Approximation

Consider an arbitrary optimization problem. Let $OPT(X)$ denote the optimal value for input X. We say algorithm A is an $\alpha(n) \geq 1$ approximation if for all X

$$\frac{OPT(X)}{A(X)} \leq \alpha(n) \quad \text{and} \quad \frac{A(X)}{OPT(X)} \leq \alpha(n)$$

Minimum Vertex Cover

Vertex Cover: A vertex subset s.t. each edge has at least one endpoint in the set.

A Greedy Algorithm:

- $C \leftarrow \emptyset$
- while G has at least one edge:
  - $v \leftarrow$ vertex with max degree in G
  - $C \leftarrow C \cup \{v\}$
  - $G \leftarrow G \setminus v$

Let's sum up d_i's for the first OPT iterations of the main loop. We have

$$\sum_{i=1}^{OPT} d_i \geq \sum_{i=1}^{OPT} \frac{|G_i|}{OPT} \geq \sum_{i=1}^{OPT} \frac{|G_{OPT}|}{OPT}$$

$$= |G_{OPT}| = |G| - \sum_{i=1}^{OPT} d_i$$

This means that

$$\sum_{i=1}^{opt} d_i \geq |G|/_2 .$$

Thm: The output of Greedy Vertex Cover is an $O(\lg n)$ approximation.

Proof: Let us define:

G_i : The graph G after i iterations of the main loop in the greedy algorithm.

d_i : max degree in graph G_i.

Let C^* be the optimal minimum vertex cover. That is, $|C^*| = OPT$. Note that C^* is a vertex cover for every G_i. Therefore, for any i,

$$\sum_{v \in C^*} \deg_{G_i}(v) \geq |G_i|$$

The # edges in graph G_i

This means that there is a vertex $v \in C^*$ s.t.

$$\deg_{G_i}(v) \geq \frac{|G_i|}{|C^*|} = \frac{|G_i|}{OPT},$$

which also implies $d_i \geq |G_i| / OPT$.

This means that within the first OPT iterations, we cover at least half of the edges, therefore repeating this for $\lg |G|$ times, we cover all edges. Total # vertices added is

$$OPT \cdot \lg |G| \leq 2 \; OPT \lg n . \quad \square$$

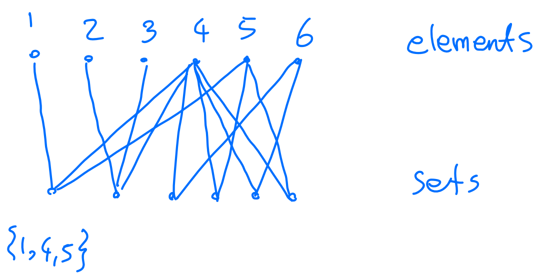
Remark: There are graphs for which this alg does indeed select $\lg n$ times more vertices than OPT.

Set Cover and Hitting Set

We are given a set system $(X, \mathcal{F})$ where X is a finite ground set and $\mathcal{F}$ is a family of subsets of X.

Set Cover: The minimum # of subsets in $\mathcal{F}$ whose union is X.

Hitting Set: The minimum # of elements in X s.t. every set in $\mathcal{F}$ has at least one element picked.



Note that given a graph $G=(V,E)$, we can let E be the ground set and for each vertex v have a set that includes all the edges adj to v. So solving this set cover instance solves min vertex cover $\Rightarrow$ Set Cover is NP-hard

---

It turns out the a simple generalization of the greedy vertex cover algorithm also $\lg n$-approximates set cover.

- $C \leftarrow \emptyset$
- while X is not empty:
  - $S \leftarrow$ set in $\mathcal{F}$ of largest size
  - $C \leftarrow C \cup \{S\}$
  - delete all elements of S from X

Thm: The output is a set cover of size at most $O(\lg n)$ times the min set cover.

Let's revisit MVC problem

Consider the following algorithm:

$C \leftarrow \emptyset$

while there exists an edge e in G:

  Put both endpoints of e in C.

  Remove endpoints of e along with their edges.

Let $e_1, e_2, e_3, \ldots, e_k$ be the edges picked to cover by the algorithm.

Observe that $e_1, \ldots, e_k$ are vertex disjoint

Therefore any vertex cover must be of size $\geq k$. Our solution has size $2k$.

Thm: There is a 2-apx alg for MVC.

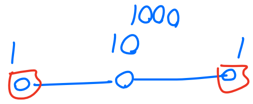
Weighted Vertex Cover

Input: $G=(V,E)$ and weights $w: V \to \mathbb{R}_+$

Goal: Select a vertex cover $U \subseteq V$

minimizing $\sum_{v \in U} w_v$.

Both alg's we discussed for the unweighted case can be arbitrarily bad



Greedy : 10^{1000}

2nd alg: $10^{1000}+1$

OPT : 2

Thm: There is a polytime 2-apx sol'n also for weighted min VC problem.

Min VC can be formulated as an integer linear program

$$\min \sum_v w_v \cdot x_v$$

$$x_u + x_v \geq 1 \qquad \forall (u,v) \in E$$

$$x_v \in \{0,1\} \qquad \forall v \in V$$

Let's relax the integer constraint.

$$\min \sum_v w_v \cdot x_v$$

$$x_u + x_v \geq 1 \qquad \forall (u,v) \in E$$

$$0 \leq x_v \leq 1 \qquad \forall v \in V$$

Rounding the LP:

$$\text{Let } x'_v = \begin{cases} 1 & \text{if } x_v \geq \frac{1}{2} \\ 0 & \text{if } x_v < \frac{1}{2} \end{cases}$$

Observe that x' indicates a vertex cover.

On the other hand $x'_v \leq 2x_v$, therefore

$$\sum_v w_v x'_v \leq 2 \sum_v w_v x_v \leq 2\,\mathrm{OPT}$$