Handout 1: Approximating minimum vertex cover and set cover

Uriel Feige

14 Nov 2018

The course has a large overlap with the book [WS11]: Williamson and Shmoys, The Design of Approximation Algorithms, 2011.

Several approximation algorithms for min-VC were presented in class, all with approximation ratio 2. They include a reduction to the bipartite case, a local ratio approach, and use of LP relaxations of the associated integer program. For rounding of the LP relaxations, we saw threshold rounding, use of complementary slackness conditions to round the solution of the dual, a primal-dual algorithm and its relation to local ratio. For bipartite graphs we showed that there are integer solutions that attain the minimum value of the LP relaxation.

For mininimum set cover, the vertex cover approximation algorithms extend to give an approximation ratio of r, where r is the maximum number of sets that cover a single item. Alternatively, there are algorithms giving an approximation ratio of roughly $\ln n$. We showed the greedy algorithm and two ways of analysing it, one of which uses LP duality. In addition, we showed randomized rounding of LPs.

We also discussed how to design negative instances that show that the analysis of the algorithms presented cannot be improved. Set cover instances that exhibit the worst behavior for the greedy algorithm are in fact instances of vertex cover.

For every $\epsilon > 0$, it is NP-hard to approximate minimum vertex cover within a ratio better than $\sqrt{2} - \epsilon$, and to approximate minimum set cover within a ratio better than $(1 - \epsilon) \ln n$.

Conventions for homework. If you solve the questions by yourself – great! If you need to consult a reference (book, web site, etc.), thats fine, but please cite the references in the appropriate place in your answer. Likewise, if you need to consult a person (e.g., another student), please give credit to that person in the appropriate place in your answer.

Homework – hand in by November 28. (Grader: Yael Hitron.)

1. (Partly overlaps Problem 1.5 in [WS11].) Recall the LP relaxation for the minimum vertex cover problem: *minimize* $\sum_{i \in V} c_i x_i$ subject to the constraints $x_i + x_j \geq 1$ for every $(i, j) \in E$, and $x_i \geq 0$ for every $i \in V$.

Recall also the reduction described in class from vertex cover in an arbitrary graph $G(V,E)$ to vertex cover in a bipartite graph G', by doubling every vertex of V and every edge of E.

(a) Show that for every graph G, by applying the reduction to G' described above, one gets a graph G' for which the optimal value of the LP relaxation is exactly twice as large as the optimal value of the LP relaxation for G.

(b) Prove that for every graph G, the optimal value of the LP relaxation is attained by at least one solution in which all variables have half-integral values (namely, $x_i \in \{0, \frac{1}{2}, 1\}$ for all $i \in V$).

(c) Show that for every half-integral optimal solution of the LP relaxation, there is an optimal integral solution (to the integral problem) that includes all vertices $i \in V$ for which $x_i = 1$, and none of the vertices $i \in V$ for which $x_i = 0$.

2. Consider the minimum vertex cover problem in r-uniform hypergraphs $H(V,E)$, where each hyperedge in E contains r vertices (the case $r = 2$ corresponds to graphs), and vertices have positive costs. Consider the LP-relaxation *minimize* $\sum_{i \in V} c_i x_i$ subject to the constraints $\sum_{i \in e} x_i \geq 1$ for every hyperedge $e \in E$, and $x_i \geq 0$ for every $i \in V$.

   (a) Suppose that in an optimal solution to the LP, $x_i > 0$ for every i. Show that in this case the value of the LP is precisely $\frac{1}{r} \sum_{i \in V} c_i$. (Hint: to show that the value is at least $\frac{1}{r} \sum_{i \in V} c_i$, use the complementary slackness conditions.)

   (b) We say that an r-uniform hypergraph is *r-colored* if we are given a coloring of its vertices by r colors, and under this coloring no hyperedge in monochromatic. Design a polynomial time algorithm that for r-colored hypergraphs approximates minimum vertex cover within a ratio of $r - 1$.

3. (Problem 1.4 in [WS11].) In the *uncapacitated facility location* problem, we have a set D of clients and a set F of facilities. For each client $j \in D$ and facility $i \in F$, there is a cost c_{ij} of assigning client j to facility i. For each facility $i \in F$ there is a cost f_i for opening it. The goal of the problem is to open a subset of facilities $S \subset F$ so as to minimize the sum of openning costs of the facilities in S and the costs of assigning each client $j \in D$ to the nearest open facility. In other words, we wish to find S that minimizes $\sum_{i \in S} f_i + \sum_{j \in D} min_{i \in S} c_{ij}$.

   (a) Show that there exists some $c > 0$ such that there is no $(c \ln |D|)$-approximation algorithm for the uncapacitated facility location problem, unless P=NP.

   (b) Give an $O(\log |D|)$-approximation algorithm for the uncapacitated facility location problem.