AM 221: Advanced Optimization — Spring 2018

Prof. Rasmus Kyng — *Section 4 — Tuesday, Feb. 19th*

1. Taking duals. Taking the dual of an optimization problem is a powerful tool for designing efficient algorithms. The solution to the dual problem of a minimization problem provides a lower bound to the solution of the primal problem (and an upper bound if the primal is a maximization problem). Given the following optimization problem:

$$\begin{aligned}\min_{\mathbf{x}\in\mathbb{R}^m}\ & z(\mathbf{x}) = \mathbf{c}^\intercal\mathbf{x} \\ \text{s.t } & A^\intercal\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0\end{aligned}$$

To find the dual, we first choose a vector $\mathbf{y}$ to create a new equality:

$$\begin{aligned}\mathbf{y}^\intercal A\mathbf{x} &= \mathbf{y}^\intercal\mathbf{b} \\ 0 &= \mathbf{y}^\intercal\mathbf{b} - \mathbf{y}^\intercal A\mathbf{x}\end{aligned}$$

We add this the objective function:

$$\begin{aligned}z(\mathbf{x}) &= \mathbf{c}^\intercal\mathbf{x} + \mathbf{y}^\intercal\mathbf{b} - \mathbf{y}^\intercal A\mathbf{x} \\ &= \mathbf{y}^\intercal\mathbf{b} + (\mathbf{c}^\intercal - \mathbf{y}^\intercal A)\mathbf{x}\end{aligned}$$

Suppose we choose $\mathbf{y}$ in such a way that $\mathbf{c}^\intercal - \mathbf{y}^\intercal A \geq 0$ and $\mathbf{x} \geq 0$ is feasible. Then we would have that $\mathbf{c}^\intercal - \mathbf{y}^\intercal A\mathbf{x} \geq 0$ and $z(\mathbf{x}) \geq \mathbf{y}^\intercal\mathbf{b}$. This second piece tells us that we now have a lower bound on the objective function. For our minimization problem we want to know the largest lower bound possible, so we will formulate this as an optimization problem. We see that:

$$\begin{aligned}\mathbf{c}^\intercal - \mathbf{y}^\intercal A &\geq 0 \\ &= \mathbf{y}^\intercal A \leq \mathbf{c}^\intercal \\ &= A^\intercal\mathbf{y} \leq \mathbf{c}\end{aligned}$$

This gives us our constraints, and we can state the new optimization problem, which we call the dual problem, as:

$$\begin{aligned}\max_{\mathbf{y}\in\mathbb{R}^n}\ & \mathbf{b}^\intercal\mathbf{y} \\ \text{s.t } & A^\intercal\mathbf{y} \leq \mathbf{c}\end{aligned}$$

Weak duality versus strong duality

Plainly stated, the weak duality theorem states that any feasible solution of one problem corresponds to a bound on the other. In the notation from the example above, this means $\mathbf{b}^\intercal\bar{\mathbf{y}} \leq \mathbf{c}^\intercal\bar{\mathbf{x}}$ for any feasible solutions $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$. The strong duality theorem takes this a step further and states that that the optimal values of the primal and dual are always equal. The difference between the optimal values of the primal and the dual is called the duality gap, and for strong duality to hold the gap must be 0. Ergo, if $\mathbf{b}^\mathsf{T}\bar{\mathbf{y}} = \mathbf{c}^\mathsf{T}\bar{\mathbf{x}}$ then $\bar{\mathbf{x}} = \bar{\mathbf{y}}$.

Relationship between primal and dual

Recall that the fundamental theorem of linear programming states that a linear program is either infeasible, unbounded, or optimal. As a corrollary we make the following three statements:

- dual has a feasible solution $\Rightarrow$ primal is bounded
- primal has a feasible solution $\Rightarrow$ dual is bounded
- both have feasible solutions $\Rightarrow$ both are optimal

Constructing the dual

The following table shows the relationship between constraint and objective function inequalities for the primal and the dual.

