ORIE 6300 Mathematical Programming I September 5, 2016

Lecture 8

Lecturer: Damek Davis *Scribe: Pamela Badian-Pessot*

1 Recap

- Finished proving strong duality
- Theorem of Alternatives

  Theorem 1 (Theorem of Alternatives) *Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$. Then exactly one of the following hold:*

  (a) $\exists x \in \mathbb{R}^n$ such that $Ax \leq b$
  (b) $\exists y \in \mathbb{R}^m$ such that $A^T y = 0$, $y \geq 0$, $b^T y < 0$.

- Value function and sensitivity analysis:

  We showed
  $$v(u) = \max\{c^T x | Ax \leq b + u\}$$
  is concave and piecewise linear.

2 Value Functions

Proposition 2 *Suppose that $v(0)$ exists and is finite. Then the following are equivalent:*

(a) $\forall u \in \mathbb{R}^n$, $v(u) \leq v(0) + u^T y^*$;

(b) $y^* \in \operatorname{argmin}\{b^T y | A^T y = c,\ y \geq 0\}$.

Proof: Let $x^*(u) \in Q(A, b+u)$ satisfy $v(u) = c^T x^*(u)$.

"$(a) \Rightarrow (b)$": Let u^* satisfy (a). We will prove that

1. $y^* \geq 0$
2. ${y^*}^T (b - Ax^*(0)) = 0$
3. $A^T y^* = c$

Which implies that y^* is dual optimal because

$$\min\{b^T y | A^T y = c,\ y \geq 0\} \geq c^T x^*(0) = (A^T y^*)^T x^*(0) = b^T y^* \geq \min\{b^T y | A^T y = c,\ y \geq 0\}.$$

Proof of 1. ($y^* \geq 0$)
$(\forall i)\ -\infty < v(0) \leq v(e_i)$ because $Ax^*(0) \leq b \leq b + e_i$. Thus, the inequality $v(e_i) \leq v(0) + e_i^T y^*$ implies

$$-e_i^T y^* \leq v(0) - v(e_i) \leq 0,$$

i.e., $y_i \geq 0$.

Proof of 2. (${y^*}^T(b - Ax^*(0)) = 0$)
First, $-\infty < v(0) \leq v(Ax^*(0) - b)$ because $x^*(0) \in Q(A, b + (Ax^*(0) - b))$. Thus the inequality $v(Ax^*(0) - b) \leq v(0) + (Ax^*(0) - b)^T y^*$ implies

$$0 \leq (b - Ax^*(0))^T y^* \leq v(0) - v(Ax^*(0) - b) \leq 0.$$

So ${y^*}^T(b - Ax^*(0)) = 0$.

Proof of 3. ($A^T y^* = c$)
Because $x \in Q(A, b + (Ax - b))$, we have

$$(\forall x \in \mathbb{R}^n) \quad c^T x + (b - Ax)^T y^* \leq v(Ax - b) + (b - Ax)^T y^* \leq v(0),$$

where the second inequality follows by assumption. Notice that

$$v(0) = c^T x^*(0) = c^T x^*(0) + (b - Ax^*(0))^T y^*.$$

Thus $f(x) = c^T x + (b - Ax)^T y^*$ is maximized at $x^*(0)$. Therefore

$$0 = \nabla f(x^*(0)) = c - A^T y^*$$

so $A^T y^* = c$ and y^* is dual optimal.

"$(a) \Leftarrow (b)$": If y^* is dual optimal then

$$v(0) = \max\{c^T x | Ax \leq b\} = b^T y^*.$$

If $v(u) = -\infty$, then the inequality is trivial. Otherwise, because y^* is dual feasible it follows that

$$v(u) = \min\{(b+u)^T y | A^T y = c,\ y \geq 0\} \leq (b+u)^T y* = v(0) + u^T y^*.$$

□

Remarks

- We have shown that the set of optimal dual solutions is exactly $-\partial[-V](0)$, i.e. the set of supgradients of v at 0. (See homework 3).
- By shifting v, we have also shown that the subgradients of v at u are exactly the solutions of $\min\{(b+u)^T y | A^T y = c,\ y \geq 0\}$.
- When v happens to be differentiable at u, we can show that

  $$\nabla v(u) = \arg\min\{(b+u)^T y | A^T y = c,\ y \geq 0\},$$

  which implies that the dual minimizer is unique.

- Clearly, piecewise linear functions are differentiable almost everywhere on their domains (the almost everywhere part comes from ties in the expression for v). Notice that

$$dom(v) = \{u|v(u) > -\infty\} \supseteq \{u|u \geq 0\}.$$

Corollary 3

$$P(\arg\min\{(b+u)^T y|y \geq 0,\ A^T y = c\} \text{ is unique}) = 1$$

where $u \sim Unif[0,1]^m$.

3 Fourier Motzkin

Several times in this class we've made the assumption that, for linear maps L,

$$LQ(A,b) = \{x|Ax \leq b\}$$

is a closed set. **This is a nontrival fact.** The next proof follows the proof from Jim Renegar's Excellent textbook [1].

Theorem 4 (Fourier Motzkin Elimination) *Let* $A \in \mathbb{R}^{m\times n}$*,* $b \in \mathbb{R}^m$ *and* $L \in \mathbb{R}^{k\times m}$*. Then the set*

$$LQ(A,b) = \{x|Ax \leq b\}$$

is a polyhedron.

