École Polytechnique Fédérale de Lausanne
Prof. Marcus

Mar 31, 2022

Math 261 – Discrete Optimization (Spring 2022)

Assignment 6

Problem 1

Recall that for a vector $\mathbf{v} \in \mathbb{R}^m$,

$$\|\mathbf{v}\|_1 = \sum_{i=1}^{m} |v_i| \quad \text{and} \quad \|\mathbf{v}\|_\infty = \max\{|v_i| : 1 \le i \le m\}$$

and for an $m \times n$ matrix $\mathbf{A}$ and vector $\mathbf{b}$, consider the problems

$$\mathcal{P} = \inf\{\|\mathbf{Ax} - \mathbf{b}\|_\infty : \mathbf{x} \in \mathbb{R}^n\}$$

and

$$\mathcal{Q} = \sup\{\boldsymbol{\lambda} \cdot \mathbf{b} : \boldsymbol{\lambda}^\intercal \mathbf{A} = \mathbf{0}^\intercal, \|\boldsymbol{\lambda}\|_1 \le 1\}$$

(a) Show that P and Q provide certificates for each other — that is,

$$\boldsymbol{\lambda} \cdot \mathbf{b} \le \|\mathbf{Ax} - \mathbf{b}\|_\infty$$

whenever $\mathbf{x}$ is feasible in $\mathcal{P}$ and $\boldsymbol{\lambda}$ is feasible in $\mathcal{Q}$.

Solution:

Let $\boldsymbol{\lambda}$ be any vector in $\mathbb{R}^m$ that satisfies $\sum_i |\lambda_i| \le 1$ and $\boldsymbol{\lambda}^\intercal \mathbf{A} = \mathbf{0}^\intercal$ and let $\mathbf{y} = \mathbf{Ax} - \mathbf{b}$. Then

$$\boldsymbol{\lambda} \cdot \mathbf{b} = \boldsymbol{\lambda} \cdot \mathbf{Ax} - \boldsymbol{\lambda} \cdot \mathbf{y} = -\boldsymbol{\lambda} \cdot \mathbf{y}$$

since $\boldsymbol{\lambda}^\intercal \mathbf{A} = \mathbf{0}$. Hence

$$|\boldsymbol{\lambda} \cdot \mathbf{y}| = \left|\sum_i \lambda_i y_i\right| \le \sum_i |\lambda_i||y_i| \le \max_i |y_i| = \|\mathbf{Ax} - \mathbf{b}\|_\infty.$$

(b) Find a linear program $\mathcal{P}'$ which has the same optimal solution as $\mathcal{P}$ and a linear program $\mathcal{Q}'$ which has the same optimal solution as $\mathcal{Q}$ such that $\mathcal{P}'$ and $\mathcal{Q}'$ are duals of each other.

Solution:
We can turn both $\mathcal{P}$ and $\mathcal{Q}$ into linear programs using the same technique we used on the first problem set). The linear program for $\mathcal{P}$ is

$$\min\{t : \mathbf{Ax} - t\mathbf{1} \le \mathbf{b}, \mathbf{Ax} + t\mathbf{1} \ge \mathbf{b}, t \ge 0, \mathbf{x} \text{ free}\}.$$

The linear program for $\mathcal{Q}$ is a bit more involved — given a vector $\boldsymbol{\lambda}$, we split it into positive and negative parts:

$$\boldsymbol{\lambda}_i^+ = \max\{0, \boldsymbol{\lambda}_i\} \quad \text{and} \quad \boldsymbol{\lambda}_i^- = \min\{0, \boldsymbol{\lambda}_i\}$$

so that $\boldsymbol{\lambda} = \boldsymbol{\lambda}^+ + \boldsymbol{\lambda}^-$ and then the linear program for Q is

$$\max\{\boldsymbol{\lambda} \cdot \mathbf{b} : \boldsymbol{\lambda}^\intercal \mathbf{A} = \mathbf{0}^\intercal, (\boldsymbol{\lambda}^+ - \boldsymbol{\lambda}^-)^\intercal \mathbf{1} \le 1, \boldsymbol{\lambda}^+ \ge 0, \boldsymbol{\lambda}^- \le 0\}$$

and we see that the two are exact duals of each other.

(c) Part (b) implies that $\mathcal{P}$ and $\mathcal{Q}$ provide optimal certificates for each other — that is, there exists a $\mathbf{x}^*$ feasible in P and $\boldsymbol{\lambda}^*$ feasible in $\mathcal{Q}$ for which

$$\boldsymbol{\lambda}^* \cdot \mathbf{b} = \|\mathbf{A}\mathbf{x}^* - \mathbf{b}\|_\infty.$$

What do the complementary slackness conditions from part (b) say?

Solution:

Note that the dual variables $\boldsymbol{\lambda}^+$ correspond to the primal constraints

$$\mathbf{A}\mathbf{x} - \mathbf{b} \geq -t\mathbf{1}$$

and the $\boldsymbol{\lambda}^-$ correspond to the primal constraints

$$\mathbf{A}\mathbf{x} - \mathbf{b} \leq t\mathbf{1}.$$

Together, these tell us that

$$\lambda_i^* \big(|\operatorname{row}_i(\mathbf{A}) \cdot \mathbf{x}^* - \mathbf{b}| - t \big) = 0$$

for all i. In the other direction, the primal variable t corresponds to the dual constraint

$$\sum_i |\lambda_i| \leq 1$$

which tells us that

$$t\,(\|\boldsymbol{\lambda}^*\|_1 - 1) = 0.$$

Problem 2

Let $\mathcal{P}$ be the linear program

$$\mathcal{P} = \max \{\mathbf{0} \cdot \mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}.$$

(a) Find the dual of $\mathcal{P}$ and show that it is always feasible.

