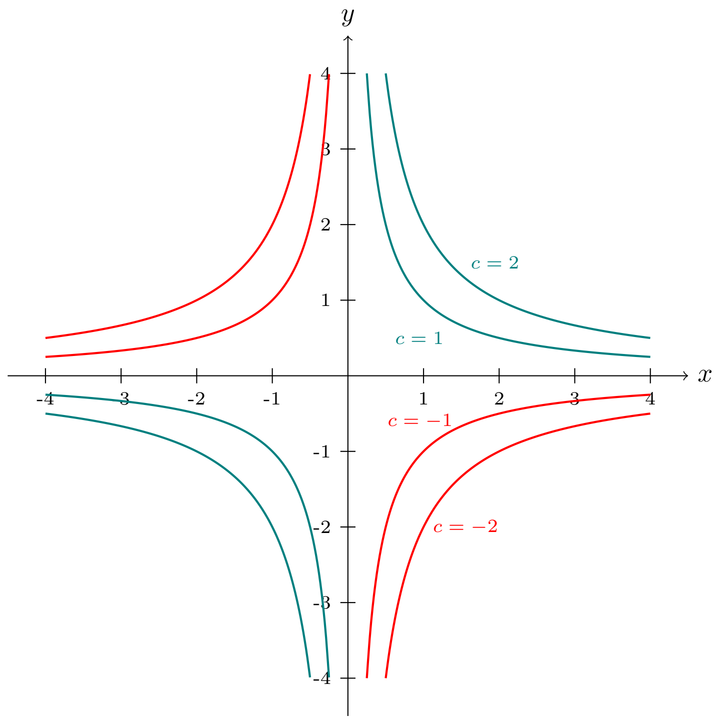


Figure 3: Curvas de nivel de $z = axy$: hipérbolas $y = \frac{c}{ax}$

5. Hipérbola: $z = x^2 - y^2$

- **Curva de nivel:**

$$x^2 - y^2 = c$$

- **Casos:**

  - $\mathbf{c > 0}$:

$$\frac{x^2}{c} - \frac{y^2}{c} = 1$$

    vértices en $x = \pm\sqrt{c}$

  - $\mathbf{c = 0}$:

$$y = x \quad \text{y} \quad y = -x$$

  - $\mathbf{c < 0}$:

$$\frac{y^2}{|c|} - \frac{x^2}{|c|} = 1$$

    vértices en $y = \pm\sqrt{|c|}$

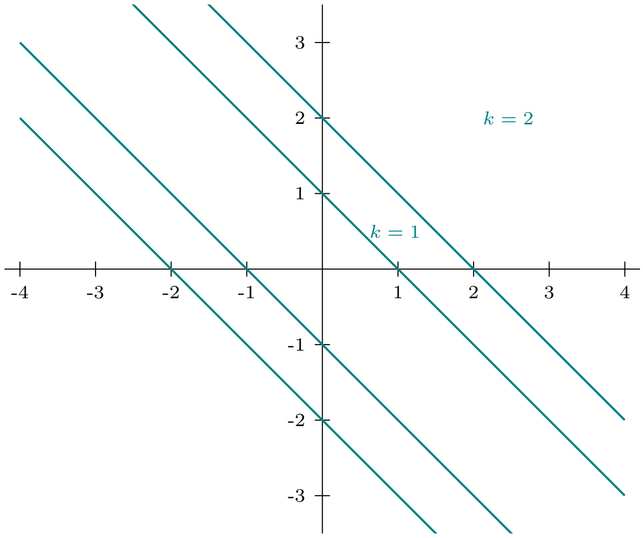


Figure 4: Curvas de nivel del plano $ax + by = cz$: rectas paralelas $ax + by = ck$ para distintos valores de k,

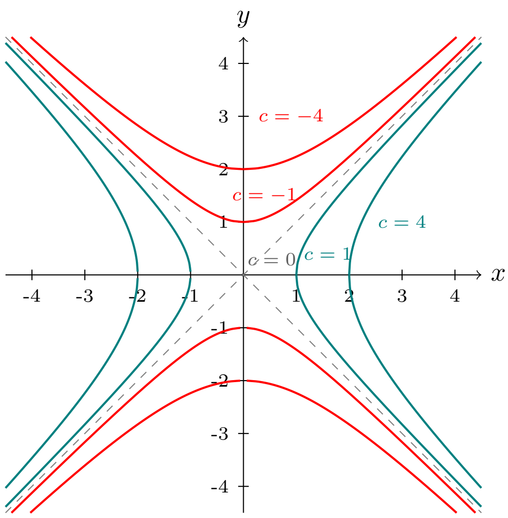


Figure 5: Curvas de nivel de $z = x^2 - y^2$: para $c > 0$ abren sobre el eje x; para $c < 0$ (rojo) abren sobre el eje y. El caso $c = 0$ son las rectas $y = \pm x$ (grises, punteadas).

2 Curvas de nivel en Economía

En esta sección se analiza la importancia de las curvas de nivel en el contexto económico, enfocándonos en su uso para representar y analizar funciones de producción y utilidad.

- Curvas de Nivel: Representan conjuntos de puntos en los cuales una función de dos variables mantiene un mismo valor. En economía, se utilizan para identificar combinaciones de insumos o bienes que generan un mismo nivel de producción o satisfacción.
- Isocuantas: En funciones de producción, las curvas de nivel se denominan isocuantas, que muestran las combinaciones de factores productivos que producen una cantidad constante de output.
- Curvas de Indiferencia: En el análisis de la utilidad, las curvas de nivel se conocen como curvas de indiferencia y representan combinaciones de bienes que proporcionan el mismo nivel de satisfacción al consumidor.

Propiedades Deseables de las Curvas de Nivel en Economía

Las curvas de nivel derivadas de funciones de utilidad deben exhibir ciertas propiedades para reflejar comportamientos económicos razonables. A continuación, se describen las características y la intuición económica asociada a cada una:

1. No se Cortan

- Intuición Económica: Cada curva de nivel representa un nivel único de utilidad. Si dos curvas se cruzaran, implicaría que una misma combinación de bienes podría proporcionar dos niveles diferentes de satisfacción, lo cual es inconsistente en el análisis de preferencias.

2. Pendiente Negativa

- Intuición Económica: Para mantener constante la utilidad, un aumento en la cantidad de un bien debe ir acompañado de una disminución en la cantidad del otro.

3. Monotonicidad

- Intuición Económica: La monotonicidad implica que al aumentar la cantidad de cualquiera de los bienes (mientras se mantiene la otra constante), la utilidad no disminuye. En la gráfica, esto se traduce en curvas que crecen hacia arriba y a la derecha, reflejando la preferencia por más cantidad de bienes.

4. Continuidad

- Intuición Económica: La continuidad asegura que pequeñas variaciones en las cantidades de bienes resultan en pequeños cambios en la utilidad. Esto permite predecir el comportamiento del consumidor sin saltos abruptos en sus niveles de satisfacción.

5. Convexidad al Origen

- Intuición Económica: A medida que se dispone de más de un bien, el consumidor está dispuesto a sacrificar cada vez menos del otro para obtener una unidad adicional del primero. Es una manifestación de la preferencia por combinaciones balanceadas de bienes, evitando consumir cantidades extremas de uno solo.

Recta Presupuestaria y Métodos de Graficación

La recta presupuestaria representa las combinaciones de dos bienes que un consumidor puede adquirir gastando todo su ingreso, dados los precios de dichos bienes.

Método 1: Despejar una Variable en Función de la Otra

Ecuación general de la restricción presupuestaria:

$$p_1q_1 + p_2q_2 = R$$

Despejando q_2 en función de q_1:

$$q_2 = \frac{R}{p_2} - \frac{p_1}{p_2}q_1$$

- Intuición: La ecuación muestra que, a medida que aumenta la cantidad de q_1, la cantidad de q_2 debe disminuir para mantener constante el gasto total R. La pendiente $-\frac{p_1}{p_2}$ refleja la tasa a la cual se puede intercambiar bienes.
- Aplicación: Este método permite graficar la recta presupuestaria como una función lineal, facilitando el análisis algebraico de las decisiones del consumidor.

Método 2: Utilizar los Interceptos

- Intercepto en el eje q_1:
$$q_1 = \frac{R}{p_1}$$
- Intercepto en el eje q_2:
$$q_2 = \frac{R}{p_2}$$
- Intuición: Estos interceptos representan las cantidades máximas de cada bien si el consumidor gasta todo su ingreso en uno solo.
- Aplicación: Conociendo ambos puntos, se traza la recta que los une. Esta gráfica ilustra de manera clara las opciones de combinación de bienes disponibles.

Curva de Costo para el Productor

La curva de isocosto representa las combinaciones de insumos que se pueden adquirir para un costo total fijo:

$$w_1x_1 + w_2x_2 = C$$

donde w_1 y w_2 son los precios de los insumos, x_1 y x_2 sus cantidades.

Isoingreso

La línea de isoingreso representa las combinaciones de productos o bienes que generan un ingreso total constante:

$$p_1q_1 + p_2q_2 = I$$

Curvas de Indiferencia Cobb-Douglas

Las curvas de indiferencia derivadas de una función de utilidad Cobb-Douglas representan combinaciones de bienes que otorgan un nivel constante de utilidad. Una forma general de esta función es:

$$U(x_1, x_2) = x_1^{\alpha}x_2^{\beta}$$

donde $\alpha, \beta > 0$.

Métodos para Graficar Curvas de Indiferencia

- Despejar una Variable:

$$x_2 = \left(\frac{U_0}{x_1^\alpha}\right)^{\frac{1}{\beta}}$$

  o, equivalentemente,

$$x_2 = \frac{U^{1/\beta}}{x_1^{\alpha/\beta}}$$

  Esto permite graficar la curva de indiferencia para un nivel dado de utilidad U, considerando distintos valores de x_1.

- Observaciones:
  - Las curvas son convexas al origen, lo que refleja una tasa marginal de sustitución decreciente.
  - Nunca tocan los ejes, ya que si $x_1 = 0$ o $x_2 = 0$, entonces la utilidad total es cero.
  - Las curvas se ubican en el primer cuadrante, ya que se supone que el consumidor sólo considera cantidades positivas de bienes.

Cada curva representa un nivel fijo de utilidad, y al aumentar dicho nivel, las curvas se desplazan hacia arriba y a la derecha, reflejando preferencias monótonas.

Nota: Las funciones de producción también suelen tener forma Cobb-Douglas, en cuyo caso las curvas de nivel se denominan isocuantas y reflejan combinaciones de factores productivos que permiten obtener una misma cantidad de producto.

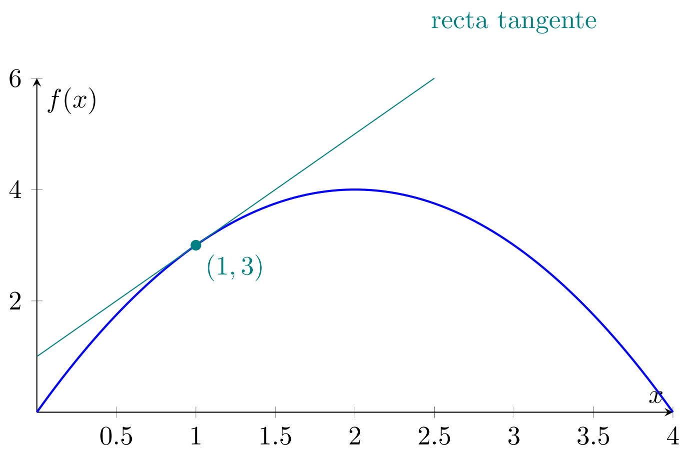
3 Derivadas parciales

Derivada para funciones de una variable

Para funciones de una variable la derivada relaciona el incremento de la función con el incremento de la variable independiente (cociente incremental) cuando el incremento de x tiende a cero

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

Gráficamente:



La función representada es:

$$f(x) = -(x-2)^2 + 4$$

En el punto $(1, 3)$, la pendiente de la recta tangente es precisamente la derivada de f respecto a x:

$$\left.\frac{df}{dx}\right|_{x=1} = 2$$

Como esta pendiente es positiva, la función en ese punto está aumentando.

Derivadas parciales

Para una función de dos variables debemos considerar el incremento de la función z asociada a cambios en las variables independientes x e y. Para ello, permitimos la variación de una variable manteniendo la otra constante.

Dada la función $z = f(x, y)$, se define la derivada parcial de f con respecto a x en el punto (x_0, y_0) como el valor del siguiente límite, si existe y es finito:

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h,\, y_0) - f(x_0, y_0)}{h}$$

Del mismo modo, se define la derivada parcial de f con respecto a y en el punto (x_0, y_0) como el siguiente límite, si existe y es finito:

$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0,\, y_0 + h) - f(x_0, y_0)}{h}$$

Intuición: Básicamente consideramos un pequeño incremento h (que luego hacemos tender a cero) en una sola variable —x o y— mientras la otra permanece fija. A ese desplazamiento le restamos el valor original de la función para medir cuánto ha cambiado f al pasar del punto viejo al nuevo, y dividimos por la distancia h recorrida. De este modo obtenemos la variación de f por unidad de desplazamiento en esa dirección, que es precisamente la derivada parcial en ese punto.

Derivadas parciales en $\mathbb{R}^n$

Para generalizar lo anterior podemos pasar a n variables: Sea $f\colon U \subseteq \mathbb{R}^n \to \mathbb{R}$ y sea $a = (a_1, \ldots, a_n) \in U$. Para cada $i = 1, \ldots, n$, la derivada parcial de f con respecto a la variable x_i en el punto a se define como

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h\to 0} \frac{f(a_1, \ldots, a_{i-1},\, a_i + h,\, a_{i+1}, \ldots, a_n) \;-\; f(a_1, \ldots, a_n)}{h}$$

Ejemplo

Sea la función

$$f(x, y) = x^2 y + 3\,x\,y^2$$

Por definición

$$\frac{\partial f}{\partial x}(1, 2) = \lim_{h\to 0} \frac{f(1+h, 2) - f(1, 2)}{h}$$

Calculamos

$$f(1+h, 2) = 2(1+h)^2 + 12(1+h) = 2 + 4h + 2h^2 + 12 + 12h = 14 + 16h + 2h^2 \quad f(1, 2) = 14$$

Luego

$$\frac{f(1+h, 2) - f(1, 2)}{h} = \frac{16h + 2h^2}{h} = 16 + 2h$$

y al tomar el límite

$$\left.\frac{\partial f}{\partial x}\right|_{(x,y)=(1,2)} = 16$$

También es posible aplicar reglas de derivadas como en el caso de funciones de una sola variable. Para esto las mismas reglas son aplicables teniendo en cuenta que las demás variables se mantienen constantes cuando derivamos con respecto a una:

Reglas de derivadas más utilizadas

- **Regla de la potencia:**

$$\frac{d}{dx}(x^n) = n\,x^{n-1}$$

- **Derivada de la función logarítmica:**

$$\frac{d}{dx}(\ln x) = \frac{1}{x} \quad (x > 0)$$

- **Derivadas de las funciones trigonométricas:**

$$\frac{d}{dx}(\sin x) = \cos x, \quad \frac{d}{dx}(\cos x) = -\sin x$$

- **Derivadas de las funciones exponenciales:**

$$\frac{d}{dx}(a^x) = a^x \ln a \quad (a > 0,\ a \neq 1)$$

- **Regla del producto:**

$$\frac{d}{dx}(u(x)\,v(x)) = \frac{du}{dx}\,v(x) + u(x)\,\frac{dv}{dx}$$

- **Regla del cociente:**

$$\frac{d}{dx}\Big(\frac{u(x)}{v(x)}\Big) = \frac{\frac{du}{dx}\,v(x) - u(x)\,\frac{dv}{dx}}{v(x)^2}$$

Siguiendo con el ejemplo anterior:

$$f(x,y) = x^2y + 3\,x\,y^2$$

Por regla

$$\frac{\partial f}{\partial x}(x,y) = 2x\,y + 3\,y^2 \quad \Longrightarrow \quad \frac{\partial f}{\partial x}(1,2) = 2\cdot 1\cdot 2 + 3\cdot 2^2 = 16$$

Otro ejemplo

Sea la función

$$f(x,y) = x^3y^2 + 5\,x\,y^3 - x^2y$$

Por definición

$$\frac{\partial f}{\partial x}(x,y) = \lim_{h\to 0}\frac{f(x+h,y) - f(x,y)}{h}$$

$$f(x+h,y) = (x+h)^3y^2 + 5(x+h)y^3 - (x+h)^2y = x^3y^2 + 3x^2y^2h + 3xy^2h^2 + y^2h^3 + 5xy^3 + 5y^3h - x^2y - 2xyh - yh^2$$

$$\frac{f(x+h,y) - f(x,y)}{h} = 3x^2y^2 + 3xy^2h + y^2h^2 + 5y^3 - 2xy - yh \xrightarrow{h\to 0} 3x^2y^2 + 5y^3 - 2xy$$

$$\frac{\partial f}{\partial y}(x,y) = \lim_{h\to 0}\frac{f(x,y+h) - f(x,y)}{h}$$

$$f(x,y+h) = x^3(y+h)^2 + 5x(y+h)^3 - x^2(y+h) = x^3y^2 + 2x^3yh + x^3h^2 + 5xy^3 + 15xy^2h + 15xyh^2 + 5xh^3 - x^2y - x^2h$$

$$\frac{f(x,y+h) - f(x,y)}{h} = 2x^3y + x^3h + 15xy^2 + 15xyh + 5xh^2 - x^2 \xrightarrow{h\to 0} 2x^3y + 15xy^2 - x^2$$

Por reglas de derivación

$$\frac{\partial f}{\partial x}(x,y) = \frac{\partial}{\partial x}(x^3y^2) + \frac{\partial}{\partial x}(5xy^3) - \frac{\partial}{\partial x}(x^2y) = 3x^2y^2 + 5y^3 - 2x\,y$$

$$\frac{\partial f}{\partial x}(x,y) = 3x^2y^2 + 5y^3 - 2xy$$

$$\frac{\partial f}{\partial y}(x,y) = \frac{\partial}{\partial y}(x^3y^2) + \frac{\partial}{\partial y}(5xy^3) - \frac{\partial}{\partial y}(x^2y) = 2x^3y + 15x\,y^2 - x^2$$

$$\frac{\partial f}{\partial y}(x,y) = 2x^3y + 15xy^2 - x^2$$

Las reglas de derivación (regla de potencia, producto, cociente, etc.) se obtienen a partir de la definición límite de derivada. Sin embargo, cuando la función está definida por tramos, esas reglas pueden no ser aplicables en los puntos de unión. En dichos casos es necesario calcular la derivada parcial directamente mediante la definición por límites

Ejemplo con función partida

Sea la función

$$f(x,y) = \begin{cases} \dfrac{x^3}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$$

Por definición

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{\frac{h^3}{h^2} - 0}{h} = \lim_{h\to 0} \frac{h}{h} = 1$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = \lim_{h\to 0} \frac{\frac{0}{h^2} - 0}{h} = \lim_{h\to 0} \frac{0}{h} = 0$$

Derivadas sucesivas

En una sola variable ya conocemos las derivadas sucesivas:

$$f'(x), \quad f''(x), \quad f^{(n)}(x)$$

que se obtienen aplicando la definición de derivada reiteradamente sobre la función resultante.

En varias variables podemos también derivar más de una vez, ya sea respecto a la misma variable o a variables distintas. Así surgen las derivadas parciales de orden dos y superiores, por ejemplo:

$$\frac{\partial^2 f}{\partial x^2} = f_{xx}, \quad \frac{\partial^2 f}{\partial y \partial x} = f_{xy}, \quad \frac{\partial^2 f}{\partial x \partial y} = f_{yx}, \quad \frac{\partial^2 f}{\partial y^2} = f_{yy}$$

Las derivadas cruzadas f_{xy} y f_{yx} se obtienen derivando primero con respecto a x y luego a y, o viceversa. En general, las derivadas sucesivas se calculan tomando la derivada parcial de la derivada parcial considerada como nueva función, es decir, si $g(x,y) = \frac{\partial f}{\partial x}(x,y)$, entonces

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{\partial}{\partial y}(g(x,y))$$

y así sucesivamente para órdenes mayores.

Ejemplo

$f(x, y) = e^x + e^y + y\sin(x)$

Derivadas parciales de primer orden

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial}{\partial x}(e^x + e^y + y\sin(x)) = e^x + y\cos(x) \quad \Longrightarrow \quad f_x(x, y) = e^x + y\cos(x)$$

$$\frac{\partial f}{\partial y}(x, y) = \frac{\partial}{\partial y}(e^x + e^y + y\sin(x)) = e^y + \sin(x) \quad \Longrightarrow \quad f_y(x, y) = e^y + \sin(x)$$

Derivadas parciales de segundo orden

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}(e^x + y\cos(x)) = e^x - y\sin(x) \quad \Longrightarrow \quad f_{xx}(x, y) = e^x - y\sin(x)$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}(e^x + y\cos(x)) = \cos(x) \quad \Longrightarrow \quad f_{xy}(x, y) = \cos(x)$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}(e^y + \sin(x)) = \cos(x) \quad \Longrightarrow \quad f_{yx}(x, y) = \cos(x)$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}(e^y + \sin(x)) = e^y \quad \Longrightarrow \quad f_{yy}(x, y) = e^y$$

Teorema de Schwarz

El teorema de Schwarz garantiza que, cuando las segundas derivadas parciales cruzadas de una función existen y son continuas, éstas coinciden:

$$f_{x_i x_j} = f_{x_j x_i}$$

Esto resulta especialmente práctico porque, tras comprobar esta igualdad, basta calcular una sola de las dos derivadas cruzadas en lugar de ambas, ahorrando tiempo.

Condiciones de aplicabilidad

Supóngase que en un entorno de un punto (x_0, y_0) se cumplen:

- Existen las derivadas parciales $\dfrac{\partial f}{\partial x}$ y $\dfrac{\partial f}{\partial y}$.
- Existe una de las derivadas parciales cruzadas (por ejemplo $\dfrac{\partial^2 f}{\partial y\, \partial x}$) y es continua en (x_0, y_0).

Bajo estas hipótesis, las derivadas parciales cruzadas van a ser iguales.

Ejemplo

Sea la función

$$f(x, y) = x^2 y + e^{xy}$$

En todo $\mathbb{R}^2$ se cumplen las condiciones:

- Existen las derivadas parciales $\dfrac{\partial f}{\partial x}$ y $\dfrac{\partial f}{\partial y}$.
- Existe la derivada parcial cruzada $\dfrac{\partial^2 f}{\partial y\, \partial x}$ y es continua en $\mathbb{R}^2$.

Por tanto, en cualquier punto (x, y) vale el teorema de Schwarz.

Cálculo de las derivadas parciales

$$f_x(x,y) = \frac{\partial}{\partial x}(x^2y + e^{xy}) = 2x\,y + y\,e^{xy} \qquad f_y(x,y) = \frac{\partial}{\partial y}(x^2y + e^{xy}) = x^2 + x\,e^{xy}$$

Derivadas cruzadas

$$f_{xy}(x,y) = \frac{\partial}{\partial y}(2x\,y + y\,e^{xy}) = 2x + e^{xy} + x\,y\,e^{xy}$$

$$f_{yx}(x,y) = \frac{\partial}{\partial x}(x^2 + x\,e^{xy}) = 2x + e^{xy} + x\,y\,e^{xy}$$

$$f_{xy}(x,y) = f_{yx}(x,y)$$

4 Diferencial

Recordatorio: cálculo univariable

Sea $f : \mathbb{R} \to \mathbb{R}$ diferenciable en x_0. La *aproximación lineal* (o *linealización*) alrededor de x_0 es

$$f(x_0 + h) \;\approx\; f(x_0) + f'(x_0)\,h \quad \text{con } h = x - x_0$$

El *diferencial* de f en x_0 se define como

$$\mathrm{d}f \;=\; f'(x_0)\,\mathrm{d}x$$

Intuición: Al "hacer zoom" sobre la gráfica de f alrededor de x_0, la curva se vuelve casi recta: la *recta tangente* con pendiente $f'(x_0)$ es la mejor aproximación local de f. Dar un pasito $\mathrm{d}x$ en x provoca un cambio real

$$\Delta f \;=\; f(x_0 + \mathrm{d}x) - f(x_0)$$

que es *aproximadamente igual* al diferencial:

$$\Delta f \;\approx\; \mathrm{d}f \;=\; f'(x_0)\,\mathrm{d}x$$

es decir, *pendiente* $\times$ *paso*

También se puede ver como

$$f(x_0 + h) = f(x_0) + f'(x_0)h + o(h) \quad \text{cuando } h \to 0$$

donde $o(h)$ denota un término tal que $\lim\limits_{h\to 0} \dfrac{o(h)}{h} = 0$; es decir, el error entre la aproximación lineal y el valor real es despreciable frente a $|h|$ cuando $h \to 0$

Definición en dos variables independientes

Sea $f : \mathbb{R}^2 \to \mathbb{R}$ diferenciable en (a, b). El *diferencial* de f en (a, b) es

$$\mathrm{d}f \;=\; f_x(a,b)\,\mathrm{d}x \;+\; f_y(a,b)\,\mathrm{d}y$$

Equivalente:

$$f(a + \mathrm{d}x,\, b + \mathrm{d}y) \;\approx\; f(a,b) \;+\; f_x(a,b)\,\mathrm{d}x \;+\; f_y(a,b)\,\mathrm{d}y$$

Geometría: en dos variables, la aproximación *no es una recta* sino un *plano tangente* al gráfico $z = f(x, y)$ en (a, b):

$$z \;\approx\; f(a,b) \;+\; f_x(a,b)\,(x-a) \;+\; f_y(a,b)\,(y-b)$$

Intuición: Lo que buscamos es aproximar el cambio de z a partir de cambios pequeños en x e y. Cada derivada parcial $f_x(a, b)$ y $f_y(a, b)$ mide cuán sensible es z a x o a y por separado. Si nos movemos $\mathrm{d}x$ en x y $\mathrm{d}y$ en y, el cambio total se aproxima sumando ambos aportes:

$$\Delta z \;\approx\; \mathrm{d}f \;=\; f_x(a,b)\,\mathrm{d}x \;+\; f_y(a,b)\,\mathrm{d}y$$

es decir, *"pendiente en x"* $\times$ *"paso en x"* + *"pendiente en y"* $\times$ *"paso en y"*

Condiciones suficientes

Si f_x y f_y existen en un entorno de (a, b) y son continuas en (a, b), entonces f es diferenciable en (a, b) y la aproximación lineal anterior es válida.

