Approximation Algorithms

decision problems ⇝ optimization problems

CNF-SAT/3SAT — Find an assignment to max # clauses satisfied

K-CLIQUE — MAX-CLIQUE → OPT(I)

$\alpha > 1$

α- Appx Algs :- Polynomial time algorithm ALG such that for every instance I of Π

$$\frac{ALG(I)}{OPT(I)} \leq \alpha$$ if our opt pgm is minimization

$$\frac{ALG(I)}{OPT(I)} \geq \frac{1}{\alpha}$$ if opt pgm is maximization

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

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

$\langle X = x_1, x_2, \ldots x_n \rangle$ = partition of J into n bundles

span of machine $i = \sum_{j \in X_i} p_j$

Goal: minimize the max-span

$$\min_{\langle X = x_1, x_2, \ldots x_n \rangle} \boxed{\max_{i \in [n]} \sum_{j \in X_i} p_j} \rightarrow \text{max-span}$$

PARTITION $\leq_p$ SCHEDULING

I: $S = \{s_1, s_2 \ldots s_n\}$ $\rightsquigarrow$ I' = job j has processing time $p_j = s_j$.

$m = 2$.

$\exists\, S_1, S_2$, $S_1 \cap S_2 = \phi$

$S_1 \cup S_2 = S$

$\exists$ partition x_1, x_2

$S_1 \cup S_2 = S$

$\sum_{s \in S_1} s = \sum_{s \in S_2} s = \frac{1}{2} \sum_{s \in S} s \iff$ $\exists$ partition X_1, X_2 s.t max-span $= \frac{1}{2} \sum_{j \in J} P_j$

Greedy-Approximation

while $J \neq \emptyset$ $\rightarrow$ $\exists$ no specific job that I choose.

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) $J = J \setminus \{j\}$

$$OPT \geq \max_j P_j$$

$$\geq \frac{1}{n} \cdot \sum_{j \in J} P_j.$$

Consider any iteration in the greedy alg job j^* is assigned to machine i^*

$$\sum_{j \in X_{i^*}} P_j \leq \frac{1}{n} \cdot \sum_{j \in J} P_j \leq OPT$$

$$+ \quad P_{j^*} \leq \max_{j \in J} P_j \leq OPT$$

$$\sum P_j \leq 2\,OPT$$

$$\sum_{j \in x_{i\alpha}} p_j \leq 2OPT$$

$$\bigcup_{j\alpha} = x'_{i\alpha}$$

When the algorithm terminates, the total processing time on any machine $\leq 2OPT$

$\Downarrow$

$$\text{maxspan} \leq 2OPT$$

2-approximation Alg

Sort the jobs, $J = \{j_1, j_2, \ldots, j_m\}$

decreasing order of processing time

ALG

For $l = 1$ to m

Assign j_l to the machine with least load.

Thm:- ALG is a 3/2-approximation alg

Assume $m > n$ (otherwise ALG is optimal)

$$OPT \geq \frac{1}{n} \cdot \sum_{j \in J} p_j$$

$$\frac{1}{2} OPT \geq p_{n+1}$$

$\frac{1}{2}$ OPT $\geq$ $\frac{1}{n+1}$

Every time after the 1^{st} n rounds, ALG assigns job j^* to machine i^*

$$\sum_{j \in X_{i^*}} p_j \leq \frac{1}{n} \cdot \sum_{j \in J} p_j \leq OPT$$

$$+ \quad p_{j^*} \leq OPT/2$$

$$\sum_{j \in X_{i^*} \cup j^*} p_j \leq \frac{3}{2} OPT$$

SET-COVER

Given :- 1) set of elements $\{e_1, e_2, \ldots, e_n\}$
2) Sets $S_1, S_2, \ldots, S_m$, s.t. $S_i \subseteq \{e_1, e_2, \ldots, e_n\}$
3) set S_i has weight w_i.

Goal :- Find $T = \{S_1, S_2, \ldots, S_d\}$ $\min \sum_{i \in [d]} w_i$
s.t. $\bigcup_{i \in [d]} S_i = \{e_1, e_2, \ldots, e_n\}$

$\downarrow$ SPECIAL CASE

VERTEX-COVER: Given :- Graph $G = (V, E)$
We say vertex v covers e if e is incident on v.
Goal :- Find $S \subset V$ of minimum

e is incident on v

Goal: $\Rightarrow$ Find $S \subseteq V$ of minimum size that covers all edges.

$$IS \leq_p \text{VERTEX-COVER}$$

S = independent set in G $\iff$ $V \setminus S$ is vertex cover

"$\Rightarrow$" S is IS, but $V \setminus S$ is not VC

$\Rightarrow$ $(u, v) \in E$ s.t. $u \notin V \setminus S$
$v \notin V \setminus S$

$\Downarrow$

$u \in S, v \in S$ ↯

"$\Leftarrow$" $\underline{V \setminus S \text{ is a VC}}$, but S is not IS

$\Rightarrow$ $u \in S$, $v \in S$ s.t. $\underline{(u, v)} \in E$

$u \notin V \setminus S$, $v \notin V \setminus S$ $\quad$ (u, v) is not covered by $V \setminus S$

$$IS \leq_p VC \leq_p \text{SET-COVER}.$$

ALG

$T \leftarrow \emptyset$

$R \leftarrow \{e_1, e_2, \ldots, e_n\}$ // elements not covered yet

while $R \neq \emptyset$

while $R \neq \phi$

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: ALG is a $O(\log n)$-approximation alg

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

Let ALG is the solution of ALG,
OPT is our optimum solution

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

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

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

if e was covered by set S_k in the ALG
& R = uncovered elements at that point

& ℓ = Uncovered elements at that point

$$c(e) = \frac{w_k}{|S_k \cap R|}$$

Claim :- $\sum_{e \in S_i} c(e) \leq H_{|S_i|} \cdot w_i$.

Assume claim is true

$$\sum_{S_i \in OPT} \sum_{e \in S_i} c(e) \leq H_{|S_i|} w_i \leq \sum_{S_i \in OPT} H_n w_i$$

$$= H_n \sum_{S_i \in OPT} w_i$$

$$= H_n \cdot COST(OPT)$$

It remains to prove the claim.

$$\sum_{S_i \in OPT} c(e) \leq H_{|S_i|} \cdot w_i$$

S_i: $\{e_1, e_2, \cdots, e_d\}$ $d = |S_i|$

ordered exactly as they are covered by our alg.

e_1 was covered by some set S_k by our alg

$$\text{cost}(e_i) = \frac{w_t}{|S_t \cap R|} \le \frac{w_i}{|S_i \cap R|} \le \frac{w_i}{d-i+1}$$

R ↳ set of uncovered elements at that time

$$\sum_{i \in [d]} \text{cost}(e_i) \le \sum_{i \in [d]} \frac{w_i}{d-i+1}$$

$$= w_i \sum_{i \in [d]} \frac{1}{d-i+1}$$

$$= w_i \left(1 + \frac{1}{2} + \cdots + \frac{1}{d} \right)$$

$$= w_i H_d.$$