LP-based approximation algorithms
Exercises and extensions

Vincenzo Bonifaci

April 3, 2012

Contents

1 Integrality of optimal Min-Cut LP solutions

Problem 1.1. Prove that an optimal solution of the Min-Cut linear program is without loss of generality an integral solution. Hint: use the dual program and the complementary slackness conditions.

Recall the Min-Cut LP for a graph $G = (V, A)$ with capacities $(c_a)_{a \in A}$:

$$\begin{aligned} \min \quad & \sum_{(i,j) \in A} c_{ij} \cdot d_{ij} \\ \text{s.t.} \quad & d_{ij} \geq p_i - p_j && \forall (i,j) \in A \\ & p_s - p_t \geq 1 \\ & d_{ij} \geq 0 && \forall (i,j) \in A \\ & p_i \geq 0 && \forall i \in V. \end{aligned} \tag{1}$$

Its dual is the Maximum Flow LP:

$$\begin{aligned} \max \quad & f_{ts} \\ \text{s.t.} \quad & f_{ij} \leq c_{ij} && \forall (i,j) \in A \\ & f(\delta^-(i)) - f(\delta^+(i)) \leq 0 && \forall i \in V \\ & f_{ij} \geq 0 && \forall (i,j) \in A. \end{aligned} \tag{2}$$

We have used the shorthands $f(\delta^-(i))$ and $f(\delta^+(i))$ for the flow entering and leaving node i, respectively: $f(\delta^-(i)) = \sum_{j:(j,i) \in A} f_{ji}$, $f(\delta^+(i)) = \sum_{j:(i,j) \in A} f_{ij}$.

Let f be an optimal solution of the dual (maximum flow) linear program. We know that a feasible solution (d, p) of the Min-Cut LP is optimal if and only if it also satisfies the complementary slackness conditions:

$$(d_{ij} > 0) \Rightarrow (f_{ij} = c_{ij}) \tag{3}$$

$$(f_{ij} > 0) \Rightarrow (d_{ij} = p_i - p_j) \tag{4}$$

$$(p_s - p_t > 1) \Rightarrow (f_{ts} = 0) \tag{5}$$

$$(f(\delta^-(i)) - f(\delta^+(i)) < 0) \Rightarrow (p_i = 0) \tag{6}$$

Consider the "residual graph" of f, which is obtained from G by only considering the arcs (i, j) such that $f_{ij} < c_{ij}$, plus the reverse arcs (j, i) such that $f_{ij} > 0$.

If there is a path from s to t in the residual graph of f, then by slightly increasing the flow along this path we get a new feasible flow of larger value. But this is impossible since f is optimal. Therefore, there t is not reachable from s in the residual graph of f.

We now define a solution (d, p) for the Min-Cut LP. Let X be the set of nodes reachable from the node s in the residual graph of f. So $s \in X$, $t \notin X$. Now define $p_i = 1$ if $i \in X$, $p_i = 0$ if $i \notin X$. Moreover, define $d_{ij} = 1$ if $i \in X$, $j \notin X$, and $d_{ij} = 0$ otherwise.

It is not difficult to check that all the constraints of the Min-Cut LP are satisfied by (d, p). Therefore it is a feasible solution. Now consider the complementary slackness conditions:

- (3) is satisfied since if $i \in X$, $j \notin X$ then $f_{ij} = c_{ij}$, and for all other arcs (i, j), $d_{ij} = 0$ (all arcs from X to $\bar{X}$ are saturated by the flow);
- (4) is satisfied since if $i \notin X$, $j \in X$ then $f_{ij} = 0$, and for all other arcs (i, j), $d_{ij} = p_i - p_j$ (all reverse arcs from $\bar{X}$ to X have no flow);
- (5) is satisfied simply because $p_s - p_t = 1$;
- (6) is satisfied simply because $f(\delta^-(i)) = f(\delta^+(i))$ for all $i \in V$ (remember that we have added an arc (t, s) to ensure flow conservation also at s and t).

