SOS intro

The basic idea is to generalize SA_d.
Again we introduce variables Y_I, For all subsets $I \subseteq [n]$, $|I| \le t$.

But now we will require that the matrix

$$\begin{array}{c|c} & J \\ \hline I & Y_{I \cup J} \end{array}$$

$|I|, |J| \le \frac{d}{2}$

$|I \cup J| \le d$

is PSD

- Since PSD matrices are convex, can still solve (via SDP) to arbitrary accuracy in time $n^{O(d)}$
- Duality to proof system (SOS) follows by extended Farkas Lemma + relationship between SOS polys and PSD matrices
- PSD condition implies the SA inequalities (so degree-d SOS tightenings are at least as good as degree-d SA_d)

SOS Preliminaries

All vectors over Reals

Vectors u, v are column vectors, so

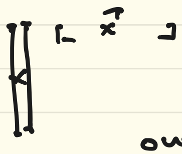
$$u^T v = \langle u, v \rangle = \text{a scalar}$$

$$u v^T = u \otimes v = \text{tensor product}$$

$$(u \otimes v)_{ij} = u_i \cdot v_j$$

Fact $x^T A x = (x x^T) \bullet A$

dot product as vectors



outerproduct

Fact $Tr(AB) = Tr(BA)$

Fact A, B symmetric $\Rightarrow Tr(AB) = A \bullet B$

Tr = sum of diagonal entries

$a_{11}b_{11} + a_{12}b_{12} + \cdots$

(dot product, viewing A, B as vectors)

$$\|v\|_p = \left(\sum_{i=1}^{n} |v_i|^p \right)^{1/p}$$

$p=2$: Euclidean norm

$$\|v\|_\infty = \max_{i \in [n]} |v_i|$$

A, B matrices, entries A_{ij} $B_{i'j'}$

$A \otimes B$, the Kronecker product is a matrix with entries $(A \otimes B)_{ii', jj'} = A_{ij} B_{i'j'}$

$$A^{\otimes k} = A \otimes A \otimes \cdots \otimes A$$

A, B square matrices

$$\text{Det}(AB) = \text{Det}(A) \cdot \text{Det}(B)$$

A diagonal matrix or (upper or lower) triangular. Then

$\text{Det}(A) =$ product of diagonal entries

Thm Let A be an $n \times n$ matrix

Then the sum of eigenvalues

of A = trace(A)

and product of eigenvalues = Det(A)

PSD matrices

Let A be an $n \times n$ symmetric matrix over Reals

A is positive-semidefinite (PSD) if $\forall v \in \mathbb{R}^n$

$$v^T A v \geq 0 \qquad \left(v^T A v = \sum_{i,j} A_{ij} v_i v_j \right)$$

A is positive definite if it is positive semi-definite and nonsingular

Notation $A \succeq 0$ means A is PSD
$A \succ 0$ means A is positive definite

Fact If $M_1, \ldots, M_\ell$ are PSD, so is any nonneg combination

$$\alpha_1 M_1 + \alpha_2 M_2 + \cdots + \alpha_\ell M_\ell \qquad \alpha_i \geq 0$$

(so PSP_n – all n-by-n PSD matrices forms a cone.)

Matrix Decompositions

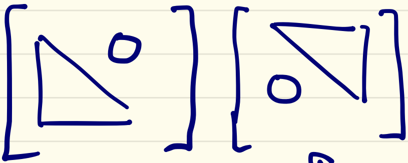
Let $A \succ 0$ (A n-by-n)

Then A has the following equivalent decompositions

1. $A = U^T U$ Cholesky decomp

U is unique

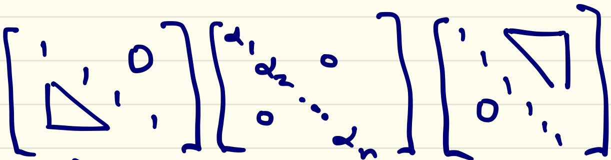
witnesses $A \succ 0$ since $x^T A x = x^T U^T U x = \|Ux\|^2 > 0$



U^T

U is upper triangular, real

2. $A = L D L^T$ LDL decomp



L lower triangular

D diagonal & nonneg

L^T

* LDL holds for any symmetric matrix, but $A \succ 0 \Rightarrow D$ has positive entries

D, L are unique

③ $A = Q \Lambda Q^T$ SVD spectral decomp (eigendecomp) (not unique)

