Introduction to Optimization

CO 250

Kevin Shu

Lecture 16

Live (interactive) version of these slides: https://ootks.github.io/co250-lectures/lecture16.html

Lecture Outline

- Review of Complementary Slackness
- Strong Duality and Farkas' Lemma
- Normal Cones

Review: Complementary Slackness

Review: Duality

- Every linear program has a dual linear program.
- Dual constraints correspond to primal variables.
- Dual variables correspond to primal constraints.
- Variable sign conditions should be thought of as being part of the variable definition and not a separate constraint for these purposes.
- Inequality constraints correspond to dual variables with sign constraints and vice versa.

Review: Complementary Slackness

- For each primal inequality constraint, the dual variable will either be 0 or nonzero.
- For an optimal solution x^* to the LP, a primal inequality is tight if the left and right sides of the inequality are equal at x^*. It is slack otherwise.
- **Theorem:** If x^* is an optimal solution to an LP, and a given primal inequality is slack, then the dual variable is equal to 0.
- Each LP is the dual of its dual LP, so the dual concept holds.
- **Theorem:** If x^* is an optimal solution to an LP, and $x_i \geq 0$ is a nonnegativty constraint, and $x_i^* > 0$, then for any optimal solution y^* to the dual LP, the dual inequality in the dual LP is tight.

Complementary Slackness Example

Primal

$$\begin{aligned} \max \quad & 3x_1 + 4x_2 - x_3 \\ \text{s.t.} \quad & x_1 + 2x_2 + x_3 \leq -1, \\ & 2x_1 + x_2 - x_3 = 1, \\ & x_1 - x_2 + 2x_3 \geq -3, \\ & x_1 \geq 0,\ x_2 \text{ free},\ x_3 \leq 0. \end{aligned}$$

Dual

$$\begin{aligned} \min \quad & -y_1 + y_2 - 3y_3 \\ \text{s.t.} \quad & y_1 + 2y_2 + y_3 \geq 3, \\ & 2y_1 + y_2 - y_3 = 4, \\ & y_1 - y_2 + 2y_3 \leq -1, \\ & y_1 \geq 0,\ y_2 \text{ free},\ y_3 \leq 0. \end{aligned}$$

The optimal primal solution: $x^* = (0, 0, -1)$, with optimal value 1.

Complementary Slackness Example

Primal		Dual
Constraints $\leftrightarrow$ *dual variables*		
$x_1 + 2x_2 + x_3 \leq -1$	$\leftrightarrow$	$y_1 \geq 0$
$2x_1 + x_2 - x_3 = 1$	$\leftrightarrow$	y_2 free
$x_1 - x_2 + 2x_3 \geq -3$	$\leftrightarrow$	$y_3 \leq 0$
Variables $\leftrightarrow$ *dual constraints*		
$x_1 \geq 0$	$\leftrightarrow$	$y_1 + 2y_2 + y_3 \geq 3$
x_2 free	$\leftrightarrow$	$2y_1 + y_2 - y_3 = 4$
$x_3 \leq 0$	$\leftrightarrow$	$y_1 - y_2 + 2y_3 \leq -1$

Complementary Slackness Example

Primal at x^*		**Dual**
Constraints $\leftrightarrow$ dual variables		
$x_1^* + 2x_2^* + x_3^* = -1 \leq -1$	$\leftrightarrow$	$y_1 \geq 0$
$2x_1^* + x_2^* - x_3^* = 1 = 1$	$\leftrightarrow$	y_2 free
$x_1^* - x_2^* + 2x_3^* = -2 \geq -3$	$\leftrightarrow$	$y_3 \leq 0$ **(tight)**
Variables $\leftrightarrow$ dual constraints		
$x_1^* = 0 \geq 0$	$\leftrightarrow$	$y_1 + 2y_2 + y_3 \geq 3$
$x_2^* = 0$ *(free)*	$\leftrightarrow$	$2y_1 + y_2 - y_3 = 4$ **(tight)**
$x_3^* = -1 \leq 0$	$\leftrightarrow$	$y_1 - y_2 + 2y_3 \leq -1$ **(tight)**

Red: inequalities with slack at $x^* = (0, 0, -1)$. Green: rows with no slack (tight).

