Assignment 1   Zonghao Li   # 1003843148

1. $f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x)$

$\nabla f(x) = \frac{\partial f}{\partial x} = -\sum_{i=1}^{m}\left[\frac{1}{b_i - a_i^T x}(-a_i)\right] = \sum_{i=1}^{m} \frac{a_i}{b_i - a_i^T x}$

$\nabla^2 f(x) = \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left[\sum_{i=1}^{m} \frac{a_i}{b_i - a_i^T x}\right] = \sum_{i=1}^{m} \frac{\partial}{\partial x}\left[a_i (b_i - a_i^T x)^{-1}\right]$

$= \sum_{i=1}^{m} a_i (-1)(b_i - a_i^T x)^{-2} \cdot (-a_i^T) = \sum_{i=1}^{m} a_i (b_i - a_i^T x)^{-2} a_i^T$

$\therefore f(x_0) = f(x_0) + \nabla f(x_0)^T (x - x_0) + \frac{1}{2}(x - x_0)^T \nabla^2 f(x)(x - x_0)$

$$\boxed{= -\sum_{i=1}^{m} \underbrace{\log(b_i - a_i^T x)}_{1\times 1} + \underbrace{\sum_{i=1}^{m} \frac{a_i^T}{b_i - a_i^T x}}_{1\times n} \cdot \underbrace{(x - x_0)}_{n\times 1} + \frac{1}{2}\underbrace{(x - x_0)^T}_{1\times n} \underbrace{\sum_{i=1}^{m} \underbrace{a_i}_{n\times 1} \underbrace{(b_i - a_i^T x)^{-2}}_{1\times 1} \underbrace{a_i^T}_{1\times n}}_{n\times n} \cdot \underbrace{(x - x_0)}_{n\times 1}}$$

2. $P_1 = \{x \in R^n \mid a^T x = b_1\}$, $P_2 = \{x \in R^n \mid a^T x = b_2\}$

pick point on hypeplane P_1 as x_1, the line that pass through P_1, intersect x_1 and normal to P_1 is: $x_2 = x_1 + at$, where $t \in R$; it intersect with P_2:

$\therefore \quad a^T(x_1 + at) = b_2 \Rightarrow t = \frac{b_2 - a^T x_1}{a^T a} = \frac{b_2 - b_1}{a^T a}$

$\therefore$ the intersect point of P_2 is: $x_2 = x_1 + a\frac{b_2 - b_1}{a^T a}$.

$$\therefore \boxed{\|x_2 - x_1\|_2 = \|a\|_2 \frac{|b_2 - b_1|}{\|a\|_2^2} = \frac{|b_2 - b_1|}{\|a\|_2}}$$

Scanned with CamScanner

3. $S_a = \{ x \mid \inf_{y \in S} \|x-y\| \le a \}$, where: $a \geq 0$.

therefore, ~~by~~ the def of convexity:, for: $\theta \in [0,1]$, $x_1, x_2 \in S_a$:

$$\inf_{y\in S} \|\theta x_1 + (1-\theta) x_2 - y\| = \inf_{y \in S} \|\theta (x_1 - y) + (1-\theta)(x_2-y)\|$$

$$\le \theta \inf_{y\in S} \|x_1 - y\| + (1-\theta) \inf_{y \in S} \|x_2 - y\| \le a.$$

therefore, $x_1 + (1-\theta)x_2 \in S_a$, therefore, S_a is a convex set.

4.

<a>. $\nabla_x^2 f(x,z) = \dfrac{\partial^2 f}{\partial x^2} \geq 0$ for convexity of x for each fixed z.

$\nabla_z^2 f(x,z) = \dfrac{\partial^2 f}{\partial z^2} \le 0$ for concavity of z for each fixed x.

<b>. when $z = \bar{z}$ (fixed), since $\nabla_x^2 f(x,z) \geq 0$,
there exists $x = \bar{x}$ so that:

for $\forall x$. ~~for $\forall x$.~~

$\exists x = \bar{x}$, so that $\{x \mid f(\bar{x}, \bar{z}) \le f(x, \bar{z})\}$ ①

same when $x = \bar{x}$ (fixed), since $\nabla_z^2 f(x,z) \le 0$, there exists $z = \bar{\bar{z}}$ so that.

