Sums of Squares on the Hypercube

Greg Blekherman [1] João Gouveia [2] James Pfeiffer [3]

[1]Georgia Tech

[2]Universidade de Coimbra

[3]University of Washington

Section 1

Introduction

Nonnegativity of a polynomial

Let $I \subseteq \mathbb{R}[x]$ be an ideal:

$$\mathcal{P}(I) = \{p \in \mathbb{R}[I] \ : \ p \text{ is nonnegative on } \mathcal{V}_{\mathbb{R}}(I)\}.$$

Efficiently checking membership in $\mathcal{P}(I)$ is important for polynomial optimization.

Nonnegativity of a polynomial

Let $I \subseteq \mathbb{R}[x]$ be an ideal:

$$\mathcal{P}(I) = \{p \in \mathbb{R}[I] \;:\; p \text{ is nonnegative on } \mathcal{V}_{\mathbb{R}}(I)\}.$$

Efficiently checking membership in $\mathcal{P}(I)$ is important for polynomial optimization.

A typical strategy is to approximate $\mathcal{P}(I)$ by

$$\Sigma(I) = \left\{ p \in \mathbb{R}[I] \;:\; p \equiv \sum_{i=1}^{t} h_i^2 \text{ for some } h_i \in \mathbb{R}[I] \right\},$$

and its truncations

$$\Sigma_k(I) = \left\{ p \in \mathbb{R}[I] \;:\; p \equiv \sum_{i=1}^{t} h_i^2 \text{ for some } h_i \in \mathbb{R}_k[I] \right\}.$$

Sums of squares

Notes

- $p \in \Sigma_k(I)$ is said to be k-sos (modulo I).

Sums of squares

Notes

- $p \in \Sigma_k(I)$ is said to be k-sos (modulo I).
- $\Sigma_1(I) \subseteq \Sigma_2(I) \subseteq \cdots \subseteq \Sigma(I) \subseteq \mathcal{P}(I)$.

Sums of squares

Notes

- $p \in \Sigma_k(I)$ is said to be k-sos (modulo I).
- $\Sigma_1(I) \subseteq \Sigma_2(I) \subseteq \cdots \subseteq \Sigma(I) \subseteq \mathcal{P}(I)$.
- Checking membership in $\Sigma_k(I)$ is doable (SDP feasibility)

Sums of squares

Notes

- $p \in \Sigma_k(I)$ is said to be k-sos (modulo I).
- $\Sigma_1(I) \subseteq \Sigma_2(I) \subseteq \cdots \subseteq \Sigma(I) \subseteq \mathcal{P}(I)$.
- Checking membership in $\Sigma_k(I)$ is doable (SDP feasibility)
- Optimizing over $\Sigma_k(I)$ is doable (SDP)

Sums of squares

Notes

- $p \in \Sigma_k(I)$ is said to be k-sos (modulo I).
- $\Sigma_1(I) \subseteq \Sigma_2(I) \subseteq \cdots \subseteq \Sigma(I) \subseteq \mathcal{P}(I)$.
- Checking membership in $\Sigma_k(I)$ is doable (SDP feasibility)
- Optimizing over $\Sigma_k(I)$ is doable (SDP)

When are sums of squares enough?

Theorem (Hilbert 1888)

$\Sigma_k(\mathbb{R}^n) = \mathcal{P}_{2k}(\mathbb{R}^n)$ *if and only if* $n = 1$, $k = 1$ *or* $(n, k) = (2, 2)$.

Hilbert's 17th problem

Theorem (Artin 1927 - *Hilbert's 17th problem*)

A polynomial is nonnegative if and only if it is a sum of squares of rational functions.

Hilbert's 17th problem

Theorem (Artin 1927 - *Hilbert's 17th problem*)

A polynomial is nonnegative if and only if it is a sum of squares of rational functions.

We can use these stronger certificates.

$p \in \mathbb{R}[I]$ is (d,k)-rsos if for $0 \neq h \in \Sigma_d(I)$ we have $hp \in \Sigma_k(I)$.

Hilbert's 17th problem

Theorem (Artin 1927 - *Hilbert's 17th problem*)

A polynomial is nonnegative if and only if it is a sum of squares of rational functions.

We can use these stronger certificates.

$p \in \mathbb{R}[I]$ is (d, k)-rsos if for $0 \neq h \in \Sigma_d(I)$ we have $hp \in \Sigma_k(I)$.

If $d = \lfloor k - \deg(p)/2 \rfloor$ we will just say p is k-rsos.

Hilbert's 17th problem

Theorem (Artin 1927 - *Hilbert's 17th problem*)

A polynomial is nonnegative if and only if it is a sum of squares of rational functions.

