Concavidad y convexidad

Aplicaciones matemáticas para economía y negocios (EAF2010)

Felipe Del Canto

Pontificia Universidad Católica de Chile

Primer semestre de 2021

Concavidad y convexidad

Motivación: Precio del cobre

- Recordemos el modelo del precio del cobre que vimos antes.

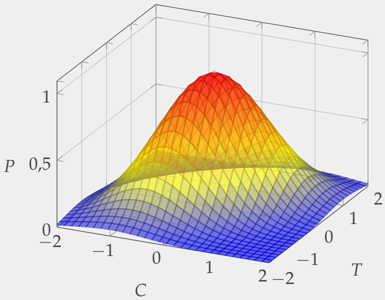
$$P(C,T) = e^{-C^2+CT-T^2}$$

- Esta función es claramente continua en $\mathbb{R}^2$.

- A priori, C y T pueden no estar acotados (por arriba).

Motivación: Precio del cobre

- Entonces, el teorema de Weierstrass no nos ayuda.
  - No nos asegura la existencia de un óptimo.
- Pero la gráfica de la función era bastante interesante:



Motivación: Precio del cobre

- La gráfica tiene una forma particular.
- Decimos que la función es cóncava.
  - Similar a como era para funciones univariadas.
- Si la función es cóncava y tiene un máximo local.
  - Entonces ese máximo debe ser global.
  - Porque se forma este "montículo", no vuelve a subir.

Concavidad, convexidad y optimización

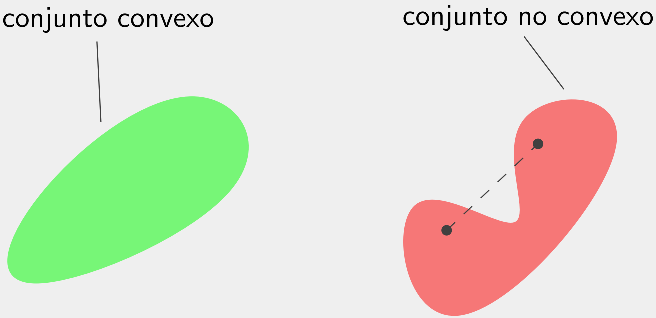
- Las funciones cóncavas y convexas son muy útiles en optimización.
  - No solamente permiten asegurar que óptimos locales serán globales.
  - Sino que además tienen propiedades útiles para efectos computacionales.
- Al igual que en el caso univariado “la segunda derivada” es importante.
  - Y tendrá una interpretación geométrica interesante.
- Pero antes, una definición.

Definición (Conjunto convexo)

Sea $D \subset \mathbb{R}^n$. Decimos que D es **convexo** si para todo $\mathbf{x}_0, \mathbf{x}_1 \in D$ y todo $\lambda \in [0,1]$, el punto $\lambda \mathbf{x}_0 + (1-\lambda)\mathbf{x}_1$ pertenece a D.

- Un punto de la forma $\lambda \mathbf{x}_0 + (1-\lambda)\mathbf{x}_1$ se conoce como **combinación convexa**.
- Geométricamente un conjunto es convexo si:
  - Los segmentos que unen dos puntos se mantienen dentro.
  - No está “abollado”, es decir, “hundido” hacia adentro.

Conjunto convexo



Ejemplo (Conjunto convexo)

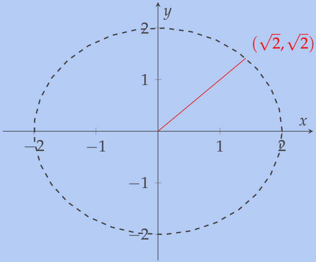
El conjunto $S = \{(x,y) \in \mathbb{R}^2 \,|\, x^2 + y^2 < 4\}$ es convexo. Por dibujo es claro (es una circunferencia), pero algebraicamente también se puede ver. Sea (x_0,y_0) y (x_1,y_1) dos puntos de S y sea $\lambda \in [0,1]$. Queremos ver si el punto $(\lambda x_0 + (1-\lambda)x_1, \lambda y_0 + (1-\lambda)y_1)$ pertenece a S. Para eso calculamos

$$
\begin{aligned}
(\lambda x_0 + (1-\lambda)x_1)^2 + (\lambda y_0 + (1-\lambda)y_1)^2 &= \lambda^2 x_0^2 + 2\lambda(1-\lambda)x_0x_1 + (1-\lambda)^2 x_1^2 \\
&\quad + \lambda^2 y_0^2 + 2\lambda(1-\lambda)y_0y_1 + (1-\lambda)^2 y_1^2 \\
&= \lambda^2(x_0^2 + y_0^2) + 2\lambda(1-\lambda)(x_0x_1 + y_0y_1) \\
&\quad + (1-\lambda)^2(x_1^2 + y_1^2)
\end{aligned}
$$

Los términos $x_i^2 + y_i^2$ son menores que 4 (por la definición de S). Para el término $x_0x_1 + y_0y_1$ deben hacer un salto de fe...

Ejemplo (Conjunto convexo)

Para cada coordenada, el valor más grande que puede tomar es **casi** 2. Sin embargo, si la coordenada x es cercana a 2, la y debe estar cerca a 0, y viceversa.



Eso significa que para aumentar al máximo el término $x_0x_1 + y_0y_1$, tenemos que hacerlos parecidos, lo que obliga a que valgan a lo más $\sqrt{2}$ cada uno. Así, $x_0x_1 + y_0y_1 < \sqrt{2}^2 + \sqrt{2}^2 = 4$.

