Taller 3 teórico

Thursday, September 23, 2021 8:18 PM

1. ¿Es $f(x) = 4 + 3x$ convexa en $\mathbb{R}$? ¿Es cóncava en $\mathbb{R}$?

Funciones convexas: si $\forall x_1, x_2 \in \Omega$, $0 \le \alpha \le 1$:

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2)$$

aplicando la definición:

$$3(\alpha x_1 + (1-\alpha)x_2) + 4 \le \alpha(4 + 3x_1) + (1-\alpha)(4+3x_2)$$

$$\cancel{3\alpha x_1} + \cancel{3x_2} - \cancel{3\alpha x_2} \le \cancel{4\alpha} + \cancel{3\alpha x_1} + 4 + \cancel{3x_2} - \cancel{4\alpha} - \cancel{3\alpha x_2}$$

$$0 \le 4$$

$f(x) = 4 + 3x$, es convexa

---

2. ¿Es $f(x) = x^2 - 2x$ convexa en $\mathbb{R}$?

$$\nabla f(x) = \begin{bmatrix} 2x - 2 \end{bmatrix}, \quad H(f(x)) = \begin{bmatrix} 2 \end{bmatrix}$$

dado que la derivada 2da es toda positiva $f(x)$ es convexa

---

3. ¿Es $f(x) = x^3$ convexa en $\mathbb{R}$? ¿Es convexa en $\mathbb{R}_+$?

$$\frac{\delta f(x)}{\delta x} = 3x^2, \quad \frac{\delta(3x^2)}{\delta x} = 6x$$

Vea que si $x < 0$, la segunda derivada deja de ser positiva.

$f(x) \rightarrow$ no es convexa.

$f(x) \to$ no es convexa.

4. ¿Es $f(x) = (x-3)^2 - 2$ convexa en $\mathbb{R}$?

$$\frac{\delta f(x)}{\delta x} = 2(x-3) = 2x-6$$

$$\frac{\delta (2x-6)}{\delta x} = 2$$

$f(x) \to$ es convexa.

5. ¿Sobre qué región de $\mathbb{R}$ es $f(x) = x^2(x^2-1)$ convexa? ¿Es estrictamente convexa sobre esta región?

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

$$\frac{\delta f(x)}{\delta x} = 2x^3 - 2x \quad , \quad \frac{\delta (2x^3-2x)}{\delta x} = 6x^2 - 2$$

No es estrictamente convexa en la región.

6. Determine cuáles de las siguientes funciones son convexas, cóncavas o ninguna.

(a)

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

a

$$\nabla f(x_1, x_2) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$$

$$H(x_1, x_2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

esta es convexa

(b) $$f(x_1, x_2) = 2x_1^2 - 4x_1x_2 - 8x_1 + 3x_2$$

Proposition 5. *Let $f \in C^2$. Then f is convex over a convex set Ω containing an interior point if and only if the Hessian matrix* **F** *of f is positive semidefinite throughout Ω.*

Definition. A function g defined on a convex set Ω is said to be *concave* if the function $f = -g$ is convex. The function g is *strictly concave* if $-g$ is strictly convex.

$$\nabla f(x_1, x_2) = \begin{bmatrix} 4x_1 - 4x_2 - 8 \\ 4x_1 + 3 \end{bmatrix}$$

$$H(x_1, x_2) = \begin{bmatrix} 4 & -4 \\ 4 & 0 \end{bmatrix}$$

→ No es convexa

$$\nabla -f(x_1, x_2) = \begin{bmatrix} -4x_1 + 4x_2 + 8 \\ -4x_1 - 3 \end{bmatrix}$$

