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

Proof. Given an affine set C, let $x_0 \in C$ be arbitrary, and recall that $V = C - x_0$ is a subspace. Then for all $x_i \in C$ and given $\sum_{i=1}^n \theta_i = 1$ for $i = 1, \ldots, n$. Observe that

$$\sum_{i=1}^n \theta_i(x_i - x_0) \in V$$
$$\sum_{i=1}^n \theta_i(x_i - x_0) + x_0 \in C$$
$$\sum_{i=1}^n \theta_i x_i - \sum_{i=1}^n \theta_i x_0 + x_0 \in C$$
$$\sum_{i=1}^n \theta_i x_i - x_0 \sum_{i=1}^n \theta_i + x_0 \in C$$
$$\sum_{i=1}^n \theta_i x_i - x_0 \cdot 1 + x_0 \in C$$
$$\sum_{i=1}^n \theta_i x_i \in C.$$

□

2. Answer the following questions:

(a) What is the distance between two parallel hyperplanes. i.e., $\left\{x | a^T x = b\right\}$ and $\left\{x | a^T x = c\right\}$?

Proof. Observe that

$$\left\{x | a^T x = b\right\}, \left\{x | a^T x = c\right\} \text{ is equivalent to } \left\{x | a^T x = |b-c|\right\} \left\{a^T x = 0\right\},$$

which geometrically gives us one hyperplane passing through the origin, and the other parallel to it. Since the shortest path from the origin to the hyperplane is a vector x_0 going directly to it (same direction as normal vector a), we can conclude that

$$a^T x_0 = |b-c|$$

$$\frac{a^T}{\|a\|} x_0 = \frac{|b-c|}{\|a\|}. \tag{1}$$

Recall that

$$\left(a^T / \|a\|\right) x_0 = \left\| \frac{a^T}{\|a\|} \right\| \|x_0\| \cos\theta = 1 \cdot \|x_0\| \cdot 1 \tag{2}$$

since $(a^T / \|a\|) x_0$ is a dot product. By (1) and (2) we conclude that

$$\|x_0\| = \frac{|b-c|}{\|a\|}.$$

□

(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.

Proof. To show the given set is a halfspace, we only need to be able to express it in form $c^T x \leq d$, and then we will be done. We can algebraically manipulate the given equation:

$$\begin{aligned}
\|x-a\|_2 &\leq \|x-b\|_2 \\
\|x-a\|_2^2 &\leq \|x-b\|_2^2 \\
x^T x - 2x^T a - a^T a &\leq^T x - 2x^T b + b^T b \\
-2x^T a + \|a\|_2^2 &\leq -2x^T b + \|b\|_2^2 \\
2(x^T b - x^T a) &\leq \|b\|_2^2 - \left\|a_2^2\right\| \\
x^T (b-a) &\leq \frac{\|b\|_2^2 - \|a\|_2^2}{2} \\
(b-a)^T x &\leq \frac{\|b\|_2^2 - \|a\|_2^2}{2}.
\end{aligned}$$

Since we have shown the given constraint is equivalent to the definition of a subspace, we are done. □

3. Which of the following sets are convex?
   (a) A slab $\left\{x \in \mathbb{R}^n | \alpha \leq a^T x \leq \beta\right\}$.
   (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:
   $$\{x| \left\|x - x_0\right\|_2 \leq x - y_2 \text{ for all } y \in S\}$$
   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 $\{x| \left\|x - a\right\|_2 \leq \theta \left\|x - b\right\|_2\}$. 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 $\epsilon = \left\{x|(x-x_c)^T P^{-1}(x-x_c) \leq 1\right\}$. 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 $\mathrm{dom}\,(f) = \mathbb{R}_{++}^n$ is concave.

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

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

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

   (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.

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

   where $A(x) = A_0 + A_1x_1 + \ldots + A_nx_n \in \mathbb{S}^n$ 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.