

אלגוריתם גאוס - 2008

ע"פ

מספרים ממשיים של מערכת משוואות ליניאריות

Systems of equations $Ax = b$

solution: an (affine) subspace of \mathbb{R}^n

$x \in \{x: Ax = b\}$ where $Ax = b$

algorithm: Gauss elimination

What about inequalities?

$$\begin{cases} Ax = b \\ x \geq 0 \end{cases}$$

$$\begin{cases} Ax \leq b \end{cases}$$

$$\begin{cases} x_1 + 2x_2 + 3x_3 \leq 4 \\ x_1 + x_2 + x_3 = 1 \\ x_1 \geq 0 \\ x_2 \geq 0 \\ x_3 \geq 0 \end{cases}$$

Question #1

Find an efficient algorithm for the following problem:

Input: a set Π of constraints
(in)equalities over variables $x_1, \dots, x_n \in \mathbb{R}$

[Each constraint is of the form:

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \begin{matrix} \leq \\ \geq \\ = \end{matrix} b_i]$$

Output: "not feasible" OR $\vec{x} \in \mathbb{R}^n$ that satisfies all the constraints

Question #2: Find an efficient algorithm for the following problem:

Input: $m \times n$ matrix $A \in \mathbb{M}_{m \times n}$
vector $b \in \mathbb{R}^m$
vector $c \in \mathbb{R}^n$

Goal: find $x \in \mathbb{R}^n$ that minimizes $c^t x \triangleq \sum_j c_j x_j$
subject to $A \cdot x = b, x \geq 0$.

Output: "not feasible" or $x^* = \operatorname{argmin} \{ c^t x / Ax = b, x \geq 0 \}$.

Both questions are about **linear programming**!

- Q1: Is a polyhedron empty?
- Q2: Linear programming in standard form.

Lots of: structure
& algorithms ...

Combinatorial Optimization Problems

find "best" solution for a combinatorial problem (as opposed to a "numerical problem").

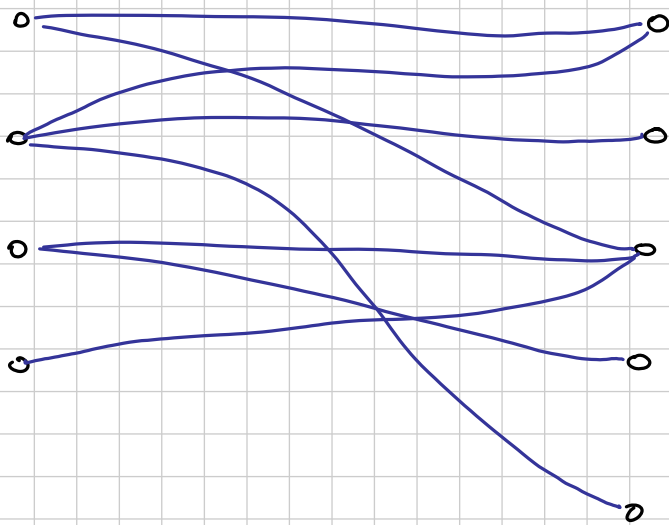
Examples:

- Set Cover
- Vertex Cover
- Minimum weight matching
- Maximum flow

Set Cover - Example

doctors

patients



Rules:

Every doctor can serve a subset of the patients.

Goal:

Choose fewest # doctors, so that each patient is served.

Definition: Set-Cover

Input: a set of elements $[1..n]$

a family of subsets: $S_1, \dots, S_m \subseteq [1..n]$

$\{S_{i_1}, S_{i_2}, \dots, S_{i_k}\}$ is a **Set Cover** if

$$S_{i_1} \cup S_{i_2} \cup \dots \cup S_{i_k} = [1..n]$$

every patient served by a doctor

Goal: find a set cover of minimum size.

History : Set Cover

- 1) Lots of applications & special cases.
- 2) Among the first problems proven to be NP-Complete [Karp]
- 3) Greedy algorithm gives $H_n \approx \ln n$ approximation.
- 4) NP-Complete to find "better" approximations [Feige, RS]

What does linear programming have to say?

