Convex Optimization

(EE227A: UC Berkeley)

Lecture 28
(Algebra + Optimization)

02 May, 2013

Suvrit Sra

Admin

- Poster presentation on **10th May** — mandatory
- HW, Midterm, Quiz — to be reweighted
- Project final report on **16th May** — upload to easychair
- Any questions / concerns: email me!
- Email me if you need to meet

Convex sets: geometry vs algebra

- Geometry of convex sets is very rich and well-understood (we didn't cover much of it)

Convex sets: geometry vs algebra

- Geometry of convex sets is very rich and well-understood (we didn't cover much of it)
- But what about (efficient) **representation** of these geometric objects?

Convex sets: geometry vs algebra

- Geometry of convex sets is very rich and well-understood (we didn't cover much of it)
- But what about (efficient) **representation** of these geometric objects?
- How do algebraic, geometric, computational aspects interact?

Convex sets: geometry vs algebra

- Geometry of convex sets is very rich and well-understood (we didn't cover much of it)
- But what about (efficient) **representation** of these geometric objects?
- How do algebraic, geometric, computational aspects interact?
- Semidefinite programming plays a major role!

Convex sets: geometry vs algebra

- Geometry of convex sets is very rich and well-understood (we didn't cover much of it)
- But what about (efficient) **representation** of these geometric objects?
- How do algebraic, geometric, computational aspects interact?
- Semidefinite programming plays a major role!

☞ A nice book for detailed development of these ideas:

G. Blekherman, P. Parrilo, R. R. Thomas. *Semidefinite optimization and convex algebraic geometry* (2012).

Polyhedral sets

Recall (convex) *polyhedra*, described by **finitely** many half-spaces

$$\left\{x \in \mathbb{R}^n \mid a_i^T x \leq b_i, \quad i = 1, \ldots, m\right\}.$$

Polyhedral sets

Recall (convex) *polyhedra*, described by **finitely** many half-spaces

$$\left\{x \in \mathbb{R}^n \mid a_i^T x \leq b_i, \quad i = 1, \ldots, m\right\}.$$

Convex polyhedra have many nice properties:

- Remain preserved under projection (Fourier-Motzkin elimination)

Polyhedral sets

Recall (convex) *polyhedra*, described by **finitely** many half-spaces

$$\left\{x \in \mathbb{R}^n \mid a_i^T x \leq b_i, \quad i = 1, \ldots, m\right\}.$$

Convex polyhedra have many nice properties:

- Remain preserved under projection (Fourier-Motzkin elimination)
- Farkas lemma / duality theory gives emptiness test

Polyhedral sets

Recall (convex) *polyhedra*, described by **finitely** many half-spaces

$$\left\{x \in \mathbb{R}^n \mid a_i^T x \leq b_i, \quad i = 1, \ldots, m\right\}.$$

Convex polyhedra have many nice properties:

- Remain preserved under projection (Fourier-Motzkin elimination)
- Farkas lemma / duality theory gives emptiness test
- Optimization over cvx polyhedra is linear programming.

Polyhedral sets

Recall (convex) *polyhedra*, described by **finitely** many half-spaces

$$\left\{x \in \mathbb{R}^n \mid a_i^T x \leq b_i, \quad i = 1, \ldots, m\right\}.$$

Convex polyhedra have many nice properties:

- Remain preserved under projection (Fourier-Motzkin elimination)
- Farkas lemma / duality theory gives emptiness test
- Optimization over cvx polyhedra is linear programming.

But getting away from linearity....

SDP, LMIs

We've seen SOCPs, SDPs as substantial generalization.

SDP, LMIs

We've seen SOCPs, SDPs as substantial generalization.

Semidefinite representations

SDP, LMIs

We've seen SOCPs, SDPs as substantial generalization.

Semidefinite representations

Which sets can be represented via SDPs?

SDP, LMIs

We've seen SOCPs, SDPs as substantial generalization.

Semidefinite representations

Which sets can be represented via SDPs?

- LP case—well-understood: if a set is polyhedral (i.e., finite number of extreme points / rays)

SDP, LMIs

We've seen SOCPs, SDPs as substantial generalization.

Semidefinite representations

Which sets can be represented via SDPs?

- LP case—well-understood: if a set is polyhedral (i.e., finite number of extreme points / rays)
- Do we have a similar nice characterization in SDP case?

SDP, LMIs

We've seen SOCPs, SDPs as substantial generalization.

Semidefinite representations

Which sets can be represented via SDPs?

