

Randomized Rounding

Previously

Probability

Events, probabilities,
conditional probabilities,
union bound

Random variables,
expected value,
linearity of expectation
Markov's inequality

Algorithms

Approx. 3-SAT,

Big idea: we only want the average

$$\begin{aligned} \text{Linearity of Expectation} \\ E[X+Y] &= E[X] + E[Y] \\ \text{"average sum = sum of averages"} \end{aligned}$$

let $Y_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{otherwise} \end{cases}$ ($i=1, \dots, m$)

$$E[Y_i] = 7/8 \quad (1 \text{ clause case})$$

$$E[\text{\# clauses satisfied}] = E\left[\sum_{i=1}^m Y_i\right] = \sum_{i=1}^m E[Y_i] = 7/8 m$$

let $Y_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{otherwise} \end{cases}$

$$Z = \sum_{i=1}^m Y_i = \text{\# clauses satisfied}$$

$$E[Z] = \frac{1}{2} E[Z | x_1 = \text{true}] + \frac{1}{2} E[Z | x_1 = \text{false}]$$

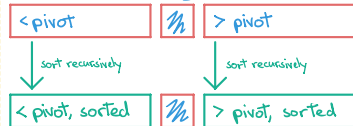
$\Rightarrow$ either $E[Z | x_1 = \text{true}]$ or $E[Z | x_1 = \text{false}]$ is $\geq E[Z]$!

QuickSort,

Random divide and conquer

① pick random 'pivot'

② Divide A into two parts



Expected running time of QuickSort

For each $i, j \in [n]$ w/ $i < j$,

$X_{ij} = \begin{cases} 1 & \text{if } i\text{th smallest element is compared to } j\text{th smallest} \\ 0 & \text{otherwise} \end{cases}$

$\sum_{i < j} X_{ij}$ = total # comparisons made by the algorithm
 $\approx$ total running time

$E[\sum_{i < j} X_{ij}]$ = average running time

$E[\sum_{i < j} X_{ij}] = \sum_{i < j} E[X_{ij}]$

now analyze $E[X_{ij}]$ in isolation

$$E[X_{ij}] = \Pr[X_{ij} = 1] = \Pr[i\text{th element is compared to } j\text{th}]$$

i th and j th elem selected before any element in between happens w/ probability $\frac{2}{j-i+1}$

$$E[X_{ij}] = \frac{2}{j-i+1}$$

by linearity of expectation

$$E[\sum_{i < j} X_{ij}] = \sum_{i < j} \frac{2}{j-i+1}$$

$$= 2 \sum_{i=1}^n \sum_{j=i+1}^n \frac{1}{j-i+1}$$

$$= 2 \sum_{k=1}^{n-1} \sum_{i=1}^{n-k} \frac{1}{k}$$

$$= 2 \sum_{k=1}^{n-1} (n-k) \frac{1}{k}$$

$$= 2n \sum_{k=1}^{n-1} \frac{1}{k} - 2 \sum_{k=1}^{n-1} 1$$

$$= 2n H_{n-1} - 2(n-1)$$

$$= 2n H_{n-1} - 2n + 2$$

$$= 2n H_{n-1} - 2n + 2$$

$$= 2n H_{n-1} - 2n + 2$$

$$= 2n H_{n-1} - 2n + 2$$

Quickselect (HW)

today

Probability (& techniques)

Approximations via
linear programming
& randomized rounding

Algorithms

Max Sat

Set Cover

Covering
Integer Programs

today

Probability (& techniques)

Approximations via
linear programming
& randomized rounding

Algorithms

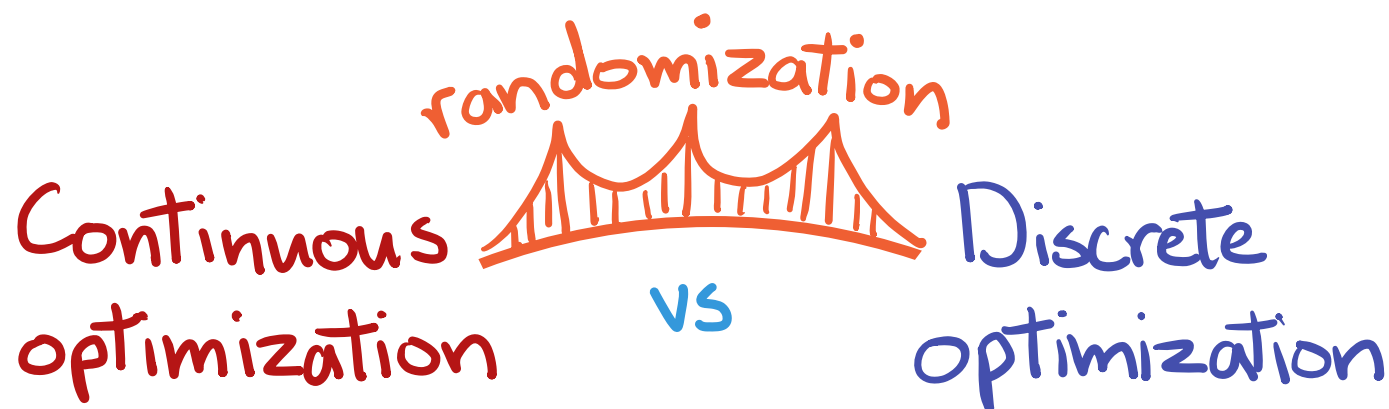
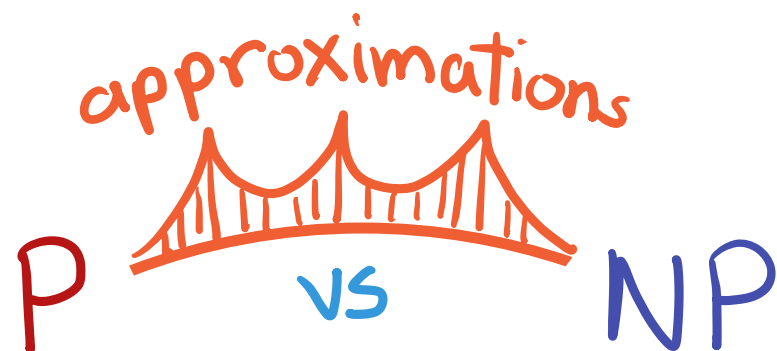
Max Sat

Set Cover

Covering
Integer Programs

Randomized Rounding

Story of



Integer Programs (IPs)

min/max $\sum_{j=1}^n c_j x_j$

linear objective

over $x \in \mathbb{Z}^n$

integral variables

s.t. $A_1 x \leq b_1$
 $A_2 x = b_2$
 $A_3 x \geq b_3$

linear
constraints

for fixed vectors c, b_1, b_2, b_3 and matrices A_1, A_2, A_3

NP-Hard

Linear programs (LPs)

similar to IPs, except no integrality constraints

$$\text{min/max} \quad \sum_{j=1}^n c_j x_j$$

linear objective

$$\text{Over } x \in \mathbb{R}^n$$

~~integral~~ ^{real} variables

$$\text{s.t.} \quad A_1 x \leq b_1$$

$$A_2 x = b_2$$

$$A_3 x \geq b_3$$

linear
constraints

for fixed vectors c, b_1, b_2, b_3 and matrices A_1, A_2, A_3

Linear programs (LPs)

similar to IPs, except no integrality constraints

$$\min/\max \quad \sum_{j=1}^n c_j x_j$$

linear objective

$$\text{Over } x \in \mathbb{R}^n$$

~~linear~~
~~integral~~ variables

$$\text{s.t. } A_1 x \leq b_1$$

$$A_2 x = b_2$$

$$A_3 x \geq b_3$$

linear
constraints

for fixed vectors c, b_1, b_2, b_3 and matrices A_1, A_2, A_3

"linear program"

$$\max \overbrace{\sum_{(s,w) \in E} f(s,w) - \sum_{(u,s) \in E} f(u,s)}^{\text{size of flow}} \quad \text{over } f \in \mathbb{R}^E$$

"capacity constraints"

$$\text{s.t. } 0 \leq f(e) \leq c(e) \quad \forall e \in E$$

"conservation of flow"

$$\sum_{(u,v) \in E} f(u,v) = \sum_{(v,w) \in E} f(v,w) \quad \forall v \notin \{s,t\}$$

"s-t flow" $f: E \rightarrow \mathbb{R}_{\geq 0}$ s.t.

(a) "capacity constraints"

$$0 \leq f(e) \leq c(e) \quad \forall e \in E$$

(b) "conservation of flow"

$$\sum_{(u,v) \in E} f(u,v) = \sum_{(v,w) \in E} f(v,w) \quad \forall v \notin \{s,t\}$$



"linear program" because objective, constraints, all linear functions of f

Linear Programs are in P (!!)

↑
(weakly polynomial)

1947 - Simplex algorithm
- Dantzig

(not polytime but practical)

1979 - Ellipsoid algorithm
- Khachiyan

polytime

1984 - Interior point algorithm
- Karmarkar

what's the difference?

Discrete
optimization

vs

Continuous
optimization

Integer
Programs

vs

Linear
Programs



NP-Hard



polynomial time!