Q is orthogonal matrix

columns of Q are unitary eigenvectors of A
ie. $QQ^T = I$

Λ: diagonal matrix corresponding to eigenvalues, all are positive

A symmetric $\Rightarrow$ eigenvectors with different eigenvalues are orthogonal:

x has eigenvalue $\lambda_x \neq 0$
y " " $\lambda_y \neq 0$

Then $x^T A y = \lambda_x x^T y$

$y^T A x = \lambda_y y^T x = \lambda_y x^T y$

so $\lambda_x x^T y = \lambda_y x^T y \Rightarrow \lambda_x = \lambda_y$

A symmetric + real $\Rightarrow$ eigenvalues are real
say x has eigenvalue λ. Then

$(Ax)^T Ax = x^T A^T Ax = x^T A Ax = \lambda^2 |x|^2$
is real + positive so λ is real

LDL Decomposition

show by induction that $A = L^T D L$

using (for $\alpha > 0$):

$$\begin{bmatrix} \alpha & v^T \\ v & C \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ v/\alpha & I \end{bmatrix} \cdot \begin{bmatrix} \alpha & 0 \\ 0 & \underbrace{C - vv^T/\alpha}_{\text{outer product}} \end{bmatrix} \cdot \begin{bmatrix} 1 & v^T/\alpha \\ 0 & I \end{bmatrix}$$

v^T : length $n-1$ vector

C, I : $(n-1)$-by-$(n-1)$ matrix

Need to show: $B = C - vv^T/\alpha$ is positive def.

Fix $u \in \mathbb{R}^{n-1}$, u nonzero

Fix $x^T = [\underbrace{-u^T v/\alpha}_{\text{scalar}}, u^T]$.

Then $u^T B u = x^T A x > 0$ so pos. definite

From LDL, get cholesky decomp. by:

$$A = L^T D L = (L^T D^{1/2})(D^{1/2} L)$$

since D is diagonal

$$= (LD^{1/2})^T (LD^{1/2})$$

$$= U^T U$$

If A is PSD, all decompositions still work. But then the entries of U + D can be zero.

Also U, L, D not unique and $\mathcal{L}$ can have zero entries

[For LDL If A is PSD, $\alpha \geq 0$.

If $\alpha = 0$ then v is also 0 since otherwise there would be a y st $y^T A y < 0$

Lemma A symmetric $n \times n$ matrix over $\mathbb{R}$

A is PSD iff $A \bullet B \geq 0$ for all PSD matrices B

Pf $\Rightarrow$ A, B PSD $\Rightarrow$ can write A as $Q^T \Lambda Q$ (spectral decomp)

Then

$$A \bullet B = Tr(AB) = Tr(Q^T \Lambda Q B)$$

$$= Tr(\Lambda Q B Q)$$

Tr(M) = sum of diag. entries

$$= Tr(\Lambda B')$$

Λ diagonal, non-neg

$B' = QBQ$ is PSD & nonneg diag. entries

$$= \sum_i \lambda_i b'_{ii}$$

$$\geq 0$$

$\Leftarrow$: say $\exists y^{\geq 0}$ s.t. $y^T A y < 0$

Fix $B = yy^T$. Then $A \bullet B = y^T A y$

B is PSD by construction

& $A \bullet B < 0$.

Lemma $A \succ 0$ Then $A \bullet B > 0$ $\forall B \succeq 0$ unless $B = 0$

A pos def.

Lemma (another characterization)

A n-b-n square, symmetric matrix over $\mathbb{R}$

A is positive definite iff determinant of all upper left submatrices are positive

---

Thm A n-b-n symmetric. TFAE: $\Leftarrow$ $\boxed{\text{summary}}$

1. A is PSD $(A \succeq 0)$
2. All eigenvalues are non-neg
3. Determinant of all upper left submatrice is non-negative
4. The quadratic polynomial $P_A(x) = \sum A_{ij} x_i x_j$ is a SOS's.

   i.e. $\exists$ linear fxns $L_1 \ldots L_n$ s.t.

   $$P_A = \sum_i (L_i)^2$$

5. $A = U^T U$ (cholesky)
6. There are correlated r.v's $X_1 \ldots X_m$ s.t. $\forall i,j$ $\mathbb{E}\, x_i x_j = A_{ij}$ and X_i distributed like $N(0, A_{ii})$

mean 0 variance A_{ij}

SOS - some history

(1) Late 1800's Minkowski / Hilbert asked:

Can every nonneg multivar poly over $\mathbb{R}$ be written as a SOS's $(p_1^2 + p_2^2 + \dots + p_r^2)$

Motzkin [1960's] - NO!

