MSBD 5004 Due: 28th Oct.

Assignment #3 Solution

Instructor: Jianfeng CAI Name: ID:

Question 1

(a) Let $\alpha=\beta=1$, $x=(0,1)$, $y=(1,0)$

then $f(\alpha x+\beta y) = \max_k(1,1) - \min_k(1,1) = 0$

$\alpha f(x)+\beta f(y) = [\max(0,1) - \min(0,1)] + [\max(1,0) - \min(1,0)]$

$= 1+1 = 2$

$f(\alpha x+\beta y) \neq \alpha f(x) + \beta f(y)$.

(b) Let $a = \text{diag}(-1, \underbrace{0, 0, \cdots, 0}_{(n-2) \text{ times.}}, 1)$

then $f(x) = x_n - x_1 = a^T x$.

Question 2

$\phi: \mathbb{R}^2 \to \mathbb{R} = f(x)+b$, where $f(x)$ is a linear function.

$\phi(1,0) = f(1,0)+b$. $\phi(1,-2) = f(1,-2)+b$.

(1) $\phi(1,-1) = f(1,-1)+b = \frac{1}{2}f(2,-2)+b = \frac{1}{2}[f(2,-2)+2b]$

$= \frac{1}{2}[f(1,0)+b+f(1,-2)+b] = 1.5$. $\Rightarrow f(2,-2) = 3-2b$.

(2) $\phi(2,-2) = f(2,-2)+b = 3-2b+b = 3-b$.

while b is unknown, we can't get $\phi(2,-2)$.

$\boxed{\text{Question 3}}$

(a) $E_{st}(X) = x_{st} = \begin{pmatrix} 0 & \cdots & & 0 \\ 0 & \cdots & & 0 \\ 0 & \cdots & 1 \cdots & 0 \\ 0 & \cdots & \uparrow \cdots & 0 \\ 0 & & \cdots & 0 \end{pmatrix} \cdot X = \langle A, X \rangle = A^T X.$

(s,t)-entry

$A = \begin{pmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ 0 & \cdots 1 \cdots & 0 \\ 0 & \cdots \uparrow \cdots & 0 \end{pmatrix}$

(t,s)-entry.

(b) $H(X) = \sum_{i+j=n+1} x_{ij} = x_{1n} + x_{2,n-1} + x_{3,n-2} + \cdots + x_{n-1,2} + x_{n,1}$

$= \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix} \cdot X = \langle A, X \rangle = A^T X.$

$A = \begin{pmatrix} & & 1 \\ & 1 & \\ 1 & & \end{pmatrix}$, only the anti-diagonals entries are 1, others 0.

(c) $f(x) = |\langle a, x \rangle|^2 = x^T a a^T x \in \mathbb{R}.$

$x^T a a^T x = tr(x^T a a^T x) = tr(\underbrace{aa^T}_{n\times n} \underbrace{xx^T}_{n\times n})$ $\quad (tr(ABC) = tr(BCA))$

if A, B have same cols and rows, $tr(A^T B) = \langle A, B \rangle$

$tr(aa^T xx^T) = \langle (aa^T)^T, xx^T \rangle = \langle aa^T, xx^T \rangle$

$A = aa^T$

Question 4

(a) $S_1 \cap S_2 = \{ x \in V \mid \langle a_1, x \rangle = b_1, \langle a_2, x \rangle = b_2 \}$

$x \in S_1 \cap S_2, \Rightarrow \langle a_1, x \rangle = b_1, \langle a_2, x \rangle = b_2$

$z \in S_1 \cap S_2, \Rightarrow \langle a_1, z \rangle = b_1, \langle a_2, z \rangle = b_2$

$x \in V \Rightarrow -tx \in V, \forall t \in \mathbb{R}$. $\quad z \in V \Rightarrow (1+t) z \in V, \forall t$

$\Rightarrow -tx + (1+t) z \in V$.

$\langle a_1, (1+t) z - tx \rangle = (1+t) \langle a_1, z \rangle - t \langle a_1, x \rangle = (1+t) b_1 - t b_1 = b_1$.

$\langle a_2, (1+t) z - tx \rangle = (1+t) \langle a_2, z \rangle - t \langle a_2, z \rangle = (1+t) b_2 - t b_2 = b_2$

Thus, $(1+t) z - tx \in S_1 \cap S_2$.

(b) proof.

If z is the solution of $\min_{x \in S_1 \cap S_2} \|x-y\|$, then $z \in S_1 \cap S_2$.

we have: $\begin{cases} \langle a_1, z \rangle = b_1, \quad \langle a_2, z \rangle = b_2 \\ (1+t) z - tx \in S_1 \cap S_2, \quad \forall x \in S_1 \cap S_2, t \in \mathbb{R}. \end{cases}$

$\|z-y\|^2 \le \|(1+t) z - tx - y\|^2 = \|(z-y) + t(z-x)\|^2$

$= \|z-y\|^2 + t^2 \|z-x\|^2 + 2t \langle z-y, z-x \rangle$

i.e. $t \langle z-y, z-x \rangle \geq -\frac{1}{2} t^2 \|z-x\|^2$

if we choose $t > 0$: $\langle z-y, z-x \rangle \geq -\frac{1}{2} t \|z-x\|^2$

let $t \to 0^+$, $\langle z-y, z-x \rangle \geq 0$

if we choose $t < 0$: $\langle z-y, z-x \rangle \leq -\frac{1}{2} t \|z-x\|^2$

let $t \to 0^-$, $\langle z-y, z-x \rangle \leq 0$

Altogether, z satisfies $\langle z-y, z-x \rangle = 0$.