	$\geq$	$\Leftrightarrow$	≥ 0	
$\min c^T x$	$=$	$\Leftrightarrow$	free	$\max b^T y$
s.t.	$\leq$	$\Leftrightarrow$	≤ 0	s.t.
$Ax \square b$	≥ 0	$\Leftrightarrow$	$\leq$	$y^T A \square c$
$x \square 0$	free	$\Leftrightarrow$	$=$	$y \square 0$
	≤ 0	$\Leftrightarrow$	$\geq$	

2. Extreme Points. Given a polyhedron $P \in \{\mathbf{x} \in \mathbb{R}^n : A^*\mathbf{x} \leq \mathbf{b}^*\}$, we can illustrate that $\bar{\mathbf{x}} \in P$ is an extreme point of P if and only if $\mathbf{x}$ is a basic feasible solution of P. We let $A\mathbf{x} = \mathbf{b}$ be the set of tight constraints for P (recall that tight constraints are those that don't need slack variables when converting inequalities to equalities).

We will first show that $\bar{\mathbf{x}}$ is a basic feasible solution if $rank(A) = n$. [We note that this is also true in the reverse, but we leave the proof as an exercise]. For a matrix A, we can say that a set of column indices B forms a basis if the matrix A_B is a square nonsingular matrix. For a system of equations $A\mathbf{x} = \mathbf{b}$, we can further say that the elements $x_i \in \mathbf{x}$ are basic when $i \in B$ and non-basic otherwise. We decompose the matrix $A\mathbf{x}$ into a sum of basis and non-basis matrices as follows:

$$\begin{aligned} A\mathbf{x} &= \sum_{i=1}^{n} x_i A_i \\ &= \sum_{i \in B} x_i A_i + \sum_{i \in N} x_i A_i \\ &= A_B \mathbf{x}_b + A_N \mathbf{x}_N \end{aligned}$$

By definition, a vector $\bar{\mathbf{x}}$ is a basic feasible solution of $A\mathbf{x} = \mathbf{b}$ if $A\bar{\mathbf{x}} = \mathbf{b}$, $\bar{\mathbf{x}}_N = 0$, and $\bar{\mathbf{x}} \geq 0$ all hold true. If $\bar{\mathbf{x}}$ is such a solution, then

$$\begin{aligned} \mathbf{b} &= A\bar{\mathbf{x}} \\ &= A_B \bar{\mathbf{x}}_B + A_N \bar{\mathbf{x}}_B \\ &= A_B \bar{\mathbf{x}}_B \end{aligned}$$

Notice that this tells us that A is a full-rank matrix for which the rows are linearly independent. This gives us $rank(A) = n$ and thus $\bar{\mathbf{x}}$ is a basic feasible solution.

Now we will suppose that $rank(A) = n$ and show that $\bar{\mathbf{x}}$ is an extreme point via contradiction. We know by definition that $\bar{\mathbf{x}} \in P$ is not an extreme point of P if and only if $\bar{\mathbf{x}} = \lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ for distinct points $\mathbf{x}_1, \mathbf{x}_2 \in P$ and $\lambda \in (0, 1)$. This means that

$$\begin{aligned} \mathbf{b} &= A\bar{\mathbf{x}} \\ &= A(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \\ &= \lambda \mathbf{b} + (1 - \lambda \mathbf{b}) \geq \lambda A\mathbf{x}_1 + (1-\lambda)A\mathbf{x}_2 \end{aligned}$$

This implies that $A\mathbf{x}_1 = A\mathbf{x}_2 = \mathbf{b}$. Since $rank(A) = n$, there exists a unique solution to $A\mathbf{x} = \mathbf{b}$. Therefore, $\bar{\mathbf{x}} = \mathbf{x}_1 = \mathbf{x}_2$, which is a contradiction. [We leave the other direction as an exercise].

