# Homework 7

## Self grades are due at 11 PM one week after the homework.

1. **Symmetric Matrices**

Recall that $\mathbb{R}^{n\times n}$ can be thought of as the vector space of all $n\times n$ matrices. As a vector space, $\mathbb{R}^{n\times n}$ has dimension $n^2$. Let $\mathbb{S}^n \subseteq \mathbb{R}^{n\times n}$ denote the set of symmetric matrices $n\times n$ matrices. Let $\mathbb{S}^n_+ \subseteq \mathbb{S}^n$ denote the set of positive semidefinite $n\times n$ matrices. Let $\mathbb{S}^n_{++} \subseteq \mathbb{S}^n_+$ denote the set of positive definite $n\times n$ matrices.

(a) Show that $\mathbb{S}^n$ is a subspace of $\mathbb{R}^{n\times n}$ of dimension $\binom{n+1}{2}$.

**Solution:**

If $A, B \in \mathbb{S}^n$, i.e. $A$ and $B$ are symmetric $n\times n$ matrices, and $\alpha, \beta$ are arbitrary real numbers, then $\alpha A + \beta B$ is an $n\times n$ symmetric matrix, i.e. $\alpha A + \beta B \in \mathbb{S}^n$. This shows that $\mathbb{S}^n$ is a subspace of $\mathbb{R}^{n\times n}$.

For $1 \le i \le n$, let $E_{ii} \in \mathbb{R}^{n\times n}$ have all entries $0$ except for the $(i,i)$ entry, which is $1$. For $1 \le i < j \le n$, let $E_{ij}$ have entries $0$ except for the $(i,j)$ and $(j.i)$ entries, each of which equals $\frac{1}{\sqrt{2}}$. Then one can check that

$$\{E_{ii}, 1 \le i \le n\} \cup \{E_{ij}, 1 \le i < j \le n\}$$

is an orthonormal basis for $\mathbb{S}^n$. (Here the inner product of two matrices $C, D \in \mathbb{R}^{n\times n}$ is defined to be $\sum_{i=1}^n \sum_{j=1}^n c_{ij} d_{ij}$.) This shows that the dimension of the vector space $\mathbb{S}^n$ is

$$n + \binom{n}{2} = \binom{n+1}{2}.$$

(b) Show that $\mathbb{S}^n_+$ is a convex subset of $\mathbb{R}^{n\times n}$.

**Solution:**

Let $A, B \in \mathbb{S}^n_+$. Then $A, B \in \mathbb{S}^n$, and we have $\vec{x}^T A \vec{x} \ge 0$ and $\vec{x}^T B \vec{x} \ge 0$ for all $\vec{x} \in \mathbb{R}^n$. Hence, for every $\theta \in [0,1]$, we have $\vec{x}^T(\theta A + (1-\theta)B)\vec{x} \ge 0$, and also $\theta A + (1-\theta)B \in \mathbb{S}^n$. This implies that $\theta A + (1-\theta)B \in \mathbb{S}^n_+$.

Since this holds for all $A, B \in \mathbb{S}^n_+$ and for every $\theta \in [0,1]$, we have verified that $\mathbb{S}^n_+$ is a convex subset of $\mathbb{R}^{n\times n}$.

(c) Show that the affine hull of $\mathbb{S}^n_+$ is $\mathbb{S}^n$.

Recall that the affine hull of a subset $A$ of a vector space $V$ is the smallest subspace of $V$ that contains $A$. It can be characterized as the set of all linear combinations of the form $\sum_{i=1}^k \theta_i \vec{x}_i$, where $k \ge 1$ is arbitrary, $\vec{x}_1, \ldots, \vec{x}_k$ are vectors in $A$, and $\theta_1, \ldots, \theta_k$ are arbitrary real numbers satisfying $\sum_{i=1}^k \theta_i = 1$. Note that, in contrast to the definition of the convex hull of $A$, the $\theta_i$ are allowed to be negative.

*HINT: Every symmetric matrix is conjugate to a diagonal matrix by an orthogonal change of basis.*

**Solution:**

We have already shown that $\mathbb{S}^n$ is a subspace of $\mathbb{R}^{n\times n}$ that contains $\mathbb{S}^n_+$. This implies that $\text{aff}(\mathbb{S}^n_+) \subseteq \mathbb{S}^n$. To show that $\mathbb{S}$ is the smallest such subspace, and hence $\mathbb{S}^n = \text{aff}(\mathbb{S}^n_+) \subseteq \mathbb{S}^n$, we need to show that every $A \in \mathbb{S}^n$ can be written as a linear combination of matrices in $\mathbb{S}^n_+$, with coefficients that sum to 1.

