p85

2014 07 08

Let us discuss problems with rational (hence integer) data. In the LP setting we have data $(A \in \mathbb{Z}^{m \times n}, b \in \mathbb{Z}^m, c \in \mathbb{Z}^n)$

$\text{size}(A, b, c) =: L$

text has proof

Proposition 4.14': Every extreme point $\bar{x}$ of

$$F := \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$$

satisfies $\text{size}(\bar{x}) \leq L$. Moreover, for every pair of extreme points $\bar{x}$ and $\hat{x}$ of F, either

$$c^T\bar{x} = c^T\hat{x}$$

or

$$|c^T\bar{x} - c^T\hat{x}| > 2^{-2L}.$$

proof also in text

In LPs with feasible region in the form of F above, if we have $\tilde{x} \in F$ then in at most $O(n^3)$ arithmetic operations we can find $\hat{x} \in F$ an extreme point of F such that

$$c^T\hat{x} \leq c^T\tilde{x}.$$

In SDP, the algorithm mentioned above has a generalization with the same claimed property. However, the claims of Prop. 4.14' are both false in the general case of SDP and currently we do not have a suitable replacement.

Consider the SDP (in dual form):

$$y_1 \geq 2, \begin{bmatrix} 1 & y_1 \\ y_1 & y_2 \end{bmatrix} \succeq 0, \begin{bmatrix} 1 & y_2 \\ y_2 & y_3 \end{bmatrix} \succeq 0, \ldots, \begin{bmatrix} 1 & y_{n-1} \\ y_{n-1} & y_n \end{bmatrix} \succeq 0$$

In every feasible solution $y_n \geq 2^{2^{n-1}}$. Hence $\text{size}(y) \geq 2^{n-1}$ for all ration feasible solutions y.
Note that $L = O(n)$.
So, in SDPs even with integer data (and even with $L = O(n)$) the following are possible.

(i) all extreme points of the feasible region has size $\geq 2^L$ (entries $\geq 2^{2^L}$)
(ii) SDP has a unique optimal solution or feasible solution with irrational entries
(iii) The feasible region of SDP has non-empty interior, but the radius of the largest ball contained in it is $2^{-2^{\Omega(L)}}$

Open problem 2: Is SDP Feasibility in P?

SDP Feasibility: Given $A_1, \ldots, A_m \in \mathbb{S}^n \cap \mathbb{Z}^{n \times n}$, $b \in \mathbb{Z}^m$, does there exist $X \in \mathbb{S}^n_+$ such that $\langle A_i, X \rangle = b_i \; \forall i \in \{1, \ldots, m\}$.

read chapters 1–4

(2014 09 08)

Open Problem 2': Is SDP feasibility in $\mathcal{NP}$?

Approximation Algorithms Based on SDP

Let $G=(V,E)$ be a simple undirected graph, $w \in \mathbb{R}^E_+$ be given.
Every $U \subseteq V$ defines a cut

$$\delta(U) := \{\{i,j\} \in E ; i \in U, j \in V \setminus U\}.$$

The weight of the cut $\delta(U)$ is

$$\sum_{\{i,j\} \in \delta(U)} w_{ij} .$$

MAX CUT: Find a cut of maximum weight in G.
Let's define $w_{ij}=0 \ \forall \{i,j\} \notin E$.

$$u_i := \begin{cases} 1 & \text{if } i \in U \\ -1 & \text{if } i \in V \setminus U \end{cases}, \quad u \in \mathbb{R}^n, \quad n := |V|.$$

Max Cut Problem is

$$\operatorname{Max} \tfrac{1}{4} \sum_{i,j} w_{ij}(1-u_i u_j)$$
$$u \in \{-1,1\}^n$$

"$X = uu^T$"

Equivalently,

$$\operatorname{Max} \tfrac{1}{4}\left(\langle W, \bar{e}\bar{e}^T \rangle - \langle W, X \rangle\right)$$
$$X = uu^T \text{ for some } u \in \{-1,1\}^n$$

