Tutorial-5

Saturday, 5 February 2022 10:19 AM

Q) Determine whether the following function is Convex :-

$$f(x) = \underbrace{(x_1 - 4x_2)^4}_{\text{convex}} + e^{\underbrace{(9x_1^2 + 6x_1x_2 + 4x_2^2)}_{\text{convex}}}$$

$e^x \to$ convex

let $f(x) = f_1(x) + f_2(x)$

$f_1(x) = (x_1 - 4x_2)^4$

$g_1(x) = x^4 \quad (\mathbb{R} \to \mathbb{R})$

$h_1(x) = x_1 - 4x_2 \quad (\mathbb{R}^2 \to \mathbb{R})$

affine (Both convex Concave)

$g_1'(x) = 4x^3$

$g_1''(x) = 12x^2 \geq 0$ (convex)

$\nabla h_1(x) = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$

$\nabla^2 h_1(x) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ convex

$f_1(x) = g_1 \circ h_1(x)$ convex

$f_2(x) = e^{(9x_1^2 + 6x_1x_2 + 4x_2^2)}$

$g_2(x) = e^x \quad (\mathbb{R} \to \mathbb{R})$

$g_2'(x) = e^x; \quad g_2''(x) = e^x > 0$

convex

$h_2(x) = 9x_1^2 + 6x_1x_2 + 4x_2^2 \quad (\mathbb{R}^2 \to \mathbb{R})$

$9x_1^2 + 3x_1x_2 + 3x_1x_2 + 4x_2^2$

$\nabla^2 H = x^T A x \quad ; \quad A$

$= \begin{bmatrix} 9x_1 + 3x_2 & 3x_1 + 4x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$A = [\quad]$

$h_2(x) \cong x^T A x$

$= [x_1, x_2] \begin{bmatrix} 2.9 & 2.3 \\ 2.3 & 2.4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$h_2(x) \cong (x^T A x)$ — $= [x_1, x_2] \begin{bmatrix} 2.9 & 2.3 \\ 2.3 & 2.4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\nabla^2 h_2(x) = 2A$

$\begin{bmatrix} 18 & 6 \\ 6 & 8 \end{bmatrix}$ convex

$A^1 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$\text{tra}(A) = a + d < 0$ —ve

$|A^1| > 0$ $\rightarrow$ +ve, −ve

$$f(x) = -(x_1 - 4x_2)^4 + e^{(9x_1^2 + 6x_1x_2 + 4x_2^2)}$$

concave f_1 + convex f_2 (Cannot say if it is convex or concave need to check)

$\nabla^2 f(x) = \nabla^2 f_1(x) + \nabla^2 f_2(x)$

$$\nabla^2 f(x) \quad (x_1 - 4x_2)^2 \begin{bmatrix} -12 & 48 \\ 48 & -192 \end{bmatrix} + e^{9x_1^2 + 6x_1x_2 + 4x_2^2} \begin{bmatrix} 18x_1 + 6x_2 + 18 & 6x_1 + 8x_2 + 6 \\ 6x_1 + 8x_2 + 6 & 6x_1 + 8x_2 + 8 \end{bmatrix}$$

at $(x_1, x_2) = (-1, 1)$ $\quad x_1 = -1 \;;\; x_2 = 1$

$\nabla^2 f(x)\big|_{(-1,1)} < 0 \quad \therefore$ not convex

Q) find local minima, for the following functions in the interval $[0.3, 1]$?

$f(x) = \underbrace{3x^4 + 5x^3 - 4x^2 + 2}_{\text{convex}}$ Global mini ??

$f'(x) = 12x^3 + 15x^2 - 8x = 0$ $\quad x = 0$ (stationary)

