15859GG: Lecture 3
Quadratic Optimization on Hypercube

Last time: $\max_{x \in \{\pm 1\}^n} f_G(x) = \max_{x \in \{\pm 1\}^n} x^T L_G x.$

L_G: Laplacian matrix, $L_G = D - A$

$$D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}, \quad A: \text{0-1 adj}$$

$L_G \succeq 0$ PSD.

Hom quadratic poly $= x^T B x.$

① take pseudo-distribution μ

② WLOG: $\tilde{\mathbb{E}}_\mu x = 0$, $\tilde{\mathbb{E}}_\mu x x^T$

③ $g \sim N(0, \tilde{\mathbb{E}}_\mu x x^T)$

Set $\hat{x}_i = \text{sign}(g_i)$ $\forall i$: $\hat{x} = \text{sign}(g)$

Algorithm: 1. Find a p.d. of deg 2. μ that maximizes $\tilde{\mathbb{E}}_\mu f_G(x)$. ✓

2. Apply above rounding algorithm:

$\tilde{\mathbb{E}}_\mu f_G(x)$

$\geq \max \tilde{\mathbb{E}} f_G(x)$

$\mu \quad -2^{-n}$

Suppose I know how to sample $N(0, \mathbb{I})$.

$\Sigma \leftarrow$ target covariance

$\Sigma : (\Sigma^{1/2})^2$

Box-Muller transform

example of quadratic optimization.

$h \sim N(0, \mathbb{I})$, $\quad \underbrace{g = \Sigma^{1/2} \cdot h}$.

Input: $B \in \mathbb{Q}^{n \times n}$

Goal: $\max_{x \in \{\pm 1\}^n} x^T B x$.

(1) B is PSD : $\frac{\pi}{2}$ approx. (Nesterov).

(2) $\text{supp}(B) : \{\{i,j\} \mid B_{ij} \neq 0\}$ (Alon-Naor).

$\downarrow$ again: constant factor approx.

$\lesssim 1.748\ldots$ (Grothendieck's inequality) (Krivine '77)

(3)

Theorem (Nesterov): Let B be PSD. Then

$\frac{\text{Opt}(B)}{C} - x^T B x$ has a deg 2 SoS cert

Further, given a p.d.of

for $c = \frac{2}{\pi} \sim 0.63$:

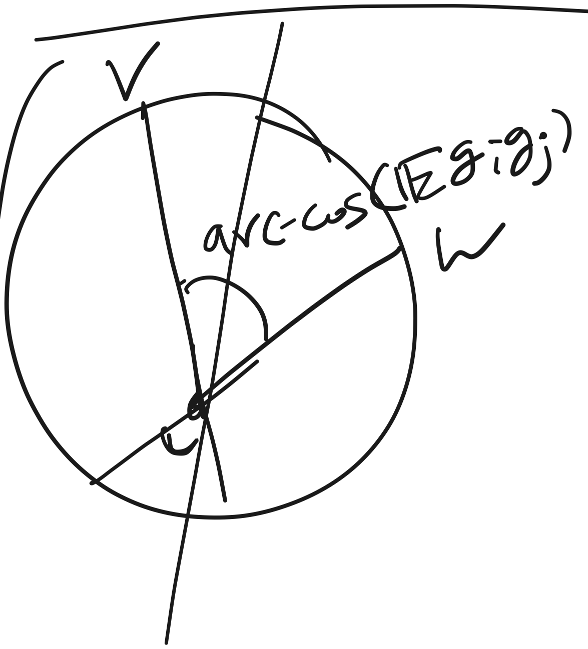
deg 2 : μ, can efficiently sample $\hat{x}$ s.t.

$$\hat{x}^T B \hat{x} \geq \left(\frac{2}{\pi} - 2^{-n}\right) \cdot \tilde{\mathbb{E}}_\mu x^T B x \quad \text{(in poly time)}$$