Equivalently,

$$\left.\begin{array}{c} \operatorname{Max} \tfrac{1}{4}\left(\langle W, \bar{e}\bar{e}^T \rangle - \langle W, X \rangle\right) \\ \operatorname{diag}(X) = \bar{e} \\ X \succeq 0 \end{array}\right\} \text{SDP relaxation:}$$

$$\boxed{\operatorname{rank}(X)=1} \text{ - relax}$$

SDP relaxation: (of the Max Cut Problem)

$$(P) \quad \text{Max} \; -\frac{1}{4}\langle W, X\rangle \left(+\frac{1}{4}\langle W, \bar{e}\bar{e}^T\rangle\right)$$

$$\text{diag}(X) = \bar{e}$$

$$X \succeq 0$$

Its dual is

$$\text{Min} \; \bar{e}^T y$$

$$\text{Diag}(y) - S = -\frac{1}{4}W$$

$$S \succeq 0$$

Both SDPs have Slater points $\bar{X} := I$, $\eta := \sum_{ij} w_{ij} + 1$, $\bar{y} := \eta\bar{e}$, $\bar{S} := \frac{1}{4}W + \text{Diag}(\bar{y}) \succ 0$.

$d := \text{rank}(X^*)$

Let X^* be an optimal solution of the primal SDP. Then $\exists B \in \mathbb{R}^{n \times d}$ such that $X^* = BB^T$.

$$B^T =: [v^{(1)}, \ldots, v^{(n)}]$$

Note that

$$X^*_{ij} = \langle v^{(i)}, v^{(j)}\rangle.$$

$$X^* = \begin{bmatrix} \langle v^{(1)}, v^{(1)}\rangle & \langle v^{(1)}, v^{(2)}\rangle & \cdots & \langle v^{(1)}, v^{(n)}\rangle \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \langle v^{(n)}, v^{(1)}\rangle & \cdots & \cdots & \langle v^{(n)}, v^{(n)}\rangle \end{bmatrix}$$

We are representing u_i's $\in \{-1,1\}^n$ by $v^{(i)}$'s $\in \mathbb{R}^d$. Note that $\langle v^{(i)}, v^{(i)}\rangle = 1$, so $\|v^{(i)}\|_2 = 1$.
(since $\text{diag}(X^*) = \bar{e}$)

Note that the following nonlinear optimization problem is solvable as an SDP: 2014 07 10

$$\text{Min/Max} \; f_0(\langle v^{(i)}, v^{(j)}\rangle, i, j)$$

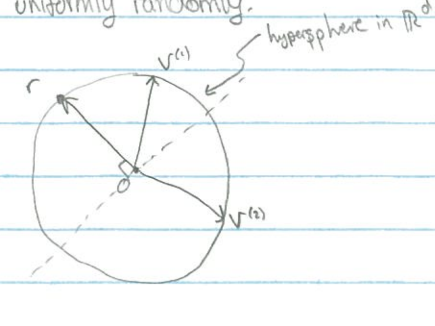
$$\text{s.t.:} \quad f_i(\langle v^{(i)}, v^{(j)}\rangle, i, j) \begin{bmatrix} \geq \\ \leq \\ = \end{bmatrix} b_i \quad \forall i$$

where f_0, f_i are linear functions.

Given these $v^{(i)}$'s in $\mathbb{R}^d$ pick r on the unit hypersphere in $\mathbb{R}^d$ uniformly randomly.

The ~~R~~ random hyperplane technique.

$$U := \{i \in V;\; \langle r, v^{(i)}\rangle \geq 0\}$$



(works perfectly when $d = 1$)

arccos: $\to [0, \pi)$

(2014 07 10)

Lemma 5.1: Let $v^{(i)}$ and r be as above. Then

$$\text{Prob}\{\text{sgn}(\langle r, v^{(i)}\rangle) \neq \text{sgn}(\langle r, v^{(j)}\rangle)\} = \theta/\pi,$$

where

$$\theta = \arccos(\langle v^{(i)}, v^{(j)}\rangle),$$

ie the angle between $v^{(i)}$ and $v^{(j)}$.

Let's approximate arccos by simpler functions.

