NEW CS 473: Theory II, Fall 2015

Approximation Algorithms using Linear Programming

Lecture 21

November 10, 2015

Weighted vertex cover

Weighted Vertex Cover problem

$G = (V, E)$.

Each vertex $v \in V$: cost c_v.

Compute a vertex cover of minimum cost.

1. vertex cover: subset of vertices V so each edge is covered.
2. **NP-Hard**
3. ...unweighted **Vertex Cover** problem.
4. ... write as an integer program (IP):
5. $\forall v \in V$: $x_v = 1 \iff v$ in the vertex cover.
6. $\forall vu \in E$: covered. $\implies x_v \vee x_u$ **true**. $\implies x_v + x_u \geq 1$.
7. minimize total cost: $\min \sum_{v \in V} x_v c_v$.

Weighted vertex cover

State as IP $\implies$ Relax $\implies$ LP

$$\begin{aligned} \min \quad & \sum_{v \in V} c_v x_v, \\ \text{such that} \quad & x_v \in \{0,1\} && \forall v \in V \qquad (1) \\ & x_v + x_u \geq 1 && \forall vu \in E. \end{aligned}$$

1. ... **NP-Hard**.
2. relax the integer program.
3. allow x_v get values $\in [0, 1]$.
4. $x_v \in \{0,1\}$ replaced by $0 \leq x_v \leq 1$. The resulting LP is

$$\begin{aligned} \min \quad & \sum_{v \in V} c_v x_v, \\ \text{s.t.} \quad & 0 \leq x_v && \forall v \in V, \\ & x_v \leq 1 && \forall v \in V, \\ & x_v + x_u \geq 1 && \forall vu \in E. \end{aligned}$$

Weighted vertex cover – rounding the LP

1. Optimal solution to this LP: $\widehat{x}_v$ value of var X_v, $\forall v \in V$.
2. optimal value of LP solution is $\widehat{\alpha} = \sum_{v \in V} c_v \widehat{x}_v$.
3. optimal integer solution: x_v^I, $\forall v \in V$ and α^I.
4. **Any valid solution to IP is valid solution for LP!**
5. $\widehat{\alpha} \leq \alpha^I$.
   Integral solution not better than LP.
6. Got fractional solution (i.e., values of $\widehat{x}_v$).
7. Fractional solution is better than the optimal cost.
8. Q: How to turn fractional solution into a (valid!) integer solution?
9. Using **rounding**.

How to round?

1. consider vertex $\mathbf{v}$ and fractional value $\widehat{x}_v$.
2. If $\widehat{x}_v = 1$ then include in solution!
3. If $\widehat{x}_v = 0$ then do not not include in solution.
4. if $\widehat{x}_v = 0.9 \implies$ LP considers $\mathbf{v}$ as being 0.9 useful.
5. The LP puts its money where its belief is...
6. ...$\widehat{\alpha}$ value is a function of this "belief" generated by the LP.
7. Big idea: Trust LP values as guidance to usefulness of vertices.
8. Pick all vertices $\geq$ threshold of usefulness according to LP.
9. $S = \left\{ v \;\middle|\; \widehat{x}_v \geq 1/2 \right\}$.
10. **Claim**: S a valid vertex cover, and cost is low.
11. Indeed, edge cover as: $\forall vu \in E$ have $\widehat{x}_v + \widehat{x}_u \geq 1$.
12. $\widehat{x}_v, \widehat{x}_u \in (0,1) \implies \widehat{x}_v \geq 1/2$ or $\widehat{x}_u \geq 1/2$.
 $\implies v \in S$ or $u \in S$ (or both).
 $\implies S$ covers all the edges of $\mathbf{G}$.

Cost of solution

Cost of S:

$$c_S = \sum_{v \in S} c_v = \sum_{v \in S} 1 \cdot c_v \leq \sum_{v \in S} 2\widehat{x}_v \cdot c_v \leq 2 \sum_{v \in V} \widehat{x}_v c_v = 2\widehat{\alpha} \leq 2\alpha',$$

since $\widehat{x}_v \geq 1/2$ as $v \in S$.
α' is cost of the optimal solution $\implies$

Theorem

The **Weighted Vertex Cover** *problem can be* 2*-approximated by solving a single* LP*. Assuming computing the* LP *takes polynomial time, the resulting approximation algorithm takes polynomial time.*

The lessons we can take away

Or not - boring, boring, boring.

