Lecture 5: SOS Proofs and the Motzkin Polynomial

Lecture Outline

- Part I: SOS proofs and examples
- Part II: Motzkin Polynomial

Part I: SOS proofs and examples

SOS proofs

- Fundamental question: What can we say about the pseudo-expectation values SOS gives us?
- In other words, which statements that are true for any expectation of an actual distribution of solutions must also be true for pseudo-expectation values?

Non-negativity of Squares

- Trivial but extremely useful: If f is a sum of squares i.e. $f = \sum_j g_j^2$ then $\tilde{E}[f] \geq 0$
- Example: If $f = x^2 - 4x + 5$ then $\tilde{E}[f] \geq 0$ as $f = (x-2)^2 + 1$. In fact, $\tilde{E}[f] \geq 1$

Single Variable Polynomials

- Theorem: For a single-variable polynomial $\mathrm{p}(\mathrm{x})$, $p(x)$ is non-negative $\Leftrightarrow$ $p(x)$ is a sum of squares.
- Proof: By induction on the degree d
- Base case $d = 0$ is trivial
- If $d > 0$, let $c \geq 0$ be the minimal value of $p(x)$ and let a be a zero of $p(x) - c$. Since $p(x) - c$ is non-negative, it has a zero of order $2k$ at a for some integer $k \geq 1$ (the order must be even).
- Write $p = (x-a)^{2k}p' + c$ where $p' = \frac{p-c}{(x-a)^{2k}}$ is non-negative and thus a sum of squares.

Degree 2 Polynomials

- Given a degree 2 polynomial f, we can write $f(x_1, x_2, \ldots, x_n) = \sum_{i,j} c_{ij} x_i x_j$ where $c_{ji} = c_{ij}$ for all i and j.
- Taking M to be the coefficient matrix where $M_{ij} = c_{ij}$, we can write $M = \sum_i \lambda_i v_i v_i^T$ where the $\{v_i\}$ are orthonormal. Now
  1. $f(x) = x^T M x$.
  2. $f(x) = \sum_i \lambda_i \, x^T v_i v_i^T x = \sum_i \lambda_i \left(\sum_{j=1}^{n} v_{ij} x_j\right)^2$

Degree 2 Polynomials

- We have that
  1. $M = \sum_i \lambda_i v_i v_i^T$ where the $\{v_i\}$ are orthonormal.
  2. $f(x) = x^T M x$
  3. $f = \sum_i \lambda_i \left(\sum_{j=1}^{n} v_{ij} x_j\right)^2$
- If $M \succcurlyeq 0$ then $\forall i, \lambda_i \geq 0$ so f is a sum of squares
- If M is not PSD then $\lambda_i < 0$ for some i. Taking $x = v_i, f(x) = v_i^T M v_i < 0$ so f is not non-negative.
- Thus if $\deg(f) = 2$, f is non-negative $\Leftrightarrow f$ is SOS

Cauchy Schwarz Inequality

- Cauchy-Schwarz inequality:

$$\left(\sum_i f_i g_i\right)^2 \leq \left(\sum_i f_i^2\right)\left(\sum_i g_i^2\right)$$

- Extremely useful
- Proof: Consider f and g as vectors. Cauchy-Schwarz is equivalent to $(f \cdot g)^2 \leq \|f\|^2 \|g\|^2$
- This is true as $(f \cdot g)^2 = \|f\|^2 \|g\|^2 \cos^2 \Theta$ where Θ is the angle between f and g.
- How about an SOS proof?

