ORIE 6300 Mathematical Programming I	September 8, 2016
Lecture 6	
Lecturer: Damek Davis	*Scribe: Angela Zhou*

1 Recap

- Every bounded polyhedron is a polytope; we proved this using set polars and the separating hyperplane theorem.
- The **normal cone** of a closed, convex set $S \subseteq \mathbb{R}^n$ is defined as

$$N_S : \mathbb{R}^n \to 2^{\mathbb{R}^n}$$

(power set on all points).

$$N_S(x) = \begin{cases} \{g \in \mathbb{R}^n | \, (\forall z \in S) \;\; g^T(z - x) \leq 0\} & \text{if } x \in S \\ \emptyset & \text{if } x \notin S \end{cases}$$

Normal cones characterize

- interiors and boundaries of convex sets:

$$N_S(x) = 0 \iff x \in int(S) \text{ and } N_S(x) \supsetneq \{0\} \iff x \in \partial S,$$

- Normal cones characterize projections: $X = P_S(y) \iff y - x \in N_S(x)$.

Theorem 1 (General Optimality Conditions) *Let $S \subseteq \mathbb{R}^n$ be a nonempty, closed, convex set and let $c \in \mathbb{R}^n$. Then the following are equivalent:*

1. *x^* solves:*

$$\max_{x \in S} c^T x \tag{1}$$

2. *$c \in N_S(x^*)$*
3. *$x^* = P_S(x^* + c)$*

Proof: Observe that x^* solves (1) if, and only if, $c^T(x - x^*) \leq 0$ which is true if, and only if, $c \in N_S(x^*)$. The inclusion $c \in N_S(x^*)$ holds if, and only if, $c = x^* + c - x^* \in N_s(x^*)$, which is true if, and only if, $x^* = P_S(x^* + c)$. □

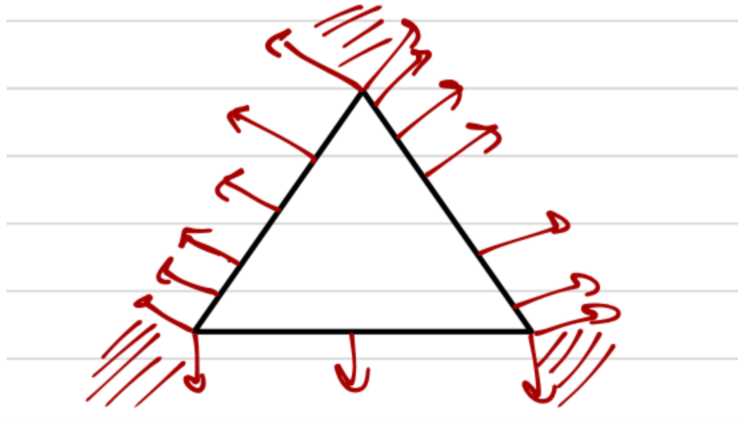
The following Corollary shows that bounded convex sets are round: whichever direction you pick, you can always find a point where that direction is in the normal cone.

Lemma 2 *Let $c \in \mathbb{R}^n$. Let S be a nonempty, closed, bounded, convex set. Then*

$$(\forall c \in \mathbb{R}^n), \; (\exists x(c) \in S) : \; c \in N_S(x(c))$$

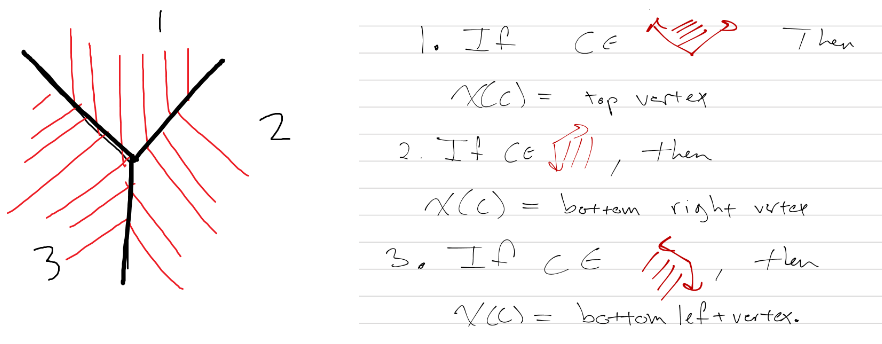
Proof: Let $x(c)$ be the maximizer of $c^T x$ over S. By the previous theorem and the Weierstrass' theorem, $c \in N_S(x(c))$. □

Example Suppose we take a triangle and write down its normal cone directions:



Given $c \in \mathbb{R}^n$, $f(x) = c^T x$ is maximized at one of three points:

Partition the space of $\mathbb{R}^n$ into three cells, each corresponding to the normal cone at a different vertex. Then depending on which cell c is in, choose the corresponding vertex at which the normal cone is located. The previous theorem then shows that $c^T x$ is maximized at that vertex.