<u>Proof:</u> ① $g \sim N(0, \tilde{\mathbb{E}}_\mu x x^T)$

② $\hat{x}_i = \text{sign}(g_i) \quad \forall\, 1 \leq i \leq n$

<u>Recall:</u> $\Pr[\hat{x}_i \neq \hat{x}_j] = \dfrac{\text{arc-cos}(\mathbb{E}\, g_i g_j)}{\pi}$



$g_i = \langle h, v \rangle \qquad h = N(0, I_2)$

$g_j = \langle h, w \rangle$

$\|v\|_2^2 = \|w\|_2^2 = 1$

$\langle v, w \rangle = \mathbb{E}\, g_i g_j$

$\Pr[\text{sgn}(g_i) \neq \text{sign}(g_j)]$

$= \dfrac{\text{arc-cos}(\mathbb{E}\, g_i g_j)}{\pi}$

<u>Prop:</u> $\mathbb{E}\, \hat{x}_i \cdot \hat{x}_j = \frac{2}{\pi} \text{arc-sin}(\mathbb{E}\, g_i g_j)$

<u>Pf:</u> $\mathbb{E}\, \hat{x}_i \cdot \hat{x}_j = (-1) \cdot \dfrac{\text{arc-cos}(\mathbb{E}\, g_i g_j)}{\pi}$

$+ 1 \cdot (1 - (\ldots$

$$= 1 - \frac{2\,\text{arc-cos}(\mathbb{E} g_i g_j)}{\pi}$$

$$= 2\left(\frac{\frac{\pi}{2} - \text{arc-cos}(\mathbb{E} g_i g_j)}{\pi}\right)$$

$$= \frac{2}{\pi}\,\text{arc-sin}(\mathbb{E} g_i g_j).$$

__Cool Lemma:__ Let $M \in \mathbb{R}^{n\times n}$ be PSD, $\text{diag}(M) = \vec{1}$.

Let $f : \mathbb{R} \to \mathbb{R}$. $\quad f[M] \in \mathbb{R}^{n\times n}$.

$$f[M]_{i,j} = f(M_{i,j}).$$

Suppose f has a +ve coeff Taylor series that's uniformly convergent on $[-1,1]$. Then,

$$f[M] \succeq 0.$$

__Schur Product Thm:__ M, M' PSD.

__Then__ $M \circ M'$: $M \circ M'(i,j) = M(i,j) \cdot M'(i,j)$

$$M \circ M' \succeq 0 \quad \text{(PSD)}.$$

Pf. of Cool lemma: $\quad [M]^2 = M \circ M.$

Pf of cool lemma : $[M]^i \succeq 0$

$[M]^i \succeq 0 \quad \forall i$

$\Rightarrow \quad \sum c_i [M]^i \succeq 0 \quad \forall c_i \geq 0$

$\Rightarrow \quad \left(\sum c_i x^i\right)[M] \succeq 0$

$\square$

Pf of Schur product thm :

$M = \sum \lambda_i \cdot v_i v_i^T \quad , \quad \lambda_i \geq 0$

$M' = \sum \lambda_i' \cdot v_i' \cdot v_i'^T$

$$M \circ M' = \sum_{i,j} \lambda_i \cdot \lambda_j' \left(v_i v_i^T\right) \circ \left(v_j' \cdot v_j'^T\right)$$

$$= \sum_{i,j} \lambda_i \cdot \lambda_j' \left(v_i \circ v_j'\right) \cdot \left(v_i \circ v_j'\right)^T$$

$\Rightarrow \quad M \circ M' \succeq 0 \qquad \square$

Corollary: $\text{arc-sin}[M] \succeq M$ whenever M is PSD with $\text{diag} = \vec{1}$.