We can use these stronger certificates.

$p \in \mathbb{R}[I]$ is (d, k)-rsos if for $0 \neq h \in \Sigma_d(I)$ we have $hp \in \Sigma_k(I)$.

If $d = \lfloor k - \deg(p)/2 \rfloor$ we will just say p is k-rsos.

We are interested in bounding how big must k be for a given polynomial to be k-rsos.

Hilbert's 17th problem

Theorem (Artin 1927 - *Hilbert's 17th problem*)

A polynomial is nonnegative if and only if it is a sum of squares of rational functions.

We can use these stronger certificates.

$p \in \mathbb{R}[I]$ is (d, k)-rsos if for $0 \neq h \in \Sigma_d(I)$ we have $hp \in \Sigma_k(I)$.

If $d = \lfloor k - \deg(p)/2 \rfloor$ we will just say p is k-rsos.

We are interested in bounding how big must k be for a given polynomial to be k-rsos.
In other words, we want to bound the degrees of the denominators in the rational functions used.

Advantages and Disadavantages

Schmudgen's Positivstellensatz

If $\mathcal{V}_{\mathbb{R}}(I)$ is compact, p positive on $\mathcal{V}_{\mathbb{R}}(I)$ implies p is k-sos for some k.

No uniform bounds on how big can k be.

Advantages and Disadavantages

Schmudgen's Positivstellensatz

If $\mathcal{V}_{\mathbb{R}}(I)$ is compact, p positive on $\mathcal{V}_{\mathbb{R}}(I)$ implies p is k-sos for some k.

No uniform bounds on how big can k be.

Stengle's Positivstellensatz

For any I, p nonnegative on $\mathcal{V}_{\mathbb{R}}(I)$ implies p is k-rsos for some k.

k is uniformly bounded depending only on $\deg(p)$ and on I.

Advantages and Disadavantages

Schmudgen's Positivstellensatz

If $\mathcal{V}_{\mathbb{R}}(I)$ is compact, p positive on $\mathcal{V}_{\mathbb{R}}(I)$ implies p is k-sos for some k.

No uniform bounds on how big can k be.

Stengle's Positivstellensatz

For any I, p nonnegative on $\mathcal{V}_{\mathbb{R}}(I)$ implies p is k-rsos for some k.

k is uniformly bounded depending only on $\deg(p)$ and on I.

No free lunches

Advantages and Disadavantages

Schmudgen's Positivstellensatz

If $\mathcal{V}_{\mathbb{R}}(I)$ is compact, p positive on $\mathcal{V}_{\mathbb{R}}(I)$ implies p is k-sos for some k.

No uniform bounds on how big can k be.

Stengle's Positivstellensatz

For any I, p nonnegative on $\mathcal{V}_{\mathbb{R}}(I)$ implies p is k-rsos for some k.

k is uniformly bounded depending only on deg(p) and on I.

No free lunches

- Checking k-rsosness is still an SDP feasibility problem.

Advantages and Disadavantages

Schmudgen's Positivstellensatz

If $\mathcal{V}_{\mathbb{R}}(I)$ is compact, p positive on $\mathcal{V}_{\mathbb{R}}(I)$ implies p is k-sos for some k.

No uniform bounds on how big can k be.

Stengle's Positivstellensatz

For any I, p nonnegative on $\mathcal{V}_{\mathbb{R}}(I)$ implies p is k-rsos for some k.

k is uniformly bounded depending only on $\deg(p)$ and on I.

No free lunches

- Checking k-rsosness is still an SDP feasibility problem.
- Optimizing over the set of all k-rsos polynomials is not as easy.

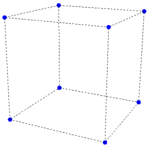
Section 2

Upper bounds on multipliers

The n-cube

We are interested in the n-cube:

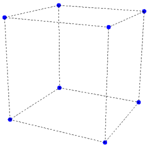
$$C_n = \{0,1\}^n = \{x \in \mathbb{R}^n \ : \ x_i^2 - x_i = 0, \ i = 1, \cdots, n\} = \mathcal{V}(I_n).$$



The n-cube

We are interested in the n-cube:

$$C_n = \{0,1\}^n = \{x \in \mathbb{R}^n \ : \ x_i^2 - x_i = 0,\ i = 1, \cdots, n\} = \mathcal{V}(I_n).$$



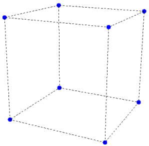
Corollary

Every nonnegative quadratic polynomial on C_n is $(\lfloor n/2 \rfloor + 1)$-rsos.

