SDPs:

A $n\times n$ symmetric matrix A is psd if :- $(A \succeq 0)$

[Löwner ordering $A \succeq B \overset{def}{\iff} A - B \succeq 0$.]

HW 5/6

(a) $x^T A x \geq 0 \quad \forall x$.

(b) All evs of A are nonnegative

(c) $A = PDP^T$ where $D = diag(\lambda_1, \lambda_2 \cdots \lambda_n)$

$= \sum_i \lambda_i v_i v_i^T$ cols of P are eigenvectors v_i of A.

define $A^{1/2} = PD^{1/2}P^T$ (as before) and hence $A = A^{1/2} A^{1/2}$.

(d) there exist jointly distributed random variables $X_1, X_2 \ldots X_n$ such that

$A_{ij} = E(X_i X_j) \quad \forall ij$

(e) there exist vectors $V_1, \ldots V_n$ such that

$A_{ij} = \langle V_i, V_j \rangle$

(f) A can be written as BB^T for some matrix B (just take rows of $B = v_i$)

~~SDP semidefinite program~~

Fact: given a matrix A, can check if it is psd in (strongly) polynomial time.

In fact if A is not psd, can return x st $x^T A x < 0$.

OK: what is a semidefinite program?

(i) It's a linear program where the variables are indexed by $i,j \in [N]$ (wlog)

say $(x_{ij})_{ij \in N}$

And we have an extra constraint $X = (x_{ij}) \succeq 0$ (psd)

(ii) It's a linear program where instead of linear constraints on variables (and linear obj fn on vars) we have constraints/obj on inner products of vector variables $\sum_{ij} a_{ij}^{\ell} \langle V_i, V_j \rangle \geq b_i \quad \forall \ell \in 1..m$.

By definition (e) these are equivalent.

Thm: Using the ellipsoid algorithm, we can solve SDPs in time $poly(\text{size of instance}, \log \frac{1}{\epsilon})$ to within $(1+\epsilon)$

(assuming bounds on the solution size !! ← fix)

Why is this useful? Two examples.

① Eigenvalues: given symmetric A suppose we want to find its largest eigenvalue λ_1.

then SPS we say

$$\boxed{\begin{array}{ll} \max & A \bullet X \\ \text{st} & X \bullet I = 1 \\ & X \succeq 0 \end{array}} \quad \text{or} \quad \boxed{\begin{array}{ll} \min & t \\ \text{st} & tI - A \succeq 0. \end{array}}$$

Here $A \bullet B = \sum_{ij}' a_{ij} b_{ij} = tr(A^T B)$.

is called the Schur product.

this is same as $\{\max \sum' a_{ij} x_{ij} \mid \sum x_{ii} = 1, \; X \succeq 0\}$

sps $X = x_i x_i^T$ (rank 1)

then to max $A \circ X$ (" $= A \bullet x_i x_i^T = x_i^T A x_i$) we'd want to take $x_i = v_1$ to get λ_1

Now ~~the max is obtained~~

for general $X = \sum \mu_i x_i x_i^T$ might as well put all mass on one μ_i, with $x_i = v_i$

~~Two natural ways to write SDPs~~: (later).

② Max-Cut: $x_v \in \{0,1\}$, $y_e \in \{0,1\}$ max $\sum' y_e$ st y_e & x are "consistant"

want $y_e \leq |x_u - x_v|$

~~OK for 0-1 solution.~~

~~But can set~~

$y_e \leq x_u + x_v$

$y_e \leq 2 - (x_u + x_v)$

OK for 0/1 soln

for $x_u = 1/2$ set $y_e = 1 \; \forall e$.

$\exists$ graphs where cut value $\approx \frac{1}{2}|E|$ but the LP has value $|E|$.

$\Rightarrow$ "Integrality gap" of 2.

So the LP is weak!

Can make it stronger by adding in constraints $\sum'_{e \in C} y_e \leq |C| - 1$ for all odd cycles C.

But still can construct graphs where LP $\geq (2-\varepsilon)$ OPTimum Max Cut.

