Discussion 7

1. **(Sp '19 Midterm 2 #3) Convexity of Sets**

   Determine if each set C given below is convex. Prove that each set is convex or provide an example to show that it is not convex. You may use any techniques used in class or discussion to demonstrate or disprove convexity.

   (a) $C = \{\vec{x} \in \mathbb{R}^2 \mid x_1 x_2 \geq 0\}$, where $\vec{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top$.

   (b) $C = \{X \in \mathbb{S}^n \mid \lambda_{\min}(X) \geq 2\}$, where $\mathbb{S}^n$ is the set of symmetric matrices in $\mathbb{R}^{n \times n}$ and $\lambda_{\min}(X)$ is the minimum eigenvalue of X.

   (c) Let $\mathcal{H}(\vec{w})$ denote the hyperplane with normal direction $\vec{w} \in \mathbb{R}^n$, i.e.,

   $$\mathcal{H}(\vec{w}) = \{\vec{x} \in \mathbb{R}^n \mid \vec{x}^\top \vec{w} = 0\}. \tag{1}$$

   Let $P : \mathbb{R}^n \to \mathbb{R}^n$ be given by

   $$P(\vec{x}) = \operatorname*{argmin}_{\vec{y} \in \mathcal{H}(\vec{w})} \|\vec{y} - \vec{x}\|_2 . \tag{2}$$

   Let

   $$C = \{P(\vec{x}) \mid \vec{x} \in \mathcal{B}\} \tag{3}$$

   where $\mathcal{B} = \{\vec{x} \in \mathbb{R}^n \mid \|\vec{x}\|_2 \leq 1\}$.

2. (Sp '20 Midterm # 5) Subspace Projection

Consider a set of points $\vec{z}_1, \ldots, \vec{z}_n \in \mathbb{R}^d$. The first principal component of the data, $\vec{w}^\star$, is the direction of the line that minimizes the sum of the squared distances between the points and their projections on $\vec{w}^\star$, i.e.,

$$\vec{w}^\star = \underset{\|\vec{w}\|_2=1}{\operatorname{argmin}} \sum_{i=1}^{n} \|\vec{z}_i - \langle \vec{w},\ \vec{z}_i \rangle \vec{w}\|^2 .$$

In this problem, we generalize to finding the r-dimensional subspace (instead of a 1-dimensional line) that minimizes the sum of the squared distances between the points $\vec{z}_i$ and their projections on the subspace. We assume that $1 \leq r \leq \min(n, d)$. We can represent an r-dimensional subspace by an orthonormal basis $(\vec{w}_1, \ldots, \vec{w}_r)$, and we want to solve:

$$(\vec{w}_1^\star, \ldots, \vec{w}_r^\star) = \underset{\substack{\|\vec{w}_i\|_2=1 \\ \langle \vec{w}_i,\ \vec{w}_j \rangle = 0\ \forall i \neq j \\ 1 \leq i,j \leq r}}{\operatorname{argmin}} \sum_{i=1}^{n} \min_{\alpha_1, \ldots, \alpha_r} \left\| \vec{z}_i - \sum_{k=1}^{r} \alpha_k \vec{w}_k \right\|^2 . \tag{4}$$

Note that the inner minimization projects the point $\vec{z}_i$ onto the subspace defined by $(\vec{w}_1, \ldots, \vec{w}_r)$. The variables $\alpha_k \in \mathbb{R}$. This means that for an arbitrary point $\vec{z}$, this inner minimization

$$(\alpha_1^\star, \ldots, \alpha_r^\star) = \underset{\alpha_1, \ldots, \alpha_r}{\operatorname{argmin}} \left\| \vec{z} - \sum_{k=1}^{r} \alpha_k \vec{w}_k \right\|^2$$

has minimizers $\alpha_k^\star = \langle \vec{w}_k,\ \vec{z} \rangle$.

(a) With the following definition of matrices Z and W:

$$Z = \begin{bmatrix} \uparrow & \cdots & \uparrow \\ \vec{z}_1 & \ldots & \vec{z}_n \\ \downarrow & \cdots & \downarrow \end{bmatrix}, \qquad W = \begin{bmatrix} \uparrow & \cdots & \uparrow \\ \vec{w}_1 & \ldots & \vec{w}_r \\ \downarrow & \cdots & \downarrow \end{bmatrix},$$

show that we can rewrite the optimization problem in Equation (4) as:

$$(\vec{w}_1^\star, \ldots, \vec{w}_r^\star) = \underset{\substack{\|\vec{w}_i\|_2=1 \\ \langle \vec{w}_i,\ \vec{w}_j \rangle = 0\ \forall i \neq j \\ 1 \leq i,j \leq r}}{\operatorname{argmin}} \left\| Z - WW^\top Z \right\|_F^2 . \tag{5}$$

Next, we will solve the optimization problem in Equation (5) using the SVD of Z.

(b) Let σ_i refer to the i^{th} largest singular value of Z, and $l = \min(n, d)$. First **show** that,

$$\min_{\substack{\|\vec{w}_i\|_2=1 \\ \langle \vec{w}_i,\ \vec{w}_j \rangle = 0\ \forall i \neq j \\ 1 \leq i,j \leq r}} \left\| Z - WW^\top Z \right\|_F^2 \geq \sum_{i=r+1}^{l} \sigma_i^2 .$$

(c) Again σ_i refers to the i^{th} largest singular value of Z, and $l = \min(n, d)$. **Show** that,

$$\min_{\substack{\|\vec{w}_i\|_2=1 \\ \langle \vec{w}_i,\ \vec{w}_j \rangle = 0\ \forall i \neq j \\ 1 \leq i,j \leq r}} \left\| Z - WW^\top Z \right\|_F^2 \leq \sum_{i=r+1}^{l} \sigma_i^2 .$$

Hint: Find a W that achieves this upper bound.