Intuición: Exigir que f_x y f_y sean continuas cerca de (a, b) garantiza que las pendientes en x y en y no cambien bruscamente. Si las pendientes varían de forma suave, entonces la superficie de f cerca de (a, b) se parece mucho a su plano tangente, y la aproximación lineal con el diferencial describe bien los cambios pequeños en z

Caso general $\mathbb{R}^n \to \mathbb{R}$

Para $f : \mathbb{R}^n \to \mathbb{R}$ diferenciable en $\mathbf{a} \in \mathbb{R}^n$,

$$\mathrm{d}f \;=\; \frac{\partial f}{\partial x_1}(\mathbf{a})\,\mathrm{d}x_1 \;+\; \frac{\partial f}{\partial x_2}(\mathbf{a})\,\mathrm{d}x_2 \;+\; \cdots \;+\; \frac{\partial f}{\partial x_n}(\mathbf{a})\,\mathrm{d}x_n$$

Es decir:

$$\mathrm{d}f \;=\; \nabla f(\mathbf{a}) \cdot \mathrm{d}\mathbf{x} \;=\; \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{a})\,\mathrm{d}x_i$$

y de forma aproximada:

$$f(a_1 + \mathrm{d}x_1, \ldots, a_n + \mathrm{d}x_n) \;\approx\; f(a_1, \ldots, a_n) \;+\; \mathrm{d}f$$

Ejemplos

Sea $f(x, y) = x^2 y + e^{xy}$. Entonces

$$f_x = 2xy + y\,e^{xy} \qquad f_y = x^2 + x\,e^{xy}$$

En $(a, b) = (1, 1)$,

$$\mathrm{d}f \;=\; (2 \cdot 1 \cdot 1 + 1 \cdot e)\,\mathrm{d}x \;+\; (1^2 + 1 \cdot e)\,\mathrm{d}y \;=\; (2 + e)\,\mathrm{d}x + (1 + e)\,\mathrm{d}y$$

Para $\mathrm{d}x = 0.01$ y $\mathrm{d}y = -0.02$,

$$\Delta f \approx \mathrm{d}f = (2 + e)(0.01) + (1 + e)(-0.02) \approx -0.0272$$

El resultado negativo indica que, con un incremento pequeño en x y una disminución en y, el valor de la función f tiende a reducirse levemente. En otras palabras, el efecto de disminuir y en un 0.02 pesa más que el incremento positivo que produce aumentar x en 0.01, de modo que el cambio neto es una pequeña disminución en f.

5 Aplicación económica de diferencial

Tasa marginal de sustitución (TMS)

Dada una función de utilidad que depende de las cantidades de dos bienes q_1 y q_2, llamamos *tasa marginal de sustitución (TMS)* a la tasa de compensación entre ellos. Es decir, es la cantidad que se estaría dispuesto a ceder de uno de los bienes para obtener una unidad adicional del otro, manteniendo constante el nivel de utilidad.

Calculamos el diferencial de la función de utilidad $u(q_1, q_2)$:

$$du = \frac{\partial u}{\partial q_1} dq_1 + \frac{\partial u}{\partial q_2} dq_2$$

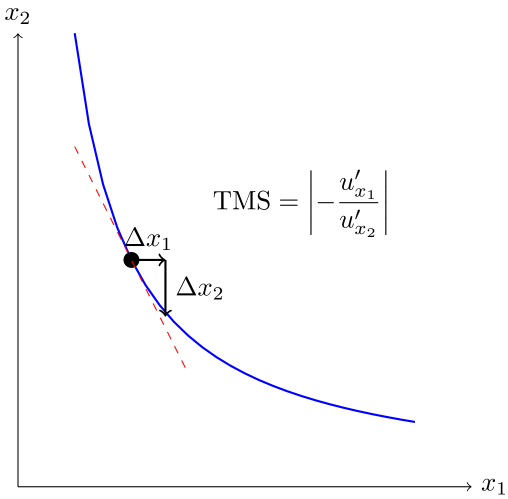
Como $du = 0$ sobre la curva de indiferencia:

$$\frac{\partial u}{\partial q_1} dq_1 + \frac{\partial u}{\partial q_2} dq_2 = 0 \Rightarrow \frac{dq_2}{dq_1} = -\frac{\frac{\partial u}{\partial q_1}}{\frac{\partial u}{\partial q_2}}$$

Esta pendiente o el módulo de la misma representa lo que se suele llamar tasa marginal de sustitución de bienes:

$$\text{TMS} = \left| \frac{dq_2}{dq_1} \right| = \frac{\frac{\partial u}{\partial q_1}}{\frac{\partial u}{\partial q_2}}$$

Graficamente



Ejemplo

Dada la función de utilidad $u(q_1, q_2) = 5q_1q_2$, calcular la TMS en el punto $(q_1 = 5, q_2 = 2)$.

$$\frac{\partial u}{\partial q_1} = 5q_2 = 5 \cdot 2 = 10$$
$$\frac{\partial u}{\partial q_2} = 5q_1 = 5 \cdot 5 = 25$$
$$\text{TMS} = \frac{10}{25} = 0.4$$

Interpretación: Para aumentar una unidad de q_1, se deben ceder 0.4 unidades de q_2 para mantener constante el nivel de utilidad

Si se invierte la tasa y se calcula $\frac{dq_1}{dq_2}$, obtenemos:

$$\frac{dq_1}{dq_2} = \frac{\frac{\partial u}{\partial q_2}}{\frac{\partial u}{\partial q_1}} = \frac{25}{10} = 2.5$$

Interpretación: Para aumentar una unidad de q_2, se deben ceder 2.5 unidades de q_1 para mantener constante el nivel de utilidad

Tasa de sustitución técnica

Dada una función de producción que depende de dos insumos x_1 y x_2, llamamos *tasa marginal de sustitución técnica (TST)* a la tasa a la que una empresa puede sustituir un factor productivo por otro, manteniendo constante el nivel de producción.

Calculamos el diferencial de la función de producción $q(x_1, x_2)$:

$$dq = \frac{\partial q}{\partial x_1} dx_1 + \frac{\partial q}{\partial x_2} dx_2$$

Como $dq = 0$ sobre la isocuanta:

$$\frac{\partial q}{\partial x_1} dx_1 + \frac{\partial q}{\partial x_2} dx_2 = 0 \Rightarrow \frac{dx_2}{dx_1} = -\frac{\frac{\partial q}{\partial x_1}}{\frac{\partial q}{\partial x_2}}$$

El módulo de esta pendiente representa la tasa marginal de sustitución técnica:

$$\text{TST} = \left| \frac{dx_2}{dx_1} \right| = \frac{\frac{\partial q}{\partial x_1}}{\frac{\partial q}{\partial x_2}}$$

Ejemplo

Dada la función de producción:

$$q(a, b) = 20 - 7a + 8b - a^2 + b^2$$

calcular la TST en el punto $a = 1.2, b = 2.2$.

$$\begin{aligned} \frac{\partial q}{\partial a} &= -7 - 2a = -7 - 2(1.2) = -9.4 \\ \frac{\partial q}{\partial b} &= 8 + 2b = 8 + 2(2.2) = 12.4 \\ \text{TST} &= \frac{-9.4}{12.4} \approx -0.758 \end{aligned}$$

Interpretación: Para aumentar una unidad de a, se deben ceder aproximadamente 0.758 unidades de b para mantener constante el nivel de producción

Si se invierte la tasa y se calcula $\frac{dx_1}{dx_2}$, obtenemos:

$$\frac{dx_1}{dx_2} = \frac{\frac{\partial q}{\partial b}}{\frac{\partial q}{\partial a}} = \frac{12.4}{-9.4} \approx -1.319$$

Interpretación: Para aumentar una unidad de b, se deben ceder aproximadamente 1.319 unidades de a para mantener constante el nivel de producción

Notación

En muchos textos se utiliza una notación abreviada del tipo $\text{TMS}(x_1/x_2)$ para referirse a las tasas marginales de sustitución.

$\text{TMS}(x_1/x_2)$ representa la tasa marginal de sustitución donde el numerador es la derivada respecto de x_2 y el denominador la derivada respecto de x_1, es decir:

$$\text{TMS}(x_1/x_2) = \frac{dx_2}{dx_1} = \frac{\frac{\partial u}{\partial x_1}}{\frac{\partial u}{\partial x_2}}$$

Interpretación: indica cuántas unidades de x_2 deben cederse para obtener una unidad adicional de x_1, manteniendo constante el nivel de utilidad

De manera inversa, $\text{TMS}(x_2/x_1)$ representa la tasa marginal de sustitución donde el numerador es la derivada respecto de x_1 y el denominador la derivada respecto de x_2, es decir:

$$\text{TMS}(x_2/x_1) = \frac{dx_1}{dx_2} = \frac{\frac{\partial u}{\partial x_2}}{\frac{\partial u}{\partial x_1}}$$

Interpretación: indica cuántas unidades de x_1 deben cederse para obtener una unidad adicional de x_2, manteniendo constante el nivel de utilidad

Interpretación intuitiva

Cuando aparece una expresión como

$$\frac{dx_2}{dx_1} = \frac{\frac{\partial u}{\partial x_1}}{\frac{\partial u}{\partial x_2}}$$

se interpreta como: "cuánto debe reducirse x_2 para aumentar en una unidad x_1, manteniendo constante la utilidad".

Una forma intuitiva de pensarlo es:

- El numerador es lo que se entrega o se cede
- El denominador es lo que se desea obtener

6 Clasificación de bienes y elasticidad

Conceptos clave

- **Funciones de demanda:** Representan la relación entre la cantidad demandada de un bien y variables como su precio propio, el precio de bienes relacionados y el ingreso del consumidor
- **Derivadas parciales:** Permiten analizar el efecto marginal de un cambio en una variable (precio o ingreso) sobre la cantidad demandada, manteniendo constantes las demás
- **Elasticidad:** Es la medida de sensibilidad de la cantidad demandada ante cambios porcentuales en el precio o ingreso, ayudando a clasificar el bien según su comportamiento

Clasificación de bienes

1. Según el precio propio

- **Bienes típicos:** Al aumentar el precio del bien, la cantidad demandada disminuye. La mayoría de bienes se comporta así:

$$\frac{\partial Q}{\partial P} < 0$$

- **Bienes Giffen:** En situaciones excepcionales, el aumento en el precio del bien provoca un aumento en su demanda, debido al efecto ingreso que supera al efecto sustitución. No confundir con los bienes Veblen, que son bienes de lujo demandados por sus características de señalizar estatus económico

$$\frac{\partial Q}{\partial P} > 0$$

Intuición: En condiciones normales, la relación inversa (mayor precio, menor demanda) es esperada; los bienes Giffen representan casos raros donde, por limitaciones presupuestarias, el aumento de precio lleva a consumir más del bien

2. Según el precio del otro bien

- **Bienes sustitutos:** Si al aumentar el precio de un bien la demanda del otro aumenta, se considera que los bienes son sustitutos. Esto ocurre cuando el consumidor opta por el bien relativamente más económico. Por ejemplo, marcas rivales: Coca Cola y Pepsi

$$\frac{\partial Q_1}{\partial P_2} > 0$$

- **Bienes complementarios:** Si el aumento en el precio de un bien reduce la demanda del otro, se clasifican como complementarios, ya que se consumen en conjunto. Por ejemplo, café y azúcar

$$\frac{\partial Q_1}{\partial P_2} < 0$$

Intuición: La existencia de bienes sustitutos permite al consumidor cambiar de producto al encarecerse uno, mientras que los bienes complementarios se consumen conjuntamente para satisfacer una necesidad

3. Según el ingreso

- **Bienes normales:** La demanda aumenta al incrementar el ingreso del consumidor, la mayoría de bienes tienden a ser normales
$$\frac{\partial Q}{\partial I} > 0$$

- **Bienes inferiores:** La demanda disminuye cuando aumenta el ingreso, ya que los consumidores optan por bienes de mayor calidad. Por ejemplo, marcas de segunda línea
$$\frac{\partial Q}{\partial I} < 0$$

Intuición: La clasificación según el ingreso refleja la capacidad del consumidor para ajustar sus hábitos de consumo en función de su poder adquisitivo

Elasticidad de la demanda

La elasticidad cuantifica la sensibilidad de la cantidad demandada ante cambios en precios o ingresos

Elasticidad precio e ingreso

$$\text{Elasticidad precio} = \frac{\Delta Q/Q}{\Delta P/P} = \frac{P}{Q}\frac{\partial Q}{\partial P}$$

- **Interpretación (en términos absolutos):**
  - Si
  $$\left|\frac{P}{Q}\frac{\partial Q}{\partial P}\right| > 1$$
  la demanda es **elástica** (sensible a cambios en el precio)
  - Si
  $$\left|\frac{P}{Q}\frac{\partial Q}{\partial P}\right| < 1$$
  la demanda es **inelástica** (poco sensible)
  - Si
  $$\left|\frac{P}{Q}\frac{\partial Q}{\partial P}\right| = 1$$
  la elasticidad es **unitaria** (la demanda se mueve exactamente en la misma proporción que el precio)
  - Si
  $$\frac{\partial Q}{\partial P} = 0 \quad \text{(lo que implica que } \left|\frac{P}{Q}\frac{\partial Q}{\partial P}\right| = 0\text{)}$$
  la demanda es **perfectamente inelástica** (la cantidad demandada permanece constante sin importar cambios en el precio)

- Aunque se presenta el concepto como **elasticidad precio**, esta noción se extiende a la **elasticidad ingreso** y a cualquier otra función que admita derivadas parciales. Por ejemplo, la elasticidad ingreso se define como:
$$\text{Elasticidad ingreso} = \frac{\Delta Q/Q}{\Delta I/I} = \frac{I}{Q}\frac{\partial Q}{\partial I}$$
donde I representa el ingreso

Bienes de lujo vs. bienes esenciales

La clasificación de bienes en términos de lujo o esencial se basa en la **elasticidad ingreso**, la cual mide la sensibilidad de la cantidad demandada ante cambios en el ingreso del consumidor. Es importante resaltar que esta clasificación se aplica únicamente a bienes **normales**

Bienes esenciales (necesidades básicas)

Para los bienes esenciales, la elasticidad ingreso se encuentra en el rango:

$$0 < E_I < 1$$

donde

$$E_I = \frac{I}{Q}\frac{\partial Q}{\partial I}$$

Esto significa que, ante un aumento porcentual en el ingreso, la cantidad demandada del bien aumenta en una proporción menor. Dichos bienes satisfacen necesidades básicas y su consumo tiende a estabilizarse incluso cuando el ingreso crece

Bienes de lujo

En cambio, los bienes de lujo se caracterizan por tener:

$$E_I > 1$$

Esto implica que un incremento porcentual en el ingreso genera un aumento porcentual mayor en la demanda del bien. Estos bienes son adquiridos en mayor medida cuando los consumidores disponen de mayores recursos, reflejando un comportamiento de consumo más discrecional

Ejemplos:

- **Bienes esenciales:** Alimentos básicos (por ejemplo, pan y arroz), medicamentos genéricos, vivienda
- **Bienes de lujo:** Ropa de diseñador, automóviles de alta gama, joyería fina

Ejemplo

Supongamos que la función de demanda de un bien se expresa como:

$$Q(P, I) = 200 + 0.3I - 4P$$

donde P es el precio del bien e I es el ingreso del consumidor

Cálculo de las derivadas parciales:

$$\frac{\partial Q}{\partial P} = -4, \quad \frac{\partial Q}{\partial I} = 0.3$$

El signo negativo en $\frac{\partial Q}{\partial P}$ indica que, al aumentar el precio, la cantidad demandada disminuye (bien típico). El signo positivo en $\frac{\partial Q}{\partial I}$ muestra que, al aumentar el ingreso, la demanda crece, lo que caracteriza al bien como normal

Evaluación en un punto específico:

Sea $P = 15$ e $I = 500$. Entonces:

$$Q(15, 500) = 200 + 0.3 \times 500 - 4 \times 15 = 200 + 150 - 60 = 290$$

Cálculo de la elasticidad precio:

La elasticidad precio se define como:

$$E_P = \frac{P}{Q}\frac{\partial Q}{\partial P}$$

Sustituyendo los valores:

$$E_P = \frac{15}{290} \times (-4) \approx -0.207$$

En valor absoluto, $|E_P| \approx 0.207$, lo que indica que la demanda es **inelástica** (poco sensible a cambios en el precio)

Cálculo de la elasticidad ingreso:

$$E_I = \frac{I}{Q}\frac{\partial Q}{\partial I}$$

Sustituyendo los valores:

$$E_I = \frac{500}{290} \times 0.3 \approx 0.517$$

Esto significa que, ante un aumento del 1% en el ingreso, la cantidad demandada incrementa aproximadamente en un 0.517%, también es inelástica respecto al ingreso

Interpretación respecto a bienes de lujo vs. esenciales:

Dado que $0 < E_I < 1$, el bien se clasifica como un **bien esencial** (o necesidad básica), ya que la respuesta de la demanda a cambios en el ingreso es relativamente baja

7 Derivadas implícitas

Derivación implícita: deducción de la fórmula y requerimientos

En muchos casos en matemáticas y economía, las relaciones entre variables se presentan de forma implícita, es decir, mediante una ecuación de la forma

$$F(x, y) = 0$$

donde y se define como función de x de forma implícita (i.e., $y = f(x)$). Para poder aplicar la derivación implícita, es fundamental cumplir con ciertos requerimientos y condiciones.

Requerimientos para la aplicación de la derivación implícita

1. Verificación de la solución: se asume que el punto de interés (x_0, y_0) satisface la ecuación, es decir, $F(x_0, y_0) = 0$ (esto implica que $y_0 = f(x_0)$)

2. Diferenciabilidad de $F(x, y)$: la función F debe ser diferenciable en un entorno del punto de interés (x_0, y_0). Esto garantiza que existen las derivadas parciales F_x y F_y

3. No nulidad de la derivada parcial respecto a y: se requiere que $F_y(x_0, y_0) \neq 0$. Esta condición es esencial para aplicar el Teorema de la Función Implícita, el cual asegura que en un entorno de (x_0, y_0) se puede expresar y como una función diferenciable de x (i.e., $y = f(x)$)

Deducción de la fórmula de la derivada implícita

Consideremos la ecuación

$$F(x, y) = 0$$

donde suponemos que $y = f(x)$ y F es diferenciable en un entorno de (x_0, y_0). Como $F(x, f(x)) = 0$ para todos los x en dicho entorno, derivamos ambos lados de la igualdad con respecto a x, aplicando la regla de la cadena:

$$\frac{d}{dx}F(x, f(x)) = F_x(x, f(x)) + F_y(x, f(x)) \cdot f'(x) = 0$$

Aquí, F_x y F_y representan las derivadas parciales de F con respecto a x y y, respectivamente. Despejamos $f'(x)$ de la siguiente forma:

$$F_y(x, f(x)) \cdot f'(x) = -F_x(x, f(x))$$

$$\implies \quad f'(x) = -\frac{F_x(x, f(x))}{F_y(x, f(x))}$$

Esta es la fórmula general para la derivada de una función definida de forma implícita.

Ejemplo

Consideremos la ecuación de la circunferencia:

$$x^2 + y^2 = 25$$

Definimos la función $F(x, y)$ como:

$$F(x, y) = x^2 + y^2 - 25 = 0$$

Observamos que para cualquier punto (x, y) sobre la circunferencia se cumple $F(x, y) = 0$

Paso 1: calcular las derivadas parciales Se tiene:

$$F_x(x, y) = 2x \quad \text{y} \quad F_y(x, y) = 2y$$

Paso 2: aplicar la fórmula de la derivada implícita Utilizando la fórmula:

$$\frac{dy}{dx} = -\frac{F_x(x,y)}{F_y(x,y)} = -\frac{2x}{2y}$$

$$\implies \quad \frac{dy}{dx} = -\frac{x}{y}$$

Derivación implícita: función de dos variables independientes

En algunos problemas en matemáticas y economía, las relaciones entre variables se presentan de forma implícita mediante una ecuación de la forma

$$F(x,y,z) = 0$$

donde z se define implícitamente como función de x e y (i.e., $z = f(x,y)$). Para aplicar la derivación implícita en este contexto, es fundamental cumplir con ciertos requerimientos y condiciones.

Requerimientos para la aplicación de la derivación implícita

1. Verificación de la solución: se asume que el punto de interés $Q_0 = (x_0, y_0, z_0)$ satisface la ecuación, es decir, $F(x_0, y_0, z_0) = 0$ (lo que implica que $z_0 = f(x_0, y_0)$)
2. Diferenciabilidad de $F(x,y,z)$: la función F y sus derivadas parciales F_x, F_y y F_z deben existir y ser continuas en un entorno del punto Q_0
3. No nulidad de la derivada parcial respecto a z: se requiere que $F_z(x_0, y_0, z_0) \neq 0$. Esta condición es esencial para poder expresar z como una función diferenciable de x e y en el entorno de Q_0

Deducción de las fórmulas de las derivadas parciales

Consideremos la ecuación

$$F(x,y,z) = 0$$

donde suponemos que $z = f(x,y)$ y F es diferenciable en un entorno del punto $Q_0 = (x_0, y_0, z_0)$. Como $F(x, y, f(x,y)) = 0$ para todos los (x,y) en dicho entorno, derivamos ambos lados de la igualdad con respecto a x e y, aplicando la regla de la cadena.

Aplicación de la regla de la cadena

Para clarificar la deducción, se puede reescribir la función $F(x,y,z)$ definiendo $u = x$, $v = y$ y $z = f(x,y)$, de forma que

$$w = F(u,v,z) = 0$$

Aplicando la regla de la cadena respecto a x:

$$\frac{\partial w}{\partial x} = \frac{\partial F}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial F}{\partial v} \cdot \frac{\partial v}{\partial x} + \frac{\partial F}{\partial z} \cdot \frac{\partial z}{\partial x} = 0$$

Como $\frac{\partial u}{\partial x} = 1$ y $\frac{\partial v}{\partial x} = 0$, se obtiene:

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial z} \cdot \frac{\partial z}{\partial x} = 0 \quad \implies \quad \frac{\partial z}{\partial x} = -\frac{F_x}{F_z}$$

El mismo procedimiento se aplica para obtener $\frac{\partial z}{\partial y}$

Estas fórmulas permiten calcular las derivadas parciales de $z = f(x,y)$ de forma implícita, siempre y cuando se cumplan los requerimientos indicados.

Ejemplo

Consideremos la función:

$$F(x,y,z) = x^2y + \sin(z) - z\cos(y) + z^2 - 1 = 0$$

Observamos que para cualquier punto (x,y,z) que satisface esta ecuación se define implícitamente z como función de x e y (i.e., $z = f(x,y)$)

Paso 1: calcular las derivadas parciales Se tiene:

$$F_x(x,y,z) = 2xy$$

$$F_y(x,y,z) = x^2 + z\sin(y)$$

$$F_z(x,y,z) = \cos(z) - \cos(y) + 2z$$

Paso 2: aplicar la fórmula de la derivada implícita Las fórmulas para las derivadas parciales de $z = f(x,y)$ son:

$$\frac{\partial z}{\partial x} = -\frac{F_x(x,y,z)}{F_z(x,y,z)} \quad \text{y} \quad \frac{\partial z}{\partial y} = -\frac{F_y(x,y,z)}{F_z(x,y,z)}$$

Paso 3: evaluar en un punto Elijamos el punto $Q_0 = (1,1,0)$. Verificamos que:

$$1^2 \cdot 1 + \sin(0) - 0 \cdot \cos(1) + 0^2 - 1 = 1 - 1 = 0$$

por lo que Q_0 pertenece a la superficie definida por $F(x,y,z) = 0$

Evaluamos las derivadas parciales en Q_0:

$$F_x(1,1,0) = 2 \cdot 1 \cdot 1 = 2$$

$$F_y(1,1,0) = 1^2 + 0 \cdot \sin(1) = 1$$

$$F_z(1,1,0) = \cos(0) - \cos(1) + 2 \cdot 0 = 1 - \cos(1) \approx 0.46$$

Por lo tanto:

$$\left.\frac{\partial z}{\partial x}\right|_{(1,1,0)} = -\frac{2}{1-\cos(1)} \approx -4.35$$

$$\left.\frac{\partial z}{\partial y}\right|_{(1,1,0)} = -\frac{1}{1-\cos(1)} \approx -2.17$$

Derivadas implícitas sucesivas

Consideremos la función implícita:

$$F(x,y,z) = 2\sin(z) - xz + y^3 - 1 = 0$$

que define a $z = z(x,y)$. Se desea calcular las derivadas de segundo orden:

$$z_{xx}, \quad z_{xy}, \quad z_{yx}, \quad z_{yy}$$

evaluadas en el punto $Q_0 = (1,1,0)$

Paso 1: derivadas de primer orden

Dado que $F(x, y, z(x, y)) = 0$, diferenciamos implícitamente respecto de x:

$$F_x + F_z \cdot z_x = 0 \quad \Rightarrow \quad z_x = -\frac{F_x}{F_z}$$

Análogamente, derivando respecto de y:

$$F_y + F_z \cdot z_y = 0 \quad \Rightarrow \quad z_y = -\frac{F_y}{F_z}$$

Calculamos ahora las derivadas parciales necesarias:

$$F_x = \frac{\partial F}{\partial x} = -z, \quad F_y = \frac{\partial F}{\partial y} = 3y^2, \quad F_z = \frac{\partial F}{\partial z} = 2\cos(z) - x$$

Entonces:

$$z_x = \frac{z}{2\cos(z) - x}, \quad z_y = -\frac{3y^2}{2\cos(z) - x}$$

A partir de estas expresiones, aplicamos la regla del cociente para obtener las segundas derivadas sin evaluar numéricamente hasta el final.