Solution:
The dual is $\mathcal{D} = \min \{\boldsymbol{\lambda} \cdot \mathbf{b} : \boldsymbol{\lambda}^\intercal \mathbf{A} \geq 0\}$ and it is always feasible because $\mathbf{0}$ is always a feasible solution.

(b) Use your answer to part (a) to prove the following lemma, one of the many versions of *Farkas' Lemma*

Lemma (Farkas)**.** *Let* $\mathbf{A}$ *be a matrix of dimension* $m \times n$ *and let* $\mathbf{b} \in \mathbb{R}^m$*. Then exactly one of the following holds:*

(I) There exists a vector $\mathbf{x} \geq \mathbf{0}$ *satisfying* $\mathbf{A}\mathbf{x} = \mathbf{b}$.
(II) There exists a vector $\boldsymbol{\lambda}$ *such that* $\boldsymbol{\lambda}^\intercal \mathbf{A} \geq \mathbf{0}^\intercal$ *and* $\boldsymbol{\lambda} \cdot \mathbf{b} < 0$.

Solution:
Showing that they cannot both be true is easy: assume (for contradiction) that this could happen. That is, there exists a $\mathbf{x}$ and $\boldsymbol{\lambda}$ satisfying both conditions. Then

$$\mathbf{x} \geq \mathbf{0} \quad \text{and} \quad \boldsymbol{\lambda}^\intercal \mathbf{A} \geq \mathbf{0}^\intercal \quad \Rightarrow \quad \boldsymbol{\lambda}^\intercal \mathbf{A}\mathbf{x} \geq 0 \quad \Rightarrow \quad \boldsymbol{\lambda}^\intercal \mathbf{b} \geq 0$$

But we assumed $\boldsymbol{\lambda} \cdot \mathbf{b} < 0$ (a contradiction).

Now we show that they cannot both be false. Assume that (I) is false — in other words, the linear program $\mathcal{P}$ is infeasible. We know that this forces $\mathcal{D}$ to be either (1) infeasible, or (2) unbounded, but we showed in part (a) that $\mathcal{D}$ has a feasible solution. Hence it must be that $\mathcal{D}$ is unbounded. But $\mathcal{D}$ is a minimization problem, so that means we can find a solution $\boldsymbol{\lambda}$ which is feasible for $\mathcal{D}$ and for which $\boldsymbol{\lambda} \cdot \mathbf{b}$ is as small as we want. In particular, we can find a $\boldsymbol{\lambda}$ for which $\boldsymbol{\lambda}^\intercal \mathbf{A} \geq \mathbf{0}$ and $\boldsymbol{\lambda} \cdot \mathbf{b} < 0$, so (II) must be true.

Hence they cannot both be true and they cannot both be false, so it must be that exactly one is true (and the other false).

Problem 3

In this problem, we will consider how we can get certificates of geometric statements. Two sets $X, Y \subseteq \mathbb{R}^n$ are said to be *separated by a hyperplane* if there exists a vector $\mathbf{v}$ and a real number c such that

$$\mathbf{v} \cdot \mathbf{x} < c \quad \text{for all } \mathbf{x} \in X \quad \text{and} \quad \mathbf{v} \cdot \mathbf{y} \geq c \quad \text{for all } \mathbf{y} \in Y$$

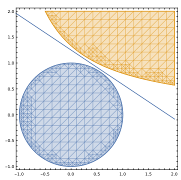
(a) Consider the regions

$$X = \left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\right\} \quad \text{and} \quad Y = \left\{(x, y) \in \mathbb{R}^2 : (x + 3/2)y \geq 2, y \geq 0\right\}$$

Show that X and Y can be separated by a hyperplane (find a valid $\mathbf{v}$ and c).

Note: You do not have to prove this formally — it would suffice to show a picture.

Solution:
The easiest way to do this is to plot both regions, draw a line between them, and then figure out what that line is. In this picture:



the line I drew has the equation $y = 5/4 - 2x/3$ which I can rearrange to get

$$8x + 12y = 15.$$

Hence if I let $\mathbf{v} = \begin{bmatrix} 8 \\ 12 \end{bmatrix}$ and $c = 15$, then (just by looking at the picture), it is clear that

$$\mathbf{v} \cdot \mathbf{x} < c \quad \text{for all } \mathbf{x} \in X \quad \text{and} \quad \mathbf{v} \cdot \mathbf{y} > c \quad \text{for all } \mathbf{y} \in Y$$

and so X and Y can be separated by a hyperplane.

(b) Let $\{\mathbf{u}_i\}_{i=1}^m$ be a collection of vectors in $\mathbb{R}^n$ and let

$$X = \text{cone}\{\mathbf{u}_1, \ldots, \mathbf{u}_m\} = \left\{\sum_i \alpha_i \mathbf{u}_i : \alpha_i \geq 0\right\}.$$

Show that, for any point $\mathbf{y} \in \mathbb{R}^n$, the following are equivalent (if and only if)

- $\mathbf{y} \notin X$
- $\mathbf{y}$ and X can be separated by a hyperplane that goes through the origin

Solution:
Let $\mathbf{A}$ be the matrix whose rows are the vectors $\mathbf{u}_i$. Then $\mathbf{y} \in X$ if and only if there exists an $\mathbf{x} \geq \mathbf{0}$ for which $\mathbf{A}\mathbf{x} = \mathbf{y}$. Similarly, $\mathbf{y}$ and X can be separated by a hyperplane which goes through the origin if and only if there exists a $\mathbf{v}$ for which $\mathbf{v}^\mathsf{T}\mathbf{A} \geq \mathbf{0}$ and $\mathbf{v} \cdot \mathbf{y} < 0$. Hence the equivalence follows from Problem 2.