Approximation Algorithms

decision problems → optimized problems

\[ \text{CNF-SAT/3SAT} \]

K-CLIQUE

\[ \frac{\alpha}{\text{OPT}(I)} \]

\( \alpha > 1 \)

\[ \frac{\text{ALG}(I)}{\text{OPT}(I)} \leq \alpha \quad \text{if opt program is minimization} \]

\[ \frac{\text{ALG}(I)}{\text{OPT}(I)} \geq \frac{1}{\alpha} \quad \text{if opt program is maximization} \]

\[ \max \left( \frac{\text{ALG}(I)}{\text{OPT}(I)}, \frac{\text{OPT}(I)}{\text{ALG}(I)} \right) \leq \alpha. \]

Approximation Alg
Approximation Alg

Greedy

DP

LP

Randomized

SCHEDULING

Given:
1) Set J of m jobs.
2) Set M of n machines.
3) Each job j \in J, has processing time \( p_j \).

\( \langle X = x_1, x_2, \ldots, x_n \rangle = \text{partition of J into n bundles} \)

\text{Span of machine } i = \sum_{j \in X_i} p_j

Goal: minimize the max-span

\min \left( \max_{1 \leq i \leq n} \sum_{j \in X_i} p_j \right)

PARTITION \leq_p SCHEDULING

I : S = \{S_1, S_2, \ldots, S_m\} \implies I' = \text{job } j \text{ has processing time } P_j = S_j.

\exists S_1, S_2, S_1 \cap S_2 = \emptyset

\implies S_1 \cup S_2 = S

\implies I' \text{ partition } x_1, x_2
\[ s \cup s = s \]
\[ s \cup s = s \]
\[ s \cup s = s \]
\[ s \cup s = s \]
\[ s \cup s = s \]

**Greedy-Approximation**

While \( J \neq \emptyset \)

1. Pick \( j \in J \)
2. Pick machine \( i \) s.t. \( \sum_{j \in x_i} p_j \) is minimum
3. \( x_i = x_i \cup \{j\} \)
4. \( J = J \setminus \{j\} \)

\[ \text{OPT} \geq \max_j p_j \]
\[ \geq \frac{1}{n} \sum_{j \in J} p_j \]

Consider any iteration in the greedy alg:

\( j^* \) is assigned to machine \( i^* \)

\[ \sum_{j \in x_{i^*}} p_j \leq \frac{1}{n} \sum_{j \in J} p_j \leq \text{OPT} \]
\[ p_{j^*} \leq \max_{j \in J} p_j \leq \text{OPT} \]

\[ \sum p_j \leq 2\text{OPT} \]
\[ \sum_{j \in J_0} p_j \leq 2\text{OPT} \]

When the algorithm terminates, the total processing time on any machine \( \leq \text{OPT} \)

\[ \text{maxspan} \leq 2\text{OPT} \]

**2-approximation Alg**

Sort the jobs, \( J = \{j_1, j_2, \ldots, j_m\} \) in decreasing order of processing time.

\[ \text{Alg} \]

For \( l = 1 \) to \( m \)

Assign \( j_l \) to the machine with least load.

**Thm:** \( \text{Alg} \) is a 3/2-approximation alg

Assume \( m > n \) (otherwise \( \text{Alg} \) is optimal)

\[ \text{OPT} \geq \frac{1}{n} \sum_{j \in J} p_j \]

\[ \frac{1}{2} \text{OPT} \geq \frac{p_{n+1}}{n+1} \]
Every time after the 1st \( n \) rounds, 
All assigns job \( j^* \) to machine \( i^* \) 
\[ \frac{1}{n} \sum_{j \in j^*} p_j \leq \frac{1}{n} \sum_{j \in j^*} p_j \leq \text{opt} \]
\[ + \quad p_{j^*} \leq \frac{1}{2} \text{opt} \]
\[ \sum_{j \in j^*} p_j \leq \frac{3}{2} \text{opt} \]