Conjunto convexo

Ejemplo (Conjunto convexo)

Juntando todo, tenemos que

$$
\begin{aligned}
(\lambda x_0 + (1-\lambda)x_1)^2 + (\lambda y_0 + (1-\lambda)y_1)^2 &= \lambda^2(x_0^2 + y_0^2) + 2\lambda(1-\lambda)(x_0x_1 + y_0y_1) \\
&\quad + (1-\lambda)^2(y_0^2 + y_1^2) \\
&< 4\lambda^2 + 4\cdot(2\lambda(1-\lambda)) + 4(1-\lambda)^2 \\
&= 4\Big(\lambda^2 + 2\lambda(1-\lambda) + (1-\lambda)^2\Big) \\
&= 4\Big((\lambda + (1-\lambda))^2\Big) \\
&= 4
\end{aligned}
$$

Es decir, por definición el punto $(\lambda x_0 + (1-\lambda)x_1, \lambda y_0 + (1-\lambda)y_1)$ pertenece al conjunto S.

Ejercicio (Conjunto convexo)

Determine algebraicamente si el conjunto $\{(x,y) \in \mathbb{R}^2 \,|\, x > 0, y > 0, xy \geq 1\}$ es convexo. Ayúdese con un dibujo si lo necesita. (Bonus: ¿Es compacto?)

Definición (Funciones cóncavas y convexas)

Sea $f : D \to \mathbb{R}$ una función, con $D \subset \mathbb{R}^n$ convexo. Decimos que

- f es **cóncava** si
$$f(\lambda \mathbf{x}_0 + (1-\lambda)\mathbf{x}_1) \geq \lambda f(\mathbf{x}_0) + (1-\lambda) f(\mathbf{x}_1)$$
para todo $\mathbf{x}_0, \mathbf{x}_1 \in D$ con $\mathbf{x}_0 \neq \mathbf{x}_1$ y para todo $\lambda \in (0,1)$. Si la desigualdad es estricta, decimos que f es **estrictamente cóncava**.

- f es **convexa** si
$$f(\lambda \mathbf{x}_0 + (1-\lambda)\mathbf{x}_1) \leq \lambda f(\mathbf{x}_0) + (1-\lambda) f(\mathbf{x}_1)$$
para todo $\mathbf{x}_0, \mathbf{x}_1 \in D$ con $\mathbf{x}_0 \neq \mathbf{x}_1$ y para todo $\lambda \in (0,1)$. Si la desigualdad es estricta, decimos que f es **estrictamente convexa**.

Notar que f es cóncava si y solo si $-f$ es convexa, y viceversa.

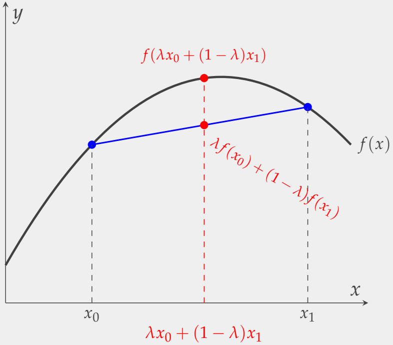
Funciones cóncavas y convexas

- También podemos ver concavidad geométricamente.
  - Segmentos que unen partes de la función quedan debajo de ella.

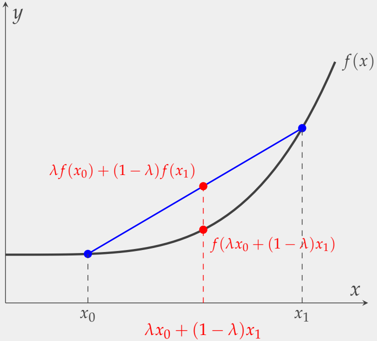
- Y viceversa con convexidad.

- Esto tendrá directa relación con el concepto de recta/plano tangente.

Funciones cóncavas y convexas



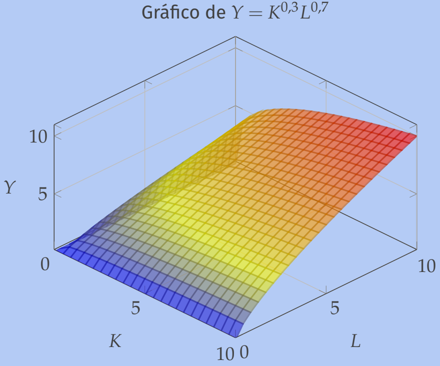
(a) Función cóncava



(b) Función convexa

Ejemplo (Funciones cóncavas y convexas)

La función $Y(K,L) = K^{0,3}L^{0,7}$ que vimos en el capítulo 1 es cóncava.



Ejemplo (Funciones cóncavas y convexas)

Pero no es estrictamente cóncava porque, como Y es homogénea de grado 1, todos los puntos de la forma (tK, tL) con $t > 0$ cumplen:

$$Y(tK, tL) = (tK)^{0{,}3}(tL)^{0{,}7} = tY(K, L),$$

es decir, Y tiene un comportamiento lineal a lo largo de rayos desde el origen y por lo tanto la desigualdad en la definición no será estricta.

Ejercicio (Funciones cóncavas y convexas)

Muestre que las funciones lineales, $f(\mathbf{x}) = \sum_{i=1}^{n} c_i x_i$, donde $(c_1, \ldots, c_n) \in \mathbb{R}^n$, son cóncavas y convexas. Muestre, pero solo para el caso univariado, que son las únicas que cumplen esta propiedad. (*Ayuda: ¡use un dibujo!*)

Ejercicio (Aplicación a loterías)

Un agente tiene función de utilidad $u(w)$, donde w es su ingreso. Inicialmente, su renta es w_0. A este agente le ofrecen jugar una lotería: si gana obtiene z y si pierde le quitan z. Cada resultado tiene $\frac{1}{2}$ de probabilidad de ocurrir. Si participa del juego, entonces su utilidad es

$$\frac{1}{2}u(w_0 - z) + \frac{1}{2}u(w_0 + z)$$

Suponga que u es estrictamente creciente.