The n-cube

We are interested in the n-cube:

$$C_n = \{0,1\}^n = \{x \in \mathbb{R}^n \ : \ x_i^2 - x_i = 0, \ i = 1, \cdots, n\} = \mathcal{V}(I_n).$$



Corollary

Every nonnegative quadratic polynomial on C_n is $(\lfloor n/2 \rfloor + 1)$-rsos.

Main Lemma

Let $\ell : \mathbb{R}[X]_{2d} \to \mathbb{R}$ be given by $\ell(f) = \sum_{v \in X} \mu_v f(v)$ with all $\mu_v \neq 0$. Suppose that ℓ is nonnegative on $\Sigma_d(X)$. Then

$$\#\{v \in X : \mu_v > 0\} \geq \dim \mathbb{R}[X]_d.$$

Positive certificates

Unfortunately, being rsos modulo I_n does not guarantee nonnegativity, since the variety is not irreducible.

Positive certificates

Unfortunately, being rsos modulo I_n does not guarantee nonnegativity, since the variety is not irreducible.

$p \in \mathbb{R}[I]$ is (d,k)-rsos with positive multipliers if for $h \in \Sigma_d(I)$ we have $(1+h)p \in \Sigma_k(I)$.

Positive certificates

Unfortunately, being rsos modulo I_n does not guarantee nonnegativity, since the variety is not irreducible.

$p \in \mathbb{R}[I]$ is (d, k)-rsos with positive multipliers if for $h \in \Sigma_d(I)$ we have $(1 + h)p \in \Sigma_k(I)$.

Equivalently

$p \in \mathbb{R}[I]$ is (d, k)-rsos with positive multipliers if for $h \in \text{int}(\Sigma_d(I))$ we have $hp \in \Sigma_k(I)$.

Positive certificates

Unfortunately, being rsos modulo I_n does not guarantee nonnegativity, since the variety is not irreducible.

$p \in \mathbb{R}[I]$ is (d,k)-rsos with positive multipliers if for $h \in \Sigma_d(I)$ we have $(1+h)p \in \Sigma_k(I)$.

Equivalently

$p \in \mathbb{R}[I]$ is (d,k)-rsos with positive multipliers if for $h \in \text{int}(\Sigma_d(I))$ we have $hp \in \Sigma_k(I)$.

Theorem

Every nonnegative quadratic polynomial on C_n is $(\lfloor n/2 \rfloor + 2)$-rsos with positive multipliers.

Positive certificates

Unfortunately, being rsos modulo I_n does not guarantee nonnegativity, since the variety is not irreducible.

$p \in \mathbb{R}[I]$ is (d,k)-rsos with positive multipliers if for $h \in \Sigma_d(I)$ we have $(1+h)p \in \Sigma_k(I)$.

Equivalently

$p \in \mathbb{R}[I]$ is (d,k)-rsos with positive multipliers if for $h \in \text{int}(\Sigma_d(I))$ we have $hp \in \Sigma_k(I)$.

Theorem

Every nonnegative quadratic polynomial on C_n is $(\lfloor n/2 \rfloor + 2)$-rsos with positive multipliers.

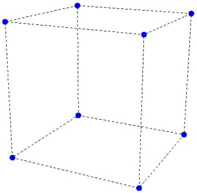
Open Question: Is the increased degree needed?

Section 3

Lower bounds on hypercube multipliers

Hypercube

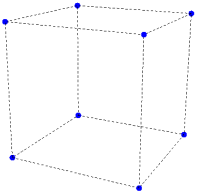
We will again focus solely on the n-cube $C_n = \{0, 1\}^n$.



Cube C_3

Hypercube

We will again focus solely on the n-cube $C_n = \{0,1\}^n$.



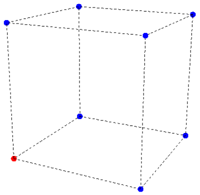
Cube C_3

S_n acts on C_n by permuting coordinates, and if p is symmetric, it will be completely characterized by its evaluation at the levels T_k of the cube:

$$T_k = \{x \in C_n \ : \ \sum x_i = k\}.$$

Hypercube

We will again focus solely on the n-cube $C_n = \{0, 1\}^n$.



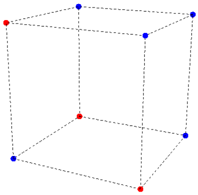
Level T_0

S_n acts on C_n by permuting coordinates, and if p is symmetric, it will be completely characterized by its evaluation at the levels T_k of the cube:

$$T_k = \{x \in C_n \ : \ \sum x_i = k\}.$$

Hypercube

We will again focus solely on the n-cube $C_n = \{0,1\}^n$.