---

**SET-COVER**

**Given:**
1. Set of elements \( \{e_1, e_2, \ldots, e_m \} \)
2. Sets \( S_1, S_2, \ldots, S_m \), \( S_i \subseteq \{e_1, e_2, \ldots, e_m \} \)
3. Set \( S_i \) has weight \( w_i \)

**Goal:** Find \( T = \{S_1, S_2, \ldots, S_d \} \) that \( \min_{i \in [d]} w_i \)

\[ \sum_{i \in [d]} S_i = \{e_1, e_2, \ldots, e_m \} \]

---

**SPECIAL CASE**

**VERTEX-COVER:** Given a graph \( G = (V, E) \)
We say vertex \( v \) covers \( e \) if \( e \) is incident on \( v \)
Find a set \( SC \) of vertices of minimum
\( e \) is incident to \( v \)

Goal: Find \( S \subseteq V \) of minimum size that covers all edges

\[ IS \leq_p \text{ VERTEX-COVER} \]

\( S \) is independent \( \iff \) \( \forall s \in S \) \( v \notin \forall s \)

\( \implies \) \( S \) is IS, but \( \forall S \) is not VC

\( \implies (u,v) \in E \iff u \notin \forall s, v \notin \forall s \)

\( \Downarrow \)

\( u \not\in S, v \not\in S \)

\( \Leftarrow S \) is VC, but \( S \) is not IS

\( \implies \forall s, v \not\in S \iff (u,v) \not\in E \)

\( \Downarrow \)

\( u \not\in \forall s, v \not\in \forall s \) \( (u,v) \) is not covered by \( \forall s \)

\[ IS \leq_p \text{ VC} \leq_p \text{ SET-COVER}. \]

\[ \text{AIC} \]

\( \text{T} \leftarrow \phi \)

\( \text{R} \leftarrow \{ e_1, e_2, \ldots, e_n \} \) || elements not covered yet

while \( L \neq \phi \)

\[ T \]
while $R \neq \emptyset$

1) let $S$ be the set with minimum value of $\frac{w_S}{|S \cap R|}$

2) $T \leftarrow T \cup S$

3) $R \leftarrow R \setminus S$

thm. $\aleph$ ALG is a $O(\log n)$-approximation

$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \in O(\log n)$

let $\text{ALG}$ be the solution of $\text{ALG}$

$\text{OPT}$ is our optimum solution

$\text{cost}(\text{OPT}) = \sum_{S \in \text{OPT}} w_i$

$\text{cost}(\text{ALG}) = \sum_{S \in \text{ALG}} w_i = \sum_{e \in \text{ALG}} c(e) \leq \sum_{e \in \text{OPT}} c(e)$

For each element $e$, we introduce a cost $c(e)$

if $e$ was covered by set $S_k$ in the $\text{ALG}$

and $e \notin \text{Uncovered elements at that point}$
\[ \text{Claim: } \sum_{e \in S_i} c(e) \leq H_{\|S_i\|} \cdot w_i \cdot \text{cost}(\text{OPT}) \]

It remains to prove the claim.

\[ \sum_{e \in S_i} c(e) \leq H_{\|S_i\|} \cdot w_i \]

\[ S_i = \{ e_1, e_2, \ldots, e_{\|S_i\|} \} \]

ordered exactly as they are covered by our alg.

\[ e_1 \text{ was covered by some set } s_b \text{ by our alg.} \]
\[
\cos t(e_i) = \frac{\omega_i}{|S_t \cap R|} \leq \frac{\omega_i}{1 + \frac{1}{d-t+1}} \leq \frac{\omega_i}{d-i+1}
\]

\[
\leq \sum_{i \in G(d)} \cos t(e_i) \leq \sum_{i \in G(d)} \frac{\omega_i}{d-i+1} = \omega_i \left( \frac{1}{1 + \frac{1}{2} + \ldots + \frac{1}{d}} + 1 \right)
\]

\[
= \omega_i \cdot H_d.
\]