$$H(x_1, x_2) = \begin{bmatrix} -4 & 4 \\ -4 & 0 \end{bmatrix}$$

→ No es cóncava

(c) $$f(x_1, x_2) = x_1 e^{-(x_1+3x_2)}$$

$$\nabla f(x_1, x_2) = \Big[e^{-(x_1+3x_2)} \quad - x_1 e^{-(x_1+3x_2)} \Big]$$

$$\nabla f(x_1, x_2) = \begin{bmatrix} e^{-(x_1+3x_2)} - x_1 e^{-(x_1+3x_2)} \\ -3x_1 e^{-(x_1+3x_2)} \end{bmatrix}$$

$$H f(x_1,x_2) = \begin{bmatrix} (x_1-2) e^{-x_1-3x_2} & 3(x_1-1) e^{-x_1-3x_2} \\ 3(x_1-1) e^{-x_1-3x_2} & 9x_1 e^{-x_1-3x_2} \end{bmatrix}$$

$$H f(x_1,x_2) = \begin{bmatrix} (x_1-2) e^{-(x_1+3x_2)} & 3(x_1-1) e^{-(x_1+3x_2)} \\ 3(x_1-1) e^{-(x_1+3x_2)} & 9x_1 e^{-(x_1+3x_2)} \end{bmatrix}$$

→ No es Ni concava ni convexa.

(d)

$$f(x_1, x_2) = -x_1^2 - 3x_2^2 + 4x_1x_2 + 10x_1 - 10x_2$$

$$\nabla f(x_1,x_2) = \begin{bmatrix} -2x_1 + 4x_2 + 10 \\ -6x_2 + 4x_1 - 10 \end{bmatrix}$$

$$H(x_1,x_2) = \begin{bmatrix} -2 & 4 \\ 4 & -6 \end{bmatrix}$$

→ No es convexa ni concava.

7. Demuestre si la siguiente función, definida sobre el conjunto S,

$$f(x_1, x_2) = 10 - 3(x_2 - x_1^2)^2$$

es cóncava o no.

(a) Considere $S = \{(x_1, x_2) : -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1\}$

(b) Considere $S = \{(x_1, x_2) : x_1^2 \geq x_2\}$

$$\nabla f(x_1, x_2) = \begin{bmatrix} 12x_1 \cdot (x_2 - x_1^2) \\ -6(x_2 - x_1^2) \end{bmatrix}$$

$$H(x_1, x_2) = \begin{bmatrix} -12(x_2 - 3x_1^2) & -12x_1 \\ -12x_1 & 6 \end{bmatrix}$$

Det>0

$$-12(x_2 - 3x_1^2) > 0$$

$$-12x_2 + 36x_1^2 > 0$$

$$12x_2 < 36x_1^2$$

$$x_2 < 3 \cdot x_1^2$$

Det > 0

$$-72(x_2 - 3x_1^2) - 24x_1^2 > 0$$

$$-72x_2 - 216x_1^2 - 24x_1^2 > 0$$

$$-72x_2 - 240x_1^2 > 0$$

$$-240x_1^2 > 72x_2$$

$$240x_1^2 < -72x_2$$

→ vea que (b) no es cóncava porque $-Hf(x_1, x_2)$ no es convexa.

→ analizando los determinantes en (a) $f(x_1, x_2)$ es cóncava porque $-Hf(x_1, x_2)$ es convexa.

8. Considere la función $f : \mathbb{R}^3 \to \mathbb{R}$ $f(x) = x'Ax$ donde

$$A = \begin{bmatrix} 2 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 2 & \theta \end{bmatrix}.$$

Determine la Hesiana de f (en cualquier punto de $\mathbb{R}^3$). ¿Para qué valores de θ es f estrictamente convexa?

11. Dado el siguiente problema de minimización,

$$\text{mín } f(x) = 3x_1^2 - x_2^2 + x_1 + 2x_2$$

(a) Encuentre la derivada direccional de f en el punto $\hat{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ en la dirección $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

(b) Encuentre todos los puntos que satisfagan la condición necesaria de primer orden para un mínimo local para f.

