MSBD 5004 Due: 14th Oct.

Assignment #2 Solution

Instructor: Jianfeng CAI Name: ID:

Question 1

(a) proof:

$\|x\| = \|(x-y)+y\| \le \|x-y\| + \|y\| \Rightarrow \|x\| - \|y\| \le \|x-y\|$

$\|y\| = \|(x-y)+x\| \le \|x-y\| + \|x\| \Rightarrow \|y\| - \|x\| \le \|x-y\|$.

$\therefore \big| \|x\| - \|y\| \big| \le \|x-y\|$.

(b) $\lim\limits_{k\to\infty} x_k = x$,

proof: $\|x_k\| - \|x\| \le \| x_k - x\|$ $\lim\limits_{k\to\infty} \big|\|x_k\| - \|x\|\big| \le \lim\limits_{k\to\infty} \|x_k - x\| = 0$

$\because \big| \|x_k\| - \|x\| \big| \ge 0 \quad \therefore \lim\limits_{k\to\infty} \big| \|x_k\| - \|x\| \big| = 0, \quad \lim\limits_{k\to\infty} \|x_k\| = \|x\|$.

(c) $\begin{cases} \lim\limits_{k\to\infty} x^{(k)} = x \Rightarrow \exists N_1 \in \mathbb{N},\ n > N_1,\ \|x^{(k)} - x\| < \varepsilon \\ \lim\limits_{k\to\infty} x^{(k)} = y. \Rightarrow \exists N_2 \in \mathbb{N},\ n > N_2,\ \|x^{(k)} - y\| < \varepsilon. \end{cases}$

Let $N = N_1 + N_2$, $\varepsilon = \frac{1}{5}\|x-y\|$, $\forall n > N$:

$\|x-y\| = \|x - x^{(k)} + x^{(k)} - y\| \le \|x - x^{(k)}\| + \|x^{(k)} - y\|$

$= \|x^{(k)} - x\| + \|x^{(k)} - y\| < \varepsilon + \varepsilon = \frac{2}{5}\|x-y\|$.

Thus $\|x-y\| = \frac{2}{5}\|x-y\| \Rightarrow x = y$ (otherwise $1 = \frac{2}{5}$, contradictory).

Question 2

(a) $\langle 0, x \rangle = \langle 0+0, x \rangle = \langle 0, x \rangle + \langle 0, x \rangle = 2\langle 0, x \rangle \Rightarrow \langle 0, x \rangle = 0$

$\langle x, 0 \rangle = \langle x, 0+0 \rangle = \langle x, 0 \rangle + \langle x, 0 \rangle = 2 \cdot \langle x, 0 \rangle \Rightarrow \langle x, 0 \rangle = 0$

Thus $\langle 0, x \rangle = \langle x, 0 \rangle = 0$ for any $x \in V$.

(b) proof.

"$\Rightarrow$": $\langle \alpha x_1 + \beta x_2, y \rangle = \alpha \langle x_1, y \rangle + \beta \langle x_2, y \rangle$.

Let $\alpha, \beta = 1$, we have: $\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$.

Let $x_2 = x_1$, we have. $\langle (\alpha+\beta) x_1, y \rangle = (\alpha+\beta) \langle x_1, y \rangle$.

"$\Leftarrow$":

$\langle \alpha x_1, y \rangle = \alpha \langle x_1, y \rangle \quad \langle \beta x_2, y \rangle = \beta \langle x_2, y \rangle$

Let $x_1 = \alpha x_1$, $x_2 = \beta x_2$, then:

$\langle \alpha x_1 + \beta x_2, y \rangle = \langle \alpha x_1, y \rangle + \langle \beta x_2, y \rangle = \alpha \langle x_1, y \rangle + \beta \langle x_2, y \rangle$

Thus they are equivalent.

Question 3

$\langle A, B \rangle = \mathrm{trace}(A^T B) = \sum_{i,j} a_{ij} b_{ij}$

① $\forall X \in \mathbb{R}^{m \times n}$, $\langle X, X \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}^2 \geq 0$.

if $\langle X, X \rangle = 0$, $\sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}^2 = 0 \Rightarrow x_{ij} = 0$, for all $i = 1, \cdots, m$ and $j = 1, \cdots, n$

$\Rightarrow X = 0$.

