Matrix Theory Assignment 1 Solutions

General Instructions

The following document contains the solutions to the theory-based questions for Assignment 1. Please note that the solutions provided may not be the only possible way to solve the questions. They indicate only one of the many (possibly) valid solutions. The solutions provided are relatively crisp and do not include all the steps that you must have. Your solution should be logical and contain all supporting arguments. Feel free to contact us via email in case of any discrepancy you find in the solutions provided.

Question 1

Given $\mathbf{x} \in \mathbb{R}^{\mathbb{N}}$ Let $\mathbf{x} = [x_1 \; x_2 \; \cdots \; x_N]^T$

a) Average, $\mu = \frac{1}{N}\sum_{n=1}^{N} x_n$ Since, the daily "price" of a stock can never be negative:

$|x_n| = x_n$

$$\begin{aligned} \mu &= \frac{1}{N}\sum_{n=1}^{N} |x_n| \\ \implies \mu &= \frac{1}{N}\|\mathbf{x}\|_1 \end{aligned}$$

b) We calculate the variance as follows:

$$\begin{aligned} \sigma^2 &= \frac{1}{N}\sum_{n=1}^{N}(x_n - \mu)^2 \\ &= \frac{1}{N}\sum_{n=1}^{N}(x_n^2 - 2x_n\mu + \mu^2) \\ &= \frac{1}{N}\left(\sum_{n=1}^{N} x_n^2 - \sum_{n=1}^{N} 2x_n\mu + \sum_{n=1}^{N}\mu^2\right) \end{aligned}$$

As we know, $|x_n^2| = x_n^2$

$$
\begin{aligned}
\implies \sigma^2 &= \frac{1}{N}\left(\sum_{n=1}^{N}|x_n|^2 - \mu\sum_{n=1}^{N}2x_n + N\mu^2\right) \\
&= \frac{1}{N}\left(\|\mathbf{x}\|_2^2 - 2\mu\|\mathbf{x}\|_1 + N.\frac{1}{N^2}\|\mathbf{x}\|_1^2\right) \\
&= \frac{1}{N}\left(\|\mathbf{x}\|_2^2 - \frac{2}{N}\|\mathbf{x}\|_1^2 + \frac{1}{N}\|\mathbf{x}\|_1^2\right) \\
&= \frac{1}{N}\|\mathbf{x}\|_2^2 - \frac{1}{N^2}\|\mathbf{x}\|_1^2
\end{aligned}
$$

c) We now find a recursive relation for μ as

$$
\begin{aligned}
\mu_{N+1} &= \frac{1}{N+1}\sum_{n=1}^{N+1}x_n \\
&= \frac{1}{N+1}\left(\left(\sum_{n=1}^{N}x_n\right) + x_{n+1}\right) \\
&= \frac{1}{N+1}(N\mu_N + x_{n+1})
\end{aligned}
$$

d) We now find a recursive relation for σ^2 as

$$
\begin{aligned}
\sigma_{N+1}^2 &= \frac{1}{N+1}\sum_{n=1}^{N+1}(x_n - \mu_{N+1})^2 \\
&= \frac{1}{N+1}\sum_{n=1}^{N+1}(x_n^2 - 2\mu_{N+1}x_n + \mu_{N+1}^2) \\
&= \frac{1}{N+1}(N\sigma_N^2 + N\mu_N^2 + x_{n+1}^2) - 2\mu_{N+1}^2 + \mu_{N+1}^2 \\
&= \frac{N}{N+1}\sigma_N^2 + \frac{N}{N+1}\mu_N^2 + \frac{1}{N+1}x_{n+1}^2 - \mu_{N+1}^2 \\
&= \frac{N}{N+1}\sigma_N^2 + \frac{N}{N+1}\mu_N^2 + \frac{1}{N+1}x_{n+1}^2 - \left(\frac{1}{N+1}\right)^2(N\mu_N + x_{n+1})^2 \\
&= \frac{1}{N+1}\left(N\sigma_N^2 + N\mu_N^2 + x_{n+1}^2 - \frac{1}{N+1}(N^2\mu_N^2 + 2N\mu_N x_{n+1} + x_{n+1}^2)\right) \\
&= \frac{N}{N+1}\left(\sigma_N^2 + \frac{1}{N+1}(\mu_N - x_{n+1})^2\right)
\end{aligned}
$$