[BTW: how to solve the LP? Ellipsoid! Exercise on optional HW #7]

[In fact No LP can get better than 2, unless has ~~exponential~~ super poly size] [CLRS 12]

OK: sdps to the rescue. Let's change the basis to $x_v \in \{-1, 1\}$. and write

$$\text{OPT as: } \max_{x_i \in \{\pm 1\}} \sum_{ij \in E}' \frac{\|x_i - x_j\|^2}{4} = \sum_{ij \in E}' \frac{\|x_i\|^2 + \|x_j\|^2 - 2\langle x_i, x_j\rangle}{4}$$

unit vectors

$$= \sum_{ij \in E}' \frac{1 - \langle x_i, x_j\rangle}{2}$$

"Relax" this to get

$$\boxed{\max_{\|v_i\|_2 = 1} \sum_{ij \in E}' \frac{1 - \langle v_i, v_j\rangle}{2}}$$

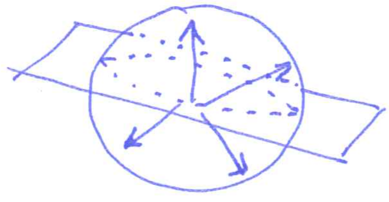
Thm: [GW] this SDP has an integrality gap of at most $\alpha_{GW} = 0.87856\ldots$

[ ] ———— at least ✓

[KKMO] Assuming the unique games conjecture, Max cut does not have any algos better than $0.878\ldots$

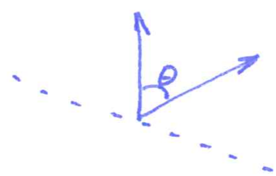
Pf: Claim: Pr[ij cut by a random hyperplane]

$$\geq \alpha \cdot SDP_{ij} \leftarrow \frac{1 - \langle v_i, v_j\rangle}{2}$$



Pf: let v_i, v_j be unit vectors with angle θ

$$\text{then } SDP_{ij} = \frac{1 - \cos\theta}{2}$$



$$\Pr[ij \text{ cut}] = \frac{\theta}{\pi}$$

$$\Rightarrow \alpha = \min_{0 \leq \theta \leq \pi} \frac{\theta}{\pi} \cdot \frac{2}{1 - \cos\theta} = 0.878\ldots \quad \blacksquare$$

Here's another way to get this SDP:-

we would like to solve $\max \sum_{ij} \left(\frac{1 - y_{ij}}{2}\right)$ st $y_{ij} = x_i x_j$, $x_i \in \{+1, -1\}$ $\forall i$

quadratic!

So cannot solve that. However, can throw down "valid" inequalities (linear) that are consistent with $y_{ij} = x_i x_j$:— $\forall c,\ c^T Y c = c^T x x^T c = (c^T x)^2 \geq 0$.

(aka $Y = xx^T$)

$\Rightarrow$ are throwing in the (infinitely many) constraints $c^T Y c \geq 0 \quad \forall c \in \mathbb{R}^n$.

use Ellipsoid to separate for these (need strong separation oracle, etc.).

Good: seen what SDPs are, how they are useful. More powerful than LPs.

Next up: · SDP duality

· a couple more examples. → ~~Max Cut~~ Indep. set & coloring?

→ SOS?

→ Grothendieck?

·

SDP Duality:

Two canonical ways of writing LPs:

$$\begin{array}{l} \min\ c^T x \\ Ax = b \\ x \geq 0 \end{array} \iff \begin{array}{l} \max\ b^T y \\ \text{st } A^T y \leq c \end{array}$$

(duals)

~~max~~ ~~Ax~~ ~~x≥0~~ $\iff$

$A = \begin{pmatrix} - a_1^T - \\ - a_i^T - \\ - a_m^T - \end{pmatrix}$

$$\begin{array}{l} \min\ c^T x \\ \bar{a}_i^T \bar{x} = b_i \\ x \geq 0 \end{array}$$

and

$$\begin{array}{l} \max\ b^T y \\ \text{st } \sum_i y_i \bar{a}_i \leq \bar{c}. \end{array}$$

Here are two canonical SDP formulations.

Let $A_1, A_2 \ldots A_m$ be symmetric matrices. then
also C

$$\boxed{\begin{array}{ll} \min\ C \bullet X & \\ \text{st } A_i \bullet X = b_i & \forall i \in [m] \\ X \succeq 0 & \end{array}} \quad \text{and} \quad \boxed{\begin{array}{l} \max \sum_{i=1}^m b_i y_i \\ \sum A_i y \preceq C. \end{array}}$$

In fact these are duals of each other: [soon.] [~~treat as~~ Lagrangian duals]

an example :-
$$\begin{array}{l} \min\ -A \bullet X \\ \text{st } -I \bullet X = -1 \\ X \succeq 0 \end{array} \Rightarrow \begin{array}{l} \max\ -y \\ \text{st } -Iy \preceq -A \end{array} \text{ or } \begin{array}{l} \min\ y \\ \text{st } yI - A \succeq 0 \end{array}$$

which we saw before.