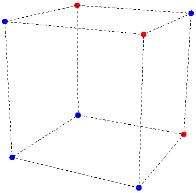
Level T_1

S_n acts on C_n by permuting coordinates, and if p is symmetric, it will be completely characterized by its evaluation at the levels T_k of the cube:

$$T_k = \{x \in C_n \ : \ \sum x_i = k\}.$$

Hypercube

We will again focus solely on the n-cube $C_n = \{0,1\}^n$.



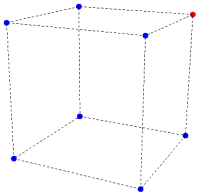
Level T_2

S_n acts on C_n by permuting coordinates, and if p is symmetric, it will be completely characterized by its evaluation at the levels T_k of the cube:

$$T_k = \{x \in C_n \ : \ \sum x_i = k\}.$$

Hypercube

We will again focus solely on the n-cube $C_n = \{0,1\}^n$.



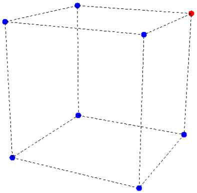
Level T_3

S_n acts on C_n by permuting coordinates, and if p is symmetric, it will be completely characterized by its evaluation at the levels T_k of the cube:

$$T_k = \{x \in C_n \ : \ \sum x_i = k\}.$$

Hypercube

We will again focus solely on the n-cube $C_n = \{0, 1\}^n$.



Level T_3

S_n acts on C_n by permuting coordinates, and if p is symmetric, it will be completely characterized by its evaluation at the levels T_k of the cube:

$$T_k = \{x \in C_n \ : \ \sum x_i = k\}.$$

Symmetric polynomials appear naturally in combinatorial optimization, and we want lower bounds for the degree of nonnegativity certificates.

The bound

Lemma

Suppose $f \in \mathbb{R}_d[I_n]$*, vanishes on* T_t*. If* $d \leq t \leq n - d$*, then* f *is properly divisible by* $\ell = t - \sum x_i$*.*

The bound

Lemma

Suppose $f \in \mathbb{R}_d[I_n]$, vanishes on T_t. If $d \leq t \leq n - d$, then f is properly divisible by $\ell = t - \sum x_i$.

(by *properly divisible* we mean $f = \ell g$ with $\deg(g) < \deg(f)$).

The bound

Lemma

Suppose $f \in \mathbb{R}_d[I_n]$, vanishes on T_t. If $d \leq t \leq n - d$, then f is properly divisible by $\ell = t - \sum x_i$.

(by *properly divisible* we mean $f = \ell g$ with $\deg(g) < \deg(f)$).

Proof: Classic representation theory + magic.

The bound

Lemma

Suppose $f \in \mathbb{R}_d[I_n]$, vanishes on T_t. If $d \leq t \leq n - d$, then f is properly divisible by $\ell = t - \sum x_i$.

(by *properly divisible* we mean $f = \ell g$ with $\deg(g) < \deg(f)$).

Proof: Classic representation theory + magic.

Theorem

Suppose $f \in \mathbb{R}_t[I_n]$ with $t \leq n/2$ is an S_n-invariant polynomial and f is properly divisible by $\ell = t - (x_1 + \cdots + x_n)$ to odd order. Then f is not d-rsos for $d \leq t$.

Quadratics

In particular we have:

Theorem

Let $k = \lfloor \frac{n}{2} \rfloor$ *and let* $f \in \mathbb{R}[I_n]$ *be given by*

$$f = (x_1 + \cdots + x_n - k)(x_1 + \cdots + x_n - k - 1).$$

Then f *is nonnegative on* C_n *but* f *is not* k*-rsos.*

Quadratics

In particular we have:

Theorem

Let $k = \lfloor \frac{n}{2} \rfloor$ and let $f \in \mathbb{R}[I_n]$ be given by

$$f = (x_1 + \cdots + x_n - k)(x_1 + \cdots + x_n - k - 1).$$

Then f is nonnegative on C_n but f is not k-rsos.

This shows our upper bound was tight.

Section 4

Applications

Globally nonnegative polynomials

We can leverage our result to obtain lower bounds for Hilbert's 17th problem.

Globally nonnegative polynomials

We can leverage our result to obtain lower bounds for Hilbert's 17th problem.

Corollary

Let $k = \lfloor \frac{n}{2} \rfloor$. There exists a polynomial p of degree 4 nonnegative on $\mathbb{R}^n$ which is not k-rsos in $\mathbb{R}[x_1, \ldots, x_n]$.