Question 2

Let two independent vectors $\mathbf{a}$ and $\mathbf{b}$. The scalar projection of vector $\mathbf{a}$ onto vector $\mathbf{b}$ is given by:

$$
\text{Scalar Projection} = \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{b}\|_2}
$$

Now, we know that the Cauchy-Schwarz inequality states that:

$$\langle \mathbf{a}, \mathbf{b} \rangle \leq \|\mathbf{a}\|_2 \, \|\mathbf{b}\|_2$$
$$\therefore \frac{\langle \mathbf{a}, \mathbf{b} \rangle}{\|\mathbf{b}\|_2} \leq \|\mathbf{a}\|_2$$

Hence, proved.

Question 3

Given $\mathbf{a} \in \mathbb{R}^+$ and $\|\mathbf{a}\|_1 = 1$. Let $\mathbf{a} = [a_1 \; a_2 \; \cdots \; a_N]^T$ such that

$$\sum_{i=1}^{n} a_i = 1$$

Let us define a new vector $\mathbf{x}$ = $\mathbf{1}$ where $\mathbf{1}$ is a vector with all its elements equal to 1.

Using Cauchy-Schwarz Inequality for $\mathbf{a}$ and $\mathbf{x}$,

$$\begin{aligned} |\langle \mathbf{a}, \mathbf{x} \rangle| &\leq \|\mathbf{a}\|_2 \|\mathbf{x}\|_2 \\ \implies |\langle \mathbf{a}, \mathbf{x} \rangle| &\leq \|\mathbf{a}\|_2 \sqrt{n} \\ \implies \sum_{i=1}^{n} a_i x_i &\leq \|\mathbf{a}\|_2 \sqrt{n} \\ \implies \sum_{i=1}^{n} a_i &\leq \|\mathbf{a}\|_2 \sqrt{n} \\ \implies \frac{1}{\sqrt{n}} &\leq \|\mathbf{a}\|_2 \end{aligned}$$

Therefore, the minimum value of $\|\mathbf{a}\|_2$ is $\frac{1}{\sqrt{n}}$

Question 4

a. Considering the function $f : \mathbb{R} \to \mathbb{R}$ as follows

$$\begin{aligned} f(\alpha_1) &= \|\mathbf{a} + \alpha_1 \mathbf{b}\|_2^2 \\ &= \langle \mathbf{a} + \alpha_1 \mathbf{b}, \mathbf{a} + \alpha_1 \mathbf{b} \rangle \\ &= \|\mathbf{a}\|_2^2 + 2\alpha_1 \langle \mathbf{a}, \mathbf{b} \rangle + \alpha_1^2 \|\mathbf{b}\|_2^2 \end{aligned}$$

The given expression can be minimized by minimizing $f(\alpha_1)$

$$\frac{\partial f(\alpha_1)}{\partial \alpha_1} = 2 \langle \mathbf{a}, \mathbf{b} \rangle + 2\alpha_1 \|\mathbf{b}\|_2^2$$

$$\frac{\partial^2 f(\alpha_1)}{\partial^2 \alpha_1} = 2\|\mathbf{b}\|_2^2 \geq 0$$

The given expression will have minima when $\frac{\partial f(\alpha_1)}{\partial \alpha_1} = 0.$