Complementary Slackness Example

We find from the previous considerations that for any dual optimal solution, $y_3 = 0$ (its primal constraint has slack), and the dual constraints for x_2 (free) and $x_3 < 0$ (sign constraint with slack) are tight:

$$y_3 = 0$$
$$2y_1 + y_2 - y_3 = 4$$
$$y_1 - y_2 + 2y_3 = -1$$

There is a unique solution to these 3 equations in 3 unknowns.

$$y^* = (y_1, y_2, y_3) = (1, 2, 0).$$

Complementary Slackness Example II

Complementary Slackness Example II

Primal

$$\begin{aligned} \max \quad & 4x_1 + 4x_2 - 3x_3 \\ \text{s.t.} \quad & x_1 + x_2 + x_3 = 2, \\ & x_1 - x_2 + 2x_3 = 5, \\ & 2x_1 + x_2 - x_3 = 2, \\ & x_1 \geq 0,\ x_2 \text{ free},\ x_3 \geq 0. \end{aligned}$$

Dual

$$\begin{aligned} \min \quad & 2y_1 + 5y_2 + 2y_3 \\ \text{s.t.} \quad & y_1 + y_2 + 2y_3 \geq 4, \\ & y_1 - y_2 + y_3 = 4, \\ & y_1 + 2y_2 - y_3 \geq -3, \\ & y_1, y_2, y_3 \text{ free}. \end{aligned}$$

The optimal primal solution: $x^* = (2, -1, 1)$, with optimal value 1.

Complementary Slackness Example II

Primal		Dual
Constraints $\leftrightarrow$ *dual variables*		
$x_1 + x_2 + x_3 = 2$	$\leftrightarrow$	y_1 free
$x_1 - x_2 + 2x_3 = 5$	$\leftrightarrow$	y_2 free
$2x_1 + x_2 - x_3 = 2$	$\leftrightarrow$	y_3 free
Variables $\leftrightarrow$ *dual constraints*		
$x_1 \geq 0$	$\leftrightarrow$	$y_1 + y_2 + 2y_3 \geq 4$
x_2 free	$\leftrightarrow$	$y_1 - y_2 + y_3 = 4$
$x_3 \geq 0$	$\leftrightarrow$	$y_1 + 2y_2 - y_3 \geq -3$

Complementary Slackness Example II

Primal at x^*		**Dual**
Constraints $\leftrightarrow$ dual variables		
$x_1^* + x_2^* + x_3^* = 2 = 2$	$\leftrightarrow$	y_1 free
$x_1^* - x_2^* + 2x_3^* = 5 = 5$	$\leftrightarrow$	y_2 free
$2x_1^* + x_2^* - x_3^* = 2 = 2$	$\leftrightarrow$	y_3 free
Variables $\leftrightarrow$ dual constraints		
$x_1^* = 2 \geq 0$	$\leftrightarrow$	$y_1 + y_2 + 2y_3 \geq 4$ **(tight)**
$x_2^* = -1$ *(free)*	$\leftrightarrow$	$y_1 - y_2 + y_3 = 4$ **(tight)**
$x_3^* = 1 \geq 0$	$\leftrightarrow$	$y_1 + 2y_2 - y_3 \geq -3$ **(tight)**

Red: inequalities with slack at $x^* = (2, -1, 1)$. Green: rows with no slack (tight).

Complementary Slackness Example II

The dual constraints associated with $x_1^*, x_3^* > 0$ and with the free variable x_2 are tight for any dual optimal solution:

$$y_1 + y_2 + 2y_3 = 4$$
$$y_1 - y_2 + y_3 = 4$$
$$y_1 + 2y_2 - y_3 = -3$$

There is a unique solution to these 3 equations in 3 unknowns.

$$y^* = (y_1, y_2, y_3) = (1, -1, 2).$$

Farkas' Lemma and Strong Duality

Farkas' Lemma

Recall Farkas' lemma.

Farkas' Lemma
For $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, exactly one of the following two things holds:

- There is some $x \in \mathbb{R}^n$ so that $Ax = b$, and $x \geq 0$.
- There is some $y \in \mathbb{R}^m$ so that $A^\mathsf{T} y \geq 0$ and $b^\mathsf{T} y < 0$.

We discussed a geometric proof of this using the separating hyperplane theorem and now we will prove it using strong duality.

Farkas' Lemma Proof

We have already seen that if there is y so that $A^\intercal y \geq 0$ and $b^\intercal y < 0$, then there cannot be an element of the set $\{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$. If there were, then

$$\underbrace{y^\intercal Ax}_{\substack{\geq 0 \\ A^\intercal y \geq 0,\ x \geq 0}} = \underbrace{y^\intercal b}_{<0}.$$

So we to prove the other direction, we will need strong duality. What LP are we going to apply the strong duality theorem to?

