Notes for Lecture 6

This lecture is based on Vazirani's book [1], chapters 2, 13, 14.

Outline

1. **Set Cover – problem definiton:**

   Universe U of n elements, subsets $\mathcal{S} = \{S_1, \ldots, S_k\}$ with costs $c(S_i) \geq 0$.

   Examples: (1) $U = a, b, c$ and $S_1 = \{a, b\}$, $S_2 = \{b, c\}$, $S_3 = \{a, c\}$, of unit cost. (2) in a museum, cover exhibits with guard positions (line of sight) (3) in a graph – cover edges with vertices or vertices with edges.

   Vertex cover as a special case – each element can be covered by exactly two sets.

2. **Linear programming relaxation:**

$$\min \sum_{S \in \mathcal{S}} c(S) x_S$$

$$\text{s.t.} \sum_{S : e \in S} x_S \geq 1, \quad e \in U$$

   Replacing $x_S \in \{0, 1\}$ with $x_S \geq 0$.

   Can be viewed as a *fractional set cover*.

   Example: In ex. (1) above, can choose $x_S = 1/2$ with total cost $3/2$.

3. **Relaxation provides a lower bound:**

   Lemma 1: The LP above can be solved in polynomial time.

   Lemma 2: LP $\leq$ OPT.

   Importance of lower bound: Proving ALG $\leq \rho \cdot$ LP will imply ALG $\leq \rho \cdot$ OPT.

4. **Threshold rounding for vertex-cover:**

   Algorithm: Let $\hat{x}$ be an optimal LP solution. Set $\bar{x}_S = 1$ if $\hat{x}_S \geq 1/2$, and $\bar{x}_S = 0$ otherwise.

   Theorem 3: The algorithm computes a set cover of cost within factor 2 of optimal.

5. **Proof:**

   (1) $\bar{x}$ covers all elements, because for every element at least one of the two sets that can cover it must have $x_S \geq 1/2$.

   (2) ALG $\leq$ 2LP because $c(\bar{x}) \leq 2c(\hat{x})$.

   Exercise 1: Show an instance of vertex-cover where OPT $\geq$ 1.99LP.

6. **Randomized rounding:**

   Algorithm: Let $\hat{x}$ be an optimal LP solution. Set $\bar{x}_S = 1$ independently with probability $\hat{x}_S$, and $\bar{x}_S = 1$ otherwise. Repeat $t = O(\log n)$ times independently and take all sets $S \in \mathcal{S}$ that were picked at least once.

   Theorem 4: With probability at least 1/2, the algorithm's output is a set cover of cost within factor $O(\log n)$ of optimal.

7. **Claim:** For every element $e \in U$, Pr[a single iteration does not cover e] $\leq 1/e$.

   Proof of claim: By elementary calculus, this probability is maximized when all relevant $\hat{x}_S$ are equal, and in turn is smaller than $1/e$.

8. **Proof:**

   (1) Pr[e is not covered in any iteration $\leq (1/e)^t \leq 1/4n$, and by a union bound, all n elements are covered with probability at least 3/4.

   (2) The expected cost of one iteration is LP (by linearity of expectation). Thus, $\mathbb{E}[\text{ALG}] \leq t\text{LP}$. Using Markov's inequality, $\Pr[\text{ALG} \geq 4t \cdot \text{LP}] \leq 1/4$.

9. **Exercises and Comments:**

   Exercise 2: Give alternate analysis by one rounding iteration where probabilites are increased by an $O(\log n)$ factor. (Hint: Chernoff)

   Exercise 3: Suppose that every element has to be covered at least $5 \log n$ times, and every set can be picked any number of times. Show that a similar algorithm achieves constant approximation ratio.

   Exercise 4: Show an instance where $\text{OPT} = \Theta(\log n)\text{LP}$.

   Note: A greedy algorithm gives similar results but with better leading constant.

References

[1] Vijay Vazirani. *Approximation Algorithms.* Springer, 2001. 1