MAS374 Optimization Theory

Homework #8

20150597 Jeonghwan Lee*

Department of Mathematical Sciences, KAIST

December 1, 2021

Problem 1 (*Exercise 11.1* in [1]**).**

(1) Let $\mathcal{L}(\mathbf{p};\mathbf{q}) := \{(1-\theta)\mathbf{p} + \theta\mathbf{q} : \theta \in [0,1]\} \subseteq \mathbb{R}^n$ for $\mathbf{p} \neq \mathbf{q}$ in $\mathbb{R}^n$. Then,

$$\begin{aligned}\mathcal{D}(\mathbf{p};\mathbf{q}) &:= (\text{the minimum distance from the origin } \mathbf{0} \in \mathbb{R}^n \text{ to the line segment } \mathcal{L}(\mathbf{p};\mathbf{q})) \\ &= \min\{\|\lambda\mathbf{p} + (1-\lambda)\mathbf{q}\|_2 : \lambda \in [0,1]\}.\end{aligned}$$

Thus, it holds for every $R \geq \mathbb{R}_+$ that

$$\begin{aligned}\mathcal{D}(\mathbf{p};\mathbf{q}) \geq R \quad &\Leftrightarrow \quad \|\lambda\mathbf{p} + (1-\lambda)\mathbf{q}\|_2 \geq R,\ \forall\lambda \in [0,1] \\ &\Leftrightarrow \quad \|\lambda\mathbf{p} + (1-\lambda)\mathbf{q}\|_2^2 \geq R^2,\ \forall\lambda \in \mathbb{R} \text{ such that } \lambda(1-\lambda) \leq 0,\end{aligned} \tag{1.1}$$

which precisely yields the desired result.

(2) From (1.1), it immediately follows that

$$\mathcal{D}(\mathbf{p};\mathbf{q}) \geq R \quad \Leftrightarrow \quad \{\lambda \in \mathbb{R} : \lambda(\lambda-1) \leq 0\} \subseteq \left\{\lambda \in \mathbb{R} : R^2 - \|\lambda\mathbf{p} + (1-\lambda)\mathbf{q}\|_2^2 \leq 0\right\}. \tag{1.2}$$

At this point, we define two quadratic functions $f_0(\cdot) : \mathbb{R} \to \mathbb{R}$ and $f_1(\cdot) : \mathbb{R} \to \mathbb{R}$ by

$$\begin{aligned}f_0(\lambda) &:= R^2 - \|\lambda\mathbf{p} + (1-\lambda)\mathbf{q}\|_2^2 = F_0\lambda^2 + 2g_0\lambda + h_0; \\ f_1(\lambda) &:= \lambda(\lambda-1) = F_1\lambda^2 + 2g_1\lambda + h_1,\end{aligned}$$

where $F_0 := -\|\mathbf{p}-\mathbf{q}\|_2^2$, $g_0 := \mathbf{q}^\top(\mathbf{q}-\mathbf{p})$, $h_0 := R^2 - \mathbf{q}^\top\mathbf{q}$, and $F_1 := 1$, $g_1 := -\frac{1}{2}$, $h_1 := 0$. Since the inequality $f_1(\lambda) \leq 0$ is strictly feasible, *i.e.*, $f_1\left(\tilde{\lambda}\right) < 0$ for some $\tilde{\lambda} \in \mathbb{R}$, the *lossless* S*-procedure* is valid for two quadratic functions $f_0(\cdot) : \mathbb{R} \to \mathbb{R}$ and $f_1(\cdot) : \mathbb{R} \to \mathbb{R}$: for every $R \in \mathbb{R}_+$,

$$\begin{aligned}\mathcal{D}(\mathbf{p};\mathbf{q}) \geq R \quad &\Leftrightarrow \quad \{\lambda \in \mathbb{R} : \lambda(\lambda-1) \leq 0\} \subseteq \left\{\lambda \in \mathbb{R} : R^2 - \|\lambda\mathbf{p} + (1-\lambda)\mathbf{q}\|_2^2 \leq 0\right\} \\ &\Leftrightarrow \quad \begin{bmatrix} -\|\mathbf{p}-\mathbf{q}\|_2^2 & \mathbf{q}^\top(\mathbf{q}-\mathbf{p}) \\ \mathbf{q}^\top(\mathbf{q}-\mathbf{p}) & R^2 - \mathbf{q}^\top\mathbf{q} \end{bmatrix} \preceq \tau \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix} \text{ for some } \tau \in \mathbb{R}_+ \\ &\Leftrightarrow \quad \begin{bmatrix} \tau + \|\mathbf{p}-\mathbf{q}\|_2^2 & \mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2} \\ \mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2} & \mathbf{q}^\top\mathbf{q} - R^2 \end{bmatrix} \in \mathcal{S}_+^2 \text{ for some } \tau \in \mathbb{R}_+,\end{aligned} \tag{1.3}$$

*E-mail: sa8seung@kaist.ac.kr

where $\mathcal{S}_+^d$ denotes the convex cone consists of $d \times d$ real, symmetric, positive semi-definite matrices, and this completes the proof of the part (2).

(3) From (1.3), it's clear that

$$[0, \mathcal{D}(\mathbf{p};\mathbf{q})] = \underbrace{\left\{ R \in \mathbb{R}_+ : \begin{bmatrix} \tau + \|\mathbf{p}-\mathbf{q}\|_2^2 & \mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2} \\ \mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2} & \mathbf{q}^\top\mathbf{q} - R^2 \end{bmatrix} \in \mathcal{S}_+^2 \text{ for some } \tau \in \mathbb{R}_+ \right\}}_{=:\,\Omega \subseteq \mathbb{R}_+}. \tag{1.4}$$

If $R \in \Omega$, then

$$\mathbf{q}^\top\mathbf{q} - R^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}^\top \begin{bmatrix} \tau + \|\mathbf{p}-\mathbf{q}\|_2^2 & \mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2} \\ \mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2} & \mathbf{q}^\top\mathbf{q} - R^2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \geq 0,$$

thereby $R \leq \left(\mathbf{q}^\top\mathbf{q}\right)^{\frac{1}{2}}$. Thus, we have $\mathcal{D}(\mathbf{p};\mathbf{q}) \leq \left(\mathbf{q}^\top\mathbf{q}\right)^{\frac{1}{2}}$.