Cauchy Schwarz: SOS Proof

- Cauchy-Schwarz: $\left(\sum_i f_i g_i\right)^2 \leq \left(\sum_i f_i^2\right)\left(\sum_i g_i^2\right)$
- Building block: For all i and j,

$$\left(f_i g_j - f_j g_i\right)^2 = f_i^2 g_j^2 + f_j^2 g_i^2 - 2 f_i g_i f_j g_j \geq 0$$

- Note that:
  1. $\sum_{i<j}(f_i^2 g_j^2 + f_j^2 g_i^2) = \left(\sum_i f_i^2\right)\left(\sum_i g_i^2\right) - \sum_i f_i^2 g_i^2$
  2. $2\sum_{i<j}(f_i g_i f_j g_j) = \left(\sum_i f_i g_i\right)^2 - \sum_i f_i^2 g_i^2$
- Final proof: $\sum_{i,j:i<j}\left(f_i g_j - f_j g_i\right)^2 =$
$\left(\sum_i f_i^2\right)\left(\sum_i g_i^2\right) - \left(\sum_i f_i g_i\right)^2 \geq 0$

SOS Proofs With Constraints

- What if we also have constraints $s_1(x_1, \ldots, x_n) = 0, s_2(x_1, \ldots, x_n) = 0$, etc.?
- An SOS proof that $h \geq c$ now takes the form $h = c + \sum_i f_i s_i + \sum_j g_j^2$
- Example: If $x^2 = 1$ then $\mathrm{x} \geq -1$. Proof:

$$x + 1 = \frac{x^2}{2} + x + \frac{1}{2} = \frac{1}{2}(x + 1)^2 \geq 0$$

Combining Proofs

- If there is an SOS proof of degree d_1 that $f \geq 0$ and an SOS proof of degree d_2 that $g \geq 0$ then:
  1. There is an SOS proof of degree $max\{d_1, d_2\}$ that $f + g \geq 0$
  2. There is an SOS proof of degree $d_1 + d_2$ that $fg \geq 0$

Products of Pseudo-expectation Values

- What if our statements involve products of pseudo-expectation values?
- Example: We showed that

$$\tilde{E}\left[\left(\sum_i f_i g_i\right)^2\right] \leq \tilde{E}\left[\left(\sum_i f_i^2\right)\left(\sum_i g_i^2\right)\right]$$

What if we instead want to show that

$$\left(\tilde{E}\left[\sum_i f_i g_i\right]\right)^2 \leq \tilde{E}\left[\sum_i f_i^2\right]\tilde{E}\left[\sum_i g_i^2\right]?$$

- Requires modified proof, see problem set
- Can often prove such statements by using $\tilde{E}$ values as constants in the proof.

Example: Variance

- For any random variable x, $E\left[x^2\right] \geq (E[x])^2$
- Also true for pseudo-expectation values, i.e. for any polynomial f, $\tilde{E}\left[f^2\right] \geq \left(\tilde{E}[f]\right)^2$
- Proof: Given $\tilde{E}$, let $c = \tilde{E}[f]$ and observe that

$$\tilde{E}\left[(f-c)^2\right] = \tilde{E}\left[f^2\right] - 2c\tilde{E}[f] + \mathrm{c}^2$$

$$= \tilde{E}\left[f^2\right] - \left(\tilde{E}[f]\right)^2 \geq 0$$

In-class exercises

1. Prove that $\tilde{E}\left[x^4 - 4x + 3\right] \geq 0$
2. Prove that $\tilde{E}\left[x^2 + 2y^2 + 6z^2 + 2xy + 2xz + 6yz\right] \geq 0$
3. Prove that if $x^2 + y^2 = 1$ then $x + y \leq \sqrt{2}$
4. Prove that if $\tilde{E}\left[x^2\right] = 0$ then for any function f of degree at most $\frac{d}{2}$, $\tilde{E}[xf] = 0$.