Set Cover : the greedy algorithm

Input : $S_1, \dots, S_m \subseteq [1..n]$

goal : find smallest set $\{S_{i_1}, \dots, S_{i_k}\}$ s.t. $S_{i_1} \cup \dots \cup S_{i_k} = [1..n]$

method : (Greedy)

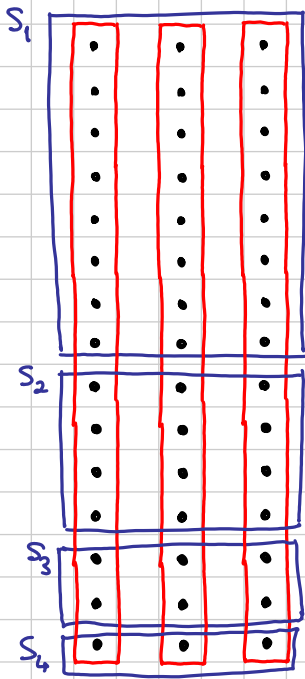
$C \leftarrow \emptyset$

while $C \neq [1..n]$ do

$S \leftarrow \operatorname{argmax} \{ |S_i \setminus C| : 1 \leq i \leq m \}$
 $C \leftarrow C \cup S$ (add S to the cover)

Question : How many sets in the cover (compared to an optimal solution) ?

Greedy - Example



Greedy picks:

$$\{S_1, S_2, S_3, S_4\}$$

But $|OPT| = 3$

Set Cover: the greedy algorithm (cost)

Theorem [Johnson]

greedy returns an (H_n) -approx.

$$H_n = 1 + \frac{1}{2} + \dots + \frac{1}{n} \\ \leq 1 + \ln n$$

$$(|\text{Greedy}| \leq (1 + \ln n) \cdot \text{OPT})$$

An Integer Programming formulation for Set Cover

For each set S define a variable $x_S \in \{0, 1\}$

$$\begin{aligned} \min \quad & \sum_S x_S \\ \text{s.t.} \quad & \sum_{\{S: e \in S\}} x_S \geq 1 \quad \forall \text{element } e \in [1..n] \end{aligned}$$

Now relax integrality constraint $x_S \in \{0, 1\}$
and use fractional constraint $0 \leq x_S \leq 1$:

$$\begin{aligned} \text{Min} \quad & \sum_S x_S \\ \text{s.t.} \quad & \sum_{\{S: e \in S\}} x_S \geq 1 \quad \forall \text{element } e \\ & 0 \leq x_S \leq 1 \quad \forall \text{set } S \end{aligned} \quad \left. \vphantom{\begin{aligned} \text{Min} \quad & \sum_S x_S \\ \text{s.t.} \quad & \sum_{\{S: e \in S\}} x_S \geq 1 \quad \forall \text{element } e \\ & 0 \leq x_S \leq 1 \quad \forall \text{set } S \end{aligned}} \right\} \begin{array}{l} \text{fractional} \\ \text{relaxation} \\ \text{for Set-Cover} \end{array}$$

Example of fractional Set Cover

n sets



n elements



\leftarrow d -regular bipartite graph (superposition of d matchings)

$$\begin{aligned} \text{Min} \quad & \sum_S x_S \\ \text{s.t.} \quad & \sum_{\{S: e \in S\}} x_S \geq 1 \quad \forall \text{element } e \\ & 0 \leq x_S \leq 1 \quad \forall \text{set } S \end{aligned}$$

Suppose $\forall \text{set } S: |S|=d$
 $\forall \text{element } e: |\{S: e \in S\}|=d$

$\Rightarrow x_S = \frac{1}{d}$ is a feasible solution.

$\sum x_S = \frac{n}{d}$ is optimal (exercise)

What can we do
with (optimal) solution
 $\{x_s^*\}_s$ of LP?

$$\begin{aligned} \text{Min} \quad & \sum_s x_s \\ \text{s.t.} \quad & \sum_{\{s: e \in s\}} x_s \geq 1 \quad \forall \text{element } e \\ & 0 \leq x_s \leq 1 \quad \forall \text{set } s \end{aligned}$$

1) Lower bound: $\sum_s x_s^* \triangleq \text{OPT}_f \leq \text{OPT}$

2) If (almost) integral then $\{x_s^*\}$ is (almost) a set cover.

3) Round $\{x_s^*\}$...

First, let's use it to analyze Greedy.