Case #1. $\mathbf{p}^\top\mathbf{q} \geq \mathbf{q}^\top\mathbf{q}$: It's clear that $\left(\mathbf{q}^\top\mathbf{q}\right)^{\frac{1}{2}} \in \Omega$, since

$$\begin{bmatrix} \tau + \|\mathbf{p}-\mathbf{q}\|_2^2 & \mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2} \\ \mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2} & 0 \end{bmatrix} \in \mathcal{S}_+^2 \quad \text{for } \tau = 2\mathbf{q}^\top(\mathbf{p}-\mathbf{q}) \in \mathbb{R}_+.$$

This implies $\mathcal{D}(\mathbf{p};\mathbf{q}) \geq \left(\mathbf{q}^\top\mathbf{q}\right)^{\frac{1}{2}}$, thereby we have

$$\mathcal{D}(\mathbf{p};\mathbf{q}) = \left(\mathbf{q}^\top\mathbf{q}\right)^{\frac{1}{2}}, \tag{1.5}$$

provided that $\mathbf{p}^\top\mathbf{q} \geq \mathbf{q}^\top\mathbf{q}$.

Case #2. $\mathbf{p}^\top\mathbf{q} \geq \mathbf{p}^\top\mathbf{p}$: Since $\underbrace{\left(\mathbf{p}^\top\mathbf{p} - \mathbf{p}^\top\mathbf{q}\right)}_{\leq\, 0} + \left(\mathbf{q}^\top\mathbf{q} - \mathbf{p}^\top\mathbf{q}\right) = \|\mathbf{p}-\mathbf{q}\|_2^2 > 0$, we find that

$$\mathbf{p}^\top\mathbf{p} \leq \mathbf{p}^\top\mathbf{q} < \mathbf{q}^\top\mathbf{q}. \tag{1.6}$$

If $\left(\mathbf{q}^\top\mathbf{q}\right)^{\frac{1}{2}} \in \Omega$, then

$$\mathsf{det}\left(\begin{bmatrix} \tau + \|\mathbf{p}-\mathbf{q}\|_2^2 & \mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2} \\ \mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2} & 0 \end{bmatrix}\right) = -\left\{\mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2}\right\}^2 \geq 0 \tag{1.7}$$

for some $\tau \in \mathbb{R}_+$. However, the inequality (1.6) cannot hold since the inequality (1.5) implies $2\mathbf{q}^\top(\mathbf{p}-\mathbf{q}) < 0$, contradiction! Thus, $\mathcal{D}(\mathbf{p};\mathbf{q}) < \left(\mathbf{q}^\top\mathbf{q}\right)^{\frac{1}{2}}$ and $\mathbf{q}^\top\mathbf{q} - R^2 > 0$ for every $R \in \Omega = [0, \mathcal{D}(\mathbf{p};\mathbf{q})]$. Due to the

Schur's complement rule, one can obtain by doing some straightforward algebra that

$$
\begin{aligned}
\Omega &= \left\{ R \in \mathbb{R}_+ : \tau + \|\mathbf{p}-\mathbf{q}\|_2^2 - \frac{\left\{\mathbf{q}^\top(\mathbf{p}-\mathbf{q}) - \frac{\tau}{2}\right\}^2}{\mathbf{q}^\top\mathbf{q} - R^2} \geq 0 \text{ for some } \tau \in \mathbb{R}_+ \right\} \\
&= \left\{ R \in \mathbb{R}_+ : \left\{\tau - 2\left(\mathbf{q}^\top\mathbf{p} - R^2\right)\right\} - 4\left(R^2 - \mathbf{p}^\top\mathbf{p}\right)\left(R^2 - \mathbf{q}^\top\mathbf{q}\right) \leq 0 \text{ for some } \tau \in \mathbb{R}_+ \right\} \\
&= \left\{ R \in \mathbb{R}_+ : \min\left\{ \left\{t - 2\left(\mathbf{q}^\top\mathbf{p} - R^2\right)\right\} - 4\left(R^2 - \mathbf{p}^\top\mathbf{p}\right)\left(R^2 - \mathbf{q}^\top\mathbf{q}\right) : t \in \mathbb{R}_+ \right\} \leq 0 \right\} \\
&\overset{\text{(a)}}{=} \left\{ R \in \left[0, \left(\mathbf{q}^\top\mathbf{p}\right)^{\frac{1}{2}}\right] : \left(R^2 - \mathbf{p}^\top\mathbf{p}\right)\left(R^2 - \mathbf{q}^\top\mathbf{q}\right) \geq 0 \right\} \\
&\quad \cup \left\{ R \in \left(\left(\mathbf{q}^\top\mathbf{p}\right)^{\frac{1}{2}}, +\infty\right) : \left(R^2 - \mathbf{q}^\top\mathbf{p}\right)^2 - \left(R^2 - \mathbf{p}^\top\mathbf{p}\right)\left(R^2 - \mathbf{q}^\top\mathbf{q}\right) \leq 0 \right\} \\
&\overset{\text{(b)}}{=} \left[0, \left(\mathbf{p}^\top\mathbf{p}\right)^{\frac{1}{2}}\right] \cup \underbrace{\left\{ R \in \left(\left(\mathbf{q}^\top\mathbf{p}\right)^{\frac{1}{2}}, +\infty\right) : R^2 \leq \mathbf{q}^\top\mathbf{q} - \frac{\left\{\mathbf{q}^\top(\mathbf{p}-\mathbf{q})\right\}^2}{\|\mathbf{p}-\mathbf{q}\|_2^2} \right\}}_{=\varnothing} \\
&\overset{\text{(c)}}{=} \left[0, \left(\mathbf{p}^\top\mathbf{p}\right)^{\frac{1}{2}}\right],
\end{aligned}
\tag{1.8}
$$

where the steps (a)–(c) holds due to the following reasons:

(a) It's straightforward that

$$
\begin{aligned}
&\min\left\{ \left\{t - 2\left(\mathbf{q}^\top\mathbf{p} - R^2\right)\right\} - 4\left(R^2 - \mathbf{p}^\top\mathbf{p}\right)\left(R^2 - \mathbf{q}^\top\mathbf{q}\right) : t \in \mathbb{R}_+ \right\} \\
&= \begin{cases} -4\left(R^2 - \mathbf{p}^\top\mathbf{p}\right)\left(R^2 - \mathbf{q}^\top\mathbf{q}\right) & \text{if } \mathbf{q}^\top\mathbf{p} - R^2 \geq 0; \\ 4\left[\left(R^2 - \mathbf{q}^\top\mathbf{p}\right)^2 - \left(R^2 - \mathbf{p}^\top\mathbf{p}\right)\left(R^2 - \mathbf{q}^\top\mathbf{q}\right)\right] & \text{otherwise.} \end{cases}
\end{aligned}
\tag{1.9}
$$

(b) We have from the relation (1.6) that

$$
\begin{aligned}
\left\{ R \in \left[0, \left(\mathbf{q}^\top\mathbf{p}\right)^{\frac{1}{2}}\right] : \left(R^2 - \mathbf{p}^\top\mathbf{p}\right)\left(R^2 - \mathbf{q}^\top\mathbf{q}\right) \geq 0 \right\} &= \left\{ R \in \left[0, \left(\mathbf{q}^\top\mathbf{p}\right)^{\frac{1}{2}}\right] : R^2 \leq \mathbf{p}^\top\mathbf{p} \text{ or } R^2 \geq \mathbf{q}^\top\mathbf{q} \right\} \\
&= \left[0, \left(\mathbf{p}^\top\mathbf{p}\right)^{\frac{1}{2}}\right].
\end{aligned}
$$

(c) It suffices to prove that $\mathbf{q}^\top\mathbf{p} \geq \mathbf{q}^\top\mathbf{q} - \frac{\left\{\mathbf{q}^\top(\mathbf{p}-\mathbf{q})\right\}^2}{\|\mathbf{p}-\mathbf{q}\|_2^2}$:

$$
\mathbf{q}^\top\mathbf{p} - \left[\mathbf{q}^\top\mathbf{q} - \frac{\left\{\mathbf{q}^\top(\mathbf{p}-\mathbf{q})\right\}^2}{\|\mathbf{p}-\mathbf{q}\|_2^2}\right] = \frac{\left(\mathbf{q}^\top\mathbf{p} - \mathbf{q}^\top\mathbf{q}\right)\left(\mathbf{p}^\top\mathbf{p} - \mathbf{p}^\top\mathbf{q}\right)}{\|\mathbf{p}-\mathbf{q}\|_2^2} \overset{\text{(d)}}{\geq} 0,
$$

where the step (d) follows from (1.6).

From (1.8), we may conclude that

$$
\mathcal{D}(\mathbf{p};\mathbf{q}) = \left(\mathbf{p}^\top\mathbf{p}\right)^{\frac{1}{2}}, \tag{1.10}
$$

when $\mathbf{p}^\top\mathbf{q} \geq \mathbf{p}^\top\mathbf{p}$.

Case #3. $\mathbf{p}^\top\mathbf{q} < \min\left\{\mathbf{p}^\top\mathbf{p}, \mathbf{q}^\top\mathbf{q}\right\}$: Since $\mathbf{q}^\top(\mathbf{p}-\mathbf{q}) = \mathbf{p}^\top\mathbf{q} - \mathbf{q}^\top\mathbf{q} < 0$, one can see that $\mathcal{D}(\mathbf{p};\mathbf{q}) < \left(\mathbf{q}^\top\mathbf{q}\right)^{\frac{1}{2}}$ and $\mathbf{q}^\top\mathbf{q} - R^2 > 0$ for every $R \in \Omega = [0, \mathcal{D}(\mathbf{p};\mathbf{q})]$ due to the same reason as in **Case #2**. Applying the Schur's complement rule, we arrive at

$$
\begin{aligned}
\Omega &= \left\{ R \in \mathbb{R}_+ : \tau + \|\mathbf{p} - \mathbf{q}\|_2^2 - \frac{\left\{\mathbf{q}^\top (\mathbf{p} - \mathbf{q}) - \frac{\tau}{2}\right\}^2}{\mathbf{q}^\top \mathbf{q} - R^2} \ge 0 \text{ for some } \tau \in \mathbb{R}_+ \right\} \\
&= \left\{ R \in \mathbb{R}_+ : \left\{ \tau - 2\left(\mathbf{q}^\top \mathbf{p} - R^2\right)\right\} - 4\left(R^2 - \mathbf{p}^\top \mathbf{p}\right)\left(R^2 - \mathbf{q}^\top \mathbf{q}\right) \le 0 \text{ for some } \tau \in \mathbb{R}_+ \right\} \\
&= \left\{ R \in \mathbb{R}_+ : \min\left\{ \left\{ t - 2\left(\mathbf{q}^\top \mathbf{p} - R^2\right)\right\} - 4\left(R^2 - \mathbf{p}^\top \mathbf{p}\right)\left(R^2 - \mathbf{q}^\top \mathbf{q}\right) : t \in \mathbb{R}_+ \right\} \le 0 \right\} \\
&\overset{\text{(e)}}{=} \left\{ R \in \left[0, \left(\mathbf{q}^\top \mathbf{p}\right)^{\frac{1}{2}}\right] : \left(R^2 - \mathbf{p}^\top \mathbf{p}\right)\left(R^2 - \mathbf{q}^\top \mathbf{q}\right) \ge 0 \right\} \\
&\quad \cup \left\{ R \in \left(\left(\mathbf{q}^\top \mathbf{p}\right)^{\frac{1}{2}}, +\infty\right) : \left(R^2 - \mathbf{q}^\top \mathbf{p}\right)^2 - \left(R^2 - \mathbf{p}^\top \mathbf{p}\right)\left(R^2 - \mathbf{q}^\top \mathbf{q}\right) \le 0 \right\} \\
&\overset{\text{(f)}}{=} \left[0, \left(\mathbf{q}^\top \mathbf{p}\right)^{\frac{1}{2}}\right] \cup \left(\left(\mathbf{q}^\top \mathbf{p}\right)^{\frac{1}{2}}, \left[\mathbf{q}^\top \mathbf{q} - \frac{\left\{\mathbf{q}^\top (\mathbf{p} - \mathbf{q})\right\}^2}{\|\mathbf{p} - \mathbf{q}\|_2^2}\right]^{\frac{1}{2}} \right] \\
&= \left[0, \left[\mathbf{q}^\top \mathbf{q} - \frac{\left\{\mathbf{q}^\top (\mathbf{p} - \mathbf{q})\right\}^2}{\|\mathbf{p} - \mathbf{q}\|_2^2}\right]^{\frac{1}{2}} \right],
\end{aligned}
\tag{1.11}
$$

