Lecture 20: Proofs to algorithms and the sum-of-squares SDP.

How can we turn a proof of existence into an efficient algorithm for finding the thing?

In the worst case, this is impossible.

- some proofs are non-constructive
  i.e. proof by contradiction
- other proofs are too complicated, and it's unclear how to extract an efficient (i.e. polynomial time) algorithm from it.

This class: a way to extract algorithms from structured proofs
"sum-of-squares" proofs.

will generalize several things we've seen so far:

- Goemans-Williamson max-cut
- Cheeger's inequality.

In particular, how to prove non-negativity of polynomials.

Q: how to show that

$$p(x) = x^4 + 12x^3 + 55x^2 - 94x \geq -130 \quad \text{for all } x?$$

One way: write

$$p(x) - 130 = \underbrace{(x-7)^2 + (x+3)^4}_{\text{sum of squares} \geq 0.}$$

$$\Rightarrow p(x) - 130 \geq 0 \Rightarrow p(x) \geq -130 \quad \forall x.$$

This is a way to certify non-negativity of a polynomial!

In 1-dimension, not such an interesting task perhaps.

But much more interesting in high dimensions.

Let $p: \mathbb{R}^n \to \mathbb{R}$ be a polynomial in n variables.

We say it has degree d if every monomial in p has total degree $\leq d$.

e.g. $x_1^2 x_2 x_3^5$ has degree 8

$x_1 x_2$ has degree 2

Def: Let $p: \mathbb{R}^n \to \mathbb{R}$ be a degree d polynomial. We say $p(x) \geq 0$ has a sum-of-squares (SOS) proof of degree k if there exist polynomials $s_1, \dots, s_t$ of degree $\leq k/2$ s.t. $p(x) = \sum_{i=1}^{t} s_i(x)^2$.

We say $p(x) \geq q(x)$ has an SOS proof of deg k if $p(x) - q(x) \geq 0$ has an SOS proof of deg k.

Fact: When $n = 1$ (i.e. in 1 dimension), $p(x) \geq 0 \Leftrightarrow$ there is an SOS proof of this, but this is not true for higher dimensions.

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

"Motzkin polynomial".

So SOS is not a "complete" proof system.

i.e. it cannot prove all true statements.

examples of multivariate sos proofs:

ex 1 $\langle a,b\rangle \le \frac{1}{2}\|a\|_2^2 + \frac{1}{2}\|b\|_2^2$ is true in deg 2 sos.

$$\frac{1}{2}\|a\|_2^2 + \frac{1}{2}\|b\|_2^2 - \langle a,b\rangle = \frac{1}{2}\|a-b\|_2^2 = \underbrace{\frac{1}{2}\sum_{i=1}^{n}(a_i-b_i)^2}_{\text{sos!}}$$

ex 2 $\langle a,b\rangle^2 \le \|a\|_2^2\,\|b\|_2^2$ ← Cauchy Schwarz!

$$\|a\|_2^2\cdot\|b\|_2^2 - \langle a,b\rangle^2 = \frac{1}{2}\sum_{ij}(a_ib_j - a_jb_i)^2$$

many "natural" proofs can be captured in sos, e.g.

Cauchy-schwarz, certain AM-GM, Hölders, etc...

SoS can also capture implications

Def: Let $f_1,\dots,f_M, g_1,\dots,g_K : \mathbb{R}^n \to \mathbb{R}$ be polynomials. We say that the axioms $\{f_i = 0\}_{i=1}^{M}$ and $\{g_i \ge 0\}_{i=1}^{K}$ sos-imply that $p \ge 0$ in deg k if there exist polynomials $s_1,\dots,s_t$, $a_1,\dots,a_M$, $b_1,\dots,b_K$ s.t. $\deg(s_i) \le k/2$, $\deg(a_i f_i) \le k$, $\deg(b_i^2 g_i) \le k$ $\forall i$

so that

$$(*)\quad p(x) = \sum_{i=1}^{t} s_i(x)^2 + \sum_{j=1}^{M} a_i(x) f_i(x) + \sum_{k=1}^{K} b_i^2(x)\, g_i(x).$$

Intuition: suppose x is a point satisfying $f_i(x) = 0$ $\forall i$, $g_i(x) \ge 0$ $\forall i$.