2 Normal cone of a polyhedral set

In this class we care most about the normal cone of a polyhedron.

Theorem 3 *Let $A \in \mathbb{R}^{m\times n}$ and let $b \in \mathbb{R}^m$. Consider the polyhedron $Q(A,b) = \{x|Ax \le b\}$. Suppose $x \in Q(A,b)$, then $N_{Q(A,b)}(x) = \{A^T y | y \in \mathbb{R}^m$ such that $y \ge 0$ and $y^T(b - Ax) = 0\}$.*

The condition that $y^T(b - Ax)$ is equivalent to the complementarity conditions $b_i - a_i x > 0 \implies y_i = 0$ and $y_i > 0 \implies b_i - a_i x > 0$. This fact is actually the key to linear programming duality.

Proof:

Let $Y = \{A^T y | y \ge 0, y^T(b - Ax) = 0\}$.

"$\supseteq$":

Suppose $y \in \mathbb{R}^m, y^T(b - Ax) = 0$. Let $z \in Q(A,b)$ and expand:

$$y^T A(z - x) = \sum_{i: a_i x \neq b_i} y_i a_i (z - x) + \sum_{j: a_j x = b_j} y_j a_j (z - x)$$

By assumption $y_i = 0$ if $a_i x \neq b_i$, so the first sum is zero. We're left with

$$= \sum_{j: a_j x = b_j} y_j(a_j z - b_j) \leq 0,$$

which implies that $A^T y \in N_{Q(A,b)}$.

"$\subseteq$":

Suppose $g \in N_{Q(A,b)}(x)$ and $g \notin Y$. We wish to reach a contradiction.

By the separating hyperplane theorem, there exists a vector $\hat{a} \in \mathbb{R} \setminus \{0\}$ and a number $\hat{b} \in \mathbb{R}$ s.t.

$$(\forall w \in Y) \ \hat{a}^T w < b < \hat{a}^T g$$

Clearly, $0 \in Y$. Therefore, $0 = \hat{a}^T 0 < b$, i.e., $b > 0$.

Moreover, because $A^T = \begin{bmatrix} a_1^T & \dots & a_m^T \end{bmatrix}$, where a_i is the ith row vector of A, we find that for any i such that $a_i x = b_i$, we have

$$(\forall \lambda \geq 0) \ \lambda \hat{a}^T a_i^T = \hat{a}^T A \lambda e_i < \hat{b},$$

where e_i denotes the all-zeros vector with a 1 in the ith component. Note that $A^T \lambda e_i$ is in the Y set because it's in the range of $A^T y$ for an appropriate choice of y. Thus, that $(\forall \lambda > 0)$, $\hat{a}^T a^T < \frac{\hat{b}}{\lambda}$. Therefore take $\lambda \to \infty$ and note that $\frac{b}{\lambda} \to 0$ to show that $\hat{a}^T a_i^T \leq 0$.

Now, for each $\epsilon > 0$, define $z(\epsilon) = x + \epsilon \hat{a}$. We claim that $\exists \epsilon > 0$ s.t.

1. $z(\epsilon) \in Q(A, b)$; and
2. $g^T(z(\epsilon) - x) > \epsilon \hat{b} > 0$ which contradicts the inclusion $g \in N_{Q(A,b)}(x)$.

Proof of 1: Suppose that $a_i x = b_i$, which implies that $a_i \hat{a} \leq 0$. Then for all $\epsilon > 0$, $a_i z(\epsilon) = a_i(x + \epsilon \hat{a}) \leq a_i x \leq b_i$.

If $a_i x < b_i$ and $a_i \hat{a} \leq 0$, then a similar argument shows that for all $\epsilon > 0$, we have $a_i z(\epsilon) \leq b_i$.

On the other hand, if $a_i x < b_i$ and $a_i \hat{a} > 0$, then for all $0 < \epsilon < \frac{b_i - a_i x}{a_i \hat{a}}$, we have

$$\begin{aligned} a_i z(\epsilon) &= a_i x + \epsilon a_i \hat{a} \\ &\leq a_i x + \frac{b_i - a_i x}{a_i \hat{a}} \\ &= b_i. \end{aligned}$$

Therefore, we set $\epsilon = \min \left\{ \frac{b_i - a_i x}{a_i \hat{a}} | a_i \hat{a} > 0 \right\}$ and find that $z(\epsilon) \in Q(A, b)$.

