Homework 3

108B (5403) Machine Learning　王傳鈞　0416047

本次作業所使用到的程式碼，都已經上傳到 GitHub 當中，若有需要參考其詳細內容，歡迎點擊連結直接前往網頁瀏覽。例如：「Q6.m」的 MATLAB source code 代表是第六題所使用到的程式碼。

第一題

證明：$f: \mathbb{R}^n \to \mathbb{R}$ 是 strictly convex function。

若 x^* 是 local minimizer，則 x^* 是唯一的 global minimizer。

- $\forall x, y \in \mathbb{R}^n \;\; f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y)$ for any $\lambda \in (0,1)$
- $\exists \Omega \subset \mathbb{R}^n, \varepsilon > 0 \;\; f(x) \geq f(x^*)$ for all $x \in \Omega \setminus \{x^*\}$ and $\|x - x^*\|_2 < \varepsilon$

1. global minimizer

   (by contradiction)

   假設 x^*_{local} 是 local minimizer、x^*_{global} 是 global minimizer。

   $\Rightarrow f\left(x^*_{\text{global}}\right) < f(x^*_{\text{local}})$ ……………………………… (relation A)

   取 $x_\lambda = \lambda x^*_{\text{local}} + (1-\lambda)x^*_{\text{global}}$，其中 $\lambda \in (0,1)$ 且 $\|x_\lambda - x^*_{\text{local}}\| < \varepsilon$

   $\Rightarrow f(x^*_{\text{local}}) < f(x_\lambda) = f\left(\lambda x^*_{\text{local}} + (1-\lambda)x^*_{\text{global}}\right)$

   $< \lambda f(x^*_{\text{local}}) + (1-\lambda) f\left(x^*_{\text{global}}\right)$ ……………… (by convexity)

   $< \lambda f(x^*_{\text{local}}) + (1-\lambda) f(x^*_{\text{local}}) = f(x^*_{\text{local}})$ …. (by relation A)

   $\Rightarrow f(x^*_{\text{local}}) < f(x^*_{\text{local}}) \;\; \rightarrow\leftarrow$

2. uniqueness

   (by contradiction)

   假設 x^*_1 和 x^*_2 都是 global minimizer，且 $x^*_1 \neq x^*_2$。

   $\Rightarrow f(x^*_1) = f(x^*_2)$ ……………………………………… (relation B)

   取 $x_\lambda = \lambda x^*_1 + (1-\lambda)x^*_2$，其中 $\lambda \in (0,1)$

$$\Rightarrow f(x_\lambda) = f(\lambda {x^*}_1 + (1-\lambda){x^*}_2)$$
$$< \lambda f({x^*}_1) + (1-\lambda) f({x^*}_2) \cdots\cdots \text{(by convexity)}$$
$$< \lambda f({x^*}_1) + (1-\lambda) f({x^*}_1) = f({x^*}_1) \cdots\cdots \text{(by relation B)}$$
$$\Rightarrow\ f(x_\lambda) < \ f({x^*}_1)\ \rightarrow\leftarrow$$

第二題

- $\forall\, x, y \in \mathcal{F}, t \in [0,1]$，考慮 $tx + (1-t)y$

(a) 若 $t = 0 \text{ or } 1$，則：$tx + (1-t)y = x \text{ or } y \in \mathcal{F}$

(b) 若 $t \in (0,1)$，則：

已知 $\mathrm{g}(x) \le 0$ 且 $\mathrm{h}(x) = 0$、$\mathrm{g}(y) \le 0$ 且 $\mathrm{h}(y) = 0$

$\because g_i$ 是 convex for $i = 1,2,\dots,m$，h_i 是 linear for $j = 1,2,\dots,k$

$$\therefore \begin{cases} g_i(tx + (1-t)y) \le t g_i(x) + (1-t) g_i(y) \le 0 \text{ for } i = 1 \sim m \\ h_j(tx + (1-t)y) = t h_j(x) + (1-t) h_j(y) = 0 \text{ for } j = 1 \sim k \end{cases}$$

$\Rightarrow\ tx + (1-t)y \in \mathcal{F}$

- 綜合以上 (a) 與 (b) 所述，所以 $\mathcal{F}$ 是一個 convex set

0416047 王傳鈞

第三題

By Farkas' lemma，取 $A \in \mathbb{R}^{m\times(n+1)}$ 和 $\vec{b} = \vec{0} \in \mathbb{R}^{n+1}$

$$\Rightarrow \left\{\begin{array}{ll} \vec{\beta} \in \mathbb{R}^{n+1} & \text{such that } A\vec{\beta} \leq \vec{0} \\ \vec{\alpha} \in \mathbb{R}^{m} & \text{such that } A^T\vec{\alpha} = \vec{0} \text{ 且 } \vec{\alpha} \geq \vec{0} \text{ 且 } \vec{\alpha} \neq \vec{0} \end{array}\right\} \begin{array}{l}\text{only one} \\ \text{case has} \\ \text{a solution}\end{array}$$

