EE364a Homework 1 solutions

2.1 Let $C \subseteq \mathbf{R}^n$ be a convex set, with $x_1, \ldots, x_k \in C$, and let $\theta_1, \ldots, \theta_k \in \mathbf{R}$ satisfy $\theta_i \geq 0$, $\theta_1 + \cdots + \theta_k = 1$. Show that $\theta_1 x_1 + \cdots + \theta_k x_k \in C$. (The definition of convexity is that this holds for $k = 2$; you must show it for arbitrary k.) *Hint.* Use induction on k.

Solution. This is readily shown by induction from the definition of convex set. We illustrate the idea for $k = 3$, leaving the general case to the reader. Suppose that $x_1, x_2, x_3 \in C$, and $\theta_1 + \theta_2 + \theta_3 = 1$ with $\theta_1, \theta_2, \theta_3 \geq 0$. We will show that $y = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \in C$. At least one of the θ_i is not equal to one; without loss of generality we can assume that $\theta_1 \neq 1$. Then we can write

$$y = \theta_1 x_1 + (1 - \theta_1)(\mu_2 x_2 + \mu_3 x_3)$$

where $\mu_2 = \theta_2/(1 - \theta_1)$ and $\mu_2 = \theta_3/(1 - \theta_1)$. Note that $\mu_2, \mu_3 \geq 0$ and

$$\mu_1 + \mu_2 = \frac{\theta_2 + \theta_3}{1 - \theta_1} = \frac{1 - \theta_1}{1 - \theta_1} = 1.$$

Since C is convex and $x_2, x_3 \in C$, we conclude that $\mu_2 x_2 + \mu_3 x_3 \in C$. Since this point and x_1 are in C, $y \in C$.

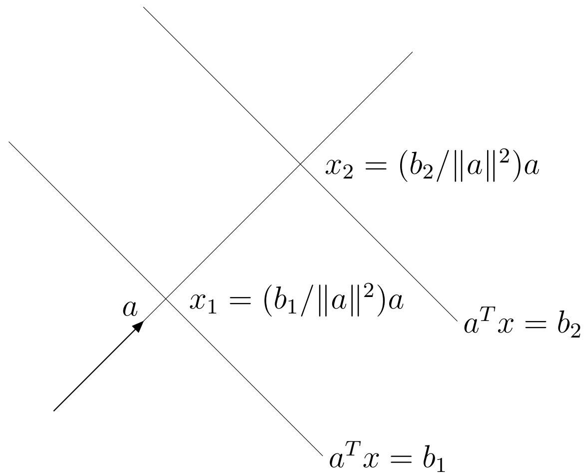
2.2 Show that a set is convex if and only if its intersection with any line is convex. Show that a set is affine if and only if its intersection with any line is affine.

Solution. We prove the first part. The intersection of two convex sets is convex. Therefore if S is a convex set, the intersection of S with a line is convex.

Conversely, suppose the intersection of S with any line is convex. Take any two distinct points x_1 and $x_2 \in S$. The intersection of S with the line through x_1 and x_2 is convex. Therefore convex combinations of x_1 and x_2 belong to the intersection, hence also to S.

2.5 What is the distance between two parallel hyperplanes $\{x \in \mathbf{R}^n \mid a^T x = b_1\}$ and $\{x \in \mathbf{R}^n \mid a^T x = b_2\}$?

Solution. The distance between the two hyperplanes is $|b_1 - b_2|/\|a\|_2$. To see this, consider the construction in the figure below.



The distance between the two hyperplanes is also the distance between the two points x_1 and x_2 where the hyperplane intersects the line through the origin and parallel to the normal vector a. These points are given by

$$x_1 = (b_1/\|a\|_2^2)a, \qquad x_2 = (b_2/\|a\|_2^2)a,$$

and the distance is

$$\|x_1 - x_2\|_2 = |b_1 - b_2|/\|a\|_2.$$

2.7 *Voronoi description of halfspace.* Let a and b be distinct points in $\mathbf{R}^n$. Show that the set of all points that are closer (in Euclidean norm) to a than b, *i.e.*, $\{x \mid \|x-a\|_2 \leq \|x-b\|_2\}$, is a halfspace. Describe it explicitly as an inequality of the form $c^T x \leq d$. Draw a picture.