$C \leftarrow \emptyset$ Greedy - with charging

while $C \neq [1..n]$ do

$$\left[\begin{array}{l} S \leftarrow \text{argmax} \{ |S_i \setminus C| : 1 \leq i \leq m \} \\ \forall e \in S \setminus C : p(e) \leftarrow \frac{1}{|S \setminus C|} \quad (\text{charge covered elements}) \\ C \leftarrow C \cup S \quad (\text{add } S \text{ to the cover}) \end{array} \right.$$

In each round $\sum_{e \in S \setminus C} p(e) = 1$ pays for S .

\Rightarrow In the end: $\sum_e p(e) = |C|$.

Why does this help?

Primal LP (covering)

$$\begin{aligned} \text{Min } & \sum_s x_s \\ \text{s.t. } & \sum_{\{s: e \in s\}} x_s \geq 1 \quad \forall \text{ element } e \\ & 0 \leq x_s \leq 1 \quad \forall \text{ set } s \end{aligned}$$

Dual LP (packing)

$$\begin{aligned} \text{max } & \sum_e y_e \\ \text{s.t. } & \sum_{e \in s} y_e \leq 1 \quad \forall s \\ & 0 \leq y_e \end{aligned}$$

Claim (weak duality) : $\sum_s x_s \geq \sum_e y_e$

$$\text{proof: } \sum_s 1 \cdot x_s \geq \sum_s \left(\sum_{e \in s} y_e \right) x_s = \sum_e \left(\sum_{s: e \in s} x_s \right) y_e \geq \sum_e y_e$$

Hope: If $\{p_e\}$ is a dual solution, then

$$|\text{Greedy}| = \sum_e p_e \leq \text{OPT}_f \leq \text{OPT} \dots$$

But, $\{p_e\}$ is NOT a dual solution... (exercise)

Lovasz: $y_e = \frac{p_e}{1 + \ln n}$ is dual solution!

$$\text{and then } |\text{Greedy}| = \sum_e p_e = (1 + \ln n) \sum_e y_e$$

$$\leq (1 + \ln n) \cdot \text{OPT}_f \leq (1 + \ln n) \cdot \text{OPT}$$

\Rightarrow Greedy is an $O(\lg n)$ -approximation.

So, now left to prove feasibility of $\{y_e\}$.

Consider any set S

$$S = \{e_1, e_2, \dots, e_l\}.$$

let $T(e) \triangleq$ iteration

in which e is covered

while $C \neq [1..n]$ do

$$S \leftarrow \operatorname{argmax} \{ |S_i \setminus C| : 1 \leq i \leq m \}$$

$$\forall e \in S \setminus C : p(e) \leftarrow \frac{1}{|S \setminus C|} \quad (\text{change covered elements})$$

$$C \leftarrow C \cup S \quad (\text{add } S \text{ to the cover})$$

assume: $T(e_1) \leq T(e_2) \leq \dots \leq T(e_l)$

$$\Rightarrow p(e_1) \leq p(e_2) \leq \dots \leq p(e_l)$$

moreover, when e_i is covered $S \setminus C \ni \{e_i, e_{i+1}, \dots, e_l\}$.

$$\Rightarrow p(e_1) \leq \frac{1}{l}, \quad p(e_2) \leq \frac{1}{l-1}, \quad \dots, \quad p(e_l) \leq 1.$$

$$\Rightarrow \forall S : \sum_{e \in S} p(e) \leq \sum_{i=1}^n \frac{1}{i} \triangleq H_n \leq 1 + \ln n.$$

$$\Rightarrow y_e = \frac{p(e)}{1 + \ln n} \text{ is dual feasible.} \quad \square$$

Recap

1) Set Cover Problem

2) Greedy Algorithm

3) Integer Programming Formulation.

4) Linear Programming Formulation.

5) Primal & Dual LP's.

6) Weak Duality

7) Analysis of approximation ratio of Greedy.

Homework (to be solved on Wed!)

1) Give an example where:

$$|\text{Greedy}| \geq \Omega(\lg n)$$

and

$$|\text{OPT}| \leq \text{constant}.$$

Hint: extend the example from slides.

* Does this show that $\{p(e)\}$ is not a dual solution?

* Can you give an example with $\frac{|\text{Greedy}|}{|\text{OPT}|} = H_n$?

