Randomized Rounding:

Set cover:

Input: - Universe U, $|U| = n$

- Sets $\mathcal{S} = \{S_1, S_2, \dots, S_m\}$ s.t. $S_i \subseteq U \; \forall i$

- costs $c : \mathcal{S} \to \mathbb{R}^+$

Feasible: $\mathcal{T} \subseteq \mathcal{S}$ s.t. $\bigcup_{S \in \mathcal{T}} S = U$

Objective: $\min \; c(\mathcal{T}) = \sum_{S \in \mathcal{T}} c(S)$

LP Relaxation:

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

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

$$0 \leq x_S \leq 1 \qquad \forall S \in \mathcal{S}$$

Thm: $OPT(ILP) = OPT(SC)$

Rounding Algorithm:

Solve LP to get fractional solution x^*

For each $S \in \mathcal{S}$ independently:

set $x'_S = 1$ with probability $\min(1, \lambda x^*_S)$

0 otherwise

$\Theta(\log n)$

$T = \{ S \in \mathcal{S} : x'_S = 1 \}$

<u>Lemma</u>: $E[c(T)] \leq \lambda \cdot LP \leq \lambda \cdot OPT$

<u>Pf</u>:

$$E[c(T)] = E\Big[\sum_{S \in \mathcal{S}} c(S)\, x'_S\Big]$$

$$= \sum_{S \in \mathcal{S}} c(S)\, E[x'_S]$$

$$= \sum_{S \in \mathcal{S}} c(S) \cdot \min(1, \lambda x^*_S)$$

$$\leq \lambda \cdot \sum_{S \in \mathcal{S}} c(S)\, x^*_S$$

$$= \lambda \cdot LP$$

Lemma: Let $u \in U$. Then

$$\Pr[e \text{ not covered by } T] \leq e^{-\lambda}$$

Pf:

$$\Pr[u \text{ not covered by } T]$$

$$= \Pr[x_S' = 0 \quad \forall S \in \mathcal{S} : u \in S]$$

$$= \prod_{S : u \in S} \Pr[x_S' = 0] \qquad \text{(independence)}$$

If $\exists S$ with $u \in S$ and $x_S^* \geq \frac{1}{\lambda}$

$$\Rightarrow x_S' = 1 \Rightarrow \Pr[x_S' = 0] = 0$$

$$\Rightarrow \Pr[u \text{ not covered by } T] = 0$$

otherwise:

$$= \prod_{S : u \in S} (1 - \lambda x_S^*)$$

$$\leq \prod_{S : u \in S} e^{-\lambda x_S^*}$$

$$= e^{-\lambda \sum_{S: u \in S} x_S^*}$$

$$\leq e^{-\lambda}$$

Def: "high probability": $1 - \frac{1}{n^c}$ for some $c \geq 1$

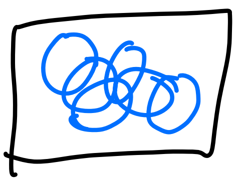
Thm: Set $\lambda = c \cdot \ln n$ for $c \geq 2$. Algorithm is an $O(\log n)$-approximation:

- $E[c(T)] \leq O(\log n) \cdot OPT$
- T feasible w.h.p.

Pf: cost from first lemma ✓

Feasible:

$Pr[T \text{ not feasible}] = Pr[\exists u \in U \text{ s.t. } T \text{ doesn't cover } u]$



$$\leq \sum_{u \in U} Pr[T \text{ doesn't cover } u] \quad \text{(union bound)}$$

$$\leq \sum_{u \in U} e^{-c \ln n}$$

$$= \sum_{u \in U} \frac{1}{n^c} = n \frac{1}{n^c} = \frac{1}{n^{c-1}}$$

Randomized bounds: why is feasible w.h.p., expectation for cost enough?

Markov's inequality: $Pr[\text{cost} > (1+\varepsilon) E[\text{cost}]] \leq \frac{1}{1+\varepsilon}$

$\Rightarrow$ repeat $\ell = \frac{2 \log n}{\log(1+\varepsilon)}$ times, take best

$Pr[\text{all have cost} > (1+\varepsilon) E[\text{cost}]]$

$$\leq \left(\frac{1}{1+\varepsilon}\right)^{\frac{2 \log n}{\log 1+\varepsilon}} = \frac{1}{n^2}$$

All feasible w.h.p. by union bound

$\Rightarrow$ w.h.p., cost $\leq (1+\varepsilon) E[\text{cost}]$, feasible

UFL:

Metric Uncapacitated Facility Location (UFL):

Input: - Metric space (V, d)

- Facility Opening costs $f: V \to \mathbb{R}^+$

Feasible solution: $S \subseteq V$ $S \neq \emptyset$

Objective: min $\text{cost}(S) = \sum_{i \in S} f(i) + \sum_{j \in V} d(j, S)$

$$\min \underbrace{\sum_{i \in V} f(i) y_i}_{= F(x,y)} + \sum_{j \in V} \underbrace{\sum_{i \in V} d(i,j) x_{ij}}_{= C(x,y)} = Z(x,y)$$

$$\text{s.t.} \quad \sum_{i \in V} x_{ij} = 1 \qquad \forall j \in V$$

$$x_{ij} \leq y_i \qquad \forall i,j \in V$$

$$0 \leq x_{ij} \leq 1 \qquad \forall i,j \in V$$

$$0 \leq y_i \leq 1 \qquad \forall i \in V$$

Def: $\Delta_j = \sum_{i \in V} d(i,j) x_{ij}$

Def: $B_j = \{ i \in V : d(i,j) \leq \alpha \Delta_j \}$

Init: $S = \emptyset$, all nodes unassigned

while $\exists$ unassigned nodes:

- Let j be unassigned node with min Δ_j
- Open $a(j) = \arg\min_{i \in B_j} f(i)$, assign j to $a(j)$
- Forall unassigned j' with $B_{j'} \cap B_j \neq \emptyset$, assign j' to $a(j)$

Let $\hat{S}$ be open facilities, integral solution $(\hat{x}, \hat{y})$

Set $\alpha = \frac{4}{3} \Rightarrow$ 4-approximation

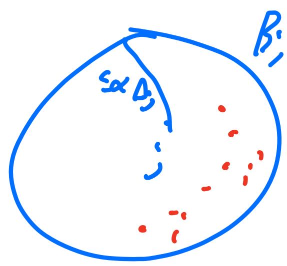
Better via <u>randomized</u> rounding?

<u>Def</u>: $M_j = \sum_{i \in B_j} x_{ij}$

$\Rightarrow M_j \leq 1$, and $M_j \geq \frac{\alpha - 1}{\alpha}$ (Markov)

<u>New alg</u>:

Same as old, but when considering j, pick $a(j)$ randomly from distribution where $i \in B_j$ has probability $\frac{x_{ij}}{M_j}$



$\Rightarrow$ set of open facilities $\hat{S}$, integral solution $(\hat{x}, \hat{y})$

Def: $\mathcal{C}$ is nodes considered by alg.

Note: $\mathcal{C}$ deterministic, and $B_j \cap B_{j'} = \emptyset \quad \forall j, j' \in \mathcal{C}, \; j \neq j'$

Def: $A(i,j) = \begin{cases} 1 & \text{if } i = a(j) \\ 0 & \text{otherwise} \end{cases} \qquad j \in \mathcal{C},\ i \in B_j$

Lemma: $E[F(\hat{x}, \hat{y})] \leq \frac{\alpha}{\alpha - 1} F(x, y)$

same as previous algorithm!

Pf:

$$E[F(\hat{x}, \hat{y})] = E\left[\sum_{i \in V} f(i) \hat{y}_i\right]$$

$$= E\left[\sum_{j \in \mathcal{C}} \sum_{i \in B_j} f(i) A(i,j)\right] \qquad \left(\begin{array}{l}\text{def of alg,} \\ B_j \cap B_{j'} = \emptyset \\ \forall j, j' \in \mathcal{C}\end{array}\right)$$

$$= \sum_{j \in C} \sum_{i \in B_j} f(i)\, E[A(i,j)]$$ (linearity of expectation)

$$= \sum_{j \in C} \sum_{i \in B_j} f(i) \frac{x_{ij}}{M_j}$$ (def of $A(i,j)$)

$$\leq \frac{\alpha}{\alpha-1} \sum_{j \in C} \sum_{i \in B_j} f(i)\, x_{ij}$$ ($M_j \geq \frac{\alpha-1}{\alpha}$)

$$\leq \frac{\alpha}{\alpha-1} \sum_{j \in C} \sum_{i \in B_j} f(i)\, y_i$$ ($x_{ij} \leq y_i$)

$$\leq \frac{\alpha}{\alpha-1} \sum_{i \in V} f(i)\, y_i$$ ($B_j \cap B_{j'} = \emptyset \;\forall j, j' \in C$)

$$= \frac{\alpha}{\alpha-1} F(x,y)$$

<u>Lemma</u>: $E[C(\hat{x},\hat{y})] \leq (2\alpha+1)\, C(x,y)$

det. alg was 3α

<u>Pf</u>: Let $j \in V$. <u>WTS</u>: $E[d(j,\hat{j})] \leq (2\alpha+1)\Delta_j$

<u>Case 1</u>: $j \in C$

Know $a(j) \in B_j \Rightarrow d(j,\hat{j}) \leq \alpha \Delta_j$: want $E[d(j,\hat{j})] \leq \Delta_j$