Solution. Since a norm is always nonnegative, we have $\|x-a\|_2 \leq \|x-b\|_2$ if and only if $\|x-a\|_2^2 \leq \|x-b\|_2^2$, so

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

Therefore, the set is indeed a halfspace. We can take $c = 2(b-a)$ and $d = b^Tb - a^Ta$. This makes good geometric sense: the points that are equidistant to a and b are given by a hyperplane whose normal is in the direction $b - a$.

2.8 Which of the following sets S are polyhedra? If possible, express S in the form $S = \{x \mid Ax \preceq b,\ Fx = g\}$.

(a) $S = \{y_1 a_1 + y_2 a_2 \mid -1 \leq y_1 \leq 1,\ -1 \leq y_2 \leq 1\}$, where $a_1, a_2 \in \mathbf{R}^n$.

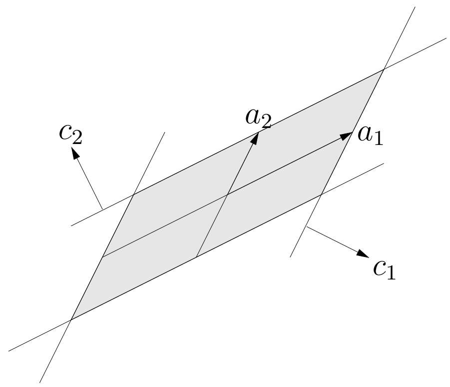
(b) $S = \{x \in \mathbf{R}^n \mid x \succeq 0,\ \mathbf{1}^T x = 1,\ \sum_{i=1}^n x_i a_i = b_1,\ \sum_{i=1}^n x_i a_i^2 = b_2\}$, where $a_1, \ldots, a_n \in \mathbf{R}$ and $b_1, b_2 \in \mathbf{R}$.

(c) $S = \{x \in \mathbf{R}^n \mid x \succeq 0,\ x^T y \leq 1 \text{ for all } y \text{ with } \|y\|_2 = 1\}$.

(d) $S = \{x \in \mathbf{R}^n \mid x \succeq 0,\ x^T y \leq 1 \text{ for all } y \text{ with } \sum_{i=1}^n |y_i| = 1\}$.

Solution.

(a) S is a polyhedron. It is the parallelogram with corners $a_1 + a_2$, $a_1 - a_2$, $-a_1 + a_2$, $-a_1 - a_2$, as shown below for an example in $\mathbf{R}^2$.



For simplicity we assume that a_1 and a_2 are independent. We can express S as the intersection of three sets:

- S_1: the plane defined by a_1 and a_2
- $S_2 = \{z + y_1 a_1 + y_2 a_2 \mid a_1^T z = a_2^T z = 0, -1 \leq y_1 \leq 1\}$. This is a slab parallel to a_2 and orthogonal to S_1
- $S_3 = \{z + y_1 a_1 + y_2 a_2 \mid a_1^T z = a_2^T z = 0, -1 \leq y_2 \leq 1\}$. This is a slab parallel to a_1 and orthogonal to S_1

Each of these sets can be described with linear inequalities.

- S_1 can be described as

$$v_k^T x = 0, \quad k = 1, \ldots, n-2$$

  where v_k are $n-2$ independent vectors that are orthogonal to a_1 and a_2 (which form a basis for the nullspace of the matrix $[a_1\ a_2]^T$).