Lemma 5.2: For every $u \in [-1, 1]$,

$$\frac{\arccos(u)}{\pi} \geq \frac{\rho}{2}(1-u),$$

write arccos as a series, see hint in textbook

$$1 - \frac{1}{\pi}\arccos(u) \geq \frac{\rho}{2}(1+u),$$

where $\rho := 0.87856\ldots$.

Theorem 5.3: The expected value of the cut U generated by the random hyperplane technique is at least

$$\frac{\rho}{4}\sum_{i,j} W_{ij}\left(1 - \langle v^{(i)}, v^{(j)}\rangle\right)$$

$$\geq \rho\left(\text{optimal objective value of (P)}\right).$$

Proof: Let X^* be an optimal solution of (P) with $v^{(i)}$'s defined as above. Then the ~~set~~ optimal ~~value~~ objective value of (P) is

$$\frac{1}{4}\sum_{i,j} W_{ij}\left(1 - X^*_{ij}\right) = \frac{1}{4}\sum_{i,j} w_{ij}\left(1 - \langle v^{(i)}, v^{(j)}\rangle\right).$$

The expected value of the cut U given by the Random Hyperplane Technique is

$$\frac{1}{4}\sum_{i,j} W_{ij}\, 2\, \text{Prob}\{\text{sgn}(\langle r, v^{(i)}\rangle) \neq \text{sgn}(\langle r, v^{(j)}\rangle)\}$$

$$\overset{\text{lem 5.1}}{=} \frac{1}{2}\sum_{i,j} W_{ij}\, \frac{\arccos(\langle v^{(i)}, v^{(j)}\rangle)}{\pi}$$

$$\overset{\text{lem 5.2}}{\geq} \frac{\rho}{4}\sum_{i,j} W_{ij}\left(1 - \langle v^{(i)}, v^{(j)}\rangle\right)$$

$$= \left(\text{optimal objective value of (P)}\right)\cdot \rho.$$

We can derandomize this algorithm.

The algorithm and the proof generalizes to Max 2-Sat. Also in Goemans & Williamson [1995].

There is a more general generalization due to Nesterov:

Let $W \in \mathbb{S}^n$.

$$\underline{f}(W) := \min_{x \in \{-1,1\}^n} x^T W x$$

$$\bar{f}(W) := \max_{x \in \{-1,1\}^n} x^T W x$$

Both problems are NP-hard.

$$\underline{f}(W) = \text{Min } \langle W, X \rangle \quad \text{diag}(X) = \bar{e}, \quad X \succeq 0, \quad \text{rank}(X) = 1$$

SDP relaxation:

strong duality thm applies

$$\underline{F}(W) := \begin{array}{l} \text{Min } \langle W, X \rangle \\ \text{s.t.: } \text{diag}(X) = \bar{e} \\ \quad X \succeq 0 \end{array} = \begin{array}{l} \text{Max } \bar{e}^T y \\ \text{s.t.: } \text{Diag}(y) \preceq W \end{array}$$

Similarly,

$$\bar{F}(W) := \begin{array}{l} \text{Max } \langle W, X \rangle \\ \text{s.t.: } \text{diag}(X) = \bar{e} \\ \quad X \succeq 0 \end{array} = \begin{array}{l} \text{Min } \bar{e}^T y \\ \text{s.t.: } \text{Diag}(y) \succeq W. \end{array}$$

Note that $\underline{f}(-W) = -\bar{f}(W)$ and $\underline{F}(-W) = -\bar{F}(W)$.

Note that we can attack Max Cut Problems with mixed (negative as well as positive) weights with this general setup.

$\underline{F}(W), \bar{F}(W)$ give us a natural "scale" or "range" for the objective values of $\bar{f}(W)$ & $\underline{f}(W)$ that is efficiently computable.

$$\underline{F}(W) \le \underline{f}(W) \le \bar{f}(W) \le \bar{F}(W).$$

We will first deal with the case that $W \in \mathbb{S}^n_+$.

not need here

Lemma 5.11: $\bar{f}(W) = \text{Max } z^T W z$

$$z = \text{Sgn}(Br)$$
$$\|r\|_2 = 1$$
$$\|B^T e_i\|_2 = 1, \quad \forall i \in \{1, \dots, n\}$$
$$B \in \mathbb{R}^{n \times n}, \ r \in \mathbb{R}^n$$

Think $B^T = [v^{(1)}, \dots, v^{(n)}]$

2014 07 12

Proof: By definition, $\mathrm{sign}(Br) \in \{-1,1\}^n$. Therefore, $\bar{f}(W) \geq$ RHS.