- Si u es cóncava, ¿participa el agente de la lotería?
- Si u es convexa ¿participa el agente de la lotería?

- En el ejemplo de la función Cobb-Douglas, determinar la concavidad es difícil.
  - Con el dibujo no, pero algebraicamente es muy difícil.
- Necesitamos más herramientas para determinar esta característica.
- La que veremos a continuación también tiene interpretación geométrica.

Teorema (Concavidad, convexidad y gradiente)

Sea $f : D \to \mathbb{R}$ una función multivariada, con D abierto y convexo. Entonces

- f es cóncava si y solo si

  $$f(\mathbf{x}) \leq f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \qquad \forall\, \mathbf{x}, \mathbf{x}_0 \in D, \quad \mathbf{x} \neq \mathbf{x}_0$$

  y es estrictamente cóncava si la desigualdad se satisface de manera estricta.

- f es convexa si y solo si

  $$f(\mathbf{x}) \geq f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \qquad \forall\, \mathbf{x}, \mathbf{x}_0 \in D, \quad \mathbf{x} \neq \mathbf{x}_0$$

  y es estrictamente convexa si la desigualdad se satisface de manera estricta.

Ejemplo (Concavidad, convexidad y gradiente)

La función $f(x,y) = x^2 + y^2$ es estrictamente convexa. Por dibujo es fácil verlo pero con el teorema anterior también. Sean $(x,y),(x_0,y_0)$ dos puntos (distintos) cualquiera de $\mathbb{R}^2$. Tenemos que

$$\nabla f(x_0,y_0) = (2x_0,2y_0),$$

luego

$$\nabla f(x_0,y_0) \cdot (x - x_0, y - y_0) = 2x_0x + 2y_0y - 2x_0^2 - 2y_0^2$$

Y así

$$f(x_0,y_0) + \nabla f(x_0,y_0) \cdot (x - x_0, y - y_0) = 2x_0x + 2y_0y - x_0^2 - y_0^2$$

Ejemplo (Concavidad, convexidad y gradiente)

Queremos comparar

$$f(x,y) \qquad vs \qquad f(x_0,y_0)+\nabla f(x_0,y_0)\cdot(x-x_0,y-y_0)$$

O bien

$$x^2+y^2 \qquad vs \qquad 2x_0x+2y_0y-x_0^2-y_0^2$$

Lo que nos da una idea de lo que debemos hacer

$$\begin{aligned}
2x_0x+2y_0y-x_0^2-y_0^2 &= 2x_0x+2y_0y-x_0^2-y_0^2+x^2+y^2-x^2-y^2\\
&= x^2+y^2-(x^2-2x_0x+x_0^2)-(y^2-2y_0y+y_0^2)\\
&= x^2+y^2-(x-x_0)^2-(y-y_0)^2\\
&\leq x^2+y^2
\end{aligned}$$

Ejemplo (Concavidad, convexidad y gradiente)

Lo último dice que

$$f(x,y) \geq f(x_0,y_0) + \nabla f(x_0,y_0) \cdot (x - x_0, y - y_0)$$

La desigualdad es estricta porque para que haya igualdad debe ocurrir que $x = x_0$ e $y = y_0$. Pero eso es imposible porque supusimos que los puntos eran distintos. Así, f es estrictamente convexa.

Ejercicio (Concavidad, convexidad y gradiente)

Determine si $Y = K^{0{,}3}L^{0{,}7}$ es (estrictamente) cóncava o convexa. (Ayuda: Puede apoyarse en esta página.)

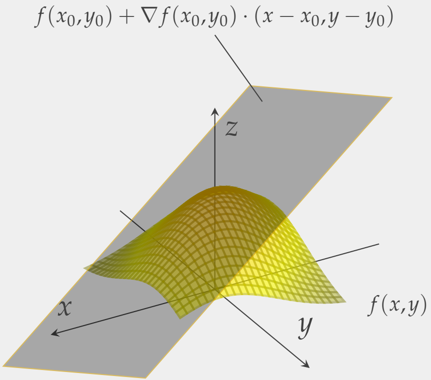
- Pensemos que f es bivariada, cóncava y con derivadas parciales continuas.

- Entonces, dado un punto (x_0, y_0) se cumple

$$f(x,y) \leq f(x_0,y_0) + \nabla f(x_0,y_0)(x - x_0, y - y_0)$$

- El lado derecho es la ecuación del plano tangente (si f es bivariada).
  - Luego, si f es cóncava, el plano tangente está sobre ella en todo punto.

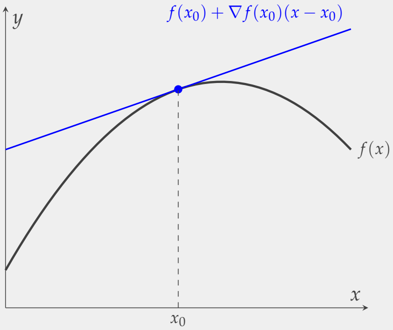
Interpretación geométrica



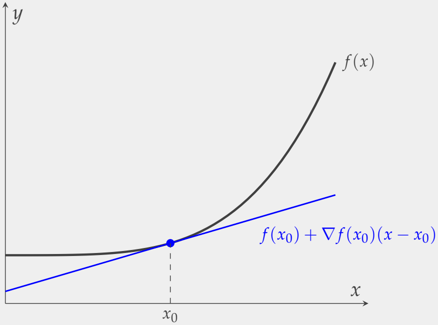
Interpretación geométrica

- Para las funciones convexas pasa lo mismo.
- En esos casos, el plano tangente (si f es bivariada) está por debajo.
- Para funciones univariadas también vale la misma interpretación.

