Yifan Luo u7351505

Exercise 1

1. False, Suppose $\vec{a}=\vec{c}$, $\vec{b}=-\vec{c}$, $\vec{c}\neq\vec{0}$, then have

$LHS=\langle\vec{a},\vec{c}\rangle=\langle\vec{c},\vec{c}\rangle>0$

$RHS=\langle\vec{a},\vec{b}\rangle+\langle\vec{b},\vec{c}\rangle=\langle\vec{c},-\vec{c}\rangle+\langle-\vec{c},\vec{c}\rangle=-2\langle\vec{c},\vec{c}\rangle<0$

$LHS > RHS$, which conflicts with $\langle\vec{a},\vec{c}\rangle\leq\langle\vec{a},\vec{b}\rangle+\langle\vec{b},\vec{c}\rangle$

2. False. Suppose $\vec{b}=\vec{0}\in V\subseteq\mathbb{R}^n$, $\langle\cdot,\cdot\rangle$ is dot product.

Let $\vec{a}=\vec{c}=\langle 1,1,\cdots,1\rangle$, then we have $\langle\vec{a},\vec{b}\rangle=\langle\vec{b},\vec{c}\rangle=0$.

But $\langle\vec{a},\vec{c}\rangle=n>0$.

3. True. $Span(S)=\{a_1\vec{v_1}+a_2\vec{v_2}+\dots+a_n\vec{v_n}\mid a_i\in\mathbb{R}, i=1,\cdots,n\}$

For $\forall\vec{y}\in Span(S)$, $\langle\vec{x},\vec{y}\rangle=\langle\vec{x},a_1\vec{v_1}+\cdots+a_n\vec{v_n}\rangle$ ①

According to the definition of $\langle\cdot,\cdot\rangle$, we have

$RHS=\langle\vec{x},a_1\vec{v_1}+\cdots+a_n\vec{v_n}\rangle=a_1\langle\vec{x},\vec{v_1}\rangle+\cdots+a_n\langle\vec{x},\vec{v_n}\rangle$

Because $\forall v\in S$, $\langle\vec{x},\vec{v}\rangle=0$, then $RHS=a_1\cdot 0+\cdots+a_n\cdot 0=0$

By reformatting ①, we have $\forall\vec{y}\in Span(S)$, $\langle\vec{x},\vec{y}\rangle=0$,

then $\vec{x}$ is orthogonal to any $\vec{y}\in Span(S)$.

4. True. Suppose $\vec{x}\in Span(S)$, $\vec{x}=a_1\vec{v_1}+a_2\vec{v_2}+\dots+a_n\vec{v_n}\in V$, then at least $\exists k\in\{1,\cdots,n\}$, $a_k\neq 0$.

Then $\forall i\in\{1,\cdots,n\}$, $\langle\vec{x},\vec{v_i}\rangle=\langle a_1\vec{v_1}+\cdots+a_n\vec{v_n},\vec{v_i}\rangle$

$=a_1\langle\vec{v_1},\vec{v_i}\rangle+\cdots+a_i\langle\vec{v_i},\vec{v_i}\rangle+\cdots+a_n\langle\vec{v_n},\vec{v_i}\rangle=\vec{0}$

Because $S=\{\vec{v_1},\cdots,\vec{v_n}\}$ is an orthonormal set, then

$RHS=a_1\cdot 0+\cdots+a_i\cdot 1+\cdots+a_n\cdot 0=\vec{0}$, then $a_1=a_2=\cdots=a_n=0$, which implies $\vec{x}=\vec{0}$

It conflicts with $\vec{x}\neq\vec{0}$

Thus, $\vec{x}\notin Span(S)$

5. True. Suppose $x \in \text{Span}(S)$, and $\forall i \in \{1, \cdots, n\}$ $\langle x, v_i \rangle = 0$, then $x \perp S$

From 1.3, we proved if $x \perp S$, then $x \perp \text{span}(S)$.

Because $x \in \text{span}(S)$, then $x \perp x$. $\langle x, x \rangle = \vec{0}$, then $\vec{x} = \vec{0}$.

which conflicts with $\vec{x} \neq \vec{0}$.

6. True. Consider two different ways to express $\vec{x} \in \text{Span}(S) = V$

$\vec{x} = a_1 \vec{x_1} + a_2 \vec{x_2} + \cdots + a_n \vec{x_n}$ and $\vec{x} = b_1 \vec{x_1} + b_2 \vec{x_2} + \cdots + b_n \vec{x_n}$

then $\vec{x} - \vec{x} = 0 = (a_1 - b_1)\vec{x_1} + \cdots + (a_n - b_n)\vec{x_n}$ holds iff. $a_i = b_i$, $i = 1, \cdots, n$.

