LP-based approximation algorithms

▷ rounding LP-solutions (*)
▷ proving approx. guarantees via LP-duality (dual solutions)

(*) Find "suitable" LP-formulation.
Solve LP-relax. in polyn. time.
"Round" fractional solution to "good" feas. one.

Simple rounding

Example: Weighted Vertex Cover:

Given an undirected graph $G=(V,E)$ with nonnegative node weights: $w: V \to \mathbb{Q}^+$, find a minimum weight vertex cover, i.e., a set $V' \subseteq V$ such that every edge in E has at least one endpoint incident at V'.

(ILP) with $x_v = \begin{cases} 1 & \text{if } v \text{ is in VC} \\ 0 & \text{othw.} \end{cases}$, $\forall v \in V$

$$\min \sum_{v \in V} w_v x_v$$

$$\text{s.t.} \quad x_u + x_v \geq 1, \quad \forall e=(u,v) \in E \qquad (1)$$

$$x_v \in \{0,1\}, \quad \forall v \in V \qquad (2)$$

(LP) $$\min \sum_{v \in V} w_v x_v$$

s.t. (1)

$$x_v \geq 0, \quad \forall v \in V \qquad (2')$$

- $x_v \leq 1$ redundant in min. obj. and $w_v \geq 0$
- (LP) can be solved optim. in polyn. time

Simple rounding alg.

- Let x^{LP} be opt. sol. of (LP).
- Round x^{LP} to x: $x_v = \begin{cases} 1 & \text{if } x^{LP} \geq \frac{1}{2} \\ 0 & \text{othw.} \end{cases}$

Thm. The simple rounding alg. is a factor 2-approximation.

Pf. (a) x is feasible for (ILP), i.e., it is a vertex cover.

Why? Consider some edge $e = (u,v) \in E$.

$x_u^{LP} + x_v^{LP} \geq 1 \Rightarrow x_u^{LP} \geq 0.5$ or $x_v^{LP} \geq 0.5$

$\Rightarrow x_u = 1$ or $x_v = 1 \Rightarrow$ edge (u,v) is covered.

(b) algorithms cost : $Alg \leq 2 \cdot z^{LP}$

Why? $x_v \leq 2\, x_v^{LP}$, $\forall v \in V$

$$\Rightarrow Alg = \sum_{v \in V} w_v \cdot x_v \leq 2 \sum_{v \in V} w_v x_v^{LP} = 2 z^{LP}.$$

$\models$ Since $z^{LP} \leq OPT$, alg. is a 2-approx. $\square$

$\rightsquigarrow$ tight ex.: $w_u = 1$, $w_v = 1$, $x_u^{LP} = 0.5$, $x_v^{LP} = 0.5$

$\rightsquigarrow$ closer look at the (LP) gives:

Thm. Any basic feas. solution to (LP) is half-integral, i.e., $x_v \in \{0, \frac{1}{2}, 1\}$, $\forall v \in V$.

Pf. Recall, a feasible solution is basic if it cannot be expressed as a convex combin. of two other feas. solutions.

Consider a feas. solution x. Let

$$V_+ := \{ v \in V \mid \tfrac{1}{2} < x_v < 1 \}$$
$$V_- := \{ v \in V \mid 0 < x_v < \tfrac{1}{2} \}.$$

Consider new solutions y and z for $\varepsilon > 0$:

$$y_v := \begin{cases} x_v + \varepsilon & \text{if } x_v \in V_+, \\ x_v - \varepsilon & \text{if } x_v \in V_-, \\ x_v & \text{otherw.} \end{cases} \qquad z_v := \begin{cases} x_v - \varepsilon & \text{if } x_v \in V_+, \\ x_v + \varepsilon & \text{if } x_v \in V_-, \\ x_v & \text{otherw.} \end{cases}$$

We can choose ε s.t. y and z are feasible solutions to (LP). ($\varepsilon < \min_{v \in V} x_v$ for $x_v \neq 0$) Clearly, $y, z \neq x$.

But $\frac{1}{2} y + \frac{1}{2} z = x$. $\Rightarrow$ x not a basic feas. solution! $\square$

[Thm. [Nemhauser, Trotter '73] All extreme points of VC polytope are half-integral.]

Thm. The alg. that solves (LP) and rounds all $x_v^{LP} \in \{\frac{1}{2}, 1\}$ to $x_v = 1$ is a 2-approx. for min. vertex cover.

Randomized rounding

Idea: interpret fractional LP values x_j as probability with which x_j is set to 1.

Def. randomized approx. alg.