- $1 + x^4y^2 + x^2y^4 - 3x^2y^2$ is nonnegative

$$\frac{1 + x^4y^2 + x^2y^4 - 3x^2y^2}{3} \geq (1 \cdot x^4y^2 \cdot x^2y^4)^{1/3}$$

so nonneg.

Can also show it cannot be written as SOS

(2) Hilbert's 17th Problem (1900 address):

Can every nonneg poly over $\mathbb{R}$ be written as sum-of-squares of Rational functions?

Artin [1927] — Yes!

More on Motzkin's polynomial

$1+x^4y^2+x^2y^4-3x^2y^2$ nonnegativity

Arithmetic Mean - Geometric Mean (AM-GM) Ineq:

AM of $x_1 \ldots x_n$: $\dfrac{x_1+x_2+\cdots+x_n}{n}$

GM of $x_1 \ldots x_n$: $(x_1 \cdot x_2 \cdot \ldots \cdot x_n)^{\frac{1}{n}}$

AM-GM : $\dfrac{x_1+\cdots+x_n}{n} \geq (x_1 \cdot x_2 \cdot \ldots \cdot x_n)^{1/n}$

[2 dimensions: perim. of x_1-by-x_2 rect is $2x_1+2x_2$

perim of square with the same area x_1x_2 is $4\sqrt{x_1x_2}$

says square has smallest perimeter of all rectangles of same area)]

Lots of proofs of AM-GM ineq

Jansen's ineq

Let f be concave function. Then

$$f\left(\frac{x_1+\dots+x_n}{n}\right) \geq \sum_i f(x_i)/n$$

f on mean of $x_1 \dots x_n$ $\geq$ mean of $f(x_1) \dots f(x_n)$

Pf of AM-GM from Jansen's Ineq:

$f = \log$

$$\log\left(\frac{\sum x_i}{n}\right) \geq \sum_i \left(\frac{1}{n}\right) \log x_i$$

$$= \sum_i \log x_i^{1/n}$$

$$= \log\left(\prod x_i^{1/n}\right)$$

Back to Nonneg of Motzkin's Poly:

$$1 + x^4y^2 + y^4x^2 - 3x^2y^2$$

apply AM-GM to $1, x^4y^2, y^4x^2$

$$-\frac{3x^2y^2}{3} + \frac{1+x^4y^2+y^4x^2}{3} \geq -x^2y^2 + \left(\frac{x^6y^6}{3}\right)^{1/3}$$

$$= \left(3^{1/3}-1\right)x^2y^2 \geq 0$$

Motzkin Poly cont'd

To see $M(x,y) = x^2y^4 + x^4y^2 + 1 - 3x^2y^2$

is Non-Neg, here's a better way.

(x^2+y^2+1) is positive

and we have

$$(x^2+y^2+1)M(x,y) =$$

$$(x^2y - y)^2 + (xy^2 - x)^2 +$$
$$(x^2y^2 - 1)^2 + \tfrac{1}{4}(xy^3 - x^3y)$$
$$+ \tfrac{3}{4}(xy^3 + x^3y - 2xy)^2$$

so $M(x,y) = \dfrac{P_1^2 + P_2^2 + P_3^2 + P_4^2 + P_5^2}{Q_1^2 + Q_2^2 + 1}$

ie. it is a sum-of-squares (always nonneg) divided by something positive

so $M(x,y)$ is nonneg!

So this poly is nonneg.

and can be written as $\dfrac{SOS}{SOS}$

But cannot be written as SOS!

Luckily

In some special cases, SOS

is equivalent to non-Negativity:

1. Univariate case $n=1$
2. Degree 2 $d=2$
3. $n=2$ and $d=4$

} Hilbert 1888

* 4. Functions over hypercube

$f: \{0,1\}^n \Rightarrow \mathbb{R}$

Univariate case

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_{2d} x^{2d}$$

$$= \begin{bmatrix} 1 & x & x^2 & \dots & x^d \end{bmatrix} \begin{bmatrix} q_{00} & q_{01} & \cdots & q_{0d} \\ q_{01} & & & q_{1d} \\ & & \vdots & \\ q_{d0} & & \cdot & q_{dd} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \\ \vdots \\ x^d \end{bmatrix}$$

$$= \sum_{i=0}^{d} \Big(\sum_{j+k=i} q_{jk} \Big) x^i$$

In univ case, SOS condition is **equivalent** to nonnegativity!