$\exists z = \bar{\bar{z}}$, so that $\{z \mid f(\bar{x}, \bar{\bar{z}}) \geq f(\bar{x}, z)\}$. ② for $\forall z$.

where point $(\bar{x}, \bar{z})$ is the common point, or the saddle point.
if $\bar{z} = \bar{\bar{z}}$, in this case.

$\therefore$ for ①: $(\bar{x}, \bar{z}) = \{x, \bar{z} \mid \inf_x f(x, z)\}$.

for ②: $(\bar{x}, \bar{\bar{z}}) = \{x = \bar{x}, z \mid \sup_z f(x, z)\}$.

when: $\bar{z} = \bar{\bar{z}}$: $\sup_z \inf_x f(x,z) = \inf_x \sup_z f(x,z)$.

<c>. $f(\bar{x},\bar{z}) \le f(x,\bar{z})$, $\forall x$ with $f(x,z)|_{z=\bar{z}}$

$\therefore \nabla_x f(x,z)|_{z=\bar{z}} = 0$.

$f(\bar{x},z) \le f(\bar{x},\bar{z})$, $\forall z$, with $f(x,z)|_{x=\bar{x}}$

$\therefore \nabla_z f(x,z)|_{x=\bar{x}} = 0$

when: $x=\bar{x}$, $z=\bar{z}$, at point $(\bar{x},\bar{z})$:

$$(\bar{x},\bar{z}) = \left\{(x,z) \,\middle|\, \begin{cases}\nabla_x f(x,z)=0 \\ \nabla_z f(x,z)=0\end{cases}\right\} \quad \boxed{\therefore \nabla f(\bar{x},\bar{z})=0.}$$

5. <a>. since e^x is convex, its affine function $f(x)$ is also $\boxed{\text{convex}}$

<b>. $$\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 x_1} \\ \frac{\partial^2 f}{\partial x_1 x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = A.$$

for: $v \in R^2$, $v^T A v = 2v_1 v_2$ can be positive, negative, zero

~~[illegible]~~ However, since every superlevel set $\{x \mid f(x) = x_1 x_2 \ge 0\}$ is convex because it is positive semi-definite ~~[illegible]~~, by definition, it is $\boxed{\text{quasiconcave}}$

<c>. $$\nabla^2 f = \begin{bmatrix} 2x_1^{-3}x_2^{-1} & x_1^{-2}x_2^{-2} \\ x_1^{-2}x_2^{-2} & 2x_1^{-1}x_2^{-3} \end{bmatrix} = A, \quad \text{for } v \in R^2, \; v^T A v = \frac{1}{x_1 x_2} \cdot \frac{2(x_1 x_2 + x_2 x_1)^2}{x_1^2 x_2^2} \ge 0.$$

$\therefore \nabla^2 f$ is positive semidefinite.

$\therefore$ $\boxed{\text{it is } \cancel{\text{concave}}}$. convex

Scanned with CamScanner

6. let: $f \cdot g = h$.

(f convex $\Rightarrow f'' \geqslant 0$; f is positive).

$$\nabla^2 h = \underbrace{f''g}_{①} + \underbrace{2f'g'}_{②} + \underbrace{f \cdot g''}_{③}.$$

we know from the question: $f''g \geqslant 0$, $f \cdot g'' \geqslant 0$.

since f, g within an interval are nondecreasing (or nonincreasing)

$$\therefore \begin{cases} f' \geqslant 0 \\ g' \geqslant 0 \end{cases} \text{ or } \begin{cases} f' \leq 0 \\ g' \leq 0 \end{cases} \Rightarrow f'g' \geqslant 0$$

$\therefore \nabla^2 h \geqslant 0 \Rightarrow$ $\boxed{h \text{ is convex within this interval}}$

7.

