Chapter 26

Approximation Algorithms using Linear Programming

OLD CS 473: Fundamental Algorithms, Spring 2015
April 30, 2015

26.0.1 Weighted vertex cover

26.0.2 Weighted vertex cover

26.0.2.1 Weighted vertex cover

Weighted Vertex Cover problem $\mathsf{G} = (\mathsf{V}, \mathsf{E})$.

Each vertex $\mathsf{v} \in \mathsf{V}$: cost c_v.

Compute a vertex cover of minimum cost.

(A) vertex cover: subset of vertices V so each edge is covered.
(B) **NP-Hard**
(C) ...unweighted **Vertex Cover** problem.
(D) ... write as an integer program (IP):
(E) $\forall \mathsf{v} \in \mathsf{V}$: $x_\mathsf{v} = 1 \iff \mathsf{v}$ in the vertex cover.
(F) $\forall \mathsf{vu} \in \mathsf{E}$: covered. $\implies x_\mathsf{v} \vee x_\mathsf{u}$ true. $\implies x_\mathsf{v} + x_\mathsf{u} \geq 1$.
(G) minimize total cost: $\min \sum_{\mathsf{v} \in \mathsf{V}} x_\mathsf{v} \mathsf{c}_\mathsf{v}$.

26.0.3 Weighted vertex cover

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

$$\begin{aligned} \min \quad & \sum_{\mathsf{v} \in \mathsf{V}} \mathsf{c}_\mathsf{v} x_\mathsf{v}, & \\ \text{such that} \quad & x_\mathsf{v} \in \{0, 1\} & \forall \mathsf{v} \in \mathsf{V} \\ & x_\mathsf{v} + x_\mathsf{u} \geq 1 & \forall \mathsf{vu} \in \mathsf{E}. \end{aligned} \tag{26.1}$$

(A) ... NP-Hard.
(B) relax the integer program.
(C) allow x_v get values $\in [0, 1]$.
(D) $x_v \in \{0, 1\}$ replaced by $0 \le x_v \le 1$. The resulting LP is

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

26.0.3.2 Weighted vertex cover – rounding the LP

(A) Optimal solution to this LP: $\widehat{x}_v$ value of var X_v, $\forall v \in V$.
(B) optimal value of LP solution is $\widehat{\alpha} = \sum_{v \in V} c_v \widehat{x}_v$.
(C) optimal integer solution: x_v^I, $\forall v \in V$ and α^I.
(D) **Any valid solution to IP is valid solution for LP!**
(E) $\widehat{\alpha} \le \alpha^I$.
Integral solution not better than LP.
(F) Got fractional solution (i.e., values of $\widehat{x}_v$).
(G) Fractional solution is better than the optimal cost.
(H) Q: How to turn fractional solution into a (valid!) integer solution?
(I) Using ***rounding***.

26.0.3.3 How to round?

(A) consider vertex v and fractional value $\widehat{x}_v$.
(B) If $\widehat{x}_v = 1$ then include in solution!
(C) If $\widehat{x}_v = 0$ then do **<u>not</u>not** include in solution.
(D) if $\widehat{x}_v = 0.9 \implies$ LP considers v as being 0.9 useful.
(E) The LP puts its money where its belief is...
(F) ...$\widehat{\alpha}$ value is a function of this "belief" generated by the LP.
(G) **Big idea**: Trust LP values as guidance to usefulness of vertices.

26.0.3.4 II: How to round?

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

(A) Pick all vertices $\ge$ threshold of usefulness according to LP.
(B) $S = \left\{ v \;\middle|\; \widehat{x}_v \ge 1/2 \right\}$.
(C) **Claim**: S a valid vertex cover, and cost is low.

(A) Indeed, edge cover as: $\forall vu \in E$ have $\widehat{x}_v + \widehat{x}_u \ge 1$.
(B) $\widehat{x}_v, \widehat{x}_u \in (0, 1)$
$\implies \widehat{x}_v \ge 1/2$ or $\widehat{x}_u \ge 1/2$.
$\implies v \in S$ or $u \in S$ (or both).
$\implies S$ covers all the edges of G.