LP-Based Approximations

① model as IP

$$\begin{array}{ll}\max & \sum_{j=1}^n c_j x_j \\ \text{Over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1\end{array}$$

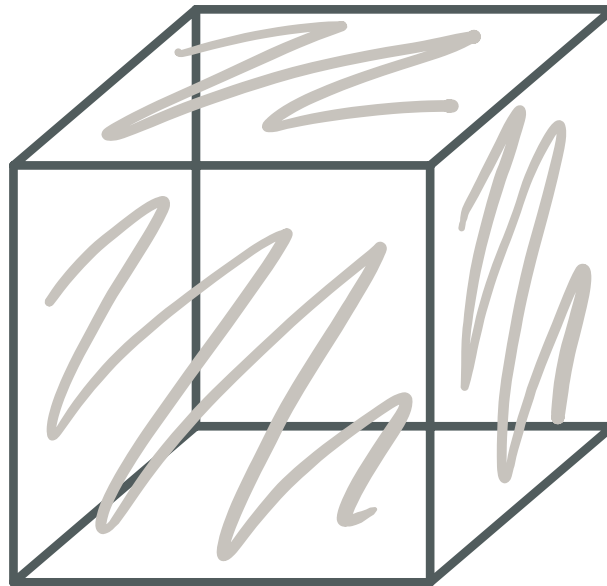


$$\begin{array}{ll}\max & \sum_{j=1}^n c_j x_j \\ \text{Over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1\end{array}$$

② relax to LP



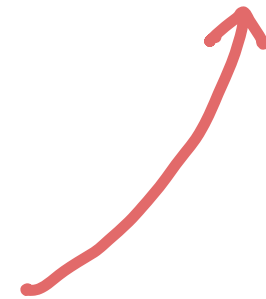
LP solver



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$



today

Probability (& techniques)

Approximations via
linear programming
& randomized rounding

Algorithms

Max Sat

Set Cover

Covering
Integer Programs

3-SAT

"conjunctive normal form"

CNF formula

$$\varphi(x_1, \dots, x_n) = (x_1 \vee \bar{x}_2 \vee x_3) \wedge (x_4 \vee \dots) \wedge \dots$$

m clauses, n variables

3 (distinct) variables per clause

The silliest APX algorithm

1. for $i = 1, \dots, n$

a. randomly set $x_i = \begin{cases} \text{TRUE} & \text{w/ prob .5} \\ \text{FALSE} & \text{w/ prob .5} \end{cases}$

We will:

① prove silly algo is $\frac{7}{8}$ -APX

② derandomize silly algo $\Rightarrow$ deterministic $\frac{7}{8}$ -APX

Even more surprising:

"probabilistically checkable proofs"

PCP Theorem

$(\frac{7}{8} + \epsilon)$ -APX is NP-Hard $\forall \epsilon > 0$

Approximating 3SAT: "Max 3-SAT"

$$\varphi(x_1, \dots, x_n) = (x_1 \vee \bar{x}_2 \vee x_3) \wedge (x_4 \vee \dots) \wedge \dots$$

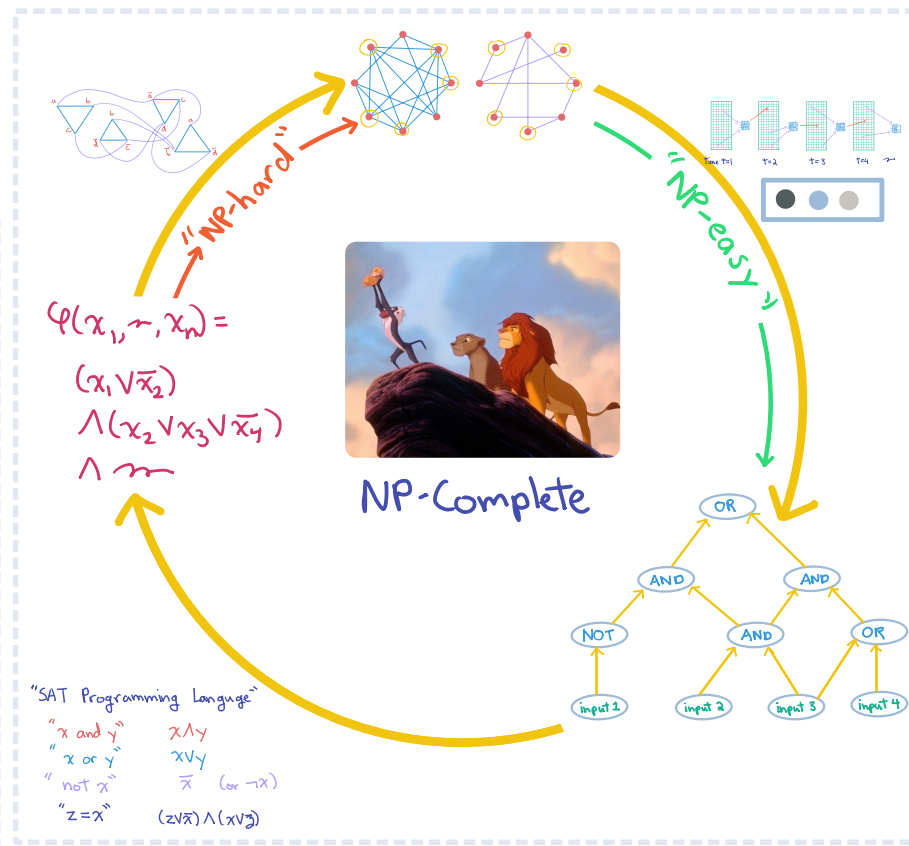
Goal: satisfy Max # clauses

Interesting even if φ is unsatisfiable

"d-Approximation algorithm" $d \in (0, 1)$

Algorithm satisfies $d \cdot \text{OPT}$ clauses

Optimum (max) value



2 disjoint clauses

① prove silly algo is $7/8$ -APX

1. for $i=1, \dots, n$ The silliest APX algorithm
 a. randomly set $x_i = \begin{cases} \text{TRUE} & \text{w/ prob. } .5 \\ \text{FALSE} & \text{w/ prob. } .5 \end{cases}$

$$\mathcal{S}(x) = (\underbrace{x_1 \vee x_2 \vee \bar{x}_3}_{C_1}) \wedge (\underbrace{x_4 \vee \bar{x}_5 \vee x_6}_{C_2})$$

E_1 = event that C_1 satisfied

E_2 = ~ C_2 ~

$$\Pr[E_1, E_2] = \frac{1}{8} \cdot \frac{1}{8}$$

$$= \Pr[E_1] \cdot \Pr[E_2]$$

2 clauses w/
1 common variable

① prove silly algo is $7/8$ -APX

1. for $i=1, \dots, n$ The silliest APX algorithm
 a. randomly set $x_i = \begin{cases} \text{TRUE} & \text{w/ prob. } .5 \\ \text{FALSE} & \text{w/ prob. } .5 \end{cases}$

$$\mathcal{S}(x) = (\underbrace{x_1 \vee \bar{x}_2 \vee x_3}_{C_1}) \wedge (\underbrace{x_3 \vee x_4 \vee \bar{x}_5}_{C_2})$$

E_1 = event that C_1 satisfied

E_2 = ~ C_2 ~

$$\Pr[E_1, E_2] =$$

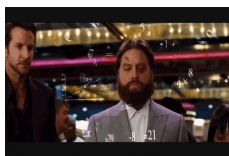
$$\underbrace{\Pr[E_1, E_2 | x_3 = T] \cdot \frac{1}{2}}_1 + \underbrace{\Pr[E_1, E_2 | x_3 = F] \cdot \frac{1}{2}}_{(\frac{3}{4})^2}$$

m clauses w/
common variables

① prove silly algo is $7/8$ -APX

1. for $i=1, \dots, n$ The silliest APX algorithm
 a. randomly set $x_i = \begin{cases} \text{TRUE} & \text{w/ prob. } .5 \\ \text{FALSE} & \text{w/ prob. } .5 \end{cases}$

$$\mathcal{S}(x) = (\underbrace{x_1 \vee \bar{x}_2 \vee x_3}_{C_1}) \wedge (\underbrace{x_3 \vee x_4 \vee \bar{x}_5}_{C_2}) \wedge (\underbrace{x_2 \vee \bar{x}_4 \vee x_6}_{C_3}) \wedge \dots$$



E_1 = event that C_1 satisfied

E_2 = ~ C_2 ~

E_3 = ~ C_3 ~

⋮

general question

how to analyze a randomized mechanism
w/ a combinatorial explosion of possibilities?

Big idea: we only want the average

Linearity of Expectation

$$E[X+Y] = E[X] + E[Y] \quad (\text{ridiculously useful})$$

"average sum = sum of averages"

let $Y_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{otherwise} \end{cases}$

($i=1, \dots, m$)

$$E[Y_i] = 7/8 \quad (1 \text{ clause case})$$

$$E[\text{\# clauses satisfied}] = E\left[\sum_{i=1}^m Y_i\right] = \sum_{i=1}^m E[Y_i] = 7/8 m$$

(actually easy!)