1. Weighted vertex cover is simple, but resulting approximation algorithm is non-trivial.
2. Not aware of any other 2-approximation algorithm does not use LP. (For the weighted case!)
3. Solving a **relaxation** of an optimization problem into a LP provides us with insight.
4. But... have to be creative in the rounding.

Revisiting **Set Cover**

1. Purpose: See new technique for an approximation algorithm.
2. Not better than greedy algorithm already seen $O(\log n)$ approximation.

Set Cover

Instance: $(S, \mathcal{F})$
 S: set of n elements
 $\mathcal{F}$: family of subsets of S, s.t. $\bigcup_{X \in \mathcal{F}} X = S$.
Question: The set $\mathcal{X} \subseteq \mathcal{F}$ such that $\mathcal{X}$ contains as few sets as possible, and $\mathcal{X}$ covers S.

Set Cover – IP & LP

$$\begin{aligned} \min \quad & \alpha = \sum_{U \in \mathcal{F}} x_U, \\ \text{s.t.} \quad & x_U \in \{0,1\} && \forall U \in \mathcal{F}, \\ & \sum_{U \in \mathcal{F}, s \in U} x_U \geq 1 && \forall s \in S. \end{aligned}$$

Next, we relax this IP into the following LP.

$$\begin{aligned} \min \quad & \alpha = \sum_{U \in \mathcal{F}} x_U, \\ & 0 \leq x_U \leq 1 && \forall U \in \mathcal{F}, \\ & \sum_{U \in \mathcal{F}, s \in U} x_U \geq 1 && \forall s \in S. \end{aligned}$$

Set Cover – IP & LP

1. LP solution: $\forall U \in \mathcal{F}$, $\widehat{x_U}$, and $\widehat{\alpha}$.
2. Opt IP solution: $\forall U \in \mathcal{F}$, x_U^I, and α^I.
3. Use LP solution to guide in rounding process.
4. If $\widehat{x_U}$ is close to 1 then pick U to cover.
5. If $\widehat{x_U}$ close to 0 do not.
6. Idea: Pick $U \in \mathcal{F}$: randomly choose U with **probability** $\widehat{x_U}$.
7. Resulting family of sets $\mathcal{G}$.
8. Z_S: indicator variable. 1 if $S \in \mathcal{G}$.
9. Cost of $\mathcal{G}$ is $\sum_{S \in \mathcal{F}} Z_S$, and the expected cost is $\mathbf{E}\left[\text{cost of } \mathcal{G}\right] = \mathbf{E}\left[\sum_{S \in \mathcal{F}} Z_S\right] = \sum_{S \in \mathcal{F}} \mathbf{E}\left[Z_S\right] = \sum_{S \in \mathcal{F}} \mathbf{Pr}\left[S \in \mathcal{G}\right] = \sum_{S \in \mathcal{F}} \widehat{x_S} = \widehat{\alpha} \leq \alpha^I$.
10. In expectation, $\mathcal{G}$ is not too expensive.
11. Bigus problumos: $\mathcal{G}$ might fail to cover some element $s \in S$.

Set Cover – Rounding continued

1. **Sol**: Repeat rounding $m = 10\lceil \lg n \rceil = O(\log n)$ times.
2. $n = |S|$.
3. $\mathcal{G}_i$: random cover computed in ith iteration.
4. $\mathcal{H} = \cup_i \mathcal{G}_i$. Return $\mathcal{H}$ as the required cover.