where the step (e) makes use of the fact (1.9), and the step (f) follows from $\mathbf{p}^\top \mathbf{q} < \min\left\{\mathbf{p}^\top \mathbf{p}, \mathbf{q}^\top \mathbf{q}\right\}$, which implies $\mathbf{q}^\top \mathbf{p} < \mathbf{q}^\top \mathbf{q} - \frac{\left\{\mathbf{q}^\top (\mathbf{p} - \mathbf{q})\right\}^2}{\|\mathbf{p} - \mathbf{q}\|_2^2}$. Hence, we obtain

$$
\mathcal{D}\left(\mathbf{p}; \mathbf{q}\right)^2 = \mathbf{q}^\top \mathbf{q} - \frac{\left\{\mathbf{q}^\top (\mathbf{p} - \mathbf{q})\right\}^2}{\|\mathbf{p} - \mathbf{q}\|_2^2}, \tag{1.12}
$$

provided that $\mathbf{p}^\top \mathbf{q} < \min\left\{\mathbf{p}^\top \mathbf{p}, \mathbf{q}^\top \mathbf{q}\right\}$.

By taking three pieces (1.5), (1.10), and (1.12) collectively, we finally get

$$
\begin{aligned}
\mathcal{D}\left(\mathbf{p}; \mathbf{q}\right)^2 &= \begin{cases} \mathbf{q}^\top \mathbf{q} & \text{if } \mathbf{p}^\top \mathbf{q} \ge \mathbf{q}^\top \mathbf{q}; \\ \mathbf{p}^\top \mathbf{p} & \text{if } \mathbf{p}^\top \mathbf{q} \ge \mathbf{p}^\top \mathbf{p}; \\ \mathbf{q}^\top \mathbf{q} - \frac{\left\{\mathbf{q}^\top (\mathbf{p} - \mathbf{q})\right\}^2}{\|\mathbf{p} - \mathbf{q}\|_2^2} & \text{if } \mathbf{p}^\top \mathbf{q} < \min\left\{\mathbf{p}^\top \mathbf{p}, \mathbf{q}^\top \mathbf{q}\right\}, \end{cases} \\
&= \begin{cases} \mathbf{q}^\top \mathbf{q} & \text{if } \mathbf{p}^\top \mathbf{q} > \mathbf{q}^\top \mathbf{q}; \\ \mathbf{p}^\top \mathbf{p} & \text{if } \mathbf{p}^\top \mathbf{q} > \mathbf{p}^\top \mathbf{p}; \\ \mathbf{q}^\top \mathbf{q} - \frac{\left\{\mathbf{q}^\top (\mathbf{p} - \mathbf{q})\right\}^2}{\|\mathbf{p} - \mathbf{q}\|_2^2} & \text{if } \mathbf{p}^\top \mathbf{q} \le \min\left\{\mathbf{p}^\top \mathbf{p}, \mathbf{q}^\top \mathbf{q}\right\}, \end{cases}
\end{aligned}
$$

and this result coincides with the result in *Exercise 9.3* in [1] as desired.

Problem 2 (*Exercise 11.7* in [1]).

(1) We first prove the "only if" direction. Assume that $\mathbf{X} \in \mathcal{S}^n$ satisfies $f_k(\mathbf{X}) = \sum_{i=1}^k \lambda_i(\mathbf{X}) \le t$. Here, $\mathcal{S}^n$ denotes the $\mathbb{R}$-vector space of all $n \times n$ real symmetric matrices. Let $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^\top$ be the spectral decomposition of $\mathbf{X} \in \mathcal{S}^n$, where $\mathbf{U} := \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} \in \mathcal{O}(n)$ and $\mathbf{\Sigma} := \mathsf{diag}\left(\lambda_1(\mathbf{X}), \lambda_2(\mathbf{X}), \cdots, \lambda_n(\mathbf{X})\right) \in \mathbb{R}^{n \times n}$ together with

$$
\lambda_1(\mathbf{X}) \ge \lambda_2(\mathbf{X}) \ge \cdots \ge \lambda_n(\mathbf{X}).
$$

Now, let $s := \lambda_k(\mathbf{X}) \in \mathbb{R}$ and

$$\mathbf{Z} = \mathbf{U}\mathsf{diag}\left(d_1, d_2, \cdots, d_n\right)\mathbf{U}^\top = \sum_{i=1}^{n} d_i \mathbf{u}_i \mathbf{u}_i^\top \in \mathcal{S}^n,$$

where

$$d_i := \begin{cases} \lambda_i(\mathbf{X}) - \lambda_k(\mathbf{X}) & \text{if } 1 \leq i \leq k; \\ 0 & \text{otherwise.} \end{cases}$$

Since d_i are all non-negative, we have

$$\mathbf{v}^\top \mathbf{Z} \mathbf{v} = \sum_{i=1}^{n} d_i \left(\mathbf{u}_i^\top \mathbf{v}\right)^2 \geq 0, \ \forall \mathbf{v} \in \mathbb{R}^n,$$

which implies that $\mathbf{Z} \succeq \mathbf{O}_{n\times n}$. Next, we claim that $t - ks - \mathsf{Trace}(\mathbf{Z}) \geq 0$. This can be justified as follows:

$$\begin{aligned} t - ks - \mathsf{Trace}(\mathbf{Z}) &= t - k\lambda_k(\mathbf{X}) - \sum_{i=1}^{k} \left\{\lambda_i(\mathbf{X}) - \lambda_k(\mathbf{X})\right\} \\ &= t - \sum_{i=1}^{k} \lambda_i(\mathbf{X}) \\ &= t - f_k(\mathbf{X}) \geq 0. \end{aligned}$$