Let $A \in \mathbb{S}^n$. We write $A = U^T \Sigma U$, where $U$ is an orthogonal matrix and $\Sigma$ is a diagonal matrix. We then write $\Sigma = \Sigma_1 - \Sigma_2$, where $\Sigma_1$ and $\Sigma_2$ are diagonal matrices, each having only nonnnegative diagonal entries. We can then write

$$A = U^T \Sigma U$$

$$= U^T\Sigma_1 U - U^T\Sigma_2 U$$
$$= 2A_1 - A_2,$$

where $A_1$ is defined to be $\frac{1}{2}U^T\Sigma_1 U$ and $A_2$ is defined to be $U^T\Sigma_2 U$. Note that $A_1, A_2 \in \mathbb{S}^n_+$, and that in the representation of $A$ as $2A_1 - A_2$ the coefficients sum to 1.

Since this procedure can be carried out for all $A \in \mathbb{S}^n$, we have shown that $\mathbb{S}^n$ is the affine hull of $\mathbb{S}^n_+$.

(d) Show that $\mathbb{S}^n_{++}$ is a convex subset of $\mathbb{R}^{n\times n}$.

**Solution:**

Let $A, B \in \mathbb{S}^n_{++}$. Then $A, B \in \mathbb{S}^n$, and we have $\vec{x}^T A\vec{x} > 0$ and $\vec{x}^T B\vec{x} > 0$ for all nonzero $\vec{x} \in \mathbb{R}^n$. Hence, for every $\theta \in [0, 1]$, we have $\vec{x}^T(\theta A + (1-\theta)B)\vec{x} \geq 0$ for all nonzero $\vec{x} \in \mathbb{R}^n$, and also $\theta A + (1-\theta)B \in \mathbb{S}^n$. This implies that $\theta A + (1-\theta)B \in \mathbb{S}^n_{++}$.

Since this holds for all $A, B \in \mathbb{S}^n_{++}$ and for every $\theta \in [0, 1]$, we have verified that $\mathbb{S}^n_{++}$ is a convex subset of $\mathbb{R}^{n\times n}$.

(e) Show that $\mathbb{S}^n_{++}$ is the relative interior of $\mathbb{S}^n_+$. For this problem, to define distances in $\mathbb{R}^{n\times n}$, it does not matter whether you use the Frobenius norm or the induced 2-norm, but use the induced 2-norm.

Recall that the relative interior of a subset $A$ of a vector space $V$ is the interior of $A$ when $A$ is viewed as a subset of its affine hull.

**Solution:**

First, we need to show that for every $A \in \mathbb{S}^n_{++}$ there is some $r > 0$ such that for every $B \in \mathbb{S}^n$ with $\|B\|_2 < r$ we have $A + B \in \mathbb{S}^n_{++}$. The reason it suffices to do this only for $B \in \mathbb{S}^n$ is that we have already shown that $\mathbb{S}^n$ is the affine hull of $\mathbb{S}^n_+$.

Let $\lambda_{\min}$ denote the smallest eigenvalue of $A$. Then, since $A \in \mathbb{S}^n_{++}$, we have $\lambda_{\min} > 0$. Let us choose $0 < r < \lambda_{\min}$. For every $B \in \mathbb{S}^n$ with $\|B\|_2 < r$ we have $A + B \in \mathbb{S}^n$. Further, for every $\vec{x} \in \mathbb{R}^n$, we have

$$\begin{aligned}
\vec{x}^T(A+B)\vec{x} &= \vec{x}^T A\vec{x} + \vec{x}^T B\vec{x} \\
&\geq \lambda_{\min}\|\vec{x}\|_2^2 - r\|\vec{x}\|_2^2 \\
&= (\lambda_{\min} - r)\|\vec{x}\|_2,
\end{aligned}$$

where the inequality follows, using the Cauchy-Schwarz inequality, because

$$|\vec{x}^T B\vec{x}| \leq \|\vec{x}\|_2\|B\vec{x}\|_2 \leq \|\vec{x}\|_2(r\|\vec{x}\|_2) = r\|\vec{x}\|_2^2.$$

Since $\lambda_{\min} - r > 0$ we have shown that $\vec{x}^T(A+B)\vec{x} > 0$ for all nonzero $\vec{x} \in \mathbb{R}^n$. This establishes that $A + B \in \mathbb{S}^n_{++}$, which is what we wanted to show.

Second, we need to show that for every $A \in \operatorname{relint}(\mathbb{S}^n_+)$, we have $A \in \mathbb{S}^n_{++}$. We proceed by contradiction: Suppose $A \notin \mathbb{S}^n_{++}$. If $A \notin \mathbb{S}^n_+$, then we automatically have $A \notin \operatorname{relint}(\mathbb{S}^n_+)$, as desired. Otherwise, $A \in \mathbb{S}^n_+$, or in words, $A$ is symmetric positive semi-definite but not symmetric positive definite. In particular, there exists some vector $\vec{v} \in N(A)$ with $\|\vec{v}\|_2 = 1$. We now claim that for each $r > 0$, there exists some matrix $B \in \mathbb{S}^n = \operatorname{aff}(\mathbb{S}^n_+)$, with $\|B\|_2 < r$, such that $A + B \notin \mathbb{S}^n_+$; this claim would establish that $A \notin \operatorname{relint}(\mathbb{S}^n_+)$, as desired. To show this, fix $r > 0$ arbitrarily, and set $B = -\frac{1}{2}r\vec{v}\vec{v}^\top$. Then $B \in \mathbb{S}^n$ with $\|B\|_2 = \frac{1}{2}r$, but $A + B \notin \mathbb{S}^n_+$, because $(A+B)\vec{v} = 0 - \frac{1}{2}r\vec{v} = -\frac{1}{2}r\vec{v}$, i.e., $-\frac{1}{2}r$ is an eigenvalue of $A + B$.