$\Updownarrow$

$\text{arc-sin}[M] - M \succeq 0$

Pf:

$$\text{arc-sin}[M] - M = \sum_{k=1}^{\infty} \frac{1 \times 3 \times 5 \times \ldots \times (2k-1)}{2^k \cdot k! \cdot (2k+1)} [M]^{2k+1}$$

$\uparrow \succeq 0$ $\square$

Pf of Nesterov's Thm:

recall: $\mathbb{E}\, \hat{x}_i \hat{x}_j = \frac{2}{\pi} \text{arc-sin}(\mathbb{E} g_i g_j)$.

$$\mathbb{E}_{\hat{x}}\, \hat{x}^T B \cdot \hat{x} = \sum_{i,j} B_{i,j} \cdot \hat{\mathbb{E}}\, \hat{x}_i \cdot \hat{x}_j$$

$$= \frac{2}{\pi} \underbrace{\sum_{i,j} B_{i,j} \cdot \text{arc-sin}(\mathbb{E} g_i g_j)}$$

$$= \frac{2}{\pi} \underbrace{\text{tr}(B \cdot \text{arc-sin}[\mathbb{E} g g^T])}.$$

Claim: $B \succeq 0$. $C \succeq 0$ $\text{tr}(B \cdot C) \succeq 0$. (self-duality)

$$\Rightarrow \frac{2}{\pi}\left(\text{tr}\, B \cdot (\text{arc-sin}[\mathbb{E} g g^T] - \mathbb{E} g g^T)\right) \succeq 0$$

$$\rightarrow \frac{2}{\pi} \text{tr}(B \cdot \text{arc-sin}[\mathbb{E} g g^T])$$

$$\succeq \frac{2}{\pi} \text{tr}(B \cdot \mathbb{E} g g^T)$$

$$= \frac{2}{\pi} \tilde{\mathbb{E}}_{\mu} \cdot x^T B x$$

$$\mathbb{E}\, \hat{x}^T \cdot B \hat{x} = \frac{2}{\pi} \tilde{\mathbb{E}}_{\mu}\, x^T B x$$

Next up: B is arbitrary $n \times n$ matrix. : $\boxed{\operatorname{diag}(B) = 0}$
(Symmetric)

$$\max_{x \in \{\pm 1\}^n} x^T B x \quad : \operatorname{opt}(B).$$

Prop (enough to optimize on $[-1,1]^n$)
Let $y \in [-1,1]^n$. Then $\exists\, \hat{y} \in \{\pm 1\}^n$ s.t

$$\hat{y}^T \cdot B \cdot \hat{y} \geq y^T \cdot B \cdot y.$$

Corollary: $\operatorname{opt}(B) \geq 0$. (Choose $y = 0$ & apply prop)

Proof: Give a dist^n on $\{\pm 1\}^n \leftarrow \hat{y}$.

s.t. $\mathbb{E}\, \hat{y}_i \cdot \hat{y}_j = y_i \cdot y_j \quad \forall\, i \neq j.$

$$\hat{y}_i = \begin{cases} 1 & \text{w.p. } \frac{1}{2} + \frac{y_i}{2} \\ -1 & \text{w.p. } \frac{1}{2} - \frac{y_i}{2} \end{cases} \qquad \text{indep}$$

Then: $\mathbb{E}\, \hat{y}_i \cdot \hat{y}_j = (\mathbb{E}\, y_i) \cdot (\mathbb{E}\, y_j)$

$$= \frac{1}{2} + \frac{y_i}{2} - \left(\frac{1}{2} - \frac{y_i}{2}\right) = \underline{y_i}$$

$$\Rightarrow \underline{\mathbb{E}\, \hat{y}^T \cdot B \hat{y} = y^T \cdot B y}$$

Prop: $\text{opt}(B) \geq \frac{1}{n} \cdot \sum_{i,j} |B_{i,j}|$. $\boxed{n \text{ is even}}$

Proof: 1. Select a random matching of the complete graph.

2. $\{i,j\}$ is in the matching.

