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

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

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

$$\sum_{i=1}^{n} \theta_i x_i - \sum_{i=1}^{n} \theta_i x_0 + x_0 \in C$$

$$\sum_{i=1}^{n} \theta_i x_i - x_0 \sum_{i=1}^{n} \theta_i + x_0 \in C$$

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

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

$\square$

2. Answer the following questions:

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

Proof. Observe that

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

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

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

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

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

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

$\square$

(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:

$$\|x - a\|_2 \leq \|x - b\|_2$$
$$\|x - a\|_2^2 \leq \|x - b\|_2^2$$
$$x^T x - 2x^T a - a^T a \leq^T x - 2x^T b + b^T b$$
$$-2x^T a + \|a\|_2^2 \leq -2x^T b + \|b\|_2^2$$
$$2(x^T b - x^T a) \leq \|b\|_2^2 - \|a_2^2\|$$
$$x^T (b - a) \leq \frac{\|b\|_2^2 - \|a\|_2^2}{2}$$
$$(b - a)^T x \leq \frac{\|b\|_2^2 - \|a\|_2^2}{2}.$$

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

3. Which of the following sets are convex?

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

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

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

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

$$\alpha_i \le y_i \le \beta_i,$$

$$\alpha_i \le z_i \le \beta_i.$$

        Then

$$\theta\alpha_i \le \theta y_i \le \theta\beta_i,$$

$$(1-\theta)\alpha_i \le (1-\theta)z_i \le (1-\theta)\beta_i,$$

        which we add to see that

$$\theta\alpha_i + (1-\theta)\alpha_i \le \theta y_i + (1-\theta)z_i \le \theta z_i + (1-\theta)z_i$$

$$\alpha_i \le \theta y_i + (1-\theta)z_i \le \beta_i.$$

        Since the convex combination is in S, we conclude S is convex. $\qquad\square$

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

$$\{x | \; \|x - x_0\|_2 \le \|x - y\|_2 \text{ for all } y \in S\}$$

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

        Proof. We can manipulate the given constraint into:

$$\|x - x_0\|_2 \le \|x - y\|_2$$

$$\|x - x_0\|_2^2 \le \|x - y\|_2^2$$

$$x^T x - 2x^T x_0 + x_0^T x_0 \le x^T x - 2x^T y + y^T y$$

$$2x^T(y - x_0) \le \|y\|_2^2 - \|x_0\|_2^2$$

$$(y - x_0)^T x \le \frac{\|y\|_2^2 - \|x_0\|_2^2}{2}.$$

        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. $\qquad\square$

    (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 $\{x | \; \|x - a\|_2 \le \theta \|x - b\|_2\}$. You can assume $a \ne b$ and $\theta \le 1$.

Proof. Again, we manipulate the constraint:

$$\|x - a\|_2 \le \theta \|x - b\|_2$$
$$\|x - a\|_2^2 \le \theta^2 \|x - b\|_2^2$$
$$x^T x - 2x^T a + a^T a \le \theta^2 \left(x^T x - 2x^T b + b^T b\right)$$
$$x^T x - \theta^2 x^T x - 2x^T a + 2\theta^2 x^T b \le \theta^2 b^T b - a^T a$$
$$(1 - \theta^2)x^T x + 2x^T(\theta^2 b - a) \le \|b\|_2^2 - \|a\|_2^2$$
$$x^T x + \frac{2x^T(\theta^2 b - a)}{1 - \theta^2} \le \frac{\|b\|_2^2 - \|a\|_2^2}{1 - \theta^2}$$
$$x^T x + \frac{2x^T * \theta^2 b - a}{1 - \theta^2} + \frac{1}{(1 - \theta^2)^2}x^T(\theta^2 b - a) \le \frac{\|b\|_2^2 - \|a\|_2^2}{1 - \theta^2} + \frac{\frac{1}{1-\theta^2}\left\|\theta^2 b - a\right\|_2^2}{1 - \theta^2}$$
$$\left\|x + \frac{1}{1 - \theta^2}(\theta^2 b - a)\right\|_2^2 \le \frac{\theta^2 \|b\|_2^2 - \|a\|_2^2 + \frac{1}{1-\theta^2}\left\|\theta^2 b - a\right\|_2^2}{1 - \theta^2}.$$

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. $\qquad\square$

4. Show the following statements.

    (a) A polyhedron, i.e. $P = \{x | Ax \succeq b, Cx = d\}$ where $A \in \mathbb{R}^{m \times n}$ and $C \in \mathbb{R}^{p \times n}$ 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 $P \in \mathbb{R}^{n \times n}$ 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

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

(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.

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

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.