Max-SAT

"conjunctive normal form"

CNF formula

$$\varphi(x_1, \dots, x_n) = (x_1 \vee \bar{x}_2 \vee x_3) \wedge (x_4 \vee \dots) \wedge \dots$$

m clauses, n variables

any # of variables per clause

goal: satisfy max # clauses

Approximation ratios for k-SAT vs max-SAT

k-sat

max-SAT

k oblivious APX

$$1 \quad 1/2$$

$$2 \quad 3/4$$

$$3 \quad 7/8$$

$$4 \quad 15/16$$

$$5 \quad 31/32$$

$\vdots$

$$k \quad 1 - 1/2^k$$

$$1/2$$

LP-Based Approximations

① model as IP

$$\begin{array}{ll}\max & \sum_{j=1}^n c_j x_j \\ \text{Over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1\end{array}$$

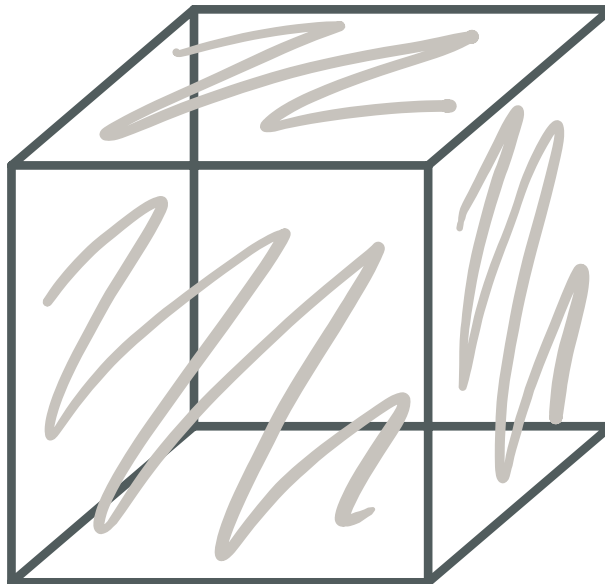
approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

LP solver

$$\begin{array}{ll}\max & \sum_{j=1}^n c_j x_j \\ \text{Over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1\end{array}$$

② relax to LP



optimal
LP solution
 $x \in \mathbb{R}^n$

IP for max-SAT

$$\max \sum_{i=1}^m z_i$$

$$\text{over } \gamma \in \{0,1\}^n, z \in \{0,1\}^m$$

$$\text{s.t. } \sum_{j: x_j \in C_i} \gamma_j + \sum_{j: \bar{x}_j \in C_i} (1 - \gamma_j) \geq z_i$$

for each clause C_i

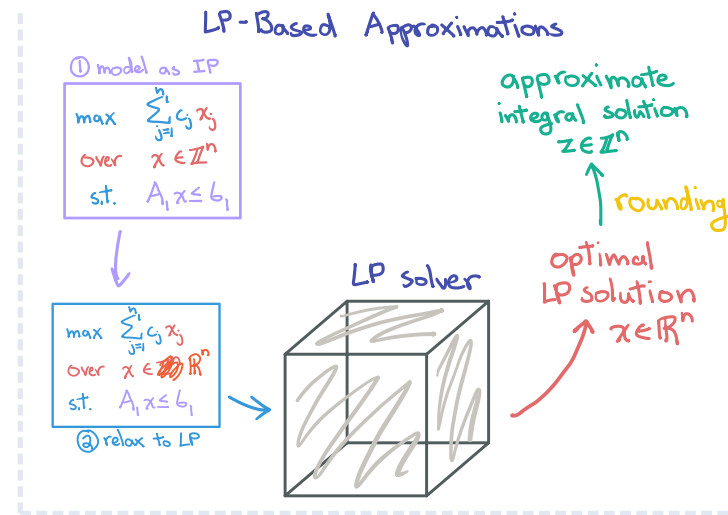
$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

$$\gamma_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

e.g. $C_i = x_1 \vee \bar{x}_2 \vee x_3$

$\Rightarrow$ constraint

$$z_i \leq \gamma_1 + 1 - \gamma_2 + \gamma_3$$



IP for max-SAT

$$\max \sum_{i=1}^m z_i$$

$$\text{over } \gamma \in \{0,1\}^n, z \in \{0,1\}^m$$

$$\text{s.t. } \sum_{j: x_j \in C_i} \gamma_j + \sum_{j: \bar{x}_j \in C_i} (1 - \gamma_j) \geq z_i$$

for each clause C_i

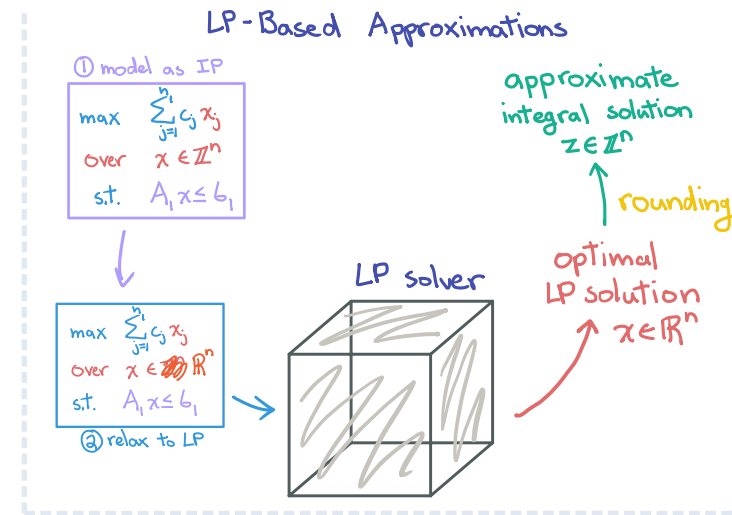
$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

$$\gamma_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

e.g. $C_i = x_1 \vee \bar{x}_2 \vee x_3$

$\Rightarrow$ constraint

$$\gamma_1 + (1 - \gamma_2) + \gamma_3 \geq z_1$$



LP for max-SAT

$$\max \sum_{i=1}^m z_i$$

$$\text{over } y \in [0,1]^n, z \in [0,1]^m$$

$$\text{s.t. } \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1 - y_j) \geq z_i \quad \text{for each clause } C_i$$

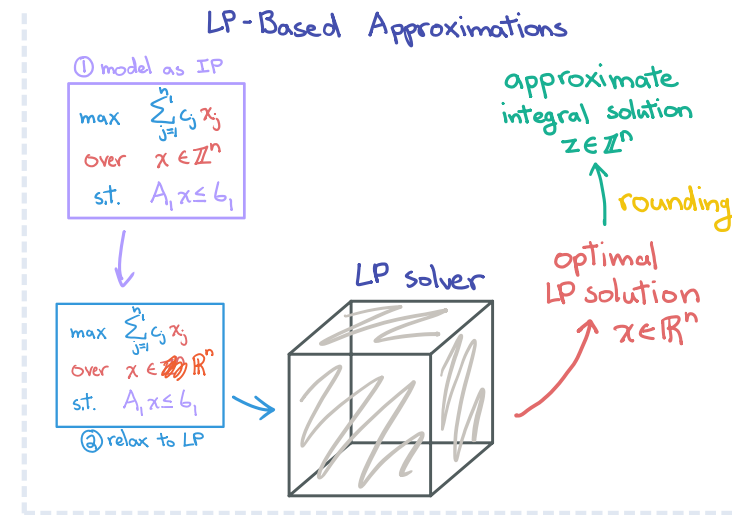
$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

$$\text{e.g. } C_i = x_1 \vee \bar{x}_2 \vee x_3$$

$\Rightarrow$ constraint

$$y_1 + (1 - y_2) + y_3 \geq z_1$$



IP for max-SAT

$$\max \sum_{i=1}^m z_i$$

over $\gamma \in \{0,1\}^n, z \in \{0,1\}^m$

$$\text{s.t. } \sum_{j: x_j \in C_i} \gamma_j + \sum_{j: \bar{x}_j \in C_i} (1 - \gamma_j) \geq z_i$$

NP-Hard

LP for max-SAT

$$\max \sum_{i=1}^m z_i$$

over $\gamma \in [0,1]^n, z \in [0,1]^m$

$$\text{s.t. } \sum_{j: x_j \in C_i} \gamma_j + \sum_{j: \bar{x}_j \in C_i} (1 - \gamma_j) \geq z_i$$

Solvable

$$\text{OPT}_{\text{IP}} \leq \text{OPT}_{\text{LP}}$$

LP-Based Approximations

① model as IP

$$\begin{array}{ll}\max & \sum_{j=1}^n c_j x_j \\ \text{Over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1\end{array}$$

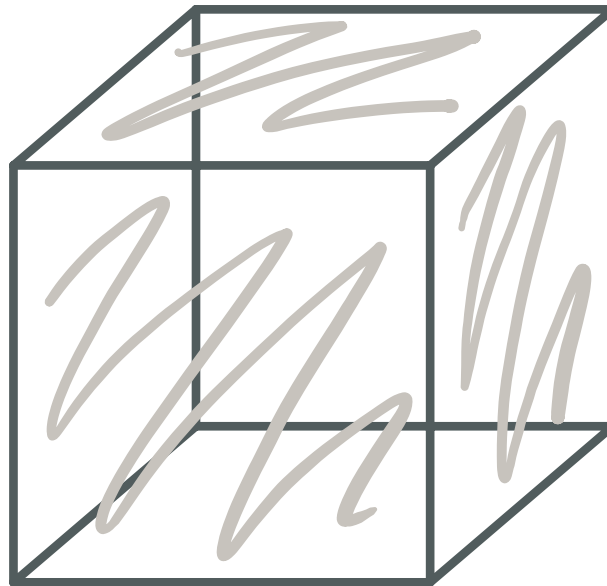


$$\begin{array}{ll}\max & \sum_{j=1}^n c_j x_j \\ \text{Over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1\end{array}$$

② relax to LP



LP solver



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$



Rounding LP solution

$$y \in [0,1]^n \rightsquigarrow x \in \{\text{TRUE}, \text{FALSE}\}^n$$

ideas?

det. rounding

e.g., how to interpret $y_i = \frac{1}{2}$?

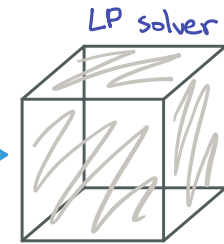
LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

↑ rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

$$\max \sum_{i=1}^m z_i$$

$$\text{over } y \in [0,1]^n, z \in [0,1]^m$$

$$\text{s.t. } \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1 - y_j) \geq z_i$$

for each clause C_i

$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

Rounding LP solution

$$y \in [0,1]^n \rightsquigarrow x \in \{\text{TRUE}, \text{FALSE}\}$$

ideas?

e.g., how to interpret $y_i = \frac{1}{2}$?

