Homework 1

MATH CS 120 Convex Optimization

Harry Coleman

April 21, 2020

Exercise 2.3

A set C is midpoint convex if whenever two points a, b are in C, the average or midpoint $(a+b)/2$ is in C. Obviously a convex set is midpoint convex. It can be proved that under mild conditions midpoint convexity implies convexity. As a simple case, prove that if C is closed and midpoint convex, then C is convex.

Let $C \in \mathbb{R}^n$ be a midpoint convex, closed set and let $a, b \in \mathbb{C}$. Let $c = \theta a + (1-\theta)b$ for some $\theta \in [0,1]$, that is, c is a convex combination of a and b. We aim to prove that $c \in C$, thereby proving C is convex. To do so, we construct a sequence of line segments betwen a_n and b_n for all $n \in \mathbb{N}$ in the following way:

$$a_0 = a, \quad b_0 = b,$$

and given $a_{n-1}, b_n - 1$, define $m_n = \dfrac{1}{2}(a_{n-1} + b_{n-1})$ and if c is on the line segment from a_n to m_n, define

$$a_n = a_{n-1}, \quad b_n = m_n.$$

Otherwise, if c is on the line segment from m_n to b_n, define

$$a_n = m_n, \quad b_n = b_{n-1}.$$

Since c is on the line segment between a_n and b_n for all $n \in \mathbb{N}$, we have that

$$|a_n - c| \leq |a_n - b_n| = \frac{1}{2^n}|a - b|.$$

So the sequence $\{a_n\}_{n=1}^{\infty}$ converges to c. And since each a_n is define by a midpoint of elements of C, each $a_n \in \mathbb{C}$. Thus c is a limit point of C, and since C is closed, $c \in \mathbb{C}$. Therefore, C is convex.

Exercise 2.8

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

(a) $S = \{y_1 a_1 + y_2 a_2 : -1 \leq y_1 \leq 1, -1 \leq y_2 \leq 1\}$, where $a_1, a_2 \in \mathbb{R}^n$.
Any point $x \in S$ corresponds to a convex combination

$$x = \theta_1(a_1 + a_2) + \theta_2(a_1 - a_2) + \theta_3(a_2 - a_1) + \theta_4(-a_1 - a_2),$$

where

$$y_1 = \theta_1 + \theta_2 - \theta_3 + \theta_4,$$
$$y_2 = \theta_1 + \theta_3 - \theta_2 + \theta_4.$$

So S can be equivalently defined as the simplex given by the convex hull of the points

$$(a_1 + a_2), (a_1 - a_2), (a_2 - a_1), (-a_1 - a_2).$$

If a_1, a_2 are affinely dependent, then this is simply a line segment. Either way, this is a polyhedron.

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

To see that S is a polyhedron, the conditions for S can simply be rewritten. If we define

$$A = \begin{bmatrix} 1 & \cdots & 1 \\ a_1 & \cdots & a_n \\ a_1^2 & \cdots & a_n^2 \end{bmatrix}, \qquad b = \begin{bmatrix} 1 \\ b_1 \\ b_2 \end{bmatrix},$$

then we can equivalently write S as

$$S = \{x \in \mathbb{R}^n : (-I_n)x \preceq 0, Ax = b\},$$

thus S is a polyhedron.

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

This set is equivalent to the nonnegative vectors of the closed unit ball centered at the origin in R^n. This can be seen since any negative x would not satisfy $x \succeq 0$ and any vector with a length greater than 1 would not satisfy $x^T y \leq 1$ for the unit vector y in the direction of x. However, any nonnegative unit vector satisfies both of these. Because of this, S cannot be expressed as the intersection of finitely many halfspaces and hyperplanes, and thus is not a polyhedron.

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

This set is equivalent to the set of vectors in $\mathbb{R}^n$ with each element between 0 and 1. Similar to (c), each x must be nonnegative, but in this case each element of x must be less than or equal to 1. For any such vector, the dot product $x^T y$ is less than or equal to the sum of all $|y_i|$, which is equal to 1. However, for any x with an element greater than 1, there is some y for which the dot product is greater than 1. Therefore, we can equivalently define

$$S = \{x \in \mathbb{R}^n : \begin{bmatrix} -I_n \\ I_n \end{bmatrix} x \preceq \begin{bmatrix} 0 \\ 1 \end{bmatrix}\}.$$

Exercise 2.10

Let $C \in \mathbb{R}^n$ be the solution set of a quadratic inequality,

$$C = \{x \in \mathbb{R}^n : x^T Ax + b^T x + c \leq 0\},$$

with $A \in S^n, b \in \mathbb{R}^n$, and $c \in \mathbb{R}$. (note: S^n is the set of $n \times n$ symmetric matrices.)