Why are these duals of each other? See Boyd & Vandenberghe, e.g. but let's at least show weak duality:-

<u>Thm</u>: if $X \succeq 0$ and $\bar{y}$ are feasible ~~duals~~ solutions for the primal and dual pair, then

$$(P) \geq (D)$$

<u>Pf</u>: We need the following fact:-

{ <u>Fact</u>: if A is symmetric then $A \bullet X \geq 0$ for all psd X $\Leftrightarrow$ A is psd. }

<u>Now</u>: Let's write

$$P = \left\{ \begin{array}{r} \min\ C \bullet X \\ A_i \bullet X = b_i \\ X \succeq 0 \end{array} \right\} \text{ and } D = \left\{ \begin{array}{r} \max \sum b_i y_i \\ \text{st } \sum A_i y_i + S = C \\ S \succeq 0 \end{array} \right\}$$

slack.

$$\text{then } (P) = C \bullet X = \left(\sum_{i=1}^{m} A_i y_i + S \right) \bullet X$$
$$= \sum_{i=1}^{m} y_i (A_i \bullet X) + S \bullet X = \sum b_i y_i + \underbrace{\geq 0}_{\text{by Fact.}}$$
$$\geq (D)$$

∎

<u>Aside</u>: dual cones. $\mathcal{C} \subseteq \mathbb{R}^d$ be a cone :- ~~contain~~ closed under positive linear combs.

then $\mathcal{C}^* \subseteq \mathbb{R}^d$ is the polar cone

$$\mathcal{C}^* = \{ x \in \mathbb{R}^d \mid \langle x, y \rangle \geq 0 \quad \forall y \in \mathcal{C} \}$$

<u>Fact</u>: $\mathcal{C}^* = \mathbb{R}^d$, $\mathcal{C}^* = \{0\}$

$\mathcal{C} = R^d_{\geq 0}$ then $\mathcal{C}^* = R^d_{\geq 0}$ } self dual

$\mathcal{C} = $ psd $\quad \mathcal{C}^* = $ psd

$$\begin{array}{l} \min\ c^T x \\ a_i x = b_i \quad \forall i \\ x \in \mathcal{C} \end{array} \Rightarrow \text{dual is} \quad \begin{array}{l} \max\ b^T y \\ \sum_i a_i y_i + s = c \\ s \in \mathcal{C}^* \end{array}$$

and same proof goes through. (Cone duality).

∎

But alas, no strong duality: there are examples where primal and dual are both feasible but $P^* \neq D^*$.

However, under certain "regularity" conditions strong duality holds. E.g. if primal has a strictly feasible sol^n (i.e. $X \succ 0$ positive definite) then dual opt is attained and primal opt = dual opt.

See BV04 and others for details on other sufficient conditions.

Now: a couple more examples —

(1) Coloring 3-colorable graphs. [~~Kar~~ Lovasz, Alon Kahale, Karger Motwani Sudan]

G be 3 colorable. Then can color it by $O(\sqrt{n})$ colors [Wigderson].

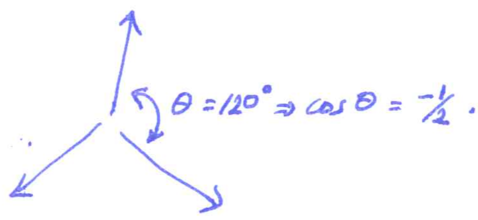
In fact, if we ~~con~~ repeatedly pick vertices of degree $\geq \Delta$ and color it and its nbrhood by 3 colors, we use $3(n/\Delta)$ + # colors to color graph with max degree Δ

$$\leq 3(n/\Delta) + \underset{\text{Brooks}}{(\Delta+1)} \leq O(\sqrt{n}).$$

Do better? Avrim: $O(n^{3/8})$ hard: 3 vs 4.

KMS: write an SDP (feasibility)

$$\begin{cases} \langle v_i \cdot v_j \rangle = -\frac{1}{2} & \forall\, i \sim j \\ \|v_i\|^2 = 1 & \forall\, i \end{cases}$$



Feasible for 3-colorable graphs.

Now: (round 1) Pick K random hyperplanes (thru origin). these divide up ball into 2^k parts.

Look at any vertex. $\Pr[v \text{ shares part with some nbr}] \leq \Delta \cdot \Pr[v, w \text{ in same part, if } u \sim v]$

$$= \Delta \cdot (1/3)^K \leq \frac{1}{2} \text{ if } K = \log_3(2\Delta)$$

$\Rightarrow$ If we throw away all nodes that share a part with a nbr, have at least $n/2$ vertices in expectation. Each part is an IS.

$\Rightarrow$ use 2^K colors, color ½ vertices. Repeat $\log n$ times.