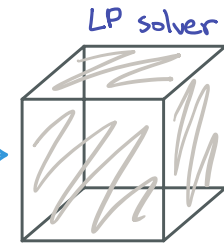
LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

$$\max \sum_{i=1}^m z_i$$

$$\text{over } y \in [0,1]^n, z \in [0,1]^m$$

$$\text{s.t. } \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1 - y_j) \geq z_i$$

for each clause C_i

$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

Rounding LP solution

$$y \in [0, 1]^n \rightsquigarrow x \in \{\text{TRUE}, \text{FALSE}\}$$

For each variable x_j , independently, randomly set $x_j = \text{t}$ with probability y_j , and $x_j = \text{f}$ otherwise.

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$



$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

③ relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

$$\max \sum_{i=1}^m z_i$$

$$\text{over } y \in [0, 1]^n, z \in [0, 1]^m$$

$$\text{s.t. } \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1 - y_j) \geq z_i$$

for each clause C_i

$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

For each variable x_j , independently, randomly set $x_j = t$ with probability y_j , and $x_j = f$ otherwise.

claim: each clause C_i is satisfied
w/ probability $(1 - 1/e) z_i$

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

$$\max \sum_{i=1}^m z_i$$

$$\text{over } y \in [0,1]^n, z \in [0,1]^m$$

$$\text{s.t. } \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1 - y_j) \geq z_i$$

for each clause C_i

$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

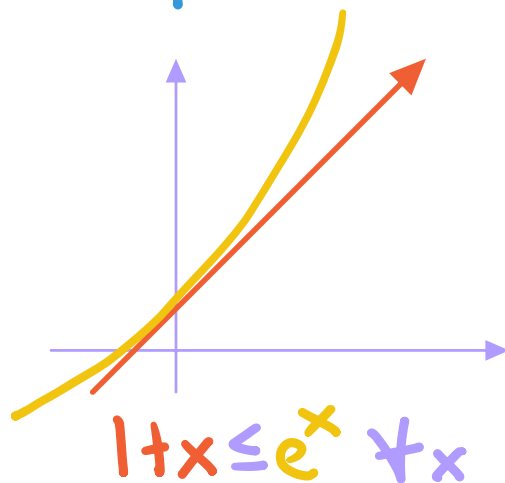
$$y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

For each variable x_j , independently, randomly set $x_j = t$ with probability y_j , and $x_j = f$ otherwise.

claim: each clause C_i is satisfied w/ probability $(1 - 1/e)z_i$

$P[C_i \text{ not satisfied}]$

$$\begin{aligned}
 &= \prod_{x_j \in C_i} (1 - y_j) \prod_{\bar{x}_j \in C_i} y_j \\
 &\leq \prod_{x_j \in C_i} e^{-y_j} \prod_{\bar{x}_j \in C_i} e^{-(1-y_j)} \\
 &= e^{-(\sum_{x_j \in C_i} y_j + \sum_{\bar{x}_j \in C_i} (1-y_j))} \leq e^{-z_i}
 \end{aligned}$$



$$e^{-z_i} \leq (1-z_i)e^0 + z_i e^1 = 1 - (1 - 1/e)z_i$$

LP-Based Approximations

① model as IP

$$\begin{aligned}
 &\max \sum_{j=1}^n c_j x_j \\
 &\text{over } x \in \mathbb{Z}^n \\
 &\text{s.t. } A_1 x \leq b_1
 \end{aligned}$$

$$\begin{aligned}
 &\max \sum_{j=1}^n c_j x_j \\
 &\text{over } x \in \mathbb{R}^n \\
 &\text{s.t. } A_1 x \leq b_1
 \end{aligned}$$

② relax to LP

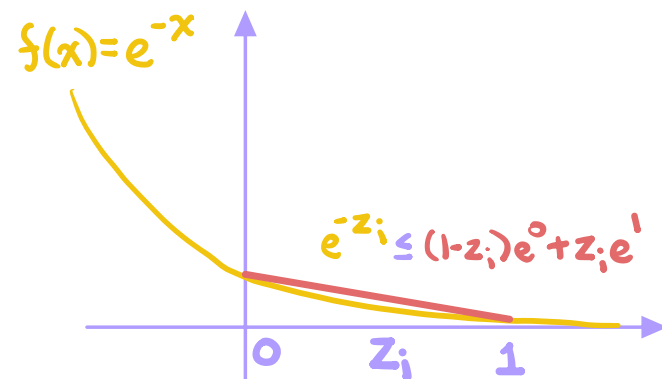


approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

$$\begin{aligned}
 &\max \sum_{i=1}^m z_i \\
 &\text{over } y \in [0,1]^n, z \in [0,1]^m \\
 &\text{s.t. } \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1-y_j) \geq z_i \\
 &\quad \text{for each clause } C_i \\
 &z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases} \\
 &y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}
 \end{aligned}$$



For each variable x_j , independently, randomly set $x_j = t$ with probability y_j , and $x_j = f$ otherwise.

claim: each clause C_i is satisfied w/ probability $(1 - 1/e) z_i$

claim:

$$E[\# \text{ sat. clauses}] \geq (1 - 1/e) \text{OPT}$$

$$A_i = \begin{cases} 1 & \text{if } i\text{th clause is sat} \\ 0 & \text{otherwise} \end{cases}$$

$$E[\sum A_i] = \sum E[A_i]$$

$$= (1 - 1/e) \sum_{i=1}^m z_i \geq (1 - 1/e) \text{OPT}$$

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_i x \leq b_i \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_i x \leq b_i \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

$$\begin{array}{ll} \max & \sum_{i=1}^m z_i \\ \text{over} & y \in [0, 1]^n, z \in [0, 1]^m \\ \text{s.t.} & \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1 - y_j) \geq z_i \end{array}$$

for each clause C_i

$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

For each variable x_j , independently, randomly set $x_j = \text{t}$ with probability y_j , and $x_j = \text{f}$ otherwise.

$$E[\# \text{ sat. clauses}] \geq (1 - 1/e) \text{OPT}$$

Better?

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

③ relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

↑ rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

$$\max \sum_{i=1}^m z_i$$

$$\text{over } y \in [0,1]^n, z \in [0,1]^m$$

$$\text{s.t. } \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1 - y_j) \geq z_i$$

for each clause C_i

$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

For each variable x_j , independently, randomly set $x_j = t$ with probability y_j , and $x_j = f$ otherwise.

claim: each clause C_i is satisfied w/ probability $(1 - 1/e) z_i$

what if C_i has 1 variable?

$$C_i = x_1 \quad P[\sim] = y_1 = z_1$$

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

↓

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

↑ rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

$$\max \sum_{i=1}^m z_i$$

$$\text{over } y \in [0,1]^n, z \in [0,1]^m$$

$$\text{s.t. } \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1 - y_j) \geq z_i$$

for each clause C_i

$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

For each variable x_j , independently, randomly set $x_j = t$ with probability y_j , and $x_j = f$ otherwise.

claim: each clause C_i is satisfied w/ probability $(1 - 1/e) z_i$

what if C_i has 1 variable?

$$P[\text{sat. } C_i] = z_i$$

~ 2 variables? e.g. $C_i = x_1 \vee x_2$

$$P[C_i \text{ not sat}] = (1 - y_1)(1 - y_2)$$

$$\leq \left(\frac{(1 - y_1) + (1 - y_2)}{2} \right)^2$$

AM-GM:

$$ab \leq \left(\frac{a+b}{2} \right)^2$$

$$\leq \left(1 - \frac{z_i}{2} \right)^2 \leq (1 - z_i) \left(1 - \frac{0}{2} \right)^2 + z_i \left(1 - \frac{1}{2} \right)^2$$

$$= 1 - z_i + z_i/4 = 1 - \frac{3}{4} z_i$$

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_i x \leq b_i \end{array}$$

② relax to LP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_i x \leq b_i \end{array}$$



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding
optimal
LP solution
 $x \in \mathbb{R}^n$

$$\max \sum_{i=1}^m z_i$$

$$\text{over } y \in [0, 1]^n, z \in [0, 1]^m$$

$$\text{s.t. } \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1 - y_j) \geq z_i$$

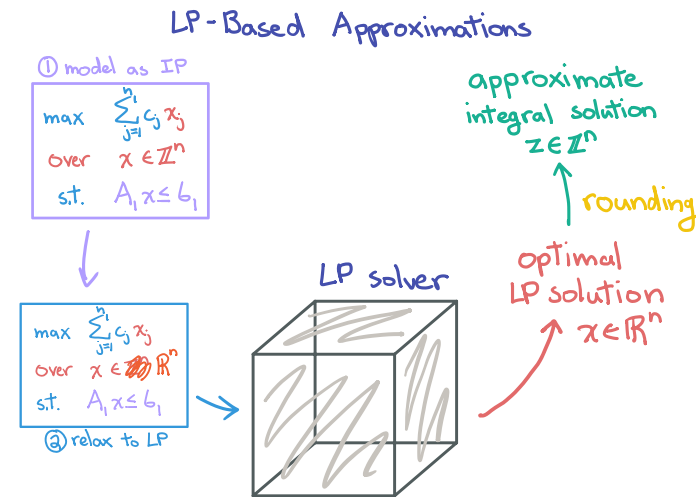
for each clause C_i

$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

For each variable x_j , independently, randomly set $x_j = \text{t}$ with probability y_j , and $x_j = \text{f}$ otherwise.