Therefore, for the required minima,

$$\frac{\partial f(\alpha_1)}{\partial \alpha_1} = 2\left\langle \mathbf{a}, \mathbf{b} \right\rangle + 2\alpha_1 |\mathbf{b}|_2^2 = 0$$
$$\implies \alpha_1 = -\frac{\left\langle \mathbf{a}, \mathbf{b} \right\rangle}{||\mathbf{b}||_2^2}$$

b. Similarly considering the function $g : \mathbb{R} \to \mathbb{R}$ as follows

$$\begin{aligned} g(\alpha_2) &= ||\mathbf{a} - \alpha_2 \mathbf{b}||_2^2 \\ &= \left\langle \mathbf{a} - \alpha_2 \mathbf{b}, \mathbf{a} - \alpha_2 \mathbf{b} \right\rangle \\ &= ||\mathbf{a}||_2^2 - 2\alpha_2 \left\langle \mathbf{a}, \mathbf{b} \right\rangle + \alpha_2^2 ||\mathbf{b}||_2^2 \end{aligned}$$

The given expression can be minimized by minimizing $g(\alpha_1)$

$$\frac{\partial g(\alpha_2)}{\partial \alpha_2} = -2\left\langle \mathbf{a}, \mathbf{b} \right\rangle + 2\alpha_2 ||\mathbf{b}||_2^2$$

$$\frac{\partial^2 g(\alpha_2)}{\partial^2 \alpha_2} = 2||\mathbf{b}||_2^2 \geq 0$$

The given expression will have maxima when $\dfrac{\partial g(\alpha_2)}{\partial \alpha_2} = 0.$

Therefore for the required minima,

$$\frac{\partial g(\alpha_2)}{\partial \alpha_2} = -2\left\langle \mathbf{a}, \mathbf{b} \right\rangle + 2\alpha_2 ||\mathbf{b}||_2^2 = 0$$
$$\implies \alpha_2 = \frac{\left\langle \mathbf{a}, \mathbf{b} \right\rangle}{||\mathbf{b}||_2^2}$$

Question 5

Given $\mathbf{x} \in \mathbb{R}^N$. Let $\mathbf{x} = [x_1 \; x_2 \; x_3 \; \cdots x_N]^T$. The L_P norm of a vector $\mathbf{x}$:

$$||\mathbf{x}||_P = \left(\sum_{i=1}^{N} |x_i|^P \right)^{\frac{1}{P}}$$

Consider $||\mathbf{x}||_1^2$ i.e., $||\mathbf{x}||_1^2 = ||\mathbf{x}||_1 ||\mathbf{x}||_1$

$$||\mathbf{x}||_1^2 = \left(\sum_{i=1}^{N} |x_i| \right) \left(\sum_{j=1}^{N} |x_j| \right)$$
$$\implies ||\mathbf{x}||_1^2 = \sum_{i=1}^{N} |x_i|^2 + \sum_{\substack{i,j \\ i \neq j}} |x_i||x_j|$$

Analyse the first summation of the RHS with respect to the L_2 norm i.e., $||\mathbf{x}||_2^2 = \sum\limits_{i=1}^{N} |x_i|^2$

The first summation is simply the L_2 norm of $\mathbf{x}$

$$\implies \|\mathbf{x}\|_1^2 = \|\mathbf{x}\|_2^2 + \sum_{\substack{i,j \\ i \neq j}} |x_i||x_j|$$

Now analyse the second summation. It is a summation of positive definites(absolute values). Hence, the summation is positive definite i.e.,

$$\sum_{\substack{i,j \\ i \neq j}} |x_i||x_j| \geq 0$$

$$\implies \|\mathbf{x}\|_1^2 \geq \|\mathbf{x}\|_2^2$$

Note that one of the criteria to be classified as a norm is the positive definite property i.e.,

$$\left(\sum_{i=1}^{N} |x_i|^P\right)^{\frac{1}{P}} \geq 0$$

$$\implies \|\mathbf{x}\|_1 \geq \|\mathbf{x}\|_2$$