Cálculo de z_{xx}

Derivamos $z_x = \frac{z}{2\cos(z)-x}$ respecto a x:

$$z_{xx} = \frac{d}{dx}\left(\frac{z}{2\cos(z) - x}\right) = \frac{\frac{dz}{dx}\,(2\cos(z) - x) - z\,\frac{d(2\cos(z)-x)}{dx}}{(2\cos(z) - x)^2}$$

$$\frac{d(2\cos(z) - x)}{dx} = -2\sin(z)\,\frac{dz}{dx} - 1 = -2\sin(z)\,z_x - 1$$

Sustituyendo:

$$z_{xx} = \frac{(z_x)(2\cos(z) - x) - z(-2\sin(z)z_x - 1)}{(2\cos(z) - x)^2}$$

$$z_{xx} = \frac{z_x(2\cos(z) - x) + z(2\sin(z)z_x + 1)}{(2\cos(z) - x)^2}$$

Evaluando en $Q_0 = (1, 1, 0)$ (con $z = 0, z_x = 0, \sin(0) = 0, \cos(0) = 1, x = 1$):

$$z_{xx}(1, 1) = \frac{(0)(2(1) - 1) + 0(2(0)(0) + 1)}{(2(1) - 1)^2} = \frac{0 + 0}{1^2} = 0$$

Cálculo de z_{xy}

Derivamos $z_x = \frac{z}{2\cos(z)-x}$ respecto a y:

$$z_{xy} = \frac{d}{dy}\left(\frac{z}{2\cos(z) - x}\right) = \frac{\frac{dz}{dy}(2\cos(z) - x) - z\frac{d(2\cos(z)-x)}{dy}}{(2\cos(z) - x)^2}$$

Calculamos las derivadas necesarias para el numerador:

$$\frac{d(2\cos(z) - x)}{dy} = -2\sin(z)\frac{dz}{dy} - 0 = -2\sin(z)z_y$$

Sustituyendo:

$$z_{xy} = \frac{(z_y)(2\cos(z) - x) - z(-2\sin(z)z_y)}{(2\cos(z) - x)^2}$$

$$z_{xy} = \frac{z_y(2\cos(z) - x) + 2z\sin(z)z_y}{(2\cos(z) - x)^2}$$

Evaluando en $Q_0 = (1, 1, 0)$ (con $z = 0, z_y = -3, \sin(0) = 0, \cos(0) = 1, x = 1$):

$$z_{xy}(1, 1) = \frac{(-3)(2(1) - 1) + 2(0)(0)(-3)}{(2(1) - 1)^2} = \frac{(-3)(1) + 0}{1^2} = -3$$

Por simetría de las derivadas cruzadas (asumiendo que se cumple el teorema de Schwarz):

$$z_{yx} = z_{xy}$$

Cálculo de z_{yy}

Derivamos $z_y = -\frac{3y^2}{2\cos(z)-x}$ respecto a y:

$$z_{yy} = \frac{d}{dy}\left(-\frac{3y^2}{2\cos(z) - x}\right) = -\frac{\frac{d(3y^2)}{dy}(2\cos(z) - x) - (3y^2)\frac{d(2\cos(z)-x)}{dy}}{(2\cos(z) - x)^2}$$

Calculamos las derivadas necesarias para el numerador:

$$\frac{d(3y^2)}{dy} = 6y$$

$$\frac{d(2\cos(z) - x)}{dy} = -2\sin(z)z_y \quad \text{(calculado antes)}$$

Sustituyendo:

$$z_{yy} = -\frac{(6y)(2\cos(z) - x) - (3y^2)(-2\sin(z)z_y)}{(2\cos(z) - x)^2}$$

$$z_{yy} = -\frac{6y(2\cos(z) - x) + 6y^2\sin(z)z_y}{(2\cos(z) - x)^2}$$

Evaluando en $Q_0 = (1, 1, 0)$ (con $y = 1, z = 0, z_y = -3, \sin(0) = 0, \cos(0) = 1, x = 1$):

$$z_{yy}(1, 1) = -\frac{6(1)(2(1) - 1) + 6(1)^2(0)(-3)}{(2(1) - 1)^2} = -\frac{6(1) + 0}{1^2} = -6$$

8 Derivadas implícitas en sistemas de ecuaciones

Derivación Implícita en Sistemas de Ecuaciones con una variable independiente

Así como una superficie puede estar definida en forma implícita por una ecuación, una curva en el espacio puede estar definida implícitamente por un sistema de dos ecuaciones

$$\begin{cases} F(x, y, z) = 0 \\ G(x, y, z) = 0 \end{cases}$$

Por ejemplo, puede ocurrir que el sistema defina implícitamente a $y = f(x)$ y $z = g(x)$ es decir, que ambas dependan de una única variable independiente. En ese caso, se dice que el sistema define a dos de las tres variables como funciones de la restante

Requerimientos para Aplicar la Derivación Implícita

Para que sea posible derivar y y z respecto a x se deben cumplir las siguientes condiciones

1. **Verificación de la solución:** El punto de interés (x_0, y_0, z_0) debe satisfacer ambas ecuaciones es decir

$$F(x_0, y_0, z_0) = 0 \quad \text{y} \quad G(x_0, y_0, z_0) = 0$$

2. **Diferenciabilidad de F y G:** Ambas funciones deben ser diferenciables en un entorno del punto, con derivadas parciales continuas

3. **No anulación del Jacobiano respecto a y y z:** El determinante

$$J = \frac{\partial(F, G)}{\partial(y, z)} = \begin{vmatrix} F_y & F_z \\ G_y & G_z \end{vmatrix} \neq 0$$

debe ser distinto de cero para que el sistema tenga solución única para $\frac{dy}{dx}$ y $\frac{dz}{dx}$

Deducción de las Fórmulas de Derivadas

Partimos de las ecuaciones implícitas

$$F(x, y, z) = 0, \quad G(x, y, z) = 0$$

Por la definición de la derivada y la regla de la cadena, al derivar la igualdad respecto a x obtenemos

$$\frac{d}{dx}F(x,\, y(x),\, z(x)) = \frac{\partial F}{\partial x}(x, y, z) + \frac{\partial F}{\partial y}(x, y, z)\,\frac{dy}{dx} + \frac{\partial F}{\partial z}(x, y, z)\,\frac{dz}{dx} = 0$$

Notamos que en este proceso se entiende que la derivada de x respecto a sí mismo es 1 Así, se llega directamente a

$$F_x(x, y, z) + F_y(x, y, z)\,\frac{dy}{dx} + F_z(x, y, z)\,\frac{dz}{dx} = 0$$

De manera similar, si además se tiene otra función $G(x, y, z) = 0$ (con y y z también funciones de x) aplicamos la regla de la cadena para obtener

$$G_x(x, y, z) + G_y(x, y, z)\,\frac{dy}{dx} + G_z(x, y, z)\,\frac{dz}{dx} = 0$$

Por lo tanto tenemos el siguiente sistema

$$\begin{cases} F_x + F_y\,\frac{dy}{dx} + F_z\,\frac{dz}{dx} = 0 \\ G_x + G_y\,\frac{dy}{dx} + G_z\,\frac{dz}{dx} = 0 \end{cases}$$

$$\begin{cases} F_y \frac{dy}{dx} + F_z \frac{dz}{dx} = -F_x \\ G_y \frac{dy}{dx} + G_z \frac{dz}{dx} = -G_x \end{cases}$$

Escribimos el sistema en forma matricial

$$\begin{pmatrix} F_y & F_z \\ G_y & G_z \end{pmatrix} \begin{pmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{pmatrix} = -\begin{pmatrix} F_x \\ G_x \end{pmatrix}$$

Resolución con Regla de Cramer

Este sistema lineal en las incógnitas $\frac{dy}{dx}$ y $\frac{dz}{dx}$ se resuelve usando determinantes

$$\frac{dy}{dx} = \frac{\begin{vmatrix} -F_x & F_z \\ -G_x & G_z \end{vmatrix}}{J} \quad \text{y} \quad \frac{dz}{dx} = \frac{\begin{vmatrix} F_y & -F_x \\ G_y & -G_x \end{vmatrix}}{J}$$

También puede escribirse de forma compacta como cocientes de derivadas parciales mixtas

$$\frac{dy}{dx} = -\frac{\partial(F,G)/\partial(x,z)}{\partial(F,G)/\partial(y,z)} \quad \text{y} \quad \frac{dz}{dx} = -\frac{\partial(F,G)/\partial(y,x)}{\partial(F,G)/\partial(y,z)}$$

Ejemplo

Consideremos el sistema de ecuaciones

$$\begin{cases} xy + z = 2 \\ x - y + z^2 = 0 \end{cases}$$

que define implícitamente a y y z como funciones de x (i.e., $y = y(x)$ y $z = z(x)$). Para aplicar la derivación implícita, definimos

$$F(x,y,z) = xy + z - 2 = 0 \qquad G(x,y,z) = x - y + z^2 = 0$$

Paso 1: Calcular las derivadas parciales Para $F(x,y,z)$

$$F_x = y, \quad F_y = x, \quad F_z = 1$$

Para $G(x,y,z)$

$$G_x = 1, \quad G_y = -1, \quad G_z = 2z$$

Paso 2: Diferenciar implícitamente y escribir el sistema Aplicando la regla de la cadena, tenemos

$$F_x + F_y \frac{dy}{dx} + F_z \frac{dz}{dx} = 0 \quad \Longrightarrow \quad y + x\frac{dy}{dx} + \frac{dz}{dx} = 0$$

$$G_x + G_y \frac{dy}{dx} + G_z \frac{dz}{dx} = 0 \quad \Longrightarrow \quad 1 - \frac{dy}{dx} + 2z\frac{dz}{dx} = 0$$

Reorganizando, obtenemos

$$\begin{cases} x\frac{dy}{dx} + \frac{dz}{dx} = -y \\ -\frac{dy}{dx} + 2z\frac{dz}{dx} = -1 \end{cases}$$

Paso 3: Escribir el sistema en forma matricial El sistema se puede expresar de la siguiente manera

$$\begin{pmatrix} F_y & F_z \\ G_y & G_z \end{pmatrix} \begin{pmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{pmatrix} = -\begin{pmatrix} F_x \\ G_x \end{pmatrix}$$

es decir,

$$\begin{pmatrix} x & 1 \\ -1 & 2z \end{pmatrix} \begin{pmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{pmatrix} = -\begin{pmatrix} y \\ 1 \end{pmatrix}$$

Paso 4: Resolver el sistema con la Regla de Cramer El determinante del sistema (Jacobiano) es

$$J = \begin{vmatrix} x & 1 \\ -1 & 2z \end{vmatrix} = 2xz + 1$$

Luego, aplicando la regla de Cramer, se tiene

Para $\dfrac{dy}{dx}$

$$\frac{dy}{dx} = \frac{\begin{vmatrix} -y & 1 \\ -1 & 2z \end{vmatrix}}{J} = \frac{(-y)(2z) - (1)(-1)}{2xz+1} = \frac{-2yz+1}{2xz+1}$$

Para $\dfrac{dz}{dx}$

$$\frac{dz}{dx} = \frac{\begin{vmatrix} x & -y \\ -1 & -1 \end{vmatrix}}{J} = \frac{x(-1) - (-y)(-1)}{2xz+1} = \frac{-x-y}{2xz+1}$$

Derivación Implícita en Sistemas de Ecuaciones (2 Variables Independientes)

Planteamos la situación en la que dos ecuaciones

$$F(x, y, u, v) = 0 \quad \text{y} \quad G(x, y, u, v) = 0$$

definen implícitamente dos funciones de dos variables independientes

$$u = h(x, y) \quad \text{y} \quad v = m(x, y)$$

Deducción de las Fórmulas de las Derivadas Parciales

Partimos de las ecuaciones implícitas

$$F(x, y, u, v) = 0 \quad \text{y} \quad G(x, y, u, v) = 0$$

Para obtener las derivadas parciales de u y v respecto de x (manteniendo y constante) procedemos de la siguiente manera

(1) Derivadas parciales respecto a x:

Consideramos $u = u(x, y)$ y $v = v(x, y)$. Entonces, aplicando la regla de la cadena a $F(x, y, u, v) = 0$ se tiene

$$\frac{d}{dx}F(x, y, u(x, y), v(x, y)) = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial x} = 0$$

Aquí se entiende que, al derivar con respecto a x, la variable y se considera constante

De manera similar, al derivar $G(x, y, u, v) = 0$ respecto a x se obtiene

$$\frac{d}{dx}G(x, y, u(x, y), v(x, y)) = \frac{\partial G}{\partial x} + \frac{\partial G}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial x} = 0$$

Podemos escribir el sistema resultante como

$$\begin{cases} F_x(x, y, u, v) + F_u(x, y, u, v)\,\frac{\partial u}{\partial x} + F_v(x, y, u, v)\,\frac{\partial v}{\partial x} = 0 \\ G_x(x, y, u, v) + G_u(x, y, u, v)\,\frac{\partial u}{\partial x} + G_v(x, y, u, v)\,\frac{\partial v}{\partial x} = 0 \end{cases}$$

Reorganizando para aislar las derivadas parciales de u y v se tiene

$$\begin{cases} F_u\,\frac{\partial u}{\partial x} + F_v\,\frac{\partial v}{\partial x} = -F_x \\ G_u\,\frac{\partial u}{\partial x} + G_v\,\frac{\partial v}{\partial x} = -G_x \end{cases}$$

Este sistema se puede escribir en forma matricial

$$\begin{pmatrix} F_u & F_v \\ G_u & G_v \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial v}{\partial x} \end{pmatrix} = - \begin{pmatrix} F_x \\ G_x \end{pmatrix}$$

(2) Derivadas parciales respecto a y:

Ahora, derivamos con respecto a y (manteniendo x constante). Aplicando la regla de la cadena a $F(x, y, u, v) = 0$ se obtiene

$$\frac{d}{dy} F(x, y, u(x, y), v(x, y)) = \frac{\partial F}{\partial y} + \frac{\partial F}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial y} = 0$$

De forma análoga, para G se tiene

$$\frac{d}{dy} G(x, y, u(x, y), v(x, y)) = \frac{\partial G}{\partial y} + \frac{\partial G}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial G}{\partial v}\frac{\partial v}{\partial y} = 0$$

El sistema resultante es

$$\begin{cases} F_y(x, y, u, v) + F_u(x, y, u, v)\, \frac{\partial u}{\partial y} + F_v(x, y, u, v)\, \frac{\partial v}{\partial y} = 0 \\ G_y(x, y, u, v) + G_u(x, y, u, v)\, \frac{\partial u}{\partial y} + G_v(x, y, u, v)\, \frac{\partial v}{\partial y} = 0 \end{cases}$$

Reorganizando se tiene

$$\begin{cases} F_u\, \frac{\partial u}{\partial y} + F_v\, \frac{\partial v}{\partial y} = -F_y \\ G_u\, \frac{\partial u}{\partial y} + G_v\, \frac{\partial v}{\partial y} = -G_y \end{cases}$$

En forma matricial

$$\begin{pmatrix} F_u & F_v \\ G_u & G_v \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial y} \end{pmatrix} = - \begin{pmatrix} F_y \\ G_y \end{pmatrix}$$

Resolución con Regla de Cramer:

Sea

$$J = \begin{vmatrix} F_u & F_v \\ G_u & G_v \end{vmatrix}$$

Entonces, aplicando la regla de Cramer para el sistema respecto a x obtenemos

$$\frac{\partial u}{\partial x} = \frac{\begin{vmatrix} -F_x & F_v \\ -G_x & G_v \end{vmatrix}}{J} \qquad \frac{\partial v}{\partial x} = \frac{\begin{vmatrix} F_u & -F_x \\ G_u & -G_x \end{vmatrix}}{J}$$

Y para el sistema respecto a y

$$\frac{\partial u}{\partial y} = \frac{\begin{vmatrix} -F_y & F_v \\ -G_y & G_v \end{vmatrix}}{J} \qquad \frac{\partial v}{\partial y} = \frac{\begin{vmatrix} F_u & -F_y \\ G_u & -G_y \end{vmatrix}}{J}$$

De esta forma, hemos obtenido las cuatro derivadas parciales que describen cómo varían u y v respecto de x e y

$$\frac{\partial u}{\partial x}, \quad \frac{\partial v}{\partial x}, \quad \frac{\partial u}{\partial y}, \quad \frac{\partial v}{\partial y}$$

Ejemplo

Consideremos el siguiente sistema de ecuaciones

$$\begin{cases} F(x, y, u, v) = x\, u + y\, v - 2 = 0 \\ G(x, y, u, v) = u^2 - v + x\, y = 0 \end{cases}$$

el cual define implícitamente a u y v como funciones de x e y es decir, $u = u(x, y)$ y $v = v(x, y)$

Paso 1: Cálculo de las Derivadas Parciales de F y G

Para la función

$$F(x, y, u, v) = x\,u + y\,v - 2$$

se tiene

$$F_x = u, \quad F_y = v, \quad F_u = x, \quad F_v = y$$

Para la función

$$G(x, y, u, v) = u^2 - v + x\,y$$

se obtiene

$$G_x = y, \quad G_y = x, \quad G_u = 2u, \quad G_v = -1$$

Paso 2: Derivadas Parciales respecto a x

Dado que $u = u(x, y)$ y $v = v(x, y)$ aplicamos la regla de la cadena a cada ecuación (manteniendo y constante)

Para $F(x, y, u, v) = 0$

$$F_x + F_u\,u_x + F_v\,v_x = 0 \quad \Longrightarrow \quad u + x\,u_x + y\,v_x = 0$$

Para $G(x, y, u, v) = 0$

$$G_x + G_u\,u_x + G_v\,v_x = 0 \quad \Longrightarrow \quad y + 2u\,u_x - v_x = 0$$

Escribimos el sistema en forma matricial

$$\begin{pmatrix} F_u & F_v \\ G_u & G_v \end{pmatrix} \begin{pmatrix} u_x \\ \\ v_x \end{pmatrix} = - \begin{pmatrix} F_x \\ \\ G_x \end{pmatrix}$$

es decir,

$$\begin{pmatrix} x & y \\ 2u & -1 \end{pmatrix} \begin{pmatrix} u_x \\ \\ v_x \end{pmatrix} = - \begin{pmatrix} u \\ \\ y \end{pmatrix}$$

Definimos el Jacobiano del sistema (en las variables dependientes u y v)

$$J = \begin{vmatrix} x & y \\ \\ 2u & -1 \end{vmatrix} = x(-1) - (y)(2u) = -x - 2u\,y$$

Aplicando la regla de Cramer

Para u_x

$$\frac{\partial u}{\partial x} = \frac{\begin{vmatrix} -F_x & F_v \\ -G_x & G_v \end{vmatrix}}{J} = \frac{\begin{vmatrix} -u & y \\ -y & -1 \end{vmatrix}}{-x - 2u\,y}$$

Calculamos el determinante

$$(-u)(-1) - y(-y) = u + y^2$$

Por lo tanto,

$$u_x = \frac{u + y^2}{-x - 2u\,y}$$

Para v_x

$$\frac{\partial v}{\partial x} = \frac{\begin{vmatrix} F_u & -F_x \\ G_u & -G_x \end{vmatrix}}{J} = \frac{\begin{vmatrix} x & -u \\ 2u & -y \end{vmatrix}}{-x - 2u\,y}$$

Calculamos el determinante

$$x(-y) - (-u)(2u) = -xy + 2u^2$$

Por lo tanto,

$$v_x = \frac{-xy + 2u^2}{-x - 2u\,y}$$

Paso 3: Derivadas Parciales respecto a y

Ahora, diferenciamos con respecto a y (manteniendo x constante)
Para $F(x,y,u,v)=0$

$$F_y + F_u\,u_y + F_v\,v_y = 0 \quad \Longrightarrow \quad v + x\,u_y + y\,v_y = 0$$

Para $G(x,y,u,v)=0$

$$G_y + G_u\,u_y + G_v\,v_y = 0 \quad \Longrightarrow \quad x + 2u\,u_y - v_y = 0$$

El sistema en forma matricial es el mismo

$$\begin{pmatrix} x & y \\ 2u & -1 \end{pmatrix} \begin{pmatrix} u_y \\ v_y \end{pmatrix} = - \begin{pmatrix} v \\ x \end{pmatrix}$$

Aplicando de nuevo la regla de Cramer
Para u_y

$$\frac{\partial u}{\partial y} = \frac{\begin{vmatrix} -F_y & F_v \\ -G_y & G_v \end{vmatrix}}{J} = \frac{\begin{vmatrix} -v & y \\ -x & -1 \end{vmatrix}}{-x - 2u\,y}$$

Determinante

$$(-v)(-1) - y(-x) = v + xy$$

Por lo tanto,

$$u_y = \frac{v + xy}{-x - 2u\,y}$$

Para v_y

$$\frac{\partial v}{\partial y} = \frac{\begin{vmatrix} F_u & -F_y \\ G_u & -G_y \end{vmatrix}}{J} = \frac{\begin{vmatrix} x & -v \\ 2u & -x \end{vmatrix}}{-x - 2u\,y}$$

Determinante

$$x(-x) - (-v)(2u) = -x^2 + 2u\,v$$

Por lo tanto,

$$v_y = \frac{-x^2 + 2u\,v}{-x - 2u\,y}$$

9 IS-LM

Introducción

El modelo IS–LM determina simultáneamente el ingreso real Y y la tasa de interés r en el corto plazo, mediante el equilibrio en dos mercados: el de bienes y servicios (curva IS) y el monetario (curva LM). Además, incorpora la relación entre ingreso y recaudación fiscal.

$$\begin{cases} Y = C(Y_d) + I(r) + g, \\ M_p = L(Y, r), \\ Y_d = Y - T(Y, \theta). \end{cases}$$

1. Mercado de bienes y servicios (IS)

$$Y = C(Y_d) + I(r) + g.$$

- Y: producción (o ingreso) agregado.
- $C(Y_d)$: consumo de los hogares, función creciente de ingreso disponible Y_d.
- $I(r)$: inversión de las empresas, función decreciente de la tasa r.
- g: gasto público, exógeno.

La curva IS ("Investment–Saving") es el conjunto de (Y, r) que satisfacen esta identidad.

2. Mercado monetario (LM)

$$M_p = L(Y, r).$$

- M_p: oferta real de dinero (stock), fijada por la autoridad monetaria.
- $L(Y, r)$: demanda de dinero, creciente en Y ($\partial L/\partial Y > 0$) y decreciente en r ($\partial L/\partial r < 0$).

La curva LM ("Liquidity preference–Money supply") es el conjunto de puntos de (Y, r) que igualan oferta y demanda de dinero.

3. Ingreso disponible

$$Y_d = Y - T(Y, \theta),$$

donde $T(Y, \theta)$ es la recaudación fiscal, función creciente de Y. El ingreso disponible Y_d es el que efectivamente determina el consumo.

Resolución del modelo IS–LM por derivación implícita

Partimos del modelo:

$$\begin{cases} Y = C(Y_d) + I(r) + g, \\ M_p = L(Y, r), \\ Y_d = Y - T(Y, \theta). \end{cases}$$

Variables endógenas	
Y	Ingreso (producción) de equilibrio
r	Tasa de interés de equilibrio

Variables exógenas	
g	Gasto público
M_p	Oferta real de dinero (stock monetario)
θ	Parámetro(s) fiscales en $T(Y,\theta)$

Hipótesis de comportamiento:

$$0 < C'(Y - T(Y,\theta)) < 1, \quad 0 < T_Y(Y,\theta) < 1, \quad I'(r) < 0, \quad L_Y(Y,r) > 0, \quad L_r(Y,r) < 0, \quad T_\theta(Y,\theta) > 0$$

Sustituyendo Y_d en la primera ecuación obtenemos dos ecuaciones en las variables endógenas Y y r:

$$\begin{cases} F(Y,r;g,\theta) \;=\; Y - C(Y - T(Y,\theta)) - I(r) - g \;=\; 0, \\ G(Y,r;M_p) \;=\; L(Y,r) - M_p \;=\; 0. \end{cases}$$

El objetivo es encontrar $\dfrac{\partial Y}{\partial g}$ y $\dfrac{\partial r}{\partial g}$. Es decir, el efecto que tiene un aumento del gasto público en el producto y en la tasa de interés.

Paso 1: Derivadas parciales

Derivamos ambas ecuaciones respecto a g, recordemos que Y depende de g, también r depende de g por lo tanto debemos utilizar regla de la cadena:

$$F'g = \frac{\partial F}{\partial Y}\frac{\partial Y}{\partial g} + \frac{\partial F}{\partial r}\frac{\partial r}{\partial g} + \frac{\partial F}{\partial g}\frac{\partial g}{\partial g} = 0$$

$$G'g = \frac{\partial G}{\partial Y}\frac{\partial Y}{\partial g} + \frac{\partial G}{\partial r}\frac{\partial r}{\partial g} + \frac{\partial G}{\partial g}\frac{\partial g}{\partial g} = 0$$

Derivamos ambas ecuaciones respecto a g, manteniendo M_p y θ constantes:

$$\frac{\partial}{\partial g}\Big[Y - C(Y - T(Y,\theta)) - I(r) - g\Big] = \Big(1 - C'(Y - T(Y,\theta))\,[1 - T_Y(Y,\theta)]\Big)\,\frac{\partial Y}{\partial g} - I'(r)\,\frac{\partial r}{\partial g} - 1 = 0,$$

$$\frac{\partial}{\partial g}[L(Y,r) - M_p] = L_Y(Y,r)\,\frac{\partial Y}{\partial g} + L_r(Y,r)\,\frac{\partial r}{\partial g} = 0.$$

$$\begin{cases} \dfrac{\partial}{\partial g}\Big[Y - C(Y - T(Y,\theta)) - I(r) - g\Big] = \Big(1 - C'(Y - T(Y,\theta))\,[1 - T_Y(Y,\theta)]\Big)\,\dfrac{\partial Y}{\partial g} - I'(r)\,\dfrac{\partial r}{\partial g} - 1 = 0, \\ \dfrac{\partial}{\partial g}[L(Y,r) - M_p] = L_Y(Y,r)\,\dfrac{\partial Y}{\partial g} + L_r(Y,r)\,\dfrac{\partial r}{\partial g} = 0. \end{cases}$$

Paso 2: Sistema matricial

De aquí obtenemos el sistema lineal

$$\begin{pmatrix} 1 - C'(Y - T(Y,\theta))\,[1 - T_Y(Y,\theta)] & -I'(r) \\ L_Y(Y,r) & L_r(Y,r) \end{pmatrix} \begin{pmatrix} \dfrac{\partial Y}{\partial g} \\ \dfrac{\partial r}{\partial g} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Paso 3: Resolver con la Regla de Cramer

El determinante del sistema es

$$J \;=\; [1 - C'(Y - T(Y,\theta))\,[1 - T_Y(Y,\theta)]]\;L_r(Y,r) \;-\; [-I'(r)]\;L_Y(Y,r).$$

Por Cramer,

$$\frac{\partial Y}{\partial g} = \frac{\begin{vmatrix} 1 & -I'(r) \\ 0 & L_r(Y,r) \end{vmatrix}}{J} = \frac{1 \cdot L_r(Y,r)}{J} = \frac{L_r(Y,r)}{J},$$

$$\frac{\partial r}{\partial g} = \frac{\begin{vmatrix} 1 - C'(Y - T(Y,\theta))\,[1 - T_Y(Y,\theta)] & 1 \\ L_Y(Y,r) & 0 \end{vmatrix}}{J} = \frac{-L_Y(Y,r)}{J}.$$

Paso 4: Signos e interpretación

Dado que

$$0 < C'(Y - T(Y,\theta)) < 1, \quad 0 < T_Y(Y,\theta) < 1, \quad I'(r) < 0, \quad L_Y(Y,r) > 0, \quad L_r(Y,r) < 0,$$

se cumple

$$1 - C'(Y - T(Y,\theta))\,[1 - T_Y(Y,\theta)] > 0, \quad -I'(r) > 0.$$

Por tanto el determinante

$$J = \underbrace{\Big[1 - \overbrace{\underbrace{C'(Y - T(Y,\theta))}_{0<\cdot<1} \times \underbrace{(1 - T_Y(Y,\theta))}_{0<\cdot<1}}^{0<\cdot<1}\Big]}_{>0} \times \underbrace{L_r(Y,r)}_{<0} - \underbrace{(-I'(r))}_{>0} \times \underbrace{L_Y(Y,r)}_{>0} \;<\; 0.$$

Como $L_r < 0,\ L_Y > 0$, el numerador de $\partial Y/\partial g$ es negativo y el de $\partial r/\partial g$ también es negativo:

$$\frac{\partial Y}{\partial g} = \frac{L_r}{J} > 0, \qquad \frac{\partial r}{\partial g} = \frac{-L_Y}{J} > 0.$$

Cuando el gobierno aumenta su gasto g, se desencadenan dos efectos:

Más demanda ⇒ más producción (Y)

- *Gasto directo*: el gobierno inyecta recursos pagando obras, sueldos, compras de insumos, etc.
- *Cadena de efectos*: esos pagos se convierten en ingresos para empresas y trabajadores, que a su vez consumen parte de ese ingreso extra, generando más ventas e ingresos en otros sectores.
- *Multiplicador*: cada unidad de gasto público produce más de una unidad de aumento en el ingreso agregado, pues el dinero circula y se vuelve a gastar varias veces.