$$E[\# \text{ sat. clauses}] \geq (1 - 1/e) \text{OPT}$$



k	oblivious	LP	average
1	1/2	z_i	$\geq (3/4)z_i$
2	3/4	$(3/4)z_i$	$\geq (3/4)z_i$
3	7/8	$(1 - 1/e)z_i$	$\geq (3/4)z_i$
4	15/16	$(1 - 1/e)z_i$	$\geq (3/4)z_i$
5	31/32	$(1 - 1/e)z_i$	$\geq (3/4)z_i$
$\vdots$			
k	$1 - 1/2^k$	$(1 - 1/e)z_i$	$\geq (3/4)z_i$

$$\begin{array}{ll} \max & \sum_{i=1}^m z_i \\ \text{over} & y \in [0,1]^n, z \in [0,1]^m \\ \text{s.t.} & \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1 - y_j) \geq z_i \\ & \text{for each clause } C_i \end{array}$$

$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$

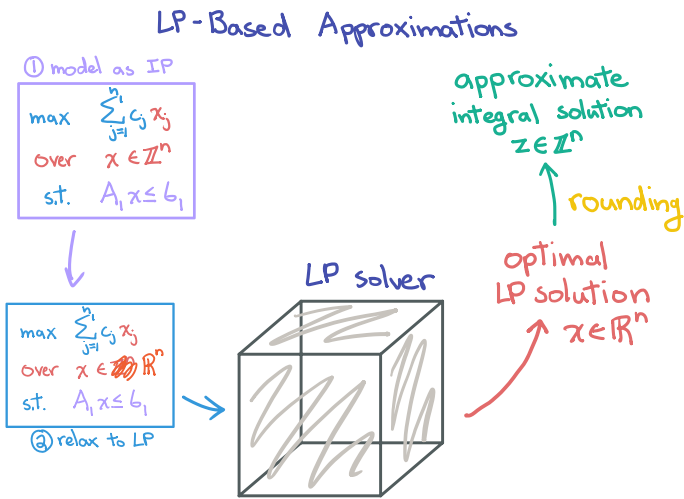
$y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$

Can we get $\frac{3}{4}$ -APX?

- ① Flip a coin
- ② If heads, do oblivious rounding
- ③ If tails, do LP rounding

For each variable x_j , independently, randomly set $x_j = t$ with probability y_j , and $x_j = f$ otherwise.

claim: each clause C_i is satisfied
w/ probability $\frac{3}{4} z_i$



$$\begin{array}{ll} \max & \sum_{i=1}^m z_i \\ \text{over} & y \in [0,1]^n, z \in [0,1]^m \\ \text{s.t.} & \sum_{j: x_j \in C_i} y_j + \sum_{j: \bar{x}_j \in C_i} (1 - y_j) \geq z_i \end{array}$$

for each clause C_i

$$z_i = \begin{cases} 1 & \text{if } i\text{th clause satisfied} \\ 0 & \text{if not} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if } x_j = \text{TRUE} \\ 0 & \text{if } x_j = \text{FALSE} \end{cases}$$

k	oblivious	LP	average
1	1/2	z_i	$\geq (3/4)z_i$
2	3/4	$(3/4)z_i$	$\geq (3/4)z_i$
3	7/8	$(1 - 1/e)z_i$	$\geq (3/4)z_i$
4	15/16	$(1 - 1/e)z_i$	$\geq (3/4)z_i$
5	31/32	$(1 - 1/e)z_i$	$\geq (3/4)z_i$
⋮			
k	$1 - 1/2^k$	$(1 - 1/e)z_i$	$\geq (3/4)z_i$

today

Probability (& techniques)

Approximations via
linear programming
& randomized rounding

Algorithms

Max Sat

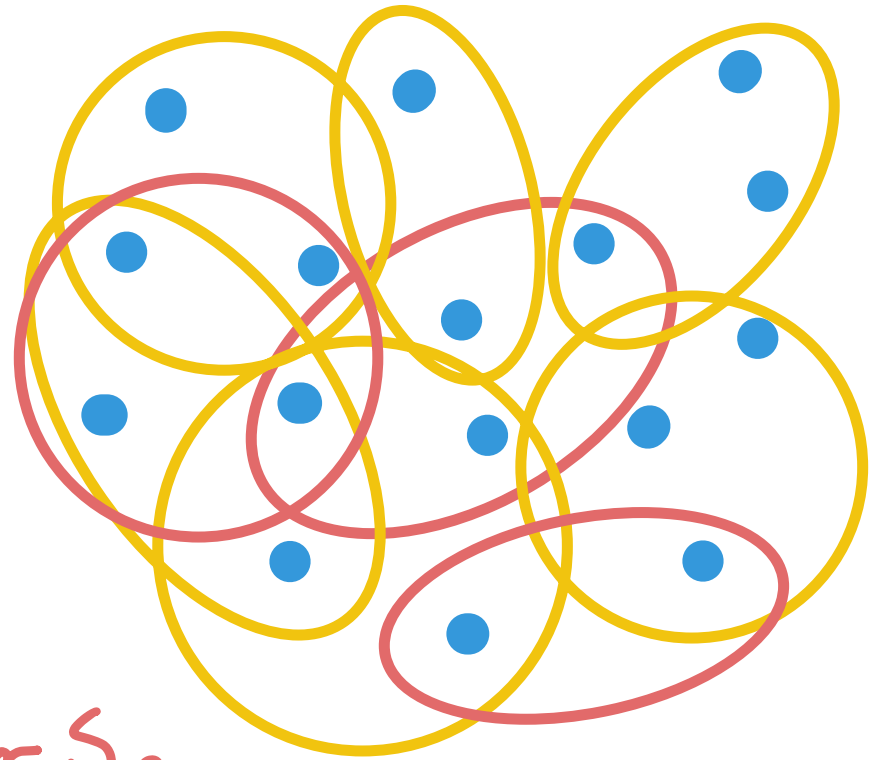
Set Cover

Covering
Integer Programs

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$



Goal: select k sets $S_{e_1}, S_{e_2}, \dots, S_{e_k}$
minimizing k ,

$$\text{s.t. } S_{e_1} \cup S_{e_2} \cup \dots \cup S_{e_k} = [m]$$

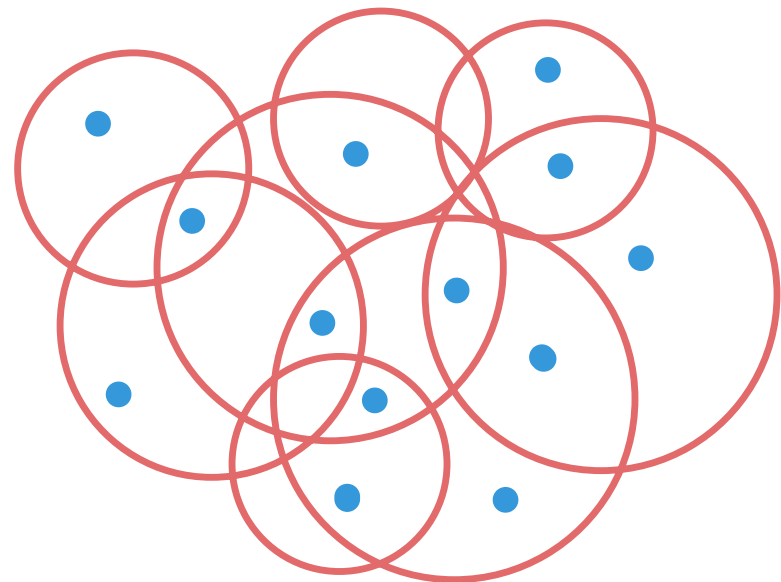
NP-Hard

$$\varphi(x_1, \dots, x_n) =$$

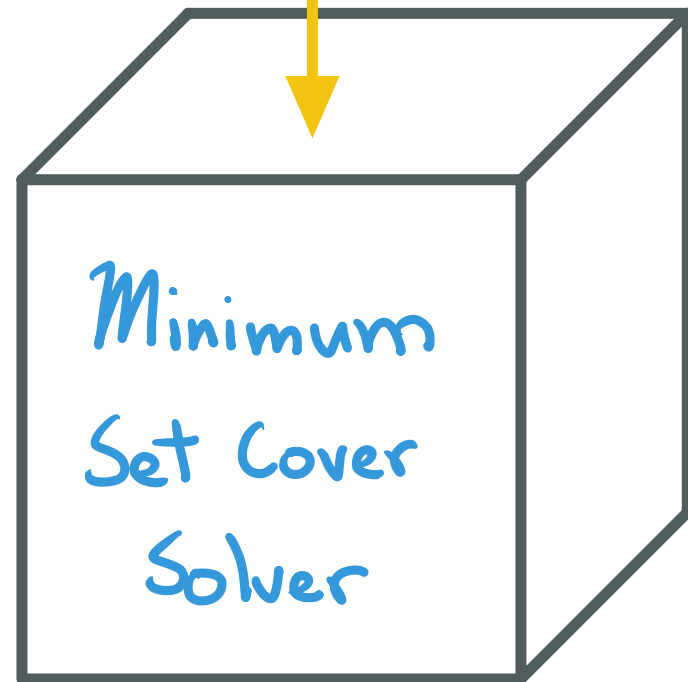
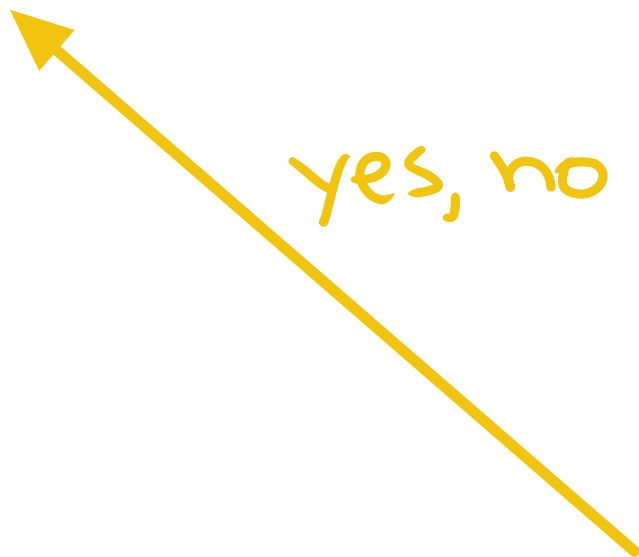
$$(x_1 \vee \bar{x}_2)$$

