Semidefinite Programming

Wednesday, March 18, 2015
3:58 PM

Linear programming:

$$\begin{aligned} \max \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

where $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$

Semidefinite programming:

Instead of $\mathbb{R}^n$, variables are from another vector space

$$SYM_n = \{ X \in \mathbb{R}^{n \times n} : x_{ij} = x_{ji}, \forall i, j \in [n] \}$$

Also, let $A \bullet B = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ij}$.

Then, an SDP looks like:

$$\boxed{\begin{aligned} \max \quad & C \bullet X \\ \text{s.t.} \quad & A_1 \bullet X = b_1 \\ & A_2 \bullet X = b_2 \\ & \quad \vdots \\ & A_m \cdot X = b_m \\ & X \succeq 0 \end{aligned}}$$

where $X \in SYM_n$, $A_1, \ldots, A_m, C \in \mathbb{R}^{n \times n}$, $b_1, \ldots, b_m \in \mathbb{R}$ and $X \succeq 0$ denotes the constraint "X is positive semidefinite"

What is positive semidefinite?

$M \in SYM_n$ is psd if all its eigenvalues are non-negative

<u>Theorem</u>: If $M \in SYM_n$, following are equivalent:

(i) M is psd

(ii) $x^T M x \geq 0 \quad \forall x \in \mathbb{R}^n$

(iii) $\exists\, U \in \mathbb{R}^{n \times n}$ s.t. $M = U^T U$ (Cholesky factorization)

<u>Pf</u>: Exercise (using diagonalization)

MAX - CUT ... SDP's

MAX-CUT

- First approx algorithms to use SDP's
- Best known approx ratio for MAX-CUT.
- MAX-CUT: Given graph $G = (V, E)$, find S maximizing $E(S, V \setminus S)$
  - Midterm asked for 0.5-factor approx (put every vertex v in S with prob. $\frac{1}{2}$ independently).
- [Goemans-Williamson '95]: 0.878-approximation
- We will assume 2 things

More later!

(1) SDP programs can be solved in poly time (true under suitable conditions)

(2) Cholesky factorization in poly time

Consider the program:

$$\max \sum_{(i,j) \in E} \frac{1 - z_i z_j}{2} \equiv \text{MAX-CUT}$$

$$\text{s.t.} \quad z_i \in \{-1, 1\} \quad \forall i$$

Replace each z_i with a vector $u_i \in \mathbb{R}^n$.

~~SDP formulation~~:

$$\text{(SDP)} \quad \max \sum_{(i,j) \in E} \frac{1 - u_i^T u_j}{2}$$

$$\text{s.t.} \quad \|u_i\| = 1$$

Relaxation of MAX-CUT: SDP $\geq$ OPT.

Why SDP? Consider:

$$\text{(SDP')} \quad \max \sum_{(i,j) \in E} \frac{1 - x_{ij}}{2}$$

$$\text{s.t.} \quad x_{ii} = 1 \quad \forall i$$

$$X \succeq 0$$

~~Claim~~: SDP = SDP'

Pf: If $x_{ij} = u_i^T u_j$, then $X = (x_{ij})$ is psd.

If $x_{ij} = u_i^T u_j$ and each u_i has unit norm.

If X is psd, $x_{ij} = u_i^T u_j$

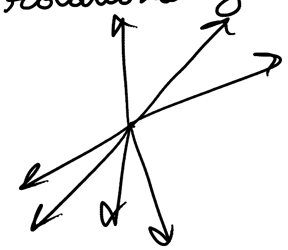
So, by assumption, we find in poly time a psd matrix X with $x^*_{i,i} = 1$ and $\sum_{(i,j)\in E} \frac{1 - x^*_{ij}}{2} \geq SDP - \varepsilon \geq OPT - \varepsilon$

By second assumption, we find in poly time a matrix U^* s.t.
$X^* = (U^*)^T U^*$ (upto tiny error). Assume factorization is exact.
So, $u_1^*, \dots, u_n^* \in \mathbb{R}^n$ are unit vectors such that:

$$\sum_{(i,j)\in E} \frac{1 - u_i^{*T} u_j^*}{2} \geq OPT - \varepsilon$$

~~Rounding the vector solution~~

Note that objective only depends on the inner product between the u_i's and is rotationally invariant.



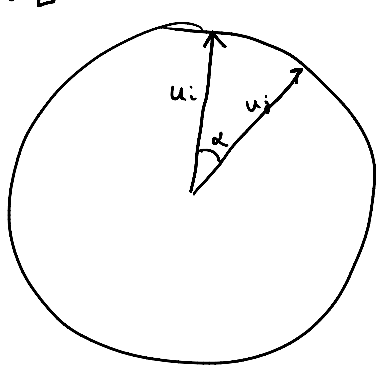
Idea: choose a random hyperplane H and round u_i to $+1$ if "above" H and -1 if "below" H.

