Lecture 12

Lecturer: Alantha Newman *Scribes: Siddhartha Brahma*

1 Primal-Dual Algorithms

Last time, we used the technique of dual-fitting to analysis the approximation ratio of algorithms. Now we explore some other approaches that use both the primal and the dual linear programs.

1.1 Complementary Slackness: Full and Approximate

Recall our canonical linear programs, where $x \in \mathbb{R}^n, y \in \mathbb{R}^m, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$.

$$\begin{array}{lll} \text{Primal } (\mathcal{P}) & \min & c^T x \\ & & Ax \geq b \\ & & x \geq 0 \\ \text{Dual } (\mathcal{D}) & \max & b^T y \\ & & A^T y \leq c \\ & & y \geq 0 \end{array}$$

Also recall that *Strong Duality* ensures that if both $(\mathcal{P})$ and $(\mathcal{D})$ have finite optima, they are equal. Assume this is the case. Then, for the optima x and y of the primal and dual programs respectively, we have

$$c^T x = \sum_{i=1}^{n} c_i x_i \geq \sum_{i=1}^{n} (\sum_{j=1}^{m} A_{ji} y_j) x_i = \sum_{j=1}^{m} (\sum_{i=1}^{n} A_{ji} x_i) y_j \geq \sum_{j=1}^{m} b_j y_j = b^T y.$$

Since $c^T x = b^T y$, the first and second inequalities should hold with equality. This implies that

$$\text{Either } c_i = \sum_{j=1}^{m} A_{ji} y_j \text{ or } x_i = 0 \quad (1)$$

$$\text{Either } b_j = \sum_{i=1}^{n} A_{ji} x_i \text{ or } y_j = 0 \quad (2)$$

In other words, for the respective optima of $(\mathcal{P})$ and $(\mathcal{D})$, either the i^{th} variable in x is zero or the corresponding constraint in $(\mathcal{D})$ is tight. Similarly, either the j^{th} variable in y is zero or the corresponding constraint in $(\mathcal{P})$ is tight. These conditions are together termed as the (full) *complementary slackness* conditions.

These definitions can be extended to more general feasible solutions of the primal and dual. Suppose x is primal feasible and y is dual feasible. Let $\alpha, \beta \geq 1$. If these feasible solutions satisfy the following conditions:

$$\text{Either } x_i = 0 \text{ or } c_i \geq \sum_{j=1}^{m} A_{ji} y_j \geq \frac{c_i}{\alpha} \quad (3)$$

$$\text{Either } y_j = 0 \text{ or } b_j \leq \sum_{i=1}^{n} A_{ji} x_i \leq \beta \cdot b_j, \quad (4)$$

then they are said to satisfy the *approximate complementary slackness* criteria. It is clear that in this case the primal feasible solution gives a $(\alpha\beta)$-approximate solution to the linear program. Indeed,

$$\sum_{i=1}^{n} c_i x_i \leq \alpha\beta \sum_{j=1}^{m} b_j y_j$$

and $\sum_{j=1}^{m} b_j y_j$ is a lower bound to the primal optimum. In the following sections, we will look at two problems where the properties of complementary slackness are used to design approximation algorithms. The class of such algorithms are called *Primal-Dual* algorithms.

1.2 Set Cover via Primal-Dual

As a first illustration, we will consider the Set-Cover problem. We have already seen several algorithms for the problem. While the procedure presented below will not give us an improved approximation ratio, it will illustrate several aspects of the Primal-Dual technique. As always, we start of with the relaxed version of the IP representing the Set-Cover instance, which is given by a set of subsets $\mathcal{S} = \{S_1, S_2, \ldots, S_n\}$ of the universe of elements $U = \{e_1, e_2, \ldots, e_m\}$. The relaxed LP and its dual are as follows. We will also assume that each element of U is a member of atmost f sets.

$$\begin{aligned}
\text{Primal } (\mathcal{P}) \qquad \min \quad & \sum_{S_i \in \mathcal{S}} x_i \\
& \sum_{S_i : e_j \in S_i} x_i \geq 1 \quad \forall e_j \in U \\
& x_i \geq 0 \\
\text{Dual } (\mathcal{D}) \qquad \max \quad & \sum_{e_j \in U} y_j \\
& \sum_{e_j \in S_i} y_j \leq 1 \quad \forall S_i \in \mathcal{S} \\
& y_j \geq 0
\end{aligned}$$

Here the primal variable vector x is of length n (the number of sets) and the dual variable vector y is of length m (the number of elements in U). In the Primal-Dual schema, we will start with a primal infeasible vector $x = 0$ and a dual feasible vector $y = 0$. As long as there is some element $e_j \in U$ that is still uncovered by the primal solution, we will look at the corresponding dual variable y_j and raise its value until some dual constraint(s) becomes tight all the time maintaining dual feasibility. Now the dual constraints correspond to sets in $\mathcal{S}$. We set the corresponding primal variables to 1 (and hence include these sets in our set cover). This is continued until all the elements are covered.