In-class exercise answers

1. Prove that $\tilde{E}\left[x^4 - 4x + 3\right] \geq 0$

Answer: $x^4 - 4x + 3 = (x-1)^2\left(x^2 + 2x + 3\right) = (x-1)^2((x+1)^2 + 2)$

In-class exercise answers

2. Prove that
$\tilde{E}\left[x^2 + 2y^2 + 6z^2 + 2xy + 2xz + 6yz\right] \geq 0$
Answer: The coefficient matrix for this polynomial is $\mathrm{M} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{bmatrix}$
One non-orthonormal factorization is $M = v_1 v_1^T + v_2 v_2^T + v_3 v_3^T$ where $v_1^T = [1 \quad 1 \quad 1]$, $v_2^T = [0 \quad 1 \quad 2]$, $v_3^T = [0 \quad 0 \quad 1]$,

In-class exercise answers

This gives us that

$$x^2 + 2y^2 + 6z^2 + 2xy + 2xz + 6yz$$
$$= (x + y + z)^2 + (y + 2z)^2 + z^2$$

In-class exercise answers

3. Prove that if we have the constraint $x^2 + y^2 = 1$ then $\tilde{E}[x + y] \leq \sqrt{2}$

Answer: $\sqrt{2} - x - y = \frac{x^2+y^2}{\sqrt{2}} - x - y + \frac{1}{\sqrt{2}} =$

$$\frac{(x-y)^2}{2\sqrt{2}} + \frac{(x+y)^2}{2\sqrt{2}} - x - y + \frac{1}{\sqrt{2}} =$$

$$\frac{(x-y)^2}{2\sqrt{2}} + \frac{1}{2\sqrt{2}}\left(x + y - \sqrt{2}\right)^2 \geq 0$$

In-class exercise answers

4. Prove that if $\tilde{E}\left[x^2\right] = 0$ then for any function f of degree at most $\frac{d}{2} - 1$, $\tilde{E}[xf] = 0$.

Answer: Observe that for any constant C,

$\tilde{E}\left[(f - Cx)^2\right] = \tilde{E}\left[f^2\right] - 2C\tilde{E}[xf] + \tilde{E}\left[x^2\right] = \tilde{E}\left[f^2\right] - 2C\tilde{E}[xf] \geq 0$

The only way this can be true for all C us if $\tilde{E}[xf] = 0$.

Part II: Motzkin Polynomial

Non-negative vs. SOS polynomials

- Unfortunately, not all non-negative polynomials are SOS.
- Are equivalent in the special cases where $n = 1$ (single-variable polynomials), $d = 2$ (quadratic polynomials), or $n = 2, d = 4$ (quartic polynomials with two variables)
- Hilbert [Hil1888]: In all other cases, there are non-negative polynomials which are not sums of squares of polynomials.
- Motzkin [Mot67] found the first explicit example.

Motzkin Polynomial

- Motzkin Polynomial:
$$p(x,y) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$$
- Question 1: Why is it non-negative?
- Question 2: How can we show it is not a sum of squares of polynomials?

AM-GM inequality

- Arithmetic mean/Geometric mean Inequality:

$\sqrt[n]{\prod_{i=1}^{n} x_i} \leq \frac{1}{n}\sum_{i=1}^{n} x_i$ if $\forall i, x_i \geq 0$ with equality if and only if all of the x_i are equal.

- Proof: Minimize $\frac{1}{n}\sum_{i=1}^{n} x_i - \sqrt[n]{\prod_{i=1}^{n} x_i}$
- Derivative with respect to x_j is $\frac{1}{n}\left(1 - \frac{\sqrt[n]{\prod_{i\neq j} x_i}}{\sqrt[n]{x_j^{n-1}}}\right)$
- Setting this to 0 for all j, $\forall j, x_j = \sqrt[n]{\prod_{i=1}^{n} x_i}$