- Let c_1 be a vector in the plane defined by a_1 and a_2, and orthogonal to a_2. For example, we can take

$$c_1 = a_1 - \frac{a_1^T a_2}{\|a_2\|_2^2} a_2.$$

  Then $x \in S_2$ if and only if

$$-|c_1^T a_1| \leq c_1^T x \leq |c_1^T a_1|.$$

- Similarly, let c_2 be a vector in the plane defined by a_1 and a_2, and orthogonal to a_1, *e.g.*,

$$c_2 = a_2 - \frac{a_2^T a_1}{\|a_1\|_2^2} a_1.$$

Then $x \in S_3$ if and only if

$$-|c_2^T a_2| \leq c_2^T x \leq |c_2^T a_2|.$$

Putting it all together, we can describe S as the solution set of $2n$ linear inequalities

$$\begin{array}{rcl} v_k^T x & \leq & 0, \quad k = 1, \ldots, n-2 \\ -v_k^T x & \leq & 0, \quad k = 1, \ldots, n-2 \\ c_1^T x & \leq & |c_1^T a_1| \\ -c_1^T x & \leq & |c_1^T a_1| \\ c_2^T x & \leq & |c_2^T a_2| \\ -c_2^T x & \leq & |c_2^T a_2|. \end{array}$$

(b) S is a polyhedron, defined by linear inequalities $x_k \geq 0$ and three equality constraints.

(c) S is not a polyhedron. It is the intersection of the unit ball $\{x \mid \|x\|_2 \leq 1\}$ and the nonnegative orthant $\mathbf{R}_+^n$. This follows from the following fact, which follows from the Cauchy-Schwarz inequality:

$$x^T y \leq 1 \text{ for all } y \text{ with } \|y\|_2 = 1 \iff \|x\|_2 \leq 1.$$

Although in this example we define S as an intersection of halfspaces, it is not a polyhedron, because the definition requires infinitely many halfspaces.

(d) S is a polyhedron. S is the intersection of the set $\{x \mid |x_k| \leq 1, \quad k = 1, \ldots, n\}$ and the nonnegative orthant $\mathbf{R}_+^n$. This follows from the following fact:

$$x^T y \leq 1 \text{ for all } y \text{ with } \sum_{i=1}^n |y_i| = 1 \iff |x_i| \leq 1, \quad i = 1, \ldots, n.$$

We can prove this as follows. First suppose that $|x_i| \leq 1$ for all i. Then

$$x^T y = \sum_i x_i y_i \leq \sum_i |x_i||y_i| \leq \sum_i |y_i| = 1$$

if $\sum_i |y_i| = 1$.

Conversely, suppose that x is a nonzero vector that satisfies $x^T y \leq 1$ for all y with $\sum_i |y_i| = 1$. In particular we can make the following choice for y: let k be an index for which $|x_k| = \max_i |x_i|$, and take $y_k = 1$ if $x_k > 0$, $y_k = -1$ if $x_k < 0$, and $y_i = 0$ for $i \neq k$. With this choice of y we have

$$x^T y = \sum_i x_i y_i = y_k x_k = |x_k| = \max_i |x_i|.$$

Therefore we must have $\max_i |x_i| \leq 1$.

All this implies that we can describe S by a finite number of linear inequalities: it is the intersection of the nonnegative orthant with the set $\{x \mid -\mathbf{1} \preceq x \preceq \mathbf{1}\}$, *i.e.*, the solution of $2n$ linear inequalities

$$\begin{array}{rcl} -x_i & \leq & 0, \quad i = 1, \ldots, n \\ x_i & \leq & 1, \quad i = 1, \ldots, n. \end{array}$$

Note that as in part (c) the set S was given as an intersection of an infinite number of halfspaces. The difference is that here most of the linear inequalities are redundant, and only a finite number are needed to characterize S.

None of these sets are affine sets or subspaces, except in some trivial cases. For example, the set defined in part (a) is a subspace (hence an affine set), if $a_1 = a_2 = 0$; the set defined in part (b) is an affine set if $n = 1$ and $S = \{1\}$; etc.

2.11 *Hyperbolic sets.* Show that the *hyperbolic set* $\{x \in \mathbf{R}^2_+ \mid x_1 x_2 \geq 1\}$ is convex. As a generalization, show that $\{x \in \mathbf{R}^n_+ \mid \prod_{i=1}^n x_i \geq 1\}$ is convex. *Hint.* If $a, b \geq 0$ and $0 \leq \theta \leq 1$, then $a^\theta b^{1-\theta} \leq \theta a + (1-\theta) b$; see §3.1.9.