Then (*) implies that $p(x) \ge 0$!

Note: k can be much larger than degree of p!

For conciseness, we let $\mathcal{A} = \{f_i = 0\} \cup \{g_i \geq 0\}$,

and if an SOS proof exists, we say

$$\mathcal{A} \vdash_k p \geq 0.$$

As before, we say $\mathcal{A} \vdash_k p \geq q$ if $\mathcal{A} \vdash_k p - q \geq 0$.

Fact: SoS proofs compose: if

$$\mathcal{A} \vdash_{k_1} \mathcal{B} \quad \text{and} \quad \mathcal{B} \vdash_{k_2} \mathcal{C}, \text{ then}$$

$$\mathcal{A} \vdash_{\max(k_1, k_2)} \mathcal{C}.$$

Why do we care about non-negativity of polynomials?

Can encode many hard problems!

e.g. max-cut

$$\max_{x_i \in \{+1,-1\}} \sum_{(i,j) \in E} \frac{1 - x_i x_j}{2}$$

$\leftarrow x_i^2 = 1$

If $\{x_i^2 = 1\}_{i=1}^n \vdash_k \sum_{(i,j) \in E} \frac{1 - x_i x_j}{2} \geq C$, then (for any k)

we can certify that the value of the optimal cut is $\geq C$.

Can encode many interesting combinatorial optimization problems in this way!

- max-cut
- 3-SAT
- sparsest cut
- ⋮

something I've worked on in the past

Can also encode many interesting learning problems

Thm: Suppose $\mathcal{A} = \{f_i = 0\}_{i=1}^{m} \cup \{g_i \geq 0\}_{i=1}^{k}$ sos implies $p \geq 0$ in deg k. $\mathcal{A} \vdash_k p \geq 0$. Then there is an algorithm which runs in time $n^{O(k)}$ which finds this proof.

proof: Let's consider the simple case where $\mathcal{A} = \emptyset$, i.e. we want to find an sos proof that

$$p \geq 0, \quad \text{if a proof exists.}$$

Key idea: monomial matrix.

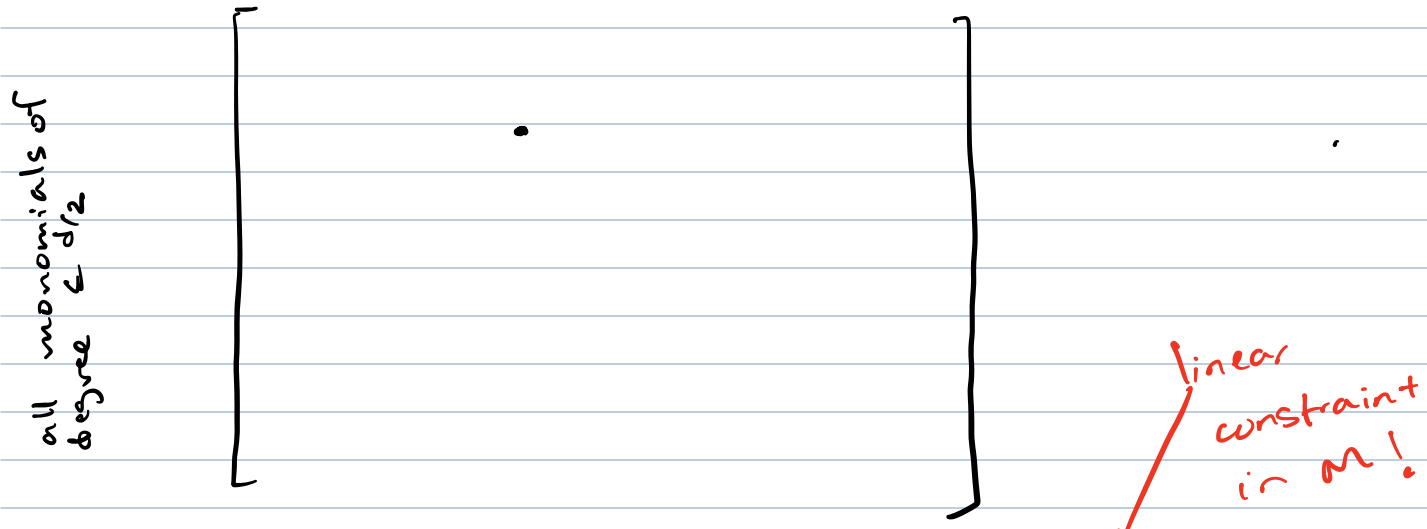
Any degree d polynomial can be encoded as a $n^{O(d)} \times n^{O(d)}$ size matrix