Thus, only a unique set of coefficients to represent $\vec{x}$.

7. False Suppose $S = \{\langle 1, 1, 0 \rangle, \langle 1, 0, 0 \rangle, \langle 0, 1, 0 \rangle\}$, $\vec{x} = \langle 1, 1, 0 \rangle \in V$

then x can be represented as $1 \cdot \langle 1, 1, 0 \rangle + 0 \cdot \langle 1, 0, 0 \rangle + 0 \cdot \langle 0, 1, 0 \rangle$

and $0 \cdot \langle 1, 1, 0 \rangle + 1 \cdot \langle 1, 0, 0 \rangle + 1 \cdot \langle 0, 1, 0 \rangle$

Thus, the coefficients are not unique.

8. False Suppose $S = \{\langle 1, 0, 0, 0 \rangle, \langle 0, 1, 0, 0 \rangle, \langle 1, 1, 0, 0 \rangle\}$

The S is pairwise linearly independent, but

$c_1 \langle 1, 0, 0, 0 \rangle + c_2 \langle 0, 1, 0, 0 \rangle + c_3 \langle 1, 1, 0, 0 \rangle = \vec{0}$ has non-trivial solution

e.g. $c_1 = c_2 = 1$, $c_3 = -1$ or $c_1 = c_2 = -1$, $c_3 = 1$

Exercise 2. To prove $\|\cdot\|$ is a norm, we prove the following three qualities:

① Positive definite: $\|\vec{x}\| \geq 0$ and $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$

Because of the definition of $\langle \cdot, \cdot \rangle$, $\langle x, x \rangle$ is positive definite.

Then we have $\langle \vec{x}, \vec{x} \rangle > 0$ for $\forall \vec{x} \in V \setminus \{\vec{0}\}$, and $\langle \vec{x}, \vec{x} \rangle = 0$ when $\vec{x} = \vec{0}$

$\|x\| = \sqrt{\langle x, x \rangle} > \sqrt{0} = 0$, and $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$. thus positive definite is proved

② Homogeneous: $\|\alpha \vec{x}\| = |\alpha| \|\vec{x}\|$, $\forall \alpha \in \mathbb{R}$, $\forall \vec{x} \in V$

Because $\langle \cdot, \cdot \rangle$ is a bilinear mapping, we have

$$\|\alpha \vec{x}\| = \sqrt{\langle \alpha \vec{x}, \alpha \vec{x} \rangle} = \sqrt{\alpha^2 \langle \vec{x}, \vec{x} \rangle} = |\alpha| \sqrt{\langle \vec{x}, \vec{x} \rangle} = |\alpha| \|x\|$$

Thus, $\|\cdot\|$ is homogeneous.

③ Triangle inequality: $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$, $\forall \vec{x}, \vec{y} \in V$

$$\|\vec{x} + \vec{y}\| = \sqrt{\langle \vec{x}+\vec{y}, \vec{x}+\vec{y} \rangle}$$

$$= \sqrt{\langle \vec{x}, \vec{x}+\vec{y} \rangle + \langle \vec{y}, \vec{x}+\vec{y} \rangle} \quad \text{(Bilinear mapping)}$$

$$= \sqrt{\langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle} \quad \text{(as above)}$$

$$= \sqrt{\langle \vec{x}, \vec{x} \rangle + 2\langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{y} \rangle}$$

$$\|\vec{x} + \vec{y}\|^2 = \langle \vec{x}, \vec{x} \rangle + 2\langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{y} \rangle$$

$\because \langle \vec{x}, \vec{y} \rangle \leq \|\vec{x}\| \|\vec{y}\| = \sqrt{\langle \vec{x}, \vec{x} \rangle \langle \vec{y}, \vec{y} \rangle}$ (Cauchy-Schwarts inequality)

$$\therefore \|\vec{x} + \vec{y}\|^2 \leq \langle \vec{x}, \vec{x} \rangle + 2\sqrt{\langle \vec{x}, \vec{x} \rangle \langle \vec{y}, \vec{y} \rangle} + \langle \vec{y}, \vec{y} \rangle$$

$$= \left(\sqrt{\langle \vec{x}, \vec{x} \rangle} + \sqrt{\langle \vec{y}, \vec{y} \rangle}\right)^2$$

$$\therefore \|\vec{x} + \vec{y}\| \leq \sqrt{\langle \vec{x}, \vec{x} \rangle} + \sqrt{\langle \vec{y}, \vec{y} \rangle} = \|\vec{x}\| + \|\vec{y}\|$$

Thus, $\|\cdot\|$ obeys triangle inequality.

Overall, $\|\cdot\|$ is a norm.

Exercise 3.

