Homework 2

MATH CS 120 Convex Optimization

Harry Coleman

June 2, 2020

4.2

Consider the optimization problem

$$\text{minimize } f_0(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x)$$

with domain $\mathbf{dom} f_0 = \{x : Ax < b\}$, where $A \in \mathbb{R}^{m \times n}$ (with rows a_i^T). We assume that $\mathbf{dom} f_0$ is nonempty. Prove the following facts

4.2.a

$\mathbf{dom} f_0$ is unbounded if and only if there exists a $v \neq 0$ with $Av \leq 0$.

Proof. Suppose $\mathbf{dom} f_0$ is unbounded. Then for any $n \in \mathbb{N}$, there exists some $x_n \in \mathbf{dom} f_0$ such that $||x_n|| > n$. Let $\{x_n\}$ be a sequence in $\mathbf{dom} f_0$ such that $||x_n|| > n$ for all $n \in \mathbb{N}$. Now consider the sequence $\{y_n\}$ given by

$$y_n = \frac{x_n}{||x_n||}$$

for each $n \in \mathbb{N}$. This is a sequence in the set of unit vectors $\{x \in \mathbb{R}^n : ||x|| = 1\}$, which is bounded. So by the Bolzano-Weierstrass theorem, there is a convergent subsequence $\{y_{n_k}\}$ with $y_{n_k} \to y$. We also take the corresponding subsequence $\{x_{n_k}\}$ of $\{x_n\}$, noting that $||x_{n_k}|| \to \infty$. We now consider for some row a_i^T of A,

$$a_i^T y_{n_k} = \frac{a_i^T x_{n_k}}{||x_{n_k}||} < b_i \frac{1}{||x_{n_k}||}.$$

Now letting $k \to \infty$, we find

$$a_i^T y \leq b_i \cdot 0 = 0.$$

Since this is true for each a_i^T row of A, we in fact have $Ay \leq 0$. Also note that since $||y_n|| = 1$ for all $n \in \mathbb{N}$, we have $||y|| = 1$, so $y \neq 0$.

Now suppose there exists some $v \neq 0$ with $Av \leq 0$. Since $\textbf{dom} f_0$ is nonempty, we also pick a point $x \in \textbf{dom} f_0$. Now for any $t \geq 0$, we have

$$A(x + tv) = Ax + tAv \leq Ax + 0 < b,$$

so $x + tv \in \textbf{dom} f_0$. And for any $M \in \mathbb{R}_+$, we can pick

$$t = \frac{\|x\| + M}{\|v\|},$$

then by the reverse triangle inequality,

$$\|x + tv\| \geq |\|x\| - t\|v\|| = \left| \|x\| - \frac{\|x\| + M}{\|v\|} \|v\| \right| = M.$$

So $\textbf{dom} f_0$ is unbounded.

$\square$

4.2.b

f_0 is unbounded below if and only if there exists a v with $Av \leq 0$, $Av \neq 0$. *Hint.* There exists v such that $Av \leq 0$, $Av \neq 0$ if and only if there exists no $z > 0$ such that $A^T z = 0$.

Proof. Suppose there exists a v with $Av \leq 0$ and $Av \neq 0$. Let $M_1 \in \mathbb{R}$ be given and let $x \in \textbf{dom} f_0$. Now for any $t \geq 0$, we have

$$A(x + tv) = Ax + tAv \leq Ax + 0 < b,$$

so $x + tv \in \textbf{dom} f_0$. We want to pick t such that

$$f_0(x + tv) = -\sum_{i=1}^{m} \log(b_i - a_i^T(x + tv)) \leq M_1.$$

This will be true if and only if

$$\sum_{i=1}^{m} \log(b_i - a_i^T(x + tv)) \geq -M_1.$$

Notice that since for each i, we have $a_i^T(x + tv) < b$, we also have $b_i - a_i^T(x + tv) > 0$, which implies $\log(b_i - a_i^T(x + tv)) > 0$. We now choose j such that $a_j^T v < 0$ since $Av \neq 0$. Then since each term is positive,

$$\sum_{i=1}^{m} \log(b_i - a_i^T(x + tv)) \geq \log(b_j - a_j^T(x + tv)).$$

Since the log function is increasing and unbounded above, we can pick some M_2 such that

$$y \geq M_2 \implies \log(y) \geq -M_1.$$

Now since

$$b_j - a_j^T(x - vt) = b_j - a_j^T - (a_j^T v)t$$

is linear with respect to t and has slope $-a_j^T v > 0$, we can pick t large enough such that

$$b_j - a_j^T(x + vt) \geq M_2.$$

This now implies that

$$\sum_{i=1}^{m} \log(b_i - a_i^T(x + tv)) \geq \log(b_j - a_j^T(x + tv)) \geq -M_1.$$

which gives us

$$f_0(x + vt) \leq M_1,$$

so f_0 is unbounded below.