3

if $z \in S_1 \cap S_2$ and $\langle z-x, z-y \rangle = 0, \ \forall x \in S_1 \cap S_2$,

$$\min_{x \in S_1 \cap S_2} \|x-y\| \Leftrightarrow \min_{x \in S_1 \cap S_2} \|x-y\|^2 = \min_{x \in S_1 \cap S_2} \|x-z+z-y\|^2$$

$$= \min_{x \in S_1 \cap S_2} \|x-z\|^2 + \|z-y\|^2 \geqslant \|z-y\|^2$$

Thus z is the solution of $\min_{x \in S_1 \cap S_2} \|x-y\|$

(c)

$$\left.\begin{array}{l} z \in S_1 \cap S_2 \Rightarrow \langle a_1, z \rangle = b_1, \ \langle a_2, z \rangle = b_2 \\ x \in S_1 \cap S_2 \Rightarrow \langle a_1, x \rangle = b_1, \ \langle a_2, x \rangle = b_2 \end{array}\right\} \Rightarrow \begin{array}{l} \langle z-x, a_1 \rangle = 0 \quad ① \\ \langle z-x, a_2 \rangle = 0 \quad ② \end{array}$$

$$z = \arg\min_{x \in S_1 \cap S_2} \|x-y\| \Rightarrow \langle z-x, z-y \rangle = 0 \quad ③$$

According to ①②③, we can get: $z-y = \alpha a_1 + \beta a_2 \Rightarrow z = y + \alpha a_1 + \beta a_2$.

$$\langle a_1, z \rangle = b_1 \Rightarrow \langle y + \alpha a_1 + \beta a_2, a_1 \rangle = b_1$$

$$\langle a_1, y \rangle + \alpha \langle a_1, a_1 \rangle + \beta \langle a_1, a_2 \rangle = b_1$$

$$\langle a_1, a_1 \rangle \alpha + \langle a_1, a_2 \rangle \beta = b_1 - \langle a_1, y \rangle \quad ④$$

$$\langle a_2, z \rangle = b_2 \Rightarrow \langle y + \alpha a_1 + \beta a_2, a_2 \rangle = b_2.$$

$$\langle a_2, y \rangle + \alpha \langle a_1, a_2 \rangle + \beta \langle a_2, a_2 \rangle = b_2$$

$$\langle a_1, a_2 \rangle \alpha + \langle a_2, a_2 \rangle \beta = b_2 - \langle a_2, y \rangle \quad ⑤$$

According to 4 and 5, we have:

$$\begin{cases} \langle a_1, a_1\rangle \alpha + \langle a_1, a_2\rangle \beta = b_1 - \langle a_1, y\rangle \\ \langle a_1, a_2\rangle \alpha + \langle a_2, a_2\rangle \beta = b_2 - \langle a_2, y\rangle \end{cases}$$

$$\alpha = - \frac{\begin{vmatrix} \langle a_1, a_2\rangle & b_1 - \langle a_1, y\rangle \\ \langle a_2, a_2\rangle & b_2 - \langle a_2, y\rangle \end{vmatrix}}{\begin{vmatrix} \langle a_1, a_1\rangle & \langle a_1, a_2\rangle \\ \langle a_1, a_2\rangle & \langle a_2, a_2\rangle \end{vmatrix}} \qquad \beta = - \frac{\begin{vmatrix} \langle a_1, a_1\rangle & b_1 - \langle a_1, y\rangle \\ \langle a_1, a_2\rangle & b_2 - \langle a_2, y\rangle \end{vmatrix}}{\begin{vmatrix} \langle a_1, a_2\rangle & \langle a_1, a_1\rangle \\ \langle a_2, a_2\rangle & \langle a_1, a_2\rangle \end{vmatrix}}$$

Thus, $z = y + \alpha \cdot a_1 + \beta \cdot a_2$, where α, β is shown above.

(d)
Suppose we have z_1, z_2 two solutions,

z_1 is a solution $\Rightarrow \langle z_1 - y, z_1 - z_2\rangle = 0$

z_2 is a solution $\Rightarrow \langle z_2 - y, z_2 - z_1\rangle = 0 \Rightarrow \langle z_2 - y, z_1 - z_2\rangle = 0$

taking difference: $\langle z_1 - z_2, z_1 - z_2\rangle = 0 \Rightarrow z_1 = z_2$.

Question 5

(a) $y = x^T a + b = a_1 x_1 + a_2 x_2 + \cdots + a_8 x_8 + b$. if $a_3 > 0$, $x_3 > 0 \Rightarrow a_3 x_3 > 0$

can't get $y > 0$. FALSE.