Solution.

(a) We prove the first part without using the hint. Consider a convex combination z of two points (x_1, x_2) and (y_1, y_2) in the set. If $x \succeq y$, then $z = \theta x + (1-\theta) y \succeq y$ and obviously $z_1 z_2 \geq y_1 y_2 \geq 1$. Similar proof if $y \succeq x$.

Suppose $y \not\succeq x$ and $x \not\succeq y$, *i.e.*, $(y_1 - x_1)(y_2 - x_2) < 0$. Then

$$\begin{array}{l} (\theta x_1 + (1-\theta) y_1)(\theta x_2 + (1-\theta) y_2) \\ \quad = \quad \theta^2 x_1 x_2 + (1-\theta)^2 y_1 y_2 + \theta(1-\theta) x_1 y_2 + \theta(1-\theta) x_2 y_1 \\ \quad = \quad \theta x_1 x_2 + (1-\theta) y_1 y_2 - \theta(1-\theta)(y_1 - x_1)(y_2 - x_2) \\ \quad \geq \quad 1. \end{array}$$

(b) Assume that $\prod_i x_i \geq 1$ and $\prod_i y_i \geq 1$. Using the inequality in the hint, we have

$$\prod_i (\theta x_i + (1-\theta) y_i) \geq \prod x_i^\theta y_i^{1-\theta} = (\prod_i x_i)^\theta (\prod_i y_i)^{1-\theta} \geq 1.$$

2.12 Which of the following sets are convex?

(a) A *slab*, *i.e.*, a set of the form $\{x \in \mathbf{R}^n \mid \alpha \leq a^T x \leq \beta\}$.

(b) A *rectangle*, *i.e.*, a set of the form $\{x \in \mathbf{R}^n \mid \alpha_i \leq x_i \leq \beta_i, \ i = 1, \ldots, n\}$. A rectangle is sometimes called a *hyperrectangle* when $n > 2$.

(c) A *wedge*, *i.e.*, $\{x \in \mathbf{R}^n \mid a_1^T x \leq b_1, \ a_2^T x \leq b_2\}$.

(d) The set of points closer to a given point than a given set, *i.e.*,

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

where $S \subseteq \mathbf{R}^n$.

(e) The set of points closer to one set than another, *i.e.*,

$$\{x \mid \mathbf{dist}(x, S) \le \mathbf{dist}(x, T)\},$$

where $S, T \subseteq \mathbf{R}^n$, and

$$\mathbf{dist}(x, S) = \inf\{\|x - z\|_2 \mid z \in S\}.$$

(f) The set $\{x \mid x + S_2 \subseteq S_1\}$, where $S_1, S_2 \subseteq \mathbf{R}^n$ with S_1 convex.

(g) The set of points whose distance to a does not exceed a fixed fraction θ of the distance to b, *i.e.*, the set $\{x \mid \|x - a\|_2 \le \theta \|x - b\|_2\}$. You can assume $a \ne b$ and $0 \le \theta \le 1$.

Solution.

(a) A slab is an intersection of two halfspaces, hence it is a convex set and a polyhedron.

(b) As in part (a), a rectangle is a convex set and a polyhedron because it is a finite intersection of halfspaces.

(c) A wedge is an intersection of two halfspaces, so it is convex and a polyhedron. It is a cone if $b_1 = 0$ and $b_2 = 0$.

(d) This set is convex because it can be expressed as

$$\bigcap_{y \in S} \{x \mid \|x - x_0\|_2 \le \|x - y\|_2\},$$

i.e., an intersection of halfspaces. (Recall from exercise 2.7 that, for fixed y, the set

$$\{x \mid \|x - x_0\|_2 \le \|x - y\|_2\}$$

is a halfspace.)

(e) In general this set is not convex, as the following example in $\mathbf{R}$ shows. With $S = \{-1, 1\}$ and $T = \{0\}$, we have

$$\{x \mid \mathbf{dist}(x, S) \le \mathbf{dist}(x, T)\} = \{x \in \mathbf{R} \mid x \le -1/2 \text{ or } x \ge 1/2\}$$

which clearly is not convex.