Geometric example

Let's consider the linear program $\max \mathbf{c}^\intercal \mathbf{x} : A\mathbf{x} \leq \mathbf{b}$ where

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

If we take $\bar{\mathbf{x}}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, we can see in Figure 1 below that $\bar{\mathbf{x}}$ corresponds to an extreme point. Notice that constraints 1 and 3 are tight, which means we can define $A_1^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. We see that $rank(A_1^*) = 2 = n$, which verifies that $\bar{\mathbf{x}}_1$ is indeed an extreme point.

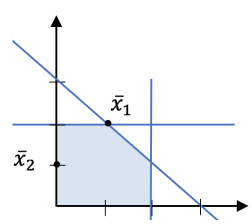

Figure 1: $\bar{\mathbf{x}}_1$ is an extreme point and a basic feasible solution.

If we take $\bar{\mathbf{x}}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then we see that constraint 4 is tight so we define $A_2^* = \begin{bmatrix} -1 & 0 \end{bmatrix}$. We see that $rank(A_2^*) = 1 < n$, so $\bar{\mathbf{x}}_2$ is not an extreme point.

3. Relaxation For an integer program we obtain its LP relaxation by removing the condition that some variables have to take integer values. This allows us to reduce many integer programming problems to linear programming. Consider two optimization problems:

$$\max \mathbf{c}^\intercal \mathbf{x} : \mathbf{x} \in P_1 \tag{1}$$

$$\max \mathbf{c}^\intercal \mathbf{x} : \mathbf{x} \in P_2 \tag{2}$$

If $P_2 \supseteq P_1$, then (2) is a relaxation of (1). If we take (1) to be an integer program, then (2) is one of its possible LP relaxations.

We then can say the following about this pair:

- if (2) is infeasible, then (1) is infeasible.

  proof: if (2) is infeasible, then $P_2 = \emptyset$ and therefore $P_1 = \emptyset$ and (1) is infeasible.

- if $\bar{\mathbf{x}}$ is an optimal solution for (2) and feasible for (1), then $\bar{\mathbf{x}}$ is also optimal for (1).

  proof: supposing $\bar{\mathbf{x}}$ is optimal for (2) and feasible for (1), then $\bar{\mathbf{x}}$ maximizes $\mathbf{c}^\intercal \forall \mathbf{x} \in P_2$ which means it also maximizes $\forall \mathbf{x} \in P_1$.

- if $\bar{\mathbf{x}}$ is an optimal solution for (2), then $c^\intercal \bar{\mathbf{x}}$ is an upper bound for (1).

  proof: since $P_2 \supseteq P_1$, the optimal value of (2) is at least as large as the optimal value of (1).

Geometric example

Let's consider the integer program $\max \mathbf{c}^\intercal \mathbf{x} : A\mathbf{x} \leq \mathbf{b}$ where

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \\ -1 & -4 \\ -1 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 7 \\ 7 \\ -4 \\ \frac{1}{2} \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and $\mathbf{x}$ is integer-valued. If we were to go about solving this problem as a linear program, we would quickly see that the optimal solution is not an integer, as shown in Figure 2 below.

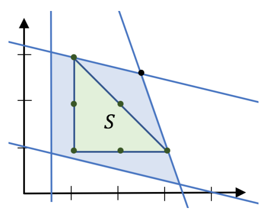

Figure 2: Relaxation of an IP (blue) to an LP (green).

We can more clearly define the set of feasible solutions by formulating an LP relaxation as $\max \mathbf{c}^\intercal \mathbf{x} : A\mathbf{x} \leq \mathbf{b}$ where

$$A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -1 \\ -1 \\ 4 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The region of space (shown in green in Figure 2) containing the set of feasible solutions for the integer program is $S = \{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}\}$. Of these, $\{\begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}\}$ is the set of optimal solutions for the integer program, which also correspond to the extreme points. Also note that every point that lies on the line between the extreme points is also an optimal solution for the LP relaxation, but by definition not an optimal solution of our original IP.