The set $\mathcal{H}$ covers S

1. For an element $s \in S$, we have that
$$\sum_{U \in \mathcal{F}, s \in U} \widehat{x_U} \geq 1, \qquad (2)$$
2. probability s not covered by $\mathcal{G}_i$ (ith iteration set).
$\mathbf{Pr}\left[s \text{ not covered by } \mathcal{G}_i\right]$
$= \mathbf{Pr}\left[\text{ no } U \in \mathcal{F}, \text{ s.t. } s \in U \text{ picked into } \mathcal{G}_i\right]$
$= \prod_{U \in \mathcal{F}, s \in U} \mathbf{Pr}\left[U \text{ was not picked into } \mathcal{G}_i\right]$
$= \prod_{U \in \mathcal{F}, s \in U} (1 - \widehat{x_U}) \leq \prod_{U \in \mathcal{F}, s \in U} \exp(-\widehat{x_U})$
$= \exp\left(-\sum_{U \in \mathcal{F}, s \in U} \widehat{x_U}\right) \leq \exp(-1) \leq \frac{1}{2}, \leq \frac{1}{2}$
3. probability s is not covered in all m iterations $\leq \left(\frac{1}{2}\right)^m < \frac{1}{n^{10}}$,
4. ...since $m = O(\log n)$.
5. probability one of n elements of S is not covered by $\mathcal{H}$ is $\leq n(1/n^{10}) = 1/n^9$.

Cost of solution

1. Have: $\mathbf{E}\left[\text{cost of } \mathcal{G}_i\right] \le \alpha^I$.
2. $\implies$ Each iteration expected cost of cover $\le$ cost of optimal solution (i.e., α^I).
3. Expected cost of the solution is

$$c_{\mathcal{H}} \le \sum_i c_{B_i} \le m\alpha^I = O\left(\alpha^I \log n\right).$$

The result

Theorem

By solving an LP one can get an $O(\log n)$-approximation to set cover by a randomized algorithm. The algorithm succeeds with high probability.

Same algorithms works for...

Corollary

By solving an LP one can get an $O(\log n)$-approximation to set cover by a randomized algorithm. The algorithm also works for the weighted case.

$$\begin{aligned} \min \quad & \alpha = \sum_{U \in \mathcal{F}} w_U x_U & \\ & 0 \le x_U \le 1 & \forall U \in \mathcal{F}, \\ & \sum_{U \in \mathcal{F}, s \in U} x_U \ge 1 & \forall s \in S. \end{aligned}$$

Rounding algorithm as before...

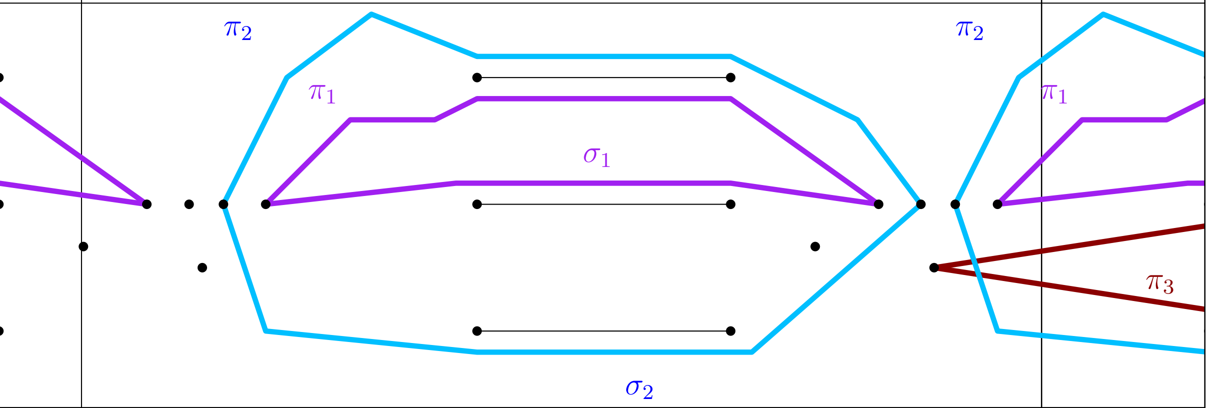
Cost of solution (weighted case)...

Same same, not the same.