Motzkin Polynomial Non-negativity

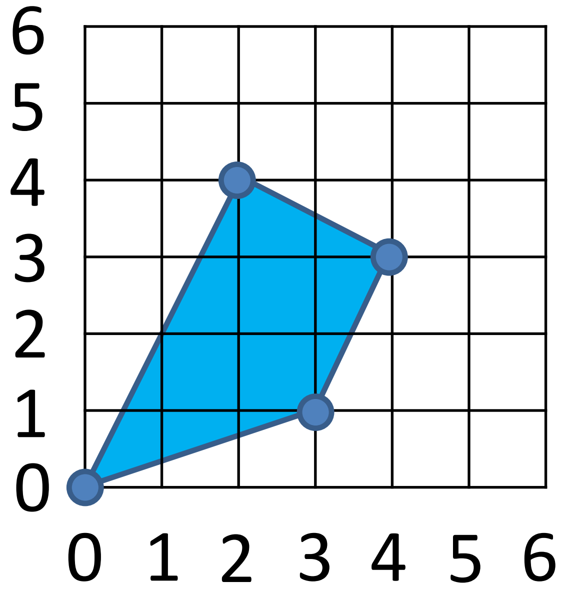
- Motzkin Polynomial:
$$p(x,y) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$$
- Applying AM-GM with x^4y^2, y^2x^4, 1,
$$x^2y^2 = \sqrt[3]{(x^4y^2)\cdot(y^2x^4)\cdot 1} \leq \frac{x^4y^2 + y^2x^4 + 1}{3}$$
- Multiplying this by 3, $p(x,y) \geq 0$

Newton Polytope

- Given a polynomial, assign a point to each monomial based on the degree of each variable. Examples:
  1. x^2y is assigned the point $(2,1)$
  2. y^5 is assigned the point $(0,5)$
  3. xy^2z^3 is assigned the point $(1,2,3)$
- The Newton polytope of a polynomial is the convex hull of the points assigned to each monomial.

Newton Polytope Example

- Example: Newton Polytope for the polynomial $p(x) = 3x^2y^4 - x^4y^3 - 2x^3y + 4$
- Note that the coefficients in front of the monomials don't change the polytope.



Newton Polytope of a Sum of Squares

- Let f be a sum of squares, i.e. $f = \sum_j g_j^2$
- Claim: The Newton polytope of f is $2X$ where X is the convex hull of all the points corresponding to some monomial in some g_j
- Proposition: If p, q are monomials with corresponding points a, b then pq corresponds to the point $a + b$
- One direction: Let X_j be the Newton polytope of g_j. The Newton polytope of $g_j^2 \subseteq 2X_j \subseteq 2X$. Thus, the Newton polytope of $f \subseteq 2X$.

Newton Polytope of a Sum of Squares

- Other direction: If p, q, r are monomials where $pr = q^2$ and a, b, c are the corresponding points, $a + c = 2b$
- Corollary: If b is a vertex of X corresponding to a monomial q then if
  1. p, r are monomials appearing in some g_j (and thus their corresponding points a, c are in X)
  2. $pr = q^2$

  then $p = r = q$.

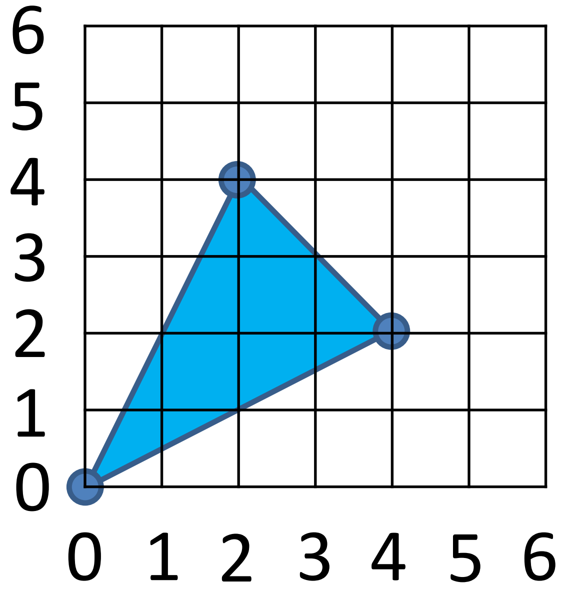
Newton Polytope of a Sum of Squares

- Corollary: If b is a vertex of X corresponding to a monomial q then q^2 appears with positive coefficient in $f = \sum_j g_j^2$.
- This implies that $2X \subseteq$ the Newton polytope of f
- Putting everthing together, the Newton polytope of f is $2X$.