② $\forall X, Y, Z \in \mathbb{R}^{m \times n}$ $\alpha, \beta \in \mathbb{R}$, $\langle \alpha X + \beta Y, Z \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} (\alpha x_{ij} + \beta y_{ij}) z_{ij}$

$= \alpha \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} z_{ij} + \beta \sum_{i=1}^{m} \sum_{j=1}^{n} y_{ij} z_{ij} = \alpha \langle X, Z \rangle + \beta \langle Y, Z \rangle$.

2

③ $\forall X, Y \in \mathbb{R}^{m\times n}$, $\langle X, Y\rangle = \mathrm{trace}(X^TY) = \mathrm{trace}(Y^TX) = \langle Y, X\rangle$

according to ① ② ③, $\langle A, B\rangle = \mathrm{trace}(A^TB)$ for $A, B \in \mathbb{R}^{m\times n}$ is indeed an inner product on $\mathbb{R}^{m\times n}$.

$\boxed{\text{Question 4}}$

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

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

$\forall$ norm $\|\cdot\|$, $\|x\|^2 \geq 0$. Thus $f(x,x) \geq 0$. $f(x,x) = 0 \Leftrightarrow \frac{1}{2}\|x\|^2 = 0 \Leftrightarrow x = 0$.

Thus: $f(x,x) \geq 0$ for any $x \in \mathbb{V}$, and $f(x,x) = 0$ iff $x = 0$.

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

(c) $f(x,y) := \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right)$. ①

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

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

$f(x,y) = \frac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$.

$f(x+y, z) = \frac{1}{4}\left(\|x+y+z\|^2 - \|x+y-z\|^2\right)$. (*)

$\|x+z+y\|^2 + \|x+z-y\|^2 = 2\|x+z\|^2 + 2\|y\|^2$.

$\|x+y+z\|^2 = 2\|x+z\|^2 + 2\|y\|^2 - \|x-y+z\|^2$ ③

$\|x+y+z\|^2 + \|y+z-x\|^2 = 2\|x\|^2 + 2\|y+z\|^2$

$\|x+y+z\|^2 = 2\|x\|^2 + 2\|y+z\|^2 - \|y+z-x\|^2$ ④

3

$\frac{1}{2}\times$(③+④): $\|x+y+z\|^2 = \|x\|^2 + \|y\|^2 + \|x+z\|^2 + \|y+z\|^2 - \frac{1}{2}\|y+z-x\|^2 - \frac{1}{2}\|x-y+z\|^2$. ⑤

using "$-z$" to substitute "z" in ⑤:

$$\|x+y-z\|^2 = \|x\|^2+\|y\|^2+\|x-z\|^2+\|y-z\|^2-\frac{1}{2}\|y-z-x\|^2-\frac{1}{2}\|x-y-z\|^2 \quad ⑥$$

take ⑤, ⑥ into (*):

$$f(x+y,z) = \frac{1}{4}\Big[\|x+z\|^2+\|y+z\|^2-\|x-z\|^2-\|y-z\|^2+\underbrace{\frac{1}{2}\|y-z-x\|^2}_{=\frac{1}{2}\|x+z-y\|^2}+\underbrace{\frac{1}{2}\|x-y-z\|^2}_{=\frac{1}{2}\|y+z-x\|^2}$$

$$-\frac{1}{2}\|x-y+z\|^2-\frac{1}{2}\|y+z-x\|^2\Big]$$

$$= \frac{1}{4}\left[\|x+z\|^2-\|x-z\|^2\right]+\frac{1}{4}\left[\|y+z\|^2-\|y-z\|^2\right]$$

$$= f(x,z)+f(y,z)$$

(d) $f(x,y) = \frac{1}{4}(\|x+y\|^2-\|x-y\|^2)$

$$f(-x,y) = \frac{1}{4}(\|-x+y\|^2-\|-x-y\|^2) = \frac{1}{4}(\|x-y\|^2-\|x+y\|^2) = -f(x,y).$$

(e) $f(x,x) = \frac{1}{2}\|x\|^2 \qquad f(y,y) = \frac{1}{2}\|y\|^2.$

$$f(x+\lambda y, x+\lambda y) = f(x,x)+\lambda f(x,y)+\lambda f(y,x)+\lambda^2 f(y,y)$$

$$= \lambda^2 f(y,y)+2\lambda f(x,y)+f(x,x) = G(\lambda)$$

① if $y\neq 0$, $f(y,y)>0$, $[2f(x,y)]^2-4f(y,y)\cdot f(x,x)\leq 0$

(because $G(\lambda)$ can only have one zero point at most)

$\Rightarrow [f(x,y)]^2 \leq f(x,x)f(y,y)$.