Suppose f_0 is unbounded below. Then let $\{x_n\}$ be a sequence in $\mathbf{dom} f_0$ such that $f_0(x_n) \to -\infty$. This implies that

$$\sum_{i=1}^{m} \log(b_i - a_i^T x_n) \to +\infty,$$

$$\sum_{i=1}^{m} (b_i - a_i^T x_n) \to +\infty.$$

Assume for contradiction that there exists some $z \in \mathbb{R}^m$, such that $z > 0$ and $A^T z = 0$. In particular, we choose z such that $z_i \geq 1$ for all i; this is possible since any positive scalar multiple of z has the same properties. Then since

$$(b_i - a_i^T x_n) > 0 \quad \text{and} \quad z_i \geq 1,$$

we have

$$\begin{aligned}
\sum_{i=1}^{m} (b_i - a_i^T x_n) &\leq \sum_{i=1}^{m} z_i(b_i - a_i^T x_n) \\
&= \sum_{i=1}^{m} z_i b_i - \sum_{i=1}^{m} z_i(a_i^T x_n) \\
&= z^T b - z^T A x_n \\
&= z^T b - (A^T z)^T x_n \\
&= z^T b - 0 x_n \\
&= z^T b.
\end{aligned}$$

This implies that $z^T b \to +\infty$ as $n \to +\infty$, which is as contradiction as $z^T b$ is constant with respect to n. Therefore, no such z exists. Then from the hint, this implies that there exists a v such that $Av \leq 0$, $Av \neq 0$.

□

4.2.c

If f_0 is bounded below then its minimum is attained, i.e., then there exists an x that satisfies the optimality condition (4.23).

Proof. Suppose f_0 is bounded below, then by the contrapositive of 4.2.b and the hint, there exists some $z > 0$ such that $A^T z = 0$. That is,

$$0 = A^T z = \sum_{i=1}^{m} z_i a_i = \sum_{i=1}^{m} \frac{1}{b_i - a_i^T w)} a_i = \nabla f_0(w)$$

if w is the vector such that

$$Aw = b - \begin{bmatrix} 1/z_1 \\ \vdots \\ 1/z_m \end{bmatrix}.$$

This vector w can be shown to exist if $\operatorname{rank} A = n$. In which case, $z > 0$ implies $Aw < b$ so $w \in \mathbf{dom} f_0$. Then we would have that w satisfies the optimality condition.

I was unable to prove $\operatorname{rank} A = n$.

□

4.2.d

The optimal set is affine: $X_{\text{opt}} = \{x^* + v : Av = 0\}$, where x^* is any optimal point.

Proof. Let x^* be an optimal point and let $v \in \ker A$. Then

$$\begin{aligned}
\nabla f_0(x^* + v) &= \sum_{i=1}^{m} \frac{1}{b_i - a_i^T(x^* + v)} a_i \\
&= \sum_{i=1}^{m} \frac{1}{b_i - a_i^T x^* + a_i^T v} a_i \\
&= \sum_{i=1}^{m} \frac{1}{b_i - a_i^T x^* + 0} a_i \\
&= \nabla f_0(x^*) \\
&= 0.
\end{aligned}$$

So $x^* + v$ is optimal. This implies that $\{x^* + v : Av = 0\} \subseteq X_{\text{opt}}$. Now suppose $y^* \in X_{\text{opt}}$. Then $y^* = x^* + z$ for some vector z. We aim to prove $z \in \ker A$, as this will imply that $y^* \in \{x^* + v : Av = 0\}$ and therefore $X_{\text{opt}} = \{x^* + v : Av = 0\}$.

I was unable to complete this.

□

4.3

Prove that $x^* = (1, 1/2, -1)$ is optimal for the optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) = \frac{1}{2}x^T P x + q^T x + r \\ \text{subject to} & -1 \leq x_i \leq 1, \quad i = 1, 2, 3, \end{array}$$

where

$$P = \begin{bmatrix} 13 & 12 & -2 \\ 12 & 17 & 6 \\ -2 & 6 & 12 \end{bmatrix}, \quad q = \begin{bmatrix} -22.0 \\ -14.5 \\ 13.0 \end{bmatrix}, \quad r = 1.$$

Proof. We first note that x^* satisfies the constraints, and is therefore feasible. Let y be another feasible point. We verify the first order optimality condition on x^*.

$$\begin{aligned} \nabla f_0(x^*)^T(y - x^*) &= (Px^* + q^T)^T(y - x^*) \\ &= \begin{bmatrix} -1 & 0 & 2 \end{bmatrix} \begin{bmatrix} y_1 - 1 \\ y_2 - 1/2 \\ y_3 + 1 \end{bmatrix} \\ &= -y_1 + 2y_3 + 4. \end{aligned}$$

Since $-1 \leq y_1, y_3 \leq 1$, then

$$\nabla f_0(x^*)^T(y - x^*) \geq -1 + 2(-1) + 4 = 1 \geq 0.$$

So x^* satisfies the first order optimality condition and is therefore optimal.