Interpretación geométrica



(a) Función cóncava



(b) Función convexa

- El teorema anterior nos permitirá formalizar la intuición del comienzo.
- Si tenemos un máximo local de una función cóncava, este debe ser global.
- La idea es que el plano tangente “no deja que la función vuelva a subir”.

Corolario (Concavidad y máximos globales)

Sea $f : D \to \mathbb{R}$ una función cóncava y con dominio $D \subset \mathbb{R}^n$ abierto y convexo. Supongamos que $\mathbf{x}_0 \in D$ es un máximo local de f. Entonces $\mathbf{x}_0$ es un máximo global de f. Si además f es estrictamente cóncava, entonces $\mathbf{x}_0$ es el único máximo global.

Demostración

Si $\mathbf{x}_0$ es un máximo local, entonces $\nabla f(\mathbf{x}_0) = 0$ (por las condiciones de primer orden). Por lo tanto, por el teorema anterior, para todo $\mathbf{x} \in D$,

$$f(\mathbf{x}) \leq f(\mathbf{x}_0) + \underbrace{\nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)}_{0} = f(\mathbf{x}_0),$$

lo que dice que $\mathbf{x}_0$ es un máximo global. Si f es estrictamente cóncava, entonces la desigualdad es estricta y $\mathbf{x}_0$ es el único máximo global.

- El resultado también vale para funciones convexas, pero con el mínimo global.

Corolario (Convexidad y mínimos globales)

Sea $f : D \to \mathbb{R}$ una función convexa y con dominio $D \subset \mathbb{R}^n$ abierto y convexo. Supongamos que $\mathbf{x}_0 \in D$ es un mínimo local de f. Entonces $\mathbf{x}_0$ es un mínimo global de f. Si además f es estrictamente convexa, entonces $\mathbf{x}_0$ es el único mínimo global.

Ejemplo (Concavidad/convexidad y optimización)

Ya vimos que la función $f(x,y) = x^2 + y^2$ es estrictamente convexa. Si esta función tuviera un mínimo local en algún punto, entonces ese punto sería mínimo global. Mirando el gradiente

$$\nabla f(x,y) = (2x,2y)$$

vemos que un candidato a óptimo local es el punto $(0,0)$, donde $f(0,0) = 0$. ¿Es un máximo local, mínimo local o nada? En este caso particular vemos que es un mínimo local, porque cerca de $(0,0)$ alguna coordenada no vale cero y, por lo tanto,

$$f(x,y) > 0 = f(0,0)$$

Luego $(0,0)$ es mínimo local y por el teorema anterior es mínimo global. Como además la convexidad es estricta, este punto es el único mínimo global de f.

Ejercicio (Concavidad/convexidad y optimización)

Encuentre un candidato a óptimo para la función $f(x,y) = 4x + 2y - x^2 - y^2 + xy$. Asuma que ese candidato es máximo local y determine si es máximo global (y si es único).

- Hay un resultado un poco más fuerte.
- Para funciones cóncavas y convexas las CPO son **necesarias** y **suficientes**.
- Esto por la misma noción anterior:
  - Si tenemos un punto crítico de f cóncava o convexa.
  - Entonces el plano tangente en ese punto deja a f a un solo lado.
  - Luego ese punto es mínimo o máximo local (y por el corolario anterior, es global).

Teorema (Convexidad, concavidad y CPO)

Sea $f : D \to \mathbb{R}$ una función con dominio $D \subset \mathbb{R}^n$ convexo. Sea $\mathbf{x}_0 \in D$ un punto interior (no está en el borde de D). Tenemos que:

- Si f es cóncava, entonces $\mathbf{x}_0$ es un máximo global de f si y solo si $\mathbf{x}_0$ es un punto crítico.
- Si f es convexa, entonces $\mathbf{x}_0$ es un mínimo global de f si y solo si $\mathbf{x}_0$ es un punto crítico.

Formas cuadráticas

- Hasta ahora sabemos relacionar concavidad/convexidad con optimización.
- Y tenemos la intuición geométrica.
- Pero determinar concavidad/convexidad puede ser difícil.
  - Es cosa de pensar en los ejemplos anteriores.

- Necesitamos un método más simple para trabajar.
- Para eso nos basaremos en “la segunda derivada”.
  - Que en este caso es la matriz Hessiana.
- **¡CUIDADO!** El método es conceptualmente simple.
  - Pero involucra calcular mucho.

- Recordemos que en una variable existía el siguiente resultado.

Teorema (Función cóncava)

Sea $f : D \to \mathbb{R}$ una función dos veces derivable con $D \subset \mathbb{R}$ abierto. Entonces f es cóncava si y solo si $f''(x) \leq 0$ para todo x y es estrictamente cóncava si la desigualdad es estricta.

- El resultado vale para convexas, dando vuelta la desigualdad.

- Este método funciona porque si f es univariada, entonces

$$f(x) \approx \underbrace{f(x_0) + f'(x_0)(x - x_0)}_{\text{recta tangente}} + \underbrace{f''(x_0)\frac{(x - x_0)^2}{2}}_{\substack{\text{término que} \\ \text{mejora la aproximación}}}$$

- Si $f''(x) \leq 0$, entonces $f(x)$ está por debajo de la recta tangente.
  - Y dijimos que eso es precisamente concavidad.

- Y viceversa con la convexidad.

- La idea del método para funciones multivariadas es similar.
- La "segunda derivada" debe ser positiva/negativa.
- Pero en este caso la aproximación se ve diferente:

$$f(\mathbf{x}) \approx \underbrace{f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)}_{\text{“plano” tangente}} + \underbrace{(\mathbf{x} - \mathbf{x}_0)^T H(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)}_{\substack{\text{término que} \\ \text{mejora la aproximación}}}$$

- Nos gustaría que el término adicional sea negativo o positivo siempre.
  - Para asegurar concavidad o convexidad, respectivamente.