Proof of 2: By assumption, take $\hat{a}^T g > \hat{b} > 0$, which implies that

$$g^T(z(\epsilon) - x) = g^T(x + \epsilon \hat{a} - x) = \epsilon g^T \hat{a} > \epsilon b > 0$$

$\square$

Normal cones arising in dual linear programs are now easy to compute.

Theorem 4 *Let $c \in \mathbb{R}^m$ and let $Y = \{y \in \mathbb{R}^m | y \geq 0, A^T y = c\}$, and let $y \in Y$. Then, we have*

$$N_Y(y) = \left\{ -b | \exists x \in \mathbb{R}^n \text{ with } Ax \leq b, y^T(b - Ax) = 0 \right\}$$

Proof: Write:

$$Y = \left\{ y \mid \begin{bmatrix} -I \\ A^T \\ -A^T \end{bmatrix} y \le \begin{bmatrix} 0 \\ c \\ -c \end{bmatrix} \right\}.$$

Then by previous theorem:

$$N_Y(y) = \left\{ \begin{bmatrix} -I & A^T & -A^T \end{bmatrix} \begin{bmatrix} s \\ t \\ w \end{bmatrix} \middle| \begin{bmatrix} s \\ t \\ w \end{bmatrix} \ge 0, \begin{bmatrix} s \\ t \\ w \end{bmatrix}^T \left(\begin{bmatrix} -I \\ A^T \\ -A^T \end{bmatrix} y - \begin{bmatrix} 0 \\ c \\ -c \end{bmatrix} \right) \ge 0 \right\}$$

Equivalently:

$$= \{-s + At - Aw | s, t, w \ge 0, s^T(-y) + t^T(A^T y - c) + w^T(-A^T y + c) = 0\}$$

Replace t, w with $z = t - w$ and let $A^T y = c$, to get

$$= \{-s + Az | s \ge 0, s^T y = 0\}$$

$$= \{\bar{b} | \bar{b} \le Az, y^T(\bar{b} - Az) = 0\},$$

where we solved for s in this expression and used the identity that $s = Az - b \ge 0$. Replace z with $-x$ and $\bar{b}$ by $-b$ to get.

$$= \{-b | Ax \le b, y^T(b - Ax) = 0\}.$$

□

3 Strong Duality

Consider the primal linear program

$$p^* = \text{maximize}_{Ax \le b}\ c^T x$$

and the dual linear program

$$d^* = \text{minimize}_{\substack{A^T y = c \\ y \ge 0}} b^T y$$

Then exactly one of the following holds:

- Both the primal and dual problems are infeasible, i.e., $p^* = -\infty$ and $d^* = \infty$.
- The maximizer of the primal and the minimizer of the dual exist, i.e., p^* and d^* are finite and $p^* = d^*$.
- The primal objective is unbounded over the feasible set and the dual is infeasible, i.e., $d^* = \infty$ and $p^* = \infty$.
- The dual objective is unbounded over the feasible set and the primal is infeasible, i.e., $p^* = -\infty$ and $d^* = -\infty$

Proof: **Proof of strong duality**

- Part 1: See Homework 2 where you provided an example of primal and dual both being infeasible.
- Part 2:

– (a) Suppose that p^* is finite and let x^* maximizes the primal program. Then, by Theorem 1, we have $c \in N_{Q(a,b)}(x^*) = \{A^T y | y \geq 0, y^T(b - Ax^*) = 0\}$. So we can take any such y from the normal cone and write $c = A^T y^*$ where $y^* \geq 0, (y^*)^T(b - Ax^*) = 0$. Then y^* is dual feasible and

$$p^* = c^T x^* = (A^T y^*)^T x^* = (y^*)^T A x^* = (y^*)^T b$$

Thus by weak duality, y^* is dual optimal and $p^* = d^*$.

– (b) Suppose d^* is finite and let y^* be a minimizer of the dual. In order to apply Theorem 1, we need to replace the dual objective by $-b^T y$. Then

$$-b \in N_{\{y|A^T y = c, y \geq 0\}}(y^*) = \{-\bar{b} | Ax \leq \bar{b}, (y^*)^T(\bar{b} - Ax)\}$$

Choose any x^* such that $Ax^* \leq b$ and $(y^*)^T(b - Ax^*) = 0$. Then

$$d^* = (y^*)^T b = (y^*)^T(Ax^*) = (A^T y^*) x^* = c^T x^*.$$

Thus, by weak duality, we conclude that x^* is primal optimal and $p^* = d^*$.

The rest of the proof will be presented in the next lecture.

□