Lemma 1 *The above algorithm is an f-approximation to Set Cover.*

Proof Let $I \subseteq \mathcal{S}$ be the collection of sets returned by the algorithm. Then, for each $S_i \in I$, the corresponding dual constraint is satisfied with equality. Therefore,

$$|I| = \sum_{S_i \in I} 1 = \sum_{S_i \in I} \sum_{e_j \in S_i} y_j = \sum_{e_j \in U} y_j \sum_{i \in I : e_j \in S_i} 1$$

Now clearly, by definition $|\sum_{i \in I : e_j \in S_i} 1| \leq f$. Therefore, if OPT denotes the size of the optimal set cover,

$$|I| \leq f \sum_{e_j \in U} y_j \leq f \cdot OPT$$

■

Notice that, the final x and y are feasible solutions for $\mathcal{P}$ and $\mathcal{D}$ respectively and the complementary slackness conditions are satisfied with $\alpha = 1$ and $\beta = f$.

1.3 The Metric Uncapacitated Facility Location Problem

In this problem, the input is a complete bipartite graph $G = (V, E)$. The vertex set V is partitioned into two sets: facilities (F) and clients (C). Each facility incurs a cost f_i to open. There is also a cost c_{ij} when client j uses facility i. The goal is to find a set $I \subseteq F$ of facilities such that the total cost of opening I and connecting every client to some facility in I is minimized. In addition, the costs c_{ij} satisfy a *metric* condition which takes the following form. For any two distinct facilities i, i' and two distinct clients j, j'

$$c_{i'j} \le c_{ij} + c_{ij'} + c_{i'j'}.$$

As always, we will start with the IP formulation of the problem and relax it to an LP. The relaxed LP and its dual are shown below. In the primal LP, there is a vector y of length $|F|$ associated with the facilities, which represents which facilities are open. There is also a variable x_{ij} for each pair of facility and client denoting whether client j is connected to facility i.

$$\begin{array}{lll} \text{Primal } (\mathcal{P}) & \min & \displaystyle\sum_{i \in F} f_i \cdot y_i + \sum_{i \in F, j \in C} c_{ij} \cdot x_{ij} \\ & & \displaystyle\sum_{i \in F} x_{ij} \ge 1 \quad \forall j \in C \\ & & y_i - x_{ij} \ge 0 \quad \forall i \in F, j \in C \\ & & y_i, x_{ij} \in [0,1] \quad \forall i \in F, j \in C \end{array}$$

The first set of constraints represents the fact that each client is connected to at least one facility. The second set of constraints enforces the requirement that a client can only be connected to an open facility. The dual LP is the following. There is a set of variables for each client α_j and a variable for each client-facility pair, β_{ij}.

$$\begin{array}{lll} \text{Dual } (\mathcal{D}) & \max & \displaystyle\sum_{j \in C} \alpha_j \\ & & \displaystyle\sum_{j \in C} \beta_{ij} \le f_i \quad \forall i \in F \\ & & \alpha_j - \beta_{ij} \le c_{ij} \quad \forall i \in F, j \in C \\ & & \alpha_j, \beta_{ij} \ge 0 \quad \forall i \in F, j \in C \end{array}$$

The Primal-Dual algorithm has two phases.

Phase I
Initially, we again start with a dual feasible solution $\{\alpha_j = 0, \beta_{ij} = 0\}$. It is beneficial to imagine the algorithm running in unit time steps starting from zero. Let t denote the current time. Raise all the α_j's uniformly (by the same amount t) until we reach a point such that $\alpha_j = c_{ij}$ for some i, j. All such edges are called *tight edges*. Now, if this happens we still have the option of raising β_{ij}'s, also at the same rate such that $\alpha_j - \beta_{ij}$ remains equal to c_{ij}. We continue doing this until we hit a constraint $\sum_{j \in C} \beta_{ij} = f_i$. It is at this point we *temporarily open* the facility i. For all the edges (i, j) such that $\alpha_j = \beta_{ij} + c_{ij}$, we connect client j to facility i. Facility i is called a *connecting witness* for all such clients j. Each edge (i, j) for which $\beta_{ij} > 0$ is called a *special* edge. Note that special edges remain tight edges and for such

edges, $\alpha_j = \beta_{ij} + c_{ij}$. Note that once a client j is connected, the value of the corresponding α_j is no longer increased.