- Esto da origen a la siguiente definición.

- Y los resultados que veremos después intentaran decirnos cuándo ocurre cual.

Definición (Matriz (semi)definida positiva)

Sea A una matriz simétrica de $n \times n$. Decimos que A es:

- **definida positiva** si $\mathbf{x}^T A\mathbf{x} > 0$ para todo $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq 0$.
- **semidefinida positiva** si $\mathbf{x}^T A\mathbf{x} \geq 0$ para todo $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq 0$.

- Matrices (semi)definidas negativas se definen de manera similar.
  - Pero con las desigualdades al revés.
- Cuidado que no todas las matrices se clasifican como alguna de estas dos.
  - A las que no son semidefinidas de ningún tipo, las llamamos **indefinidas**.

Ejemplo (Matriz (semi)definida positiva)

Ya vimos que $f(x,y) = x^2 + y^2$ es estrictamente convexa. Notar que la matriz Hessiana es simétrica (teorema de Young)

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Y así, tomando un vector $(x,y) \in \mathbb{R}^2$ con $x,y \neq 0$, tenemos

$$(x,y)^T H(x,y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 2x \\ 2y \end{pmatrix} = 2x^2 + 2y^2 > 0$$

Y por lo tanto H es definida positiva.

Ejercicio (Matriz (semi)definida positiva)

Determine si $Y(L,K) = L^{0,3}K^{0,7}$ tiene matriz Hessiana (semi)definida positiva o negativa. (*Ayuda: Recuerde que* $K, L \geq 0$.)

Teorema (Concavidad, convexidad y matriz Hessiana)

Sea $f : D \to \mathbb{R}$ una función dos veces diferenciable, con $D \subset \mathbb{R}^n$ convexo y abierto. Entonces:

- f es cóncava si y solo si $H(\mathbf{x})$ es semidefinida negativa para todo $\mathbf{x} \in D$.
- f es convexa si y solo si $H(\mathbf{x})$ es semidefinida positiva para todo $\mathbf{x} \in D$.

Además, f es estrictamente cóncava/convexa si $H(x)$ es definida negativa/positiva para todo $\mathbf{x} \in D$.

- La demostración es solo formalizar el argumento anterior de la aproximación.

- El teorema anterior es extremadamente útil.
- Pero a veces determinar si H es (semi)definida positiva o negativa no es fácil.
- Queremos una “receta” para hacerlo más fácil.
  - Existe, pero necesitamos unas definiciones adicionales.

Definición (Menor principal dominante)

Sea A una matriz de $n \times n$. Llamamos menor principal dominante de orden k de A (D_k) al determinante de la matriz que queda cuando mantenemos las primeras k filas y columnas, y eliminamos el resto, es decir

$$D_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}$$

Dicho de otra forma, lo que estamos haciendo es quitar las últimas $n-k$ filas y columnas de A y calculando el determinante de la matriz que resulta.

Ejemplo (Menor principal dominante)

Si

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$$

Entonces

$$D_1 = \begin{vmatrix} 2 \end{vmatrix}, \qquad D_2 = \begin{vmatrix} 2 & 3 \\ 3 & 5 \end{vmatrix}$$

Luego

$$D_1 = 2, \qquad D_2 = 2 \cdot 5 - 3 \cdot 3 = 1.$$

Teorema (Menores y matrices definidas positivas/negativas)

Sea A una matriz simétrica de $n \times n$. Tenemos que

- A es definida positiva si y solo si $D_k > 0$ para todo k.
- A es definida negativa si y solo si $(-1)^k D_k > 0$ para todo k.

Dicho de otra manera

- A es definida positiva si y solo si sus menores principales dominantes son positivos.
- A es definida negativa si y solo si sus menores principales dominantes alternan signo, y el primero es negativo.

Ejemplo (Menores y matrices definidas positivas/negativas)

En el ejemplo anterior,

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$$

Y

$$D_1 = 2, \qquad D_2 = 2 \cdot 5 - 3 \cdot 3 = 1.$$

Luego A es definida positiva.

Ejercicio (Menores y matrices definidas positivas/negativas)

Determine si las siguientes matrices son definidas positivas, definidas negativas o no se sabe. (*Ayuda: Si no recuerda cómo calcular determinates de 3x3 puede usar esta página*)

- $A = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$.
- $A = \begin{pmatrix} 1 & 4 \\ 4 & 1 \end{pmatrix}$.
- $A = \begin{pmatrix} -3 & 0 & 3 \\ 0 & -1 & -2 \\ -3 & -2 & -8 \end{pmatrix}$.

- Nos falta un resultado similar, pero para (semi)definición.
- Aquí los menores principales dominantes no bastan.
- Necesitamos una definición adicional.

Definición (Menor principal)

Sea A una matriz de $n \times n$. Llamamos menor principal de orden k de A ($\tilde{D}_k$) al determinante de la matriz que queda cuando mantenemos k filas y las mismas k columnas, eliminando el resto.

Dicho de otra forma, lo que estamos haciendo es quitar $n-k$ filas y las mismas $n-k$ columnas de A, para luego calcular el determinante de la matriz que resulta.

- Es fácil darse cuenta que calcular **TODOS** los menores principales es engorroso.
  - Pero técnicamente se puede hacer con calma.