令 $\vec{\beta} = \begin{bmatrix} \vec{x} \\ t \end{bmatrix}$，其中 $\vec{x} \in \mathbb{R}^n$、t 是 scalar (i.e. 1×1 matrix)

$A = \begin{bmatrix} B & \vec{w} \end{bmatrix}$，其中 $B \in \mathbb{R}^{m\times n}$、$\vec{w} \in \mathbb{R}^{m\times1}$

$\Rightarrow$ case (1)： $A\vec{\beta} \leq \vec{0} \Rightarrow B\vec{x} + t\vec{w} \leq \vec{0}$

$\Rightarrow B\vec{x} + z\vec{1} < B\vec{x} + t\vec{w} \leq \vec{0}$ for $z > 0$

$\Rightarrow B\vec{x} < \vec{0}$

case (2)： $A^T\vec{\alpha} = \vec{0}$ 且 $\vec{\alpha} \geq \vec{0}$ & $\vec{\alpha} \neq \vec{0}$

$\Rightarrow \begin{bmatrix} B^T\vec{\alpha} \\ \vec{w}^T\vec{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow B^T\vec{\alpha} = \vec{0}$ 且 $\vec{\alpha} \geq \vec{0}$ & $\vec{\alpha} \neq \vec{0}$

0416047 王傳鈞

第四題

pf :

$(\Rightarrow)$ 已知 $\vec{x} \in \mathbb{R}^n$ such that $A\vec{x} = \vec{b}$

$\therefore$ $\boxed{A\vec{x} \le \vec{b} \text{ 自然成立}}$

$\because$ $A\vec{x} = \vec{b}$ $\therefore$ $A\vec{x} = [(A\vec{x})_1 \ (A\vec{x})_2 \ \cdots\cdots \ (A\vec{x})_m]^T$

$= \vec{b} = [b_1 \ b_2 \ \cdots\cdots \ b_m]^T$

$\Rightarrow \sum_{i=1}^{m}(A\vec{x})_i = \sum_{i=1}^{m} b_i \Rightarrow \boxed{\vec{1}^T A\vec{x} = \vec{1}^T\vec{b}}$

$(\Leftarrow)$ 已知 $A\vec{x} \le \vec{b}$ 且 $\vec{1}^T A\vec{x} \ge \vec{1}^T\vec{b}$

$$\Rightarrow \begin{cases} (A\vec{x})_i \le b_i \text{ for } i = 1 \sim m & \text{——— ①} \\ \sum_{i=1}^{m}(A\vec{x})_i = \sum_{i=1}^{m} b_i & \text{——— ②} \end{cases}$$

對①同取 $\sum_{i=1}^{m}$: $\sum_{i=1}^{m}(A\vec{x})_i \le \sum_{i=1}^{m} b_i$

by ② $\therefore \sum_{i=1}^{m}(A\vec{x})_i = \sum_{i=1}^{m} b_i$

$\Rightarrow \vec{1}^T A\vec{x} = \vec{1}^T\vec{b}$

$\Rightarrow \langle A\vec{x}, \vec{1} \rangle = \langle \vec{b}, \vec{1} \rangle$ $\therefore$ $\boxed{A\vec{x} = \vec{b}}$

第五題

- Primal problem：$\underset{\vec{x}}{MAX}\ p^T\vec{x}$ subject to $\mathrm{A}\vec{x} = \mathrm{b}$
- Dual problem：$\underset{\vec{y}}{min}\ \mathrm{b}^{\mathrm{T}}\vec{y}$ subject to $\mathrm{A}^{\mathrm{T}}\vec{y} \ge p$

第六題

- 以下符號採用：$A=\begin{bmatrix}1 & 2\\ 2 & -1\\ 1 & 1\\ 4 & -1\\ 1 & -1\\ 1 & -2\end{bmatrix}$、$\vec{x}=\begin{bmatrix}x_1\\ x_2\end{bmatrix}$、$b=\begin{bmatrix}2\\ -2\\ 1\\ 1\\ -2\\ -1\end{bmatrix}$，並尋找 $\min\limits_{\vec{x}}\|A\vec{x}-b\|_p$

(a) $\min\limits_{\vec{x}}\|A\vec{x}-b\|_1$：

$\vec{x}^*=(0.4444, 0.7777)^T$、$\|A\vec{x}-b\|_1{}^*=4.1111$

(b) $\min\limits_{\vec{x}}\|A\vec{x}-b\|_2$：

$\vec{x}^*=(0.2381, 0.9524)^T$、$\|A\vec{x}-b\|_2{}^*=2.3401$

(c) $\min\limits_{\vec{x}}\|A\vec{x}-b\|_\infty$：

$\vec{x}^*=(0.2500, 1.2500)^T$、$\|A\vec{x}-b\|_\infty{}^*=1.2500$

- 以下繪圖六條方程式，與 $min\|A\vec{x}-b\|_1$、$min\|A\vec{x}-b\|_2$、$min\|A\vec{x}-b\|_\infty$ 各自的 $\vec{x}^*$ 座標點