- LP case—well-understood: if a set is polyhedral (i.e., finite number of extreme points / rays)
- Do we have a similar nice characterization in SDP case?
- We've seen a few SDRs in Lecture 6 (polyhedra, matrix norms, second order cones, etc.)

SDP, LMIs

We've seen SOCPs, SDPs as substantial generalization.

Semidefinite representations

Which sets can be represented via SDPs?

- LP case—well-understood: if a set is polyhedral (i.e., finite number of extreme points / rays)
- Do we have a similar nice characterization in SDP case?
- We've seen a few SDRs in Lecture 6 (polyhedra, matrix norms, second order cones, etc.)
- Preserved under standard "convex algebra": affine transformations, convex hulls, taking polars, etc.

SDP, LMIs

We've seen SOCPs, SDPs as substantial generalization.

Semidefinite representations

Which sets can be represented via SDPs?

- LP case—well-understood: if a set is polyhedral (i.e., finite number of extreme points / rays)
- Do we have a similar nice characterization in SDP case?
- We've seen a few SDRs in Lecture 6 (polyhedra, matrix norms, second order cones, etc.)
- Preserved under standard "convex algebra": affine transformations, convex hulls, taking polars, etc.
- See lecture notes by A. Nemirovski for SDR (and conic) calculus

Semidefinite representations

Can $\mathcal{S}$ be represented via SDPs?

- $\mathcal{S}$ must be convex and **semialgebraic**

Semidefinite representations

Can $\mathcal{S}$ be represented via SDPs?

- $\mathcal{S}$ must be convex and **semialgebraic**
  $\mathcal{S}$ can be defined using a finite number of polynomial inequalities.

Semidefinite representations

Can $\mathcal{S}$ be represented via SDPs?

- $\mathcal{S}$ must be convex and **semialgebraic**
  $\mathcal{S}$ can be defined using a finite number of polynomial inequalities.
- Exact or approx. representations (also, relaxing nonconvex $\mathcal{S}$)

Semidefinite representations

Can $\mathcal{S}$ be represented via SDPs?

- $\mathcal{S}$ must be convex and **semialgebraic**
  $\mathcal{S}$ can be defined using a finite number of polynomial inequalities.
- Exact or approx. representations (also, relaxing nonconvex $\mathcal{S}$)
- Example ("direct" representation)

$$x \in \mathcal{S} \quad \Leftrightarrow \quad A_0 + \sum_i x_i A_i \succeq 0$$

Semidefinite representations

Can $\mathcal{S}$ be represented via SDPs?

- $\mathcal{S}$ must be convex and **semialgebraic**
  $\mathcal{S}$ can be defined using a finite number of polynomial inequalities.
- Exact or approx. representations (also, relaxing nonconvex $\mathcal{S}$)
- Example ("direct" representation)

$$x \in \mathcal{S} \quad \Leftrightarrow \quad A_0 + \sum_i x_i A_i \succeq 0$$

- "Lifted" representation (recall HW2), can use extra variables

$$x \in \mathcal{S} \quad \Leftrightarrow \exists y \text{ s.t. } A(x) + B(y) \succeq 0.$$

- This "projection" / lifting technique can be very useful.

Classic example

Lifting / projection

Classic example n-dimensional ℓ_1-unit ball (crosspolytope).
Requires 2^n inequalities of the form

$$\pm x_1 \pm x_2 \cdots \pm x_n \leq 1.$$

Lifting / projection

Classic example n-dimensional ℓ_1-unit ball (crosspolytope).
Requires 2^n inequalities of the form

$$\pm x_1 \pm x_2 \cdots \pm x_n \leq 1.$$

But we can efficiently represent it as a *projection*:

$$\left\{(x, y) \in \mathbb{R}^{2n} \mid \sum_i y_i = 1, \quad -y_i \leq x_i \leq y_i, \quad i = 1, \ldots, n\right\}.$$

Just $2n$ variables and $2n + 1$ constraints

Lifting / projection

Classic example n-dimensional ℓ_1-unit ball (crosspolytope).
Requires 2^n inequalities of the form

$$\pm x_1 \pm x_2 \cdots \pm x_n \leq 1.$$

But we can efficiently represent it as a *projection*:

$$\left\{(x,y) \in \mathbb{R}^{2n} \mid \sum_i y_i = 1, \quad -y_i \leq x_i \leq y_i, \quad i = 1,\ldots,n\right\}.$$

Just $2n$ variables and $2n+1$ constraints

Moral: When playing with convexity, rather than eliminating variables, often nicer to add new variables with which description of set can become simpler!