Ejemplo (Menor principal)

Para la matriz de antes,

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$$

Los menores principales dominantes son a su vez menores principales. Pero además se agrega el que resulta de quitar la primera fila y la primera columna:

$$\tilde{D}_1 = \begin{vmatrix} 5 \end{vmatrix} = 5$$

Teorema (Menores y matrices semidefinidas positivas/negativas)

Sea A una matriz simétrica de $n \times n$. Tenemos que

- A es semidefinida positiva si y solo si todo menor principal es no negativo (≥ 0).
- A es semidefinida negativa si y solo si todo menor principal de orden par (k es par) es no negativo (≥ 0) y todo menor principal de orden impar (k impar) es no positivo (≤ 0).

- Este teorema es en esencia igual al anterior, pero con dos diferencias.
  - Las desigualdades no son estrictas.
  - Los menores principales dominantes cambiaron por menores principales.

Aplicación: Desigualdad de Jensen

Desigualdad de Jensen (versión discreta)

- Recordemos la definición de función cóncava.

Definición (Función cóncava)

Sea $f : D \to \mathbb{R}$ una función, con $D \subset \mathbb{R}^n$ convexo. Decimos que f es cóncava si

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \geq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

para todo $\mathbf{x}, \mathbf{y} \in D$ con $\mathbf{x} \neq \mathbf{y}$ y para todo $\lambda \in (0,1)$

- Esto también puede generalizarse a más de dos puntos $\mathbf{x}$ e $\mathbf{y}$.

Desigualdad de Jensen (versión discreta)

Teorema (Desigualdad de Jensen, versión discreta)

Sea $f : D \to \mathbb{R}$ una función, con $D \subset \mathbb{R}^n$ convexo. Entonces f es cóncava si y solo si para todo $\mathbf{x}_1, \ldots, \mathbf{x}_m \in S$, para todo $\lambda_1, \ldots, \lambda_m \geq 0$ con $\lambda_1 + \cdots + \lambda_m = 1$ se cumple que

$$f(\lambda_1 \mathbf{x}_1 + \cdots + \lambda_m \mathbf{x}_m) \geq \lambda_1 f(\mathbf{x}_1) + \cdots + \lambda_m f(\mathbf{x}_m)$$

- Observar que cuando $m = 2$, entonces rescatamos la definición anterior.
- La gracia del teorema es que podemos incorporar más de dos puntos.
  - Ojo que con eso la interpretación geométrica se vuelve más compleja.

Idea de la demostración (Desigualdad de Jensen, versión discreta)

Veamos primero el caso "solo si", es decir, que f cóncava implica la desigualdad. Partamos con $m=3$. Tomemos $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \in D$ cualquiera y $\lambda_1, \lambda_2, \lambda_3 \geq 0$ con $\lambda_1 + \lambda_2 + \lambda_3 = 1$. Supongamos que $\lambda_2 + \lambda_3 \neq 0$ (si no fuera así, solo nos quedamos con λ_1 y el resultado sería "obvio"). Observen que

$$\lambda_1 = 1 - (\lambda_2 + \lambda_3)$$

y podemos escribir

$$\begin{aligned} f(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \lambda_3 \mathbf{x}_3) &= f\left(\lambda_1 \mathbf{x}_1 + (\lambda_2 + \lambda_3)\left(\frac{\lambda_2}{\lambda_2 + \lambda_3}\mathbf{x}_2 + \frac{\lambda_3}{\lambda_2 + \lambda_3}\mathbf{x}_3\right)\right) \\ &= f\left(\lambda_1 \mathbf{x}_1 + (1 - \lambda_1)\left(\frac{\lambda_2}{\lambda_2 + \lambda_3}\mathbf{x}_2 + \frac{\lambda_3}{\lambda_2 + \lambda_3}\mathbf{x}_3\right)\right) \end{aligned}$$

Idea de la demostración (Desigualdad de Jensen, versión discreta)

Como D es convexo, entonces $\frac{\lambda_2}{\lambda_2+\lambda_3}\mathbf{x}_2 + \frac{\lambda_3}{\lambda_2+\lambda_3}\mathbf{x}_3$ es parte de D. Luego, por la concavidad (y reemplazando $1-\lambda_1 = \lambda_2+\lambda_3$)

$$f(\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \lambda_3\mathbf{x}_3) \geq \lambda_1 f(\mathbf{x}_1) + (\lambda_2+\lambda_3) f\left(\frac{\lambda_2}{\lambda_2+\lambda_3}\mathbf{x}_2 + \frac{\lambda_3}{\lambda_2+\lambda_3}\mathbf{x}_3\right)$$

Y aplicando la definición de concavidad en el segundo término y simplificando

$$f(\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \lambda_3\mathbf{x}_3) \geq \lambda_1 f(\mathbf{x}_1) + \lambda_2 f(\mathbf{x}_2) + \lambda_3 f(\mathbf{x}_3)$$

El resultado general se obtiene haciendo inducción matemática sobre la cantidad de términos (no es importante para este curso eso). Finalmente, probar que la desigualdad implica que f sea cóncava es solamente la definición de concavidad (tomando $m=2$).

Desigualdad de Jensen (versión continua)

- Este resultado es interesante, pero no tan útil.
- Hay un resultado un poco más útil.
  - Sobre todo para probabilidades.
- Para deducirlo vamos a avanzar desde el teorema anterior.

Desigualdad de Jensen (versión continua)

- Pensemos que f es cóncava y univariada, con dominio $D = [a,b]$. Entonces

$$f(\lambda_1 x_1 + \cdots + \lambda_m x_m) \geq \lambda_1 f(x_1) + \cdots + \lambda_m f(x_m)$$

- Algo curioso pasa si:
  - m empieza a crecer mucho (sin repetir puntos).
  - Los λ_i se acercan a $\frac{a-b}{m}$.