(c) Determine el conjunto de puntos $x \in \mathbb{R}^2$ que satisfacen la condición necesaria de segundo orden para un mínimo local. ¿Tiene esta función mínimos locales?

a) la derivada direccional es:

$$f'(\hat{x};d) = \lim_{\lambda \to 0} \frac{f(\hat{x} + \lambda d) - f(\hat{x})}{\lambda}$$

$\hat{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, d = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$$= \lim_{\lambda \to 0} \frac{(3\lambda^2 - \cancel{(1)^2} + \lambda + 2) - (\cancel{-(1)^2} + 2)}{\lambda}$$

$x = [1], d = [0]$

$z = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \cdot \lambda = \begin{bmatrix} \lambda \\ 1 \end{bmatrix}$

$$= \lim_{\lambda \to 0} \frac{(\ldots)}{\lambda}$$

$$= \lim_{\lambda \to 0} \frac{3\lambda^2 + \lambda}{\lambda} = \lim_{\lambda \to 0} 3\lambda + 1 = 1$$

lo anterior es lo mismo que

$$\nabla f = \begin{bmatrix} 6x_1 + 1 \\ -2x_2 + 2 \end{bmatrix}$$

$$d' \nabla f(\hat{x}) = [1 \ 0] \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1$$

b) Encuentre todos los puntos que satisfagan la condición necesaria de primer orden para un mínimo local para f.

$$\nabla f = \begin{bmatrix} 6x_1 + 1 \\ -2x_2 + 2 \end{bmatrix} \qquad H = \begin{bmatrix} 6 & 0 \\ 0 & -2 \end{bmatrix}$$

$$6x_1 + 1 = 0, \quad x_1 = -\frac{1}{6}$$

$$-2x_2 + 2 = 0, \quad x_2 = \frac{-2}{-2} = 1$$

$$x^* = \begin{bmatrix} -1/6 \\ 1 \end{bmatrix}$$

Analicemos la condición necesaria de primer orden que dice:

Teorema

$\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, $f \in \mathcal{C}^1$ (primera derivada continua). Si x^* es un mínimo local de f en Ω, entonces, para toda dirección factible d en x^*,

$$d' \nabla f(x^*) \geq 0$$

$$d = \begin{bmatrix} \lambda \\ \theta \end{bmatrix}, \quad [\lambda \ \ \theta] \nabla f(x^*)$$

$$= [\lambda \ \ \theta] \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$= \lambda(0) + \theta(0) = 0$$

c) Determine el conjunto de puntos $x \in \mathbb{R}^2$ que satisfacen la condición necesaria de segundo orden para un mínimo local. ¿Tiene esta función mínimos locales?

Teorema

$\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, $f \in \mathcal{C}^2$. *Si x^* es un mínimo local de f en Ω, y d una dirección factible en x^* tal que $d'\nabla f(x^*) = 0$, entonces*

$$d' H(x^*) d \geq 0$$

Corolario

$\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$, $f \in \mathcal{C}^2$. *Si x^* es un mínimo local de f en Ω y x^* es un punto interior, entonces*

$$\nabla f(x^*) = 0 \quad y \quad H(x^*) \text{ es semidefinida positiva}$$

$$x^* = \begin{bmatrix} -1/6 \\ 1 \end{bmatrix}, \quad d = \begin{bmatrix} \lambda \\ \theta \end{bmatrix}, \quad H(x) = \begin{bmatrix} 6 & 0 \\ 0 & -2 \end{bmatrix}$$

$$d'\left(\begin{bmatrix} 6 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} \lambda \\ \theta \end{bmatrix} \right) \geq 0$$

$$= [\lambda \ \ \theta] \begin{bmatrix} 6\lambda \\ -2\theta \end{bmatrix} = 6\lambda^2 - 2\theta^2 \geq 0$$

Con esto:

$$6\lambda^2 \geq 2\theta^2$$
$$3\lambda^2 \geq \theta^2$$
$$\sqrt{3}|\lambda| \geq |\theta|$$

d que satisface la condición necesaria de 2do orden si es $d = \begin{bmatrix} \lambda \\ \theta \end{bmatrix}$, $\sqrt{3}|\lambda| \geq |\theta|$