(a) Show that C is convex if $A \succeq 0$. Let $x, y \in C$ and $\theta \in [0, 1]$. We aim to prove that the convex combination

$$x + \theta(y - x)$$

is in C. To show this, we substitute the point into the inequality for C to find

$$\begin{aligned}&(x + \theta(y - x))^T A(x + \theta(y - x)) + b^T(x + \theta(y - x)) + c \\ &= (y - x)^T A(y - x)\theta^2 + (2(y - x)^T Ax + b^T(y - x))\theta + x^T Ax + b^T x + c.\end{aligned}$$

Notice that this is is a quadratic function on θ, which is less than or equal to zero for values of $\theta = 0$ and $\theta = 1$, since these values correspond to the inequality for C on x and y, respectively. Therefore, this expression is less than or equal to zero if and only if the a_2 term is positive. Since

$$a_2 = (y - x)^T A(y - x)$$

and A is positive semidefinite, we in fact have $a_2 \geq 0$. Thus, the expression is less than or equal to 0, giving us $x + \theta(y - x) \in C$. The converse, however, is not true, since we could pick $A = -I_n, b = 0, c = 0$, and then the condition for C becomes

$$-x^T x \leq 0,$$

which is true for all $x \in \mathbb{R}^n$, making $C = R^n$, which is convex.

(b) Show that the intersecton of C and the hyperplane defined by $g^T x + h = 0$ (where $g \neq 0$) is convex if $A + \lambda gg^T \succeq 0$ for some $\lambda \in \mathbb{R}$.

Similar to (a), we let $x, y \in C \cap \{x \in \mathbb{R}^n : g^T + h = 0\}$, $\theta \in \mathbb{R}$ and we aim to prove

$$x + \theta(y - x) \in C \cap \{x \in \mathbb{R}^n : g^T x + h = 0\}.$$

Since the hyperplane is convex, we already know that

$$x + \theta(y - x) \in \{x \in \mathbb{R}^n : g^T x + h = 0\}.$$

So we just need to show that the point is in C. And similar to (a), this is the case if

$$(y - x)^T A(y - x) \geq 0.$$

To show this is satisfied, we assume $A + \lambda gg^T \succeq 0$ for some $\lambda \in \mathbb{R}$. So

$$\begin{aligned}0 &\leq (y - x)^T (A + \lambda gg^T)(y - x) \\ &= (y - x)^T A(y - x) + \lambda(g^T y - g^T x)^2 \\ &= (y - x)^T A(y - x) + \lambda(-h + h)^2 \\ &= (y - x)^T A(y - x).\end{aligned}$$

Thus, $x + \theta(y - x) \in C \cap \{x \in \mathbb{R}^n : g^T x + h = 0\}$, so the intersection is convex.

Exercise 2.20

Suppose $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, with $b \in R(A)$. Show that there exists an x satisfying

$$x \succ 0, \quad Ax = b$$

if and only if there exists no λ with

$$A^T\lambda \succeq 0, \quad A^T\lambda \neq 0, \quad b^T\lambda \leq 0.$$

Hint. First prove the following fact from linear algebra: $c^Tx = d$ for all x satisfying $Ax = b$ if and only if there is a vector λ such that $c = A^T\lambda, d = b^T\lambda$

We first prove that $c^Tx = d$ for all x satisfying $Ax = b$ if and only if there is a vector λ such that $c = A^T\lambda, d = b^T\lambda$. Suppose there is a vector λ such that $c = A^T\lambda, d = b^T\lambda$. Then for any x such that $Ax = b$, we find

$$c^Tx = A^T\lambda x = \lambda^T Ax = \lambda^T b = b^T\lambda = d.$$

Now suppose that $c^Tx = d$ for all x such that $Ax = b$. Let v be such that $Av = b$, then for any x such that $Ax = b$, there is some $u \in \ker A$ such that $x = u + v$. Now since $c^Tv = d$ and $c^Tx = d$, we find

$$d = c^t(u + v) = c^Tu + c^Tv,$$

which implies that $c^Tu = 0$ for all $u \in \ker A$. Now since c is orthogonal to all the vectors in the kernel of A, it must be in the orthogonal complement to $\ker A$ which is given by the rowspace of A. So c is a linear combination of the rows of A, that is, $c = A^T\lambda$ for some λ. From this, we find

$$d = c^Tx = (A^T\lambda)^Tx = \lambda^T Ax = \lambda^T b = b^T\lambda.$$

Thus $c^Tx = d$ for all x satisfying $Ax = b$ if and only if there is a vector λ such that $c = A^T\lambda, d = b^T\lambda$.