Assume that after one round of the above procedure, the set of temporarily open facilities is $L = \{i_1, i_2, \ldots, i_k\}$. All the clients that get connected to facilities in L are marked as "connected". We continue the procedure with the remaining unconnected clients. Now that L is not a null set, it may happen that for a client j_1, the edge (i, j_1) for some $i \in L$ becomes tight. In this case, j_1 is declared connected and i is called the *connecting witness* for j_1. Note that in that case, $\beta_{ij_1} = 0$. The set L does not change. The other possibility is that a new facility i_{k+1} becomes open, in which case it is added to L and all the clients corresponding to tight edges with i_{k+1} are marked connected. This procedure is continued till all the clients are marked connected.

To get some intuition, we can think of α_j to be the amount client j is paying towards the total cost of the solution. The β_{ij}'s can be thought of as the amount client j is contributing towards opening facility i. Thus, for the tight and special edges, α_j is split between the cost c_{ij} of connecting to a facility i and the contribution β_{ij} to its opening. Further, the fact that $\sum_{j \in C} \beta_{ij} = f_i$ for the temporarily open facilities ensures that the cost f_i of opening a facility has been fully paid for by all the clients that connect to it.

Phase II

After the end of Phase I, let F_t be the set of temporarily open facilities. Our goal is to permanently open a subset of F_t. To this end, consider the subgraph of T of G where there is an edge between client j and facility i iff $\beta_{ij} > 0$. Let T^2 be the graph such that contains edges (u, v) iff the distance between u and v in T is at most 2. Finally, let H be the subgraph of T^2 induced by F_t. The final step of the algorithm is to pick a *maximal independent set* in H called I and make those facilities *permanently open*.

It remains to specify which facility each client gets connected to. For a client j, let $\phi(j)$ denote the facility in I to which it is assigned. Also, for client j define F_j as

$$F_j = \{i \in F_t |\ \text{edge } (i, j) \text{ special, i.e. } \beta_{ij} > 0\}.$$

The following cases may arise for each client j.

1. $I \cap F_j \neq \emptyset$. Since I is an independent set, it follows in this case that $|I \cap F_j| = 1$. Set $\phi(j) = i$, where $i \in I \cap F_j$. Call this client *directly connected*. Clearly, in this case edge (i, j) is special and $\beta_{ij} > 0$.

2. $I \cap F_j = \emptyset$. In this case, if the connecting witness for client j is $i' \in I$, set $\phi(j) = i'$. Call this client directly connected, too. Notice that, in this case the edge (i', j) is tight and $\beta_{i'j} = 0$, since $i' \notin F_j$.

3. Finally, there may be clients such that $I \cap F_j = \emptyset$ and the connecting witness for client j is $i' \notin I$, then let i be any neighbor of i' in H such that $i \in I$. Set $\phi(j) = i$. Call the client j *indirectly connected*.

1.3.1 Analysis of the algorithm

In the analysis, we will show how the dual solution α_j accounts for the costs of opening and connecting clients to facilities. Let

$$\alpha_j = \alpha_j^f(\text{contribution for opening facility}) + \alpha_j^c(\text{contribution for connection cost}).$$

If client j is directly connected, then $\alpha_j = \beta_{ij} + c_{ij}$ where $i = \phi(j)$. In this case, we set $\alpha_j^f = \beta_{ij}$ and $\alpha_j^c = c_{ij}$. If the client is indirectly connected, then set $\alpha_j^f = 0$ and $\alpha_j^c = \alpha_j$. We have the following lemma.