26.0.3.5 Cost of solution

Cost of S:

$$\mathsf{c}_S = \sum_{\mathsf{v}\in S} \mathsf{c}_\mathsf{v} = \sum_{\mathsf{v}\in S} 1 \cdot \mathsf{c}_\mathsf{v} \le \sum_{\mathsf{v}\in S} 2\widehat{x}_\mathsf{v} \cdot \mathsf{c}_\mathsf{v} \le 2 \sum_{\mathsf{v}\in \mathsf{V}} \widehat{x}_\mathsf{v}\mathsf{c}_\mathsf{v} = 2\widehat{\alpha} \le 2\alpha^I,$$

since $\widehat{x}_\mathsf{v} \ge 1/2$ as $\mathsf{v} \in S$.
α^I is cost of the optimal solution $\implies$

Theorem 26.0.1. *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.*

26.0.4 The lessons we can take away

26.0.4.1 Or not - boring, boring, boring.

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

26.0.5 Revisiting Set Cover

26.0.5.1 Revisiting Set Cover

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

Problem: Set Cover

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

26.0.5.2 Set Cover – IP & LP

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

Next, we relax this IP into the following LP.

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

26.0.5.3 Set Cover – IP & LP

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

26.0.5.4 Set Cover – Rounding continued

(A) **Solution**: Repeat rounding stage $m = 10\lceil \lg n \rceil = O(\log n)$ times.
(B) $n = |S|$.
(C) $\mathcal{G}_i$: random cover computed in ith iteration.
(D) $\mathcal{H} = \cup_i \mathcal{G}_i$. Return $\mathcal{H}$ as the required cover.

26.0.5.5 The set $\mathcal{H}$ covers S

(A) For an element $s \in S$, we have that

$$
\sum_{U \in \mathcal{F}, s \in U} \widehat{x_U} \geq 1, \tag{26.2}
$$

(B) probability s not covered by $\mathcal{G}_i$ (ith iteration set).

$$
\begin{aligned}
& \mathbf{Pr}\left[s \text{ not covered by } \mathcal{G}_i\right] \\
& = \mathbf{Pr}\left[\text{ no } U \in \mathcal{F}, \text{ s.t. } s \in U \text{ picked into } \mathcal{G}_i\right] \\
& = \prod_{U \in \mathcal{F}, s \in U} \mathbf{Pr}\left[U \text{ was not picked into } \mathcal{G}_i\right] \\
& = \prod_{U \in \mathcal{F}, s \in U} \left(1 - \widehat{x_U}\right) \leq \prod_{U \in \mathcal{F}, s \in U} \exp\left(-\widehat{x_U}\right) \\
& = \exp\left(-\sum_{U \in \mathcal{F}, s \in U} \widehat{x_U}\right) \leq \exp(-1) \leq \tfrac{1}{2}, \leq \tfrac{1}{2}
\end{aligned}
$$

26.0.6 The set $\mathcal{H}$ covers S

26.0.6.1 Probability of a single item to be covered

(A) $\mathbf{Pr}\Big[s \text{ not covered by } \mathcal{G}_i\Big] \leq 1/2$.
(B) Number of iterations of rounding $m = O(\log n)$.
(C) Covering with sets in $\mathcal{G}_1, \ldots, \mathcal{G}_m$.
(D) probability s is not covered in all m iterations

$$\begin{aligned} P_s &= \mathbf{Pr}\Big[s \text{ not covered by } \mathcal{G}_1, \ldots, \mathcal{F}_m\Big] \\ &\leq \mathbf{Pr}\Big[(s \notin \mathcal{F}_1) \cap (s \notin \mathcal{F}_2) \cap \ldots \cap (s \notin \mathcal{F}_m)\Big] \\ &\leq \mathbf{Pr}\Big[s \notin \mathcal{F}_1\Big]\,\mathbf{Pr}\Big[s \notin \mathcal{F}_2\Big] \cdots \mathbf{Pr}\Big[s \notin \mathcal{F}_m\Big] \\ &= \frac{1}{2} \times \frac{1}{2} \times \cdots \times \frac{1}{2} = \left(\frac{1}{2}\right)^m < \frac{1}{n^{10}}, \end{aligned}$$