$12x^2 + 15x - 8 = 0$

$f''(x) = 36x^2 + 30x - 8$ $\quad > 0$ min, < 0 max

$x_1 = \cdot$
$x_2 = \cdot$

$$x = \frac{-30 \pm \sqrt{900 + 4(36)(8)}}{2 \cdot 36}$$

$x = -0.5$

$x=0 \Rightarrow = -8$

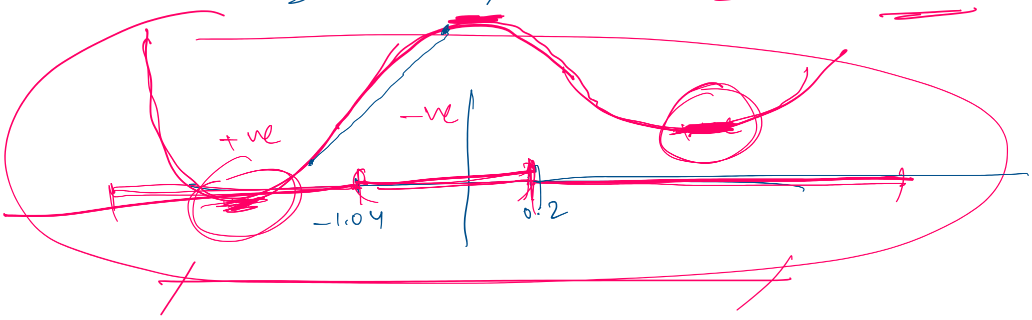
$9 - 15 - 8 < 0$

$x=0$ $\quad f''(x) < 0$

$9 + 15 - 8 < 0$ $\quad -8 \quad$ 2.36

$x=0$ local max $\quad f''(x) < 0$ local min

$= -1.04, \; 0.2$



TUTORIAL 5

Q1. <u>Tutorial 4</u> :

$$f : (-\infty, \infty) \to (-\pi/2, \pi)$$

$$f(x) = \tan^{-1}(x)$$

Lipschitz gradient exists

$f : S \to \mathbb{R}$ $\quad C^1$

Defn

$$\| \nabla f(y) - \nabla f(x) \|_2 \leq L \| y - x \|_2$$

for some constant L

$$f(x) = \tan^{-1}(x)$$

$$\frac{d f(x)}{dx} = \frac{1}{1+x^2}$$

$x \in \mathbb{R}$

$$\left\| \frac{df}{dx}(y_2) - \frac{df}{dx}(y_1) \right\| \le \underline{L} \, \| y_2 - y_1 \|$$

M.V.T $\quad (y, x) \quad z \quad \forall (x,y) \in \mathbb{R}$

$$f'(z) = \frac{f(y) - f(x)}{y - x}$$

$$\left[\frac{1}{1+y^2} - \frac{1}{1+x^2} \right] = (y - x) \; f'(z)$$

L.H.S $\quad \le L \, |y - x|$

$$\left| \frac{1}{1+y^2} - \frac{1}{1+x^2} \right| = |f'(z)|\,|y-x| \rightarrow (1)$$

$z \in \mathbb{R}$

$f''(z)$

$$\max_{z \in \mathbb{R}} f''(z)$$

$$\max_{z \in \mathbb{R}} \left(\frac{-1(2x)}{(1+x^2)^2} \right) = g(x)$$

first order - necessary - condition,

$g'(x) = 0$

$$g'(x) = -2 \left[\frac{(1+x^2)^2(1) - x(2)(1+x^2)(2x)}{(1+x^2)^4} \right] \quad x^*$$

$$g'(x) = 0 \Rightarrow \frac{-2}{(1+x^2)^4} \left[(1+x^2)\left(-4x^2 + x^2 + 1\right)\right]$$

$$\boxed{x = \pm 1/\sqrt{3}}$$

second-order-sufficient-condition

$$g''(x) > 0 \quad \text{(Local minima)}$$

$$< 0 \quad \text{(maxima)}$$

$$g''(x) = -2\left[\frac{(1+x^2)^4\left[(1+x^2)(-6x) + (1-3x^2)(2x)\right] - (1+x^2)(1-3x^2)(4)(1+x^2)^3(2x)}{(1+x^2)^8}\right]$$

$$g''(x) = \frac{-2}{(1+x^2)^8}\left[(1+x^2)^4\left\{(1+x^2)(-6x) + (1-3x^2)(2x) - (1-3x^2)(4)(2x)\right\}\right]$$

$$= \frac{-2}{(1+x^2)^4}\left[-6x - 6x^3 + (1-3x^2)(-6x)\right]$$

$$= -\frac{2}{(1+x^2)^4}\left(-12x - 6x^3 + 18x^3\right)$$

$$= \frac{-2}{(1+x^2)^4}\left(12x^3 - 12x\right)$$

$$= \frac{-2}{(1+x^2)^4}\, 12x(x^2-1)$$

$$< 0$$

$$g'\left(x^* = \frac{-1}{\sqrt{3}}\right)$$

$x = \frac{-1}{\sqrt{3}}$ is a maxima.

