

①

SDPs:

A $n \times n$ symmetric matrix A is psd if :-

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

(b) All evs of A are nonnegative

(c) $A = P D P^T$ where $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$
 $= \sum_i \lambda_i v_i v_i^T$ col of P are eigenvectors v_i of A .

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

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

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

(e) there exist vectors $v_1, \dots, v_n$ such that

$$A_{ij} = \langle v_i, v_j \rangle$$

(f) A can be written as $B B^T$ for some matrix B

(Just take rows of $B = v_i$)

~~SDP is 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 $(x_{ij})_{i, j \in N}$

And we have an extra constraint $X = (x_{ij}) \geq 0$ (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 $\sum_{ij} a_{ij}^l \langle v_i, v_j \rangle \geq b_i \quad \forall l \in 1 \dots m.$

By definition (e) these are equivalent.

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

(assuming bounds on the solution size !! $\leftarrow$ fix)

why is this useful? Two examples.

(2)

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

then SPS we say

$$\begin{array}{ll} \max & A \cdot X \\ \text{st} & X \cdot I = 1 \\ & X \geq 0 \end{array}$$

or

$$\begin{array}{ll} \min & t \\ \text{st} & tI - A \geq 0. \end{array}$$

Have $A \cdot B = \sum_{ij} a_{ij} b_{ij} = \text{tr}(A^T B)$.
is called the ~~Schur~~ ^{Frobenius inner product} product.

this is same as $\{\max \sum_{ij} a_{ij} x_{ij} \mid \sum x_{ii} = 1, x \geq 0\}$

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

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

Now the max is obtained

for general $X = \sum x_i x_i x_i^T$ 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 $y_e \leq x_u \oplus x_v$ are "consistent"

want $y_e \leq |x_u - x_v|$



OK for 0-1 solution.
~~But consistent~~

$y_e \leq x_u + x_v$ OK for 0-1 soln
 $y_e \leq 2 - (x_u + x_v)$ for $x_u = 1/2$
def $y_e = 1 \forall e$.

$\exists$ 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 .

(3)

But still can construct graphs where $LP \geq (2-\epsilon)$ 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 time] [CLRS 14]

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

$$\begin{aligned} \text{OPT as: } \max_{\substack{x_i \in \{-1, 1\} \\ \uparrow \\ \text{unit vectors}}} \sum_{i,j \in E} \frac{\|x_i - x_j\|^2}{4} &= \sum_{i,j \in E} \frac{\|x_i\|^2 + \|x_j\|^2 - 2\langle x_i, x_j \rangle}{4} \\ &= \sum_{i,j \in E} \frac{1 - \langle x_i, x_j \rangle}{2} \end{aligned}$$

"Relax" this to get $\boxed{\max_{\|v_i\|_2=1} \sum_{i,j \in E} \frac{1 - \langle v_i, v_j \rangle}{2}}$

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

[] _____ at least ✓

[KKMO] Assuming the unique games conjecture, Max cut does not have any approx better than 0.878...

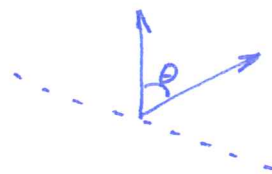
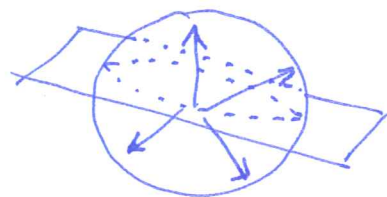
Pf: Claim: $\Pr[\text{cut by a random hyperplane}] \geq \alpha \cdot \text{SDP}_{ij} \leftarrow \frac{1 - \langle v_i, v_j \rangle}{2}$

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

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

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

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



Here's another way to ~~get~~ get this SDP:-

we would like to solve $\max_{i,j} \sum \left(\frac{1 - y_{ij}}{2} \right)$ st $y_{ij} = x_i x_j$
 $x_i \in \{-1, 1\} \forall i$
 $\uparrow$ quadratic!

So cannot solve that. However, can throw down "valid" inequalities (linear) (4)

that are consistent with $y_{ij} = x_i^T x_j$:— $\forall c, c^T Y c = c^T x x^T c = (c^T x)^2 \geq 0$.

(aka $Y = x x^T$)

↑

⇒ 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?
→ SDS?
→ Grothendieck?

