# Homework 1

A bonus of 20% of the full score will be awarded for the assignment submitted two days before the due date, November 28, 2023. The total score will not exceed 20 points.

1. (7 points)

   Suppose we have a set of data points $X = \{x_1, \dots, x_m\}$ in $\mathbb{R}^n$.

   (a) (2 points) Let $x_c$ be the centroid of $X$, i.e., the average of all the data points. We want to find the point $y$ in the convex hull $H$ of $X$ that is farthest from $x_c$.

   Formulate the optimization problem to find $y$ and prove that it is a convex optimization problem.

   Prove that the optimal solution is a vertex of the convex hull $H$.

   (b) (1 point) Let $u$ be a vector that is orthogonal to $(y - x_c)$, where $y$ is the optimal solution from (a). We want to find the point $x$ in the convex hull $H$ that maximizes the inner product $(x - x_c)^T u$.

   Formulate the optimization problem to find $x$ and prove that the optimal solution is another vertex of the convex hull $H$.

   (c) (2 points) Let $E_x = \{x \in \mathbb{R}^n : (x-c)^T Q(x-c) \leq n\}$ be the ellipsoid with the smallest volume that contains all the points in $X$.

   Formulate the optimization problem to find $E_x$, and prove that the convex hull $H$ of $X$ is contained in $E$.

   (d) (1 point) Define $z_k = [x_k, 1]^T$, and $E_z = \{z \in \mathbb{R}^{n+1} : z^T P z \leq n+1\}$ the ellipsoid with smallest volume which contains all $z_k$.

   Prove that

   $$P = \begin{bmatrix} Q & -Qc \\ -(Qc)^T & 1 + c^T Qc \end{bmatrix}$$

   (e) (1 point) Show an example to illustrate finding the ellipsoid $E_x$ from $E_z$.

2. Derive optimality conditions and analytical solution (if it exists) of the following optimization problems, solve them. Demonstrate the problem and compare the results with those obtained by CVX or CVXpy.

   $\boldsymbol{b} \in \mathbb{R}^n$ is vectorization of an image corrupted by Gaussian noise, $\mathbf{A}$ is of size $n \times m$, $m > n$, $\boldsymbol{x} \in \mathbb{R}^m$. $\mathbf{A}$ can be concatenation of discrete cosine transform matrix and wavelets transform operators.

(a) (2 points)

$$\min_{\boldsymbol{x}} \quad \|\boldsymbol{b} - \boldsymbol{x}\|_2^2 \quad \text{s.t.} \quad \boldsymbol{x} \geq 0,\ \mathbf{1}^T\boldsymbol{x} = 1$$

(b) (2 points)

$$\min_{\boldsymbol{x}} \quad \|\boldsymbol{b} - \boldsymbol{x}\|_2^2 \quad \text{s.t.} \quad \boldsymbol{x} \geq 0,\ \|\boldsymbol{x}\|_\infty \leq 1$$

(c) (2 points)

$$\min_{\boldsymbol{x}} \quad \|\boldsymbol{b} - \mathbf{A}\boldsymbol{x}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{x}\|_2^2 \leq 1$$

3. (4 points) **Clustering**

   Consider images in the MNIST dataset for three digits 0, 1, and 2. Denote by $\mathbf{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_K]$ a matrix of vectorization of the images.

   (a) Solve the following optimization problem to find an orthogonal matrix $\mathbf{U}$ and a feature matrix, $\mathbf{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K]$, of size $R \times K$

   $$\begin{aligned} \min \quad & \|\mathbf{Y} - \mathbf{U}\mathbf{X}\|_F^2 \\ \text{s.t.} \quad & \mathbf{U}^T\mathbf{U} = \mathbf{I}_R \\ & \mathbf{X} \geq 0, \quad \mathbf{X}^T\mathbf{1}_R = \mathbf{1}_K \end{aligned}$$

   (b) On the basis of extracted features, $\boldsymbol{x}_k$, apply the K-means algorithm to predict categorical labels of images.

4. (2 points) Solve the following optimization problem by applying optimality condition

   $$\begin{aligned} \min_{\boldsymbol{x} \in \mathbb{R}^n} \quad & \sum_{i=1}^{n} \alpha_i \ln x_i \\ \text{s.t.} \quad & \sum_{i=1}^{n} x_i = 1, \quad \boldsymbol{x} \geq 0 \end{aligned}$$

   where $\alpha_i < 0$.

5. (2 points) Prove that the function $f(x) = \log(\sum_n \exp(x_k))$

   - is convex
   - but not strictly convex