Therefore, (d, p) is a 0/1 optimal solution.

2 Existence of an integrality gap for Set Cover

Problem 2.1. Show an example where a fractional set cover is better than an integral set cover.

Recall the Set Cover ILP:

$$\begin{aligned} \min \quad & \sum_{S \in \mathcal{S}} c(S) \cdot x_S \\ & \sum_{S: e \in S} x_S \geq 1 \quad && \forall\, e \in U \\ & x_S \in \{0,1\} && \forall\, S \in \mathcal{S}. \end{aligned} \tag{7}$$

The LP relaxation is the following:

$$\begin{aligned} \min \quad & \sum_{S \in \mathcal{S}} c(S) \cdot x_S \\ & \sum_{S: e \in S} x_S \geq 1 \quad && \forall\, e \in U \\ & x_S \geq 0 && \forall\, S \in \mathcal{S}. \end{aligned} \tag{8}$$

Solution

Consider the following instance: $U = \{a, b, c\}$, $\mathcal{S} = \{S_1, S_2, S_3\}$, $S_1 = \{a, b\}$, $S_2 = \{b, c\}$, $S_3 = \{a, c\}$, $c(S_1) = c(S_2) = c(S_3) = 1$. An integral set cover has cost at least 2. On the other hand, if we set $x_{S_1} = x_{S_2} = x_{S_3} = 1/2$ we get a feasible solution to the LP, of cost 3/2. So the integrality gap is at least 4/3.

The example can be extended to arbitrarily large instances (how?).

3 Lower bound on Greedy for Set Cover

Problem 3.1. Find an example where Greedy is $\Omega(\log n)$-approximate for unweighted Set Cover.

(Recall that n denotes the size of the universe set and that in the unweighted case the cost of every set is 1.)

Solution

Consider the following construction. We have $U = \{0, 1, \ldots, 3 \cdot 2^k - 1\}$ (so $n = \Theta(2^k)$). In the collection $\mathcal{S}$ there are three sets $B_1 = \{0, \ldots, 2^k - 1\}$, $B_2 = \{2^k, \ldots, 2 \cdot 2^k - 1\}$, $B_3 = \{2 \cdot 2^k, \ldots, 3 \cdot 2^k - 1\}$. Furthermore, $\mathcal{S}$ contains also $k + 1$ sets $S_0, \ldots, S_k$ where

$$S_0 = \{0, 2^k, 2 \cdot 2^k\}$$

and, for $i \in [1, k]$,

$$S_i = \{e \in U : (e \bmod 2^k) \in [2^{i-1}, 2^i)\}.$$

See Figure 1 for an illustration when $k = 3$. The sets B_1, B_2, B_3 are the black sets, the sets $S_0, \ldots, S_k$ are the red sets.

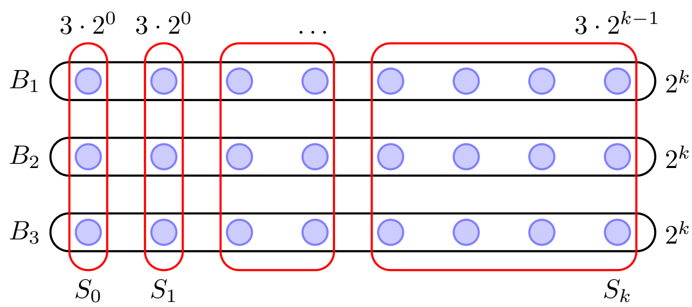

Figure 1: Counterexample for the Greedy Set Cover algorithm

The solution that picks B_1, B_2, B_3 has cost 3, and opt = 3. Since for all i, $3 \cdot 2^{i-1} > 2^i$, at each step the Greedy algorithm will select a red set and there will be $k+1$ steps. So the cost of the greedy solution is $k+1 = \Omega(\log n)$, and the approximation ratio is $\Omega(\log n)/3 = \Omega(\log n)$.