(f) This set is convex. $x + S_2 \subseteq S_1$ if $x + y \in S_1$ for all $y \in S_2$. Therefore

$$\{x \mid x + S_2 \subseteq S_1\} = \bigcap_{y \in S_2} \{x \mid x + y \in S_1\} = \bigcap_{y \in S_2} (S_1 - y),$$

the intersection of convex sets $S_1 - y$.

(g) The set is convex, in fact a ball.

$$\begin{array}{rl} & \{x \mid \|x-a\|_2 \leq \theta \|x-b\|_2\} \\ = & \{x \mid \|x-a\|_2^2 \leq \theta^2 \|x-b\|_2^2\} \\ = & \{x \mid (1-\theta^2)x^Tx - 2(a-\theta^2 b)^T x + (a^Ta - \theta^2 b^T b) \leq 0\} \end{array}$$

If $\theta = 1$, this is a halfspace. If $\theta < 1$, it is a ball

$$\{x \mid (x-x_0)^T(x-x_0) \leq R^2\},$$

with center x_0 and radius R given by

$$x_0 = \frac{a - \theta^2 b}{1-\theta^2}, \qquad R = \left(\frac{\theta^2 \|b\|_2^2 - \|a\|_2^2}{1-\theta^2} + \|x_0\|_2^2 \right)^{1/2}.$$

2.15 *Some sets of probability distributions.* Let x be a real-valued random variable with $\mathbf{prob}(x = a_i) = p_i$, $i = 1, \ldots, n$, where $a_1 < a_2 < \cdots < a_n$. Of course $p \in \mathbf{R}^n$ lies in the standard probability simplex $P = \{p \mid \mathbf{1}^T p = 1,\ p \succeq 0\}$. Which of the following conditions are convex in p? (That is, for which of the following conditions is the set of $p \in P$ that satisfy the condition convex?)

(a) $\alpha \leq \mathbf{E} f(x) \leq \beta$, where $\mathbf{E} f(x)$ is the expected value of $f(x)$, *i.e.*, $\mathbf{E} f(x) = \sum_{i=1}^n p_i f(a_i)$. (The function $f : \mathbf{R} \to \mathbf{R}$ is given.)

(b) $\mathbf{prob}(x > \alpha) \leq \beta$.

(c) $\mathbf{E}\,|x^3| \leq \alpha\,\mathbf{E}\,|x|$.

(d) $\mathbf{E}\,x^2 \leq \alpha$.

(e) $\mathbf{E}\,x^2 \geq \alpha$.

(f) $\mathbf{var}(x) \leq \alpha$, where $\mathbf{var}(x) = \mathbf{E}(x - \mathbf{E}\,x)^2$ is the variance of x.

(g) $\mathbf{var}(x) \geq \alpha$.

(h) $\mathbf{quartile}(x) \geq \alpha$, where $\mathbf{quartile}(x) = \inf\{\beta \mid \mathbf{prob}(x \leq \beta) \geq 0.25\}$.

(i) $\mathbf{quartile}(x) \leq \alpha$.

Solution. We first note that the constraints $p_i \geq 0$, $i = 1, \ldots, n$, define halfspaces, and $\sum_{i=1}^n p_i = 1$ defines a hyperplane, so P is a polyhedron.

The first five constraints are, in fact, linear inequalities in the probabilities p_i.

(a) $\mathbf{E}\, f(x) = \sum_{i=1}^n p_i f(a_i)$, so the constraint is equivalent to two linear inequalities

$$\alpha \leq \sum_{i=1}^n p_i f(a_i) \leq \beta.$$

(b) $\mathbf{prob}(x \geq \alpha) = \sum_{i:\, a_i \geq \alpha} p_i$, so the constraint is equivalent to a linear inequality

$$\sum_{i:\, a_i \geq \alpha} p_i \leq \beta.$$

(c) The constraint is equivalent to a linear inequality

$$\sum_{i=1}^n p_i(|a_i^3| - \alpha |a_i|) \leq 0.$$

(d) The constraint is equivalent to a linear inequality

$$\sum_{i=1}^n p_i a_i^2 \leq \alpha.$$

(e) The constraint is equivalent to a linear inequality

$$\sum_{i=1}^n p_i a_i^2 \geq \alpha.$$

The first five constraints therefore define convex sets.