SDRs and LMIs

- Does every convex semialgebraic set $\mathcal{S}$ have a direct SDR?

SDRs and LMIs

- Does every convex semialgebraic set $\mathcal{S}$ have a direct SDR? (answer known in 2-dimensions)
- Does ever basic convex semialgebraic set have a lifted SDR?

SDRs and LMIs

- Does every convex semialgebraic set $\mathcal{S}$ have a direct SDR? (answer known in 2-dimensions)

- Does ever basic convex semialgebraic set have a lifted SDR?

Answers to both are unknown as of now

SDRs and LMIs

- Does every convex semialgebraic set $\mathcal{S}$ have a direct SDR? (answer known in 2-dimensions)

- Does ever basic convex semialgebraic set have a lifted SDR?

Answers to both are unknown as of now

Some partial results known. See references

SDRs and LMIs

- Does every convex semialgebraic set $\mathcal{S}$ have a direct SDR? (answer known in 2-dimensions)

- Does ever basic convex semialgebraic set have a lifted SDR?

Answers to both are unknown as of now

Some partial results known. See references

Let us look at SDR and approx SDR for polynomials

Polynomials

Def. (Polynomial). Let $\mathbb{K}$ be a field and $x_1, \ldots, x_n$ be indeterminates. A polynomial f with coefficients in a field $\mathbb{K}$ is a **finite** linear combination of **monomials**:

$$f = \sum_{\alpha} c_\alpha x^\alpha = \sum_{\alpha} x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \quad c_\alpha \in \mathbb{K};$$

we sum over finite n-tuples $\alpha = (\alpha_1, \ldots, \alpha_n)$, each $\alpha_i \in \mathbb{N}_0$.

- **Degree:** $d = \sum_i \alpha_i$ (largest such sum over all α)

Polynomials

Def. (Polynomial). Let $\mathbb{K}$ be a field and $x_1, \ldots, x_n$ be indeterminates. A polynomial f with coefficients in a field $\mathbb{K}$ is a **finite** linear combination of **monomials**:

$$f = \sum_{\alpha} c_\alpha x^\alpha = \sum_{\alpha} x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \quad c_\alpha \in \mathbb{K};$$

we sum over finite n-tuples $\alpha = (\alpha_1, \ldots, \alpha_n)$, each $\alpha_i \in \mathbb{N}_0$.

- **Degree:** $d = \sum_i \alpha_i$ (largest such sum over all α)

Def. Ring of all polynomials $\mathbb{K}[x_1, \ldots, x_n]$

Polynomials

Def. (Polynomial). Let $\mathbb{K}$ be a field and $x_1, \ldots, x_n$ be indeterminates. A polynomial f with coefficients in a field $\mathbb{K}$ is a **finite** linear combination of **monomials**:

$$f = \sum_{\alpha} c_\alpha x^\alpha = \sum_{\alpha} x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \quad c_\alpha \in \mathbb{K};$$

we sum over finite n-tuples $\alpha = (\alpha_1, \ldots, \alpha_n)$, each $\alpha_i \in \mathbb{N}_0$.

- **Degree:** $d = \sum_i \alpha_i$ (largest such sum over all α)

Def. Ring of all polynomials $\mathbb{K}[x_1, \ldots, x_n]$

Eg: **Univariate** polynomials with real coefficients $\mathbb{R}[x]$

Nonnegativity

- We care about whether $p(x) \geq 0$ for all x

Nonnegativity

- We care about whether $p(x) \geq 0$ for all x
- Question equivalent to SDR for univariate polynomials

Nonnegativity

- We care about whether $p(x) \geq 0$ for all x
- Question equivalent to SDR for univariate polynomials
- For multivariate polynomials, question remains very important

Nonnegativity

- We care about whether $p(x) \geq 0$ for all x
- Question equivalent to SDR for univariate polynomials
- For multivariate polynomials, question remains very important
- (Nonnegativity intimately tied to convexity (formally real fields, algebraic closure, ordered property etc.))

Nonnegativity

- We care about whether $p(x) \geq 0$ for all x
- Question equivalent to SDR for univariate polynomials
- For multivariate polynomials, question remains very important
- (Nonnegativity intimately tied to convexity (formally real fields, algebraic closure, ordered property etc.))

☞ If $p(x) \geq 0$, then degree of p must be even

☞ Set of nonnegative polynomials quite interesting.

