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 $i = 1, \dots, n$, 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 $i = 1, \dots, n$. Observe that

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

□

2. Answer the following questions:

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

Proof. Observe that

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

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

$$a^Tx_0 = |b-c|$$
$$\frac{a^T}{\|a\|}x_0 = \frac{|b-c|}{\|a\|}. \tag{1}$$

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 $c^Tx \leq d$. Draw a picture.

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

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

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

3. Which of the following sets are convex?

(a) A slab $\{x \in \mathbb{R}^n | \alpha \leq a^T x \leq \beta\}$.

Proof. Since $S = \{x \in \mathbb{R}^n | \alpha \leq a^T x \leq \beta\}$ is the intersection of two halfspaces (which are convex), we know that S is convex. □

(b) A rectangle $\{x \in \mathbb{R}^n | \alpha_i \leq x_i \leq \beta_i, i = 1, \ldots, n\}$.

Proof. For $S = \{x \in \mathbb{R}^n | \alpha_i \leq x_i \leq \beta_i, i = 1, \ldots, n\}$, let $y, z \in S$. Observe that

$$\alpha_i \leq y_i \leq \beta_i,$$

$$\alpha_i \leq z_i \leq \beta_i.$$

Then

$$\theta\alpha_i \leq \theta y_i \leq \theta\beta_i,$$
$$(1-\theta)\alpha_i \leq (1-\theta) z_i \leq (1-\theta)\beta_i,$$

which we add to see that

$$\begin{aligned} \theta\alpha_i + (1-\theta)\alpha_i \leq \theta y_i + (1-\theta) z_i &\leq \theta z_i + (1-\theta) z_i \\ \alpha_i \leq \theta y_i + (1-\theta) z_i &\leq \beta_i. \end{aligned}$$

Since the convex combination is in S, we conclude S is convex. □

(c) The set of points closer to a given point than a given set:

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

where $S \subset \mathbb{R}^n$.

Proof. We can manipulate the given constraint into:

$$\begin{aligned} \left\|x - x_0\right\|_2 &\leq \left\|x - y\right\|_2 \\ \left\|x - x_0\right\|_2^2 &\leq \left\|x - y\right\|_2^2 \\ x^T x - 2x^T x_0 + x_0^T x_0 &\leq x^T x - 2x^T y + y^T y \\ 2x^T (y - x_0) &\leq \left\|y\right\|_2^2 - \left\|x_0\right\|_2^2 \\ (y - x_0)^T x &\leq \frac{\left\|y\right\|_2^2 - \left\|x_0\right\|_2^2}{2}. \end{aligned}$$

Since any individual y yields a halfspace, to consider all y we look at the intersection of the halfspaces, which we know to be convex. □

(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 $\theta \leq 1$.

Proof. Again, we manipulate the constraint:

$$
\begin{aligned}
\|x-a\|_2 &\leq \theta \|x-b\|_2 \\
\|x-a\|_2^2 &\leq \theta^2 \|x-b\|_2^2 \\
x^T x - 2x^T a + a^T a &\leq \theta^2 \left(x^T x - 2x^T b + b^T b\right) \\
x^T x - \theta^2 x^T x - 2x^T a + 2\theta^2 x^T b &\leq \theta^2 b^T b - a^T a \\
(1-\theta^2) x^T x + 2x^T(\theta^2 b - a) &\leq \|b\|_2^2 - \|a\|_2^2 \\
x^T x + \frac{2x^T(\theta^2 b - a)}{1-\theta^2} &\leq \frac{\|b\|_2^2 - \|a\|_2^2}{1-\theta^2} \\
x^T x + \frac{2x^T * \theta^2 b - a}{1-\theta^2} + \frac{1}{(1-\theta^2)^2} x^T(\theta^2 b - a) &\leq \frac{\|b\|_2^2 - \|a\|_2^2}{1-\theta^2} + \frac{\frac{1}{1-\theta^2}\left\|\theta^2 b - a\right\|_2^2}{1-\theta^2} \\
\left\|x + \frac{1}{1-\theta^2}(\theta^2 b - a)\right\|_2^2 &\leq \frac{\theta^2 \|b\|_2^2 - \|a\|_2^2 + \frac{1}{1-\theta^2}\left\|\theta^2 b - a\right\|_2^2}{1-\theta^2}.
\end{aligned}
$$

We have rewritten our constraint into the form for a ball (since RHS is constant when fixing a, b, θ), and we already know the ball to be convex. Thus our original set is convex. $\square$

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.

   Proof. Let $x, y \in P, \theta \in [0,1]$, and $z = \theta x + (1-\theta)y$. Then

$$\begin{aligned} Az &= A(\theta x + (1-\theta)y) \\ &= \theta Ax + (1-\theta)Ay \\ &\succeq 0 \end{aligned}$$

   since $Ax \succeq 0, Ay \succeq 0, \theta, 1-\theta \geq 0$. We also have that

$$\begin{aligned} Cz &= C(\theta x + (1-\theta)y) \\ &= \theta Cx + (1-\theta)Cy \\ &= \theta d + (1-\theta)d \\ &= d. \end{aligned}$$

   We conclude P is convex. □

   (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.

   Proof. Let $P = Q\Lambda Q^T$ where Q is orthogonal. Let also $y = Q^T(x - x_c)$, and note that $P^{-1} = (Q\Lambda Q^T)^{-1} = Q\Lambda^{-1}Q^T$. Then we can rewrite

$$\begin{aligned} (x-x_c)^T P^{-1}(x-x_c) &= (x-x_c)^T Q\Lambda^{-1}Q^T(x-x_c) \\ &= y^T\Lambda^{-1}y \\ &= \sum_{i=1}^n \frac{y_i^2}{\lambda_i^2} = 1. \end{aligned}$$

   If we let $z_i = y_i/\lambda_i$ (for ease of representation), then we end up with $sum_{i=1}^n z_i^2 = 1$. Now observe that

$$\begin{aligned} \|y\|_2^2 &= y^T y \\ &= (x-x_c)^T QQ^T(x-x_c) \\ &= (x-x_c)^T(x-x_c) \\ &= \|x-x_c\|^2. \end{aligned}$$

   Since $\|x - x_c\| = \|y\|$, we also have that

$$\|x-x_c\|^2 = \sum_{i=1}^n y_i^2 = \sum_{i=1}^n \lambda_i^2 z_i^2.$$

   Since $\lambda_n^2 \leq \ldots \leq \lambda_1^2$, we have that

$$\begin{aligned} \sum_{i=1}^n \lambda_n^2 \leq \sum_{i=1}^n \lambda_i^2 z_i^2 &\leq \sum_{i=1}^n \lambda_1^2 z_i^2 \\ \lambda_n^2 \leq \sum_{i=1}^n \lambda_i^2 z_i^2 &\leq \lambda_1^2 \\ \lambda_n^2 \leq \|x-x_c\|^2 &\leq \lambda_1^2 \\ \lambda_n \leq \|x-x_c\| &\leq \lambda_1. \end{aligned}$$

   We conclude that the largest and smallest distances to a point on the boundary of the ellipse correspond to the largest and smallest eigenvalues 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 $A(x) = A_0 + A_1 x_1 + \ldots + A_n x_n \in \mathbb{S}^n$ 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.