- Tenemos que,

$$\lambda_1 x_1 + \cdots + \lambda_m x_m \approx \int_a^b x\,dx, \qquad \lambda_1 f(x_1) + \cdots + \lambda_m f(x_m) \approx \int_a^b f(x)\,dx$$

Desigualdad de Jensen (versión continua)

- Es decir, que si f es cóncava, entonces

$$f\left(\int_a^b x\,dx\right) \geq \int_a^b f(x)\,dx$$

- En palabras burdas:

Si f es cóncava, podemos "pasar" la función dentro de la integral, pagando el precio de obtener un resultado menor.

- Se puede obtener un resultado un poco más general que este.
  - Pero que mantiene la misma esencia.

Desigualdad de Jensen (versión continua)

Teorema (Desigualdad de Jensen, versión continua)

Sea $f : D \to \mathbb{R}$ una función cóncava, con $D \subset \mathbb{R}$ convexo. Sean $x(t)$, $\lambda(t)$ dos funciones continuas, con dominio $[a,b]$. Supongamos que $\lambda(t) \geq 0$ y que $\int_a^b \lambda(t)\,dt = 1$. Entonces

$$f\left(\int_a^b \lambda(t)x(t)\,dt\right) \geq \int_a^b \lambda(t)f(x(t))\,dt$$

- A diferencia de la discusión anterior, λ y x dependen de otra variable.
- Estos resultados también valen para funciones convexas.
  - Invirtiendo las desigualdades.

Desigualdad de Jensen (versión continua)

- La aplicación usual es cuando $\lambda(t)$ es una cierta función de densidad.
- Luego x y $f(x)$ pueden verse como variables aleatorias y

$$f(\mathbb{E}[x]) \geq \mathbb{E}[f(x)]$$

- Si ninguna de estas palabras les suenan **no desesperen**.
  - Aparecerán en algún momento de su vida.

Desigualdad de Jensen (versión continua)

Ejercicio Revisar el Ejemplo 3 en el capítulo 17.7 del libro SHC.

Funciones cuasicóncavas y cuasiconvexas

FUNCIONES CUASICÓNCAVAS Y CUASICONVEXAS

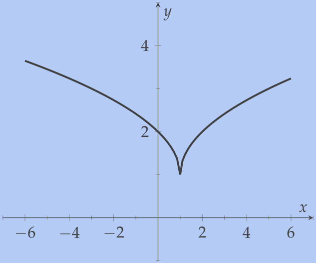
- Muchas veces nos topamos con funciones ni cóncavas ni convexas.

- Pero que tienen propiedades similares.
  - Relación con óptimos locales o globales.

- Y en muchas aplicaciones eso nos va a bastar.

Ejemplo (Función no convexa)

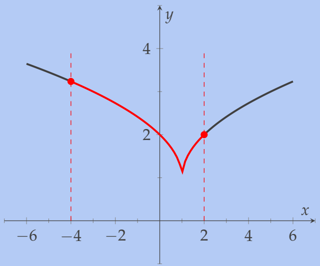
Consideremos la siguiente función



Esta función no es convexa (y no es cóncava). Pero tiene una propiedad interesante.

Ejemplo (Función no convexa)

Si consideramos dos puntos cualquiera del dominio, todos los valores de la función dentro de ese “pedazo” son menores (o iguales) que el valor máximo en los extremos



Si esta función se restringe a un conjunto convexo y compacto (como el del dibujo), ¡entonces su máximo (global) está en los extremos!

Funciones cuasicóncavas y cuasiconvexas

- Funciones como la del ejemplo se conocen como "cuasicónvexas".
  - Como dijimos, son útiles en algunas aplicaciones.

- Su definición intenta "generalizar" el concepto de convexidad.

- Lo que nos va a traer recuerdos de todo lo anterior.

Definición (Funciones cuasicóncavas y cuasiconvexas)

Sea $f : D \to \mathbb{R}$ con dominio $D \subset \mathbb{R}^n$ convexo. Decimos que f es:

- **cuasiconvexa** si para todo par de puntos $\mathbf{x}, \mathbf{y} \in D$, con $\mathbf{x} \neq \mathbf{y}$ y para todo $\lambda > 0$,

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leq \max\{f(\mathbf{x}), f(\mathbf{y})\}$$

- **cuasicóncava** si para todo par de puntos $\mathbf{x}, \mathbf{y} \in D$, con $\mathbf{x} \neq \mathbf{y}$ y para todo $\lambda > 0$,

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \geq \min\{f(\mathbf{x}), f(\mathbf{y})\}$$

- Notar que f es cuasicóncava si y solo si $-f$ es cuasiconvexa, y viceversa.

- Supuestamente estas son generalizaciones de concavidad/convexidad.

- Para eso, debe ser cierto que:
  - Cóncava $\Longrightarrow$ cuasicóncava.
  - Convexa $\Longrightarrow$ cuasiconvexa.
  - Las implicancias para el otro lado no son verdaderas.

- Esto es cierto y lo resumimos en el siguiente resultado.

Proposición (Funciones cóncavas, convexas, cuasicóncavas y cuasiconvexas)

Sea $f : D \to \mathbb{R}$ con dominio $D \subset \mathbb{R}^n$ convexo. Entonces:

1. f cóncava $\Longrightarrow f$ cuasicóncava.
2. f convexa $\Longrightarrow f$ cuasiconvexa.

Y las implicancias inversas no son ciertas.

Ejercicio Demuestre la proposición (use el ejemplo anterior de ayuda)

Ejercicio (Funciones cóncavas y convexas)

Muestre que las funciones lineales, $f(\mathbf{x}) = \sum_{i=1}^{n} c_i x_i$, donde $(c_1, \ldots, c_n) \in \mathbb{R}^n$, son cuasicóncavas y cuasiconvexas. Muestre, solo para el caso univariado, que NO son las únicas que cumplen esta propiedad.