Proof: Observation: It suffices to prove the result for a special kind of matrix, namely for P that project onto the first $n-1$ components of x.

$$P = \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & 1 & 0 \end{bmatrix}}_{n\text{-columns}} \qquad (n-1)\text{-rows}$$

Why? First by induction, if we show that $PQ(A,b)$ is a polyhedron, then we can repeatedly remove any number of coordinates of x and still get a polyhedron. For example we could remove $x_n, x_{n-1}, \ldots, x_{n-k}$ and the resulting set is still a polyhedron.

Then write down the set

$$\bar{Q} = \left\{ \begin{bmatrix} y \\ x \end{bmatrix} \middle| Ax \leq b, Lx - y \leq 0, -Lx + y \leq 0 \right\}.$$

Notice that $\begin{bmatrix} y & x \end{bmatrix}^T \in \bar{Q}$ if, and only if, $Ax \leq b$ and $Lx = y$.

If we remove the last $n+m$ coordinates of all vectors in $\bar{Q}$, the resulting set is still a polyhedron. Upon noticing that this set is $LQ(A,b) = \left\{ Lx \middle| Ax \leq b \right\}$ the conclusion of the theorem follows.

So lets work with P which projects onto the first $n-1$ coordinates. Let

$$\begin{aligned} I_0 &= \{i \in \{1,\ldots,m\} | a_{in} = 0\} \\ I_+ &= \{i \in \{1,\ldots,m\} | a_{in} > 0\} \\ I_- &= \{i \in \{1,\ldots,m\} | a_{in} < 0\} \end{aligned}$$

Then create a new matrix $\tilde{A} \in \mathbb{R}^{m \times n}$ and a new vector $\tilde{b} \in \mathbb{R}^m$.

$$\tilde{A} = \begin{bmatrix} \tilde{a}_1 \\ \vdots \\ \tilde{a}_m \end{bmatrix}, \qquad \tilde{b} = \begin{bmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_m \end{bmatrix}$$

where

$$\tilde{a}_i = \begin{cases} a_i & \text{if } i \in I_0 \\ \frac{a_i}{a_{in}} & \text{if } i \in I_+ \cup I_- \end{cases}, \qquad \tilde{b}_i = \begin{cases} b_i & \text{if } i \in I_0 \\ \frac{b_i}{a_{in}} & \text{if } i \in I_+ \cup I_- \end{cases}$$

Then $\tilde{a}_i = \begin{bmatrix} \gamma_i, \tau_i \end{bmatrix} \in R^{1 \times m}$ with $\tau_i \in \{1, 0\}$. Let $\bar{x} = (x_1, \ldots, x_{n-1})$. Then $Q(A, b)$ is described by three types of inequalities.

$$\begin{aligned} \gamma_i \bar{x} + 0x_n &\leq \tilde{b}_i & i \in I_0 \\ \gamma_i \bar{x} + x_n &\leq \tilde{b}_i & i \in I_+ \\ \gamma_i \bar{x} + x_n &\geq \tilde{b}_i & i \in I_-. \end{aligned}$$

Then rearrange,

$$\begin{aligned} 0 &\leq \tilde{b}_i - \gamma_i \bar{x} & i \in I_0 \\ x_n &\leq \tilde{b}_i - \gamma_i \bar{x} & i \in I_+ \\ x_n &\geq \tilde{b}_i - \gamma_i \bar{x} & i \in I_-. \end{aligned}$$

A point $\bar{x}$ is in $PQ(A, b)$ if, and only if, $\exists\ x_n \in \mathbb{R}$ s.t

$$0 \leq \tilde{b}_i - \gamma_i \bar{x}, \qquad i \in I_0$$

$$\max_{i \in I_-} \{\tilde{b}_i - \gamma_i \bar{x}\} \leq \min_{i \in I_+} \{\tilde{b}_i - \gamma_i \bar{x}\}$$

i.e. if, and only if,

$$0 \leq \tilde{b}_i - \gamma_i \bar{x}, \qquad i \in I_0$$

$$\tilde{b}_k - \gamma_k \overline{x} \leq x_n \leq \tilde{b}_i - \gamma_i \bar{x}, \qquad \forall k \in I_-,\ i \in I_+.$$

Thus $\bar{x} \in PQ(A, b)$ if and only if it satisfies a system of linear inequalities. Thus, $PQ(A, b)$ is a polyhedron. $\square$

Corollary 5 *Consider two polyhedra $Q(A_1, b_1)$, $Q(A_2, b_2) \subseteq \mathbb{R}^n$, then $Q(A_1, b_1) + Q(A_2, b_2)$ and $Q(A_1, b_1) - Q(A_2, b_2)$ are polyhedra and therefore closed.*

Remarks:

1. We implicitly used the fact that $\{A^T|y \geq 0, y^T(b - Ax) = 0\}$ was closed when proving the normal cone identities.

2. Calvin explicitly used the closedness of the above set to prove Farkas' Lemma.

3. Fourier-Motzkin elimination is a specific example of something called "Quantifier elimination." Used to prove certain classes of sets are closed under a variety of operations.

References

[1] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization.* Society for Industrial and Applied Mathematics, 2001.