Suppose there exists $x \succ 0$ such that $Ax = b$. We now assume for the purpose of contradiction that there is some λ such that

$$A^T\lambda \succeq 0, \quad A^T\lambda \neq 0, \quad b^T\lambda \leq 0.$$

Then we have

$$A^T\lambda x = b^T\lambda.$$

However, since x is strictly positive, $A^T\lambda$ is nonnegative, and $b^T\lambda$ is strictly positive. Therefore, no such λ exists.

Lastly suppose that there does not exist an $x \succ 0$ such that $Ax = b$. In other words, the set $G = \{x \in \mathbb{R}^n : Ax = b\}$ is disjoint with the set $H = \{x \in \mathbb{R}^n : x \succ 0\}$. Since both of these sets are convex, there is a hyperplane defined by $c \in \mathbb{R}^n, k \in \mathbb{R}$ such that

$$\begin{cases} c^Tg \leq k, & \forall g \in G; \\ c^Th \geq k, & \forall h \in H. \end{cases}$$

From this, we see that c must be nonnegative, since otherwise we could find $c^T h < k$ for large enough $h \in H$. And since $c^T h$ can be made arbitrarily close to zero for small h, we must have $k \leq 0$. And since G is affine, $c^T g$ must be the same for all $g \in G$, since otherwise we could picking two points with different values find a third along the line with an arbitrarily large value for $c^T g$. Call this value d, so $c^T x = d$ for all $x \in G$. Then the desired λ is given by $c = A^T \lambda$ and $d = b^T \lambda$.

Exercise 2.21

Suppose that C and D are disjoint subsets of $\mathbb{R}^n$. Consider the set of $(a, b) \in \mathbb{R}^{n+1}$ for which $a^T x \leq b$ for all $x \in C$, and $a^T x \geq b$ for all $x \in D$. Show that this set is a convex cone (which is the singleton $\{0\}$ if there is no hyperplane that separates C and D).

Let $(a, b), (c, d) \in \mathbb{R}^{n+1}$ define separating hyperplanes for C and D. We consider for some $\theta_1, \theta_2 \geq 0$ the conic combination

$$(e, f) = (\theta_1 a + \theta_2 c,\ \theta_1 b + \theta_2 d).$$

Then for any $x \in C$, we have

$$e^T x = (\theta_1 a + \theta_2 b)^T x = \theta_1 a^T x + \theta_2 c^T x \leq \theta_1 b + \theta_2 d = f.$$

And for any $x \in D$, we have

$$e^T x = (\theta_1 a + \theta_2 b)^T x = \theta_1 a^T x + \theta_2 c^T x \geq \theta_1 b + \theta_2 d = f.$$

So in fact, (e, f) defines a separating hyperplane of C and D. Thus the set of separating hyperplanes is a convex cone.

Exercise 2.25

Let $C \in \mathbb{R}^n$ be a closed convex set, and suppose that $x_1, \ldots, x_K$ are on the boundary of C. Suppose that for each i, $a_i^T(x - x_i) = 0$ defines a supporting hyperplane for C at X_i, i.e., $C \subseteq \{x : a_i^T(x - x_i) \leq 0\}$. Consider the two polyhedra

$$P_{\text{inner}} = \mathbf{conv}\{x_1, \ldots, x_K\}, \quad P_{\text{outer}}(x : a_i^T(x - x_i) \leq 0, i = 1, \ldots, K\}.$$

Show that $P_{\text{inner}} \subseteq C \subseteq P_{\text{outer}}$. Draw a picture illustrating this.

Let $x \in P_{\text{inner}}$. Since C is closed, the points $x_1, \ldots, x_K \in C$. And since x is a convex combination of points in C, which is a convex set, we have $x \in C$. Thus, $P_{\text{inner}} \subseteq C$.

Now let $x \in C$. Then for each of the defined supporting hyperplanes of C, we have $a_i^T(x - x_i) \leq 0$, which is precisely the condition for P_{outer}, so $x \in P_{\text{outer}}$. Thus $C \subseteq P_{\text{outer}}$.