(b) $a_2 = 0$, $y = a_1 x_1 + a_3 x_3 + \cdots + a_8 x_8 + b$. doesn't depend on x_2

TRUE.

(c) $a_6 = -0.8$, $y = M - 0.8 x_6$, where $M = a_1 x_1 + \cdots + a_5 x_5 + a_7 x_7 + a_8 x_8 + b$.

$y_1 = M - 0.8 x_6$

$y_2 = M - 0.8(x_6 + a)$, $a > 0$

$y_1 - y_2 = a > 0 \Rightarrow y_1 > y_2$

TRUE.

5

Question 6

(a) Let $a = a_s + \sum_{i=1}^{N} c_i x_i$, where $c = [c_1; \cdots, c_N]^T \in \mathbb{R}^N$, $\langle a_s, x_i \rangle = 0$ for $i = 1, 2, \cdots, N$.

proof of the decomposition:

① For $N=1$, $S = \{v \mid \langle v, x_1 \rangle = 0\}$ is a hyperplane, co-dim $= 1$.

For any a, $a = P_S a + (a - P_S a)$, where $P_S a$ is a projection of a onto S.

$$\langle a - P_S a, P_S a - v \rangle = 0 \xRightarrow{\text{let } v=0} \langle a - P_S a, P_S a \rangle = 0 \Rightarrow a - P_S a \perp P_S a.$$

Also, $a - P_S a = \frac{\langle a, x_1 \rangle - 0}{\|a\|^2} x_1$, define $c_1 \equiv \frac{\langle a, x_1 \rangle}{\|a\|^2}$, then

$a - P_S a = c_1 x_1$, $a = P_S a + c_1 x_1$. define $a_s \equiv P_S a$, $a = a_s + c_1 x_1$.

② For $N \geq 2$, $S = \{v \mid \langle v, x_i \rangle = 0, i = 1, 2, \cdots, N\}$ is a hyperplane, co-dim $= N$.

similarly, we can get $a - P_S a = \sum_{i=1}^{N} c_i x_i$.

(The calculation for $N=2$ is the same as Question 4(c).)

Thus $a = a_s + \sum_{i=1}^{N} c_i x_i$.

Thus $\sum_{i=1}^{N} (\langle a, x_i \rangle - y_i)^2 + \lambda \|a\|_2^2$

$$= \sum_{i=1}^{N} \left(\langle a_s + \sum_{i=1}^{N} c_i x_i, x_i \rangle - y_i \right)^2 + \lambda \|a_s + \sum_{i=1}^{N} c_i x_i\|_2^2.$$

$$= \sum_{i=1}^{N} \left(\sum_{j=1}^{N} c_j \langle x_i, x_j \rangle - y_i \right)^2 + \lambda \left(\|a_s\|_2^2 + \sum_{i=1}^{N} \sum_{j=1}^{N} c_i c_j \langle x_i, x_j \rangle \right)$$

$$= \|K^T c - y\|^2 + \lambda c^T K c + \lambda \|a_s\|_2^2.$$

$$K = \begin{bmatrix} \langle x_1, x_1 \rangle & \cdots & \langle x_1, x_N \rangle \\ \vdots & & \\ \langle x_N, x_1 \rangle & \cdots & \langle x_N, x_N \rangle \end{bmatrix} \in \mathbb{R}^{N \times N}.$$

Let $F_1(c) = \|K^T c - y\|^2 + \lambda c^T K c$, $F_2(a_s) = \lambda \|a_s\|^2$.

$$\min_{\substack{c \in \mathbb{R}^N \\ \langle a_s, x_i \rangle = 0 \\ i=1,\cdots,N}} F_1(c) + F_2(a_s) \iff \min_{c \in \mathbb{R}^N} F_1(c) \text{ and } \min_{\substack{\langle a_s, x_i \rangle = 0 \\ i=1,\cdots,N}} F_2(a_s).$$

$$\min_{\substack{\langle a_s, x_i\rangle = 0\\ i=1,\cdots,N}} F_2(a_s) \text{ can be solved as } a_s = 0.$$

so the solution must be in the form of $a = \sum_{i=1}^{N} c_i x_i$,

where $c = [c_1, \cdots, c_N] \in \mathbb{R}^N = \underset{c \in \mathbb{R}^N}{\operatorname{argmin}} \|K^T c - y\|^2 + \lambda c^T K c$.

(b)

$$\min_{a \in \mathbb{R}^n} \sum_{i=1}^{N} (\langle a, x_i\rangle - y_i)^2 + \lambda \|a\|_2^2 \iff \min_{c \in \mathbb{R}^N} \|K^T c - y\|^2 + \lambda c^T K c.$$

where $K = \begin{bmatrix} \langle x_1, x_1\rangle & \cdots & \langle x_1, x_N\rangle \\ \vdots & & \\ \langle x_N, x_1\rangle & \cdots & \langle x_N, x_N\rangle \end{bmatrix}$.