$$f(x,y) = x^2 + y^2 + \beta xy + x + 2y.$$

$$\frac{\partial f}{\partial x} = 2x + \beta y + 1 = 0 \Rightarrow x = \frac{-1-\beta y}{2} \quad ①$$

$$\frac{\partial f}{\partial y} = 2y + \beta x + 2 = 0 \quad ②$$

$$① \to ②: \quad 2y + \beta \cdot \frac{-1-\beta y}{2} + 2 = 0 \Rightarrow y = \frac{\beta - 4}{4 - \beta^2}. \quad ③$$

$$③ \to ①: \quad x = \frac{2\beta - 2}{4 - \beta^2} \quad ④. \qquad (\beta \neq \pm 2)$$

put: ③, ④ back to $f(x,y)$

$$f(x,y) = f(\beta) = \frac{6}{b^2-4} - \frac{8b}{b^4-8b^2+16} \cdot \frac{3b}{b^2-4} + \frac{20}{b^4-8b^2+16} - \frac{5b^2}{b^4-8b^2+16} + \frac{2b^3}{b^4-8b^2+16}.$$

$$\frac{\partial f(\beta)}{\partial \beta} = 0 \Rightarrow \beta = 1 \pm \sqrt{3} j$$

Scanned with CamScanner

8. (a). since $m > 0$, therefore mI is positive definite with $\lambda = m$; since $\nabla^2 f(x) \succeq mI$, therefore $\nabla^2 f(x)$'s minimum eigenvalue $\lambda_n \geq m > 0$, therefore. $\nabla^2 f(x)$ is positive definite. Therefore, $f(x)$ is a strictly convex function.

(b) since $\nabla^2 f(x) > 0$, λ_n, the smallest eigenvalue of $\nabla^2 f(x)$, Is larger than zero to ensure it is positive definite. However, λ_n, has to be larger than m to ensure $\nabla^2 f(x) \succeq mI$, which is arbitrary. Therefore, a strictly convex function may not be a strongly convex function.

The distinction between convex, strictly convex, and strongly convex can be subtle at first glance. If f is twice continuously differentiable and the domain is the real line, then we can characterize it as follows:

- f convex if and only if $f''(x) \geq 0$ for all x.
- f strictly convex if $f''(x) > 0$ for all x (note: this is sufficient, but not necessary).
- f strongly convex if and only if $f''(x) \geq m > 0$ for all x.

For example, let f be strictly convex, and suppose there is a sequence of points (x_n) such that $f''(x_n) = \frac{1}{n}$. Even though $f''(x_n) > 0$, the function is not strongly convex because $f''(x)$ will become arbitrarily small.

9. <a> $y_i = ax_i^2 + bx_i + c + n_i \Rightarrow n_i = y_i - ax_i^2 - bx_i - c$

$n_i \sim N(0,1)$

$$\sim \frac{1}{\sqrt{2\pi\sigma^2}} \exp\frac{-n_i^2}{2\sigma^2} = \frac{1}{\sqrt{2\pi}} \exp\frac{-n_i^2}{2} = \frac{1}{\sqrt{2\pi}} \exp\frac{-(y_i - ax_i^2 - bx_i - c)^2}{2}$$

$$\therefore a^*, b^*, c^* = \underset{a,b,c}{\arg\max} \prod_{i=1}^{100} \frac{1}{\sqrt{2\pi}} \exp\frac{-(y_i - ax_i^2 - bx_i - c)^2}{2}$$

$$= \underset{a,b,c}{\arg\max} \frac{1}{\sqrt{2\pi}} \exp\frac{\sum_{i=1}^{100}[-(y_i - ax_i^2 - bx_i - c)^2]}{2}$$

$$= \underset{a,b,c}{\arg\min} \sum_{i=1}^{100} (y_i - ax_i^2 - bx_i - c)^2, \text{ let: } \sum_{i=1}^{100} (y_i - ax_i^2 - bx_i - c)^2 = f$$

$$\frac{\partial f}{\partial a} = \sum_{i=1}^{100}[2(y_i - ax_i^2 - bx_i - c)\cdot(-x_i^2)] = 0 \Rightarrow \sum_{i=1}^{100}[(y_i - ax_i^2 - bx_i - c)x_i^2] = 0$$