Finally, it remains to show that $\mathbf{Z} - \mathbf{X} + s\mathbf{I}_n \succeq \mathbf{O}_{n\times n}$. We first observe that

$$\begin{aligned} \mathbf{Z} - \mathbf{X} + s\mathbf{I}_n &= \mathbf{U}\mathsf{diag}\left(d_1 - \lambda_1(\mathbf{X}) + s, d_2 - \lambda_2(\mathbf{X}) + s, \cdots, d_n - \lambda_n(\mathbf{X}) + s\right)\mathbf{U}^\top \\ &= \sum_{i=1}^{n} \left\{d_i - \lambda_i(\mathbf{X}) + s\right\} \mathbf{u}_i \mathbf{u}_i^\top. \end{aligned} \tag{2.1}$$

Since

$$d_i - \lambda_i(\mathbf{X}) + s = \begin{cases} 0 & \text{if } 1 \leq i \leq k; \\ \lambda_k(\mathbf{X}) - \lambda_i(\mathbf{X}) & \text{otherwise,} \end{cases}$$

$d_i - \lambda_i + s \geq 0$ for every $i \in [n]$. So for every $\mathbf{v} \in \mathbb{R}^n$, we have from (2.1) that

$$\mathbf{v}^\top \left(\mathbf{Z} - \mathbf{X} + s\mathbf{I}_n\right) \mathbf{v} = \sum_{i=1}^{n} \left(d_i - \lambda_i + s\right) \left(\mathbf{u}_i^\top \mathbf{v}\right)^2 \geq 0, \ \forall \mathbf{v} \in \mathbb{R}^n,$$

thereby $\mathbf{Z} - \mathbf{X} + s\mathbf{I}_n \succeq \mathbf{O}_{n\times n}$ as desired.

It's time to prove the "if" part. We know that

$$\mathbf{Z} = \left(\mathbf{Z} - \mathbf{X} + s\mathbf{I}_n\right) + \left(\mathbf{X} - s\mathbf{I}_n\right)$$

and $\mathbf{Z} - \mathbf{X} + s\mathbf{I}_n \succeq \mathbf{O}_{n\times n}$. By applying inequality (4.6) in [1] (which is an immediate consequence of *Corollary 4.2* in [1]), we obtain

$$\lambda_i(\mathbf{Z}) \geq \lambda_i\left(\mathbf{X} - s\mathbf{I}_n\right) = \lambda_i(\mathbf{X}) - s, \ \forall i \in [n]. \tag{2.2}$$

Thus, we arrive at

$$
\begin{aligned}
f_k(\mathbf{X}) &= \sum_{i=1}^{k} \lambda_i(\mathbf{X}) \\
&\overset{\text{(a)}}{\leq} \sum_{i=1}^{k} \lambda_i(\mathbf{Z}) + ks \\
&\overset{\text{(b)}}{\leq} \sum_{i=1}^{n} \lambda_i(\mathbf{Z}) + ks \\
&= \mathsf{Trace}(\mathbf{Z}) + ks \\
&\leq t,
\end{aligned}
$$

where the step (a) follows from the inequality (2.2), and the step (b) holds since $\mathbf{Z} \geq \mathbf{O}_{n\times n}$, which implies that $\lambda_{\min}(\mathbf{Z}) \geq 0$. This completes the proof of part (1) of Problem 2.

(2) We claim that $f_k(\cdot) : \mathcal{S}^n \to \mathbb{R}$ is a convex function. To begin with, it's evident that $\mathsf{dom}\,(f_k) = \mathcal{S}^n$ is convex in $\mathcal{S}^n$. We claim that the epigraph of the function $f_k(\cdot) : \mathcal{S}^n \to \mathbb{R}$,

$$
\mathsf{epi}\,(f_k) := \{(\mathbf{X}, t) \in \mathcal{S}^n \times \mathbb{R} : f_k(\mathbf{X}) \leq t\},
$$

is a convex subset of the $\mathbb{R}$-vector space $\mathcal{S}^n \times \mathbb{R}$. To this end, we choose any $(\mathbf{X}_0, t_0), (\mathbf{X}_1, t_1) \in \mathsf{epi}\,(f_k)$. Due to the "only if" part of (1), there exist $\mathbf{Z}_0, \mathbf{Z}_1 \in \mathcal{S}^n$ and $s_0, s_1 \in \mathbb{R}$ such that

$$
t_i - ks_i - \mathsf{Trace}\,(\mathbf{Z}_i) \geq 0; \quad \mathbf{Z}_i \succeq \mathbf{O}_{n\times n}; \quad \mathbf{Z}_i - \mathbf{X}_i + s_i \mathbf{I}_n \succeq \mathbf{O}_{n\times n}, \ \forall i \in \{0, 1\}. \tag{2.3}
$$

So it follows from the fact (2.3) that for every $\theta \in [0, 1]$,

$$
\begin{aligned}
\{(1-\theta)\,t_0 + \theta t_1\} - k\,\{(1-\theta)\,s_0 + \theta s_1\} - \mathsf{Trace}\,\{(1-\theta)\,\mathbf{Z}_0 + \theta \mathbf{Z}_1\} &\geq 0; \\
(1-\theta)\,\mathbf{Z}_0 + \theta \mathbf{Z}_1 &\succeq \mathbf{O}_{n\times n}; \\
\{(1-\theta)\,\mathbf{Z}_0 + \theta \mathbf{Z}_1\} - \{(1-\theta)\,\mathbf{X}_0 + \theta \mathbf{X}_1\} + \{(1-\theta)\,s_0 + \theta s_1\}\,\mathbf{I}_n &\succeq \mathbf{O}_{n\times n}.
\end{aligned}
$$

Hence, we have $f_k\,\{(1-\theta)\,\mathbf{X}_0 + \theta \mathbf{X}_1\} \leq (1-\theta)\,t_0 + \theta t_1$, which implies

$$
((1-\theta)\,\mathbf{X}_0 + \theta \mathbf{X}_1, (1-\theta)\,t_0 + \theta t_1) = (1-\theta)\,(\mathbf{X}_0, t_0) + \theta\,(\mathbf{X}_1, t_1) \in \mathsf{epi}\,(f_k)
$$

for every $\theta \in [0, 1]$, by the "if" part of (1). Hence, the epigraph of the function $f_k(\cdot) : \mathcal{S}^n \to \mathbb{R}$ is a convex subset of the $\mathbb{R}$-vector space $\mathcal{S}^n \times \mathbb{R}$. As a result, this leads us to the convexity of the function $f_k(\cdot) : \mathcal{S}^n \to \mathbb{R}$.