1. $\nabla_{\vec\theta} \mathcal{L}(\vec\theta,\vec c) = \frac{d\mathcal{L}(\vec\theta,\vec c)}{d\vec\theta} = \frac{d}{d\vec\theta}\left(\|\vec y - X\vec\theta - \vec c\|_A^2 + \|\vec\theta\|_B^2 + \|\vec c\|_A^2\right)$

$= \frac{d}{d\vec\theta} \langle \vec y - X\vec\theta - \vec c, y - X\vec\theta - \vec c\rangle_A + \langle\vec\theta,\vec\theta\rangle_B + \langle\vec c,\vec c\rangle_A$

$= \frac{d}{d\vec\theta} (\vec y - X\vec\theta - \vec c)^T A (\vec y - X\vec\theta - \vec c) + \vec\theta^T B\vec\theta + \vec c^T A\vec c$

$= (\vec y - X\vec\theta - \vec c)^T (A + A^T)\cdot(-X) + \theta^T(B + B^T)$

$= -y^T AX - y^T A^T X + \vec\theta^T X^T AX + \vec\theta^T X^T A^T X + \vec c^T AX + \vec c^T A^T X$

2. Let $\nabla_{\vec\theta}\mathcal{L}(\vec\theta,\vec c) = \vec 0$, we have $-y^T AX - y^T A^T X + \vec\theta^T X^T AX + \vec\theta^T X^T A^T X + \vec c^T AX + \vec c^T A^T X = 0$

then, $\theta^T = (y^T AX + y^T A^T X - C^T AX - C^T A^T X)(X^T AX + X^T A^T X + B + B^T)^{-1}$

Prove $X^T AX + X^T A^T X + B + B^T$ is a positive define matrix:

Because A is a positive define matrix, then according to $\langle\vec x,\vec y\rangle = \vec x^T A\vec y$ iff. A is a positive define matrix, we have $X^T AX = \langle X, X\rangle_A$, $\langle X, X\rangle_A$ is positive define according to the definition of $\langle\cdot,\cdot\rangle$, then $X^T A^T X = (X^T AX)^T$ is also a positive define matrix.

Because B is a positive define matrix, then $B^T = (B)^T$ is also a positive define matrix.

Thus, $X^T AX + X^T A^T X + B + B^T$ is a positive define matrix.

$\theta = (\theta^T)^T = (X^T AX + X^T A^T X + B + B^T)^{-1}(X^T A^T y + X^T Ay - X^T A^T C - X^T AC)$

3. $\nabla_{\vec c}\mathcal{L}(\vec\theta,\vec c) = \frac{d}{d\vec c}(\vec y - X\vec\theta - \vec c)^T A(\vec y - X\vec\theta - \vec c) + \vec\theta^T B\vec\theta + \vec c^T A\vec c$

$= (\vec y - X\vec\theta - \vec c)^T (A + A^T)(-1) + \vec c^T (A + A^T)$

$= -\vec y^T A - \vec y^T A^T + \theta^T X^T A + \theta^T X^T A^T + C^T A + C^T A^T + C^T A + C^T A^T$

4. Let $\nabla_{\vec c}\mathcal{L}(\vec\theta,\vec c) = 0$, we have $-\vec y^T A - \vec y^T A^T + \theta^T X^T A + \theta^T X^T A^T + C^T A + C^T A^T + C^T A + C^T A^T = 0$

then $C^T = (\vec y^T A + \vec y^T A^T - \theta^T X^T A - \theta^T X^T A^T)(2A + 2A^T)^{-1}$

Prove $2A + 2A^T$ is a positive define matrix:

Because A is a positive define matrix, then $A = A^T$, thus $2A + 2A^T = 4A$ is also a positive define matrix.

Then. $C = (C^T)^T = (2A + 2A^T)^{-1}(A^T\vec y + A\vec y - A^T X\theta - AX\theta)$

5. From 3.2, we have:

$$\theta = (\theta^T)^T = (X^T A X + X^T A^T X + B + B^T)^{-1} (X^T A^T y + X^T A y - X^T A^T C - X^T A C)$$

Let $A = I \in \mathbb{R}^{N\times N}$, $\vec{C} = \vec{0} \in \mathbb{R}^N$, $B = \lambda I \in \mathbb{R}^{D\times D}$, then

$$A = A^T = I \quad B^T = (B)^T = \lambda I^T = \lambda I$$

then

$$\begin{aligned} RHS &= (2X^T X + 2\lambda I)^{-1} (2X^T y - \vec{0}) \\ &= (2X^T X + 2\lambda I)^{-1} \cdot 2X^T y \\ &= \tfrac{1}{2} (X^T X + \lambda I)^{-1} \cdot 2 X^T y \\ &= (X^T X + \lambda I)^{-1} \cdot X^T y \end{aligned}$$

Thus, the solution from 3.2 agrees with the analytic solution.