□

4.8

Give an explicit solution of each of the following LP's.

4.8.a

Minimizing a linear function over an affine space.

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b. \end{array}$$

Proof. If the system $Ax = b$ is inconsistent, then the feasible set is empty and the minimization problem has no solution. Otherwise, $Ax = b$ is consistent, and we let x_0 be such that $Ax = b$. Then we can express by

$$X = \{x_0 + y : y \in \ker A\}$$

the feasible set. If $c \perp \ker A$, then for any $x \in X$, we have $x = x_0 + y$ where $y \in \ker A$. So

$$\begin{aligned} c^T x &= c^T(x_0 + y) \\ &= c^T x_0 + c^T y \\ &= c^T x_0 + 0 \\ &= c^T x_0. \end{aligned}$$

This tells us that $c^T x = c^T x_0$ for all $x \in X$, so $c^T x_0$ is the solution. Now if $c \not\perp \ker A$, then for some $y \in \ker A$, we have $c^T y \neq 0$. Now for any $t \in \mathbb{R}$,

$$A(x_0 + ty) = Ax_0 + tAy = b + 0 = b.$$

So $x_0 + ty \in X$ for all $t \in \mathbb{R}$. Consider now

$$c^T(x_0 + ty) = c^T x_0 + (c^T y)t.$$

So if $c^T y > 0$, then

$$\lim_{t \to -\infty} c^T(x_0 + ty) = -\infty.$$

And if $c^T y < 0$, then

$$\lim_{t \to +\infty} c^T(x_0 + ty) = -\infty.$$

So $c^T x$ is unbounded below for $x \in X$. In conclusion, the optimal value is

$$\begin{cases} c^T x, \text{ for any } x \in X & \text{if } Ax = b \text{ is consistent and } c \perp \ker A, \\ -\infty & \text{if } Ax = b \text{ is consistent and } c \not\perp \ker A, \\ \text{none} & \text{otherwise.} \end{cases}$$

□

4.8.b

Minimizing a linear function over a halfspace.

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a^T x \leq b, \end{array}$$

where $a \neq 0$.

Proof. Let $X = \{x : a^T x \leq b\}$ denote the feasible set. For each x, we write $x = t_x a + d_x$, where $d_x \perp a$. In other words, $t_x a$ is the projection of x onto a. Then

$$a^T x = a^T(t_x a + d_x) = t_x(a^T a) + a^T d_x = t_x(a^T a) + 0 = t_x(a^T a).$$

So $x \in X$ if and only if $t_x \leq b/(a^T a)$. Consider now $c = t_c a + d_c$. If $t_c > 0$, then for any $t \leq b/(a^T a)$, we have $ta \in X$ and

$$c^T(ta) = (t_c a + d_c)^T(ta) = t(t_c a^T a).$$

Then

$$\lim_{t\to-\infty} c^T(ta) = \lim_{t\to-\infty} t(t_c a^T a) = -\infty,$$

so $c^T x$ is unbounded below for $x \in X$. If $t_c < 0$, then for any $x \in X$,

$$c^T x = (t_c a + d_c)^T (t_x a + d_x) = t_x(t_c a^T a) + d_x d_c.$$

If $d_c = 0$, then $c^T x$ decreases as t_x increases, and since $t_x \le b/(a^T a)$, we know that $c^T x$ attains a minimum at $t_x = b/(a^T a)$, that is

$$\frac{b}{a^T a}(t_c a^T a) = b t_c = c^T x = (t_c a)^T x,$$

which implies $a^T x = b$. So $c^T x$ is minimized by any point on the hyperplane $a^T x = b$. If $d_c \ne 0$, then for any $x \in X$ and $k \in \mathbb{R}$,

$$a^T(x + kd_c) = a^T x + k a^T d_c = a^T x + 0 \le b.$$

So $x + kd_c \in X$ for any $x \in X$ and $k \in \mathbb{R}$. Then

$$c^T(x + kd_c) = c^T x + k c^T d_c,$$

which is linear with respect to k, and is therefore unbounded below. In conclusion, the optimal value is

$$\begin{cases} c^T x, \text{ for any } x \text{ s.t } ax = b & \text{if } a||c \text{ and } a^T c > 0, \\ -\infty & \text{otherwise.} \end{cases}$$

□

4.8.c

Minimizing a linear function over a rectangle.

