

# 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

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

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

*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.,  $\{x|a^T x = b\}$  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

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

Recall that

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

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:

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

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.  $\square$

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

*Proof.* For  $S = \{x \in \mathbb{R}^n | \alpha_i \leq x_i \leq \beta_i, i = 1, \dots, n\}$ , 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

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

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

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

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

*Proof.* We can manipulate the given constraint into:

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

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

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

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

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

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.  $\square$

(d) The set of points whose distance to  $a$  does not exceed a fixed fraction  $\theta$  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  $\theta \leq 1$ .

*Proof.* Again, we manipulate the constraint:

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

We have rewritten our constraint into the form for a ball (since RHS is constant when fixing  $a, b, \theta$ ), 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  $A \in \mathbb{R}^{m \times n}$  and  $C \in \mathbb{R}^{p \times n}$  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.  $\square$

(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  $\lambda_1^2, \dots, \lambda_n^2$  in descending order. Show that the largest and smallest distances from any point on the boundary of the ellipsoid to  $x_c$  are  $\lambda_1$  and  $\lambda_n$  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

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

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  $\lambda_n^2 \leq \dots \leq \lambda_1^2$ , 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.  $\square$

5. Show the following statements.

- (a) In machine learning, we are often given training samples in the form of  $(x_i, y_i)$  for  $i = 1, \dots, n$  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:

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

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

*Proof.* Let  $X \in \mathbb{R}^{n \times d}, y^d$  be such that the  $i$ -th row of  $X$  is  $x_i$  and  $y_i$  respectively. Then

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n (y_i - a^T x_i)^2 &= \frac{1}{n} \|y - Xa\|^2 \\ &= \frac{1}{n} (y^T y - 2a^T A^T y + a^T A^T A a) \\ \nabla_a f &= \frac{1}{n} (2A^T A a - 2A^T y) = \frac{2}{n} (A^T A a - A^T y) \\ \nabla_a^2 f &= \frac{2}{n} A^T A. \end{aligned}$$

Thus we are done if we can show that  $A^T A$  is PSD. Consider  $v \in \mathbb{R}^d$ . We compute

$$v^T A^T A v = (Av)^T Av = \|Av\|^2 \geq 0.$$

Thus  $\nabla_a^2 f \succeq 0$ , and  $f$  is convex. □

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

*Proof.* Let  $S = \sum_{i=1}^n x_i^p$ , and note that  $S > 0$ . We first directly compute the gradient of  $f$ :

$$\frac{f}{x_k} = \frac{1}{p} S^{1/p-1} \cdot p x_k^{p-1} = S^{1/p-1} x_k^{p-1}.$$

Denote  $v = (x_1^{p-1}, \dots, x_n^{p-1})^T$  for convenience, and rewrite  $\nabla f = S^{1/p-1} v$ .

Now compute the hessian.

$$\begin{aligned} \frac{f}{x_j x_k} &= \frac{1}{p} (1-p) S^{1/p-2} \cdot p x_j^{p-1} x_k^{p-1} + (p-1) x_k^{p-2} S^{p-1} \\ &= (1-p) S^{1/p-2} x_j^{p-1} x_k^{p-1} + (p-1) x_k^{p-2} S^{1/p-1}. \end{aligned}$$

Again for notational ease, set  $D = \text{diag}(x_1^{p-2}, \dots, x_n^{p-2})^T$ . Rewrite the hessian as

$$\nabla^2 f = (1-p) S^{1/p-2} v v^T + (p-1) S^{1/p-1} D = (1-p) S^{1/p-2} (v v^T - S D).$$

Notice that the sign is determined by  $v v^T - S D$ , particularly that

$$S D - v v^T \succeq 0 \longrightarrow v v^T - S D \preceq 0 \longrightarrow \nabla^2 f \preceq 0$$

which means that  $f$  is concave. It is sufficient then, to prove that  $SD - vv^T \succeq 0$ . To show this, observe that for  $z \in \mathbb{R}^d$ ,

$$z^T(SD - vv^T)z = S \sum_{i=1}^n x_i^{p-2} z_i^2 - \left( \sum_{i=1}^n x_i^{p-1} z_i \right)^2,$$

which we move on to investigate the squared sum on the LHS.

$$\begin{aligned} \left( \sum_{i=1}^n x_i^{p-1} z_i \right)^2 &= \left( \sum_{i=1}^n x_i^{p/2} x_i^{p/2-1} z_i \right)^2 \\ &\leq \sum_{i=1}^n \left( x_i^{p/2} \right)^2 \sum_{i=1}^n \left( x_i^{p/2-1} z_i \right)^2 \text{ by Cauchy-Schwarz Inequality} \\ &= \left( \sum_{i=1}^n x_i^p \right) \left( \sum_{i=1}^n x_i^{p-2} z_i^2 \right) \\ &= S \sum_{i=1}^n x_i^{p-2} z_i^2. \end{aligned}$$

This tells us that

$$S \sum_{i=1}^n x_i^{p-2} z_i^2 - \left( \sum_{i=1}^n x_i^{p-1} z_i \right)^2 \geq 0,$$

so  $SD - vv^T \succeq 0$ . □

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

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

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

(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.

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

where  $A(x) = A_0 + A_1 x_1 + \dots + A_n x_n \in \mathbb{S}^n$  and  $A_i \in \mathbb{S}^n, i = 1, \dots, n$ . 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.