Choose random $p \in \mathbb{R}^n$ s.t. $\|p\| = 1$. Let $s_i = 1$ if $p^T u_i \geq 0$, -1 o.w.

Let α be the angle between u_i and u_j. Then:

~~Lemma~~: $\Pr[s_i \neq s_j] = \frac{\alpha}{\pi} = \frac{1}{\pi} \cos^{-1}(u_i^T u_j)$

Pf:



$\Pr[H \text{ falls between } u_i \text{ and } u_j] = \frac{\alpha}{\pi} = \frac{1}{\pi} \cos^{-1}(u_i^T u_j)$

So, $\mathbb{E}[\text{cut}] = \sum_{(i,j)\in E} \frac{\cos^{-1} u_i^T u_j}{\pi} \geq 0.8785672 \sum_{(i,j)\in E} \frac{1 - u_i^T u_j}{2}$

$\geq 0.878 \cdot OPT$

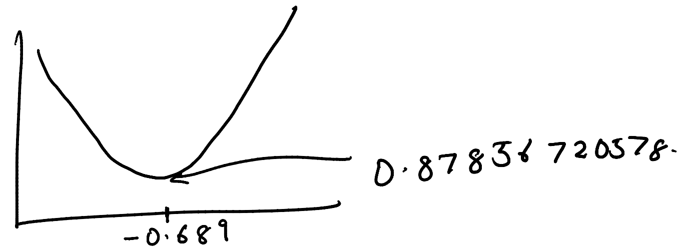
Lemma: For all $z \in [-1, 1]$:

$1 \dots 0.878567 \frac{1 - z}{}$

Lemma: For all $z \in [-1,1]$:

$$\frac{\cos^{-1} z}{\pi} \geq 0.878567 \, \frac{1-z}{2}$$

Pf: Graph $\dfrac{2\cos^{-1} z}{\pi(1-z)}$.



Important note:

We assumed SDP can be solved exactly and Cholesky decomposition can be done exactly. Not true in Turing machine model: W s.t. $U^T U = [3]$ has irrational entries. In TM model, U can be found upto arbitrarily small approx ε. Similarly for solving SDP's. We can incorporate this error into approximation factor.

Complexity of solving SDP's

Ellipsoid method: Given a convex set C contained in a ball of radius R and a poly time separation oracle, there is a poly time (in n and R) algorithm for max/min any given linear function over C.

Feasible solutions of an SDP convex? ✓

Separation oracle?

For any infeasible X, need to give a matrix $S \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}$ s.t. $S \cdot Y \leq b$ for all feasible Y but $S \cdot X > b$.

How can X not be feasible?

- X is not symmetric? Then, $X_{ij} > X_{ji}$, so $Y_{ij} - Y_{ji} \leq 0$ is violating
- X violates some linear constraint? immediate
- X is not psd. Then, $u^T X u < 0$ where u is the eigenvector of X with neg eigenvalue. Then $u^T Y u \geq 0$ is a violating constraint.

R bounded?

Depends on ~~application!~~

For MAX-CUT, $\sqrt{\sum X_{ij}^2} \leq n$

~~MAX 2-SAT~~

- SAT where each clause contains ≤ 2 literals.
- 0.75 - approximation from a random assignment
- We show $\alpha_{GW} = 0.878567$ approx using SDP's!
- For i'th variable, let $y_i \in \{\pm 1\}$. Also, have a "reference" variable $y_0 \in \{\pm 1\}$. If $y_i = y_0$, interpret setting var i to TRUE, o/w FALSE.
- For a clause C, let $V(C) = 1$ if satisfied, 0 o/w.

  If $C = v_i \vee v_j$,

$$V(C) = 1 - V(\overline{v_i}) \cdot V(\overline{v_j})$$
$$= 1 - \frac{1 - y_0 y_i}{2} \cdot \frac{1 - y_0 y_j}{2}$$
$$= \frac{1 + y_0 y_i}{4} + \frac{1 + y_0 y_j}{4} + \frac{1 - y_i y_j}{4}$$

- In total,

$$\max \sum a_{ij} (1 + y_i y_j) + b_{ij} (1 - y_i y_j)$$
$$\text{s.t.} \quad y_i \in \{\pm 1\}$$

- As SDP,

$$\max \sum a_{ij} (1 + u_i \cdot u_j) \; u_{nj} \, (1 - u_i \cdot u_j)$$
$$\text{s.t.} \quad u_i \cdot u_i = 1, \quad u_i \in \mathbb{R}$$

- Similar to MAX-CUT. Will show in Exercise that $\mathbb{E}[\#\text{ clauses satisfied}] \geq \alpha_{GW} \, OPT$.