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \ell \le x \le u, \end{array}$$

where ℓ and u satisfy $\ell \le u$.

Proof. For each index $i = 1, \ldots, n$, if $c_i > 0$, then $c_i x_i$ is increasing with respect to x_i, and attains a minimum at $x_i = \ell_i$. If $c_i < 0$, the $c_i x_i$ is decreasing with respect to x_i, and attains a minimum at u_i. If $c_i = 0$, then $c_i x_i$ is constant for all x_i. We choose x^* in the feasible set such that

$$x_i^* = \begin{cases} \ell_i & \text{if } c_i \ge 0, \\ u_i & \text{otherwise.} \end{cases}$$

Then for any y in the feasible set,

$$\begin{aligned} c^T y - c^T x^* &= \sum_{i=1}^n c_i y_i - \sum_{i=1}^n c_i x_i^* \\ &= \sum_{i=1}^n c_i y_i - c_i^* x_i^*) & \ge 0, \end{aligned}$$

Since each term is minimized by x^*. So the solution is $c^T x^*$.

$\square$

4.8.d

Minimizing a linear function over the probability simplex.

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{1}^T x = 1, \quad x \geq 0. \end{array}$$

What happens if the equality is replaces by an inequality $\mathbf{1}^T x \leq 1$?

Proof. Define x^* to be the vector with

$$x_i = \begin{cases} 1 & \text{if } c_i = \min\{c_j : j = 1, \ldots, n\} \text{ and } i \leq j \text{ for all } c_i = c_j, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, x^* is the vector with all zeros except for a 1 at the same index as the first index of the minimum value of c. The given feasibility conditions mean that for any feasible x, $c^T x$ is a convex combination of the elements of c. And any x which has a nonzero value at an index which is not a minimum of c will have a greater value than x^*. So

$$c^T x^* = \min\{c_i : i = 1, \ldots, n\}$$

is the solution.

If the equality is replaced with an inequality, then we construct x^* similarly, except if all the elements of c are positive, we take $x^* = 0$. So the solution is

$$\min(\{c_i : i = 1, \ldots, n\} \cup \{0\}).$$

$\square$

4.8.e

Minimizing a linear function over a unit box with a total budget constraint.

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \mathbf{1}^T x = \alpha, \quad 0 \leq x \leq \mathbf{1}. \end{array}$$

where α is an integer between 0 and n. What happens if α is not an integer (but satisfies $0 \leq \alpha \leq n$)? What if we change the equality to an inequality $\mathbf{1}^T x \leq \alpha$?

Proof. Similar to 4.8.d, for any feasible x, $c^T x$ is a linear combination of the elements of c, with each index of x limited between 0 and 1. In this case, we pick x^* such that it has 1's at the α indices which are least in c. In other words, if we sort the indices of c such that $c_{i_1} \leq \cdots \leq c_{i_n}$, then x^* will have 1's at indices $i_1, \ldots, i_\alpha$ and 0's elsewhere.

If α is not an integer, then we construct x^* similarly for $\lfloor \alpha \rfloor$ and define $x_{i_{\alpha+1}} = \alpha - \lfloor \alpha \rfloor$. That is, we similarly 'distribute' a total value of α across the indices of x^* starting with the indices which are minimum in c. Then as each index of x is 'filled up' to 1, we move to the next index which is the next smallest in c. In either case, the solution is

$$\sum_{k=1}^{\lfloor \alpha \rfloor} c_{i_k} + (\alpha - \lfloor \alpha \rfloor) c_{i_{\lfloor \alpha \rfloor + 1}},$$

where $(i_k)_{k=1}^n$ is a permutation of $\{1, \ldots, n\}$ with $c_{i_1} \leq \cdots \leq c_{i_n}$.

Also similar to 4.8.d, if the equality is replaced with an inequality, then we only 'fill up' indices of x so long as they are nonpositive indices of c. So the solution is

$$\sum_{k=1}^{\lfloor \alpha \rfloor} \min\{0, c_{i_k}\} + (\alpha - \lfloor \alpha \rfloor) \min\{0, c_{i_{\lfloor \alpha \rfloor + 1}}\},$$

where $(i_k)_{k=1}^n$ is a permutation of $\{1, \ldots, n\}$ with $c_{i_1} \leq \cdots \leq c_{i_n}$. □

4.8.f

Minimizing a linear function over a unit box with a weighted budget constraint.