SOS over Boolean Hypercube

Lemma Every nonnegative function $f : \{0,1\}^n \Rightarrow \mathbb{R}$ has a degree $2n$ SOS certificate

pf Let $g : \{0,1\}^n \Rightarrow \mathbb{R}$ be the (unique) multilinear function that agrees with $\sqrt{f}$ on the hypercube. Then $f = g^2$ over $\{0,1\}^n$ and has degree $\leq 2n$.

We'll soon see a better, more constructive algorithm to find an SOS certificate (but still exp time in worst case)

SOS certificates via PSD matrices

Lemma

Let $f(x_1 \ldots x_n) \in \mathbb{R}[x_1 \ldots x_n]_{2d}$ be a degree $\underline{\underline{2d}}$ poly

Let z be a vector of all degree $\leq d$ monomials.

Then f is SOS iff $\exists Q$ $f(x) = z^T Q z$, $Q \succeq 0$

$$= \sum_{ij} z_i z_j Q_{ij}$$

Proof

$\boxed{\Leftarrow}$ Factorize $Q = U^T U$.

Then $f(x) = z^T U^T U z = \|Uz\|^2$

$$= \sum_i (Uz)_i^2$$

The terms in decomposition are given by $g_i = (Uz)_i$

The # of squares = rank of Q

i.e. $(Uz)^T = y_1 \cdots y_r$ r poly's

$$(Uz)^T (Uz) = [y_1 \ldots y_r] \begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix} = y_1^2 + \ldots + y_r^2$$

So $f = (Uz)^T \cdot (Uz) = y_1^2 + \dots + y_r^2$

the i^{th} row of U_i are coeff's of y_i (since highest terms in y_i cannot cancel, degrees of y_i's are at most d)

$\boxed{\Rightarrow}$ Say $f = y_1^2 + y_2^2 + \dots + y_r^2$

create matrix Q s.t.

$$Q_{ij} = \hat{y}_i \cdot \hat{y}_j$$

$\langle$ coeff's of $y_i \cdot$ coeffs of $y_j \rangle$

[y_i's over $\vec{z}$. Can think of $|\vec{z}| = \mathbb{Z}^n$, where nonzero z_i's correspond to degree $\le d$ monomials in x vars]

<u>But</u> $f : \mathbb{R}[x_1, \dots x_n]_{2d} \Rightarrow \mathbb{R}$ nonneg is <u>NOT</u> equivalent to f being SOS

__Example__ Let $p = 2x^4 + 2x^3y - x^2y^2 + 5y^4$

If p a sos, since p is homogeneous, the poly's in the sos's representation have degree 2

[so terms are x^2, y^2, xy]

$$\begin{array}{ccc} X_1 & X_2 & X_3 \\ \end{array}$$

$$\begin{pmatrix} x^2 & y^2 & xy \end{pmatrix} \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} \begin{pmatrix} x^2 \\ y^2 \\ xy \end{pmatrix} = \sum_{i,j} q_{ij} X_i X_j$$

$$= q_{11} X_1 X_1 + q_{22} X_2 X_2 + q_{33} X_3 X_3 +$$

$$2q_{12} X_1 X_2 + 2q_{23} X_2 X_3 + 2q_{13} X_1 X_3$$

$$= q_{11} x^4 + q_{22} y^4 + q_{33} x^2 y^2 + 2q_{12} x^2 y^2$$

$$+ 2q_{23} x y^3 + 2q_{13} x^3 y$$

$$= q_{11} x^4 + q_{22} y^4 + (q_{33} + 2q_{12}) x^2 y^2 + 2q_{23} x y^3$$

$$+ 2q_{13} x^3 y$$

The existence of a PSD Q is equivalent to feasibility of an SDP in standard primal form:

$$Q \succeq 0 \quad \text{s.t.} \quad \begin{aligned} q_{11} &= 2 \\ q_{22} &= 5 \\ 2q_{13} &= 2 \\ 2q_{23} &= 0 \\ q_{33} + 2q_{12} &= -1 \end{aligned}$$

Matching coefficients of p

$$p = \underbrace{2}_{q_{11}}x^4 + \underbrace{2}_{2q_{13}}x^3y \underbrace{-1}_{}x^2y^2 + \underbrace{5}_{q_{22}}y^4 + \underbrace{0}_{2q_{23}}xy^3$$

gives q_{11} $2q_{13}$ q_{22} $2q_{23}$

$$-1 = q_{33} + 2q_{12}$$

solving SDP gives

$$Q = \begin{pmatrix} 2 & -3 & 1 \\ -3 & 5 & 0 \\ 1 & 0 & 5 \end{pmatrix} \leftarrow \text{Rank } 2 \text{ so } p \text{ is the sum of } 2 \text{ squares}$$

