Solution to Homework 2

September 12, 2007

1 Problem 1

Let $x = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$ be an arbitrary real vector. We have

$$\|Dx\|_p = (\ (|d_1| \cdot |x_1|)^p + \cdots + (|d_n| \cdot |x_n|)^p\ )^{\frac{1}{p}}\ .$$

It follows that

$$\begin{aligned} \|Dx\|_p &\leq \left((\max_i |d_i| \cdot |x_1|)^p + \cdots + (\max_i |d_i| \cdot |x_n|)^p \right)^{\frac{1}{p}} \\ &= (\ \max_i |d_i|\ ) \cdot (\ \|x\|_p\ ). \end{aligned}$$

Furthermore, we have

$$\begin{aligned} \|D\|_p &= \sup_x \frac{\|Dx\|_p}{\|x\|_p} \\ &\leq \sup_x \frac{(\ \max_i |d_i|\ ) \cdot (\ \|x\|_p\ )}{\|x\|_p} \\ &= \max_i |d_i|. \end{aligned}$$

On the other hand, we choose a particular vector $y = (y_1, \cdots, y_n)^T \in \mathbb{R}^n$, where $y_i = 1$ if $i = \arg\max_i |d_i|$, and $y_i = 0$ otherwise. We thus have

$$\|D\|_p \geq \frac{\|Dy\|_p}{\|y\|} = \max_i |d_i|.$$

It shows that $\|D\|_p = \max_i |d_i|$, where $1 \leq p \leq \infty$.

2 Problem 2

2.1 $A + B$ is positive semi-definite

Let $x \in \mathbb{R}^n$ be an arbitrary vector. Since both A and B are positive semi-definite (PSD), it follows that $x^T(A + B)x = x^T Ax + x^T Bx \geq 0$, which shows that $A + B$ is PSD.

2.2 AB is NOT necessarily positive semi-definite

We provide a counter example. Let two matrices A and B be constructed as follows :

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}.$$

It can be verified that both A and B are PSD. However, for a particularly chosen vector $x = (1, 0)^T$, we have

$$x^T ABx = (1 \;\; 2) \begin{pmatrix} 1 \\ -2 \end{pmatrix} = -3 < 0.$$

It shows that AB is not necessarily PSD.

2.3 B^T is positive semi-definite

Let $x \in \mathbb{R}^n$ be an arbitrary vector. Given B being PSD, we have $x^T B^T x = (x^T Bx)^T \geq 0$. It shows that B^T is PSD.

3 Problem 3

3.1 trace(AB) =trace(BA)

Denote matrices A and B as follows:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n1} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{n1} \end{pmatrix}.$$

It follows that

$$\text{trace}(AB) = \sum_{i=1}^{n} a_{1i} b_{i1} + \cdots + \sum_{i=1}^{n} a_{ni} b_{in} = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ji} b_{ij}.$$

Similarly, we have

$$\text{trace}(BA) = \sum_{i=1}^{n} b_{1i} a_{i1} + \cdots + \sum_{i=1}^{n} b_{ni} a_{in} = \sum_{j=1}^{n} \sum_{i=1}^{n} b_{ji} a_{ij}.$$

It shows that trace(AB) = trace(BA).

3.2 $\|A\|_F^2$ = trace(AA^T)

Let $A = (a_1, a_2, \cdots, a_n) \in \mathbb{R}^{n \times n}$, where $a_i \in \mathbb{R}^n$. It follows that

$$\begin{aligned} \text{trace}(AA^T) &= \text{trace}(A^T A) = a_1^T a_1 + \cdots + a_n^T a_n \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} (a_{ij})^2 = \|A\|_F^2, \end{aligned}$$

which completes the proof.

3.3 Show $\|QA\|_F = \|A\|_F$

Given Q being an orthogonal matrix, it follows that $Q^TQ = Q^TQ = I$. Applying the results from Sections 3.1 and 3.2, we have

$$\begin{aligned}\|QA\|_F^2 &= \text{trace}(QA(QA)^T)\\ &= \text{trace}((QA)^TQA)\\ &= \text{trace}(A^TQ^TQA)\\ &= \text{trace}(A^TA)\\ &= \|A\|_F^2,\end{aligned}$$

which completes the proof.