26.0.7 The set $\mathcal{H}$ covers S

26.0.7.1 Probability of all items to be covered

(A) $n = |S|$,
(B) Probability of $s \in S$, not to be in $\mathcal{G}_1 \cup \ldots \cup \mathcal{F}_m$ is

$$P_s < \frac{1}{n^{10}}.$$

(C) probability one of n elements of S is not covered by $\mathcal{H}$ is

$$\sum_{s \in S} \mathbf{Pr}[s \notin \mathcal{G}_1 \cup \ldots \cup \mathcal{F}_m] = \sum_{s \in S} P_s < n(1/n^{10}) = 1/n^9.$$

XXX

26.0.7.2 Reminder: LP for Set Cover

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

(A) Solve the LP.
(B) $\widehat{x_U}$: Value of x_u in the optimal LP solution.
(C) Fractional solution: $\widehat{\alpha} = \sum_{U \in \mathcal{F}} \widehat{x_U}$.
(D) Integral solution (what we want): $\alpha^I \geq \widehat{\alpha}$.

26.0.7.3 Cost of solution

(A) $(S, \mathcal{F})$: Given instance of **Set Cover**.
(B) For $U \in \mathcal{F}$, $\widehat{x_U}$: LP value for set U in optimal solution.
(C) For $\mathcal{G}_i$: Indicator variable $Z_u = 1 \iff U \in \mathcal{G}_i$.
(D) Expected number of sets in the ith sample:
$\mathbf{E}[|\mathcal{G}_i|] = \mathbf{E}\left[\sum_{U \in \mathcal{F}} Z_U\right] = \sum_{U \in \mathcal{F}} \mathbf{E}[Z_U] = \sum_{U \in \mathcal{F}} \widehat{x_U}$
$= \widehat{\alpha} \leq \alpha^I$.
(E) $\implies$ Each iteration expected cost of cover $\leq$ cost of optimal solution (i.e., α^I). XXX
(F) Expected size of the solution is

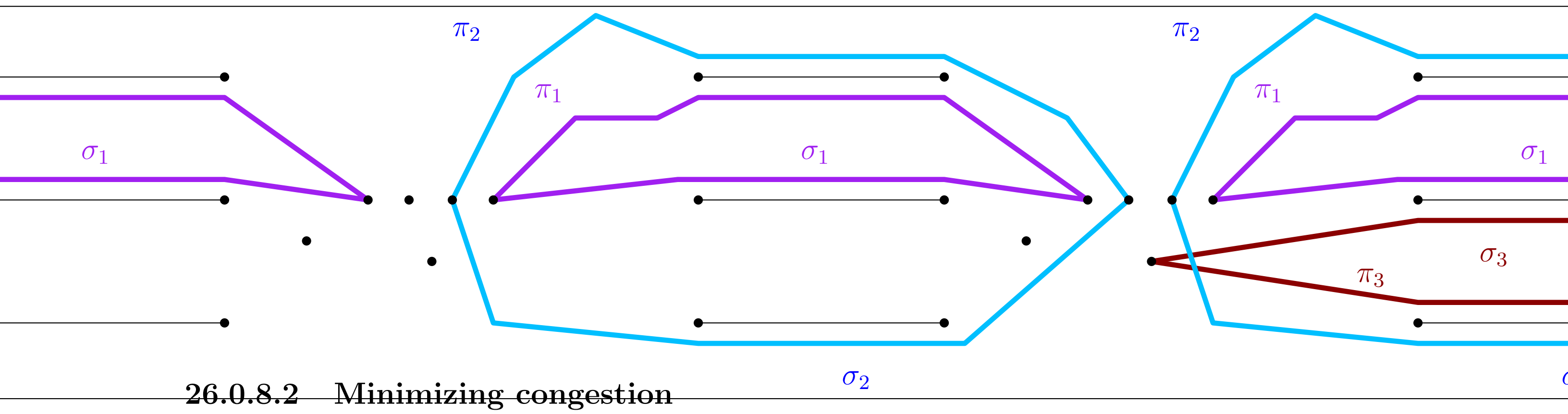
$$\mathbf{E}[|\mathcal{H}|] = \mathbf{E}[|\cup_i \mathcal{G}_i|] \leq \mathbf{E}\left[\sum_i |\mathcal{G}_i|\right] \leq m\alpha^I = O(\alpha^I \log n).$$