Farkas' Lemma Proof

Consider the linear program

$$\begin{aligned} \max \quad & \vec{0}^\intercal x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

This is infeasible precisely when the set $\{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$ is empty.

Strong duality implies that the dual LP does not have an optimal solution (if it did, then the primal LP would have an optimal solution by strong duality).

So the dual LP is either unbounded or infeasible.

Farkas' Lemma Proof

The dual LP is

$$\begin{aligned} \min \quad & b^\intercal y \\ \text{s.t.} \quad & A^\intercal y \geq 0 \end{aligned}$$

Why is this feasible?

It's feasible since $y = 0$ is feasible.

Since the dual LP is unbounded, we have seen that there must in particular be a feasible y so that $b^\intercal y < 0$.

We conclude that there is some y so that $A^\intercal y \geq 0$ and $b^\intercal y < 0$, which is what we wanted.

Farkas' Lemma Variants

We can use more general linear programming duality results to get other versions of Farkas' lemma:

Farkas' Lemma Version II
For $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, exactly one of the following two things holds:

- There is some $x \in \mathbb{R}^n$ so that $Ax \leq b$.
- There is some $y \in \mathbb{R}^m$ so that $y \geq 0$, $A^\intercal y = 0$ and $b^\intercal y < 0$.

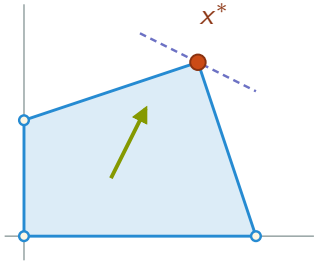
We can prove this by noting that the dual of the linear program

$$\begin{array}{ll} \max & \vec{0}^\intercal x \\ \text{s.t.} & Ax \leq b \end{array} \quad \text{is} \quad \begin{array}{ll} \min & b^\intercal y \\ \text{s.t.} & A^\intercal y = 0 \\ & y \geq 0 \end{array}$$

Normal Cones

Normal Cones

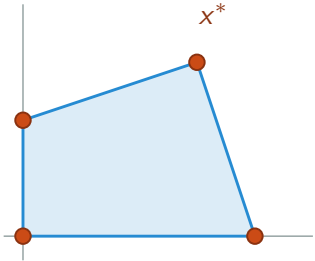
The feasible region of a linear program is a polytope, and we have seen that (typically) the unique maximizer of a linear function over a polytope is a vertex of that polytope.



$$\begin{aligned} \max \quad & x_1 + 2x_2 \\ \text{s.t.} \quad & x_1 \geq 0, \\ & x_2 \geq 0, \\ & 3x_1 + x_2 \leq 12, \\ & -x_1 + 3x_2 \leq 6. \end{aligned}$$

Normal Cones

So while there are infinitely many possible objective directions c, there are only finitely many possible outcomes of the optimal solution of the LP.



$$\begin{aligned} \max \quad & c^\intercal x \\ \text{s.t.} \quad & x_1 \geq 0, \\ & x_2 \geq 0, \\ & 3x_1 + x_2 \leq 12, \\ & -x_1 + 3x_2 \leq 6. \end{aligned}$$

Interactive Demo: Normal Cones in the Plane

Interactive demo — open in your browser

Normal Cones 2D Explorer

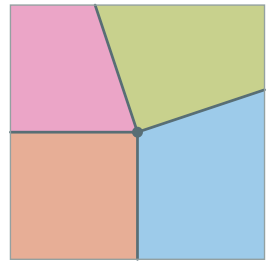
Interactive quadrilateral and the plane of objective vectors, colored by optimal vertex. It is embedded in the live slide and cannot run inside a PDF.

https://ootks.github.io/co250-lectures/demos/NormalCones2D.html

Normal Cones

For each vertex v in a polytope P, we define the normal cone N_v to be the set of c so that the maximizer of the linear program $\max\{c^\intercal x : x \in P\}$ is v:

$$N_v = \{c \in \mathbb{R}^n : \text{For any } x \in P,\ c^\intercal v \geq c^\intercal x\}.$$



Normally, with linear programming, we are given the objective direction c and want to find a vertex $v \in P$ maximizing $c^\intercal v$; the normal cone flips this idea on its head.

Can we get a more explicit description of the normal cone of a polytope?

Normal Cones

Normal Cone Characterization (SEF Polytopes)
Suppose that $P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$, and let $v \in P$ be a vertex (i.e. a basic feasible solution). Then $c \in N_v$ if and only if there exists some $y \in \mathbb{R}^m$ so that

$$c \leq A^\intercal y,$$

and $c_i = (A^\intercal y)_i$ whenever $v_i > 0$.