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

SDP Duality:

Two canonical ways of writing LPs:

$$\begin{aligned} \min \quad & c^T x \\ \text{st} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

(duals)

$$\begin{aligned} \max \quad & b^T y \\ \text{st} \quad & A^T y \leq c \end{aligned}$$

$$\begin{aligned} \max \quad & c^T x \\ \text{st} \quad & Ax \leq b \\ & x \geq 0 \end{aligned}$$

$$\begin{aligned} \min \quad & c^T x \\ \text{st} \quad & \bar{a}_i^T x = b_i \\ & x \geq 0 \end{aligned}$$

$$\begin{aligned} \max \quad & b^T y \\ \text{st} \quad & \sum_i y_i \bar{a}_i \leq \bar{c} \end{aligned}$$

Here are two canonical SDP formulations.

Let $A_1, A_2, \dots, A_m$ be symmetric matrices. then

$$\begin{aligned} \min \quad & C \bullet X \\ \text{st} \quad & A_i \bullet X = b_i \quad \forall i \in [m] \\ & X \succeq 0 \end{aligned}$$

and

$$\begin{aligned} \max \quad & \sum_{i=1}^m b_i y_i \\ \text{st} \quad & \sum_i A_i y_i \preceq C. \end{aligned}$$

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

an example :-

$$\begin{aligned} \min \quad & -A \bullet X \\ \text{st} \quad & -I \bullet X = -1 \\ & X \succeq 0 \end{aligned}$$

⇒

$$\begin{aligned} \max \quad & -y \\ \text{st} \quad & -I y \preceq -A \end{aligned} \quad \text{or} \quad \begin{aligned} \min \quad & y \\ \text{st} \quad & y I - A \succeq 0 \end{aligned}$$

which we saw before.

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

(5)

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

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

Pf: we need the following fact:-

Fact: if A is symmetric then $A \bullet X \geq 0$ for all $X \succeq 0$ $\Leftrightarrow A$ is psd.

Now: Let's write $P = \left\{ \begin{array}{l} \min C \bullet X \\ A_i \bullet X = b_i \\ X \succeq 0 \end{array} \right\}$ and $D = \left\{ \begin{array}{l} \max \sum b_i y_i \\ \sum A_i y_i + S = C \\ S \succeq 0 \end{array} \right\}$
 $\uparrow$ slack.

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

Aside: dual cones. $\ell \subseteq \mathbb{R}^d$ be a cone :- ~~convex~~ closed under positive linear combs. then ℓ^* is the polar cone

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

Fact: $\ell^* = \mathbb{R}^d$, $\ell^* = \{0\}$

$\ell = \mathbb{R}_{\geq 0}^d$ then $\ell^* = \mathbb{R}_{\geq 0}^d$ } self dual
 $\ell = \text{psd}$ $\ell^* = \text{psd}$

$\min C^T x$ $\Rightarrow$ dual is $\max b^T y$
 $A_i x = b_i \quad \forall i$ $\sum A_i y_i + S = C$
 $x \in \ell$ $S \in \ell^*$

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. Eg. if primal has a strictly feasible solⁿ (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. [Lasserre, Alon Kahale, Karim M. Sidiqi]

Let G be 3-colorable. Then can color it by $O(n)$ colors [Wigderson].

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

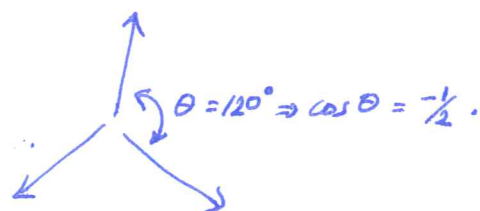
$$\leq 3(n/\Delta) + (\Delta+1) \leq O(n).$$

↑ Brooks

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

KMS: write an SDP (feasibility)

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



Feasible for 3-colorable graphs.

Now: (round 1) Pick K random hyperplanes. These divide up ball into 2^K parts. (thru orig)

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

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

$\Rightarrow$ if we throw away all nodes that share a part with a nbr, have at least $\frac{1}{2}$ vertices in expectation. Each part is an IS.

$\Rightarrow$ use 2^K colors, color $\frac{1}{2}$ vertices. Repeat $\log n$ times.

$$\Rightarrow E[\# \text{ colors in } a_1] = \underbrace{2^k}_{\text{rounds}} \cdot \log n = 2^{\log_2 \Delta} \cdot \log n = O(\Delta^{\log_2 2} \cdot \log n) \\ = \tilde{O}(\Delta^{0.631...})$$

$\Rightarrow$ By W's trick: $\frac{n}{\Delta} + \Delta^{0.631} \Rightarrow$ gives $n^{0.36...}$ not better than Arrim's $n^{0.375}...$

OK: harder work.

Thm: Find an IS of size $\left(\frac{n}{\Delta^{1/3}}\right) \Rightarrow$ gives a coloring ~~with~~ with $\Delta^{1/3} \log n$ colors
 $\Rightarrow$ way W's trick $\Rightarrow n^{1/4}$ much better!!

