APPROXIMATION ALGORITHMS

RANDOM SAMPLING & RANDOMIZED ROUNDING OF LPS

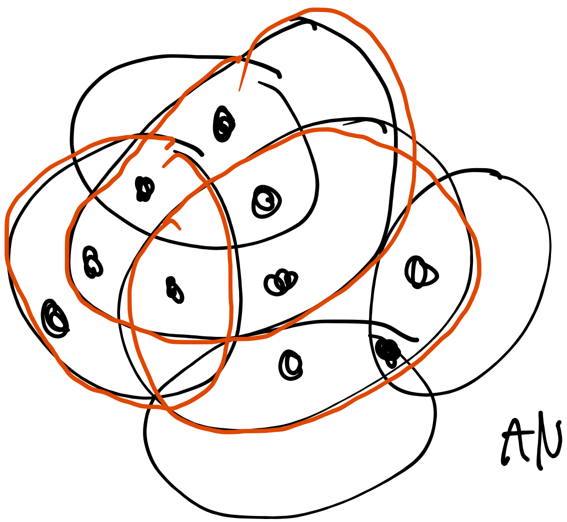
RASMUS PAGH

UNIVERSITY OF COPENHAGEN

TODAY

- CASE STUDY: WEIGHTED SET COVER
- CASE STUDY: WEIGHTED MAX SAT
  - RANDOM ASSIGNMENT BASELINE
  - DERANDOMIZING
  - SOLUTION BY RANDOMIZED ROUNDING OF LP

WEIGHTED SET COVER



FIND SUBCOLLECTION THAT "COVERS" ALL POINTS AND HAS MIN. WEIGHT

INPUT:

$S_1, \dots, S_m \subseteq \{1, \dots, n\}$

$w_1, \dots, w_m \in \mathbb{R}$

OBJECTIVE:

MINIMIZE $\sum_{i=1}^{m} w_i x_i$ ← INDICATES IF S_i IS IN SUBCOLL.

UNDER $\sum_{i \in B_j} x_i \geq 1$,

WHERE $B_j = \{ i \mid j \in S_i \}$

$\cancel{x_i \in \{0, 1\}}$ ← IP FORM.

$x_i \in [0, 1]$ ← LP RELAX.

OPTIMAL SOLUTION x^*

DET. ROUNDING:

$$\hat{x}_i = \begin{cases} 1 & \text{if } x_i^* \geq 1/f \\ 0 & \text{otherwise} \end{cases}$$

$f = \max_j |B_j|$, LARGEST #SETS COVERING POINT.

"WORST ROUNDING" $1/f \to 1$, GIVES f-APPROX.

RANDOMIZED ROUNDING:

$$\hat{x}_i = \begin{cases} 0 & \text{with prob. } (1 - x_i^*)^{5 \ln(n)} \\ 1 & \text{otherwise} \end{cases}$$

ANALYSIS OF SET COVER RANDOMIZED ROUNDING

RANDOMIZED ROUNDING:

$$\hat{X}_i = \begin{cases} 0 & \text{WITH PROB. } (1-x_i^*)^{5\ln(n)} > 1 - x_i^* \cdot 5\ln(n) \\ 1 & \text{OTHERWISE} \end{cases}$$

KNOW THAT $\sum_{i \in B_j} x_i^* \geq 1$

EXERCISE

1) VALID COVERING?

$$\Pr\left[\sum_{i \in B_j} \hat{x}_i = 0\right] \overset{\text{INDEP.}}{=} \prod_{i \in B_j} \Pr[\hat{X}_i = 0] = \left[\prod_{i \in B_j} (1-x_i^*)\right]^{5\ln(n)} < (1/e)^{5\ln(n)} < \frac{1}{n^5}$$

$$\Pr\left[\exists j : \sum_{i \in B_j} \hat{X}_i = 0\right] \leq \sum_{j=1}^{n} \Pr\left[\sum_{i \in B_j} \hat{X}_i\right] < 1/n^4.$$

2) ~~EXPECTED~~ APPROXIMATION FACTOR?

$$E\left[\sum_i w_i \hat{X}_i\right] \overset{\text{LINEARITY OF EXPECTATION}}{=} \sum_i \Pr[\hat{X}_i = 1] w_i < \sum_i x_i^* \cdot 5\ln(n) w_i = 5\ln(n) \underbrace{\sum_i x_i^* w_i}_{\text{LP LOWER BOUND}} \overset{\text{LP RELAX.}}{\leq} 5\ln(n) \cdot OPT$$

$\uparrow$ $O(\log n)$-APPR.

CAN REPEAT UNTIL WITHIN FACTOR $10\ln(n)$ OF

$O(1)$ ATTEMPTS IN EXPECT. (MARKOV'S INEQ.)

