Convex Optimization HW 1

Brandyn Tucknott

Due 17 Oct. 2025

1. Assume that C is an affine set. By definition, we know that for any $x_1, x_2 \in C$, we have

$$\theta x_1 + (1-\theta)x_2 \in C, \text{ for all } \theta \in \mathbb{R}.$$

   Building upon the definition, show that if $x_i \in C$ for $i = 1, \ldots, n$, then we have

$$\theta_1 x_1 + \ldots + \theta_n x_n,$$

   where $\sum_{i=i}^{n} \theta_i = 1$.

2. Answer the following questions:
   (a) What is the distance between two parallel hyperplanes. i.e., $\{x | a^T x = b\}$ and $\{x | a^T x = c\}$?
   (b) Let a, b be distinct points in $\mathbb{R}^n$. Show that the set of all points that are closer to a than b, i.e., $\{x | \, \|x - a\|_2 \leq \|x - b\|_2\}$, is a halfspace. Describe it explicitly as an inequality of the form $c^T x \leq d$. Draw a picture.

3. Which of the following sets are convex?
   (a) A slab $\{x \in \mathbb{R}^n | \alpha \leq a^T x \leq \beta\}$.
   (b) A rectangle $\{x \in \mathbb{R}^n | \alpha_i \leq x_i \leq \beta_i, i = 1, \ldots, n\}$.
   (c) The set of points closer to a given point than a given set:

$$\{x | \, \|x - x_0\|_2 \leq x - y_2 \text{ for all } y \in S\}$$

   where $S \subset \mathbb{R}^n$.
   (d) The set of points whose distance to a does not exceed a fixed fraction θ of the distance to b, i.e. the set $\{x | \, \|x - a\|_2 \leq \theta \, \|x - b\|_2\}$. You can assume $a \neq b$ and $\theta \leq 1$.

4. Show the following statements.
   (a) A polyhedron, i.e. $P = \{x | Ax \succeq b, Cx = d\}$ where $A \in \mathbb{R}^{mxn}$ and $C \in \mathbb{R}^{pxn}$ is a convex set.
   (b) Consider an ellipsoid $\epsilon = \{x | (x - x_c)^T P^{-1} (x - x_c) \leq 1\}$. Assume that the eigenvalues of $P \in \mathbb{R}^{nxn}$ is $\lambda_1^2, \ldots, \lambda_n^2$ in descending order. Show that the largest and smallest distances from any point on the boundary of the ellipsoid to x_c are λ_1 and λ_n respectively.

5. Show the following statements.
   (a) In machine learning, we are often given training samples in the form of (x_i, y_i) for $i = 1, \ldots, n$ where $x_i \in \mathbb{R}^d$ is the feature vector and $y_i \in \mathbb{R}$ is the label of this example. The empirical risk of Euclidean distance based linear regression can be expressed as follows:

$$f(a) = \frac{1}{n} \sum_{i=1}^{n} \left(y_i - a^T x_i\right)^2.$$

   Show the function $f(a)$ is convex in a.

(b) Suppose $p < 1$, $p \neq 0$. Show that the function

$$f(x) = \left(\sum_{i=1}^{n} x_i^p \right)^{1/p}$$

with $\operatorname{dom}(f) = \mathbb{R}_{++}^n$ is concave.

(c) Show that $f(X) = \operatorname{tr}\left(X^{-1}\right)$ is convex on $\operatorname{dom}(f) = \mathbb{S}_{++}^n$.

6. Show the conjugate of $f(X) = \operatorname{tr}\left(X^{-1}\right) with \operatorname{dom}(f) = \mathbb{S}_{++}^n$ is given by

$$f^*(Y) = -2\operatorname{tr}(-Y)^{1/2}, \operatorname{dom}(f) = -\mathbb{S}_{++}^n.$$

(Hint: for unconstrained and differentiable convex problems, min / max can be found by looking for where the function has zero gradient.)

7. Show that the following function is convex.

$$f(x) = x^T (A(x))^{-1} x, \operatorname{dom}(f) = \{x | A(x) \succ 0\},$$

where $A(x) = A_0 + A_1 x_1 + \ldots + A_n x_n \in \mathbb{S}^n$ and $A_i \in \mathbb{S}^n$, $i = 1, \ldots, n$. Hint: you are allowed to use a special form of Schur complement, described as follows: Suppose $A \succ 0$. then

$$\begin{pmatrix} A & b \\ b^T & c \end{pmatrix} \succ 0 \leftrightarrow c - b^T A^{-1} b \geq 0.$$

You will need to study "epigraph" from chapter 3 of the textbook to answer this question.