Pf: Pick a random ~~vector~~ gaussian $g \in \mathbb{R}^n$
 and take all vertices i st $\langle g, v_i \rangle \geq t$ for a threshold t .

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

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

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

• Claim: conditioned on $i \in \text{Cap}$, a nbr j of i is in the Cap w.p. $\leq N(t)^3$.

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

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

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

$$\leq \Pr[g_2 \geq \sqrt{3}t] \leq N(t)^3. \quad \blacksquare$$

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

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

SOS: A polynomial (multivariate) is positive if $P(x) \geq 0 \forall x \in \mathbb{R}^n$. (also called psd)

8

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

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

• Opposite is not true. (for $n \geq 2$)

any univariate poly is sum of 2 squares!

(though: P is positive iff $P = \text{ratio of SOS polynomials}$). [Artin].

(sums of squares of rational fns) Hilbert's 17th problem

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

$P = \sum_{\alpha \in \mathbb{Z}_{\geq 0}^n, \|\alpha\|_1 \leq d} c_\alpha x^\alpha$. Suppose $P = h(x)^2$

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

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

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

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

So ask: does $\exists Q$ st

$Q \bullet \text{Test}_2 = c$ ~~st~~

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

$Q \geq 0$.

Similarly if $P = \sum h_i(x)^2$
 $\Rightarrow P = \sum (a_i a_i^T) \bullet X$
etc. want this to be 0.

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

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

SOS proof systems:

Motzkin's form: $x^4 y^2 + y^4 x^2 + z^6 - 3 x^2 y^2 z^2$ is psd

univariate poly

$\geq 3 \sqrt[3]{x^6 y^6 z^6}$
AM-GM

but not SOS.

Example:

Is $p(x) = 2x_1^4 + 2x_1^3x_2 - x_1^2x_2^2 + 5x_2^4$ positive? sos?

univariate so that two are identical

$$p(x) = \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix}^T \begin{pmatrix} 2 & -\alpha & 1 \\ -\alpha & 5 & 0 \\ 1 & 0 & -1+2\alpha \end{pmatrix} \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix}$$

Q:

Now: want $Q \geq 0$ for some set of α

In this case $\alpha = 3$, $Q \geq 0$. (has eigenvalues 7, 5, 0)

BTW: univariate poly $p(x)$, positive $\Rightarrow$ roots are either 0 or complex conjugate

$$\Rightarrow p(x) = x^c \cdot \prod_{a \in S} (x-a)(x-\bar{a}) = x^c \prod_{a \in S} (x^2 + a^2)$$

c is even

else $p(x) \rightarrow -\infty$

$p(-x) \rightarrow -\infty$



multiply it out: sum of squares.

Grothendieck's Inequality:

①

Given numbers $a_{ij} \in \mathbb{R}$ (both positive or negative) want to figure out how.

$$OPT := \max_{\substack{x_i \in \{\pm 1\} \\ y_j \in \{\pm 1\}}} \sum a_{ij} x_i y_j$$

relates to

$$SDP := \max_{\substack{\|u_i\|_2=1 \\ \|v_j\|_2=1}} \sum a_{ij} \langle u_i, v_j \rangle \quad \leftarrow \text{the SDP relaxation.}$$

Fact: $SDP \geq OPT$ (it is a relaxation)

Thm! $OPT \geq 0.56 \cdot SDP$

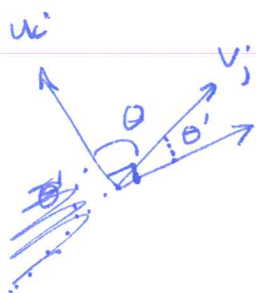
[Krivine, can do better...]

Pf: Suppose we try the simplest idea. These vectors u_i, v_j are on unit ball.
So take the random hyperplane, and set $\hat{x}_i = 1$ if $\langle \bar{g}, u_i \rangle \geq 0$
 $= -1$ if $\langle \bar{g}, u_i \rangle < 0$ etc

$$\Rightarrow \hat{x}_i = \text{sign}(\bar{g} \cdot u_i), \quad \hat{y}_j = \text{sign}(\bar{g} \cdot v_j).$$

$$\text{then } E[\hat{x}_i \hat{y}_j] = -1 \cdot \Pr(i, j \text{ separated}) + 1 \cdot \Pr(i, j \text{ not separated})$$

$$= -1 \cdot \frac{\theta}{\pi} + 1 \cdot (1 - \frac{\theta}{\pi}) = 1 - \frac{2\theta}{\pi} = \frac{2}{\pi} (\frac{\pi}{2} - \theta) = \frac{2}{\pi} \cdot \theta'$$



$$SDP_{ij} = \langle u_i, v_j \rangle = \cos \theta = \sin \theta'$$

$$\Rightarrow E[A_{ij}] = E[\sum a_{ij} \hat{x}_i \hat{y}_j] = SDP$$

Here's the problem: for Max Cut we were happy to say that $E[A_{ij}]$ was bigger than SDP_{ij} . But now we have weights that are negative, so cannot do this!