Mayor Y ⇒ mayor demanda de dinero ⇒ sube r

- *Demanda de dinero transaccional*: al crecer Y, hogares y empresas realizan más transacciones y necesitan más liquidez.
- *Curva LM*: $L(Y,r)$ crece con Y y cae con r. Con M_p fijo, un aumento de L empuja la tasa r al alza hasta reequilibrar oferta y demanda monetaria.
- *Crowding-out parcial*: al subir r, el crédito se encarece y la inversión privada se modera, atenuando algo el impulso inicial sobre Y, pero sin anularlo.

10 Homogeneidad

Definición de Homogeneidad

Una función $F : \mathbb{R}^n_+ \to \mathbb{R}$ se dice *homogénea de grado* k si, para todo $\lambda > 0$ y todo $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n_+$, se cumple

$$F(\lambda x_1, \dots, \lambda x_n) = \lambda^k F(x_1, \dots, x_n).$$

Cuando $k = 0$, se dice que la función es *homogénea de grado cero*.

Ejemplos

1. **Homogeneidad de grado 1.** Sea

$$F(x, y) = 3x + 5y.$$

   Entonces, para cualquier $\lambda > 0$,

$$F(\lambda x, \lambda y) = 3(\lambda x) + 5(\lambda y) = \lambda (3x + 5y) = \lambda^1 F(x, y).$$

   Por tanto, F es homogénea de grado $k = 1$.

2. **Homogeneidad de grado 0 (relación de precios).** Sea

$$G(x, y) = \frac{x}{y}.$$

   Para $\lambda > 0$,

$$G(\lambda x, \lambda y) = \frac{\lambda x}{\lambda y} = \frac{x}{y} = \lambda^0 G(x, y).$$

   Así, G es homogénea de grado $k = 0$.

Propiedades de las Funciones Homogéneas

1. **Multiplicación de escala:** Si F es homogénea de grado k, entonces para todo $\lambda, \mu > 0$

$$F(\lambda\mu\, \mathbf{x}) = (\lambda\mu)^k F(\mathbf{x}) = \lambda^k F(\mu\, \mathbf{x}).$$

2. **Combinación lineal:** Si F y G son homogéneas de grado k, y $a, b \in \mathbb{R}$, entonces

$$H(\mathbf{x}) = a\, F(\mathbf{x}) + b\, G(\mathbf{x})$$

   es homogénea de grado k.

3. **Producto:** Si F es homogénea de grado k y G de grado m, entonces

$$(F \cdot G)(\mathbf{x}) = F(\mathbf{x})\, G(\mathbf{x})$$

   es homogénea de grado $k + m$.

4. **Cociente:** Si $G(\mathbf{x}) \neq 0$ y F, G son homogéneas de grado k y m respectivamente, entonces

$$\left(\frac{F}{G}\right)(\mathbf{x}) = \frac{F(\mathbf{x})}{G(\mathbf{x})}$$

   es homogénea de grado $k - m$.

5. **Derivadas:** Si F es homogénea de grado k y diferenciable, entonces todas sus derivadas parciales de orden m son homogéneas de grado $k - m$.

Teorema de Euler para Funciones Homogéneas

Teorema (Euler). Sea $F : \mathbb{R}^n_+ \to \mathbb{R}$ una función diferenciable y homogénea de grado k. Entonces

$$\sum_{i=1}^{n} x_i \, \frac{\partial F}{\partial x_i}(x_1, \ldots, x_n) = k \, F(x_1, \ldots, x_n).$$

Demostración. Dado que F es homogénea de grado k, para todo $\lambda > 0$ se cumple

$$F(\lambda x_1, \lambda x_2, \ldots, \lambda x_n) = \lambda^k \, F(x_1, x_2, \ldots, x_n).$$

Definamos la función de una sola variable

$$\Phi(\lambda) \;=\; F(\lambda x_1, \lambda x_2, \ldots, \lambda x_n).$$

Por la homogeneidad, también

$$\Phi(\lambda) \;=\; \lambda^k \, F(x_1, \ldots, x_n).$$

Como F es diferenciable, Φ es diferenciable y podemos derivar ambas expresiones respecto a λ.

- Por la regla de la cadena,

$$\Phi'(\lambda) = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}(\lambda x_1, \ldots, \lambda x_n) \, \frac{d}{d\lambda}(\lambda x_i) = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}(\lambda x_1, \ldots, \lambda x_n) \, x_i.$$

- Por derivar $\lambda^k F(x_1, \ldots, x_n)$,

$$\Phi'(\lambda) = \frac{d}{d\lambda}(\lambda^k \, F(x_1, \ldots, x_n)) = k \, \lambda^{k-1} \, F(x_1, \ldots, x_n).$$

Igualando ambas expresiones para $\Phi'(\lambda)$,

$$\sum_{i=1}^{n} x_i \, \frac{\partial F}{\partial x_i}(\lambda x_1, \ldots, \lambda x_n) = k \, \lambda^{k-1} \, F(x_1, \ldots, x_n).$$

Finalmente, evaluamos en $\lambda = 1$:

$$\sum_{i=1}^{n} x_i \, \frac{\partial F}{\partial x_i}(x_1, \ldots, x_n) = k \, F(x_1, \ldots, x_n),$$

como queríamos demostrar. ■

Ejemplo

Consideremos la función polinómica

$$F(x, y) \;=\; x^3 \;+\; 2 \, x \, y^2.$$

1. **Verificación de homogeneidad.** Observamos que cada término es de grado 3:

$$(\lambda x)^3 = \lambda^3 x^3, \qquad 2 \, (\lambda x) \, (\lambda y)^2 = 2 \, \lambda^3 \, x \, y^2.$$

   Por tanto,

$$F(\lambda x, \lambda y) = (\lambda x)^3 + 2 \, (\lambda x) \, (\lambda y)^2 = \lambda^3 (x^3 + 2 \, x \, y^2) = \lambda^3 \, F(x, y),$$

   y concluimos que F es homogénea de grado $k = 3$.

2. **Derivadas parciales.**

$$F_x(x,y) = \frac{\partial}{\partial x}(x^3 + 2\,x\,y^2) = 3x^2 + 2\,y^2,$$
$$F_y(x,y) = \frac{\partial}{\partial y}(x^3 + 2\,x\,y^2) = 4\,x\,y.$$

3. **Comprobación del Teorema de Euler.** El teorema de Euler establece que, si F es homogénea de grado 3, entonces

$$x\,F_x(x,y) \;+\; y\,F_y(x,y) = 3\,F(x,y).$$

Calculemos el lado izquierdo:

$$x(3x^2 + 2y^2) \;+\; y(4xy) = 3x^3 + 2x\,y^2 + 4x\,y^2 = 3x^3 + 6x\,y^2.$$

Y el lado derecho:

$$3\,F(x,y) = 3(x^3 + 2x\,y^2) = 3x^3 + 6x\,y^2.$$

Como coinciden,

$$x\,F_x + y\,F_y = 3\,F(x,y),$$

confirmamos que el teorema de Euler se cumple.

Interpretación Económica

En economía, la homogeneidad y el Teorema de Euler tienen aplicaciones directas en teoría de la producción y en demanda:

- **Funciones de producción con rendimientos a escala:** Si la función de producción $Y = F(K,L)$ (capital K, trabajo L) es homogénea de grado k:

$$F(\lambda K, \lambda L) = \lambda^k F(K,L),$$

entonces:

  - $k = 1$ implica *rendimientos constantes a escala*: duplicar todos los insumos duplica la producción.
  - $k > 1$ implica *rendimientos crecientes a escala*: duplicar insumos más que duplica la producción.
  - $k < 1$ implica *rendimientos decrecientes a escala*: duplicar insumos menos que duplica la producción.

- **Demanda homogénea de grado cero en precios e ingreso:** Una función de demanda $x_i(p_1,\dots,p_n,M)$ (precios p_j, ingreso M) es homogénea de grado cero:

$$x_i(\lambda p_1,\dots,\lambda p_n,\lambda M) = x_i(p_1,\dots,p_n,M).$$

Esto significa que si todos los precios y el ingreso cambian en la misma proporción, las cantidades demandadas no varían: sólo importan los precios relativos y el poder de compra real.

- **Elasticidades parciales y grado de homogeneidad:** Sea $F(K,L)$ una función de producción diferenciable y homogénea de grado k. Definimos las elasticidades parciales de la producción con respecto a cada factor como

$$\varepsilon_K = \frac{\partial F}{\partial K}\,\frac{K}{F(K,L)}, \qquad \varepsilon_L = \frac{\partial F}{\partial L}\,\frac{L}{F(K,L)}.$$

Aplicando el Teorema de Euler:

$$K\,F_K(K,L) + L\,F_L(K,L) = k\,F(K,L),$$

y dividiendo ambos lados por $F(K,L)$, obtenemos

$$\varepsilon_K + \varepsilon_L = k.$$

***Interpretación:* la suma de las elasticidades parciales de la producción respecto al capital y al trabajo coincide con el grado de homogeneidad de la función. Económicamente, esto significa que los rendimientos a escala coinciden con la suma de elasticidades parciales**

- Si $\varepsilon_K + \varepsilon_L = k = 1$, rendimientos constantes a escala.
- Si $\varepsilon_K + \varepsilon_L = k > 1$, rendimientos crecientes a escala.
- Si $\varepsilon_K + \varepsilon_L = k < 1$, rendimientos decrecientes a escala.

11 Autovalores y autovectores

Definición de autovalor y autovector

Sea $A \in \mathbb{K}^{n\times n}$

Un escalar $\lambda \in \mathbb{K}$ es autovalor de A si existe un vector $v \in \mathbb{K}^n \setminus \{0\}$ tal que

$$A\,v = \lambda\,v$$

El vector v que satisface esta igualdad se llama autovector asociado al autovalor λ

Polinomio característico

Definición

Sea $A \in \mathbb{K}^{n\times n}$. El *polinomio característico* de A se define como

$$\chi_A(\lambda) = \det(A - \lambda I)$$

Cómo hallarlo

1. Formar la matriz $A - \lambda I$
2. Calcular $\det(A - \lambda I)$ como función de λ
3. Simplificar el determinante para obtener un polinomio en λ

Ejemplo para matriz 2×2

Sea

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Entonces

$$A - \lambda I = \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}$$

y

$$\chi_A(\lambda) = \det \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = (a-\lambda)\,(d-\lambda) - b\,c$$

Cálculo de autovalores y autovectores

Procedimiento

1. Resolver $\chi_A(\lambda) = 0$ para obtener los valores propios $\lambda_1, \lambda_2, \lambda_3, ...$
2. Para cada λ_i, resolver $(A - \lambda_i I)\,v = 0$ buscando vectores no nulos

Ejemplo numérico para matriz 2×2

Sea
$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$
Entonces
$$A - \lambda I = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix}$$
por lo que
$$\chi_A(\lambda) = \det(A - \lambda I) = \det \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3$$
Resolvemos
$$\lambda^2 - 4\lambda + 3 = 0$$
obteniendo
$$\lambda_1 = 1, \quad \lambda_2 = 3$$

Cálculo paso a paso de los autovectores

Para $\lambda_1 = 1$ Calculamos
$$A - I = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$
Resolvemos el sistema
$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
Lo que equivale a la ecuación única
$$x + y = 0$$
De aquí $y = -x$, por lo que un autovector es
$$v_1 = \begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad x \neq 0$$
Tomamos $v_1 = (1, -1)^T$

Para $\lambda_2 = 3$ Calculamos
$$A - 3I = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$
Resolvemos el sistema
$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
Equivalente a
$$-x + y = 0$$
$$y = x$$
Por lo tanto,
$$v_2 = \begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad x \neq 0$$
Tomamos $v_2 = (1, 1)^T$

En conclusión los autovectores asociados son $v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ para $\lambda = 1$ y $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ para $\lambda = 3$

Relación entre coeficientes del polinomio característico, traza y determinante

Sea el polinomio característico de A de grado n:

$$\chi_A(\lambda) = \det(A - \lambda I) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$$

Coeficientes y autovalores

Si $\lambda_1, \ldots, \lambda_n$ son las raíces de χ_A:

$$c_{n-1} = -\sum_{i=1}^{n} \lambda_i, \quad c_0 = (-1)^n \prod_{i=1}^{n} \lambda_i$$

En palabras: el coeficiente c_{n-1} representa el negativo de la suma de todos los autovalores, reflejando que la traza de A coincide con dicha suma. Por su parte, c_0 corresponde al producto de los autovalores multiplicado por $(-1)^n$.

Traza y determinante

Recordando que:

$$\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}, \quad \det(A) = \prod_{i=1}^{n} \lambda_i$$

Existe una relación entre autovalores y traza y determinante:

$$\sum_{i=1}^{n} \lambda_i = \operatorname{tr}(A), \qquad \prod_{i=1}^{n} \lambda_i = \det(A)$$

En particular, para $n = 2$ y $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$:

$$\chi_A(\lambda) = \lambda^2 - (a + d)\lambda + (ad - bc),$$

por lo que

$$\operatorname{tr}(A) = a + d, \quad \det(A) = ad - bc$$

Ejemplo numérico

Para la matriz

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

sus autovalores son $\lambda_1 = 1$ y $\lambda_2 = 3$. En este caso,

$$\operatorname{tr}(A) = 2 + 2 = 4 = \lambda_1 + \lambda_2 = 1 + 3, \quad \det(A) = 2 \cdot 2 - 1 \cdot 1 = 3 = \lambda_1\lambda_2 = 1 \cdot 3$$

Además, el polinomio característico calculado fue

$$\chi_A(\lambda) = \lambda^2 - 4\lambda + 3$$

De donde se observa que:

- El coeficiente de λ es -4, que coincide con $-(\lambda_1 + \lambda_2) = -(1 + 3)$.
- El término constante es 3, que coincide con el producto $\lambda_1 \cdot \lambda_2 = 1 \cdot 3$.

Raíces iguales y multiplicidad

Multiplicidad algebraica

Sea $A \in \mathbb{K}^{n\times n}$ y λ un autovalor de A.

- La *multiplicidad algebraica* de λ, denotada $\mathrm{mult}_{\mathrm{alg}}(\lambda)$, es el grado con que λ aparece como raíz del polinomio característico $\chi_A(\lambda)$.

$$\chi_A(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k}, \quad \mathrm{mult}_{\mathrm{alg}}(\lambda_i) = m_i.$$

- Obsérvese que $\sum_{i=1}^{k} m_i = n$, es decir, la suma de todas las multiplicidades algebraicas es el grado del polinomio característico.

Multiplicidad geométrica

Para cada autovalor λ, definimos su *multiplicidad geométrica* como

$$\mathrm{mult}_{\mathrm{geom}}(\lambda) = \dim(\ker(A - \lambda I))$$

Donde ker() es el conjunto de todos los vectores que se envían al vector cero (espacio nulo). **Es el número de vectores linealmente independientes que generan el espacio propio asociado a λ.**

Relación entre ambas multiplicidades

Para todo autovalor λ de A se tiene siempre:

$$1 \leq \mathrm{mult}_{\mathrm{geom}}(\lambda) \leq \mathrm{mult}_{\mathrm{alg}}(\lambda)$$

Ejemplo numérico

Considérese

$$B = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

Su polinomio característico es

$$\chi_B(\lambda) = (2-\lambda)^2$$

luego $\lambda = 2$ tiene $\mathrm{mult}_{\mathrm{alg}}(2) = 2$. En cambio,

Cálculo de los autovectores para $\lambda = 2$ Partimos de

$$B - 2I = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

y buscamos todos los vectores $v = (x, y)^T$ tales que

$$(B - 2I)\, v = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

De la igualdad $(y, 0)^T = (0, 0)^T$ se obtiene la única ecuación

$$y = 0$$

Así, x es libre y los autovectores asociados a $\lambda = 2$ son

$$v = \begin{pmatrix} x \\ 0 \end{pmatrix}, \quad x \neq 0$$

Por ejemplo, para $x = 1$ podemos tomar como autovector base

$$v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Y por lo tanto tenemos multiplicidad geométrica igual a 1, porque solamente hay un vector linealmente independiente.

Otro ejemplo numérico

$$C = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Polinomio característico

$$\chi_C(\lambda) = \det(C - \lambda I) = \begin{vmatrix} 2-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} = (2-\lambda)^2$$

Entonces,

$$\lambda = 2 \quad \text{con} \quad \text{mult}_{\text{alg}}(2) = 2$$

Autovectores para $\lambda = 2$

$$C - 2I = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow (C - 2I)v = 0 \quad \text{para todo } v \in \mathbb{R}^2$$

Por lo tanto, todos los vectores no nulos son autovectores y el espacio propio tiene dimensión 2:

$$\text{mult}_{\text{geom}}(2) = 2$$

Conclusión: En este caso,

$$\text{mult}_{\text{alg}}(2) = \text{mult}_{\text{geom}}(2) = 2$$

Un conjunto base de autovectores puede ser:

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

12 Formas cuadráticas

Definición

Una forma cuadrática en las variables $x_1, \ldots, x_n$ es un polinomio homogéneo de grado dos:

$$q(x_1, \ldots, x_n) = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2(a_{12}x_1x_2 + a_{13}x_1x_3 + \cdots + a_{n-1,n}x_{n-1}x_n)$$

También se puede expresar de forma compacta mediante sumatorias:

$$q(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2 \sum_{1\leq i<j\leq n} a_{ij}x_ix_j$$

Representación matricial

Escribiendo

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{y} \quad Q = (a_{ij})_{i,j=1}^{n}$$

con Q simétrica (es decir, $a_{ij} = a_{ji}$), la forma cuadrática se expresa de modo compacto como

$$q(x) = x^T\, Q\, x$$

En esta matriz, cada entrada $q_{ii} = a_{ii}$ multiplica x_i^2 y cada par de entradas $Q_{ij} = Q_{ji} = a_{ij}$ multiplica $2\, x_ix_j$.

Ejemplo para 3 variables

Consideremos la forma cuadrática en (x_1, x_2, x_3):

$$q(x_1, x_2, x_3) = 3x_1^2 + 4x_2^2 + 5x_3^2 + 2x_1x_2 + 4x_1x_3 + 6x_2x_3$$

La matriz asociada es

$$Q = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 4 & 3 \\ 2 & 3 & 5 \end{pmatrix},$$

con lo que

$$q(x) = x^TQx = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 3 & 1 & 2 \\ 1 & 4 & 3 \\ 2 & 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Clasificación de formas cuadráticas según autovalores

Sea una forma cuadrática $q(x) = x^TQx$ con matriz simétrica Q y autovalores $\lambda_1, \ldots, \lambda_n$.

- **Definida positiva:** $q(x) > 0$ para todo $x \neq 0$. Se cumple si y solo si $\lambda_i > 0 \quad \forall i$.
- **Semidefinida positiva:** $q(x) \geq 0$ para todo x. Se cumple si y solo si $\lambda_i \geq 0 \quad \forall i$.
- **Definida negativa:** $q(x) < 0$ para todo $x \neq 0$. Se cumple si y solo si $\lambda_i < 0 \quad \forall i$.
- **Semidefinida negativa:** $q(x) \leq 0$ para todo x. Se cumple si y solo si $\lambda_i \leq 0 \quad \forall i$.
- **Indefinida:** $q(x)$ toma valores positivos y negativos. Ocurre cuando $\exists\, i, j:\ \lambda_i > 0,\ \lambda_j < 0$.

Ejemplo práctico

Estudiar el signo de

$$\varphi(x_1, x_2, x_3) = x_1^2 \; - \; 2x_1x_2 \; + \; 3x_2^2 \; + \; x_3^2$$

sujeta a

$$\begin{cases} x_1 + x_2 - x_3 = 0, \\ -x_1 + x_3 = 0. \end{cases}$$

1. Parametrización de las restricciones.
De $-x_1 + x_3 = 0$ obtenemos $x_3 = x_1$. Sustituyendo en $x_1 + x_2 - x_3 = 0$ queda

$$x_1 + x_2 - x_1 = 0 \implies x_2 = 0.$$

Luego todo vector (x_1, x_2, x_3) que cumple las restricciones es

$$(x_1, x_2, x_3) = (x_1, 0, x_1) = y \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

con $y = x_1 \in \mathbb{R}$.

2. Forma restringida.
Sustituyendo $(x_1, 0, x_1)$ en φ:

$$\varphi(y, 0, y) = y^2 - 0 + 0 + y^2 = 2\,y^2.$$

Es decir, la forma sobre la variable libre y queda

$$\psi(y) = 2\,y^2.$$

3. Conclusión.
Como $2\,y^2 > 0$ para todo $y \neq 0$, la forma cuadrática φ es *positiva definida* en el subespacio definido por las dos restricciones.

Método matricial (sin sustitución directa)

Queremos estudiar

$$\varphi(x) = x^T A x,$$

sujeta a las restricciones lineales $C\,x = 0$.

1. Matriz A y restricción C. De

$$\varphi(x_1, x_2, x_3) = x_1^2 - 2x_1x_2 + 3x_2^2 + x_3^2$$

leemos

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

y las ecuaciones

$$\begin{cases} x_1 + x_2 - x_3 = 0, \\ -x_1 + x_3 = 0 \end{cases} \implies C = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}.$$

2. Base del espacio restringido (cálculo de $\ker C$**).** Partimos de

$$C = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

y resolvemos la ecuación lineal

$$C\,x = 0 \iff \begin{cases} x_1 + x_2 - x_3 = 0, \\ -x_1 + x_3 = 0. \end{cases}$$

De la segunda ecuación se obtiene

$$x_3 = x_1.$$

Sustituyendo en la primera:

$$x_1 + x_2 - x_1 = 0 \implies x_2 = 0.$$

Así, todo vector x en $\ker C$ tiene la forma

$$x = \begin{pmatrix} x_1 \\ 0 \\ x_1 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Por tanto una base de $\ker C$ está dada por la columna

$$B = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

y cualquier x que satisfaga las restricciones se puede escribir $x = B\,y$, donde $y \in \mathbb{R}$ es la variable libre.

3. Forma cuadrática restringida. La matriz de la forma en la variable libre y es

$$M = B^T A B = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 2.$$

Por tanto la forma restringida es

$$\psi(y) = y^T M y = 2\,y^2.$$

4. Conclusión. Como $M = 2 > 0$, la forma φ es *positiva definida* sobre el subespacio dado por las restricciones.

13 Menores

Definiciones de Menores

Sea $A = (a_{ij})_{i,j=1}^{n} \in \mathbb{K}^{n\times n}$ una matriz simétrica (aunque las definiciones de menores se aplican a cualquier matriz cuadrada, el contexto de formas cuadráticas usualmente implica simetría).

Menor (general) de orden k. Un *menor (general) de orden k* de A es el determinante de una submatriz $k \times k$ de A. Esta submatriz se forma seleccionando un conjunto $I = \{i_1, \dots, i_k\}$ de k índices de filas (con $i_1 < i_2 < \cdots < i_k$) y un conjunto $J = \{j_1, \dots, j_k\}$ de k índices de columnas (con $j_1 < j_2 < \cdots < j_k$). Si $A[I, J]$ denota la submatriz formada por las filas en I y las columnas en J, es decir, $A[I, J] = (a_{i_p j_q})_{p,q=1}^{k}$, entonces el menor es $\det(A[I, J])$.

Menor principal de orden k. Un *menor principal de orden k* es un menor de orden k donde el conjunto de índices de las filas seleccionadas es el mismo que el conjunto de índices de las columnas seleccionadas. Es decir, si $I = \{i_1, \dots, i_k\} \subset \{1, 2, \dots, n\}$ con $i_1 < i_2 < \cdots < i_k$, el menor principal asociado a I (a veces denotado M_I o Δ_I) es:

$$\Delta_I = \det(A[I, I]),$$

donde $A[I, I] = (a_{i_p i_q})_{p,q=1}^{k}$. Estos son los determinantes de las submatrices principales de A.

Menor principal inicial (o dominante) de orden k. Un *menor principal inicial* o *menor principal líder* de orden k (usualmente denotado Δ_k) es el menor principal obtenido al seleccionar las primeras k filas y las primeras k columnas de A. Es decir, corresponde a tomar $I = \{1, 2, \dots, k\}$:

$$\Delta_k = \det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}.$$

En resumen, para una matriz A

- *Menor (general) de orden k*: Determinante de cualquier submatriz $k \times k$ (los índices de las k filas elegidas pueden ser distintos de los índices de las k columnas elegidas).
- *Menor principal de orden k* (Δ_I): Determinante de una submatriz $k \times k$ formada tomando el mismo conjunto de k índices para las filas y para las columnas.
- *Menor principal líder (o de orden inicial) de orden k* (Δ_k): Determinante de la submatriz $k \times k$ superior izquierda de A. Es un caso particular de menor principal.

Ejemplo: menores de una matriz 3×3

Sea

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

1. Menores de orden 1 (todos los elementos de A):

$$\begin{aligned} M_{11} &= 2, & M_{12} &= -1, & M_{13} &= 0, \\ M_{21} &= -1, & M_{22} &= 3, & M_{23} &= 1, \\ M_{31} &= 0, & M_{32} &= 1, & M_{33} &= 2. \end{aligned}$$

2. Menores principales de orden 1:

$$M_{11} = 2, \quad M_{22} = 3, \quad M_{33} = 2.$$

3. Menores de orden 2 (todas las submatrices 2×2):

	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$
$\{1,2\}$	$\det\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix} = 5$	$\det\begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix} = 2$	$\det\begin{pmatrix} -1 & 0 \\ 3 & 1 \end{pmatrix} = -1$
$\{1,3\}$	$\det\begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix} = 2$	$\det\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 4$	$\det\begin{pmatrix} -1 & 0 \\ 1 & 2 \end{pmatrix} = -2$
$\{2,3\}$	$\det\begin{pmatrix} -1 & 3 \\ 0 & 1 \end{pmatrix} = -1$	$\det\begin{pmatrix} -1 & 1 \\ 0 & 2 \end{pmatrix} = -2$	$\det\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} = 5$

4. Menores principales de orden 2: Sólo los subíndices coinciden:

$$\det A[\{1,2\},\{1,2\}] = 5, \quad \det A[\{1,3\},\{1,3\}] = 4, \quad \det A[\{2,3\},\{2,3\}] = 5.$$

5. Menores principales líderes (o menores principales de orden inicial) :

$$\Delta_1 = \det([2]) = 2, \quad \Delta_2 = \det\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix} = 5, \quad \Delta_3 = \det A = 8.$$

Otro ejemplo

Para

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 3 & 1 \\ 0 & 1 & 4 \end{pmatrix}$$

los menores principales líderes son

$$A_1 = (2) \quad , \quad A_2 = |\begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}| \quad , \quad A_3 = |A|$$

Clasificación de formas cuadráticas mediante menores

Sea la forma cuadrática

$$Q(x) = x^T A\, x$$

con A simétrica.

Menores principales iniciales

Denotamos por Δ_k el menor principal inicial de orden k, es decir

$$\Delta_k = \det(A[\{1,\ldots,k\},\{1,\ldots,k\}]),$$

y por Δ_I el menor principal cualquiera, correspondiente al subconjunto de índices $I \subset \{1,\ldots,n\}$.[1]

[1] Para un subconjunto $I \subset \{1,\ldots,n\}$, definimos $\Delta_I := \det(A[I,I])$ como el *menor principal* asociado a I. Por ejemplo, $\Delta_{\{1,3\}} = \det\begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix}$. Cuando $I = \{1,\ldots,k\}$ escribimos simplemente Δ_k, el *menor principal líder* de orden k.

Criterio de Sylvester (definitud estricta)

- Q es *definida positiva* si y solo si
$$\Delta_k > 0 \quad \text{para todo } k = 1, \ldots, n.$$
- Q es *definida negativa* si y solo si
$$(-1)^k \Delta_k > 0 \quad \text{para todo } k = 1, \ldots, n.$$

Criterio para semidefinitud

- Q es *semidefinida positiva* si y solo si todos los menores principales $\Delta_I \geq 0$ para todo $I \subset \{1, \ldots, n\}$.
- Q es *semidefinida negativa* si y solo si $(-1)^{|I|} \Delta_I \geq 0$ para todo $I \subset \{1, \ldots, n\}$.

Criterio para formas cuadráticas indefinidas

Si existen índices $I, J \subset \{1, \ldots, n\}$ tales que los menores principales satisfacen $\Delta_I > 0$ y $\Delta_J < 0$, entonces Q es *indefinida*

Observaciones

- Las condiciones del criterio de Sylvester son necesarias y suficientes para la *definitud* de Q, pero para la *semidefinitud* son solo necesarias.
- Una matriz que es definida positiva es semidefinida positiva y si es negativa también es semidefinida negativa.

Algunos casos posibles con una matriz simétrica 4x4

Sea $|A_i|$ el menor principal líder de orden i.[2]

(a) Si $|A_1| > 0$, $|A_2| > 0$, $|A_3| > 0$ y $|A_4| > 0$, entonces A es definida positiva.

(b) Si $|A_1| < 0$, $|A_2| > 0$, $|A_3| < 0$ y $|A_4| > 0$, entonces A es definida negativa.

(c) Si $|A_1| > 0$, $|A_2| > 0$, $|A_3| = 0$ y $|A_4| < 0$, entonces A es indefinida debido al cuarto menor principal.

(d) Si $|A_1| < 0$, $|A_2| < 0$, $|A_3| < 0$ y $|A_4| < 0$, entonces A es indefinida debido al segundo y cuarto menor principal.

(e) Si $|A_1| = 0$, $|A_2| < 0$, $|A_3| > 0$ y $|A_4| = 0$, entonces A es indefinida debido al segundo menor principal.

(f) Si $|A_1| > 0$, $|A_2| = 0$, $|A_3| > 0$ y $|A_4| > 0$, entonces A no es definida. No es semidefinida negativa, pero puede ser semidefinida positiva. Para comprobar la semidefinitud positiva, es necesario verificar los 15 menores principales de A, no sólo los cuatro primeros. Si ninguno es negativo, A es semidefinida positiva; si al menos uno es negativo, A es indefinida.

(g) Si $|A_1| = 0$, $|A_2| > 0$, $|A_3| = 0$ y $|A_4| > 0$, entonces A no es definida, pero puede ser semidefinida positiva o semidefinida negativa. Para decidirlo, es necesario comprobar nuevamente los 15 menores principales de A.

[2] En esta sección, A_i denota la *submatriz principal líder* de orden i, obtenida tomando las primeras i filas y columnas de A. En consecuencia, $|A_i| = \det(A[\{1, \ldots, i\}, \{1, \ldots, i\}])$, que también denotamos por Δ_i.

Ejemplo

Sea

$$E = \begin{pmatrix} -4 & 1 & 0 \\ 1 & -3 & 1 \\ 0 & 1 & -2 \end{pmatrix}$$

Menor principal líder de orden 1

$$\Delta_1 = \det(E[\{1\},\{1\}]) = \det([-4]) = -4$$

Menor principal líder de orden 2

$$\Delta_2 = \det \begin{pmatrix} -4 & 1 \\ 1 & -3 \end{pmatrix} = (-4)\,(-3) - 1 \cdot 1 = 12 - 1 = 11$$

Menor principal líder de orden 3

$$\begin{aligned} \Delta_3 = \det(E) &= -4 \det \begin{pmatrix} -3 & 1 \\ 1 & -2 \end{pmatrix} - 1 \det \begin{pmatrix} 1 & 1 \\ 0 & -2 \end{pmatrix} + 0 \det \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix} \\ &= -4((-3)(-2) - 1 \cdot 1) - 1(1 \cdot (-2) - 1 \cdot 0) \\ &= -4(6-1) - 1(-2) = -20 + 2 = -18 \end{aligned}$$

Verificación del criterio de Sylvester

$$(-1)^1\,\Delta_1 = -(-4) = 4 > 0$$

$$(-1)^2\,\Delta_2 = 11 > 0$$

$$(-1)^3\,\Delta_3 = -(-18) = 18 > 0$$

Por tanto E es definida negativa

Otro ejemplo

Sea

$$C = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Menores principales de orden 1

$$\det(C[\{1\},\{1\}]) = 1 \quad \det(C[\{2\},\{2\}]) = 1 \quad \det(C[\{3\},\{3\}]) = 0$$

Menores principales de orden 2

$$\det(C[\{1,2\},\{1,2\}]) = \det \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 0$$

$$\det(C[\{1,3\},\{1,3\}]) = \det \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 0$$

$$\det(C[\{2,3\},\{2,3\}]) = \det \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 0$$

Menor principal de orden 3

$$\det(C) = 0$$

Como todos los menores principales son ≥ 0, C es semidefinida positiva según el criterio completo de semidefinitud positiva

Otro ejemplo

Sea

$$D = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Menores principales de orden 1

$$\det(D[\{1\},\{1\}]) = 4$$
$$\det(D[\{2\},\{2\}]) = 1$$
$$\det(D[\{3\},\{3\}]) = 0$$

Menores principales de orden 2

$$\det(D[\{1,2\},\{1,2\}]) = \det \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix} = 4 \cdot 1 - 2 \cdot 2 = 0$$

$$\det(D[\{1,3\},\{1,3\}]) = \det \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix} = 4 \cdot 0 - 0 \cdot 0 = 0$$

$$\det(D[\{2,3\},\{2,3\}]) = \det \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1 \cdot 0 - 0 \cdot 0 = 0$$

Menores principales de orden 3

$$\det(D) = 0$$

Como todos los menores principales son ≥ 0 la matriz es semidefinida positiva.

Otro ejemplo

Sea

$$F = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Menores principales de orden 1

$$\det(F[\{1\},\{1\}]) = -1$$
$$\det(F[\{2\},\{2\}]) = -2$$
$$\det(F[\{3\},\{3\}]) = 0$$

Menores principales de orden 2

$$\det(F[\{1,2\},\{1,2\}]) = \det \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} = (-1) \cdot (-2) - 0 \cdot 0 = 2$$

$$\det(F[\{1,3\},\{1,3\}]) = \det \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} = (-1) \cdot 0 - 0 \cdot 0 = 0$$

$$\det(F[\{2,3\},\{2,3\}]) = \det \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix} = (-2) \cdot 0 - 0 \cdot 0 = 0$$

Menor principal de orden 3

$$\det(F) = 0$$

Verificación del criterio de semidefinitud negativa

$$(-1)^1\,\Delta_1:\ (-1)\cdot(-1)=1\geq 0$$

$$(-1)^1\,\Delta_1^{(2,2)}:\ (-1)\cdot(-2)=2\geq 0$$

$$(-1)^1\,\Delta_1^{(3,3)}:\ (-1)\cdot 0=0\geq 0$$

$$(-1)^2\,\Delta_2^{\{1,2\}}:\ 2\geq 0$$

$$(-1)^2\,\Delta_2^{\{1,3\}}:\ 0\geq 0$$

$$(-1)^2\,\Delta_2^{\{2,3\}}:\ 0\geq 0$$

$$(-1)^3\,\Delta_3:\ (-1)\cdot 0=0\geq 0$$

Por tanto F es semidefinida negativa.

14 Concavidad y convexidad

Definición de convexidad y concavidad para funciones de una variable

Sea $f\colon I \subseteq \mathbb{R} \to \mathbb{R}$, donde I es un intervalo convexo. Decimos que:

- f es *convexa* si para todo $x_1, x_2 \in I$ y todo $t \in [0,1]$ se cumple

$$f((1-t)x_1 + t\,x_2) \;\leq\; (1-t)f(x_1) + t\,f(x_2)$$

  Intuición geométrica: El término

$$(1-t)f(x_1) + t\,f(x_2)$$

  coincide con el valor que adopta la recta secante en la posición intermedia $(1-t)x_1 + t\,x_2$. Por tanto, la desigualdad

$$f((1-t)x_1 + t\,x_2) \;\leq\; (1-t)f(x_1) + t\,f(x_2)$$

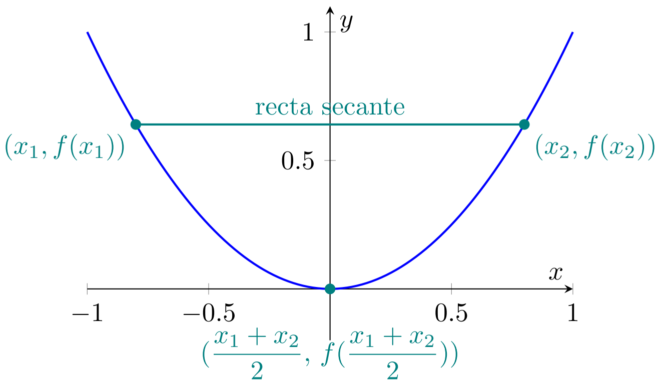
  expresa que, en cada punto entre x_1 y x_2, la curva de f permanece siempre por debajo de esa secante

- f es *cóncava* si para todo $x_1, x_2 \in I$ y todo $t \in [0,1]$ se cumple

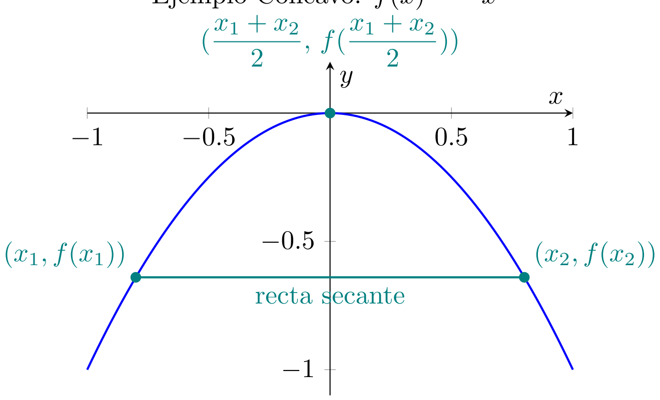
$$f((1-t)x_1 + t\,x_2) \;\geq\; (1-t)f(x_1) + t\,f(x_2)$$

  Intuición geométrica: Aquí la desigualdad se invierte: el gráfico de f "queda por encima" de la recta secante entre los puntos dados

Ejemplo Convexo: $f(x) = x^2$



Ejemplo Cóncavo: $f(x) = -x^2$

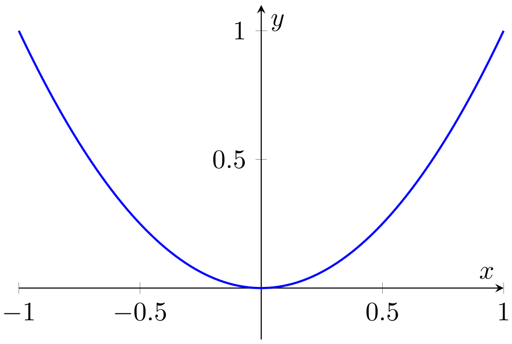


Relacionando la concavidad y convexidad con conjuntos convexos

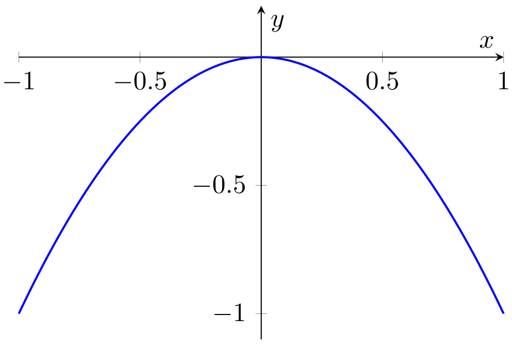
- *Convexidad:* la región
$$\{(x, y) \mid y \geq f(x)\}$$
es un conjunto convexo. Esto significa que si tomas dos puntos (x_1, y_1) y (x_2, y_2) por encima del gráfico de f, el segmento que los une permanece siempre por encima de la curva

- *Concavidad:* la región
$$\{(x, y) \mid y \leq f(x)\}$$
es un conjunto convexo. En este caso, cualquier segmento trazado entre dos puntos por debajo del gráfico de f queda enteramente dentro de esa región

Ejemplo Convexo: $f(x) = x^2$



Ejemplo Cóncavo: $f(x) = -x^2$



Concavidad y convexidad en ^{2}

Definición:

- f es *convexa* si para todo (x_1, y_1), (x_2, y_2) y todo $t \in [0, 1]$ se cumple
$$f((1-t)(x_1, y_1) + t(x_2, y_2)) \ \leq \ (1-t)\, f(x_1, y_1) + t\, f(x_2, y_2)$$

- f es *cóncava* si $-f$ es convexa, es decir la desigualdad se invierte

Intuición geométrica: La intuición se mantiene, el conjunto que está por debajo de la gráfica de una función cóncava es un conjunto convexo y el conjunto que está por arriba de una función convexa es un conjunto convexo

Regla en base a derivadas segundas

Recordatorio: criterio para funciones de una variable
Sea $g\colon I \subset \mathbb{R} \to \mathbb{R}$ de clase C^2 en un intervalo I. Entonces

$$\begin{cases} g''(t) \geq 0 \quad \forall\, t \in I & \Longrightarrow \; g \text{ es convexa en } I \\ g''(t) \leq 0 \quad \forall\, t \in I & \Longrightarrow \; g \text{ es cóncava en } I \end{cases}$$

Sea $f\colon \mathbb{R}^2 \to \mathbb{R}$ de clase C^2.

Tomemos dos puntos (x_1, y_1) y (x_2, y_2) en ^{2} y sea

$$g(t) \;=\; f((1-t)x_1 + t\,x_2,\; (1-t)y_1 + t\,y_2) \quad t \in [0,1]$$

Intuición sobre g. Para entender qué hace g, primero observemos que

$$(x(t), y(t)) \;=\; ((1-t)x_1 + t\,x_2,\; (1-t)y_1 + t\,y_2)$$

es la parametrización del segmento rectilíneo que une los puntos (x_1, y_1) y (x_2, y_2) en el plano Por tanto,

$$g(t) \;=\; f(x(t), y(t))$$

No es más que el valor de f al desplazarnos a lo largo de ese segmento

- Cuando $t = 0$ estamos en (x_1, y_1)
- Cuando $t = 1$ estamos en (x_2, y_2)
- Para $t \in (0,1)$ recorremos los puntos intermedios

Por la regla de la cadena aplicada a dos variables:

$$\begin{aligned} g'(t) &= f_x(x(t), y(t))\,\frac{d}{dt}((1-t)x_1 + t\,x_2) + f_y(x(t), y(t))\,\frac{d}{dt}((1-t)y_1 + t\,y_2) \\ &= f_x(x(t), y(t))\,(x_2 - x_1) + f_y(x(t), y(t))\,(y_2 - y_1) \end{aligned}$$

$$\begin{aligned} g''(t) &= \frac{d}{dt}[f_x(x(t), y(t))]\,(x_2 - x_1) + \frac{d}{dt}[f_y(x(t), y(t))]\,(y_2 - y_1) \\ &= [f_{xx}(x(t), y(t))\,x'(t) + f_{xy}(x(t), y(t))\,y'(t)]\,(x_2 - x_1) \\ &\quad + [f_{yx}(x(t), y(t))\,x'(t) + f_{yy}(x(t), y(t))\,y'(t)]\,(y_2 - y_1) \end{aligned}$$

Sabiendo que $x'(t) = x_2 - x_1$ y $y'(t) = y_2 - y_1$:

$$\begin{aligned} g''(t) &= f_{xx}(x(t), y(t))\,(x_2 - x_1)\,(x_2 - x_1) + f_{xy}(x(t), y(t))\,(y_2 - y_1)\,(x_2 - x_1) \\ &\quad + f_{yx}(x(t), y(t))\,(x_2 - x_1)\,(y_2 - y_1) + f_{yy}(x(t), y(t))\,(y_2 - y_1)\,(y_2 - y_1) \end{aligned}$$

Usando $f_{xy} = f_{yx}$, se agrupan los términos mixtos:

$$g''(t) = f_{xx}(x(t), y(t))\,(x_2 - x_1)^2 + 2\,f_{xy}(x(t), y(t))\,(x_2 - x_1)(y_2 - y_1) + f_{yy}(x(t), y(t))\,(y_2 - y_1)^2$$

Redefiniendo para mayor claridad:

$$w = x_2 - x_1 \qquad z = y_2 - y_1$$

$$\begin{aligned} g''(t) &= f_{xx}(x(t), y(t))\,w^2 + 2\,f_{xy}(x(t), y(t))\,w\,z + f_{yy}(x(t), y(t))\,z^2 \\ &= \begin{pmatrix} w & z \end{pmatrix} \begin{pmatrix} f_{xx}(x(t), y(t)) & f_{xy}(x(t), y(t)) \\ f_{xy}(x(t), y(t)) & f_{yy}(x(t), y(t)) \end{pmatrix} \begin{pmatrix} w \\ z \end{pmatrix} \end{aligned}$$

Ahora, si la matriz Hessiana $H_f(x,y) = [f_{ij}(x,y)]$ es semidefinida positiva en todo punto, entonces para cada $t \in [0,1]$

$$g''(t) \ \geq \ 0$$

Lo cual implica que g es convexa

Definamos la *recta secante* entre $(0, g(0))$ y $(1, g(1))$:

$$L(t) \ = \ g(0) \ + \ t(g(1) - g(0)) \ = \ (1-t)\, g(0) \ + \ t\, g(1)$$

Si la Hessiana $H_f(x,y)$ es semidefinida positiva, entonces $g''(t) \geq 0$ en $[0,1]$, lo que por teoría de funciones de una variable implica

$$g(t) \ \leq \ L(t) \ = \ (1-t)\, g(0) + t\, g(1)$$

Finalmente, usando $g(0) = f(x_1,y_1)$ y $g(1) = f(x_2,y_2)$, concluimos

$$f((1-t)(x_1,y_1) + t\,(x_2,y_2)) \ = \ g(t) \ \leq \ (1-t)\, f(x_1,y_1) \ + \ t\, f(x_2,y_2)$$

que es precisamente la condición de convexidad de f en $\mathbb{R}^2$

De forma análoga, si H_f es semidefinida *negativa*, entonces

$$g''(t) \leq 0 \quad \Longrightarrow \quad g(t) \geq (1-t)\, g(0) + t\, g(1)$$

y por tanto f es *cóncava*

Extensión a ^{n}

El mismo argumento—considerar la función unidimensional $g(t) = f((1-t)x + t\,y)$—se aplica en ^{n}. Si la matriz de derivadas segundas (matriz hessiana) es semidefinida positiva para todo x, entonces en cada dirección $y - x$ la derivada segunda $g''(t) \geq 0$, lo que implica la convexidad de g y, por tanto, la convexidad de f. Análogamente, semidefinida negativa da concavidad de f en ^{n}

Además, si la Hessiana $H_f(x)$ es *definida positiva* en todo punto, entonces f es *estrictamente convexa*; y si $H_f(x)$ es *definida negativa* en todo punto, entonces f es *estrictamente cóncava*

En resumen

Hessiana $H_f(x)$	**Condición** ($\forall x \in D$ convexo, $\forall h \in \mathbb{R}^n$)	**Función** f
Semidefinida positiva	$h^T H_f(x)\, h \ \geq \ 0$	Convexa
Definida positiva	$h^T H_f(x)\, h \ > \ 0$ para todo $h \neq 0$	Estr. convexa
Semidefinida negativa	$h^T H_f(x)\, h \ \leq \ 0$	Cóncava
Definida negativa	$h^T H_f(x)\, h \ < \ 0$ para todo $h \neq 0$	Estr. cóncava

(Aquí D denota el dominio de f.)

Ejemplo

Consideremos la función

$$f(x,y) \ = \ x^2 \ + \ 2\,x\,y \ + \ 3\,y^2 \qquad (x,y) \in^2$$

Cálculo del Hessiano

Las segundas derivadas parciales son

$$f_{xx} = 2, \quad f_{xy} = 2, \quad f_{yy} = 6$$

Por tanto la matriz Hessiana es

$$H_f(x,y) = \begin{pmatrix} 2 & 2 \\ 2 & 6 \end{pmatrix}$$

Comprobación de definitud (criterio de Sylvester).
Calculamos los menores principales iniciales:

$$\Delta_1 = f_{xx} = 2 > 0 \qquad \Delta_2 = \det H_f = (2)(6) - (2)^2 = 12 - 4 = 8 > 0$$

Como $\Delta_1 > 0$ y $\Delta_2 > 0$, H_f es *definida positiva* en todo ^{2}

Conclusión:
Al ser su Hessiana definida positiva para todo (x, y), la función f es *estrictamente convexa* en ^{2}

Otro ejemplo

Consideremos

$$f(x, y) = -x^2 \qquad (x, y) \in^2$$

Cálculo del Hessiano

$$f_{xx} = -2, \quad f_{xy} = 0, \quad f_{yy} = 0$$

Por tanto la matriz Hessiana es

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

Comprobación de semidefinición.
Los menores principales iniciales son

$$\Delta_1 = f_{xx} = -2 < 0 \qquad \Delta_2 = \det H_f = (-2) \cdot 0 - 0^2 = 0$$

Como $\Delta_1 < 0$ y $\Delta_2 = 0$, debemos analizar un menor principal más: $\Delta_1^{2,2} = 0 \leq 0$, entonces H_f es *semidefinida negativa* (pero no definida negativa)

Conclusión.
Al ser su Hessiana semidefinida negativa en todo ^{2}, f es *cóncava*. Además, f no es convexa, ni estrictamente cóncava, ni estrictamente convexa

Cuasiconcavidad y cuasiconvexidad

Primero, observemos que:

- Toda función *cóncava* es *cuasicóncava*
- Toda función *convexa* es *cuasiconvexa*

Caracterización de cuasiconvexidad y cuasiconcavidad

- f es *cuasiconvexa* si y solo si todos sus *subniveles* $L_\alpha(f)$ son conjuntos convexos
- f es *cuasicóncava* si y solo si todos sus *superniveles* $U^\alpha(f)$ son conjuntos convexos

Interpretación en curvas de nivel
Recordemos que las *curvas de nivel* de f para un valor α son los conjuntos

$$\{x \in^n : f(x) = \alpha\}$$

Los conjuntos de nivel (subniveles y superniveles) son las regiones delimitadas por esas curvas:

$$L_\alpha(f) = \{x \in^n : f(x) \leq \alpha\}$$
$$U^\alpha(f) = \{x \in^n : f(x) \geq \alpha\}$$

Criterio de hessiano orlado para cuasiconvexidad y cuasiconcavidad

Las siguientes condiciones son válidas cuando evaluamos las funciones en números positivos (es decir en el ortante no negativo):
Sea $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ de clase C^2 en un dominio convexo U, y supongamos $\nabla f(x) \neq 0$ para todo $x \in U$. Definimos la *matriz hessiano orlado* en x como

$$B(x) = \begin{pmatrix} 0 & \nabla f(x)^T \\ \nabla f(x) & H_f(x) \end{pmatrix}$$

donde $H_f(x)$ es la Hessiana usual de f
Entonces se tiene la caracterización siguiente:
Condiciones necesarias:

- f es cuasiconvexa en U solo si los menores principales de orden inicial (sin contar el 0) son todos menores o iguales a 0
- f es cuasicóncava en U solo si los menores principales de orden inicial (sin contar el 0) alternan de signo (o son iguales a 0) empezando por el negativo

Condiciones suficientes y que además aseguran cuasiconcavidad y cuasiconvexidad estricta:

- f es cuasiconvexa en U si los menores principales de orden inicial (sin contar el 0) son todos menores a 0
- f es cuasicóncava en U si los menores principales de orden inicial (sin contar el 0) alternan de signo empezando por el negativo

Para que $z = f(x_1, \ldots, x_n)$ sea cuasicóncava en el n-ortante no negativo, es necesario que

$$|B_1| \leq 0, \quad |B_2| \geq 0, \quad \ldots, \quad |B_n| \begin{cases} \leq 0 & n \text{ impar} \\ \geq 0 & n \text{ par} \end{cases}$$

Una condición suficiente para que f sea *cuasicóncava* en el n-ortante no negativo es que

$$|B_1| < 0, \quad |B_2| > 0, \quad \ldots, \quad |B_n| \begin{cases} < 0 & n \text{ impar} \\ > 0 & n \text{ par} \end{cases}$$

Ejemplo

Sea

$$f: A \to, \qquad A = \{(x, y) \in^2 : x > 0,\ y > 0\}, \quad f(x, y) = x\,y$$

Derivadas parciales y Hessiano

$$f_x = y, \quad f_y = x, \qquad H_f(x, y) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Hessiano orlado
Definimos las siguientes matrices:

$$B_1 = \begin{pmatrix} 0 & y \\ y & 0 \end{pmatrix} \qquad B_2 = \begin{pmatrix} 0 & y & x \\ y & 0 & 1 \\ x & 1 & 0 \end{pmatrix}$$

Cálculo de determinantes

$$\det B_1 = -\,y^2 < 0 \qquad \det B_2 = 2\,x\,y \ > \ 0$$

Conclusión
Para cuasiconcavidad en el dominio convexo A se exige

$$D_1 < 0, \quad D_2 \geq 0$$

Como $x, y > 0$ garantizan $D_1 < 0$ y $D_2 > 0$, concluimos que

$$f(x, y) = x\,y \text{ es cuasicóncava en } A$$

Relación y diferencias entre cuasiconcavidad y concavidad (o cuasiconvexidad y convexidad)

- **Relación entre ambas:**
  - Toda función *cóncava* es *cuasicóncava*
  - La recíproca no es cierta: hay funciones cuasicóncavas que no son cóncavas
  - Toda función *convexa* es *cuasiconvexa*
  - La recíproca no es cierta: hay funciones cuasiconvexas que no son convexas
- **Operaciones y composición:**
  - Suma y combinación lineal (con coeficientes no negativos) de cóncavas/convexas produce una función cóncava/convexa
  - La suma de funciones cuasicóncavas/cuasiconvexas no siempre es cuasicóncava/cuasiconvexa
  - La cuasiconcavidad/cuasiconvexidad se conserva bajo composición con funciones estrictamente crecientes
  - Además, sea
    $$h(x) = (g \circ f)(x)$$
    con
    $$f \colon M \subseteq \mathbb{R}^n \to \mathbb{R}, \quad g \colon I \subseteq \mathbb{R} \to \mathbb{R}$$
    donde M es convexo e I es un intervalo. Entonces:
    1. Si f es convexa y g es creciente y convexa, entonces h es convexa
    2. Si f es cóncava y g es creciente y cóncava, entonces h es cóncava

Ejemplo

Anteriormente probamos que $f(x, y) = xy$ es cuasicóncava. Veamos una transformación de dicha función: sea

$$g(x, y) = \ln(xy) = \ln x + \ln y$$

$$g_x = \frac{1}{x}, \quad g_y = \frac{1}{y}, \qquad g_{xx} = -\frac{1}{x^2},\ g_{xy} = 0,\ g_{yy} = -\frac{1}{y^2}$$

$$B_1(x, y) = \begin{pmatrix} 0 & g_x \\ g_x & g_{xx} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{x} \\ \frac{1}{x} & -\frac{1}{x^2} \end{pmatrix}$$

$$D_1(x, y) = \det B_1 = -\frac{1}{x^2} < 0$$

$$B_2(x,y) = \begin{pmatrix} 0 & g_x & g_y \\ g_x & g_{xx} & g_{xy} \\ g_y & g_{xy} & g_{yy} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{x} & \frac{1}{y} \\ \frac{1}{x} & -\frac{1}{x^2} & 0 \\ \frac{1}{y} & 0 & -\frac{1}{y^2} \end{pmatrix}$$

$$D_2(x,y) = \det B_2 = \frac{2}{x^2\, y^2} > 0$$

Conclusión

$$D_1 < 0, \quad D_2 > 0$$

Por tanto $g(x,y) = \ln(xy)$ es *cuasicóncava* en A

15 Optimización sin restricciones

Optimización sin restricciones: cálculo univariable

Sea $f\colon I \subseteq \mathbb{R} \to \mathbb{R}$ una función de clase C^2. Para encontrar extremos locales de f aplicamos el siguiente procedimiento:

1. **Condición de primer orden:** Calcular la derivada primera y resolver

$$f'(x^*) = 0$$

   Los puntos x^* que satisfacen esta ecuación son los *puntos críticos*.

2. **Condición de segundo orden:** Evaluar la derivada segunda en cada x^*:

$$f''(x^*) \begin{cases} > 0 & \Rightarrow \; x^* \text{ es un mínimo local} \\ < 0 & \Rightarrow \; x^* \text{ es un máximo local} \\ = 0 & \Rightarrow \; \text{criterio inconcluso (punto de inflexión posible)} \end{cases}$$

Ejemplo rápido: Para $f(x) = x^3 - 3x$:

$$f'(x) = 3x^2 - 3,\; f'(x) = 0 \Rightarrow x = \pm 1; \quad f''(x) = 6x.$$

$$f''(1) = 6 > 0 \;\Rightarrow\; x = 1 \text{ mínimo local}, \qquad f''(-1) = -6 < 0 \;\Rightarrow\; x = -1 \text{ máximo local.}$$

Pasando a n variables

Denotemos

$$x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$$

y sea

$$x^* = (x_1^*, x_2^*, \ldots, x_n^*)$$

un punto crítico en $\mathbb{R}^n$.

$$F\colon A \subseteq \mathbb{R}^n \;\to\; \mathbb{R}$$

la *condición de primer orden* generaliza a

$$\nabla F(x^*) = \mathbf{0}$$

es decir, todas las derivadas parciales en x^* deben anularse.

La *condición de segundo orden* se extrae de la matriz Hessiana $H = \nabla^2 F(x)$, que reúne todas las derivadas segundas de F. De la misma manera que $f''(x^*)$ en una variable nos indica concavidad o convexidad, la semidefinición positiva o negativa de H nos permitirá determinar si F tiene un mínimo o máximo local en x^*.

Condiciones suficientes y criterio de la matriz Hessiana

Sea

$$F\colon U \subseteq \mathbb{R}^n \;\longrightarrow\; \mathbb{R}$$

de clase C^2 y sea $x^* \in U$ un punto crítico, es decir,

$$\nabla F(x^*) = \mathbf{0}.$$

- **Óptimo local (criterio de la Hessiana).**

- Si $H(x^*)$ es *definida negativa*, entonces x^* es *máximo local.*
- Si $H(x^*)$ es *definida positiva*, entonces x^* es *mínimo local.*
- Si $H(x^*)$ es *indefinida*, entonces x^* no es ni mínimo ni máximo local (punto silla).

- **Óptimo global (funciones convexas o cóncavas).**
  - Si F es convexa en U y $\nabla F(x^*) = 0$, entonces x^* es *mínimo global.*
  - Si F es cóncava en U y $\nabla F(x^*) = 0$, entonces x^* es *máximo global.*

Criterio de los menores principales para óptimo local.

Sea

$$F: U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$$

de clase C^2 y supongamos que $x^* \in U$ es un punto crítico:

$$\nabla F(x^*) = \mathbf{0}$$

Denotemos Δ_k por el menor principal inicial de orden k de la matriz Hessiana evaluada en el punto crítico.

Condición para máximo local. Si los menores principales alternan de signo comenzando negativo:

$$\Delta_1 < 0, \quad \Delta_2 > 0, \quad \Delta_3 < 0, \ \ldots, \ (-1)^n \Delta_n > 0$$

entonces x^* es un *máximo local* de F.

Condición para mínimo local . Si todos los menores principales son positivos:

$$\Delta_1 > 0, \quad \Delta_2 > 0, \quad \Delta_3 > 0, \ \ldots, \ \Delta_n > 0$$

entonces x^* es un *mínimo local* de F.

Criterio de los menores principales para óptimo global

Mínimo global. Para que F tenga un *mínimo global* en U, es suficiente que la matriz Hessiana H sea semidefinida positiva en *todos* los puntos $x \in U$. **En términos de menores principales, esto equivale a que para *cualquier* submatriz principal (es decir, para cualquier selección de filas y columnas con los mismos índices) el determinante sea mayor o igual a cero.** Cuando esto se cumple, F es convexa y cualquier punto crítico es automáticamente un mínimo global.

Máximo global. Análogamente, F tiene un *máximo global* en U si la Hessiana H es semidefinida negativa en todo U. **En lenguaje de menores principales, esto significa que para cada submatriz principal de orden k, el determinante de dicha submatriz multiplicado por $(-1)^k$ es mayor o igual a cero.** Bajo esta condición, F es cóncava y cada punto crítico será un máximo global.

Diferencia entre óptimos locales y globales

La principal diferencia entre óptimos que son locales y óptimos globales es que al momento de realizar el cálculo de segundas derivadas (con la matriz hessiana), para asegurar un óptimo local tenemos que evaluar el hessiano en un punto particular, de tal forma de mostrar la función tiene un comportamiento similar al de una función estrictamente convexa/cóncava en la vecindad inmediata de x^*. Por otro lado, para poder confirmar que un óptimo es global es necesario analizar la forma de la función, analizando el hessiano sin evaluar en ningún punto particular y ver si la función es convexa o cóncava.

Ejemplo: $F(x,y) = x^3 - y^3 + 9xy$

1. Derivadas parciales y puntos críticos

Definimos

$$F(x,y) = x^3 - y^3 + 9\,x\,y$$

Calculamos las derivadas parciales:

$$F_x = \frac{\partial F}{\partial x} = 3x^2 + 9y \qquad F_y = \frac{\partial F}{\partial y} = -3y^2 + 9x$$

Para hallar los puntos críticos resolvemos

$$\begin{cases} 3x^2 + 9y = 0 \\ -3y^2 + 9x = 0 \end{cases}$$

De la primera ecuación:

$$y = -\frac{x^2}{3}$$

Sustituyendo en la segunda:

$$-3\Big(-\frac{x^2}{3}\Big)^2 + 9x = 0 \implies -3\,\frac{x^4}{9} + 9x = 0 \implies -\frac{x^4}{3} + 9x = 0 \implies -x^4 + 27x = 0 \implies x\,(-x^3 + 27) = 0$$

De aquí obtenemos

$$x = 0 \quad \implies \quad y = 0, \qquad -x^3 + 27 = 0 \implies x^3 = 27 \implies x = 3 \implies y = -\frac{3^2}{3} = -3$$

Por tanto, los puntos críticos son

$$(0,0) \quad \text{y} \quad (3,-3)$$

2. Hessiano y clasificación de extremos

El Hessiano es

$$\nabla^2 F(x,y) = \begin{pmatrix} F_{xx} & F_{xy} \\ F_{yx} & F_{yy} \end{pmatrix} = \begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix}$$

1. **En** $(0,0)$**:**

$$\nabla^2 F(0,0) = \begin{pmatrix} 0 & 9 \\ 9 & 0 \end{pmatrix}, \quad \Delta_1 = 0, \quad \Delta_2 = -81 < 0$$

   Como $\Delta_2 < 0$, es un punto silla (saddle point).

2. **En** $(3,-3)$**:**

$$\nabla^2 F(3,-3) = \begin{pmatrix} 18 & 9 \\ 9 & 18 \end{pmatrix}, \quad \Delta_1 = 18 > 0, \quad \Delta_2 = 18 \cdot 18 - 9 \cdot 9 = 243 > 0$$

   Aquí la Hessiana es definida positiva, por lo que $(3,-3)$ es un mínimo local.

Conclusión:

- $(0,0)$ es un punto silla.
- $(3,-3)$ es un mínimo local de F.
- No hay máximos locales.

3. Concavidad/convexidad global de F

Para determinar si los extremos locales son también globales necesitamos ver si F es convexa o cóncava en todo ^{2}. Calculamos de nuevo el Hessiano general:

$$\nabla^2 F(x,y) = \begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix}$$

Los menores principales siguen un patrón ya que por ejemplo $6x$, puede ser negativo o positivo. Lo mismo con $-6y$ y $|H|$. No podemos afirmar que tenemos un óptimo global.

Ejemplo: $f(x,y) = x^2 + y^2$

1. Derivadas parciales y punto crítico

Sea

$$f(x,y) = x^2 + y^2, \qquad (x,y) \in^2$$

Calculamos las derivadas parciales:

$$f_x = 2x, \quad f_y = 2y$$

El único punto crítico se obtiene resolviendo

$$2x = 0, \quad 2y = 0 \implies (x,y) = (0,0)$$

2. Hessiano y clasificación de extremos locales

El Hessiano es constante:

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

Sus menores principales son

$$\Delta_1^{1,1} = 2 > 0, \qquad \Delta_1^{2,2} = 2 > 0, \qquad \Delta_2 = \det \nabla^2 f = 4 > 0$$

de modo que la matriz es *definida positiva* en todo ^{2}. Por el criterio de segundo orden, $(0,0)$ es un *mínimo local.*

3. Convexidad global y mínimo global

Ya que $\nabla^2 f(x,y) \geq 0$ en todo ^{2}, f es *convexa* en ^{2}. En toda función convexa cualquier mínimo local es también *mínimo global.*

Óptimo local y global únicos

Teorema: Sea $f : D \to \mathbb{R}$ con $D \subseteq \mathbb{R}^n$, abierto y convexo, $f \in C^1(D)$ y sea $\mathbf{x}_0 \in D$ un punto crítico de f. Se verifica:

1. Si f es convexa en D entonces f presenta en $\mathbf{x}_0$ un mínimo global.
2. Si f es estrictamente convexa en D entonces f presenta en $\mathbf{x}_0$ un mínimo global único.
3. Si f es cóncava en D entonces f presenta en $\mathbf{x}_0$ un máximo global.
4. Si f es estrictamente cóncava en D entonces f presenta en $\mathbf{x}_0$ un máximo global único.

Preservación de máximos y mínimos locales en dos variables

Sea $D \subseteq \mathbb{R}$ un conjunto abierto. Sea $\varphi : D \to \mathbb{R}$ una función de clase C^2 tal que $\varphi'(t) > 0$ para todo $t \in D$. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ una función de clase C^2 tal que su imagen está contenida en D (es decir, $f(x, y) \in D$ para todo (x, y) en el dominio de f o en la región de interés). Entonces:

1. (x_0, y_0) es punto crítico de f entonces (x_0, y_0) es punto crítico de $\varphi \circ f$.
2. Si la matriz Hessiana $D^2 f(x_0, y_0)$ es definida negativa (máximo local) o definida positiva (mínimo local), la Hessiana de $\varphi \circ f$ en (x_0, y_0) mantiene la misma definitud.

Ejemplos de φ

- $\varphi(t) = a\,t + b$, con $a > 0$.
- $\varphi(t) = \mathrm{e}^t$.
- $\varphi(t) = \log(t + c)$, con $c > 0$ (dominio $t > -c$).
- $\varphi(t) = t^3$ (con $t \neq 0$).
- $\varphi(t) = \sqrt{t}$ (en $t > 0$).

16 Aplicaciones económicas de optimización

Teorema de la envolvente (optimización sin restricciones)

Consideremos el problema de maximización sin restricciones

$$\max_{x,y} U = f(x, y, \phi)$$

donde x, y son variables endógenas y ϕ es un parámetro exógeno.

Condiciones de óptimo

Las condiciones necesarias de primer orden en el óptimo $(x^*(\phi), y^*(\phi))$ son

$$\frac{\partial f}{\partial x}(x^*(\phi), y^*(\phi), \phi) = 0, \qquad \frac{\partial f}{\partial y}(x^*(\phi), y^*(\phi), \phi) = 0. \tag{1}$$

Bajo condiciones de segundo orden adecuadas, estas ecuaciones definen implícitamente $x^*(\phi)$ y $y^*(\phi)$.

$$x^* = x^*(\phi) \qquad y^* = y^*(\phi)$$

Función de valor indirecta

Sustituyendo en la función objetivo obtenemos la función de valor máximo (función objetivo indirecta)

$$V(\phi) = f(x^*(\phi), y^*(\phi), \phi)$$

Esta es una función que depende en última instancia de ϕ, a diferencia de la función objetivo que depende de x y de y.

Derivada de la función de valor

Al diferenciar V con respecto a ϕ usando la regla de la cadena:

$$\frac{dV}{d\phi} = f_x \frac{\partial x^*}{\partial \phi} + f_y \frac{\partial y^*}{\partial \phi} + f_\phi$$

donde todos los términos f_x, f_y, f_ϕ se evalúan en $(x^*(\phi), y^*(\phi), \phi)$.

Pero por las condiciones de primer orden (1) se tiene $f_x = f_y = 0$ en el óptimo, y por tanto

$$\frac{dV}{d\phi} = f_\phi(x^*(\phi), y^*(\phi), \phi)$$

Este es el **teorema de la envolvente**: la derivada de la función de valor máximo con respecto al parámetro ϕ equivale al efecto directo de ϕ sobre la función objetivo, ignorando los efectos indirectos vía $x^*(\phi)$ y $y^*(\phi)$.

Interpretación

- La función de valor $V(\phi)$ "envuelve" la familia de funciones $f(x, y, \phi)$ optimizadas en $(x^*(\phi), y^*(\phi))$ al variar ϕ.
- El resultado muestra que, en el óptimo, no hace falta calcular $\partial x^*/\partial\phi$ ni $\partial y^*/\partial\phi$ para conocer $\frac{dV}{d\phi}$: basta con el efecto directo f_ϕ.

Intuición: ¿Por qué "teorema de la envolvente"?

El nombre proviene de la idea geométrica de que la función de valor indirecta

$$V(\phi) = \max_{x,y} f(x, y, \phi)$$

es la *envolvente* de la familia de gráficas $\{ z = f(x, y, \phi) : (x, y) \in^2 \}$ al variar el parámetro ϕ. A continuación presentamos varias perspectivas:

Envolvente de familias de curvas/superficies

Para cada ϕ se tiene un punto en el plano (ϕ, z):

$$z = f(x^*(\phi), y^*(\phi), \phi)$$

Estos puntos conforman una curva al variar ϕ. La curva límite que toca tangencialmente a cada una de estas es la *envolvente* Gráficamente, $V(\phi)$ "abraza" o "envuelve" el punto más alto de cada miembro de la familia

Ejemplo de envolvente con $f(x; t) = t\,x^2 + x + 10$

Consideremos la familia de funciones

$$f(x; t) = t\,x^2 + x + 10 \qquad x \in,\; t < 0$$

donde x es la variable de decisión y t un parámetro exógeno. Tomamos $t < 0$ para garantizar que la maximización en x sea bien comportada (coeficiente de x^2 negativo)

Función valor

Definimos la función de valor

$$V(t) = \max_{x\in} f(x; t)$$

Condición de primer orden

Para cada $t < 0$, el óptimo $x^*(t)$ satisface

$$\frac{\partial f}{\partial x} = 2\,t\,x + 1 = 0 \quad \Longrightarrow \quad x^*(t) = -\frac{1}{2t}$$

Cálculo de la envolvente

Sustituyendo $x^*(t)$ en f:

$$V(t) = f(x^*(t); t) = t\Big(-\frac{1}{2t}\Big)^2 + \Big(-\frac{1}{2t}\Big) + 10 = \frac{1}{4t} - \frac{1}{2t} + 10 = 10 - \frac{1}{4t}$$

Teorema de la envolvente

Directamente,

$$V'(t) = \frac{d}{dt}\Big(10 - \frac{1}{4t}\Big) = \frac{1}{4t^2}$$

Por el teorema de la envolvente, también debe cumplirse

$$V'(t) = \frac{\partial f}{\partial t}(x^*(t); t) = x^*(t)^2 = \Big(-\frac{1}{2t}\Big)^2 = \frac{1}{4t^2}$$

confirmando que los efectos indirectos vía $x^*(t)$ se anulan en la derivada de la función valor

Interpretación

- Cada f es una parábola cóncava en x (porque $t < 0$)
- La envolvente $V(t) = 10 - \dfrac{1}{4t}$ toca tangencialmente a cada parábola en $x = x^*(t)$
- El teorema de la envolvente nos dice que para medir el impacto de t sobre el máximo, basta con $\partial f/\partial t$ evaluada en el óptimo

Para obtener la envolvente como función de la *variable* x, invertimos esta relación:

$$x = -\frac{1}{2t} \quad \Longrightarrow \quad t = -\frac{1}{2x} \qquad x > 0$$

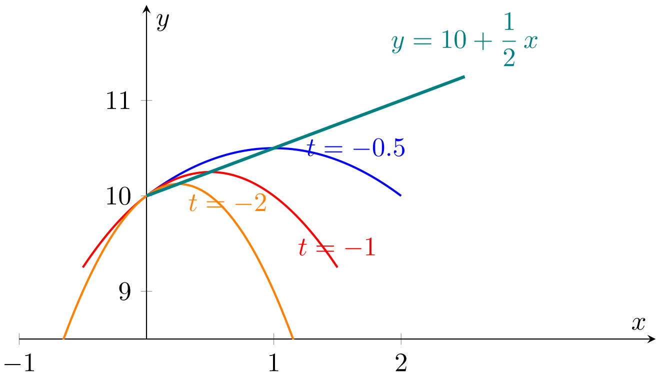
Sustituyendo $t(x)$ en f hallamos

$$V(x) = f(x;\ t(x)) = \left(-\frac{1}{2x}\right)x^2 + x + 10 = -\frac{x}{2} + x + 10 = \frac{x}{2} + 10$$

Por tanto la envolvente viene dada por

$$y = V(x) = 10 + \frac{1}{2}x \qquad x > 0$$

Gráficamente esta función toca todos los puntos máximos de la familia de parábolas: $f(x,t) = tx^2 + x + 10$



Ejemplo económico: maximización de beneficios sin restricciones

Consideremos una empresa que produce una cantidad $q \geq 0$ de un bien y enfrenta un precio de mercado $p > 0$. Su beneficio π viene dado por

$$\pi(q;p) \ = \ p\,q \ - \ C(q)$$

donde $C(q)$ es su función de costes, asumida de clase C^2 y estrictamente convexa ($C''(q) > 0$)

Problema de optimización

Para cada precio p la empresa elige q para

$$\max_{q \geq 0} \ \pi(q;p)$$

Las condiciones de primer y segundo orden son:

$$\frac{\partial \pi}{\partial q} = p - C'(q) = 0 \quad \Longrightarrow \quad C'(q^*) = p$$

$$\frac{\partial^2 \pi}{\partial q^2} = -\,C''(q^*) < 0 \quad \Longrightarrow \quad \pi \text{ es cóncava en } q \text{ y } q^*(p) \text{ es el único máximo}$$

Función de beneficio indirecto

Definimos la función valor (beneficio máximo) como

$$\Pi(p) \;=\; \pi(q^*(p);p) = p\,q^*(p) - C(q^*(p))$$

Teorema de la envolvente

Al derivar $\Pi(p)$ con respecto a p, aplicando el teorema de la envolvente se obtiene

$$\frac{d\Pi}{dp} = \frac{\partial \pi}{\partial p}\Big(q^*(p);p\Big) = q^*(p)$$

sin necesidad de calcular $\dfrac{dq^*}{dp}$, pues $\partial\pi/\partial q = 0$ en el óptimo

Función concreta

Supongamos $C(q) = \dfrac{1}{2}\,c\,q^2$ con $c > 0$. Entonces

$$C'(q) = c\,q \quad C''(q) = c > 0$$

y la condición de primer orden $p = c\,q$ da

$$q^*(p) = \frac{p}{c} \qquad \Pi(p) = p\cdot\frac{p}{c} - \frac{1}{2}\,c(\frac{p}{c})^2 = \frac{p^2}{2c}$$

Verifiquemos la envolvente:

$$\frac{d\Pi}{dp} = \frac{d}{dp}\Big(\frac{p^2}{2c}\Big) = \frac{p}{c} = q^*(p)$$

Interpretación

- La condición de envolvente $\partial\pi/\partial q = 0$ nos da la oferta óptima $q^*(p)$
- El teorema de la envolvente permite calcular la sensibilidad del beneficio máximo al precio p simplemente con $\partial\pi/\partial p = q$, sin derivar la curva de oferta $q^*(p)$. En este caso el resultado nos sugiere que ante un aumento del precio de venta, el beneficio máximo se incrementa.

Otro ejemplo económico: minimización de costes sujeto a restricción.