However, the function $f_k(\cdot) : \mathcal{S}^n \to \mathbb{R}$ is NOT a norm on $\mathcal{S}^n$ since it can attain negative real-values. For instance, it's straightforward to see that $-\mathbf{I}_n \in \mathcal{S}^n$ and

$$
f_k\,(-\mathbf{I}_n) = \sum_{i=1}^{k} \lambda_i\,(-\mathbf{I}_n) = -k < 0.
$$

(3) For every $k \in [\min\{m, n\}]$, we define a function $\varphi_k(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}_+$ by

$$
\varphi_k(\mathbf{X}) := \sum_{i=1}^{k} \sigma_i(\mathbf{X}),
$$

where $\sigma_i(\mathbf{X})$ denotes the i-th largest singular value of $\mathbf{X} \in \mathbb{R}^{m\times n}$ for $i \in [\min\{m,n\}]$. Also, we define a map $\Phi(\cdot) : \mathbb{R}^{m\times n} \to \mathcal{S}^{m+n}$ by

$$\Phi(\mathbf{X}) := \begin{bmatrix} \mathbf{O}_{m\times m} & \mathbf{X} \\ \mathbf{X}^\top & \mathbf{O}_{n\times n} \end{bmatrix}.$$

It's clear that $\Phi(\cdot) : \mathbb{R}^{m\times n} \to \mathcal{S}^{m+n}$ is an $\mathbb{R}$-linear map. At this point, we would like to prove the following crucial result:

Lemma 1. *For every $\mathbf{X} \in \mathbb{R}^{m\times n}$, the eigenvalues of $\Phi(\mathbf{X}) \in \mathcal{S}^{m+n}$ are given by*

$$\sigma_1(\mathbf{X}) \geq \sigma_2(\mathbf{X}) \geq \cdots \geq \sigma_r(\mathbf{X}) > \underbrace{0 = 0 = \cdots = 0}_{(m+n-2r)\text{ zeroes}} > -\sigma_r(\mathbf{X}) \geq \cdots \geq -\sigma_2(\mathbf{X}) \geq -\sigma_1(\mathbf{X}), \tag{2.4}$$

where $r := \mathsf{rank}(\mathbf{X}) \leq \min\{m,n\}$.

Proof of Lemma 1. To begin with, let $\mathbf{X} = \mathbf{U}\tilde{\mathbf{\Sigma}}\mathbf{V}^\top$ be a singular value decomposition of $\mathbf{X} \in \mathbb{R}^{m\times n}$, where $\mathbf{U} \in \mathcal{O}(m)$, $\mathbf{V} \in \mathcal{O}(n)$, and $\tilde{\mathbf{\Sigma}} := \begin{bmatrix} \mathbf{\Sigma} & \mathbf{O}_{r\times(n-r)} \\ \mathbf{O}_{(m-r)\times r} & \mathbf{O}_{(m-r)\times(n-r)} \end{bmatrix} \in \mathbb{R}^{m\times n}$, $\mathbf{\Sigma} := \mathsf{diag}\left(\sigma_1(\mathbf{X}), \sigma_2(\mathbf{X}), \cdots, \sigma_r(\mathbf{X})\right) \in \mathbb{R}^{r\times r}$. Then it holds that

$$\begin{aligned} \begin{bmatrix} \mathbf{U} & \mathbf{O}_{m\times n} \\ \mathbf{O}_{n\times m} & \mathbf{V} \end{bmatrix}^\top \Phi(\mathbf{X}) \begin{bmatrix} \mathbf{U} & \mathbf{O}_{m\times n} \\ \mathbf{O}_{n\times m} & \mathbf{V} \end{bmatrix} &= \begin{bmatrix} \mathbf{U}^\top & \mathbf{O}_{m\times n} \\ \mathbf{O}_{n\times m} & \mathbf{V}^\top \end{bmatrix} \begin{bmatrix} \mathbf{O}_{m\times m} & \mathbf{U}\tilde{\mathbf{\Sigma}}\mathbf{V}^\top \\ \mathbf{V}\tilde{\mathbf{\Sigma}}^\top\mathbf{U}^\top & \mathbf{O}_{n\times n} \end{bmatrix} \begin{bmatrix} \mathbf{U} & \mathbf{O}_{m\times n} \\ \mathbf{O}_{n\times m} & \mathbf{V} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{O}_{m\times m} & \tilde{\mathbf{\Sigma}} \\ \tilde{\mathbf{\Sigma}}^\top & \mathbf{O}_{n\times n} \end{bmatrix} \\ &= \sum_{i=1}^{r} \sigma_i(\mathbf{X}) \left\{ \mathbf{e}_i^{(m+n)} \left(\mathbf{e}_{m+i}^{(m+n)}\right)^\top + \mathbf{e}_{m+i}^{(m+n)} \left(\mathbf{e}_i^{(m+n)}\right)^\top \right\}, \end{aligned} \tag{2.5}$$

where $\mathbf{e}_j^{(m+n)} \in \mathbb{R}^{m+n}$ denotes the j-th unit vector in $\mathbb{R}^{m+n}$ for each $j \in [m+n]$. Set $\mathbf{Q} := \begin{bmatrix} \mathbf{U} & \mathbf{O}_{m\times n} \\ \mathbf{O}_{n\times m} & \mathbf{V} \end{bmatrix} \in \mathcal{O}(m+n)$. Then from the equation (2.5), one can obtain the following observations:

$$\left(\mathbf{Q}^\top \Phi(\mathbf{X}) \mathbf{Q}\right) \mathbf{e}_i^{(m+n)} = \begin{cases} \sigma_i(\mathbf{X}) \cdot \mathbf{e}_{m+i}^{(m+n)} & \text{if } 1 \leq i \leq r; \\ \mathbf{0} \in \mathbb{R}^{m+n} & \text{if } r+1 \leq i \leq m, \end{cases} \tag{2.6}$$

and

$$\left(\mathbf{Q}^\top \Phi(\mathbf{X}) \mathbf{Q}\right) \mathbf{e}_{m+j}^{(m+n)} = \begin{cases} \sigma_i(\mathbf{X}) \cdot \mathbf{e}_j^{(m+n)} & \text{if } 1 \leq j \leq r; \\ \mathbf{0} \in \mathbb{R}^{m+n} & \text{if } r+1 \leq j \leq n. \end{cases} \tag{2.7}$$