Cholesky factorization:

$$Q = \frac{1}{\sqrt{2}}\underbrace{\begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}}_{U}^{T} \frac{1}{\sqrt{2}}\begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\therefore U\begin{pmatrix} x^2 \\ y^2 \\ xy \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}\begin{matrix} x^2 \\ y^2 \\ xy \end{matrix} = \begin{bmatrix} \frac{1}{\sqrt{2}}(2x^2 - 3y^2 + xy) \\ \frac{1}{\sqrt{2}}(y^2 + 3xy) \end{bmatrix}\begin{matrix} \leftarrow y_1 \\ \leftarrow y_2 \end{matrix}$$

so $p = y_1^2 + y_2^2$

Now let's consider $f: \{0,1\}^n \to \mathbb{R}$. (over hypercube)
Here, f non-neg is equivalent to f having an SOS certificate

But — now SOS certificate could have much larger degree than f.

Let $f: \{0,1\}^n \Rightarrow \mathbb{R}$.
Let $\vec{z}$ be vector of all multilinear monomials over $x_1 \ldots x_n$ ($|\vec{z}| = 2^n$)

Thm $f: \{0,1\}^n \Rightarrow \mathbb{R}$ is non-neg. iff f has an SOS certificate iff $\exists Q$ s.t. $z^T Q z = f(x)$, $Q \succeq 0$

Pf ① We already saw f non neg iff f has an SOS certificate

② Same as previous pf but now can't restrict to monomials of degree $\leq \deg(f)$

Example

Here's an example of a function $f:\{0,1\}^n \Rightarrow \mathbb{Z}$
that is NON-Neg, has low degree (3)
but requires SOS certificates of degree $\Omega(n)$

Start with random UNSAT 3-CNF
[random mod 2 eqns,
or Tseitin on degree-3 expander]

$f = C_1 \wedge C_2 \wedge \dots \wedge C_m$ over $x_1 \dots x_n$

Convert each clause $C \to P_c$:

$C = (x_1 \vee \bar{x}_2 \vee x_3) \to P_c = (1-x_1)(x_2)(1-x_3)$

$\forall \alpha$ falsifying $C \to P_c(\alpha) = 1$

$\forall \alpha$ satisfying $C \to P_c(\alpha) = 0$

Let $P_f \stackrel{d}{=} \sum_{c \in F} P_c - 1$ ← degree 3 and Non-Negative if f is UNSAT

Thm (Grigoriev) P_f requires SOS degree $\Omega(n)$ [We will do this soon]

To find SOS certificate, find a feasible solution to:

$$P_f = \sum Q_{ij} z_i z_j$$

$$Q \succeq 0$$

this is a bunch of equations equating coeff's of P_f to poly's (as in example)

variables are $q_{i,j}$

where i,j correspond to $I, J \subseteq [n]$

So already we can see that sum-of-squares forms the basis of a proof system (for nonnegativity of a single function over hypercube.)

Let $f: \{0,1\}^n \Rightarrow \mathbb{Z}$

Then $f < 0 \equiv -f-1 \geq 0$ is unsolvable over $\{0,1\}^n$ iff $\exists$ sos g s.t. $g = f$.

over $\mathbb{R}$ need to consider some $\varepsilon < 0$

if f has poly bit length, $\varepsilon \approx \frac{1}{2^n}$

[$\Leftarrow$] Let g be a sos, $g = f$. Then $g \geq 0$, $-f-1 \geq 0$

$\Rightarrow g - f - 1 = -1 \geq 0$ #]

[$\Rightarrow$] If $f < 0$ unsolvable over $\{0,1\}^n$ then $f \geq 0$, so by completeness $\exists$ sos g s.t. $g = f$

[degree of g is $\leq 2n$]

More general proof System [more than one poly]

[Krivine, Stengle] 60-70's

Positivstellensatz (generalizes Nullstellensatz)

Take a set of poly inequalities asserting all are nonnegative

$P_1 \geq 0, \ldots, P_m \geq 0, 1 \geq 0$

[opposite as above - flip P to up]

If unsolvable (on $\mathbb{R}$) then they imply -1

$$\left[\sum P_i \cdot Q_i = -1, \text{ where } Q_i\text{'s are SOS's} \right]$$

This is an SOS proof