4 Problem 4

Let $A = (a_1, a_2, \cdots, a_n) \in \mathbb{R}^{m\times n}$, where $a_i \in \mathbb{R}^m$. Let $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^d$ be an arbitrary real vector. It follows that

$$\begin{aligned}\|Ax\|_1 &= (|x_1| \cdot \|a_1\|_1) + (|x_2| \cdot \|a_2\|_1) + \cdots + (|x_n| \cdot \|a_n\|_1)\\ &\leq (\max_j \|a_j\|_1) \cdot (\ |x_1| + |x_2| + \cdots + |x_n|\ )\\ &= (\max_j \|a_j\|_1) \cdot \|x\|_1.\end{aligned}$$

Moreover, we have

$$\|A\|_1 = \sup_x \frac{\|Ax\|_1}{\|x\|_1} \leq \sup_x \frac{(\max_j \|a_j\|_1) \cdot \|x\|_1}{\|x\|_1} = \max_j \|a_j\|_1.$$

On the other hand, we choose a particular real vector $y = (y_1, \cdots, y_n)^T$, where $y_j = 1$ if $j = \arg\max_j \|a_j\|_1$, and 0 otherwise. We further have

$$\|A\|_1 = \sup_x \frac{\|Ax\|_1}{\|x\|_1} \geq \frac{\|Ay\|_1}{\|y\|_1} = \max_j \|a_j\|_1.$$

Since $\|a_j\|_1 = \sum_{i=1}^m |a_{ij}|$, it shows $\|A\|_1 = \max_j \sum_{i=1}^m |a_{ij}|$.

5 Problem 5

5.1 $I_n - A$ is nonsingular

We consider the product of two matrices as follows

$$\begin{aligned}(I_n - A) \cdot (I_n - A^T) &= I_n - A - A^T + AA^T\\ &= I_n + AA^T,\end{aligned}$$

where the second equality follows from $A^T = -A$. It can be verified that $I_n + AA^T$ is positive definite and thus nonsingular, which also implies $I_n - A$ is nonsingular.

5.2 $(I_n - A)^{-1}(I_n + A)$ is orthogonal

In order to show $(I_n - A)^{-1}(I_n + A)$ is orthogonal, we compute :

$$\begin{aligned}
& \left((I_n - A)^{-1}(I_n + A)\right)^T (I_n - A)^{-1}(I_n + A) \\
= \; & (I_n + A^T)(I_n - A^T)^{-1}(I_n - A)^{-1}(I_n + A) \\
= \; & (I_n + A^T)\left((I_n - A)(I_n - A^T)\right)^{-1}(I_n + A) \\
= \; & (I_n + A^T)\left(I_n + AA^T\right)^{-1}(I_n + A). \qquad (1)
\end{aligned}$$

Consider the last two terms $\left(I_n + AA^T\right)^{-1}$ and $(I_n + A)$ in Eq. (1) as follows:

$$\begin{aligned}
& (I_n + AA^T)^{-1}(I_n + A) \\
= \; & (I_n + AA^T)^{-1}\left((I_n + A)(I_n + A^T A)\right)(I_n + A^T A)^{-1} \\
= \; & (I_n + AA^T)^{-1}\left((I_n + AA^T)(I_n + A)\right)(I_n + A^T A)^{-1} \\
= \; & (I_n + A)(I_n + A^T A)^{-1}, \qquad (2)
\end{aligned}$$

where the first equality follows that $I_n + A^T A$ is nonsingular, and the second equality follows from $A^T = -A$ and thus $AA^T = A^T A$. Applying the result of Eq. (2) into Eq. (1), we have

$$\begin{aligned}
& (I_n + A^T)\left(I_n + AA^T\right)^{-1}(I_n + A) \\
= \; & (I_n + A^T)(I_n + A)(I_n + A^T A)^{-1} \\
= \; & (I_n + A + A^T + A^T A)(I_n + A^T A)^{-1} \\
= \; & (I_n + A^T A)(I_n + A^T A)^{-1} \\
= \; & I_n.
\end{aligned}$$

Note that the third equality follows from $A^T = -A$. It completes the proof.