This is really a rephrasing of complementary slackness.

Normal Cones

Example of Theorem

Consider a polytope defined by

$$P = \{x \in \mathbb{R}^4 : Ax = b,\ x \geq 0\} \text{ with } A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ 6 \end{pmatrix}.$$

Here is a vertex of this polytope: $v = (3, 1, 0, 0)$.

Let's use the theorem to check that $c = (3, 5, 1, 0) \in N_v$ (i.e. v is a maximizer of $c^\mathsf{T} x$ in P).

Let $y = (2, 1)$, and note that $A^\mathsf{T} y = (3, 5, 2, 1)$. We see that

$$c = A^\mathsf{T} y - (0, 0, 1, 1).$$

Normal Cones

Proof of Theorem

c is in N_v if and only if v is an optimal solution of the following linear program:

$$\max\{c^\mathsf{T}x : Ax = b, x \geq 0\}$$

The dual linear program is

$$\min\{b^\mathsf{T}y : A^\mathsf{T}y \geq c\}.$$

Strong duality implies that the dual linear program is feasible, so there is some y so that $A^\mathsf{T}y \geq c$, and by complementary slackness, we know that $(A^\mathsf{T}y)_i = c_i$ whenever $v_i > 0$.

Normal Cones

The above theorem shows one characterization of the normal cone, but it does not make it easy to check if a given c is in the normal cone without solving a linear program. We can give a simpler characterization (at least when the vertex is nondegenerate).

Normal Cone Characterization (SEF Polytopes)
Let $P = \{x : Ax = b,\ x \geq 0\}$, and let v be a vertex of P with basis B with m positive entries (so v is nondegenerate). For each c, we can break up its coordinates into those coordinates in the basis B and those outside it: $c = \begin{pmatrix} c_B & c_N \end{pmatrix}$.

$$N_v = \{c \in \mathbb{R}^n : c - A^\intercal A_B^{-\intercal} c_B \leq 0\}.$$

Normal Cones

Proof of Theorem This can be seen from the terminal condition of the simplex method.

We can also see it from complementary slackness; using these dual LPs:

$$\max\{c^\mathsf{T}x : Ax = b, x \geq 0\} \qquad \min\{b^\mathsf{T}y : A^\mathsf{T}y \geq c\}.$$

Complementary slackness implies that v is an optimal solution to the primal LP if and only if there is a feasible y^* for the dual LP that satisfies $(A^\mathsf{T}y)_i = c_i$ whenever $v_i > 0$.

This implies that the unique such y^* is precisely $y^* = A_B^{-\mathsf{T}} c_B$.

Plugging in, we get that v is an optimal solution to the primal LP if and only if $A^\mathsf{T} A_B^{-\mathsf{T}} c_B \geq c$, which is what we wanted to show.

Normal Cones

We can also characterize normal cones for polytopes described by inequalities.

Normal Cone Characterization (Polytopes $Ax \leq b$)
Suppose that $P = \{x \in \mathbb{R}^n : Ax \leq b\}$, and let $v \in P$ be a vertex. Then $c \in N_v$ if and only if there exists some $y \in \mathbb{R}^m$ so that

$$c = A^\mathsf{T} y, \qquad y \geq 0,$$

and $y_i = 0$ whenever $(Av)_i < b_i$.

One way to think of this is as follows: for each v in P, there is a collection of tight constraints, $a_1^\mathsf{T} x \leq b_1, \ldots, a_k^\mathsf{T} x \leq b_k$ where $k \geq n$, and where $a_i^\mathsf{T} v = b_i$ for each i. We then have that the normal cone is

$$N_v = \Big\{ \sum_{i=1}^{k} y_i a_i : y \geq 0 \Big\}.$$

Normal Cone Example

Let's recall that the n dimensional hypercube is the polytope defined by

$$P = \{x \in \mathbb{R}^n : 0 \leq x_i \leq 1 \text{ for } i = 1, \ldots, n\}.$$

So $A = \begin{pmatrix} I \\ -I \end{pmatrix}$, and $b = \begin{pmatrix} \vec{1} \\ \vec{0} \end{pmatrix}$.

A vertex v to P is a point where n of the constraints are tight, i.e. a point where every entry is either 0 or 1.

The tight constraints at v have normals e_i (when $v_i = 1$) and $-e_i$ (when $v_i = 0$), so

$N_v = \{c \in \mathbb{R}^n : c_i \geq 0 \text{ if } v_i = 1,\ c_i \leq 0 \text{ if } v_i = 0\}$.