Let $\bar{x} \in \{-1,1\}^n$ attaining $\bar{f}(W)$. Let $r \in \mathbb{R}^n$ be such that $\|r\|_2 = 1$. Then define

$$B^T e_i := \begin{cases} r & \text{if } \bar{x}_i = 1, \\ -r & \text{if } \bar{x}_i = -1, \end{cases}$$

$\forall i \in \{1, \dots, n\}$. With

$$\xi = \mathrm{sgn}(Br) = \bar{x},$$

the objective value of this feasible solution is

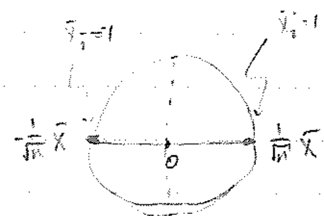
$$\xi^T W \xi = \bar{x}^T W \bar{x} = \bar{f}(W). \qquad \square$$

We will think about $r \in \mathbb{R}^n$, $\|r\|_2 = 1$ as the random variable vector.

Lemma 5.12. For every $W \in \mathbb{S}^n$

$$\begin{aligned} \bar{f}(W) = \mathrm{Max} \quad & E_r(\xi^T W \xi) \\ \text{s.t.:} \quad & \xi = \mathrm{sign}(Br) \\ & \|B^T e_i\|_2 = 1 \quad \forall i \in \{1, \dots, n\} \\ & \|r\|_2 = 1 \\ & B \in \mathbb{R}^{n \times n},\ r \in \mathbb{R}^n \end{aligned}$$

(Note the optimization variables are B and ξ.)



Proof: By the previous lemma, $\bar{f}(W) \geq$ RHS.

Let $\bar{x} \in \{-1,1\}^n$ attain $\bar{f}(W)$. Consider $B := \frac{1}{\sqrt{n}} \bar{x}\bar{x}^T$. Note

$$B^T e_i = \begin{cases} \frac{1}{\sqrt{n}} \bar{x} & \text{if } \bar{x}_i = 1 \\ -\frac{1}{\sqrt{n}} \bar{x} & \text{if } \bar{x}_i = -1 \end{cases}$$

and so $\|B^T e_i\|_2 = 1 \ \forall i$.

$$\xi^T W \xi = \sum_{i,j} W_{ij} \xi_i \xi_j = \sum W_{ij} \mathrm{sgn}(\langle r, B^T e_i \rangle) \mathrm{sgn}(\langle r, B^T e_j \rangle)$$

$$\to E_r\left(\mathrm{sgn}(\langle r, B^T e_i \rangle) \mathrm{sgn}(\langle r, B^T e_j \rangle)\right) = \begin{cases} 1, & \text{if } \bar{x}_i = \bar{x}_j, \\ -1, & \bar{x}_i \neq \bar{x}_j. \end{cases}$$

measure theory

For this choice of B,

$$E(\xi^T W \xi) = \sum_{i,j} W_{ij} \bar{x}_i \bar{x}_j = \bar{x}^T W \bar{x} = \bar{f}(W).$$

For $X \in [-1,1]^{n\times n}$ we define $\arcsin(X) \in \mathbb{R}^{n\times n}$ by
$$(\arcsin(X))_{ij} := \arcsin(X_{ij}).$$

Theorem 5.13: For every $W \in \mathbb{S}^n$,
$$\bar{f}(W) = \text{Max } \tfrac{2}{\pi}\langle W, \underline{\underline{\arcsin(X)}}\rangle$$
$$\text{s.t.: } \operatorname{diag}(X) = \bar{e}$$
$$X \succeq 0$$

we saw earlier $\longrightarrow$ $\left(\text{Note } \bar{f}(W) = \text{Max } \langle W, X\rangle;\ \operatorname{diag}(X) = \bar{e},\ X \succeq 0,\ \boxed{\operatorname{rank}(X) = 1.}\right)$

Proof: Note that the RHS is well-defined and attained. (Feasible region is non-empty, compact; $\forall$ feas. X, $X \in [-1,1]^{n\times n}$, the objective function is well-defined & continuous over the feasible region.)