Consideremos una empresa cuya tecnología de producción viene dada por la función Cobb–Douglas

$$y = F(L,K) = L^{\alpha}K^{1-\alpha} \qquad \alpha \in (0,1)$$

donde L y K son trabajo y capital, y $y > 0$ es el nivel de output

Problema de minimización de costes

Dados los precios de los factores w (salario) y r (renta del capital) y el nivel y, la empresa resuelve

$$\min_{L,K} \;\; C = w\,L \;+\; r\,K \quad \text{sujeto a} \quad L^{\alpha}K^{1-\alpha} = y$$

Demanda condicionada y coste mínimo

Escribimos el Lagrangiano

$$\mathcal{L} = wL + rK + \mu(y - L^\alpha K^{1-\alpha})$$

Las condiciones de primer orden son

$$\frac{\partial \mathcal{L}}{\partial L} = w - \mu\,\alpha\,L^{\alpha-1}K^{1-\alpha} = 0 \qquad \frac{\partial \mathcal{L}}{\partial K} = r - \mu\,(1-\alpha)\,L^{\alpha}K^{-\alpha} = 0$$

junto con la restricción $L^\alpha K^{1-\alpha} = y$. De las dos primeras:

$$\frac{w}{\alpha}L^{1-\alpha}K^{\alpha-1} = \frac{r}{1-\alpha}L^{-\alpha}K^{\alpha} \quad \Longrightarrow \quad \frac{L^*}{K^*} = \frac{\alpha}{1-\alpha}\,\frac{r}{w}$$

Sustituyendo en $L^\alpha K^{1-\alpha} = y$ y resolviendo se obtienen las demandas condicionadas

$$L^*(w,r,y) = y\left(\frac{\alpha}{w}\right)^{\alpha}\left(\frac{1-\alpha}{r}\right)^{1-\alpha}\alpha \quad K^*(w,r,y) = y\left(\frac{\alpha}{w}\right)^{\alpha}\left(\frac{1-\alpha}{r}\right)^{1-\alpha}(1-\alpha)$$

El coste mínimo resultante es

$$C(w,r,y) = w\,L^*(w,r,y) + r\,K^*(w,r,y) = y\left(\frac{w}{\alpha}\right)^{\alpha}\left(\frac{r}{1-\alpha}\right)^{1-\alpha}$$

Lema de Shephard

La función de coste $C(w,r,y)$ es la envolvente de la familia de funciones $wL + rK$ sujeta a $F(L,K) = y$ Por el teorema de la envolvente, al derivar C con respecto a w (tratando y, r constantes) basta con tomar el efecto directo:

$$\frac{\partial C}{\partial w} = L^*(w,r,y)$$

De hecho, si derivamos la expresión encontrada:

$$\frac{\partial}{\partial w}\left[y\,(w/\alpha)^{\alpha}(r/(1-\alpha))^{1-\alpha}\right] = y\,\alpha\,(w/\alpha)^{\alpha-1}(r/(1-\alpha))^{1-\alpha} = L^*(w,r,y)$$

Análogamente,

$$\frac{\partial C}{\partial r} = K^*(w,r,y)$$

Interpretación

- $C(w,r,y)$ es la curva de coste de largo plazo: "envuelve" todas las rectas $wL + rK$ tangentes a las iso-productivas $L^\alpha K^{1-\alpha} = y$
- La envolvente simplifica el cálculo de las *demandas condicionadas*: no necesitamos derivar $L^*(w,r,y)$ ni $K^*(w,r,y)$ respecto a w o r para obtener $\partial C/\partial w$, $\partial C/\partial r$
- **Este resultado se conoce como el Lema de Shephard**. El cual nos dice que una vez que tenemos la función de costo indirecta basta con derivar esta con respecto a los precios de los insumos para encontrar la demanda de insumos condicionada.

Otro ejemplo: Monopolista multiproducto con dos bienes

Sea un monopolista que ofrece dos bienes $q_1, q_2 \geq 0$ con demanda inversa lineal

$$p_1(q_1,q_2) = a - b\,q_1 - c\,q_2 \quad p_2(q_1,q_2) = d - e\,q_1 - f\,q_2$$

y costos lineales

$$C(q_1,q_2) = g\,q_1 + h\,q_2$$

Entonces el beneficio es

$$\pi(q_1,q_2) = p_1\,q_1 + p_2\,q_2 - C(q_1,q_2) = (a - bq_1 - cq_2)q_1 + (d - eq_1 - fq_2)q_2 - (gq_1 + hq_2)$$

Condiciones de primer orden

Para encontrar el máximo, derivamos la función de beneficio respecto a cada cantidad e igualamos a cero. **Es crucial incluir todos los términos cruzados.**

La derivada respecto a q_1 es:

$$\frac{\partial \pi}{\partial q_1} = (a - 2b\, q_1 - c\, q_2) + (-e\, q_2) - g = a - g - 2b\, q_1 - (c+e)q_2 = 0$$

La derivada respecto a q_2 es:

$$\frac{\partial \pi}{\partial q_2} = (-c\, q_1) + (d - e\, q_1 - 2f\, q_2) - h = d - h - (c+e)q_1 - 2f\, q_2 = 0$$

Esto nos da un sistema de dos ecuaciones lineales:

$$\begin{cases} 2b\, q_1 + (c+e)q_2 = a - g \\ (c+e)q_1 + 2f\, q_2 = d - h \end{cases}$$

Resolviendo este sistema (por ejemplo, usando la regla de Cramer o sustitución) se obtienen las cantidades óptimas:

$$q_1^* = \frac{2f\,(a-g) - (c+e)(d-h)}{4bf - (c+e)^2}$$

$$q_2^* = \frac{2b\,(d-h) - (c+e)(a-g)}{4bf - (c+e)^2}$$

Condición de segundo orden (Hessiano)

El Hessiano de π respecto a (q_1, q_2) es:

$$\nabla^2 \pi = \begin{pmatrix} \dfrac{\partial^2 \pi}{\partial q_1^2} & \dfrac{\partial^2 \pi}{\partial q_1 \partial q_2} \\ \dfrac{\partial^2 \pi}{\partial q_2 \partial q_1} & \dfrac{\partial^2 \pi}{\partial q_2^2} \end{pmatrix} = \begin{pmatrix} -2b & -(c+e) \\ -(c+e) & -2f \end{pmatrix}$$

Para que (q_1^*, q_2^*) sea un máximo local, este Hessiano debe ser **definido negativo**:

$$-2b < 0 \quad \Longrightarrow \quad b > 0$$

y su determinante debe ser positivo:

$$\det(\nabla^2 \pi) = (-2b)(-2f) - (-(c+e))(-(c+e)) = 4\,b\,f - (c+e)^2 > 0$$

Note que este denominador es el mismo que el utilizado para calcular q_1^* y q_2^*.

Sensibilidad respecto de b y d

Al aplicar el **teorema de la envolvente** al beneficio máximo $\Pi(b, d)$, basta con derivar la función de beneficio π respecto a los parámetros b y d, y luego evaluar en el punto óptimo (q_1^*, q_2^*).

$$\frac{\partial \pi}{\partial b} = -\, q_1^2 \quad \Longrightarrow \quad \frac{d\Pi}{db} = -\, (q_1^*)^2$$

$$\frac{\partial \pi}{\partial d} = q_2 \quad \Longrightarrow \quad \frac{d\Pi}{dd} = q_2^*$$

Interpretación

- Un aumento de b (la demanda del bien 1 es más sensible a su propia cantidad) **reduce el beneficio máximo**, ya que $\frac{d\Pi}{db} = -\left(q_1^*\right)^2 < 0$.
- Un aumento de d (mayor disposición a pagar por el bien 2) **incrementa el beneficio máximo**, ya que se asume que la cantidad óptima $\frac{d\Pi}{dd} = q_2^*$ será positiva.

17 Optimización con restricciones de igualdad

Optimización con restricciones y Lagrangiano

En muchos problemas de optimización no basta con minimizar (o maximizar) una función objetivo $f: \mathbb{R}^n \to \mathbb{R}$ libremente: es preciso hacerlo sujeto a ciertas restricciones que deben cumplirse. En este apartado vamos a revisar el método de los *multiplicadores de Lagrange* (o método del Lagrangiano) que es una herramienta fundamental para abordar problemas con restricciones de igualdad

Formulación del problema

Consideremos el problema

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{sujeto a} \quad g_i(x) = 0, \quad i = 1, \ldots, k$$

donde $f: \mathbb{R}^n \to \mathbb{R}$, $g_i: \mathbb{R}^n \to \mathbb{R}$ son funciones (C^1) y las curvas $g_i(x) = 0$ definen el conjunto factible.

El Lagrangiano

Definimos la función Lagrangiana asociada al problema como

$$\mathcal{L}(x, \lambda) = f(x) + \sum_{i=1}^{k} \lambda_i \, g_i(x)$$

El lagrangiano también puede formularse como:

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i=1}^{k} \lambda_i \, g_i(x)$$

Ya que el término λ es una variable artificial que puede tener signo negativo o positivo. El resultado es indiferente de si aparece sumando o restando (de hecho véase que la restricción $g_i(x) = 0$ puede multiplicarse ambos lados por -1 y sigue siendo válida).

Condiciones de primer orden

Un par (x^*, λ^*) que resuelve el problema debe satisfacer las siguientes condiciones, que incluyen

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) + \sum_{i=1}^{k} \lambda_i^* \, \nabla g_i(x^*) \; = \; 0$$

$$g_i(x^*) = 0, \quad i = 1, \ldots, k$$

Estas condiciones son las mismas para maximizar o minimizar un problema de optimización libre. Por lo que el método del lagrangiano nos proporciona una simplificación del problema a cambio de agregar un multiplicador lagrangiano por cada restricción.

Interpretación geométrica

Geométricamente, en el punto óptimo x^* el gradiente de la función objetivo debe poder expresarse como combinación lineal de los gradientes de las restricciones

$$\nabla f(x^*) \in \operatorname{span}\{\nabla g_1(x^*), \ldots, \nabla g_k(x^*)\}$$

Ejemplo: Minimización con restricción lineal

Consideremos el problema

$$\min_{x,y\in\mathbb{R}} f(x,y) = x^2 + y^2$$
$$\text{sujeto a} \quad g(x,y) = x + y - 1 = 0$$

El Lagrangiano asociado es

$$\mathcal{L}(x,y,\lambda) = x^2 + y^2 - \lambda\,(x + y - 1)$$

Las condiciones de primer orden son

$$\frac{\partial \mathcal{L}}{\partial x} = 2x - \lambda = 0 \quad , \quad \frac{\partial \mathcal{L}}{\partial y} = 2y - \lambda = 0 \quad , \quad \frac{\partial \mathcal{L}}{\partial \lambda} = -(x + y - 1) = 0$$

Es conveniente recordar que la condición $\frac{\partial \mathcal{L}}{\partial \lambda} = 0$ no es más que $g(x,y) = 0$.
De las dos primeras ecuaciones se obtiene

$$x = \frac{\lambda}{2} \quad , \quad y = \frac{\lambda}{2}$$

y sustituyendo en la restricción $x + y = 1$ resulta $\lambda = 1$ Por tanto

$$x^* = y^* = \frac{1}{2}$$

Este sería el candidato a óptimo ya que cumple las condiciones necesarias.

Otro ejemplo

Consideremos el problema

$$\max_{x,y\in\mathbb{R}} f(x,y) = x\,y$$
$$\text{sujeto a} \quad h(x,y) = x^2 + y^2 - 1 = 0$$

El Lagrangiano es

$$\mathcal{L}(x,y,\lambda) = x\,y - \lambda(x^2 + y^2 - 1)$$

Las condiciones de primer orden son

$$\frac{\partial \mathcal{L}}{\partial x} = y - 2\lambda\,x = 0 \tag{2}$$

$$\frac{\partial \mathcal{L}}{\partial y} = x - 2\lambda\,y = 0 \tag{3}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = -(x^2 + y^2 - 1) = 0 \tag{4}$$

De (2) y (3) se deduce

$$y = 2\lambda\,x \quad , \quad x = 2\lambda\,y \implies (2\lambda)^2 = 1 \implies \lambda = \pm\frac{1}{2}$$

Por la restricción $x^2 + y^2 = 1$ tenemos además

$$y = \pm x \quad , \quad x = \pm\frac{1}{\sqrt{2}} \quad , \quad y = \pm\frac{1}{\sqrt{2}}$$

Así, los puntos críticos son

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

Que serían candidatos a máximos y mínimos.

Ejemplo económico: Maximización de utilidad con restricción presupuestaria

Consideremos un consumidor que elige cantidades $x_1, x_2 > 0$ de dos bienes para maximizar su utilidad

$$u(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha} \quad , \quad \alpha \in (0,1)$$

sujeto a su restricción presupuestaria

$$p_1\, x_1 + p_2\, x_2 = M$$

donde $p_1, p_2 > 0$ son los precios y $M > 0$ su ingreso

El Lagrangiano es

$$\mathcal{L}(x_1, x_2, \lambda) = x_1^{\alpha} x_2^{1-\alpha} + \lambda(M - p_1 x_1 - p_2 x_2)$$

Las condiciones de primer orden son

$$\frac{\partial \mathcal{L}}{\partial x_1} = \alpha\, x_1^{\alpha-1} x_2^{1-\alpha} - \lambda\, p_1 = 0 \tag{5}$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = (1-\alpha)\, x_1^{\alpha} x_2^{-\alpha} - \lambda\, p_2 = 0 \tag{6}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = M - p_1 x_1 - p_2 x_2 = 0 \tag{7}$$

De (5) y (6) obtenemos

$$\frac{\alpha}{1-\alpha}\,\frac{x_2}{x_1} = \frac{p_1}{p_2} \implies x_2 = \frac{1-\alpha}{\alpha}\,\frac{p_1}{p_2}\,x_1$$

Sustituyendo en (7) y resolviendo, se llega a las demandas óptimas

$$x_1^* = \frac{\alpha\, M}{p_1} \quad , \quad x_2^* = \frac{(1-\alpha)\, M}{p_2}$$

Condiciones de segundo orden

Suponemos además que f y las g_i son de clase C^2 en un entorno de x^*.

Caso concreto: 2 variables independientes y una restricción

Sea

$$f :^2 \to, \quad g :^2 \to, \quad \mathcal{L}(x, y, \lambda) = f(x, y) + \lambda\, g(x, y).$$

En un punto estacionario (x^*, y^*, λ^*) se cumplen las condiciones de primer orden:

$$\nabla f(x^*, y^*) \;=\; \lambda^*\, \nabla g(x^*, y^*), \qquad g(x^*, y^*) = 0.$$

Definimos las segundas derivadas de la función Lagrangiana $\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda\, g(x, y)$ en el punto crítico (x^*, y^*, λ^*) como

$$\nabla^2_{(x,y,\lambda)} \mathcal{L}(x^*, y^*, \lambda^*) = \begin{pmatrix} \mathcal{L}_{xx} & \mathcal{L}_{xy} & \mathcal{L}_{x\lambda} \\ \mathcal{L}_{yx} & \mathcal{L}_{yy} & \mathcal{L}_{y\lambda} \\ \mathcal{L}_{\lambda x} & \mathcal{L}_{\lambda y} & \mathcal{L}_{\lambda\lambda} \end{pmatrix}_{(x^*, y^*, \lambda^*)} .$$

Aquí,

$$\begin{aligned} \mathcal{L}_{xx} &= f_{xx} + \lambda^*\, g_{xx}, & \mathcal{L}_{xy} &= f_{xy} + \lambda^*\, g_{xy}, \\ \mathcal{L}_{yx} &= f_{yx} + \lambda^*\, g_{yx}, & \mathcal{L}_{yy} &= f_{yy} + \lambda^*\, g_{yy}, \\ \mathcal{L}_{x\lambda} &= g_x, & \mathcal{L}_{y\lambda} &= g_y, \\ \mathcal{L}_{\lambda x} &= g_x, & \mathcal{L}_{\lambda y} &= g_y, \\ \mathcal{L}_{\lambda\lambda} &= 0 \end{aligned}$$

Por tanto, al evaluar:

$$\bar{H} = \nabla^2_{(x,y,\lambda)}\mathcal{L}(x^*, y^*, \lambda^*) = \begin{pmatrix} f_{xx} + \lambda^* g_{xx} & f_{xy} + \lambda^* g_{xy} & g_x \\ f_{yx} + \lambda^* g_{yx} & f_{yy} + \lambda^* g_{yy} & g_y \\ g_x & g_y & 0 \end{pmatrix}_{(x^*,y^*,\lambda^*)}.$$

La submatriz central de orden 2×2 es precisamente las derivadas del lagrangiano con respecto a las variables x, y, y los elementos g_x, g_y crean el "borde" que la convierte en el hessiano orlado.

Restricción lineal

Cuando las restricciones son lineales el hessiano orlado se simplifica aún más: Supongamos

$$g(x, y) = a\,x + b\,y + c$$

con a, b, c constantes. Entonces:

- $\nabla g(x, y) = (g_x, g_y) = (a, b)$: las primeras derivadas $g_x = a$, $g_y = b$ son constantes.
- Las segundas derivadas se anulan:

$$g_{xx} = \frac{\partial}{\partial x}(g_x) = 0, \quad g_{xy} = \frac{\partial}{\partial y}(g_x) = 0, \quad g_{yy} = \frac{\partial}{\partial y}(g_y) = 0$$

Por tanto,

$$\bar{H} = \nabla^2_{(x,y,\lambda)}\mathcal{L}(x^*, y^*, \lambda^*) = \begin{pmatrix} f_{xx} + \lambda^* g_{xx} & f_{xy} + \lambda^* g_{xy} & g_x \\ f_{yx} + \lambda^* g_{yx} & f_{yy} + \lambda^* g_{yy} & g_y \\ g_x & g_y & 0 \end{pmatrix}_{(x^*,y^*,\lambda^*)}.$$

los términos $\lambda\, g_{jk}$ desaparecen, y queda

$$\bar{H} = \nabla^2_{(x,y,\lambda)}\mathcal{L}(x^*, y^*, \lambda^*) = \begin{pmatrix} f_{xx} & f_{xy} & g_x \\ f_{yx} & f_{yy} & g_y \\ g_x & g_y & 0 \end{pmatrix}_{(x^*,y^*,\lambda^*)}.$$

Condiciones suficientes para máximo o mínimo

Sea en el punto crítico (x^*, y^*, λ^*) el hessiano orlado

$$\bar{H}(x^*, y^*, \lambda^*) = \begin{pmatrix} 0 & g_x & g_y \\ g_x & f_{xx} + \lambda^* g_{xx} & f_{xy} + \lambda^* g_{xy} \\ g_y & f_{yx} + \lambda^* g_{yx} & f_{yy} + \lambda^* g_{yy} \end{pmatrix}_{(x^*,y^*,\lambda^*)}.$$

Las condiciones suficientes para máximos y mínimos son:

$$\det \bar{H}(x^*, y^*, \lambda^*) > 0 \implies (x^*, y^*) \text{ es un máximo local sujeto a } g = 0$$

$$\det \bar{H}(x^*, y^*, \lambda^*) < 0 \implies (x^*, y^*) \text{ es un mínimo local sujeto a } g = 0$$

Condiciones suficientes: caso general

Sea

$$\begin{aligned} &\min_{x\in^n} && f(x), \\ &\text{sujeto a} && g_i(x) = 0, \quad i = 1, \dots, k \end{aligned}$$

o

$$\begin{aligned} &\max_{x\in^n} && f(x), \\ &\text{sujeto a} && g_i(x) = 0, \quad i = 1, \dots, k \end{aligned}$$

y sea el Lagrangiano

$$\mathcal{L}(x, \lambda) \; = \; f(x) \; + \; \sum_{i=1}^{k} \lambda_i \, g_i(x), \qquad \lambda = (\lambda_1, \dots, \lambda_k)$$

Denotemos

$$\nabla g(x) \; = \; \begin{pmatrix} \nabla g_1(x) & \cdots & \nabla g_k(x) \end{pmatrix} \; \in \; ^{n\times k}$$

Suponiendo que $n > k$ (más variables independientes que restricciones). Entonces el *hessiano orlado* se define como la matriz $(k + n) \times (k + n)$

$$\bar{H}(x, \lambda) = \begin{pmatrix} 0_{k\times k} & \nabla g(x) \\ \nabla g(x)^T & \nabla^2_{xx}\mathcal{L}(x, \lambda) \end{pmatrix}$$

Este hessiano orlado tiene asociado $n + k$ menores principales de orden inicial pero solo basta con analizar estos $n - k$ últimos menores principales de orden inicial:

- **Mínimo local**: si los últimos menores principales de orden inicial (líderes) tienen el signo de $(-1)^k$. Entonces x^* es *mínimo* local sujeto a $g_i = 0$.
- **Máximo local**: Si los signos de los menores principales de orden inicial (líderes) se alternan terminando en el signo de $(-1)^n$ entonces x^* es *máximo* local sujeto a $g_i = 0$.

  Veamos ejemplos de esto:

En el caso de dos variables y una restricción, $k = 1$, $n = 2$ (analizado antes)

Solo tenemos que analizar $2 - 1 = 1$ menor principal de orden inicial:

- **Mínimo local**: si:

  $$|\bar{H}|$$

  Tiene el signo de $(-1)^k = (-1)^1$ Es decir:

  $$|\bar{H}| < 0$$

- **Máximo local**: si:

  $$|\bar{H}|$$

  Tiene el signo de $(-1)^n = (-1)^2$ Es decir:

  $$|H| > 0$$

En el caso de tres variables y una restricción, $k = 1$, $n = 3$

El hessiano orlado asociado es el siguiente:

$$\bar{H}(x^*, y^*, z^*, \lambda^*) = \begin{pmatrix} 0 & g_x & g_y & g_z \\ g_x & f_{xx} + \lambda^* g_{xx} & f_{xy} + \lambda^* g_{xy} & f_{xz} + \lambda^* g_{xz} \\ g_y & f_{yx} + \lambda^* g_{yx} & f_{yy} + \lambda^* g_{yy} & f_{yz} + \lambda^* g_{yz} \\ g_z & f_{zx} + \lambda^* g_{zx} & f_{zy} + \lambda^* g_{zy} & f_{zz} + \lambda^* g_{zz} \end{pmatrix}_{(x^*, y^*, z^*, \lambda^*)}$$

Solo tenemos que analizar $3-1=2$ menores principales de orden inicial. Sean Δ_1 y Δ_2 los últimos dos menores principales de orden inicial:

$$\Delta_1 = \det\begin{pmatrix} 0 & g_x & g_y \\ g_x & f_{xx}+\lambda^* g_{xx} & f_{xy}+\lambda^* g_{xy} \\ g_y & f_{yx}+\lambda^* g_{yx} & f_{yy}+\lambda^* g_{yy} \end{pmatrix}$$

$$\Delta_2 = \det\begin{pmatrix} 0 & g_x & g_y & g_z \\ g_x & f_{xx}+\lambda^* g_{xx} & f_{xy}+\lambda^* g_{xy} & f_{xz}+\lambda^* g_{xz} \\ g_y & f_{yx}+\lambda^* g_{yx} & f_{yy}+\lambda^* g_{yy} & f_{yz}+\lambda^* g_{yz} \\ g_z & f_{zx}+\lambda^* g_{zx} & f_{zy}+\lambda^* g_{zy} & f_{zz}+\lambda^* g_{zz} \end{pmatrix}_{(x^*,y^*,z^*,\lambda^*)}$$

- **Mínimo local**: si:

  $$\Delta_1, \Delta_2$$

  Tienen el mismo signo que $(-1)^k = (-1)^1$ es decir

  $$\Delta_1, \Delta_2 < 0$$

- **Máximo local**: si:

  $$\Delta_1, \Delta_2$$

  Alternan el signo terminando con el signo de $(-1)^n = (-1)^3$ Es decir:

  $$\Delta_1 > 0$$

  $$\Delta_2 < 0$$

En el caso de tres variables y dos restricciones, $k=2$, $n=3$

El hessiano orlado asociado en $(x^*, y^*, z^*, \lambda_1^*, \lambda_2^*)$ es

$$\bar{H}(x^*, y^*, z^*, \lambda_1^*, \lambda_2^*) = \begin{pmatrix} 0 & 0 & g_{1x} & g_{1y} & g_{1z} \\ 0 & 0 & g_{2x} & g_{2y} & g_{2z} \\ g_{1x} & g_{2x} & f_{xx}+\lambda_1^* g_{1xx}+\lambda_2^* g_{2xx} & f_{xy}+\lambda_1^* g_{1xy}+\lambda_2^* g_{2xy} & f_{xz}+\lambda_1^* g_{1xz}+\lambda_2^* g_{2xz} \\ g_{1y} & g_{2y} & f_{yx}+\lambda_1^* g_{1yx}+\lambda_2^* g_{2yx} & f_{yy}+\lambda_1^* g_{1yy}+\lambda_2^* g_{2yy} & f_{yz}+\lambda_1^* g_{1yz}+\lambda_2^* g_{2yz} \\ g_{1z} & g_{2z} & f_{zx}+\lambda_1^* g_{1zx}+\lambda_2^* g_{2zx} & f_{zy}+\lambda_1^* g_{1zy}+\lambda_2^* g_{2zy} & f_{zz}+\lambda_1^* g_{1zz}+\lambda_2^* g_{2zz} \end{pmatrix}$$

Solo tenemos que analizar $(3-2)=1$ menor principal de orden inicial es decir $|\bar{H}|$.

- **Mínimo local**: si:

  $$|\bar{H}|$$

  Tiene el signo de $(-1)^k = (-1)^2$ Es decir:

  $$|\bar{H}| > 0$$

- **Máximo local**: si:

  $$|\bar{H}|$$

  Tiene el signo de $(-1)^n = (-1)^3$ Es decir:

  $$|H| < 0$$

Condiciones para máximos y mínimos globales y únicos

Al igual que en optimización no restringida es necesario especificar si la función es cóncava o convexa para la existencia de un máximo o mínimo global, con la optimización restringida ocurre lo mismo. Sea

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

la función objetivo y sea

$$C = \{ x \in \mathbb{R}^n : h_j(x) = 0, \; j = 1, \ldots, p \}$$

el conjunto factible.

Condiciones para máximo global y único

Si f es estrictamente cuasicóncava y el conjunto factible C es convexo, entonces cualquier óptimo local del problema

$$\max_{x \in C} \; f(x)$$

es también óptimo global y único.

Condiciones para mínimo global y único

Si f es estrictamente cuasiconvexa y el conjunto factible C es convexo, entonces cualquier óptimo local del problema

$$\min_{x \in C} \; f(x)$$

es también óptimo global y único.

18 Optimización con restricciones de desigualdad

Problema con una sola variable independiente

Consideremos el problema

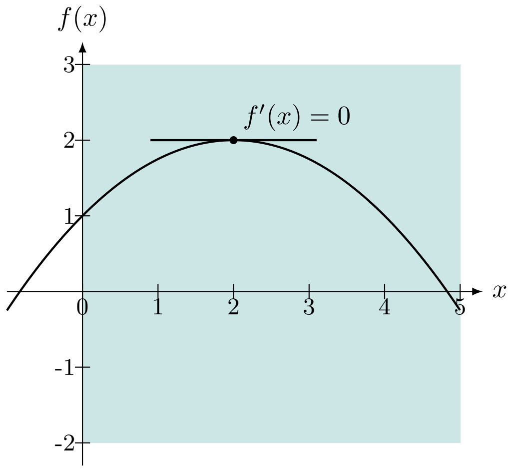
$$\max_{x \geq 0} f(x), \qquad f \in C^1(\mathbb{R}).$$

Un punto x^* es máximo local solo en una de estas dos situaciones:

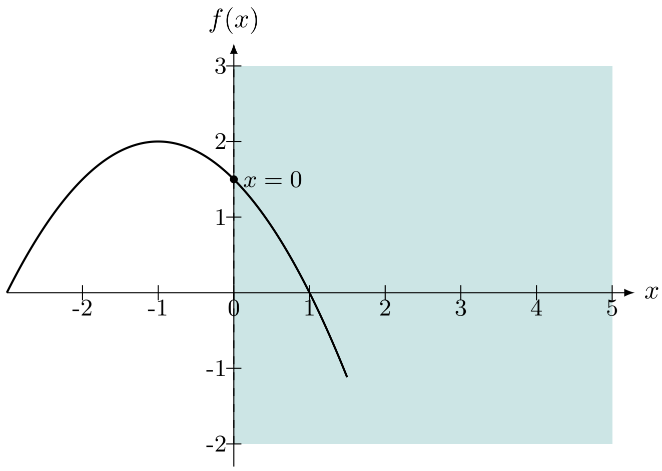
- **Interior:** $x^* > 0$ y $f'(x^*) = 0$. Aquí la restricción no interviene y se cumple la condición clásica de primer orden.
- **Frontera:** $x^* = 0$ y $f'(0) \leq 0$.