$$f''\left(x^* = -\frac{1}{\sqrt{3}}\right) = \frac{-2x - 1/\sqrt{3}}{(1+1/3)^2} = \frac{2}{\sqrt{3}} \div \frac{16}{9} = \frac{18}{16\sqrt{3}} = \frac{9}{8\sqrt{3}}$$

$$\left|\frac{1}{1+y^2} - \frac{1}{1+x^2}\right| \leq \frac{9}{8\sqrt{3}} \cdot \left(|y-x|\right)$$

Lipschitz - gradient exists.

Q Tutorial 3 :

$$-(x_1^2 + x_2^2)$$

Q1. (c) $f(x) = x_1 x_2 e^{-(x_1^2 + x_2^2)}$

$f : \mathbb{R}^2 \to \mathbb{R}$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} x_2 \left[x_1 e^{-(x_1^2+x_2^2)} (-2x) + e^{-(x_1^2+x_2^2)} (1) \right] \\ x_1 \left[x_2 e^{-(x_1^2+x_2^2)} (2x_2) + e^{-(x_1^2+x_2^2)} (1) \right] \end{bmatrix}$$

$$= \begin{bmatrix} x_2 (1 - 2x_1^2) e^{-(x_1^2+x_2^2)} \\ (1 - 2x_2^2) x_1 e^{-(x_1^2+x_2^2)} \end{bmatrix}$$

$$\nabla f = \begin{bmatrix} 0 \\ \end{bmatrix}$$

$$0$$

$$\begin{bmatrix} (1-2x_1^2)\, x_2\, e^{-(x_1^2+x_2^2)} \\ (1-2x_2^2)\, x_1\, e^{-(x_1^2+x_2^2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$x_2(1-2x_1^2) = 0$$

$$x_1(1-2x_2^2) = 0 \leftarrow 2$$

$x_1 = 0,\ x_2 = 0$

$$\left.\begin{aligned} x_1^2 &= 1/2 \\ x_2^2 &= 1/2 \end{aligned}\right]$$

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} -1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

→ <u>local minima or maxima?</u>

$$x_2(1-2x_1^2) = 0 \rightarrow (1)$$

$$x_1\underbrace{(1-2x_2^2)} = 0 \quad -$$

$$x_2 = 0 \qquad\qquad x_1^2 = \cancel{\ }\ 1/2$$

$$x_1 = 0 \qquad\qquad x_2^2 = 1/2$$

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

$$\frac{\partial^2 f}{\partial x_1^2} = \left[x_2 \left[(1-2x_1^2)\, e^{-(x_1^2+x_2^2)}(-2x_1) + e^{-(x_1^2+x_2^2)}(-4x_1) \right] \right]$$

$$\partial^2 f \quad = \quad e^{-(x_1^2+x_2^2)} ([illegible])(2x_1^2 - 3)$$

$$\frac{\partial f}{\partial x_1^2} = e \;\ldots\; (2x_1 x_2)(2x_1 \ldots)$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = (1 - 2x_1^2)(1 - 2x_2^2)\, e^{-(x_1^2 + x_2^2)}$$

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

$$\frac{\partial^2 f}{\partial x_2^2} = 2x_1 x_2\, e^{-(x_1^2 + x_2^2)} (2x_2^2 - 3)$$

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

$$\left(\frac{1}{\sqrt{2}},\ \frac{-1}{\sqrt{2}} \right)$$

$$\nabla^2 f = \begin{bmatrix} 2/e = \lambda_1 & 0 \\ 0 & 2/e = \lambda_2 \end{bmatrix}$$

> 0

second-order condition,
$\{$<u>Local minima</u>$\}$

Q. <u>Tutorial 3</u>:

Q.5. (b) convex, strictly convex, strongly convex $\{x^T A x\}$

$f: \mathbb{R}^2 \to \mathbb{R}$

$$f(x) = (x_1 - 3x_2)^2 + (x_1 - 2x_2)^2$$

$$f(x) = x_1^2 - 6x_1x_2 + 9x_2$$
$$+ x_1^2 - 4x_1x_2 + 4x_2^2$$

$$= 2x_1^2 - 10x_1x_2 + 13x_2^2$$

$$= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & -5 \\ -5 & 13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$f(x) = x^T A X$$