$$\wedge (x_2 \vee x_3 \vee \bar{x}_4)$$

$$\wedge m$$



yes, no



[pause here and try
to think of a
reduction from 3-SAT]

$$f(x) = (x_1 \vee \bar{x}_2 \vee x_3) \wedge (x_2 \vee x_4 \vee x_5) \wedge \dots$$

assign. "x₁ = true" → Set S₁
 "x₁ = false" → Set T₁

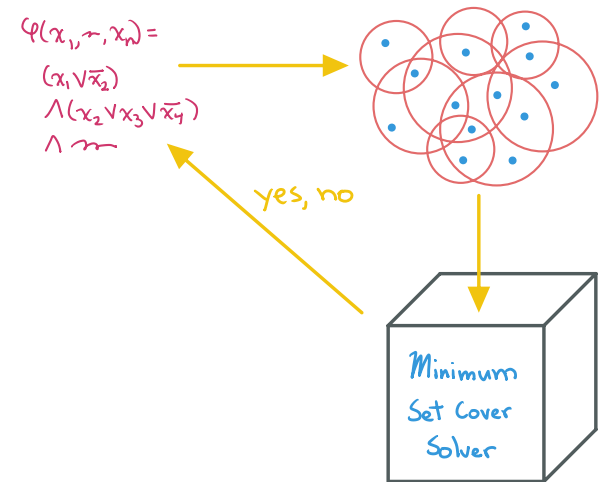
clauses → points

C_i → P_i

$$S_1 = \{P_i : "x_1" \text{ in } C_i\}$$

if min # set cover ≤ k, return "S satisfiable"

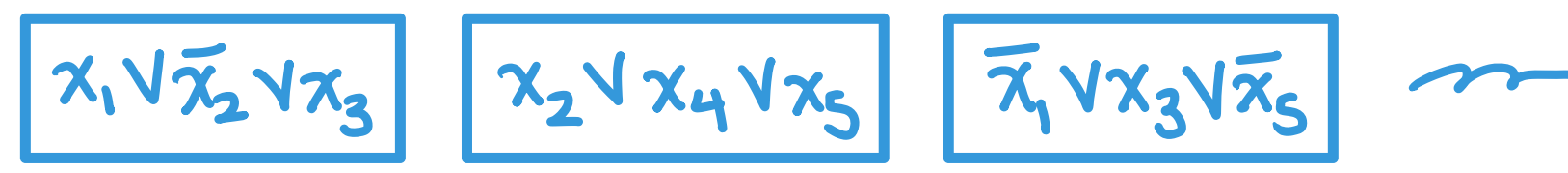
(n variables,
m clauses)



$$f(x) = (x_1 \vee \bar{x}_2 \vee x_3) \wedge (x_2 \vee x_4 \vee x_5) \wedge \dots$$

(n variables,
m clauses)

elements
per
clause



sets =
assignments



elements
per variable



f satisfiable $\Leftrightarrow$ set cover of size n

IP for set cover

$$\min \sum_{j=1}^n x_j$$

$$\text{over } x \in \{0, 1\}^n$$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \quad \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$



$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



optimal LP solution $x \in \mathbb{R}^n$

rounding

approximate integral solution $z \in \mathbb{Z}^n$

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$

minimizing k ,

s.t. $S_{e_1} \cup \dots \cup S_{e_k} = [m]$

LP for set cover

$$\min \sum_{j=1}^n x_j$$

$$\text{over } x \in [0,1]^n$$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \quad \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

LP-Based Approximations

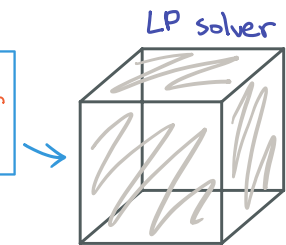
① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$



$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$

minimizing k ,

s.t. $S_{e_1} \cup \dots \cup S_{e_k} = [m]$

IP for set cover

$$\min \sum_{j=1}^n x_j$$

over $x \in \{0,1\}^n$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \\ \forall i \in [m]$$

NP-Hard

LP for set cover

$$\min \sum_{j=1}^n x_j$$

over $x \in [0,1]^n$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \\ \forall i \in [m]$$

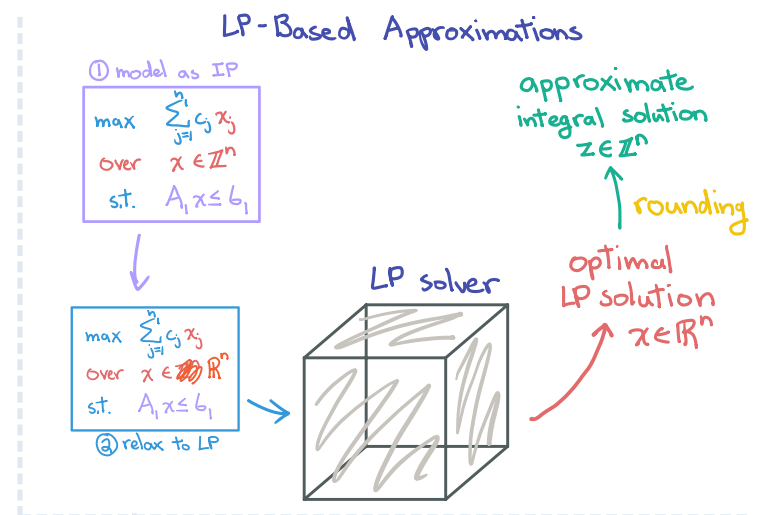
Polytime solvable

Rounding LP solution

$$x \in [0,1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$$

Ideas?

How to interpret $x_i = .5$?



Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$
minimizing k ,

$$\text{s.t. } S_{e_1} \cup \dots \cup S_{e_k} = [m]$$

$$\min \sum_{j=1}^n x_j$$

$$\text{over } x \in [0,1]^n$$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \quad \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

Rounding LP solution

$$x \in [0,1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$$

For each set S_j , independently, take S_j with probability x_j .

$$E[|F|] = \sum_j E[A_j] = \sum_j x_j = \text{OPT}_{LP}$$

$$A_j = \begin{cases} 1 & \text{if } S_j \in F \\ 0 & \text{otherwise} \end{cases}$$

LP-Based Approximations

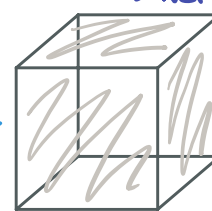
① model as IP

$$\begin{aligned} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_j x \leq b_j \end{aligned}$$



$$\begin{aligned} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_j x \leq b_j \end{aligned}$$

② relax to LP



LP solver

approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

Minimum Set Cover
 m points $[m] \equiv \{1, \dots, m\}$
 n sets $S_1, \dots, S_n \subseteq [m]$
 Goal: select k sets $S_{e_1}, \dots, S_{e_k}$
 minimizing k ,
 s.t. $S_{e_1} \cup \dots \cup S_{e_k} = [m]$

$$\begin{aligned} \min & \sum_{j=1}^n x_j \\ \text{over} & x \in [0,1]^n \\ \text{s.t.} & \sum_{j: i \in S_j} x_j \geq 1 \\ & \forall i \in [m] \end{aligned}$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

Rounding LP solution

$$x \in [0,1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$$

For each set S_j , independently, take S_j with probability x_j .

$$E[|F|] = \sum_j x_j = \text{OPT} \quad (\text{good!})$$

Fix a point i .

$$\begin{aligned} E[\text{sets in } F \text{ covering } i] &= E\left[\sum_{S_j \ni i} A_j\right] \\ &= \sum_{S_j \ni i} E[A_j] = \sum_{S_j \ni i} x_j \geq 1 \end{aligned}$$

$$A_j = \begin{cases} 1 & \text{if } S_j \in F \\ 0 & \text{otherwise} \end{cases}$$

LP-Based Approximations

① model as IP

$$\begin{aligned} \max \quad & \sum_{j=1}^n c_j x_j \\ \text{over } & x \in \mathbb{Z}^n \\ \text{s.t. } & A_j x \leq b_j \end{aligned}$$



$$\begin{aligned} \max \quad & \sum_{j=1}^n c_j x_j \\ \text{over } & x \in \mathbb{R}^n \\ \text{s.t. } & A_j x \leq b_j \end{aligned}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$
minimizing k ,

$$\text{s.t. } S_{e_1} \cup \dots \cup S_{e_k} = [m]$$

$$\min \sum_{j=1}^n x_j$$

over $x \in [0,1]^n$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \quad \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

Rounding LP solution

$$x \in [0,1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$$

For each set S_j , independently, take S_j with probability x_j .

$$E[|F|] = \sum_j x_j = \text{OPT} \quad (\text{good!})$$

Fix a point i .