$\hat{x}_i$: random entry in $\{\pm 1\}$

$\rightarrow \hat{x}_j = \hat{x}_i$ iff $B_{i,j} \geq 0$

$\text{Sign}(\hat{x}_i \cdot \hat{x}_j) = B_{ij}$ for every edge $\{i,j\}$ in the matching.

if $\{i,j\}$ is not in matching then $\mathbb{E}\, \hat{x}_i \cdot \hat{x}_j = 0$.

$$\mathbb{E}\, \mathbb{E}\, \hat{x}^T B \hat{x} \geq \mathbb{E} \sum_{i,j} Q_{i,j} |B_{i,j}|$$

$$\geq \frac{1}{n} \cdot \sum_{i,j} |B_{i,j}|$$

<u>Theorem (Megretski '01, Charikar–Wirth '04)</u>

Let n be large enough (>60). Then for $C = O(\log(n))$. ($C \lesssim 4\log(n)$). Then.

$$\frac{opt(B)}{C} - x^T B x \text{ has a deg 2 SoS cert.}$$

<u>Further</u>, $\forall$ p.d. μ on deg 2 on $\{\pm1\}^n$, $\exists$ a dist^n on $\{\pm 1\}^n$ (eff samplable) such that

$$\mathbb{E}\, \hat{x}^T B \hat{x} \geq \frac{1}{C} \cdot \tilde{\mathbb{E}}_\mu\, x^T B x$$

<u>Proof</u>: ① enough to round to $[-1,1]^n$. $\in$

ⓘ choose $g \sim N(0, \tilde{\mathbb{E}}_\mu x x^T)$

<u>Obs^n</u>: $\Pr[\exists i : |g_i| \geq 10\sqrt{\log(n)}] \leq 1/n^2$

$(\Rightarrow$ w.p. $1 - 1/n^2$, $|g_i| \leq 10\sqrt{\log(n)}\ \forall i$

Choose $v = \frac{g_i}{10\sqrt{\log(n)}}$.

then w.p. $1 - 1/n^2$, $v \in [-1,1]^n$.

What happens to quadratic form?

$$\mathbb{E}\left[v^T B v \mid |g_i| \le 10\sqrt{\log(n)}\right]$$

$$= \frac{1}{100\log(n)} \mathbb{E}\left[g^T B g \mid |g_i| \le 10\sqrt{\log(n)}\right]$$

relate to $\mathbb{E}[g^T B g]$.

Final rounding:

1. For every i: let

$$\hat{x}_i = \frac{g_i}{10\sqrt{\log(n)}} \quad \text{if } |g_i| \le 10\sqrt{\log(n)}$$

o/w: $\hat{x}_i = \frac{g_i}{|g_i|}$.

Can prove that $\mathbb{E}\, \hat{x}_i \cdot \hat{x}_j \geq \mathbb{E}\, g_i g_j$

Can prove ... $\mathbb{E}$ ... $\frac{}{10 \log(n)}$

$- \mathbb{E}[\Delta_{i,j}]$.

& can prove that: $\mathbb{E} \sum_{i,j} B_{i,j} \cdot \Delta_{i,j} \leq \frac{opt(B)}{n^2}$.

Special Case of: $R \cdot PR^2$ rounding

Feige-Langberg

(1) $\tilde{\mathbb{E}}_\mu xx^T$

$\downarrow$ modify

(2)

(3) Sample $g \sim$ modified $(\tilde{\mathbb{E}}_\mu xx^T)$

(4) Scale g to put it in $[-1,1]^n$

(5) randomized rounding

Suppose you want the best possible rounding for Max-Cut.

→ search over all possible RPR^2

→ output the best

Hardness: $C = \Omega(\log(n))$

$U.G.C. \Rightarrow \Theta(\log(n)^r)$ is a lower bound

Exploiting structure in support(B).

Support(B) = bipartite

$$\begin{pmatrix} x \\ y \end{pmatrix} \begin{array}{c} \leftarrow n \\ \leftarrow n \end{array} \quad \begin{array}{cc} n & n \\ \end{array}$$

$$\begin{pmatrix} 0 & B' \\ B'^T & 0 \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \end{pmatrix}^T \cdot B \begin{pmatrix} x \\ y \end{pmatrix} = 2 x^T B' y$$

↳ equivalently: $B \in \mathbb{R}^{n \times n}$ arbitrary

$$\boxed{\text{Goal: } \max_{x,y \in \{\pm 1\}^n} \underline{x}^T \cdot B \underline{y}}$$: operator norms