A is symmetric

$(A^T = A$

$$\nabla f(x) = 2AX$$

$$\nabla^2 f(x) = 2A$$

$$= \begin{bmatrix} 4 & -10 \\ -10 & 26 \end{bmatrix}$$

P.D strictly convex Defn

$$\text{eigen values } \underbrace{\begin{bmatrix} 4 & -10 \\ -10 & 26 \end{bmatrix}}_{\nabla^2 f(x)} > 0$$

<u>strong convexity</u>

Q. can we find an $m > 0$

$$\nabla^2 f(x) - mI \succeq 0$$

$$\begin{bmatrix} 4 & -10 \\ -10 & 26 \end{bmatrix} - \begin{bmatrix} m & 0 \\ 0 & m \end{bmatrix} \geq 0$$

$$\begin{bmatrix} 4-m & -10 \\ -10 & 25-m \end{bmatrix} \qquad \{m = 0.1\}$$

$\boxed{f(x) \text{ is strongly convex with parameter } m = 0.1}$

Q.) Tutorial 1

(Q2.) verify the parallelogram-Law,

$$\|x+y\|_2^2 + \|x-y\|_2^2 = 2(\|x\|_2^2 + \|y\|_2^2)$$

$$\|x\|_2^2 = x^T x$$

L.H.S

$$(x+y)^T(x+y) + (x-y)^T(x-y)$$

$$= x^T x + \cancel{x^T y} + \cancel{y^T x} + y^T y$$

$$+ \; x^T x - \cancel{x^T y} - \cancel{y^T x} + y^T y$$

$$= 2\left\{ x^T x + y^T y \right\}$$

$$= 2\left\{ \|x\|_2^2 + \|y\|_2^2 \right\}$$

<u>proved.</u>

Q. Tutorial $\quad \nabla(x^T A x) = (A + A^T)x,$

A is not symmetric

$$A = \begin{bmatrix} \leftarrow a_1^T \rightarrow \\ \leftarrow a_2^T \rightarrow \\ \end{bmatrix} \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

$$\begin{bmatrix} \leftarrow a_n^T \rightarrow \end{bmatrix} \in \mathbb{R}^{n \times n}$$

$$Ax = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_n^T x \end{bmatrix} \equiv \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ x_n \end{pmatrix}$$

$$x^T A x = \underbrace{x_1 (a_1^T x)} + x_2 (a_2^T x) + \cdots x_n (a_n^T x)$$

$$\nabla (x^T A x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

$$\begin{bmatrix} \dfrac{\partial f}{\partial x_1} = x_1 \left[\underbrace{a_{11}}\right] + a_1^T x + x_2 (a_{21}) + \cdots x_n (a_{n1}) \\ \dfrac{\partial f}{\partial x_2} = x_1 (a_{12}) + x_2 (a_{22}) + a_2^T x + \cdots x_n a_{n2} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} = x_1 (a_{1n}) + x_2 a_{n2} + \cdots + a_n^T x + x_n a_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} x_1 + a_1^T x + x_2 a_{21} + \cdots x_n a_{n1} \\ x_1 (a_{12}) + x_2 a_{22} + a_2^T x + \cdots x_n a_{n2} \\ \\ x_1 a_{1n} + x_2 a_{n2} + \cdots a_n^T x + x_n a_{nn} \end{bmatrix} = \nabla f$$

$$= \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ \vdots \end{bmatrix} + \begin{bmatrix} a_{11} x_1 + x_2 a_{21} + \cdots x_n a_{n1} \\ a_{12} x_1 + x_2 a_{22} + \cdots x_n a_{n2} \\ \\ \cdots x_n a_{nn} \end{bmatrix}$$

$$\begin{bmatrix} \vdots \\ a_n^T x \end{bmatrix} \qquad \begin{bmatrix} \vdots \\ a_{1n} x_1 + x_2 a_{n2} + \cdots \end{bmatrix}$$

$$a_1^T = (a_{11}, a_{12}, \cdots, a_{1n})$$

$$= Ax + \begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ a_1^T & a_2^T & \cdots & a_n^T \\ \downarrow & \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= Ax + A^T x$$

$$\nabla f = \underbrace{(A + A^T)}_{A^T = A} x$$

$$= 2Ax$$