② if $y=0$, $[f(x,0)]^2 = 0 \leq f(x,x)f(0,0) = 0$.

Thus, for all $x,y\in V$, $[f(x,y)]^2 \leq f(x,x)f(y,y)$

Question 5

k_1, k_2 are kernel functions. $\Rightarrow \begin{cases} k_i(x,y) = k_i(y,x), \ i=1,2. \\ C^T K_1 C \geq 0, \ C^T K_2 C \geq 0, \ \forall C \in \mathbb{R}^m \end{cases}$

(a)

$$C^T K(x,y) C = C^T [K_1(x,y) + k_2(x,y)] C = C^T K_1 C + C^T K_2 C \geq 0.$$

$$K(x,y) = k_1(x,y) + k_2(x,y) = k_1(y,x) + k_2(y,x) = K(y,x).$$

(b) $k(x,y) = a \cdot k_1(x,y) = a \cdot k_1(y,x) = k(y,x)$

$$C^T k(x_i, x_j) C = C^T \cdot a k_1(x_i, x_j) C = a \cdot C^T k_1(x_i, x_j) C.$$

$\because a > 0, \ C^T K_1 C \geq 0 \quad \therefore C^T k(x_i, x_j) C \geq 0.$

Question 6

$$\langle \phi(x), \phi(y) \rangle = 2x_1y_1 + 2x_2y_2 + x_1^2y_1^2 + x_2^2y_2^2 + 2x_1x_2y_1y_2 + 1 = (x_1y_1 + x_2y_2 + 1)^2$$

$$k(x,y) = (x^Ty+1) = (\langle x, y \rangle + 1) = (x_1y_1 + x_2y_2 + 1)^2$$

$\underline{k(x,y) = \langle \phi(x), \phi(y) \rangle}$.

$\forall C \in \mathbb{R}^6$, $C^T K C = C^T (x_i^T x_j + 1)^2 C = C^T (x_i^T x_j + 1)^T (x_i^T x_j + 1) C$

$$= [C(x_i^T x_j + 1)]^T [C(x_i^T x_j + 1)] \geq 0.$$

$$K(x,y) = (x^Ty+1)^2 = (y^Tx+1)^2 = K(y,x).$$

$\therefore K$ is SPSD.

5

Question 7

$$k(x,y) = y^T A x = x^T A^T y = x^T A y = k(y, x).$$

A is SPSD matrix, $\exists B \in \mathbb{R}^d$, s.t $A = B^T B$

$$k(x,y) = y^T A x = y^T B^T B x = (By)^T (Bx) = \langle Bx, By \rangle. \quad \phi(x) = Bx.$$

$$C^T k(x_i, x_j) C = \sum_{i=1}^{m} \sum_{j=1}^{m} c_i c_j \sum_{k=1}^{d} \sum_{p=1}^{d} \sum_{q=1}^{d} b_p x_{ik} \, b_q x_{jk}$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{m} \sum_{k=1}^{d} \sum_{p=1}^{d} \sum_{q=1}^{d} c_i c_j b_p x_{ik} b_q x_{jk} = \sum_{k=1}^{d} \left(\sum_{i=1}^{m} \sum_{p=1}^{d} c_i b_p x_{ik} \right) \left(\sum_{j=1}^{m} \sum_{q=1}^{d} c_j b_q x_{jk} \right)$$

$$= \sum_{k=1}^{d} \left(\sum_{i=1}^{m} \sum_{p=1}^{d} c_i b_p x_{ik} \right)^2 \geq 0.$$

$\therefore$ $k(x,y)$ is a kernel function.