Operator Norms: Usual op norm: $\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2$

$$\|A\| = \inf\{c : \underbrace{\|Ax\| \leq c \cdot \|x\|} \ \forall x \in \mathbb{R}^n\}$$

$$\|A\|_{\infty \to 1} : \max_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|_1}{\|x\|_\infty}.$$

$$\boxed{\|y\|_1 = \sum |y_i| \quad , \quad \|y\|_\infty : \max_i |y_i|.}$$

$$\|A\|_{\infty \to 1} = \max_{x, y \in \{\pm 1\}^n} x^T A y$$

$$\max_{y \in \{\pm 1\}^n} \langle Ax, y \rangle = \|Ax\|_1$$

<u>Theorem:</u> $\exists$ a constant K_G

$K_G \cdot opt(B) - x^T B y$ has a deg 2 SoS cert.

$$\mathbb{E}\, \hat{x}^T \cdot B \hat{x} \geqslant \frac{1}{K_G} \cdot opt(B).$$

do not know what K_G is.

$$\frac{\pi}{2} \leq K_G < \underbrace{\frac{\pi}{2 \cdot \ln(1+\sqrt{2})}}$$

<u>Krivine's Constant</u>

Krivine '73: conjectured tight!

Krivine ++ : Conjectured tight!!

{ Braverman – Makarychev – Makarychev – Naor '11 refuted Krivine's conj.

Proof: $M = \tilde{\mathbb{E}}_\mu x x^T$, $\mathbb{E}_\mu x = 0$

1. Let $M' = \begin{pmatrix} \sinh[c \cdot \tilde{\mathbb{E}}_\mu x x^T] & \sin[c \cdot \tilde{\mathbb{E}} x y^T] \\ \sin[c \cdot \tilde{\mathbb{E}} y x^T] & \sinh[c \cdot \tilde{\mathbb{E}}_\mu y y^T] \end{pmatrix}$

has diagonals $\vec{1}$. $c = \dfrac{1}{\ln(1+\sqrt{2})}$ $(M' \succeq 0)$

2. Choose $(g, h) \sim N(0, M')$.

3. Output: $\hat{x} = \operatorname{sign}(g)$, $\hat{y} = \operatorname{sign}(h)$

Main claim:
1. $M' \succeq 0$
2. $\operatorname{diag}(M') = \vec{1}$
3. $\underbrace{\tilde{\mathbb{E}}_\mu x \cdot y^T}_{\text{p.d.}} = K \cdot \mathbb{E}_{g,h} \underbrace{\operatorname{sign}(g) \operatorname{sign}(h)^T}_{\text{boolean dist}^n}$

Proof of Thm using main claim:

$$\operatorname{tr}(C\cdot D) = \Sigma_{i,j} C_{i,j}\cdot D_{i,j}$$

$$\underbrace{\tilde{\mathbb{E}}_{\mu}\, x^T B y} = \operatorname{tr}\left(B\cdot \tilde{\mathbb{E}}_{\mu}\cdot x\cdot y^T\right)$$

$$= K \operatorname{tr}\left(B\cdot \mathbb{E}_{g,h}\, \operatorname{sign}(g)\cdot \operatorname{sign}(h)^T\right)$$

$$= K\cdot \underbrace{\mathbb{E}\, \hat{x}^T\cdot B\cdot \hat{y}}\,.$$

Pf of Main claim:

<u>Recall</u>: $\mathbb{E}_{g,h}\, \operatorname{sign}(g)\cdot \operatorname{sign}(h)^T$

$$= \frac{2}{\pi}\, \underbrace{\text{arc-sin}\left[\mathbb{E}\, g\cdot h^T\right]}$$

$$\overset{??}{=} K\cdot \tilde{\mathbb{E}}_{\mu}\, x y^T$$

Then must choose

$$\mathbb{E}\, g h^T = \sin\left[c\cdot \tilde{\mathbb{E}}_{\mu}\, x\cdot y^T\right].$$