Example: MAX SAT

- m clauses $C_1, \dots, C_m$ and n boolean variables v_j, $j = 1, \dots, n$
- $\phi = \underbrace{(x_{11} \vee x_{12})}_{C_1} \wedge \underbrace{(x_{21} \vee x_{22} \vee x_{23})}_{C_2} \wedge \dots \wedge \underbrace{(x_{m1} \vee x_{m2} \vee \dots)}_{C_m}$

  $x_{ij} \in \{v, \bar{v}\}$ literals, $v \in \{v_1, v_2, \dots, v_n\}$ boolean variables
- each clause C_i has non-negative weight w_i.

Task: Find a truth assignment to the boolean variables such that the total weight of satisfied clauses is maximized.

Notation:

- Let $k(i)$ denote the number of literals in clause C_i.

  Note: - problem with $k(i) \le k$, $\forall i = 1, \dots, m$, also called Max-kSAT.
  - Max-2SAT is NP-hard, while 2SAT is in P.
- Let W be a random variable that denotes the total weight of satisfied clauses in a randomized truth assignment.
- Let W_i be the weight contributed by C_i.

1.) A trivial algorithm (no LP's) [Johnson '74]

Alg: Set every boolean var. v_j to TRUE indep. with probab. $\frac{1}{2}$.

Thm: Alg yields a 2-approx. for MAX-SAT.

Lemma: $E[W_i] = \left(1 - \frac{1}{2^{k(i)}}\right) \cdot w_i$.

Pf. C_i satisfied iff not all literals are FALSE.
Probability for this event:

$$\mathrm{Prob}[C_i \text{ satisfied}] = 1 - \frac{1}{2^{k(i)}}$$

$\square$

Pf. (of Thm.): For $\bar{k} \geq 1$ $\forall$ s.t. $k(i) \geq \bar{k}$: $1 - \frac{1}{2^{k(i)}} \geq 1 - \frac{1}{2^{\bar{k}}}$

From Lemma then follows:

$$E[W] = \sum_{i=1}^{m} E[W_i] \geq \left(1 - \frac{1}{2^{\bar{k}}}\right) \sum_{i=1}^{m} w_i \geq \left(1 - \frac{1}{2^{\bar{k}}}\right) OPT$$

$$\geq \frac{1}{2} OPT. \qquad \square$$

$\bar{k} = \min_i k(i)$

$\bar{k} = 1 \Rightarrow$ 2-approx.

$\bar{k} = 2 \Rightarrow$ 4/3-approx.

$\vdots \qquad \vdots$

$\bar{k} \nearrow \Rightarrow c \searrow$

2. Randomized rounding of LP solutions

0/1-var. y_j for bool. var. v_j, $j = 1, \dots, n$

0/1-var. z_i for clause C_i, $i = 1, \dots, m$

For any clause C_i, let $J^+(i)$ ($J^-(i)$) denote the set of indices of non-negated (negated) variables in C_i.

(ILP) $\max \sum_{i=1}^{m} w_i z_i$

s.t. $\sum_{j \in J^+(i)} y_j + \sum_{j \in J^-(i)} (1 - y_j) \geq z_i$, $i = 1, \dots, m$ (1)

$z_i \in \{0,1\}$, $i = 1, \dots, m$ (2)

$y_j \in \{0,1\}$, $j = 1, \dots, n$ (3)

(LP) $\max \sum_{i=1}^{m} w_i z_i$

s.t. (1)

$0 \leq z_i \leq 1$, $i = 1, \dots, m$

$0 \leq y_j \leq 1$, $j = 1, \dots, n$

Alg. • solve LP relaxation $\longrightarrow (y^{LP}, z^{LP})$

• set y_j to TRUE (=1) with probab. y_j^{LP} indep., $j = 1, \dots, n$

Thm. Alg. yields an approx. factor of $\frac{e}{e-1} \approx 1.57$ for MAX-SAT.

[Goemans, Williamson '93]

<u>Lemma</u> $E[W_i] \geq \left(1 - \left(1 - \frac{1}{k(i)}\right)^{k(i)}\right) w_i \cdot z_i^{LP}$

<u>Pf.</u>
- Wlog. all literals in C_i appear non-negated. (Othw. replace v_j by $\bar{v}_j$ in all clauses; no effect on W_i or z^{LP}.)
- Rename variables s.t. $C_i = (v_1 \vee v_2 \vee \dots \vee v_{k(i)})$

C_i satisfied if not all $v_1, \dots, v_{k(i)}$ are set to FALSE.

$$\Rightarrow \text{Prob}[C_i \text{ satisfied}] = 1 - \prod_{j=1}^{k(i)} (1 - y_j^{LP})$$

$$\geq 1 - \left(\frac{\sum_{j=1}^{k(i)} (1 - y_j^{LP})}{k(i)}\right)^{k(i)}$$

geometric mean ≥ arithmetic mean, i.e. $\frac{\sum a_i}{n} \geq \sqrt[n]{\prod a_i}$

