Linear Programming Duality

Vazirani Chapter 13

We just saw one method of using linear programming to design approximation algorithms:

- Write an integer LP describing the problem
- Relax the Integer LP and solve it
- *Randomly Round* the relaxed LP solution

This technique relied upon the fact that optimum value of relaxed LP was a bound on OPT. In this set of lectures we will see another general approach to designing approximation algorithms:

Linear Programming Duality

This technique

- Writes an integer LP describing the problem
- Relaxs the Integer LP
- Writes the *dual* of the relaxed LP
- Finds a feasible solution to dual LP

The technique relies upon the fact that *feasible solutions to the dual bound optimal solutions to the primal.*

Definitions and Background

Linear Programming is the problem of minimizing/maximizing a linear function subject to linear inequality constraints.

The function being optimized is called the *objective function.*

$$\begin{array}{ll} \text{minimize} & 7x_1 + x_2 + 5x_3 \\ \text{subject to} & x_1 - x_2 + 3x_3 \geq 10 \\ & 5x_1 + 2x_2 - x_3 \geq 6 \\ & x_1,\, x_2,\, x_3 \geq 0 \end{array}$$

A problem with all constraints of the form $\geq$ and objective function *minimize* is said to be in *standard form*

A solution that satisfies all of the constraints of the LP is a *feasible solution.*

We are now going to see a *dual* LP that will "help" us solve the original one.

$$\begin{array}{ll} \text{minimize} & 7x_1 + x_2 + 5x_3 \\ \text{subject to} & x_1 - x_2 + 3x_3 \geq 10 \\ & 5x_1 + 2x_2 - x_3 \geq 6 \\ & x_1,\, x_2,\, x_3 \geq 0 \end{array}$$

Let $z^* = \min 7x_1 + x_2 + 5x_3$ be the minimum value possible for a feasible solution.
For any given α we can ask "Is $z^* \leq \alpha$." A *Yes* answer has an easy proof. Just give a feasible solution with objective function value $\leq \alpha$.

Example: Is $z^* \leq 30$. The answer is yes and a "proof" is $x = (2, 1, 3)$.

How can we prove a NO answer?

Since

$$7x_1 + x_2 + 5x_3 \geq x_1 - x_2 + 3x_3 \geq 10$$

we have a proof that $z^* \not< 10$.
Even more, we can prove $z^* \not< 16$:

$$7x_1 + x_2 + 5x_3 \geq (x_1 - x_2 + 3x_3) + (5x_1 + 2x_2 - x_3) \geq 16.$$

$$\begin{array}{ll} \text{minimize} & 7x_1 + x_2 + 5x_3 \\ \text{subject to} & x_1 - x_2 + 3x_3 \geq 10 \\ & 5x_1 + 2x_2 - x_3 \geq 6 \\ & x_1,\, x_2,\, x_3 \geq 0 \end{array}$$

In fact, if $y_1, y_2 \geq 0$ have the property that

$$7x_1 + x_2 + 5x_3 \geq y_1(x_1 - x_2 + 3x_3) + y_2(5x_1 + 2x_2 - x_3)$$

then

$$z^* \geq 10y_1 + 6y_2.$$

The best lower bound is achieved by choosing the best y_1, y_2.

Another way of writing this is

$$\begin{array}{ll} \text{maximize} & 10y_1 + 6y_2 \\ \text{subject to} & y_1 + 5y_2 \leq 7 \\ & -y_1 + 2y_2 \leq 1 \\ & 3y_1 - y_2 \leq 5 \\ & y_1,\, y_2 \geq 0 \end{array}$$

This new linear program is the *dual* of the original, which is known as the *primal.*

$$
\begin{array}{ll}
\text{minimize} & 7x_1 + x_2 + 5x_3 \\
\text{subject to} & x_1 - x_2 + 3x_3 \geq 10 \\
 & 5x_1 + 2x_2 - x_3 \geq 6 \\
 & x_1,\, x_2,\, x_3 \geq 0
\end{array}
$$

$$
\begin{array}{ll}
\text{maximize} & 10y_1 + 6y_2 \\
\text{subject to} & y_1 + 5y_2 \leq 7 \\
 & -y_1 + 2y_2 \leq 1 \\
 & 3y_1 - y_2 \leq 5 \\
 & y_1,\, y_2 \geq 0
\end{array}
$$

Every feasible solution of the dual is a lower bound on the minimum value of the primal. In particular the maximum objective function achievable for the dual is $\leq$ the minimum achievable for the primal.

If there are feasible solutions for the primal and dual with the same objective function value then both are optimal solutions.

Example: $x = (7/4, 0, 11/4)$ and $y = (2, 1)$ are feasible solutions for the primal and the dual. Their respective objective functions are both 26 so they are both optimal solutions.

THIS ALWAYS HAPPENS

Given primal

$$
\boxed{
\begin{array}{lll}
\text{minimize} & \sum_{j=1}^{n} c_j x_j & \\
\text{subject to} & \sum_{j=1}^{n} a_{ij} x_j \geq b_i, & i = 1, \ldots, m \\
 & x_j \geq 0, & j = 1, \ldots, n
\end{array}
}
$$

the dual is

$$
\boxed{
\begin{array}{lll}
\text{maximize} & \sum_{i=1}^{m} b_i y_i & \\
\text{subject to} & \sum_{i=1}^{m} a_{ij} y_i \leq c_j, & j = 1, \ldots, n \\
 & y_i \geq 0, & i = 1, \ldots, m
\end{array}
}
$$

The Weak-Duality Theorem: If $x = (x_1, \ldots, x_n)$ is a feasible solution to the primal and $y = (y_1, \ldots, y_m)$ a feasible solution to the dual then

$$\sum_{j=1}^{n} c_j x_j \geq \sum_{i=1}^{m} b_i y_i.$$

The LP-Duality Theorem: The primal program has a finite optimum if and only if the dual has a finite optimum. Furthermore if $x^* = (x_1^*, \ldots, x_n^*)$ is an optimal solution for the primal and $y^* = (y_1^*, \ldots, y_m^*)$ an optimal solution for the dual then

$$\sum_{j=1}^{n} c_j x_j^* = \sum_{i=1}^{m} b_i y_i^*.$$

An LP Duality Approach to Set-Cover

We *again* return to the set cover problem from Lecture 2 with addition that every set $F \in \mathcal{F}$ now has associated $Cost(F) \geq 0$. The cost of a collection of sets is just the sum of their costs. The *Weighted Set Cover problem* is to find a Set Cover of Minimum Weight.
The generalization of the Greedy set cover algorithm we saw before is

A Greedy Set-Cover algorithm
$U \leftarrow X$
$\mathcal{C} \leftarrow \emptyset$
  select a $F \in \mathcal{F}$ that minimizes $\frac{Cost(F)}{|F \cap U|}$
  $U \leftarrow U - F$
  $\mathcal{C} \leftarrow \mathcal{C} \cup \{F\}$
  For each $e \in F \cap U$ set $price(e) = \frac{Cost(F)}{|F \cap U|}$
return($\mathcal{C}$)

The value $\frac{Cost(F)}{|F \cap U|}$ is the *cost-effectiveness* of the set. It is the average cost of adding each item in $|F \cap U|$ to the cover. $price(e)$ will contain this average value (used later in the analysis). Note that the set cover constructed has cost $\Sigma_{e \in U}\, price(e)$.

Note that if $Cost(F) = 1$ for all sets F then the algorithm always picks F that *maximizes* $|F \cap U|$, and collapses to the greedy algorithm we saw previously.

The integer LP will be

$$
\boxed{
\begin{array}{ll}
\multicolumn{2}{l}{\textbf{Minimize } \Sigma_{F \in \mathcal{F}} C(F) x_F} \\
\multicolumn{2}{l}{\text{subject to conditions}} \\
\forall e \in U, & \Sigma_{e \in F}\, x_F \geq 1 \\
\forall F \in \mathcal{F} & x_F \in \{0, 1\}
\end{array}
}
$$

The relaxation of the LP is

$$
\boxed{
\begin{array}{ll}
\multicolumn{2}{l}{\textbf{Minimize } \Sigma_{F \in \mathcal{F}} C(F) x_F} \\
\multicolumn{2}{l}{\text{subject to conditions}} \\
\forall e \in U, & \Sigma_{e \in F}\, x_F \geq 1 \\
\forall F \in \mathcal{F}, & x_F \geq 0
\end{array}
}
$$

We now introduce a variable y_e for all $e \in U$. The *dual* of the relaxed LP is then

$$
\boxed{
\begin{array}{ll}
\multicolumn{2}{l}{\textbf{Maximize } \Sigma_{e \in U}\, y_e} \\
\multicolumn{2}{l}{\text{subject to conditions}} \\
\forall F \in \mathcal{F}, & \Sigma_{e \in F}\, y_e \leq C(F) \\
\forall e \in U, & y_e \geq 0
\end{array}
}
$$

Maximize $\Sigma_{e \in U} y_e$
subject to conditions
$\forall F \in \mathcal{F}, \qquad \Sigma_{e \in F} y_e \leq C(F)$
$\forall e \in U, \qquad y_e \geq 0$

Theorem: The setting $y_e = \frac{price(e)}{H_n}$ is a feasible solution to the dual problem where $n = |U|$.

Proof: Consider some set $F \in \mathcal{F}$. Let $k = |F|$. Number the elements of F in the order in which they are covered by the algorithm as $e_1, e_2, \ldots, e_k$, breaking ties arbitrarily.
Let us examine the step of the algorithm at which item e_i is covered. Before this step starts F contains at least $k - i + 1$ uncovered elements.

Therefore, at this step F itself can cover e_i with cost-effectiveness at most $\frac{c(F)}{k-i+1}$. Since the algorithm chooses a set F' with minimal cost-effectiveness this implies $price(e_i) \leq \frac{c(F)}{k-i+1}$. Thus

$$y_{e_i} \leq \frac{1}{H_n} \cdot \frac{C(F)}{k - i + 1}$$

so

$$\sum_{i=1}^{k} y_{e_i} \leq \frac{C(F)}{H_n} \left(\frac{1}{k} + \frac{1}{k-1} + \cdots \frac{1}{1} \right) = \frac{H_k}{H_n} \cdot C(F) \leq C(F)$$

proving the theorem.

Theorem: The approximation algorithm is an H_n-approximation algorithm.

Proof:
Recall that the cost of the set cover picked by the algorithm is $\Sigma_{e \in U} \, price(e)$.

- The theorem on the previous page states that the setting $\forall e \in U$, $y_e = \frac{price(e)}{H_n}$ is a feasible solution for the dual LP. The objective function for the dual was $\Sigma_{e \in U} y_e$ which for this setting has value $w = \Sigma_{e \in U} \frac{price(e)}{H_n}$.
- From the *Weak Duality Theorem* we have that $w \leq z^*$ where z^* is optimal solution of the primal.
- $z^* \leq OPT$ by definition of LP relaxation.
- Then $w \leq z^* \leq OPT$ or $\Sigma_{e \in U} \frac{price(e)}{H_n} \leq OPT$.
- Recall that $Cost(\mathcal{C}) = \Sigma_{F \in \mathcal{C}} Cost(F) = \Sigma_{e \in U} \, price(e)$. This proves $Cost(\mathcal{C}) \leq H_n \cdot OPT$.