$$E[\text{sets in } F \text{ covering } i] = \sum_{j: i \in S_j} x_j \geq 1$$

$$P[F \text{ not cover } i] = \prod_{S_j \ni i} (1 - x_j) \leq e^{-\sum_{S_j \ni i} x_j} \leq e^{-1}$$

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding
optimal
LP solution
 $x \in \mathbb{R}^n$

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$
minimizing k ,

$$\text{s.t. } S_{e_1} \cup \dots \cup S_{e_k} = [m]$$

$$\min \sum_{j=1}^n x_j$$

over $x \in [0,1]^n$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \quad \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

Rounding LP solution

$$x \in [0,1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$$

For each set S_j , independently, take S_j with probability x_j .

$$E[|F|] = \sum_j x_j = \text{OPT} \quad (\text{good!})$$

Fix a point i .

$$E[\text{sets in } F \text{ covering } i] = \sum_{j: i \in S_j} x_j \geq 1$$

$$P[F \text{ not cover } i] = \prod_{j: i \in S_j} (1 - x_j) \leq \prod_{j: i \in S_j} e^{-x_j}$$

$$= e^{-\sum_{j: i \in S_j} x_j} \leq e^{-1}$$

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$
minimizing k ,

$$\text{s.t. } S_{e_1} \cup \dots \cup S_{e_k} = [m]$$

$$\min \sum_{j=1}^n x_j$$

over $x \in [0,1]^n$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \quad \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

Rounding LP solution

$$x \in [0,1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$$

For each set S_j , independently, take S_j with probability x_j .

$$E[|F|] = \sum_j x_j = \text{OPT} \quad (\text{good!})$$

$$P[F \text{ not cover } i] = e^{-1}$$

so F covers each point w/
constant probability

$\Rightarrow$ F is a set cover w/ constant prob.

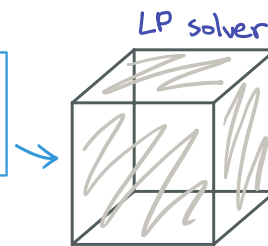
LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$
minimizing k ,

$$\text{s.t. } S_{e_1} \cup \dots \cup S_{e_k} = [m]$$

$$\min \sum_{j=1}^n x_j$$

$$\text{over } x \in [0,1]^n$$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \quad \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

Rounding LP solution

$$x \in [0,1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$$

2. For each set S_j independently, take S_j with probability $\min\{1, \alpha x_j\}$.

for $d > 1$ TBD.

$$E[|F|] = \sum_j d x_j = d \sum x_j = d \text{OPT}$$

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$
minimizing k ,

$$\text{s.t. } S_{e_1} \cup \dots \cup S_{e_k} = [m]$$

$$\min \sum_{j=1}^n x_j$$

$$\text{over } x \in [0,1]^n$$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \quad \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

Rounding LP solution

$$x \in [0,1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$$

2. For each set S_j independently, take S_j with probability $\min\{1, \alpha x_j\}$.

for $d > 1$ TBD.

$$E[|F|] = d \sum_j x_j = d \text{OPT}$$

$$P[F \text{ not cover } i] = \prod_{S_j \ni i} (1 - d x_j)$$

$$\leq e^{-d \sum_{S_j \ni i} x_j} \leq e^{-d}$$

$$\leq \frac{1}{m^2} \text{ for } d = 2 \ln(m)$$

$$1 - x \leq e^{-x}$$

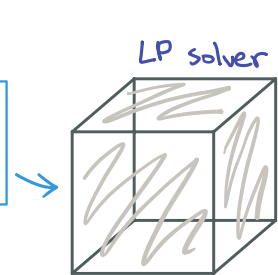
LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$
minimizing k ,

$$\text{s.t. } S_{e_1} \cup \dots \cup S_{e_k} = [m]$$

$$\min \sum_{j=1}^n x_j$$

$$\text{over } x \in [0,1]^n$$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \quad \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

Rounding LP solution

$$x \in [0,1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$$

2. For each set S_j independently, take S_j with probability $\min\{1, \alpha x_j\}$.

for $d > 1$ TBD.

$$E[|F|] = d \sum_j x_j = d \text{OPT}$$

$$P[F \text{ not cover } i] = \prod_{S_j \ni i} (1 - d x_j) \leq e^{-d \sum_{S_j \ni i} x_j} \leq e^{-d}$$

$$P[F \text{ not set cover}] \leq \sum_i P[F \text{ not cover } i] \leq m e^{-d} \leq m \cdot \frac{1}{m^2} = \frac{1}{m}$$

$$P_r[A \cup B] \leq P_r[A] + P_r[B] \quad d = 2 \ln m$$

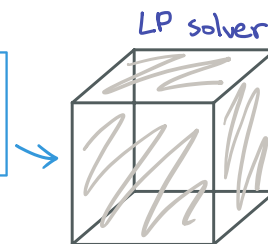
LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$
minimizing k ,

$$\text{s.t. } S_{e_1} \cup \dots \cup S_{e_k} = [m]$$

$$\min \sum_{j=1}^n x_j$$

over $x \in [0,1]^n$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \quad \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

Rounding LP solution

$$x \in [0,1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$$

2. For each set S_j independently, take S_j with probability $\min\{1, \alpha x_j\}$.

for $d > 1$ TBD.

$$E[|F|] = d \sum_j x_j = d \text{OPT}$$

$$P[F \text{ not cover } i] = \prod_{S_j \ni i} (1 - d x_j) \\ \leq e^{-d \sum x_j} \leq e^{-d}$$

$$P[F \text{ not set cover}] \leq \sum_i P[F \text{ not cover } i] \\ \leq m \cdot e^{-d} \leq \frac{1}{m^{\alpha d}} \text{ for } d = O(\log m)$$

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_j x \leq b_j \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_j x \leq b_j \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$
minimizing k ,

$$\text{s.t. } S_{e_1} \cup \dots \cup S_{e_k} = [m]$$

$$\min \sum_{j=1}^n x_j \\ \text{over } x \in [0,1]^n$$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \\ \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

Rounding LP solution

$$x \in [0,1]^n \rightsquigarrow F \subseteq \{S_1, \dots, S_n\}$$

2. For each set S_j independently, take S_j with probability $\min\{1, \alpha x_j\}$.

for $d > 1$ TBD.

$$E[|F|] = d \sum_j x_j = d \text{OPT}$$

$$P[F \text{ not set cover}] \leq \frac{1}{m^{\alpha d}}$$

$$\text{for } d = O(\log m)$$

$$\Rightarrow O(\log m)\text{-APX}$$

LP-Based Approximations

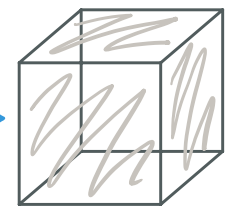
① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$



$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



LP solver

optimal LP solution $x \in \mathbb{R}^n$

rounding

approximate integral solution $z \in \mathbb{Z}^n$

Minimum Set Cover

m points $[m] \equiv \{1, \dots, m\}$

n sets $S_1, \dots, S_n \subseteq [m]$

Goal: select k sets $S_{e_1}, \dots, S_{e_k}$

minimizing k ,

$$\text{s.t. } S_{e_1} \cup \dots \cup S_{e_k} = [m]$$

$$\min \sum_{j=1}^n x_j$$

$$\text{over } x \in [0,1]^n$$

$$\text{s.t. } \sum_{j: i \in S_j} x_j \geq 1 \quad \forall i \in [m]$$

$$x_j = \begin{cases} 1 & \text{if } S_j \text{ in set cover,} \\ 0 & \text{otherwise} \end{cases}$$

today

Probability (& techniques)

Approximations via
linear programming
& randomized rounding

Algorithms

Max Sat

Set Cover

Covering
Integer Programs

Covering IPs

$$\min \sum_j c_j x_j$$

$$\text{over } x \in \mathbb{Z}_{\geq 0}^n$$

$$\text{s.t. } \sum_{j=1}^n A_{ij} x_j \geq b_i \text{ for } i \in [m]$$

$$\text{where } A \in \mathbb{R}_{\geq 0}^{m \times n}, \quad b \in \mathbb{R}_{> 0}^m, \quad c \in \mathbb{R}_{> 0}^n$$

"demands" "costs"

WLOG $A_{ij} \leq b_i \quad \forall i, j$ (why?)