26.0.7.4 The result

Theorem 26.0.2. *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.*

26.0.8 Minimizing congestion

26.0.8.1 Minimizing congestion by example



26.0.8.2 Minimizing congestion

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

Definition 26.0.3. *Given a set X of paths in a graph G, the* **congestion** *of X is the maximum number of paths in X that use the same edge.*

26.0.8.3 Minimizing congestion

(A) IP $\implies$ LP:

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

(B) $\widehat{x_i}$: value of x_i in the optimal LP solution.
(C) $\widehat{w}$: value of w in LP solution.
(D) Optimal congestion must be bigger than $\widehat{w}$.
(E) X_i: random variable one with probability $\widehat{x_i}$, and zero otherwise.
(F) If $X_i = 1$ then use π to route from v_i to u_i.
(G) Otherwise use σ_i.

26.0.8.4 Minimizing congestion

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

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

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

26.0.8.5 Minimizing congestion - continued

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

26.0.8.6 Minimizing congestion - continued

(A) By Chernoff inequality:

$$\mathbf{Pr}\big[Y_\mathsf{e} \geq (1+\delta)\alpha_\mathsf{e}\big] \leq \exp\left(-\frac{\alpha_\mathsf{e}\delta^2}{4}\right) \leq \exp\left(-\frac{\widehat{w}\delta^2}{4}\right).$$

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

$$\mathbf{Pr}\Big[Y_{\mathsf{e}} \geq (1+\delta)\alpha_{\mathsf{e}}\Big] \leq \exp\left(-\frac{\delta^2 \widehat{w}}{4}\right) \leq \frac{1}{t^{100}},$$

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

26.0.8.7 Minimizing congestion - continued

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

$$\mathbf{Pr}\Big[Y_{\mathsf{e}} \geq (1+\delta)\alpha_{\mathsf{e}}\Big] \leq \exp\left(-\frac{\delta^2 \widehat{w}}{4}\right) \leq \frac{1}{t^{100}},$$

(B) 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.

26.0.8.8 Minimizing congestion: result

Theorem 26.0.4. *(A) G: Graph n vertices.*
(B) $(s_1, t_1), \ldots, (s_t, t_t)$: pairs o vertices
(C) π_i, σ_i: two different paths connecting s_i to t_i
(D) $\widehat{w}$: Fractional congestion at least $n^{1/2}$.
(E) opt: Congestion of optimal solution.
(F) $\implies$ In polynomial time (LP solving time) choose paths
  (A) congestion $\forall$ edges: $\leq (1+\delta)\mathrm{opt}$
  (B) $\delta = \sqrt{\frac{20}{\widehat{w}} \ln t}$.

26.0.8.9 When the congestion is low

(A) Assume $\widehat{w}$ is a constant.
(B) Can get a better bound by using the Chernoff inequality in its more general form.
(C) set $\delta = c \ln t / \ln \ln t$, where c is a constant. For $\mu = \alpha_{\mathsf{e}}$, we have that

$$\begin{aligned}\mathbf{Pr}\Big[Y_{\mathsf{e}} \geq (1+\delta)\mu\Big] &\leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu \\ &= \exp\Big(\mu\big(\delta - (1+\delta)\ln(1+\delta)\big)\Big) \\ &= \exp\Big(-\mu c' \ln t\Big) \leq \frac{1}{t^{O(1)}},\end{aligned}$$

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

26.0.8.10 When the congestion is low

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

26.0.9 Reminder about Chernoff inequality

26.0.9.1 The Chernoff Bound — General Case
26.0.9.2 Chernoff inequality

Problem 26.0.5. *Let $X_1, \dots X_n$ be n independent Bernoulli trials, where*

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

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

26.0.9.3 Chernoff inequality

Theorem 26.0.6 (Chernoff inequality). *For any $\delta > 0$,*

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

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

$$\mathbf{Pr}\Big[Y > (1+\delta)\mu\Big] < \exp\big(-\mu\delta^2/4\big),$$

and

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

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

26.0.9.4 More Chernoff...

Theorem 26.0.7. *Under the same assumptions as the theorem above, we have*

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