Take an optimal solution of the RHS problem in lemma 5.12. Then
$$E_r(\operatorname{sgn}(\langle r, B^T e_i\rangle)\operatorname{sgn}(\langle r, B^T e_j\rangle)) = \operatorname{Prob}(\operatorname{sgn}(\langle r, B^T e_i\rangle) = \operatorname{sgn}(\langle r, B^T e_j\rangle))$$
$$- \operatorname{Prob}(\ \text{"} \neq \text{"}\ )$$
$$= 1 - 2\operatorname{Prob}(\ \text{"} \neq \text{"}\ )$$
$$\overset{\text{lemma 5.1}}{=} 1 - 2/\pi \cdot \arccos(\langle B^T e_i, B^T e_j\rangle)$$
$$= \tfrac{2}{\pi}\arcsin((BB^T)_{ij})$$

Since $BB^T \succeq 0$ and $\operatorname{diag}(BB^T) = \bar{e}$, $\bar{X} := BB^T$ is a feasible solution of the RHS.
$$\tfrac{2}{\pi}\langle W, \arcsin(\bar{X})\rangle = \sum_{i,j} W_{ij} E_r(\operatorname{sgn}(\langle r, B^T e_i\rangle)\operatorname{sgn}(\langle r, B^T e_j\rangle)) \underset{\text{previous lemma}}{=} \bar{f}(W).$$

Therefore $\bar{f}(W) \leq \text{RHS}$.

For the other direction, let $\hat{X}$ be an optimal solution of the RHS. Define $\hat{X} =: BB^T$, $B \in \mathbb{R}^{n\times n}$. (can do as $\hat{X} \succeq 0$)

Note $\|B^T e_i\|_2^2 = e_i^T BB^T e_i = \hat{X}_{ii} = 1\ \forall i$. Therefore B is a feasible solution of the optimization problem in the RHS of lemma 5.12. Its objective value is
$$\underbrace{\tfrac{2}{\pi}\sum_{ij} W_{ij}\arcsin((BB^T)_{ij})}_{\substack{\wedge\!\mid \text{ — lemma 5.12} \\ \bar{f}(W)}} = \tfrac{2}{\pi}\langle W, \arcsin(\hat{X})\rangle$$

■

see text for details

Lemma 5.14: For every $X \in \mathbb{S}^n_+$ such that $|X_{ij}| \leq 1\ \forall i,j$, we have
$$\arcsin(X) \succeq X.$$

for the proof all
you have to do is,
look at the board and smile

(2014 07 12)

Theorem 5.15: For every $W \in \mathbb{S}^n_+$, we have — first time using $W \succeq 0$

$$\bar{f}(W) \geq \frac{2}{\pi} \operatorname{Max} \langle W, X \rangle \quad \begin{array}{l} \operatorname{diag}(X) = \bar{e} \\ X \succeq 0 \end{array}$$

Proof: By ~~5.14~~ 5.13,

$$\bar{f}(W) = \frac{2}{\pi} \operatorname{Max} \langle W, \arcsin(X) \rangle \quad \begin{array}{l} \operatorname{diag}(X) = \bar{e} \\ X \succeq 0 \end{array} \quad \geq \quad \frac{2}{\pi} \operatorname{Max} \langle W, X \rangle \quad \begin{array}{l} \operatorname{diag}(X) = \bar{e} \\ X \succeq 0 \end{array}$$

since $W \succeq 0$ and by Lem 5.14 ∎

Max Cut is a special case via

$$W = \begin{cases} -w_{ij} & \text{if } \{i,j\} \in E \\ \sum_k w_{ik} & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

For all $w \in \mathbb{R}^E_+$, $W \succeq 0$ because it is diagonally dominant.