$$\sin\left(c\cdot \tilde{\mathbb{E}}\, x y^T\right)$$

$$M' : \begin{pmatrix} \sinh(c\,\tilde{\mathbb{E}}_\mu xx^T) & \sin(c\,\tilde{\mathbb{E}}_\mu xy^T) \\ \sin(c \cdot \tilde{\mathbb{E}}_\mu yx^T) & \sinh(c\,\tilde{\mathbb{E}}_\mu yy^T) \end{pmatrix} \succeq_M \begin{pmatrix} \tilde{\mathbb{E}}_\mu xx^T & \tilde{\mathbb{E}}_\mu xy^T \\ \tilde{\mathbb{E}}_\mu yx^T & \tilde{\mathbb{E}}_\mu yy^T \end{pmatrix}$$

<u>Key Fact</u>: $\sin(t) : \sum_{k=0}^{\infty} a_k \cdot t^k$

<u>then</u>: $\sinh(t) : \sum_{k=0}^{\infty} |a_k| \cdot t^k.$

<u>Prop</u>: $M \succeq 0$.

Then so is: $$\begin{pmatrix} v \\ w \end{pmatrix}^T \begin{pmatrix} [M_{11}]^i & -[M_{12}]^i \\ -[M_{21}]^i & [M_{22}]^i \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix}$$

<u>Pf</u>: $$= \begin{pmatrix} v \\ -w \end{pmatrix}^T \begin{pmatrix} [M_{11}]^i & [M_{12}]^i \\ \cdots & \cdots \end{pmatrix} \begin{pmatrix} v \\ -w \end{pmatrix}$$

$$\left(\vdots \ W \right) \left([M_{12}]^i \ [M_{22}]^i \right) \left(\ \right)$$

<u>Corollary</u>: $\forall \ p(t) = \sum_{i=1}^{k} a_i \cdot t^i$, $p_+ = \sum_{i=1}^{k} |a_i| \cdot t^i$

If $M \succeq_c 0$, then

$$\begin{pmatrix} p_+(M_{11}) & p(M_{12}) \\ p(M_{21}) & p_+(M_{22}) \end{pmatrix} \succeq_c 0$$

$\hookrightarrow$ can extend this to uniformly convergent Taylor series on $(-1,1)$.

<u>Corollary</u>:

$$\begin{pmatrix} \sinh(cM_{11}) & \sin(cM_{12}) \\ \sin(cM_{21}) & \sinh(cM_{22}) \end{pmatrix} \succeq_c 0.$$

<u>For part 1</u>: must choose c so that

$$\sinh(c) = \underline{1}$$

the c needs to be $\ln(1+\sqrt{2})$

$$\boxed{K_G \lesssim \frac{\pi}{2\ln(1+\sqrt{2})}}$$

Know: $K_G < \dfrac{\pi}{2\ln(1+\sqrt{2})}$.

Cut-Norm (B): $B \in \mathbb{R}^{n\times n}$.

$$\|B\|_{cut} = \max_{\substack{S\subseteq[n]\\ T\subseteq[n]}} \sum_{\substack{i\in S\\ j\in T}} B_{i,j}$$

Fact: $\|B\|_{\infty\to 1}$ is within a factor 4 of $\|B\|_{cut}$

Alon-Makarychev-Makarychev-Naor ??

max $x^T B x$ over

The best approx. to $\max_{x \in \{\pm 1\}^n}$ ...

all matrices B that have support graph G

is "~~bet~~": $\left[\Theta(\log(\omega(G))), \Theta(\log \chi(G))\right]$

<u>Megretski '01</u> : $\Theta(\log(\vartheta(\bar{G})))$.

<u>~~RPR~~^{2} is optimal for Max-Cut</u>

α_{opt} : <u>O'donnell-Wu '03</u>

<u>Constant-model roundings</u> : <u>Raghavendra '08</u>

<u>UGC</u> $\Rightarrow$ NP-hard to beat the performance of

$\longrightarrow$ "Constant model roundings" of deg 2 p.d.

$\forall$ generalized CSPs.