Krivine's Trick:

Lemma: Given $u_1, \dots, u_n, v_1, \dots, v_m$, can get $\tilde{u}_i, \tilde{v}_j$ such that

$$\frac{\pi}{2} E[\underbrace{\text{sign}(\langle \tilde{u}_i, g \rangle)}_{\hat{x}_i} \cdot \underbrace{\text{sign}(\langle \tilde{v}_j, g \rangle)}_{\hat{y}_j}] \overset{\text{equality!!}}{=} c \cdot \langle u_i, v_j \rangle$$

unit vectors

N.b. LHS = $\frac{\pi}{2} \cdot A[g]_{ij} \Rightarrow E[A[g]] = E[\sum_{ij} a_{ij} \hat{x}_i \hat{y}_j] = \sum_{ij} a_{ij} \cdot \frac{c}{(\pi/2)} \langle u_i, v_j \rangle = \left(\frac{c}{\pi/2}\right) \cdot \text{SDP}.$

Also: LHS, by calculation above = $\sin^{-1}(\langle \tilde{u}_i, \tilde{v}_j \rangle).$

So we want to prove that ~~that~~ $\langle \tilde{u}_i, \tilde{v}_j \rangle = \sin(c \langle u_i, v_j \rangle).$

Pf: $\sin(c \langle u_i, v_j \rangle) = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{c^{2k+1}}{(2k+1)!} (u_i \cdot v_j)^{2k+1}$

existence

$$= \sum_{k=0}^{\infty} (-1)^k \cdot \frac{c^{2k+1}}{(2k+1)!} (u_i^{\otimes 2k+1} \cdot v_j^{\otimes 2k+1})$$

So define ~~$\tilde{u}_i$~~ to be a vector whose k^{th} set of coordinates

are $(-1)^k \cdot \sqrt{\frac{c^{2k+1}}{(2k+1)!}} \cdot u_i^{\otimes 2k+1}$

$\tilde{v}_j$ has $\sqrt{\frac{c^{2k+1}}{(2k+1)!}} v_j^{\otimes 2k+1}$

$(u \cdot v)^2 \text{ etc.}$
 $= u^{\otimes 2} \cdot v^{\otimes 2}$

then by defⁿ $\langle \tilde{u}_i, \tilde{v}_j \rangle = \sin(c \langle u_i, v_j \rangle).$

What are lengths of $\tilde{u}_i$? $\|\tilde{u}_i\|^2 = \sum_{k=0}^{\infty} \frac{c^{2k+1}}{(2k+1)!} \|u_i\|^{(2k+1)^2} = \sinh(c \|u_i\|^2)$
 $= \sinh(c)$

So choose $c = \sinh^{-1}(1) = \ln(1 + \sqrt{2})$
 $= 1.$

similarly $\|\tilde{v}_j\|^2 = \sinh(c) = 1.$

Finally: this was an infinite construction. But this shows $\exists$ vectors that achieve this.

How do you find these vectors? write an SDP for vectors $\tilde{u}_i, \tilde{v}_j$ s.t. (3)

$$\langle \tilde{u}_i, \tilde{v}_j \rangle = \sin(c \cdot \langle u_i, v_j \rangle) \quad \forall i, j$$

↑ inputs!

this SDP gives the solution you want.

Further work: This gave the Grothendieck constant to be ≈ 0.56 . ← Krivine's bound.

[Alon Makarychev², Naor]. Given a general graph, the gap between OPT and SDP

$$\text{OPT} = \max_{\{x_i\}} \sum_{ij \in E} a_{ij} \langle x_i, x_j \rangle \quad \text{vs} \quad \text{SDP} = \max_{\|v_i\|=1} \sum_{ij \in E} a_{ij} \langle v_i, v_j \rangle$$

↑
same set of vars

is at most $O(\log \chi(\bar{G})) \leftarrow K(G)$

↑ the Lovasz theta fn on the complement of G .

For bipartite graphs (which are perfect)

$$\chi(\bar{G}) = \chi(G), \text{ get constant factor } \leftarrow \text{we gave explicit bound.}$$

↑ is true in general

(*) They also show that the Krivine bound is not tight, can beat it by a little bit.

Open Question: What is true bound for K_G ? Current bounds:-

$$\Omega(\log \omega(G)) \leq K_G \leq O(\log \chi(\bar{G}))$$