2) Prove that if a set-system has n sets & n elements in which $\forall \text{set } |S| = d$ and $\forall \text{element } |\{S : e \in S\}| = d$, then $\text{OPT}_f = \frac{n}{d}$.

Hint: weak duality.

3) Find an example in which

$$\text{OPT} \geq \Omega(\lg n) \cdot \text{OPT}_f.$$

Hint: Elements $U = \{0,1\}^k - 0^k$ (all nonzero strings)

Sets $\{S_i\}_{i \in U}$

where

$$S_i = \left\{ u \in U \mid \sum_{j=1}^k u_j \cdot i_j \text{ is odd} \right\}$$

$$\text{show: } \forall i: |S_i| = \frac{2^k}{2}$$

$$\forall u: |\{S_i: u \in S_i\}| = \frac{2^k}{2}$$

use prob. (2) to prove $\text{OPT}_f \leq 2$.

prove $\text{OPT} \geq k$ (system of $< k$ equations in $\text{GF}(2)$ has nonzero solution)

Problems

4) Suppose that $\forall \text{element } e: |\{S: e \in S\}| \leq d$.

Find an approximation algorithm with

$$|\text{ALG}| \leq d \cdot \text{OPT}.$$

5) Improve the analysis of the approximation

ratio of the Greedy algorithm if

$\forall \text{Set } S: |S| \leq k$. (say, $k=3$).
(what if $k=2$???)

Problems

6) Weak duality

<u>Primal</u>	<u>Dual</u>
$\min c^t x$	$\max y^t b$
s.t. $Ax \geq b$	s.t. $y^t A \leq c$
$x \geq 0$	$y \geq 0$

prove that: if x is a primal solution & y is a dual solution, then $c^t x \geq y^t b$.

Encore: Rounding a fractional set-cover

Suppose $\{x_s\}$ is an optimal fractional set-cover,

$$\text{I.E., } \forall e: \sum_{e \in S} x_s \geq 1 \text{ and } \sum_s x_s = \text{OPT}_f.$$

How can we obtain an integral set-cover

$$\hat{x}_s \in \{0, 1\} \text{ from } x_s ?$$

2 Goals:

feasibility: $\forall e \exists S: e \in S \text{ \& } \hat{x}_s = 1$

cost: $\frac{\sum_s \hat{x}_s}{\sum_s x_s}$ is small ($\sum_s \hat{x}_s \leq O(\lg n) \cdot \text{OPT}_f$)

Randomized Rounding

For each set S , randomly flip a "coin" \hat{x}_S such that $\Pr[\hat{x}_S = 1] = x_S$.

Now: \forall element e :

$$\Pr[e \text{ is covered}] = \Pr[\exists S: e \in S \ \& \ \hat{x}_S = 1]$$

$$= 1 - \Pr[\forall S: e \in S \Rightarrow \hat{x}_S = 0]$$

$$= 1 - \prod_{S: e \in S} (1 - x_S)$$

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

$$\geq 1 - e^{-\sum_{S: e \in S} x_S} \geq 1 - e^{-1} \approx 0.632$$

\Rightarrow Each element is covered with constant prob.

Repeat this experiment $\ln(4n)$ times

(S is in the cover if $\hat{x}_S = 1$ in one of experiments)

$$\Pr(e \text{ is not covered}) \leq \left(\frac{1}{e}\right)^{\ln(4n)} = \frac{1}{4n}$$

$$\Rightarrow \Pr(\exists e \text{ not covered}) \leq \sum_e \Pr(e \text{ not covered}) = \frac{1}{4}$$

\Rightarrow with prob $\leq \frac{1}{4}$ every element is covered if we repeat $\ln(4n)$ times.

Amplify prob. of success by repeating $k \cdot \ln(4n)$ times $\implies \Pr(\text{Success}) \geq 1 - \left(\frac{1}{4}\right)^k$.

What about cost?

$$E\left[\sum_s \hat{x}_s\right] = \sum_s E[\hat{x}_s] = \sum_s x_s$$

We repeat $\ln(4n)$ times $\Rightarrow E(\text{cover}) = \ln(4n) \cdot \text{OPT}_f$.

Markov Inequality:

$$\Pr(\text{cover} \geq 4 \cdot \ln(4n) \text{OPT}_f) \leq \frac{1}{4}.$$