1. Fractional LP solution. Target: $\widehat{\alpha}$ $\quad \forall U \in \mathcal{F}$: $\widehat{x_U} \in [0,1]$.
2. Integral opt solution. Target: α^I. $\quad \forall U \in \mathcal{F}$: $x_U^I \in \{0,1\}$.
3. $\alpha^I = \sum_{U \in \mathcal{F}} w_U x_U^I$.
4. Rounding. $\forall U \in \mathcal{F}$: $\mathbf{Pr}[X_U = 1] = \widehat{x_U}$. $\mathbf{E}\left[\text{cost } \mathcal{G}_i\right] = \sum_{U \in \mathcal{F}} \mathbf{E}[w_U X_U] = \sum_{U \in \mathcal{F}} w_U \widehat{x_U} = \widehat{\alpha} \le \alpha^I$.
5. Have: $\mathbf{E}\left[\text{cost of } \mathcal{G}_i\right] \le \alpha^I$.
6. $\implies$ Each iteration expected cost of cover $\le$ cost of optimal solution (i.e., α^I).
7. Expected cost of the solution is

$$c_{\mathcal{H}} \le \sum_{i=1}^{O(\log n)} c_{B_i} \le m\alpha^I = O\left(\alpha^I \log n\right).$$

Minimizing congestion by example



Minimizing congestion

1. $\mathbf{G}$: graph. n vertices.
2. π_i, σ_i paths with the same endpoints $\mathbf{v}_i, \mathbf{u}_i \in \mathbf{V}(\mathbf{G})$, for $i = 1, \ldots, t$.
3. Rule I: Send one unit of flow from $\mathbf{v}_i$ to $\mathbf{u}_i$.
4. Rule II: Choose whether to use π_i or σ_i.
5. Target: No edge in $\mathbf{G}$ is being used too much.

Definition

Given a set X of paths in a graph $\mathbf{G}$, the **congestion** of X is the maximum number of paths in X that use the same edge.

Minimizing congestion

1. IP $\implies$ LP:

$$
\begin{aligned}
\textbf{min} \quad & w & \\
\text{s.t.} \quad & x_i \geq 0 & i = 1, \ldots, t, \\
& x_i \leq 1 & i = 1, \ldots, t, \\
& \sum_{\mathsf{e} \in \pi_i} x_i + \sum_{\mathsf{e} \in \sigma_i} (1 - x_i) \leq w & \forall \mathsf{e} \in E.
\end{aligned}
$$

2. $\widehat{x}_i$: value of x_i in the optimal LP solution.
3. $\widehat{w}$: value of w in LP solution.
4. Optimal congestion must be bigger than $\widehat{w}$.
5. X_i: random variable one with probability $\widehat{x}_i$, and zero otherwise.
6. If $X_i = 1$ then use π to route from $\mathbf{v}_i$ to $\mathbf{u}_i$.
7. Otherwise use σ_i.

Minimizing congestion

1. Congestion of $\mathbf{e}$ is $Y_\mathsf{e} = \sum_{\mathsf{e} \in \pi_i} X_i + \sum_{\mathsf{e} \in \sigma_i} (1 - X_i)$.
2. And in expectation

$$
\begin{aligned}
\alpha_\mathsf{e} = \mathbf{E}\Big[Y_\mathsf{e}\Big] &= \mathbf{E}\left[\sum_{\mathsf{e} \in \pi_i} X_i + \sum_{\mathsf{e} \in \sigma_i} (1 - X_i)\right] \\
&= \sum_{\mathsf{e} \in \pi_i} \mathbf{E}\Big[X_i\Big] + \sum_{\mathsf{e} \in \sigma_i} \mathbf{E}\Big[(1 - X_i)\Big] \\
&= \sum_{\mathsf{e} \in \pi_i} \widehat{x}_i + \sum_{\mathsf{e} \in \sigma_i} (1 - \widehat{x}_i) \leq \widehat{w}.
\end{aligned}
$$

3. $\widehat{w}$: Fractional congestion (from LP solution).

Minimizing congestion - continued

1. $Y_e = \sum_{e \in \pi_i} X_i + \sum_{e \in \sigma_i} (1 - X_i)$.
2. Y_e is just a sum of independent $0/1$ random variables!
3. Chernoff inequality tells us sum can not be too far from expectation!