Teorema (Funciones cuasicóncavas)

Sea $f : D \to \mathbb{R}$ con dominio $D \subset \mathbb{R}^n$ convexo. Entonces las siguientes afirmaciones son equivalentes:

1. f es cuasicóncava.
2. Para todo $\mathbf{x}, \mathbf{y} \in D$ y todo $\lambda \in [0,1]$, si $f(\mathbf{x}) \geq f(\mathbf{y})$, entonces
$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \geq f(\mathbf{y})$$
3. El conjunto $P_c = \{\mathbf{x} \in D \mid f(\mathbf{x}) \geq c\}$ (el conjunto sobrenivel) es convexo para todo c.

Teorema (Funciones cuasiconvexas)

Sea $f : D \to \mathbb{R}$ con dominio $D \subset \mathbb{R}^n$ convexo. Entonces las siguientes afirmaciones son equivalentes:

1. f es cuasiconvexa.
2. Para todo $\mathbf{x}, \mathbf{y} \in D$ y todo $\lambda \in [0,1]$, si $f(\mathbf{x}) \geq f(\mathbf{y})$, entonces
$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leq f(\mathbf{x})$$
3. El conjunto $P^c = \{\mathbf{x} \in D \,|\, f(\mathbf{x}) \leq c\}$ (el conjunto bajonivel) es convexo para todo c.

- Las equivalencias entre 1 y 2 es “simple”.
- Lo interesante son las equivalencias entre 1 y 3.
- Porque conectan la forma de la función con la de (una porción) del dominio.

Ejercicio (Funciones cuasicóncavas y cuasiconvexas)

Compruebe que para la función que vimos en el ejemplo del comienzo se cumple 3.

El Hessiano orlado

- Estas funciones también se pueden identificar usando la matriz Hessiana.
- Pero se requieren algunas modificaciones.
- Vamos a ver una definición y un teorema.

Definición (Hessiano orlado o ampliado)

Sea $f : D \to \mathbb{R}$ dos veces diferenciable, con dominio $D \subset \mathbb{R}^n$ convexo y abierto. El Hessiano orlado (o ampliado) de f es la matriz Hessiana ampliada que se muestra a continuación

$$\overline{H} = \begin{pmatrix} 0 & f_{x_1} & f_{x_2} & \cdots & f_{x_n} \\ f_{x_1} & f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_n} \\ f_{x_2} & f_{x_2x_1} & f_{x_2x_2} & \cdots & f_{x_2x_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{x_n} & f_{x_nx_1} & f_{x_nx_2} & \cdots & f_{x_nx_n} \end{pmatrix}$$

O, escrito en forma de “bloque”,

$$\overline{H} = \begin{pmatrix} 0 & \nabla f \\ \nabla f^T & H \end{pmatrix}$$

El Hessiano orlado

Teorema (Cuasiconcavidad y Hessiano orlado)

Sea $f : D \to \mathbb{R}$ dos veces diferenciable, con dominio $D \subset \mathbb{R}^n$ convexo y abierto. Entonces

- Si f es cuasicóncava, entonces para $\overline{H}$, los D_k de orden par son no positivos (≤ 0) y los de orden impar (mayores que 1) son no negativos (≥ 0).
- Si para $\overline{H}$, los D_k de orden par son negativos y los de orden impar (mayores que 1) son positivos, entonces entonces f es cuasicóncava.

- Observar que esto no es un si y solo si.
- Observar además que el D_k de orden 1 no se revisa (¡y es 0 por definición!).

- Un resultado similar vale para las cuasiconvexas.

Teorema (Cuasiconvexidad y Hessiano orlado)

Sea $f : D \to \mathbb{R}$ dos veces diferenciable, con dominio $D \subset \mathbb{R}^n$ convexo y abierto. Entonces

- Si f es cuasiconvexa, entonces para $\overline{H}$, los D_k (con $k > 1$) son no positivos (≤ 0).
- Si para $\overline{H}$, los D_k (con $k > 1$) son negativos, entonces entonces f es cuasiconvexa.

Ejercicio (Cuasiconcavidad y hessiano orlado)

Muestre que la función $u(x,y) = \ln(x) + \ln(y)$ es cuasicóncava.

Funciones (cuasi)cóncavas, convexas y transformaciones

Algunas propiedades interesantes

- Para cerrar, veremos algunos resultados interesantes.
- Que nos permitirán clasificar ciertas funciones.
  - Como (cuasi)cóncavas y convexas.
- Cuando estas sean transformaciones de otras funciones que sí conozcamos.

Teorema (Concavidad, convexidad y transformaciones)

- La suma de funciones cóncavas (convexas) es cóncava (convexa).
- Si f es cóncava (convexa) y F es concava (convexa) y creciente, entonces $F(f(\mathbf{x}))$ es cóncava (convexa).
- Si f es cóncava (convexa) y F es convexa (cóncava) y decreciente, entonces $F(f(\mathbf{x}))$ es convexa (cóncava).

Teorema (Cuasiconcavidad, cuasiconvexidad y transformaciones)

- Una suma de funciones cuasicóncavas (cuasiconvexas) no es necesariamente cuasicóncava (cuasiconvexa).
- Si f es cuasicóncava (cuasiconvexa) y F es estrictamente creciente, entonces $F(f(\mathbf{x}))$ es cuasicóncava (cuasiconvexa).
- Si f es cuasicóncava (cuasiconvexa) y F es estrictamente decreciente, entonces $F(f(\mathbf{x}))$ es cuasiconvexa (cuasicóncava).

- Para la primera, consideren $f(x) = -x$ y $g(x) = x^3$.