In general, when $W \in \mathbb{S}^n$, note $\forall y \in \mathbb{R}^n$

$$\bar{f}(W + \operatorname{Diag}(y)) = \bar{f}(W) + \bar{e}^T y$$

$x^T \operatorname{Diag}(y_1, \ldots, y_n) x = \sum_{i=1}^n y_i x_i^2 = \bar{e}^T y$

$$\bar{F}(W + \operatorname{Diag}(y)) = \bar{F}(W) + \bar{e}^T y$$

For all feas. X,

$$\langle W + \operatorname{Diag}(y), X \rangle = \langle W, X \rangle + \langle y, \operatorname{diag}(X) \rangle = \langle W, X \rangle + \bar{e}^T y$$

2014 07 15

(recall)

$W \in \mathbb{S}^n$

$$\underline{F}(W) := \min \langle W, X \rangle \quad \begin{array}{l} \operatorname{diag}(X) = \bar{e} \\ X \succeq 0 \end{array} \quad = \operatorname{Max} \bar{e}^T y \quad \operatorname{Diag}(y) \preceq W$$

$$\bar{F}(W) := \operatorname{Max} \langle W, X \rangle \quad \begin{array}{l} \operatorname{diag}(X) = \bar{e} \\ X \succeq 0 \end{array} \quad = \operatorname{Min} \bar{e}^T y \quad \operatorname{Diag}(y) \succeq W$$

$\forall y \in \mathbb{R}^n$, $\bar{f}(W + \operatorname{Diag}(y)) = \bar{e}^T y + \bar{f}(W)$, same for $\underline{f}, \bar{F}, \underline{F}$.

finish reading chapter 5

Theorem 5.16: For every $W \in \mathbb{S}^n$ we have

$$\underline{F}(W) \le \underline{f}(W) \le \tfrac{2}{\pi}\underline{E}(W) + (1-\tfrac{2}{\pi})\overline{F}(W) \le (1-\tfrac{2}{\pi})\underline{E}(W) + \tfrac{2}{\pi}\overline{F}(W) \le \overline{f}(W) \le \overline{F}(W).$$

Proof: Let $\bar{y}$ be an optimal solution of the dual SDP defining $\overline{F}(W)$. Then

$$\begin{aligned}
\overline{F}(W) - \underline{f}(W) &= \bar{e}^T\bar{y} - \underline{f}(W) \\
&\cancel{\ge \underline{f}(\mathrm{Diag}(\bar{y}) - W)} \quad \cancel{\succeq 0 \text{ by choice of } \bar{y}} \\
&= \bar{e}^T\bar{y} + \overline{f}(-W) \\
&= \overline{f}(\underbrace{\mathrm{Diag}(\bar{y}) - W}_{\succeq 0}) \\
\text{thm 5.15} \rightarrow &\ge \tfrac{2}{\pi}\overline{F}(\mathrm{Diag}(\bar{y}) - W) \\
&= \tfrac{2}{\pi}(\bar{e}^T\bar{y}) + \tfrac{2}{\pi}\overline{F}(-W) \\
&\cancel{= \tfrac{2}{\pi}(\bar{e}^T\bar{y})} \\
&= \tfrac{2}{\pi}(\overline{F}(W) - \underline{F}(W))
\end{aligned}$$

We proved

$$\underline{f}(W) \le \tfrac{2}{\pi}\underline{E}(W) + (1-\tfrac{2}{\pi})\overline{F}(W)$$

To prove the remaing non-trivial inequality, estimate $\overline{f}(W) - \underline{E}(W)$ as above. ∎

read some text at this point

Note that we can also handle linear terms:

$$\max_{x \in \{-1,1\}^n} \; x^T W x + w^T x$$

$$\widetilde{W} := \begin{bmatrix} 0 & \tfrac{1}{2}w^T \\ \tfrac{1}{2}w & W \end{bmatrix}$$

Consider $\displaystyle \max_{\begin{bmatrix} x_0 \\ x \end{bmatrix} \in \{-1,1\}^n} [x_0, x^T]\, \widetilde{W}\, [x_0\ x]^T$.