(f) The constraint

$$\mathbf{var}(x) = \mathbf{E}\, x^2 - (\mathbf{E}\, x)^2 = \sum_{i=1}^n p_i a_i^2 - (\sum_{i=1}^n p_i a_i)^2 \leq \alpha$$

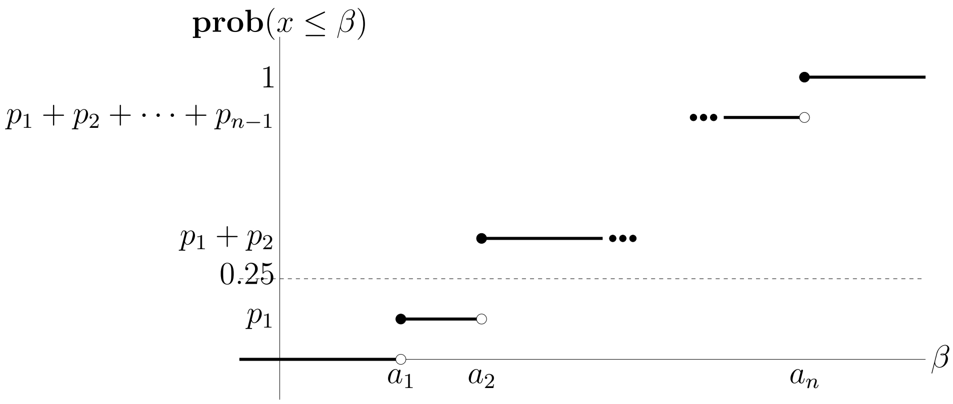
is not convex in general. As a counterexample, we can take $n = 2$, $a_1 = 0$, $a_2 = 1$, and $\alpha = 1/5$. $p = (1, 0)$ and $p = (0, 1)$ are two points that satisfy $\mathbf{var}(x) \leq \alpha$, but the convex combination $p = (1/2, 1/2)$ does not.

(g) This constraint is equivalent to

$$\sum_{i=1}^n a_i^2 p_i + (\sum_{i=1}^n a_i p_i)^2 = b^T p + p^T A p \leq \alpha$$

where $b_i = a_i^2$ and $A = aa^T$. This defines a convex set, since the matrix aa^T is positive semidefinite.

Let us denote $\mathbf{quartile}(x) = f(p)$ to emphasize it is a function of p. The figure illustrates the definition. It shows the cumulative distribution for a distribution p with $f(p) = a_2$.



(h) The constraint $f(p) \ge \alpha$ is equivalent to

$$\mathbf{prob}(x \le \beta) < 0.25 \text{ for all } \beta < \alpha.$$

If $\alpha \le a_1$, this is always true. Otherwise, define $k = \max\{i \mid a_i < \alpha\}$. This is a fixed integer, independent of p. The constraint $f(p) \ge \alpha$ holds if and only if

$$\mathbf{prob}(x \le a_k) = \sum_{i=1}^{k} p_i < 0.25.$$

This is a strict linear inequality in p, which defines an open halfspace.

(i) The constraint $f(p) \le \alpha$ is equivalent to

$$\mathbf{prob}(x \le \beta) \ge 0.25 \text{ for all } \beta \ge \alpha.$$

Here, let us define $k = \max\{i \mid a_i \le \alpha\}$. Again, this is a fixed integer, independent of p. The constraint $f(p) \le \alpha$ holds if and only if

$$\mathbf{prob}(x \le a_k) = \sum_{i=1}^{k} p_i \ge 0.25.$$

If $\alpha \le a_1$, then no p satisfies $f(p) \le \alpha$, which means that the set is empty.