Approximation Algorithms Lecture 19

Contents

1 Recap

Lower bounds on approximation guarantees, f-approximation for set cover.

2 Content

2.1 Primal-dual algorithms

We'll look at some algorithms that use the primal-dual method.

2.2 Steiner forest problem

The steiner tree problem says this: find a minimum cost connected subgraph which includes all terminals (T).

The steiner forest problem says this: Given k pairs of terminals (s_i, t_i), find a minimum cost subgraph in which each pair s_i, t_i is connected.

The Steiner forest problem is a generalization of Steiner tree: if I have to find a Steiner tree over a set of terminals then this can be formulated as a Steiner forest problem by defining pairs $(t_1, t), \ldots, (t_k, t)$ where $\{t, t_1, \ldots, t_k\} = T$.

Let's formulate this as an integer program.

Integer program for Steiner forest

For edge $e \in E$, let x_e be defined as 1 if e is picked in the Steiner forest, and 0 otherwise.

The cost of the Steiner forest becomes $\sum_{e \in E} c_e \cdot x_e$, where c_e is the cost of edge e.

Now we need to capture the connectivity relations. Consider a set S which has s such that $\overline{S}$ has t. Then for s being connected to t, it is necessary that at least one of the edges between S and $\overline{S}$ should be in the Steiner forest. Call these edges $\delta(s)$.

Our condition becomes $\sum_{e \in \delta(S)} x_e \geq 1$, $\forall S : \exists i \ni (s_i, t_i)$ is separated by S.

This is clearly necessary. Suppose E' be the edges which have $x_e = 1$. Suppose E' does not form a Steiner forest. There exists i such that there is no path from s_i to t_i using edegs of E'. Consider the set of vertices S which are reachable from s_i using edges of E'. Then clearly, $\overline{S}$ has t_i, else there is a path from s_i to t_i. So S separates s_i and t_i. Also note that no edge of $\delta(S)$ has $x_e = 1$ (since we could have added another vertex into this set S). So this implies that $\sum_{e \in \delta(S)} x_e = 0$, and this is a violation of constraint, which is a contradiction.

Let $r : 2^V \to \{0, 1\}$ be a requirement function on sets where $r(s) = 1$ iff S separates an (s_i, t_i).

Now we will look at a LP-relaxation.

LP relaxation of IP

$$\min \sum_{e \in E} c_e x_e$$

subject to the constraints

$$\forall S \subseteq V, \sum_{e \in \delta(S)} x_e \geq r(S)$$

$$0 \leq x_e \leq 1$$

The $x_e \leq 1$ constraint is redundant, so we can drop this constraint.

Of course, this LP has exponentially many equations, so we won't actually solve this (but use this as a conceptual tool).

Consider the dual of this LP. Then we have the dual program as:

$$\max \sum_{S \in 2^V} r(S) \cdot y_S$$

subject to the constraints

$$\forall e \in E, \sum_{S: e \in \delta(S)} y_S \leq c_e$$

$$0 \leq y_S$$

Remember that the dual is written so that y_S should be such that for each variable x_e for the primal LP, coefficient of x_e in the linear combination should not exceed its coefficient in the objective function, i.e., c_e.

These constraints are called **packing constraints**. Let's consider the Steiner forest problem when we have only one pair (s_1, t_1). This is exact the shortest path problem.

In this setting, the primal LP would be the same. We can look at how Dijkstra works on this algorithm, and we'll eventually try to assign y_S using consecutive sets.