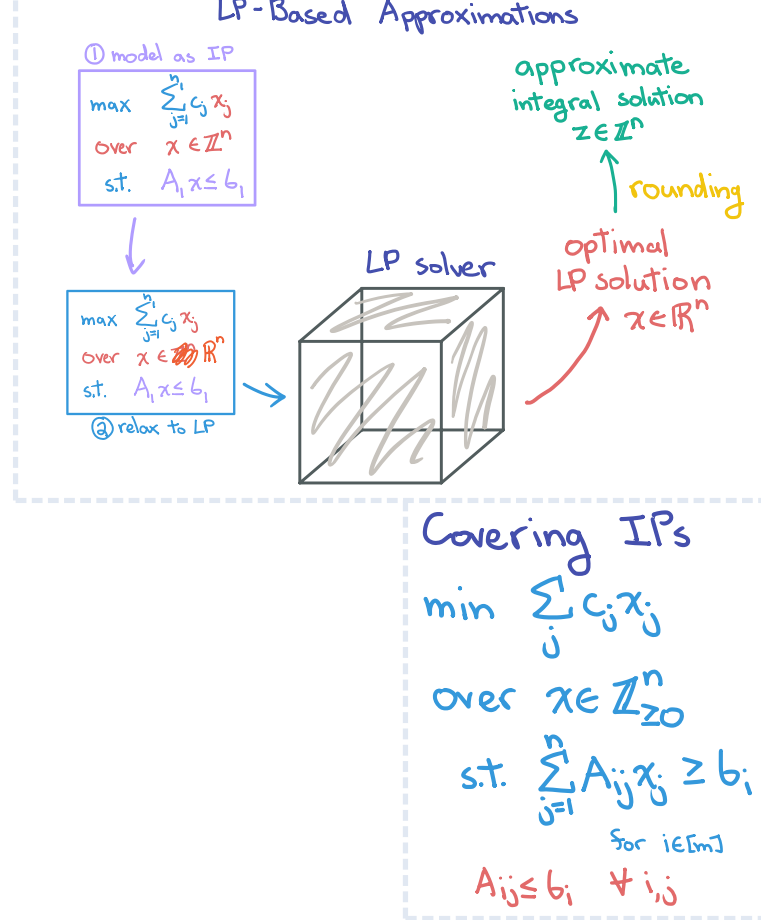
Set cover is special case w/ $A_{ij} \in \{0, 1\}$, $b_i = 1$, $c_j = 1$

LP relaxation

$$\min \sum_j c_j x_j$$

$$\text{over } x \in \mathbb{Z}_{\geq 0}^n$$

$$\text{s.t. } \sum_{j=1}^n A_{ij} x_j \geq b_i \text{ for } i \in [m]$$



Rounding LP solution

$$x \in \mathbb{R}_{\geq 0}^n \rightsquigarrow z \in \mathbb{Z}_{\geq 0}^n$$

Ideas?

LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$



$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

↑ rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

LP relaxation

$$\begin{array}{ll} \min & \sum_j c_j x_j \\ \text{over} & x \in \mathbb{R}_{\geq 0}^n \\ \text{s.t.} & \sum_{j=1}^n A_{ij} x_j \geq b_i \\ & \text{for } i \in [m] \end{array}$$

Covering IPs

$$\begin{array}{ll} \min & \sum_j c_j x_j \\ \text{over} & x \in \mathbb{Z}_{\geq 0}^n \\ \text{s.t.} & \sum_{j=1}^n A_{ij} x_j \geq b_i \\ & \text{for } i \in [m] \\ & A_{ij} \leq b_i \quad \forall i, j \end{array}$$

Rounding LP solution

$$x \in \mathbb{R}_{\geq 0}^n \rightsquigarrow z \in \mathbb{Z}_{\geq 0}^n$$

For each $j \in [n]$, independently, let

$$z_j = \begin{cases} \lceil \alpha x_j \rceil & \text{with probability } \alpha x_j - \lfloor \alpha x_j \rfloor \\ \lfloor \alpha x_j \rfloor & \text{otherwise.} \end{cases}$$

$$E[z_j] = \alpha x_j$$

$$E\left[\sum_j c_j z_j\right] = \alpha \sum_j c_j x_j = \alpha \text{OPT}$$

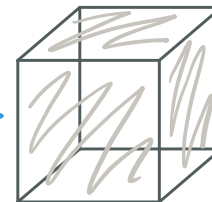
LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

↑ rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

↑

LP relaxation

$$\begin{array}{ll} \min & \sum_j c_j x_j \\ \text{over} & x \in \mathbb{R}_{\geq 0}^n \\ \text{s.t.} & \sum_{j=1}^n A_{ij} x_j \geq b_i \end{array}$$

for $i \in [m]$

Covering IPs

$$\begin{array}{ll} \min & \sum_j c_j x_j \\ \text{over} & x \in \mathbb{Z}_{\geq 0}^n \\ \text{s.t.} & \sum_{j=1}^n A_{ij} x_j \geq b_i \end{array}$$

for $i \in [m]$

$A_{ij} \leq b_i \quad \forall i, j$

For each $j \in [n]$, independently, let

$$z_j = \begin{cases} \lceil \alpha x_j \rceil & \text{with probability } \alpha x_j - \lfloor \alpha x_j \rfloor \\ \lfloor \alpha x_j \rfloor & \text{otherwise.} \end{cases}$$

$$P[\sum_j A_{ij} z_j \leq b_i]$$

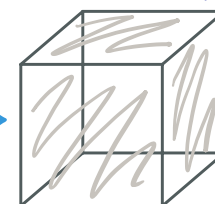
LP-Based Approximations

① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_1 x \leq b_1 \end{array}$$

② relax to LP



LP solver

approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding
optimal
LP solution
 $x \in \mathbb{R}^n$

LP relaxation

$$\begin{array}{ll} \min & \sum_j c_j x_j \\ \text{over} & x \in \mathbb{R}_{\geq 0}^n \\ \text{s.t.} & \sum_{j=1}^n A_{ij} x_j \geq b_i \\ & \text{for } i \in [m] \end{array}$$

Covering IPs

$$\begin{array}{ll} \min & \sum_j c_j x_j \\ \text{over} & x \in \mathbb{Z}_{\geq 0}^n \\ \text{s.t.} & \sum_{j=1}^n A_{ij} x_j \geq b_i \\ & \text{for } i \in [m] \\ & A_{ij} \leq b_i \quad \forall i, j \end{array}$$

For each $j \in [n]$, independently, let

$$z_j = \begin{cases} \lceil \alpha x_j \rceil & \text{with probability } \alpha x_j - \lfloor \alpha x_j \rfloor \\ \lfloor \alpha x_j \rfloor & \text{otherwise.} \end{cases}$$

$$P[\sum_j A_{ij} z_j \leq b_i]$$

LP-Based Approximations

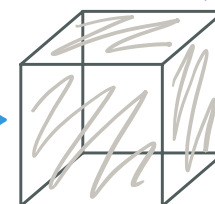
① model as IP

$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{Z}^n \\ \text{s.t.} & A_i x \leq b_i \end{array}$$



$$\begin{array}{ll} \max & \sum_{j=1}^n c_j x_j \\ \text{over} & x \in \mathbb{R}^n \\ \text{s.t.} & A_i x \leq b_i \end{array}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

↑ rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

LP relaxation

$$\begin{array}{ll} \min & \sum_j c_j x_j \\ \text{over} & x \in \mathbb{R}_{\geq 0}^n \\ \text{s.t.} & \sum_{j=1}^n A_{ij} x_j \geq b_i \\ & \text{for } i \in [m] \end{array}$$

Covering IPs

$$\begin{array}{ll} \min & \sum_j c_j x_j \\ \text{over} & x \in \mathbb{Z}_{\geq 0}^n \\ \text{s.t.} & \sum_{j=1}^n A_{ij} x_j \geq b_i \\ & \text{for } i \in [m] \\ & A_{ij} \leq b_i \quad \forall i, j \end{array}$$

Multiplicative Chernoff bound

let $X_1, \dots, X_n \in [0, 1]$ be independent

$$\mu = E[X_1 + \dots + X_n] \in (0, 1)$$

$$P[X_1 + \dots + X_n \leq (1 - \epsilon)\mu] \leq e^{-\epsilon^2 \mu / 2}$$

For each $j \in [n]$, independently, let

$$z_j = \begin{cases} \lceil \alpha x_j \rceil & \text{with probability } \alpha x_j - \lfloor \alpha x_j \rfloor \\ \lfloor \alpha x_j \rfloor & \text{otherwise.} \end{cases}$$

LP-Based Approximations

① model as IP

$$\begin{aligned} \max \quad & \sum_{j=1}^n c_j x_j \\ \text{over } & x \in \mathbb{Z}^n \\ \text{s.t. } & A_1 x \leq b_1 \end{aligned}$$

$$\begin{aligned} \max \quad & \sum_{j=1}^n c_j x_j \\ \text{over } & x \in \mathbb{R}^n \\ \text{s.t. } & A_1 x \leq b_1 \end{aligned}$$

② relax to LP



approximate
integral solution
 $z \in \mathbb{Z}^n$

rounding

optimal
LP solution
 $x \in \mathbb{R}^n$

LP relaxation

$$\begin{aligned} \min \quad & \sum_j c_j x_j \\ \text{over } & x \in \mathbb{R}_{\geq 0}^n \\ \text{s.t. } & \sum_{j=1}^n A_{ij} x_j \geq b_i \\ & \text{for } i \in [m] \end{aligned}$$

Covering IPs

$$\begin{aligned} \min \quad & \sum_j c_j x_j \\ \text{over } & x \in \mathbb{Z}_{\geq 0}^n \\ \text{s.t. } & \sum_{j=1}^n A_{ij} x_j \geq b_i \\ & \text{for } i \in [m] \\ & A_{ij} \leq b_i \quad \forall i, j \end{aligned}$$

Multiplicative Chernoff bound
let $X_1, \dots, X_n \in [0, 1]$ be independent

$$\mu = \mathbb{E}[X_1 + \dots + X_n] \in (0, 1)$$

$$P[X_1 + \dots + X_n \leq (1 - \epsilon)\mu] \leq e^{-\epsilon^2 \mu / 2}$$

Multiplicative Chernoff bound

let $X_1, \dots, X_n \in [0, 1]$ be independent

$$\mu = E[X_1 + \dots + X_n]$$

$$\epsilon \in (0, 1)$$

$$\bullet \quad P[X_1 + \dots + X_n \geq (1 + \epsilon)\mu] \leq e^{-\epsilon^2 \mu / 3}$$

$$\bullet \quad P[X_1 + \dots + X_n \leq (1 - \epsilon)\mu] \leq e^{-\epsilon^2 \mu / 2}$$

(proof at end if time permits)