$$\Rightarrow \underbrace{\sum_{i=1}^{100} x_i^2 y_i}_{J1} - a\underbrace{\sum_{i=1}^{100} x_i^4}_{J2} - b\underbrace{\sum_{i=1}^{100} x_i^3}_{J3} - c\underbrace{\sum_{i=1}^{100} x_i^2}_{J4} = 0 \quad ①$$

$$\frac{\partial f}{\partial b} = \sum_{i=1}^{100}[2(y_i - ax_i^2 - bx_i - c)(-x_i)] = 0 \Rightarrow \sum_{i=1}^{100}[(y_i - ax_i^2 - bx_i - c)(x_i)] = 0$$

$$\Rightarrow \underbrace{\sum_{i=1}^{100} x_i y_i}_{K1} - a\underbrace{\sum_{i=1}^{100} x_i^3}_{K2} - b\underbrace{\sum_{i=1}^{100} x_i^2}_{K3} - c\underbrace{\sum_{i=1}^{100} x_i}_{K4} = 0 \quad ②$$

$$\frac{\partial f}{\partial c} = \sum_{i=1}^{100} 2(y_i - ax_i^2 - bx_i - c)(-1) = 0 \Rightarrow \sum_{i=1}^{100}(y_i - ax_i^2 - bx_i - c) = 0.$$

$$\Rightarrow \underbrace{\sum_{i=1}^{100} y_i}_{L1} - a\underbrace{\sum_{i=1}^{100} x_i^2}_{L2} - b\underbrace{\sum_{i=1}^{100} x_i}_{L3} - \underbrace{100}_{L4}c = 0 \quad ③.$$

$$\left.\begin{array}{l} ① \\ ② \\ ③ \end{array}\right\} \Rightarrow \quad a = \frac{J_1K_3L_4 - J_1K_4L_3 - J_3K_1L_4 + J_3K_4L_1 + J_4K_1L_3 - J_4K_3L_1}{J_2K_3L_4 - J_2K_4L_3 - J_3K_2L_4 + J_3K_4L_2 + J_4K_2L_3 - J_4K_3L_2}$$

$$b = -\frac{J_1K_2L_4 - J_1K_4L_2 - J_2K_1L_4 + J_2K_4L_1 + J_4K_1L_2 - J_4K_2L_1}{J_2K_3L_4 - J_2K_4L_3 - J_3K_2L_4 + J_3K_4L_2 + J_4K_2L_3 - J_4K_3L_2}$$

$$c = \frac{J_1K_2L_3 - J_1K_3L_2 - J_2K_1L_3 + J_2K_3L_1 + J_3K_1L_2 - J_3K_2L_1}{J_2K_3L_4 - J_2K_4L_3 - J_3K_2L_4 + J_3K_4L_2 + J_4K_2L_3 - J_4K_3L_2}.$$

Above can be done by Matlab code. (see attachment).

$$\Rightarrow \begin{cases} a = 0.012 \\ b = 3.129 \\ c = -42.355 \end{cases} \quad \therefore y_i = 0.012x_i^2 + 3.129x_i - 42.355 + n_i$$

<b>. $n_i \sim e^{-z}$

$$a^*, b^*, c^* = \underset{a,b,c}{\arg\max} \prod_{i=1}^{100} e^{-(y_i - ax_i^2 - bx_i - c)}$$

$$\Rightarrow \quad = \underset{a,b,c}{\arg\min} \sum_{i=1}^{100} y_i - ax_i^2 - bx_i - c$$

let objective function: $\underset{a,b,c}{\min} f = \underset{a,b,c}{\min} \left(-\sum_{i=1}^{100} x_i^2\right)a - \left(\sum_{i=1}^{100} x_i\right)b - 100c$

Subject to: $y_i - ax_i^2 - bx_i - c \geq 0$

$$\Rightarrow x_i^2 a + x_i b + c \leq y_i$$

since Matlab take "min" as objective and "$\leq$" as constraint.

by applying linear programming in Matlab:

$$\begin{cases} a = 0.0113 \\ b = 3.1734 \\ c = -153.6072 \end{cases}$$

See attached Matlab code and figures.