Minimizing congestion - continued

1. By Chernoff inequality:

$$\Pr\left[Y_e \geq (1+\delta)\alpha_e\right] \leq \exp\left(-\frac{\alpha_e \delta^2}{4}\right) \leq \exp\left(-\frac{\widehat{w}\delta^2}{4}\right).$$

2. Let $\delta = \sqrt{\frac{400}{\widehat{w}} \ln t}$. We have that

$$\Pr\left[Y_e \geq (1+\delta)\alpha_e\right] \leq \exp\left(-\frac{\delta^2 \widehat{w}}{4}\right) \leq \frac{1}{t^{100}},$$

3. If $t \geq n^{1/50} \implies \forall$ edges in graph congestion $\leq (1+\delta)\widehat{w}$.
4. t: Number of pairs, n: Number of vertices in $\mathbf{G}$.

Minimizing congestion - continued

1. Got: For $\delta = \sqrt{\frac{400}{\widehat{w}} \ln t}$. We have

$$\Pr\left[Y_e \geq (1+\delta)\alpha_e\right] \leq \exp\left(-\frac{\delta^2 \widehat{w}}{4}\right) \leq \frac{1}{t^{100}},$$

2. Play with the numbers. If $t = n$, and $\widehat{w} \geq \sqrt{n}$. Then, the solution has congestion larger than the optimal solution by a factor of

$$1 + \delta = 1 + \sqrt{\frac{20}{\widehat{w}} \ln t} \leq 1 + \frac{\sqrt{20 \ln n}}{n^{1/4}},$$

which is of course extremely close to 1, if n is sufficiently large.

Minimizing congestion: result

Theorem

1. $\mathbf{G}$: *Graph n vertices.*
2. $(s_1, t_1), \ldots, (s_t, t_t)$: *pairs o vertices*
3. π_i, σ_i: *two different paths connecting s_i to t_i*
4. $\widehat{w}$: *Fractional congestion at least $n^{1/2}$.*
5. opt: *Congestion of optimal solution.*
6. $\implies$ *In polynomial time (LP solving time) choose paths*
   - 6.1 *congestion $\forall$ edges:* $\leq (1+\delta)\mathrm{opt}$
   - 6.2 $\delta = \sqrt{\frac{20}{\widehat{w}} \ln t}$.

When the congestion is low

1. Assume $\widehat{w}$ is a constant.
2. Can get a better bound by using the Chernoff inequality in its more general form.
3. set $\delta = c \ln t / \ln \ln t$, where c is a constant. For $\mu = \alpha_e$, we have that

$$\Pr\left[Y_e \geq (1+\delta)\mu\right] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$$
$$= \exp\Big(\mu(\delta - (1+\delta)\ln(1+\delta))\Big)$$
$$= \exp\Big(-\mu c' \ln t\Big) \leq \frac{1}{t^{O(1)}},$$

where c' is a constant that depends on c and grows if c grows.

When the congestion is low

1. Just proved that...
2. if the optimal congestion is $O(1)$, then...
3. algorithm outputs a solution with congestion $O(\log t / \log \log t)$, and this holds with high probability.

Chernoff inequality

Problem

Let $X_1, \ldots X_n$ be n independent Bernoulli trials, where

$$\Pr\left[X_i = 1\right] = p_i, \qquad \Pr\left[X_i = 0\right] = 1 - p_i,$$
$$Y = \sum_i X_i, \qquad \text{and} \qquad \mu = \mathbf{E}\left[Y\right].$$

We are interested in bounding the probability that $Y \geq (1+\delta)\mu$.

Chernoff inequality

Theorem (Chernoff inequality)

For any $\delta > 0$,

$$\Pr\left[Y > (1+\delta)\mu\right] < \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu.$$

Or in a more simplified form, for any $\delta \leq 2e - 1$,

$$\Pr\left[Y > (1+\delta)\mu\right] < \exp(-\mu\delta^2/4),$$

and

$$\Pr\left[Y > (1+\delta)\mu\right] < 2^{-\mu(1+\delta)},$$

for $\delta \geq 2e - 1$.

More Chernoff...

Theorem

Under the same assumptions as the theorem above, we have

$$\mathbf{Pr}\left[Y < (1-\delta)\mu\right] \leq \exp\left(-\mu\frac{\delta^2}{2}\right).$$