Motzkin Polynomial Newton Polytope

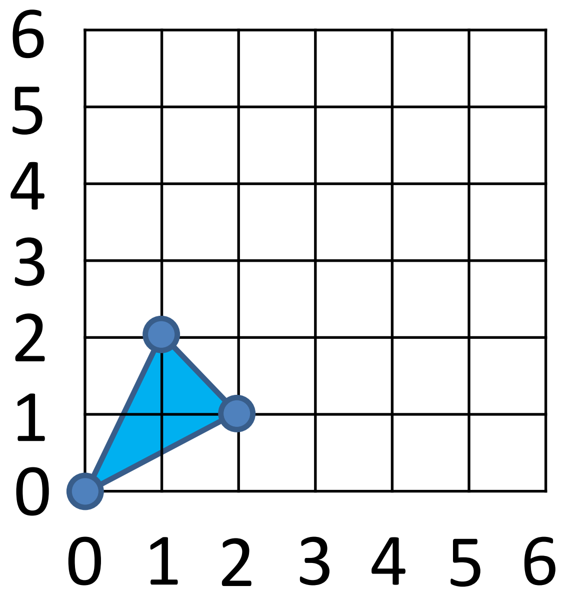
- Motzkin polynomial:

$$p(x) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$$



Motzkin Polynomial Newton Polytope

- If $p(x)$ were a sum of squares of polynomials, their corresponding points would have to be inside the following polytope.



Motzkin is not a Sum of Squares

- If $p(x) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$ were a sum of squares of polynomials, it would have to be a sum of terms of the form
$$\left(ax^2y + bxy^2 + cxy + d\right)^2$$
- However, no such term has a negative coefficient of x^2y^2. Contradiction.

Showing Polynomials are not SOS

- Is there a more general way to show a polynomial is not a sum of squares?
- Observation: By definition, if $f = \sum_j g_j^2$ then for any valid pseudo-expectation values,

$$\tilde{E}[f] = \sum_j \tilde{E}\left[g_j^2\right] \geq 0$$

- Thus, if we can find pseudo-expectation values such that $\tilde{E}[f] < 0$, then f is not a sum of squares of polynomials.

Motzkin is a Rational Function of Sums of Squares

- $p(x) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$
- $(x^2 + y^2 + 1)p(x) = x^6y^2 + 2y^4x^4 + x^2y^6 - 2x^4y^2 - 2x^2y^4 - 3x^2y^2 + x^2 + y^2 + 1$
- This is a sum of squares. The components are:

1. $2\left(\frac{1}{2}x^3y + \frac{1}{2}xy^3 - xy\right)^2 = \frac{1}{2}\left(x^6y^2 + 2y^4x^4 + x^2y^6\right) - 2x^4y^2 - 2x^2y^4 + 2x^2y^2$
2. $\left(x^2y - y\right)^2 = x^4y^2 - 2x^2y^2 + y^2$
3. $\left(xy^2 - x\right)^2 = x^2y^4 - 2x^2y^2 + x^2$
4. $\frac{1}{2}\left(x^3y - xy\right)^2 = \frac{1}{2}x^6y^2 - x^4y^2 + \frac{1}{2}x^2y^2$
5. $\frac{1}{2}\left(xy^3 - xy\right)^2 = \frac{1}{2}x^2y^6 - x^2y^4 + \frac{1}{2}x^2y^2$
6. $\left(x^2y^2 - 1\right)^2 = x^4y^4 - 2x^2y^2 + 1$

Can SOS use Rational Functions?

- $p(x) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$
- $p(x) = \frac{\sum_j g_j^2}{x^2+y^2+1} \geq 0$
- Can the SOS hierarchy use such reasoning?
- Yes and no… (see problem set)

References

- [Hil1888] D. Hilbert. Uber die darstellung definiter formen als summe von formen-quadraten. Annals of Mathematics 32:342–350, 1888.
- [Mot67] T. Motzkin. The arithmetic-geometric inequality. In Proc. Symposium on Inequalities p. 205–224, 1967.