Sea x^* un mínimo $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\nabla f(x^*) = \begin{bmatrix} 6x_1 + 1 \\ -2x_2 + 2 \end{bmatrix}$

$$d' \nabla f(x^*) = [\lambda \ \ \theta] \begin{bmatrix} (6x_1 + 1) \\ (-2x_2 + 2) \end{bmatrix} = 0$$

$$d'\nabla f(x^*) = [\lambda \quad \theta] \begin{bmatrix} (6x_1+1) \\ (-2x_2+2) \end{bmatrix} = 0$$

$$= \lambda(6x_1+1) + \theta(-2x_2+2) = 0$$

$$\lambda(6x_1+1) + \sqrt{3}\lambda(-2x_2+2) = 0$$

$$(6x_1+1) + \sqrt{3}(-2x_2+2) = 0$$

$$6x_1 = 2\sqrt{3}x_2 - 2\sqrt{3} - 1$$

$$x_1 = \frac{\sqrt{3}}{3}x_2 - \frac{\sqrt{3}}{3} - \frac{1}{6}$$

los puntos de la forma $x = \begin{bmatrix} \frac{\sqrt{3}}{3}(x_2-1) - \frac{1}{6} \\ x_2 \end{bmatrix}$

son mínimos locales ya que fueron creados a partir de la condición necesaria de 2do orden y la primera.

---

12. Suponga que tiene n números reales $x_1, x_2, \ldots, x_n$. Encuentre el número $\bar{x} \in \mathbb{R}$ que minimiza la suma de diferencias al cuadrado entre los números dados y $\bar{x}$. Es decir, determine una expresión para $\bar{x}$ en términos de los números dados.

$$\min \sum_{i=1}^{n} (x_i - \bar{x})^2 = \underbrace{[1 \cdots 1]}_{\text{Costos}} \begin{bmatrix} (x_i - \bar{x})^2 \\ \vdots \\ (x_{i+n} - \bar{x})^2 \end{bmatrix}$$

$$\nabla f = \begin{bmatrix} 2(x_i - \bar{x}) \\ \vdots \\ 2(x_{i+n} - \bar{x}) \end{bmatrix}, \quad H = \begin{bmatrix} 2 & 0 & \cdots & 0 \\ 0 & 2 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & & & 2 \end{bmatrix}$$

$$= 2 \cdot \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$= 2 \cdot I_{n \times n}$$

Vea que dado que H es semidefinida positiva, la función es convexa por lo que cuando minimice el gradiente minimizo también

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 \doteq \sum (x_i - \bar{x})^2$$

$$\frac{\delta \sum^{n} (x_i - \bar{x})^2}{\delta x_i} = \sum_{i=1}^{n} \frac{\delta (x_i - \bar{x})^2}{\delta x_i}$$

$$= \sum_{i=1}^{n} 2(x_i - \bar{x})$$

$$= \sum^{n} 2x_i - \sum^{n} 2\bar{x}$$

$$= 2\sum^{n} x_i - 2n\bar{x} \quad \Big] g$$

trataremos de minimizar g, vea que en el mejor de los casos esta se hace cero luego:

$$\min \quad 2\sum^{n} x_i - 2n\bar{x} = 0$$

$$\cancel{2}\sum^{n} x_i = \cancel{2}n\bar{x}$$

$$\sum^{n} \frac{x_i}{n} = \bar{x}$$

Así $\bar{x}$ debe ser $\sum^{n} \frac{x_i}{n}$. es decir el promedio.

13. Se quiere minimizar la función $f(x) : \mathbb{R}^2 \to \mathbb{R}$ donde $f(x) = 4x_1^2 - x_2^2$.

(a) ¿Satisface el punto $\hat{x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ la condición necesaria de primer orden?