Intuition: know Δ_j was expected connection cost for j originally, and just shifted probability mass (closer)

$$E[d(j,\hat{s})] = E[d(j,a(j))]$$

$$= E\Big[\sum_{i \in B_j} d(i,j)\, A(i,j)\Big]$$

$$= \sum_{i \in B_j} d(i,j)\, \frac{x_{ij}}{\mu_j}$$

$$= \sum_{i \in B_j} \Big(d(i,j) + \Big(\frac{1}{\mu_j} - 1\Big) d(i,j) \Big) x_{ij}$$

$$= \sum_{i \in B_j} d(i,j) x_{ij} + \sum_{i \in B_j} d(i,j) \Big(\frac{1}{\mu_j} - 1\Big) x_{ij}$$

$$\leq \sum_{i \in B_j} d(i,j) x_{ij} + \sum_{i \in B_j} \alpha \Delta_j \Big(\frac{1}{\mu_j} - 1\Big) x_{ij}$$

$$= \sum_{i \in B_j} d(i,j) x_{ij} + \alpha \Delta_j \Big(\frac{1}{\mu_j} - 1\Big) \mu_j$$

$$= \sum_{i \in B_j} d(i,j) x_{ij} + \alpha \Delta_j (1 - \mu_j)$$

$$= \sum_{i \in B_j} d(i,j) x_{ij} + \alpha \Delta_j \sum_{i \notin B_j} x_{ij}$$

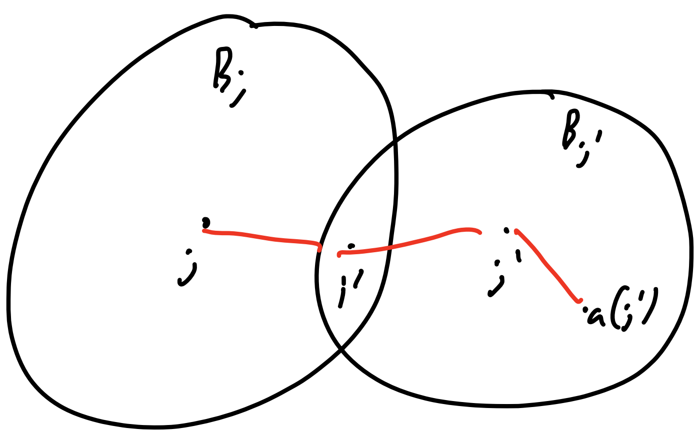
$$\leq \sum_{i \in B_j} d(i,j) x_{ij} + \sum_{i \notin B_j} d(i,j) x_{ij}$$

$$= \sum_{i \in U} d(i,j) x_{ij}$$

$$= \Delta_j$$

<u>Case 2</u>: $j \notin C$

Just like deterministic!



$\Delta_{j'} \leq \Delta_j$

$$E[d(j,\hat{s})] \leq E[d(j,i') + d(i',j') + d(j',a(j'))]$$

$$\leq d(j,i') + d(i',j') + E[d(j',a(j'))]$$

$$\leq \alpha \Delta_j + \alpha \Delta_{j'} + \Delta_{j'}$$

$$\leq (2\alpha + 1) \Delta_j$$

So instead of $\max\left(\frac{\alpha}{\alpha-1}, 3\alpha\right)$-approx, get

$\max\left(\frac{\alpha}{\alpha-1}, 2\alpha+1\right)$-approx

$$\frac{\alpha}{\alpha-1} = 2\alpha+1 \Rightarrow \alpha = (2\alpha+1)(\alpha-1)$$

$$\Rightarrow 2\alpha^2 - 2\alpha - 1 = 0$$

$$\Rightarrow \alpha = \frac{2 \pm \sqrt{4+8}}{4} = \frac{2+2\sqrt{3}}{4} = \frac{1}{2}(1+\sqrt{3})$$

$\Rightarrow$ approx of $2\alpha+1 = 2+\sqrt{3} \approx 3.73$