Lemma 2 *For each $i \in I$, $\sum_{j:\phi(j)=i} \alpha_j^f = f_i$.*

Proof Since $i \in I$ is temporarily open at the end of Phase I, we have

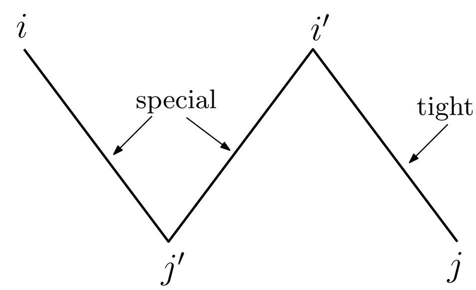
$$\sum_{j:(i,j) \text{ is special}} \beta_{ij} = f_i$$

Notice that each client j, who has contributed to f_i in the above sum, is directly connected and hence $\alpha_j^f = \beta_{ij}$ for such clients. The other clients connected directly through tight edges have $\alpha_j^f = \beta_{ij} = 0$. Finally, the indirectly connected clients have again $\alpha_j^f = 0$ by definition. ■

Intuitively speaking, this lemma shows that the cost of opening the facilities is completely accounted for by the α_j^f's. It remains to bound the cost incurred by the α_j^c's. We have the following lemma.

Lemma 3 *For an indirectly connected client j and facility $i = \phi(j)$, $c_{ij} \leq 3\alpha_j^c$.*

Proof Let i' be the connecting witness for city j. Since j is indirectly connected to i, (i, i') must be an edge in H. Since H was constructed by taking connecting vertices in F at a distance of at most 2 in T, there must be a client j' such that (i, j') and (i', j') are both special edges (recall T consisted only of special edges). Further, let t_1 and t_2 be the two time instances when i and i' were declared temporarily open. Since edge (i', j) is tight, $\alpha_j \geq c_{i'j}$. Since edges (i, j') and (i', j') are also special edges, $\alpha_{j'} \geq c_{ij'}$



and $\alpha_{j'} \geq c_{i'j'}$. Clearly, both these edges must have become special at a time before either i or i' was temporarily opened in Phase I. Further, $\alpha_{j'}$ would have stopped increasing after the opening of either i or i'. Therefore, $\alpha_{j'} \leq \min(t_1, t_2)$. Finally, since i' is the connecting witness of j, $\alpha_j \geq t_2$. Thus we have following inequalities:

$$\alpha_j \geq c_{i'j}$$

$$\alpha_j \geq t_2 \geq \min(t_1, t_2) \geq \alpha_{j'} \geq c_{ij'}, c_{i'j'}.$$

Combining them, and using the fact that $c_{ij} \leq c_{i'j} + c_{i'j'} + c_{ij'}$ (by the metric assumption), we can conclude that $c_{ij} \leq 3\alpha_j^c$ (recall for indirectly connected client, $\alpha_j^c = \alpha_j$). ■

The final result, showing a 3-approximation guarantee follows easily.

Theorem 4 *The above algorithm gives a 3-approximation to the metric uncapacitated facility location problem.*

Proof The total cost of the solution is given by

$$PC = \sum_{j \in C} c_{\phi(j),j} + \sum_{i \in I} f_i$$

The cost of the dual solution, which is a lower bound to PC is

$$DC = \sum_{j \in C} \alpha_j = \sum_{j \in C} \alpha_j^c + \sum_{j \in C} \alpha_j^f$$

The cost of the dual solution is a lower bound to the primal optimal. But we know that $\sum_{j:\phi(j)=i} \alpha_j^f = f_i$ for each $i \in I$. Therefore, the second terms in the above expressions are equal. As for the first terms, we can divide the clients into the directly and indirectly connected sets. For the directly connected set, $c_{\phi(j),i} = \sum_j \alpha_j^c$. For the indirectly connected set, the above proved lemma shows that $c_{\phi(j),i} \leq 3\alpha_j^c$. Combining everything, we have

$$PC = \sum_{j \in C} c_{\phi(j),i} + \sum_{i \in I} f_i \leq 3 \sum_{j \in C} \alpha_j = 3 \cdot DC.$$

■

Finally, we note that for all for all dual constraints of the form $\alpha_j - \beta_{ij} \leq c_{\phi(j),j}$ that are not tight, we have shown:

$$\begin{aligned} \frac{1}{3} \cdot c_{\phi(j),j} &\leq \alpha_j \quad \Rightarrow \\ \frac{1}{3} \cdot c_{\phi(j),j} &\leq \alpha_j - \beta_{ij}, \end{aligned}$$

since $\beta_{ij} = 0$ for these constraints. And for all tight constraints, we have:

$$\alpha_j - \beta_{ij} = c_{\phi(j),j}.$$

Thus, we have shown that the approximate complementary slackness conditions (3) holds for $\alpha = 3$, which directly implies a factor 3 approximation.