(f) Show that if $n > 1$ then the interior of $\mathbb{S}^n_+$ is empty. Here again, to define distances in $\mathbb{R}^{n\times n}$, it does not matter whether you use the Frobenius norm or the induced 2-norm, but use the induced 2-norm.

**Solution:**

Let $n > 1$ be given, and let $A \in \mathbb{S}^n_+$. No matter how small we make $r > 0$, we can find $B \in \mathbb{R}^{n\times n}$ with $\|B\|_2 < r$ such that $A + B \notin \mathbb{S}^n_+$. To see this, it suffices to define $B$ such that its $(1, 2)$-entry is $\frac{r}{2}$, and its other entries are all 0. Indeed, with this choice we have $\|B\|_2 = \frac{r}{2}$ and $A + B$ will not be symmetric, so it will not be in $\mathbb{S}^n_+$.

**2. Distance between polytopes as a quadratic program**

Let $\vec{p}^{(1)}, \ldots, \vec{p}^{(r)}$ and $\vec{q}^{(1)}, \ldots, \vec{q}^{(s)}$ be vectors in $\mathbb{R}^d$, where $r, s \geq 1$. Let $\mathcal{P}$ denote the polytope defined as the convex hull of $\{\vec{p}^{(1)}, \ldots, \vec{p}^{(r)}\}$, and $\mathcal{Q}$ the polytope defined as the convex hull of $\{\vec{q}^{(1)}, \ldots, \vec{q}^{(s)}\}$. Thus every point in $\mathcal{P}$ can be written as $\sum_{i=1}^{r} x_i \vec{p}^{(i)}$ for some $x_i \geq 0$, $1 \leq i \leq r$ such that $\sum_{i=1}^{r} x_i = 1$, and every point in $\mathcal{Q}$ can be written as $\sum_{j=1}^{s} x_{r+j} \vec{q}^{(j)}$ for some $x_j \geq 0$, $r+1 \leq j \leq r+s$ such that $\sum_{j=r+1}^{r+s} x_j = 1$. Let us define $n = r + s$.

Define the matrix $C \in \mathbb{R}^{d \times n}$ whose $i$-th column is $\vec{p}^{(i)}$, $1 \leq i \leq r$ and whose $r+j$-th column is $-\vec{q}^{(j)}$, $1 \leq j \leq s$.

Pose the problem of finding the minimum squared $\ell_2$ distance between points in $\mathcal{P}$ and points in $\mathcal{Q}$ as a quadratic program with objective function $\|C\vec{x}\|_2^2$, viewed as a function on $\mathbb{R}^n$.

*NOTE*: A quadratic program is a convex optimization problem where the objective function is a quadratic function and the constraints are linear equalities and inequalities. Recall that a quadratic convex function on $\mathbb{R}^n$ is one of the form $\vec{x}^T H \vec{x} + \vec{a}^T \vec{x} + \vec{b}$ where $b \in \mathbb{R}$, $\vec{a} \in \mathbb{R}^n$, and $H$ is a positive semidefinite matrix in $\mathbb{R}^{n \times n}$ (i.e. $H \in \mathbb{S}_+^n$).

**Solution:**

$$
\begin{aligned}
\min_{\vec{x}} \quad & \vec{x}^T C^T C \vec{x} \\
\text{subject to:} \quad & \sum_{i=1}^{r} x_i = 1, \\
& \sum_{j=1}^{s} x_{r+j} = 1, \\
& x_i \geq 0, \; i = 1, \ldots, n,
\end{aligned}
$$

where we recall that $n := r + s$. To see this, note that the objective is

$$
\|\sum_{i=1}^{r} x_i \vec{p}^{(i)} - \sum_{j=1}^{s} x_{r+j} \vec{q}^{(j)}\|_2^2,
$$

which is the squared $\ell_2$ distance between the points $\sum_{i=1}^{r} x_i \vec{p}^{(i)}$ and $\sum_{j=1}^{s} x_{r+j} \vec{q}^{(j)}$, which lie in $\mathcal{P}$ and $\mathcal{Q}$ respectively, because of the constraints on $\vec{x}$. Further, as $\vec{x}$ ranges over the feasible set, the pair $(\sum_{i=1}^{r} x_i \vec{p}^{(i)}, \sum_{j=1}^{s} x_{r+j} \vec{q}^{(j)})$ ranges over all possible pairs of point in $\mathcal{P} \times \mathcal{Q}$ (possibly with redundancy).