Globally nonnegative polynomials

We can leverage our result to obtain lower bounds for Hilbert's 17th problem.

Corollary

Let $k = \lfloor \frac{n}{2} \rfloor$. There exists a polynomial p of degree 4 nonnegative on $\mathbb{R}^n$ which is not k-rsos in $\mathbb{R}[x_1, \ldots, x_n]$.

Globally nonnegative polynomials

We can leverage our result to obtain lower bounds for Hilbert's 17th problem.

Corollary

Let $k = \lfloor \frac{n}{2} \rfloor$. There exists a polynomial p of degree 4 nonnegative on $\mathbb{R}^n$ which is not k-rsos in $\mathbb{R}[x_1, \ldots, x_n]$.

This is proven by a perturbed extension of the polynomial on the previous theorem:

$$p = (x_1 + \cdots + x_n - k)(x_1 + \cdots + x_n - k - 1)$$

Globally nonnegative polynomials

We can leverage our result to obtain lower bounds for Hilbert's 17th problem.

Corollary

Let $k = \lfloor \frac{n}{2} \rfloor$. There exists a polynomial p of degree 4 nonnegative on $\mathbb{R}^n$ which is not k-rsos in $\mathbb{R}[x_1, \ldots, x_n]$.

This is proven by a perturbed extension of the polynomial on the previous theorem:

$$p = (x_1 + \cdots + x_n - k)(x_1 + \cdots + x_n - k - 1) + \varepsilon$$

Globally nonnegative polynomials

We can leverage our result to obtain lower bounds for Hilbert's 17th problem.

Corollary

Let $k = \lfloor \frac{n}{2} \rfloor$. There exists a polynomial p of degree 4 nonnegative on $\mathbb{R}^n$ which is not k-rsos in $\mathbb{R}[x_1, \ldots, x_n]$.

This is proven by a perturbed extension of the polynomial on the previous theorem:

$$p = (x_1 + \cdots + x_n - k)(x_1 + \cdots + x_n - k - 1) + \varepsilon + A \sum_i ({x_i}^2 - x_i)^2.$$

MaxCut

The maxcut problem over K_n can be reduced to

Binary polynomial formulation of MaxCut

$$\max p(x) = \sum_{i \neq j} (1 - x_i) x_j \text{ s.t. } x \in C_n$$

MaxCut

The maxcut problem over K_n can be reduced to

Binary polynomial formulation of MaxCut

$$\max p(x) = \sum_{i \neq j} (1 - x_i) x_j \text{ s.t. } x \in C_n$$

Laurent has proved that Lassere relaxations are of limited use.

Laurent

For $n = 2k + 1$, $p_{\text{sos}}^k > p_{\max}$.

MaxCut

The maxcut problem over K_n can be reduced to

Binary polynomial formulation of MaxCut

$$\max p(x) = \sum_{i \neq j} (1 - x_i) x_j \text{ s.t. } x \in C_n$$

Laurent has proved that Lassere relaxations are of limited use.

Laurent

For $n = 2k + 1$, $p^k_{\text{sos}} > p_{\text{max}}$.

Note that p attains its maximum in C_n at T_k and T_{k+1} so

Theorem

For $n = 2k + 1$, $p^k_{\text{rsos}} > p_{\text{max}}$.

MaxCut 2

Consider the weighted maxcut formulation.

Binary polynomial formulation of MaxCut

$$\max p_\omega(x) = \sum_{i \neq j} \omega_{ij}(1 - x_i)x_j \text{ s.t. } x \in C_n.$$

MaxCut 2

Consider the weighted maxcut formulation.

Binary polynomial formulation of MaxCut

$$\max p_\omega(x) = \sum_{i \neq j} \omega_{ij}(1 - x_i)x_j \text{ s.t. } x \in C_n.$$

Conjecture (Laurent)

If $n = 2k + 1$, $(p_\omega)_{\max} = (p_\omega)_{\text{sos}}^{k+1}$ for all weights.

MaxCut 2

Consider the weighted maxcut formulation.

Binary polynomial formulation of MaxCut

$$\max p_\omega(x) = \sum_{i \neq j} \omega_{ij}(1 - x_i)x_j \text{ s.t. } x \in C_n.$$

Conjecture (Laurent)

If $n = 2k + 1$, $(p_\omega)_{\max} = (p_\omega)^{k+1}_{\text{sos}}$ for all weights.

A weaker version can now be proved.

Theorem

If $n = 2k + 1$, $(p_\omega)_{\max} = (p_\omega)^{k+1}_{\text{rsos}}$ for all weights or $(p_\omega)^{k+2}_{\text{rsos}}$ if we want positive multipliers.

Thank You