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & d^T x = \alpha, \quad 0 \leq x \leq \mathbf{1}. \end{array}$$

with $d > 0$, and $0 \leq \alpha \leq \mathbf{1}^T d$.

Proof. This is equivalent to minimizing $c^T x$ subject to $\mathbf{1}^T x = \alpha, 0 \leq x \leq d$. In other words, this is the same as 4.8.e, except instead of each index of x being able to hold a maximum value of 1, each index can hold a maximum value of the corresponding index of d. So we have a similar solution as 4.8.e, but each term is multiplied by it's weight in d:

$$\sum_{k=1}^{\lfloor \alpha \rfloor} c_{i_k} d_{i_k} + (\alpha - \lfloor \alpha \rfloor) c_{i_{\lfloor \alpha \rfloor + 1}} d_{i_{\lfloor \alpha \rfloor + 1}},$$

where $(i_k)_{k=1}^n$ is a permutation of $\{1, \ldots, n\}$ with $c_{i_1} \leq \cdots \leq c_{i_n}$.

□

4.9

Consider the LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b, \end{array}$$

with A square and nonsingular. Show that the optimal value is given by

$$p^* = \begin{cases} c^T A^{-1} b & \text{if } A^{-T} c \leq 0, \\ -\infty & \text{otherwise.} \end{cases}$$

Proof. Denote the original minimization problem by (1). Denote by (2) the problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax + z = b, \\ & z \geq 0. \end{array}$$

We claim problems (1) and (2) are equivalent. Let X and Z be the feasible sets for (1) and (2), respectively. If $x \in X$, then $Ax \leq b$, which implies $Ax + z = b$ for some $z \geq 0$. Then the pair $(x, z) \in Z$. Likewise, if $(x, z) \in Z$, then $z \geq 0$ and $Ax + z = b$, so $Ax \leq b$. Then $x \in X$. This gives us a correspondence between the feasible sets of (1) and (2) and the objective function of each are the same. Thus, problems (1) and (2) are equivalent.

Now given some $z \geq 0$, we can find the necessary x such that $(x, z) \in Z$, by solving for x in the equality constraint. That is,

$$\begin{aligned} Ax + z &= b \\ Ax &= b - z \\ x &= A^{-1}(b - z). \end{aligned}$$

So for any $z \geq 0$, $(A^{-1}(b - z), z) \in Z$. So we can now write problem (3), equivalent to (2), which is found by substituting the solved value of x into the objective function:

$$\begin{array}{ll} \text{minimize} & f_0(z) = c^T A^{-1}(b - z) \\ \text{subject to} & z \geq 0. \end{array}$$

We claim that if $A^{-T}c \leq 0$, then $z = 0$ minimizes (3). To prove this, suppose $A^{-T}c \leq 0$ and let $y \geq 0$ be feasible for (3). Then consider

$$\begin{aligned} f_0(y) - f_0(z) &= c^T A^{-1}(b - y) - c^T A^{-1}(b - z) \\ &= c^T A^{-1}b - c^T A^{-1}y - c^T A^{-1}b + c^T A^{-1}0 \\ &= -c^T A^{-1}y. \end{aligned}$$

Now since $A^{-T}c \leq 0$, then $-c^T A^{-1} = -(A^{-T}c)^T \geq 0$. And since $y \geq 0$, we have

$$f_0(y) - f_0(z) \geq 0.$$

So $f_0(z) \leq f_0(y)$ for all feasible y. Thus, $z = 0$ is an optimal point for (3). This now implies that $(A^{-1}b, 0)$ is and optimal point for (2), and that $A^{-1}b$ is an optimal point for (1). Therefore, the optimal value of (1) is

$$c^T A^{-1}b.$$

Now if it is not the case that $A^{-T}c \leq 0$, then for some index i, we have $(A^{-T}c)_i = (c^T A^{-1})_i > 0$. Then for any $t \geq 0$, the point te_i is feasible for (3). Consider now

$$c^T A^{-1}(b - te_i) = c^T A^{-1}b - c^T A^{-1}te_i = c^T A^{-1}b - (c^T A^{-1})_i t,$$

which goes to $-\infty$ as $t \to \infty$. Thus (3) is unbounded below, and similarly, (2) and (1). Thus the optimal value of (1) is

$$p^* = \begin{cases} c^T A^{-1}b & \text{if } A^{-T}c \leq 0, \\ -\infty & \text{otherwise.} \end{cases}$$

□