$\Rightarrow E[\#\text{colors in all}] = \underbrace{2^k}_{} \cdot \log n = 2^{\log_3 2\Delta} \cdot \log n = O(\Delta^{\log_3 2} \cdot \log n)$
rounds

$= \tilde{O}(\Delta^{0.631\ldots})$

$\Rightarrow$ By W's trick: $\frac{n}{\Delta} + \Delta^{0.631}$ $\Rightarrow$ gives $n^{0.38\ldots}$ not better than Arrim's $n^{0.375}\ldots$

OK: harder work.

Thm: Find an IS of size $\Omega\left(\frac{n}{\Delta^{1/3}}\right)$ $\Rightarrow$ gives a coloring ~~of size~~ with $\Delta^{1/3}\log n$ colors

$\Rightarrow$ using W's trick $\Rightarrow$ $n^{1/4}$ $\boxed{\text{much better!!}}$

Pf: Pick a random ~~vector~~ gaussian $g \in \mathbb{R}^n$

and take all vertices st $\langle g \cdot v_i \rangle \geq t$ for a threshold t.

- Recall: $\langle g \cdot v_i \rangle \sim N(0, \|v_i\|_2^2) = N(0,1)$.
  g: multivariate gaussian; v_i: unit vector

$\Rightarrow \Pr[\langle g, v_i \rangle \geq t] \leq \mathcal{N}(t)$ ← upper tail of gaussian.

$\Rightarrow E[\text{Cap} \cap V] = n \cdot \mathcal{N}(t)$.

- Claim: conditioned on $i \in$ Cap, a nbr j of i is in the Cap w/p $\lesssim \mathcal{N}(t)^3$.

Pf: angle is $120°$. So if use rotational symmetry, $v_i = e_i$
$v_j = -\frac{1}{2}e_i + \frac{\sqrt{3}}{2}e_j$

$\Rightarrow \Pr[\langle g \cdot v_j \rangle \geq t \mid \langle g \cdot v_i \rangle \geq t]$

$= \Pr\left[-\frac{g_1}{2} + \frac{g_2 \cdot \sqrt{3}}{2} \geq t \mid g_1 \geq t\right]$

$\leq \Pr[g_2 \geq \sqrt{3}t] \approx \mathcal{N}(t)^3$. ∎

$\Rightarrow$ Set t st $\mathcal{N}(t)^3 \approx \frac{1}{2\Delta}$ $\Rightarrow \Pr[i \text{ has no nbrs in cap} \mid i \text{ in cap}] \geq \frac{1}{2}$

$\Rightarrow E[\text{size of IS in Cap}] \geq \frac{1}{2} n \mathcal{N}(t) = \Omega\left(\frac{n}{\Delta^{1/3}}\right)$.

∎

<u>SoS</u> : A polynomial (multivariate, n-vars.) is <u>positive</u> if

$$P(\bar{x}) \geq 0 \quad \forall\, x \in \mathbb{R}^n.$$

: A poly is SoS if $\exists$ polys $h_1, h_2 \dots h_t$ st $P(\bar{x}) = \sum_{i=1}^{t} (h_i(x))^2$.

Clearly P is SOS $\Rightarrow$ P is positive.

- <u>Opposite is not true</u>. (for $n \geq 2$)

(<u>though</u>: P is positive iff P = ratio of SOS polynomials). [Emil Artin].
(sums of squares of rational fns) Hilbert's 17th problem

<u>Q</u>: Can we check if P is an SOS polynomial? Let $P(x_1, x_2 \dots x_n)$ of degree d.
Yes, via SDPs.

$$P = \sum_{\substack{\alpha \in \mathbb{Z}^n_{\geq 0} \\ \|\alpha\|_1 \leq d}} c_\alpha \bar{x}^\alpha .$$

Suppose $P = h(x)^2$

where $h(x) = \sum a_\alpha x^\alpha$

$$\Rightarrow h(x)^2 = \sum_{\alpha,\beta} a_\alpha a_\beta \cdot x^{\alpha+\beta}$$

$$= (aa^T) \bullet X$$

psd matrix. $X_{\alpha,\beta} = x^{\alpha+\beta}$.

<u>So ask</u>: does $\exists Q$ st → of size $\binom{n}{\leq d} \times \binom{n}{\leq d}$

$$Q \bullet Test_\gamma = c_\gamma \quad \forall \gamma$$

$Test_\gamma$ has 1's in locations α, β st $\alpha + \beta = \gamma$

$$Q \succeq 0 .$$

has size $\leq n^{O(d)}$.

$\boxed{\text{Similarly if } P = \sum h_i(x)^2 \Rightarrow P = \underbrace{\sum (a_i a_i^T)}_{\text{etc.}} \bullet X \text{ want this to be } Q.}$

<u>N.b.</u> h_i must have degree $\leq d$ else cannot cancel the high degree terms.

<u>SoS proof systems</u>: