Primal-Dual Schema

Mikola Lysenko

May 5, 2009

Motivation So far we have been talking about linear programs and it's applications to finding approximate solutions for various integer programs. This lecture introduces a method for analyzing these types of approximations. First a quick review:

Linear Programs A linear program is given by a tuple (A, b, c) where A is an $n \times m$ matrix, b is a $n \times 1$ column vector, and c a $1 \times m$ row vector and defines the following (potentially empty) convex set:

$$\underset{x \in \Re^{m+} | Ax \leq b}{\operatorname{argmin}} \; cx$$

The dual of a linear program is defined as

$$(A, b, c)^* = (-A^T, -c^T, -b^T)$$

Which gives the set

$$\underset{y \in \Re^{n+} | -A^T y \leq -c^T}{\operatorname{argmin}} - b^T y = \underset{y \in \Re^{n+} | A^T y \geq c^T}{\operatorname{argmax}} \; b^T y$$

A solution to a program is a point within the aforementioned set. The set $\{x | Ax \leq b\}$ is known as the feasible space of a linear program. Recall that if a pair of points x, y are a solution for $(A, b, c), (A, b, c)^*$ respectively if and only if

$$\forall j \in [1, m] \; : x_j \sum_{i=1}^{n} A_{i,j} y_i \quad = \quad c_j x_j \tag{1}$$

$$\forall i \in [1, n] \; : y_i \sum_{j=1}^{n} A_{i,j} x_j \quad = \quad b_i y_i \tag{2}$$

Approximate Complementary Slackness Equations 1 and 2 are called the primal and dual complementary slackness conditions and are both necessary and sufficient conditions for optimality. We now introduce a somewhat less strict condition on a (possibly infeasible) solution pair (x, y) for (A, b, c) known as approximate slackness; given scalars $\alpha, \beta \geq 1$

$$\forall j \in [1, m] \; : \frac{c_j}{\alpha} x_j \leq \quad x_j \sum_{i=1}^{n} A_{i,j} y_i \quad \leq c_j x_j \tag{3}$$

$$\forall i \in [1, n] \; : b_i y_i \leq \quad y_i \sum_{j=1}^{n} A_{i,j} x_j \quad \leq \beta b_i y_i \tag{4}$$

Now suppose that equations 3, 4 hold for some (x, y); then we know:

$$\begin{aligned}
\sum_{j=1}^{m} \frac{c_j}{\alpha} x_j &\leq \sum_{j=1}^{m} x_j \sum_{i=1}^{n} A_{i,j} y_i \\
\sum_{j=1}^{m} x_j \sum_{i=1}^{n} A_{i,j} y_i &= \sum_{i=1}^{n} y_i \sum_{j=1}^{m} A_{i,j} x_j \\
\sum_{i=1}^{n} y_i \sum_{j=1}^{m} A_{i,j} x_j &\leq \sum_{i=1}^{n} \beta b_i y_i \\
\sum_{j=1}^{m} \frac{c_j}{\alpha} x_j &\leq \sum_{i=1}^{n} \beta b_i y_i
\end{aligned}$$

And thus we arrive at the following condition:

$$\sum_{j=1}^{m} c_j x_j \leq \alpha\beta \sum_{i=1}^{n} b_i y_i \tag{5}$$

The values α, β determine the amount of deviation from an optimality. In the case where $\alpha = \beta = 1$, then the exact complementary slackness of equations 1, 2 is satisified and we would have an optimal solution. As α, β grow larger, the bounds on the quality of the approximation gets worse.

Example: Set Cover The set cover problem can be stated as follows: given a collection of sets $\mathcal{S}$, some points $U \subset \bigcup_{S \in \mathcal{S}} S$ and a scalar cost function $c(S) : \mathcal{S} \to \Re^{+}$, find the collection of sets $X \subset \mathcal{S}$ such that $U \subseteq \bigcup_{S \in X} S$ minimizing $\sum_{S \in X} c(S)$. This problem can be concisely written as an integer program

$$\text{Primal:} \quad \underset{x \in \mathbb{Z}^{\mathcal{S}} | e \in U \to \sum_{e \in S} x_S \geq 1}{\operatorname{argmin}} c(S) x_S \qquad\qquad \text{Dual:} \quad \underset{y \in \Re^{U} | S \in \mathcal{S} \to \sum_{e \in S} y_e \leq c(S)}{\operatorname{argmax}} y_e$$

Now consider the following approximation algorithm for set-cover:

$x \leftarrow 0, y \leftarrow 0$
while $\exists y_e = 0$
  do Pick an uncovered e
    $S_{min} \leftarrow \emptyset, y_{min} \leftarrow \infty$
    for each $S \ni e$
      do $y' \leftarrow \sum_{e' \in S - \{e\}} y_{e'}$
        if $y' < y_m in$ and $y' > 0$
          then $S_{min} \leftarrow S, y_{min} \leftarrow y'$
    $x_S \leftarrow 1, y_e \leftarrow y_{min}$

This algorithm achieves an approximation factor of f, where f is the maximum number of times a given element is covered. To show this, first observe that it is obvious all elements are eventually covered, and so there are no overpacked sets (more formally, $\forall S : \sum_{e \in S} y_e \leq c(S)$). Thus, the final solution is both primal and dual feasible. Pick approximation factors $\alpha = 1, \beta = f$, thus we have a f-factor approximation.

To show this bound is tight, consider the following example. Let

$$U = \{e_0, ...e_n\}$$

$$S = \{\{e_0, e_1\}, \{e_0, e_2\}, ..., \{e_0, e_{n-1}\}\} \cup U$$

$$c(\{e_0, e_i\}) = 1, \forall i \in 1..n-1$$

$$c(U) = 1 + \epsilon, \epsilon > 0$$

The optimal solution is to pick the cover $\{U\}$ with total cost $1 + \epsilon$, while the above algorithm will end up picking all sets in S for a total cost of $n + \epsilon$.