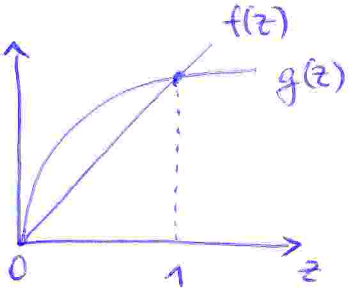
$$= 1 - \left(1 - \frac{\sum_{j=1}^{k(i)} y_j^{LP}}{k(i)}\right)^{k(i)}$$

(*) $\geq z_i^{LP}$ by (1) in (LP)

$$\geq 1 - \left(1 - \frac{z_i^{LP}}{k(i)}\right)^{k(i)}$$

Now $g(z) := 1 - (1 - \frac{z}{k})^k$ is a concave function in z with $g(0) = 0$ and $g(1) = 1 - (1 - \frac{1}{k})^k$.
And $f(z) := \left(1 - (1 - \frac{1}{k})^k\right) \cdot z$ is a lin. function in z with $f(0) = g(0)$ and $f(1) = g(1)$.

$\Rightarrow$ $g(z) \geq f(z)$ on interval $[0,1]$ since $g(z) \geq f(z)$ at endpoints of interval.



$$\Rightarrow \text{Prob}[C_i \text{ satisfied}] \geq \left(1 - \left(1 - \frac{1}{k(i)}\right)^{k(i)}\right) \cdot z_i.$$

□

<u>Pf. (of Thm).</u>

$$E[W] = \sum_{i=1}^{m} E[W_i] \geq \sum_{i=1}^{m} \underbrace{\left(1 - \left(1 - \frac{1}{k(i)}\right)^{k(i)}\right)}_{\text{decreasing fct. of } k(i)} z_i^{LP} \cdot w_i$$

$(k \geq k(i))$

$$\geq \left(1 - \left(1 - \frac{1}{k}\right)^k\right) \sum_{i=1}^{m} z_i^{LP} \cdot w_i = \left(1 - \left(1 - \frac{1}{k}\right)^k\right) z^{LP}$$

$$\geq \left(1 - \underbrace{\left(1 - \frac{1}{k}\right)^k}_{\forall k \in \mathbb{Z}^+}\right) \text{OPT}. \geq \frac{e-1}{e} \cdot \text{OPT}.$$

see other

□

3.) A combined algorithm.

Alg: Flip a fair coin to decide which of the two previous algorithms to run.

Remark: - actually we set variable y_j to TRUE with probab. $\frac{1}{2}\cdot\frac{1}{2}+\frac{1}{2}\cdot y_j^{LP}$
- but they are not set independently any more

Thm. The combined alg. is a 4/3 - approximation for MAX-SAT.

Pf. $$\text{Prob. } [C_i \text{ satisfied}] \geq \underbrace{\frac{1}{2}\left(1-\frac{1}{2^{k(i)}}\right)} + \frac{1}{2}\left(1-\left(1-\frac{1}{k(i)}\right)^{k(i)}\right) z_i^{LP}$$

$$\geq \frac{1}{2}\left(1-\frac{1}{2^{k(i)}}\right) z_i^{LP} \quad \text{since } z_i^{LP} \leq 1$$

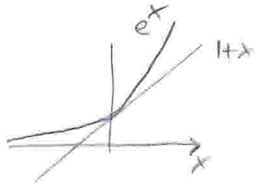
$$\geq f(k(i)) \cdot z_i^{LP} \quad \text{with } f(k(i)) := \frac{1}{2}\left(1-\frac{1}{2^{k(i)}}\right) + \frac{1}{2}\left(1-\left(1-\frac{1}{k(i)}\right)^{k(i)}\right)$$

analyze function $f(k(i))$:
- $f(1) = f(2) = 3/4$
- for $k(i) \geq 3$: $f(k(i)) \geq \frac{1}{2}\cdot\frac{7}{8} + \frac{1}{2}\cdot\left(1-\frac{1}{e}\right) \geq 3/4$

$$\Longrightarrow E[W] = \sum_{i=1}^{m} E[W_i] = \sum_{i=1}^{m} f(k(i)) \cdot z_i^{LP} \cdot w_i$$

$$\geq \frac{3}{4} z^{LP} \geq \frac{3}{4} \text{OPT}. \qquad \square$$

Why $\left(1-\frac{1}{k}\right)^k \leq \frac{1}{e}$, $\forall k \in \mathbb{Z}^+$?



$1+x \leq e^x$, $\forall x \in \mathbb{R}$

$x := -\frac{1}{k}$

$\Rightarrow 1-\frac{1}{k} \leq e^{-1/k} \quad \Rightarrow \quad \left(1-\frac{1}{k}\right)^k \leq \frac{1}{e}$