From the observations (2.6) and (2.7), one can see that

$$\left(\mathbf{Q}^\top \Phi(\mathbf{X}) \mathbf{Q}\right) \cdot \frac{1}{\sqrt{2}} \left(\mathbf{e}_i^{(m+n)} + \mathbf{e}_{m+i}^{(m+n)}\right) = \sigma_i(\mathbf{X}) \cdot \frac{1}{\sqrt{2}} \left(\mathbf{e}_i^{(m+n)} + \mathbf{e}_{m+i}^{(m+n)}\right), \ \forall i \in [r], \tag{2.8}$$

and

$$\left(\mathbf{Q}^\top \Phi(\mathbf{X}) \mathbf{Q}\right) \cdot \frac{1}{\sqrt{2}} \left(\mathbf{e}_j^{(m+n)} - \mathbf{e}_{m+j}^{(m+n)}\right) = -\sigma_j(\mathbf{X}) \cdot \frac{1}{\sqrt{2}} \left(\mathbf{e}_j^{(m+n)} - \mathbf{e}_{m+j}^{(m+n)}\right), \ \forall j \in [r]. \tag{2.9}$$

By taking four pieces (2.6)–(2.9) collectively, we may conclude that

$$\begin{aligned}\mathcal{B} := & \left\{ \frac{1}{\sqrt{2}} \left(\mathbf{e}_i^{(m+n)} + \mathbf{e}_{m+i}^{(m+n)} \right) : i \in [r] \right\} \cup \left\{ \mathbf{e}_i^{(m+n)} : i \in [r+1; m] \right\} \\ & \cup \left\{ \mathbf{e}_{m+j}^{(m+n)} : j \in [r+1; n] \right\} \cup \left\{ \frac{1}{\sqrt{2}} \left(\mathbf{e}_j^{(m+n)} - \mathbf{e}_{m+j}^{(m+n)} \right) : j \in [r] \right\}\end{aligned}$$

forms an orthonormal basis for $\mathbb{R}^{m+n}$ consisting of eigenvectors of the matrix $\mathbf{Q}^\top \Phi(\mathbf{X}) \mathbf{Q} \in \mathcal{S}^{m+n}$, where

$$[a; b] := \{a, a+1, \cdots, b-1, b\}, \ \forall a \leq b \text{ in } \mathbb{Z},$$

with the corresponding eigenvalues given in (2.4). So, the eigenvalues of $\Phi(\mathbf{X}) \in \mathcal{S}^{m+n}$ are given as (2.4) as we claimed.

□

By Lemma 1, we have for every $k \in [\min\{m, n\}]$,

$$\varphi_k(\mathbf{X}) = \sum_{i=1}^{k} \sigma_i(\mathbf{X}) = \sum_{i=1}^{k} \lambda_i (\Phi(\mathbf{X})) = f_k (\Phi(\mathbf{X})), \ \forall \mathbf{X} \in \mathbb{R}^{m \times n}. \tag{2.10}$$

From the fact (2.10), one can easily derive the following conclusions, which are also generalized versions of the results in part (1) and (2):

(i) For every $k \in [\min\{m, n\}]$, one has

$$\begin{aligned}\left\{ \mathbf{X} \in \mathbb{R}^{m \times n} : \varphi_k(\mathbf{X}) \leq t \right\} & \overset{\text{(c)}}{=} \left\{ \mathbf{X} \in \mathbb{R}^{m \times n} : f_k (\Phi(\mathbf{X})) \leq t \right\} \\ & \overset{\text{(d)}}{=} \Big\{ \mathbf{X} \in \mathbb{R}^{m \times n} : \exists \tilde{\mathbf{Z}} \in \mathcal{S}^{m+n} \ \& \ s \in \mathbb{R} \text{ such that } t - ks - \mathsf{Trace}\left(\tilde{\mathbf{Z}}\right) \geq 0; \\ & \qquad \tilde{\mathbf{Z}} \succeq \mathbf{O}_{(m+n)\times(m+n)}; \ \tilde{\mathbf{Z}} - \Phi(\mathbf{X}) + s\mathbf{I}_{m+n} \succeq \mathbf{O}_{(m+n)\times(m+n)} \Big\},\end{aligned}$$

where the step (c) holds due to the equation (2.10), and the step (d) follows by applying the part (1) for $\Phi(\mathbf{X}) \in \mathcal{S}^{m+n}$;

(ii) The function $\varphi_k(\cdot) : \mathbb{R}^{m \times n} \to \mathbb{R}_+$ is a convex function. To this end, we first take a closer inspection on the epigraph $\mathsf{epi}\,(\varphi_k)$ of the function $\varphi_k(\cdot) : \mathbb{R}^{m \times n} \to \mathbb{R}_+$:

$$\begin{aligned}\mathsf{epi}\,(\varphi_k) & = \left\{ (\mathbf{X}, t) \in \mathbb{R}^{m \times n} \times \mathbb{R} : \varphi_k(\mathbf{X}) \leq t \right\} \\ & = \left\{ (\mathbf{X}, t) \in \mathbb{R}^{m \times n} \times \mathbb{R} : f_k (\Phi(\mathbf{X})) \leq t \right\} \\ & = \left\{ (\mathbf{X}, t) \in \mathbb{R}^{m \times n} \times \mathbb{R} : (\Phi(\mathbf{X}), t) \in \mathsf{epi}\,(f_k) \right\}.\end{aligned} \tag{2.11}$$