MAX SAT

$$\underbrace{\underbrace{(x_3 \vee x_7 \vee \bar{x}_{13})}_{\text{CLAUSE}},\ (\quad),\ (\quad)}_{m \text{ CLAUSES}}$$

WEIGHT $w_1, w_2, \dots, w_m$

SATISFIED IF $x_3 = \text{true}$, $x_7 = \text{true}$ OR $x_{13} = \text{false}$

$\boxed{\text{DEFINE } l_i \text{ AS \# LITERALS IN CLAUSE } i}$

GOAL: CHOOSE $x_1, \dots, x_n$ TO MAXIMIZE $\sum_{i=1}^{m} w_i y_i$

~~WHERE~~ $y_i = \begin{cases} 1 & \text{IF CLAUSE } i \text{ IS SATISFIED} \\ 0 & \text{OTHERWISE} \end{cases}$

BASELINE: ASSIGN $x_1, \dots, x_n \in \{\overset{\text{false}}{0}, \overset{\text{true}}{1}\}$ i.i.d. AT RANDOM.

$$E\left[\sum w_i y_i\right] = \sum_{i=1}^{m} w_i \Pr[\text{CLAUSE } i \text{ IS SATISFIED}]$$

$$= \sum_{i=1}^{m} w_i (1 - 2^{-l_i}) \geq \frac{1}{2} \sum_{i=1}^{m} w_i \geq \frac{1}{2} \cdot OPT.$$

DERANDOMIZATION

GOAL: GET $\sum_{i=1}^{m} w_i y_i \geq \frac{1}{2} \cdot \sum_{i=1}^{m} w_i \geq \frac{1}{2} OPT$ WITHOUT MAKING RANDOM CHOICES.

IDEA: CONDITIONAL EXPECTATIONS; COMPUTE

$$E\left[\sum_{i=1}^{m} w_i y_i \mid x_1 = 0\right]$$
$$E\left[\sum_{i=1}^{m} w_i y_i \mid x_1 = 1\right]$$

CHOOSE VALUE OF x_1 THAT MAXIMIZES EXPECTATION, REPEAT.

DOES NOT DECREASE EXPECTATION:

$$E\left[\sum_{i=1}^{m} w_i y_i\right] = \frac{1}{2} \cdot \underbrace{E\left[\sum_{i=1}^{m} w_i y_i \mid x_1 = 0\right] + \frac{1}{2} E\left[\sum_{i=1}^{m} w_i y_i \mid x_1 = 1\right]}_{\text{CANNOT BOTH BE SMALLER}}$$

RANDOMIZED ROUNDING

P_j: POSITIVE VARIABLES IN CLAUSE j

N_j: NEGATIVE —— || ——

HELPER VARIABLE $z_j \in \{0,1\}$: IS CLAUSE j SATISFIED?

~~INTEGER~~ LINEAR PROGRAM:

MAXIMIZE $\sum_{j=1}^{m} w_j z_j$

SUBJ. TO $\cancel{x_i, z_j \in \{0,1\}}$

$$\sum_{i \in P_j} x_i + \sum_{i \in N_j} (1-x_i) \geq z_j, \ \forall j$$

$$x_i \in [0,1], \quad z_j \in [0,1]$$

INTERPRET AS PROB. OF SETTING $x_i = 1$

INTERPRET AS PROBABILITY CLAUSE j IS TRUE

RANDOMIZED ROUNDING ALGORITHM:

$$\hat{x}_i = \begin{cases} 1 & \text{WITH PROB. } x_i^* \\ 0 & \text{OTHERWISE} \end{cases}$$

OPTIMAL SOLUTION RELAXED LP

ANALYSIS:

A.G MEAN INEQ. LP CONSTRAINT

$$\Pr[\overbrace{\text{CLAUSE } j \text{ NOT SAT.}}^{Y_j = 0}] = \prod_{i \in P_j} (1-x_i^*) \prod_{i \in N_j} x_i^* \leq \left[\frac{1}{l_j}\left(\sum_{i \in P_j}(1-x_i^*) + \sum_{i \in N_j} x_i^*\right)\right]^{l_j} \leq \left(1 - \frac{z_j^*}{l_j}\right)^{l_j}$$

$z_j^* \leq 1$

$$E\left[\sum_{j=1}^{m} w_j Y_j\right] = \sum_{j=1}^{m} w_j \Pr[Y_j = 1] \geq \sum_{j=1}^{m} w_j \left(1 - \left(1 - \frac{z_j^*}{l_j}\right)^{l_j}\right) \geq \sum_{j=1}^{m} z_j^* w_j \left(1 - \frac{1}{e}\right) \geq \left(1 - \frac{1}{e}\right) \cdot OPT$$