A deg-m monomial is indexed by a multi-index S w/ $|S| \leq m$.

$$x_1^2 x_3^3 x_{10}^7 \longleftrightarrow (2, 0, 3, \cdots, \underset{\uparrow}{7}, 0, \cdots, 0)$$
$$\text{shorthand} \hookrightarrow x^S \qquad 10^{\text{th}} \text{ position}$$

$$q \overset{\Longleftarrow}{\Longrightarrow} M_q$$

all monomials of degree $\leq d/2$



Constraint: $\forall T$:

$$\sum_{(R,S):\ |R|,|S| \leq d/2,\ R+S=T} (M_q)_{R,S} = \text{coefficient of } x^T \text{ in } q.$$

and M is symmetric

Why do we care? Let $s(x) = \sum_{|S| \le d/2} a_S x^S$ be some degree $\le d/2$ polynomial. Then:

$$s^2(x) = \left(\sum_{|S| \le d/2} a_S x^S \right)^2$$

$$= \sum_{\substack{R, S \\ |R|, |S| \le d/2}} a_R a_S x^R x^S = \sum_{\substack{R, S \\ |R|, |S| \le d/2}} a_R a_S x^{R+S}$$

$$= \sum_{|T| \le d/2} x^T \left(\sum_{\substack{|R|, |S| \le d/2 \\ R+S=T}} a_R a_S \right)$$

Associate S w/ a <u>vector</u> indexed by all monomials w/ deg $\le d/2$.

$$S \iff \begin{bmatrix} a_\phi \\ a_{\{1\}} \\ \vdots \\ a_R \\ \vdots \\ S \end{bmatrix} = v.$$

$$vv^T = \begin{bmatrix} & \vdots & \\ R & \cdots\cdots a_S a_R & \\ & & \end{bmatrix}$$

so vv^T is a valid matrix for $s^2(x)$!

rank-1 outer products $\iff$ squared polynomials.

Note also vv^T is PSD!

$$\forall x: \quad x^T(vv^T)x = \langle v, x\rangle^2 \geq 0.$$

So a sum of squares corresponds to

$v_1, \dots, v_t$ s.t.

$$M = \sum_{i=1}^{t} v_i v_i^T = \begin{bmatrix} & V & \end{bmatrix}\begin{bmatrix} & V^T & \end{bmatrix}$$

Claim: such a decomposition exists iff M is PSD!

pf: $\Leftarrow$ If M is PSD:

$$M = U^T D U = U^T D^{1/2} D^{1/2} U$$

allowed b/c $D_{ii} \geq 0$

V $\quad$ V^T

$\Rightarrow$ If $M = \sum v_i v_i^T$

$$x^T M x = \sum \langle x, v_i\rangle^2 \geq 0 \Rightarrow M \succcurlyeq 0.$$

So the following SDP finds a sos proof if it exists:

find $M \succcurlyeq 0$ s.t. M satisfies the linear constraints

$$\forall T, \quad \sum_{\substack{|R|,|S| \leq d/2 \\ R+S=T}} M_{R,S} = c_T.$$

Ph.D at UW!

Raghavendra's Theorem: The sos hierachy achieves optimal approximation ratios for almost all "nice"

constraint satisfaction problems, assuming the unique games conjecture.