We know from the part (2) that the epigraph $\mathsf{epi}\,(f_k)$ of the function $f_k(\cdot) : \mathcal{S}^{m+n} \to \mathbb{R}$ is a convex subset of $\mathcal{S}^{m+n} \times \mathbb{R}$ due to the convexity of $f_k(\cdot) : \mathcal{S}^{m+n} \to \mathbb{R}$. As $\Phi(\cdot) : \mathbb{R}^{m \times n} \to \mathcal{S}^{m+n}$ is an $\mathbb{R}$-linear map, it's clear that the map

$$(\mathbf{X}, t) \in \mathbb{R}^{m \times n} \times \mathbb{R} \mapsto (\Phi(\mathbf{X}), t) \in \mathcal{S}^{m+n} \times \mathbb{R} \tag{2.12}$$

is also an $\mathbb{R}$-linear map. So from (2.11), it follows that $\mathsf{epi}\,(\varphi_k)$ is a convex subset of $\mathbb{R}^{m \times n} \times \mathbb{R}$ since it is an inverse image of the convex subset $\mathsf{epi}\,(f_k)$ of $\mathcal{S}^{m+n} \times \mathbb{R}$ under the $\mathbb{R}$-linear map (2.12). Hence, the function $\varphi_k(\cdot) : \mathbb{R}^{m \times n} \to \mathbb{R}_+$ is a convex function;

(iii) Finally, unlike the case of the function $f_k(\cdot)$, the function $\varphi_k(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}_+$ is not only just a convex function, but also a norm on $\mathbb{R}^{m\times n}$:

- **Positive definiteness of the map $\varphi_k(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}_+$:** It's evident that if $\mathbf{X} = \mathbf{O}_{m\times n}$, then $\varphi_k(\mathbf{X}) = 0$ for every $k \in [\min\{m,n\}]$. Conversely, assume that $\varphi_k(\mathbf{X}) = 0$ for $\mathbf{X} \in \mathbb{R}^{m\times n}$. Then we have $\sigma_1(\mathbf{X}) = 0$, which clearly implies that $\sigma_i(\mathbf{X}) = 0$ for all $i \in [\min\{m,n\}]$. Thus, we have $\mathbf{X} = \mathbf{O}_{m\times n}$ by considering its singular value decomposition. This establishes the positive definiteness of the map $\varphi_k(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}_+$;

- **Absolute homogeneity of the map $\varphi_k(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}_+$:** Let $\mathbf{X} = \mathbf{U}\tilde{\mathbf{\Sigma}}\mathbf{V}^\top$ be a singular value decomposition of $\mathbf{X} \in \mathbb{R}^{m\times n}$, where $\mathbf{U} \in \mathcal{O}(m)$, $\mathbf{V} \in \mathcal{O}(n)$, and $\tilde{\mathbf{\Sigma}} := \begin{bmatrix} \mathbf{\Sigma} & \mathbf{O}_{r\times(n-r)} \\ \mathbf{O}_{(m-r)\times r} & \mathbf{O}_{(m-r)\times(n-r)} \end{bmatrix} \in \mathbb{R}^{m\times n}$, $\mathbf{\Sigma} := \mathsf{diag}\left(\sigma_1(\mathbf{X}), \sigma_2(\mathbf{X}), \cdots, \sigma_r(\mathbf{X})\right) \in \mathbb{R}^{r\times r}$. Then for any $\alpha \in \mathbb{R}$, $\alpha\mathbf{X} = (\mathsf{sign}(\alpha)\mathbf{U})\left(|\alpha|\,\tilde{\mathbf{\Sigma}}\right)\mathbf{V}^\top$ is a singular value decomposition of $\alpha\mathbf{X} \in \mathbb{R}^{m\times n}$, where

$$\mathsf{sign}(\alpha) := \begin{cases} +1 & \text{if } \alpha > 0; \\ 0 & \text{if } \alpha = 0; \\ -1 & \text{otherwise.} \end{cases}$$

  Thus, we obtain $\sigma_i(\alpha\mathbf{X}) = |\alpha| \cdot \sigma_i(\mathbf{X})$ for all $i \in [\min\{m,n\}]$. Hence, it holds that for every $k \in [\min\{m,n\}]$,

$$\varphi_k(\alpha\mathbf{X}) = \sum_{i=1}^{k} \sigma_i(\alpha\mathbf{X}) = \sum_{i=1}^{k} |\alpha| \cdot \sigma_i(\mathbf{X}) = |\alpha|\,\varphi_k(\mathbf{X}) \tag{2.13}$$

  for every $\alpha \in \mathbb{R}$ and $\mathbf{X} \in \mathbb{R}^{m\times n}$. This establishes the absolute homogeneity of $\varphi_k(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}_+$;

- **Sub-additivity of the map $\varphi_k(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}_+$:** Choose any matrices $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{m\times n}$. Then we obtain

$$\varphi_k(\mathbf{X}+\mathbf{Y}) \overset{\text{(e)}}{=} 2\varphi_k\left(\frac{\mathbf{X}+\mathbf{Y}}{2}\right) \overset{\text{(f)}}{\le} 2 \cdot \frac{\varphi_k(\mathbf{X}) + \varphi_k(\mathbf{Y})}{2} = \varphi_k(\mathbf{X}) + \varphi_k(\mathbf{Y}),$$

  where the step (e) follows from the absolute homogeneity (2.13) of the map $\varphi_k(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}_+$, and the step (f) makes use of the convexity of the function $\varphi_k(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}_+$ (which was argued in (ii)). Hence, the map $\varphi_k(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}_+$ is sub-additive.

Thanks to the above observations, one can conclude that the function $\varphi_k(\cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}_+$ is a norm on $\mathbb{R}^{m\times n}$, while we have shown that the function $f_k(\cdot) : \mathcal{S}^n \to \mathbb{R}$ cannot be a norm on $\mathcal{S}^n$.

References

[1] Giuseppe C Calafiore and Laurent El Ghaoui. *Optimization models*. Cambridge university press, 2014.