Obsérvese que $f'(0) > 0$ no puede corresponder a un máximo, pues implicaría pendiente creciente al inicio de la región factible.

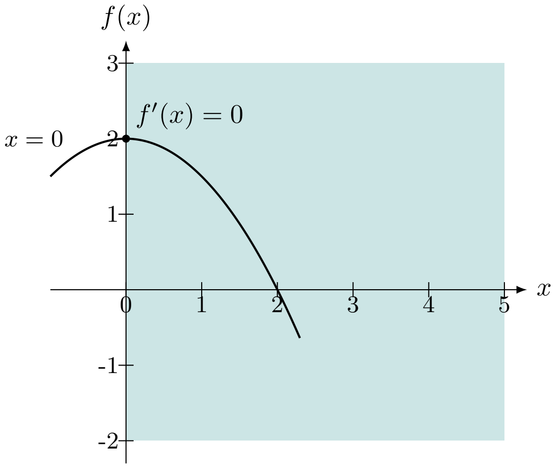
Gráficamente los casos son los siguientes: $f'(x) = 0$ por lo que la restricción no está activa.



Por otro lado $f'(x) < 0$ implica que la restricción está activa: $x = 0$.



Y por último también es posible que sucedan las dos cosas al mismo tiempo: $f'(x) = 0$ y $x = 0$.



Estas condiciones anteriores pueden resumirse:

$$f'(x) \leq 0, \quad x \geq 0, \quad x\,f'(x) = 0.$$

De esas condiciones se deduce que necesariamente

$$x = 0 \quad \text{o} \quad f'(x) = 0,$$

pues la última igualdad obliga a que al menos uno de los factores sea cero.

- $x \geq 0$:
- $f'(x) \leq 0$:
- $x\,f'(x) = 0$: *holgura complementaria.*

Minimización

Análogamente, para

$$\min_{x \geq 0} f(x),$$

las condiciones necesarias son

$$f'(x) \geq 0, \quad x \geq 0, \quad x\,f'(x) = 0,$$

Ejemplo

Consideremos el problema

$$\max_{x \geq 0} f(x), \quad f(x) = -x^2 - x - 1,$$

Caso 1: la restricción está inactiva (óptimo interior).
Si la restricción no actuase, existiría un máximo interior $x^* > 0$ y por lo tanto por la condición de holgura complementaria $xf'(x) = 0$ tenemos que:

$$f'(x^*) = 0.$$

Pero

$$f'(x) = -2x - 1,$$

luego

$$-2x^* - 1 = 0 \quad \Longrightarrow \quad x^* = -\frac{1}{2},$$

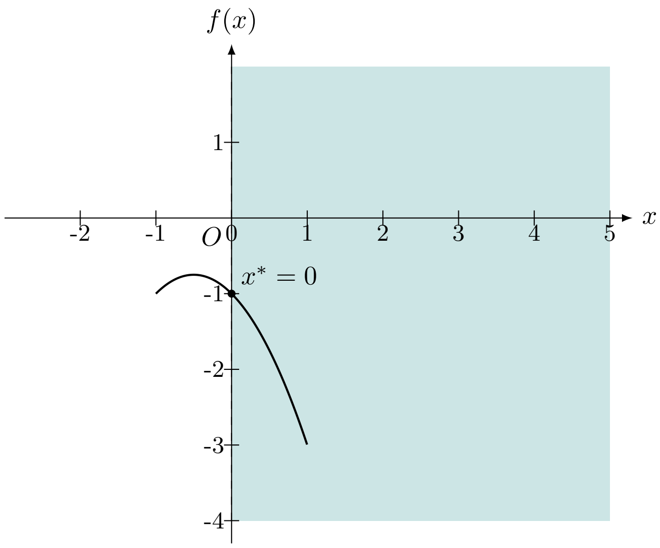
que contradice $x^* > 0$. Por tanto no puede haber óptimo interior en la región factible.

Caso 2: la restricción está activa (óptimo en la frontera).
La única posibilidad restante es que $x^* = 0$. Para verificar que es un máximo local basta comprobar la condición de primer orden en la frontera:

$$f'(0) = -2 \cdot 0 - 1 = -1 \leq 0.$$

Conclusión: En este ejemplo la restricción $x \geq 0$ está activa en el óptimo ($x^* = 0$), y el valor óptimo es $f(0) = -1 < 0$.

Gráficamente:



Problema con n variables independientes y restricciones de no negatividad

Consideremos

$$\max_{x \in \mathbb{R}^n_+} f(x_1, \ldots, x_n), \qquad f \in C^1(\mathbb{R}^n),$$

es decir, sujeto a $x_j \geq 0$ para $j = 1, \ldots, n$.

Las condiciones de primer orden adaptadas a estas restricciones son, para cada $j = 1, \ldots, n$:

$$\frac{\partial f}{\partial x_j} \leq 0, \qquad x_j \geq 0, \qquad x_j \frac{\partial f}{\partial x_j} = 0,$$

Problema con tres variables y dos restricciones de desigualdad

Planteamos primero el problema en forma de desigualdades:

$$\max_{x_1, x_2, x_3} f(x_1, x_2, x_3)$$

sujeto a

$$\begin{cases} g^1(x_1, x_2, x_3) \leq r_1, \\ g^2(x_1, x_2, x_3) \leq r_2, \\ x_1,\ x_2,\ x_3 \geq 0. \end{cases}$$

Al añadir las variables ficticias $s_1, s_2 \geq 0$ podemos convertir cada desigualdad en una igualdad:

$$\begin{cases} g^1(x_1, x_2, x_3) + s_1 = r_1, \\ g^2(x_1, x_2, x_3) + s_2 = r_2, \\ x_1,\ x_2,\ x_3,\ s_1,\ s_2 \geq 0. \end{cases}$$

Lagrangiano y condiciones de primer orden

A partir del problema con variables de holgura

$$\begin{aligned} \max_{x_1,x_2,x_3,s_1,s_2} \quad & f(x_1,x_2,x_3) \\ \text{sujeto a} \quad & g^1(x_1,x_2,x_3)+s_1=r_1, \\ & g^2(x_1,x_2,x_3)+s_2=r_2, \\ & x_j \geq 0 \ (j=1,2,3), \quad s_i \geq 0 \ (i=1,2), \end{aligned}$$

definimos el *Lagrangiano*

$$\mathcal{L}(x,s,\lambda) \ = \ f(x_1,x_2,x_3) \ + \ \lambda_1[r_1-g^1(x_1,x_2,x_3)-s_1] \ + \ \lambda_2[r_2-g^2(x_1,x_2,x_3)-s_2].$$

Aquí $\lambda_1,\lambda_2 \in \mathbb{R}$ son los multiplicadores asociados a las igualdades.

Condiciones de primer orden

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_1} &= f_1(x)-\lambda_1\, g_1^1(x)-\lambda_2\, g_1^2(x) \ \leq \ 0, \quad x_1 \ \geq \ 0, \quad x_1\,\frac{\partial \mathcal{L}}{\partial x_1}=0, \\ \frac{\partial \mathcal{L}}{\partial x_2} &= f_2(x)-\lambda_1\, g_2^1(x)-\lambda_2\, g_2^2(x) \ \leq \ 0, \quad x_2 \ \geq \ 0, \quad x_2\,\frac{\partial \mathcal{L}}{\partial x_2}=0, \\ \frac{\partial \mathcal{L}}{\partial x_3} &= f_3(x)-\lambda_1\, g_3^1(x)-\lambda_2\, g_3^2(x) \ \leq \ 0, \quad x_3 \ \geq \ 0, \quad x_3\,\frac{\partial \mathcal{L}}{\partial x_3}=0, \\ \frac{\partial \mathcal{L}}{\partial s_1} &= -\lambda_1 \ \leq \ 0, \quad s_1 \ \geq \ 0, \quad s_1\,\frac{\partial \mathcal{L}}{\partial s_1}=0, \\ \frac{\partial \mathcal{L}}{\partial s_2} &= -\lambda_2 \ \leq \ 0, \quad s_2 \ \geq \ 0, \quad s_2\,\frac{\partial \mathcal{L}}{\partial s_2}=0, \\ \frac{\partial \mathcal{L}}{\partial \lambda_1} &= r_1-g^1(x_1,x_2,x_3)-s_1=0, \\ \frac{\partial \mathcal{L}}{\partial \lambda_2} &= r_2-g^2(x_1,x_2,x_3)-s_2=0. \end{aligned}$$

Eliminación de las variables ficticias de las condiciones de primer orden

Tomemos la condición respecto a un multiplicador de lagrange:

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} \ = \ r_i-g^i(x_1,x_2,x_3)-s_i \ = \ 0,$$

se despeja inmediatamente

$$s_i \ = \ r_i-g^i(x_1,x_2,x_3), \qquad i=1,2.$$

Ahora sustituimos s_i en las tres condiciones

$$\frac{\partial \mathcal{L}}{\partial s_i} \leq 0, \quad s_i \geq 0, \quad s_i\,\frac{\partial \mathcal{L}}{\partial s_i}=0,$$

recordando que $\partial\mathcal{L}/\partial s_i=-\lambda_i$. Obtenemos:

$$\begin{aligned} -\lambda_i \leq 0 \quad &\Longrightarrow \quad \lambda_i \geq 0, \\ r_i-g^i(x_1,x_2,x_3) \geq 0 & \\ (r_i-g^i(x_1,x_2,x_3))\,(-\lambda_i)=0 \quad &\Longrightarrow \quad \lambda_i\,(r_i-g^i(x_1,x_2,x_3))=0. \end{aligned}$$

En consecuencia, las condiciones de Kuhn-Tucker quedan:

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial x_j} \;\leq\; 0, \\ x_j \;\geq\; 0, & (j = 1, 2, 3) \\ x_j \, \dfrac{\partial \mathcal{L}}{\partial x_j} \;=\; 0 \end{cases}$$

$$\begin{cases} r_i - g^i(x_1, x_2, x_3) \;\geq\; 0, \\ \lambda_i \;\geq\; 0, & (i = 1, 2) \\ \lambda_i \left[r_i - g^i(x_1, x_2, x_3) \right] \;=\; 0 \end{cases}$$

Véase que $r_i - g^i(x_1, x_2, x_3)$ coincide con la derivada del lagrangiano sin incluir las variables ficticias s_i. Entonces podemos expresar las condiciones de la siguiente manera:

$$\begin{array}{lll} \dfrac{\partial \mathcal{L}}{\partial x_j} \leq 0, & x_j \geq 0, & x_j \, \dfrac{\partial \mathcal{L}}{\partial x_j} = 0, \\ \dfrac{\partial \mathcal{L}}{\partial \lambda_i} \geq 0, & \lambda_i \geq 0, & \lambda_i \, \dfrac{\partial \mathcal{L}}{\partial \lambda_i} = 0, \end{array}$$

Resumen de las condiciones

Sea

$$L(x, \lambda) = f(x_1, \ldots, x_n) + \sum_{i=1}^{k} \lambda_i [r_i - g^i(x_1, \ldots, x_n)]$$

el Lagrangiano asociado a

$$\max_{x \in \mathbb{R}^n_+} f(x) \quad \text{sujeto a} \quad g^i(x) \leq r_i \; (i = 1, \ldots, k)$$

donde

$$\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_j \geq 0 \; \forall j = 1, \ldots, n\}$$

Las condiciones de Kuhn–Tucker para un máximo son

$$\begin{cases} \dfrac{\partial L}{\partial x_j} \leq 0, \quad x_j \geq 0, \quad x_j \, \dfrac{\partial L}{\partial x_j} = 0 \quad j = 1, \ldots, n \\ \dfrac{\partial L}{\partial \lambda_i} \geq 0, \quad \lambda_i \geq 0, \quad \lambda_i \, \dfrac{\partial L}{\partial \lambda_i} = 0 \quad i = 1, \ldots, k \end{cases}$$

Condiciones Kuhn-Tucker generales para un mínimo

Sea

$$L(x, \lambda) = f(x_1, \ldots, x_n) + \sum_{i=1}^{k} \lambda_i [r_i - g^i(x_1, \ldots, x_n)]$$

el Lagrangiano asociado a

$$\min_{x \in \mathbb{R}^n_+} f(x) \quad \text{sujeto a} \quad g^i(x) \geq r_i \; (i = 1, \ldots, k)$$

donde

$$\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_j \geq 0 \; \forall j = 1, \ldots, n\}$$

Las condiciones de Kuhn–Tucker necesarias para un mínimo son

$$\begin{cases} \dfrac{\partial L}{\partial x_j} \geq 0, \quad x_j \geq 0, \quad x_j \, \dfrac{\partial L}{\partial x_j} = 0 \quad j = 1, \ldots, n \\ \dfrac{\partial L}{\partial \lambda_i} \leq 0, \quad \lambda_i \geq 0, \quad \lambda_i \, \dfrac{\partial L}{\partial \lambda_i} = 0 \quad i = 1, \ldots, k \end{cases}$$

Ejemplo maximización

Consideremos

$$\max_{x,y} U(x,y) = xy \quad \text{sujeto a} \quad \begin{cases} x+y \le 100, \\ x \le 40, \\ x \ge 0,\ y \ge 0. \end{cases}$$

Asignamos multiplicadores $\lambda_1, \lambda_2 \ge 0$ a las dos primeras desigualdades y construimos el Lagrangiano

$$\mathcal{L}(x,y,\lambda_1,\lambda_2) = xy + \lambda_1(100-x-y) + \lambda_2(40-x).$$

Las condiciones de Kuhn-Tucker se agrupan en:

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x} &= y-\lambda_1-\lambda_2 \ \le 0, \quad x \ \ge 0, \quad x(y-\lambda_1-\lambda_2)=0, \\
\frac{\partial \mathcal{L}}{\partial y} &= x-\lambda_1 \ \le 0, \quad y \ \ge 0, \quad y(x-\lambda_1)=0, \\
\frac{\partial \mathcal{L}}{\partial \lambda_1} &= 100-x-y \ \ge 0, \quad \lambda_1 \ \ge 0, \quad \lambda_1(100-x-y)=0, \\
\frac{\partial \mathcal{L}}{\partial \lambda_2} &= 40-x \ \ge 0, \quad \lambda_2 \ \ge 0, \quad \lambda_2(40-x)=0.
\end{aligned}$$

Analizando las regiones:
En el óptimo $x^*, y^* > 0$ ya que de lo contrario la utilidad sería 0, entonces las dos primeras desigualdades se vuelven igualdades (por las condiciones de holgura complementaria)

$$y^* - \lambda_1 - \lambda_2 = 0, \quad x^* - \lambda_1 = 0.$$

Analizando las regiones, consideremos las cuatro combinaciones posibles de λ_1 y λ_2:

Caso 1: $\lambda_1 = 0,\ \lambda_2 = 0.$
Entonces

$$\frac{\partial \mathcal{L}}{\partial x} = y \le 0, \quad \frac{\partial \mathcal{L}}{\partial y} = x \le 0 \quad \Longrightarrow \quad x=y=0,$$

con $x, y \ge 0$. Esto da $U = 0$, no es máximo interior con $U > 0$. *(Descartado.)*

Caso 2: $\lambda_1 > 0,\ \lambda_2 = 0.$
Por holgura complementaria en $x, y > 0$:

$$y-\lambda_1 = 0, \quad x-\lambda_1 = 0 \quad \Longrightarrow \quad x=y=\lambda_1,$$

y por la restricción activa $x + y = 100$:

$$2\lambda_1 = 100 \ \Rightarrow \ x=y=50,$$

pero viola $x \le 40$. *(Descartado.)*

Caso 3: $\lambda_1 = 0,\ \lambda_2 > 0.$
Holgura complementaria obliga

$$40-x=0 \ \Rightarrow \ x=40,$$

Y yendo a la segunda condición tenemos: $x - \lambda_1 = 0$. Pero como $\lambda_1 = 0$ y $x = 40$, esta condición no se cumple. *(Descartado.)*

Caso 4: $\lambda_1 > 0,\ \lambda_2 > 0.$
Ambas restricciones activas:

$$x+y=100, \quad x=40 \quad \Longrightarrow \quad (x,y)=(40,60).$$

De las derivadas:

$$x-\lambda_1 = 0 \ \Rightarrow \ \lambda_1 = 40, \quad y-\lambda_1-\lambda_2=0 \ \Rightarrow \ \lambda_2 = 20 > 0,$$

todo consistente. $x^* = 40,\ y^* = 60,\ U = 2400.$

Ejemplo minimización

Consideremos el problema

$$\min_{x_1,x_2} \; C(x_1,x_2) \;=\; (x_1-4)^2 \;+\; (x_2-4)^2$$

sujeto a

$$\begin{cases} 2x_1+3x_2 \;\geq\; 6, \\ -3x_1-2x_2 \;\geq\; -12, \\ x_1\geq 0,\; x_2\geq 0. \end{cases}$$

Reescribimos las dos primeras como

$$g_1(x)\equiv 6-2x_1-3x_2 \;\leq\; 0, \quad g_2(x)\equiv 3x_1+2x_2-12 \;\leq\; 0,$$

y asignamos multiplicadores $\lambda_1,\lambda_2\geq 0$. El Lagrangiano es

$$\mathcal{L}=(x_1-4)^2+(x_2-4)^2+\lambda_1\,(6-2x_1-3x_2)+\lambda_2\,(3x_1+2x_2-12).$$

Condiciones de Kuhn-Tucker

$$\begin{cases} \dfrac{\partial\mathcal{L}}{\partial x_1}=2(x_1-4)\;-\;2\lambda_1\;+\;3\lambda_2\;\geq\;0,\; x_1\geq 0,\; x_1[2(x_1-4)-2\lambda_1+3\lambda_2]=0, \\ \dfrac{\partial\mathcal{L}}{\partial x_2}=2(x_2-4)\;-\;3\lambda_1\;+\;2\lambda_2\;\geq\;0,\; x_2\geq 0,\; x_2[2(x_2-4)-3\lambda_1+2\lambda_2]=0, \\ g_1(x)=6-2x_1-3x_2\leq 0,\; \lambda_1\geq 0,\; \lambda_1\,(6-2x_1-3x_2)=0, \\ g_2(x)=3x_1+2x_2-12\leq 0,\; \lambda_2\geq 0,\; \lambda_2\,(3x_1+2x_2-12)=0. \end{cases}$$

Análisis de casos Primero pensemos el caso de $(4,4)$ que dada la función de costos puede ser un potencial óptimo. Sin embargo, $(4,4)$ viola g_2. Por lo que vamos a analizar los otros casos posibles, suponiendo que $x_1>0$ y $x_2>0$

$$x_1>0 \implies \frac{\partial\mathcal{L}}{\partial x_1}=0, \quad x_2>0 \implies \frac{\partial\mathcal{L}}{\partial x_2}=0,$$

y de que las dos restricciones pueden estar activas o inactivas. Hay cuatro combinaciones de $\lambda_1,\lambda_2\geq 0$:

Caso 1: $\lambda_1=0,\; \lambda_2=0.$
De la derivada con respecto a x_1: $2(x_1-4)=0\Rightarrow x_1=4$ y $2(x_2-4)=0\Rightarrow x_2=4$. Pero la segunda restricción $3x_1+2x_2\leq 12$ se viola $(20>12)$. *Descartado.*

Caso 2: $\lambda_1>0,\; \lambda_2=0.$
De las condiciones obtenemos que:

$$2(x_1-4)-2\lambda_1=0 \quad (x_1=4+\lambda_1), \qquad 2(x_2-4)-3\lambda_1=0 \quad (x_2=4+\frac{3}{2}\lambda_1).$$

Además, $\lambda_1>0$ implica $2x_1+3x_2=6$, que no admite $\lambda_1>0$. *Descartado.*

Caso 3: $\lambda_1=0,\; \lambda_2>0.$

$$2(x_1-4)+3\lambda_2=0 \;\Rightarrow\; x_1=4-\frac{3}{2}\lambda_2,$$
$$2(x_2-4)+2\lambda_2=0 \;\Rightarrow\; x_2=4-\lambda_2,$$

y $\lambda_2>0$ obliga $3x_1+2x_2=12$. Resolviendo:

$$x_1=\frac{28}{13}, \quad x_2=\frac{36}{13}, \quad \lambda_2=\frac{16}{13}>0.$$

Se verifica $6-2x_1-3x_2<0$ (la primera restricción es inactiva). Solución válida.

Caso 4: $\lambda_1 > 0,\ \lambda_2 > 0$.
Ambas restricciones activas:

$$2x_1 + 3x_2 = 6, \quad 3x_1 + 2x_2 = 12 \implies (x_1, x_2) = (\frac{24}{5}, -\frac{6}{5}),$$

que viola $x_2 \geq 0$. *Descartado.*

Por tanto, el único caso factible es $\lambda_1 = 0,\ \lambda_2 = \frac{16}{13}$, con

$$(x_1^*, x_2^*) = \left(\frac{28}{13}, \frac{36}{13}\right).$$

Condiciones suficientes

Como en el caso de optimización con restricciones y optimización libre, las condiciones suficientes van a estar asociadas a la concavidad de las funciones analizadas. Vamos a ver dos tipos de condiciones, una sobre concavidad y convexidad y otra más débil sobre cuasiconcavidad y cuasiconvexidad.

Condiciones suficientes para máximo (y global)

Supongamos que el problema a resolver es:

$$\begin{aligned} &\text{Maximizar} \quad f(x) \\ &\text{sujeto a} \qquad g^i(x) \ \leq \ r_i \quad (i = 1, 2, \ldots, k), \\ &\qquad\qquad\qquad x \ \geq \ 0. \end{aligned}$$

Con $x = (x_1, x_2, ...)$ Asumiendo que se satisfacen las condiciones de Kuhn-Tucker. Si se satisfacen las siguientes condiciones en un punto x^*

1. La función objetivo $f(x)$ es diferenciable y *cóncava* en el cuadrante n-dimensional no negativo.
2. Cada función de restricción $g^i(x)$ es diferenciable y *convexa* en el cuadrante n-dimensional no negativo.

Entonces x^* da un máximo global de $f(x)$.
Existe un teorema asociado a condiciones más débiles para poder determinar que estamos ante un máximo, estas están asociados a la cuasiconcavidad y cuasiconvexidad de las funciones:

Teorema de Arrow-Enthoven para máximo

Supongamos que el problema a resolver es:

$$\begin{aligned} &\text{Maximizar} \quad f(x) \\ &\text{sujeto a} \qquad g^i(x) \ \leq \ r_i \quad (i = 1, 2, \ldots, k), \\ &\qquad\qquad\qquad x \ \geq \ 0. \end{aligned}$$

Asumiendo que se satisfacen las condiciones de Kuhn-Tucker. Si se satisfacen las siguientes condiciones en un punto x^*

1. La función objetivo $f(x)$ es diferenciable y *cuasiconcava* en el cuadrante n-dimensional no negativo.
2. Cada función de restricción $g^i(x)$ es diferenciable y *cuasiconvexa* en el cuadrante n-dimensional no negativo.
3. Se satisface cualquiera de los siguientes:
   (a) $f_j(x^*) < 0$ para al menos una variable x_j.

(b) $f_j(x^*) > 0$ para alguna variable x_j relevante, esto quiere decir una variable que en el conjunto factible (que cumpla con las restricciones), esa variable tome un valor positivo.

(c) Las n derivadas $f_j(x^*)$ no son todas cero, y la función $f(x)$ es dos veces diferenciable en la vecindad de x^* (es decir, todas las derivadas parciales de segundo orden de $f(x)$ existen para x^*).

(d) La función $f(x)$ es cóncava.

Entonces x^* da un máximo global de $f(x)$.

Condiciones suficientes para mínimo (y global)

Supongamos que el problema a resolver es:

$$\begin{array}{lll} \text{Minimizar} & f(x) & \\ \text{sujeto a} & g^i(x) \geq r_i & (i = 1, 2, \ldots, k), \\ & x \geq 0. & \end{array}$$

Con $x = (x_1, x_2, ...)$ Asumiendo que se satisfacen las condiciones de Kuhn-Tucker. Si se satisfacen las siguientes condiciones en un punto x^*

1. La función objetivo $f(x)$ es diferenciable y *convexa* en el cuadrante n-dimensional no negativo.
2. Cada función de restricción $g^i(x)$ es diferenciable y *cóncava* en el cuadrante n-dimensional no negativo.

Entonces x^* da un mínimo global de $f(x)$.

Teorema de Arrow-Enthoven para mínimo

Supongamos que el problema a resolver es:

$$\begin{array}{lll} \text{Minimizar} & f(x) & \\ \text{sujeto a} & g^i(x) \geq r_i & (i = 1, 2, \ldots, k), \\ & x \geq 0. & \end{array}$$

Asumiendo que se satisfacen las condiciones de Kuhn-Tucker. Si se satisfacen las siguientes condiciones en un punto x^*

1. La función objetivo $f(x)$ es diferenciable y *cuasiconvexa* en el cuadrante n-dimensional no negativo.
2. Cada función de restricción $g^i(x)$ es diferenciable y *cuasicóncava* en el cuadrante n-dimensional no negativo.
3. Se satisface cualquiera de los siguientes:

   (a) $f_j(x^*) > 0$ para al menos una variable x_j.

   (b) $f_j(x^*) < 0$ para alguna variable x_j relevante, esto quiere decir una variable que en el conjunto factible (que cumpla con las restricciones), esa variable tome un valor positivo.

   (c) Las n derivadas $f_j(x^*)$ no son todas cero, y la función $f(x)$ es dos veces diferenciable en la vecindad de x^* (es decir, todas las derivadas parciales de segundo orden de $f(x)$ existen para x^*).

   (d) La función $f(x)$ es convexa.

   Entonces x^* da un mínimo global de $f(x)$.