Theorem Let $\mathscr{P}_n$ denote the set of all nonnegative univariate polynomials of degree $\leq n$. Identifying a polynomial with its $n+1$ coefficients $(p_n, \ldots, p_0)$, the set $\mathscr{P}_n$ is a closed, convex, pointed cone in $\mathbb{R}^{n+1}$

Testing nonnegativity

Def. (SOS). A univariate polynomial $p(x)$ is a sum of squares (SOS) if there exist $q_1, \ldots, q_m \in \mathbb{R}[x]$ such that

$$p(x) = \sum_{k=1}^{m} q_k^2(x).$$

Testing nonnegativity

Def. (SOS). A univariate polynomial $p(x)$ is a sum of squares (SOS) if there exist $q_1, \ldots, q_m \in \mathbb{R}[x]$ such that

$$p(x) = \sum_{k=1}^{m} q_k^2(x).$$

Theorem A univariate polynomial is nonneg **if and only if** it is SOS

Proof: Obviously, if $p(x)$ is SOS, then $p(x) \geq 0$.

Testing nonnegativity

Def. (SOS). A univariate polynomial $p(x)$ is a sum of squares (SOS) if there exist $q_1, \ldots, q_m \in \mathbb{R}[x]$ such that

$$p(x) = \sum_{k=1}^{m} q_k^2(x).$$

Theorem A univariate polynomial is nonneg **if and only if** it is SOS

Proof: Obviously, if $p(x)$ is SOS, then $p(x) \geq 0$. For converse, recall by the fundamental theorem of algebra, we can factorize

$$p(x) = p_n \prod_j (x - r_j)^{n_j} \prod_k (x - z_k)^{m_k} (x - \bar{z}_k)^{m_k},$$

where r_j and z_k are real and complex roots, respectively.

Testing nonnegativity

Def. (SOS). A univariate polynomial $p(x)$ is a sum of squares (SOS) if there exist $q_1, \ldots, q_m \in \mathbb{R}[x]$ such that

$$p(x) = \sum_{k=1}^{m} q_k^2(x).$$

Theorem A univariate polynomial is nonneg **if and only if** it is SOS

Proof: Obviously, if $p(x)$ is SOS, then $p(x) \geq 0$. For converse, recall by the fundamental theorem of algebra, we can factorize

$$p(x) = p_n \prod_j (x - r_j)^{n_j} \prod_k (x - z_k)^{m_k} (x - \bar{z}_k)^{m_k},$$

where r_j and z_k are real and complex roots, respectively.
Since $p(x) \geq 0$, $p_n > 0$, multiplicities n_j of real roots are even. Also, note $(x - z)(x - \bar{z}) = (x - a)^2 + b^2$, if $z = a + ib$.

Testing nonnegativity

Def. (SOS). A univariate polynomial $p(x)$ is a sum of squares (SOS) if there exist $q_1, \dots, q_m \in \mathbb{R}[x]$ such that

$$p(x) = \sum_{k=1}^{m} q_k^2(x).$$

Theorem A univariate polynomial is nonneg **if and only if** it is SOS

Proof: Obviously, if $p(x)$ is SOS, then $p(x) \geq 0$. For converse, recall by the fundamental theorem of algebra, we can factorize

$$p(x) = p_n \prod_j (x - r_j)^{n_j} \prod_k (x - z_k)^{m_k} (x - \bar{z}_k)^{m_k},$$

where r_j and z_k are real and complex roots, respectively.
Since $p(x) \geq 0$, $p_n > 0$, multiplicities n_j of real roots are even. Also, note $(x - z)(x - \bar{z}) = (x - a)^2 + b^2$, if $z = a + ib$. Thus, we have

$$p(x) = \prod_j (x - r_j)^{2s_j} \prod_k [(x - a_k)^2 + b_k^2]^{m_k}.$$

Expand out above product of SOS into a sum to see that $p(x)$ is SOS.

SOS

Exercise: Show that in fact if $p(x) \geq 0$, then it can be written as a sum of just two squares, i.e., $p(x) = q_1^2(x) + q_2^2(x)$. (*Hint:* It may help to notice $(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ac + bd)^2$)

SOS

Exercise: Show that in fact if $p(x) \geq 0$, then it can be written as a sum of just two squares, i.e., $p(x) = q_1^2(x) + q_2^2(x)$. (*Hint:* It may help to notice $(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ac + bd)^2$)

Unfortunately, for multivariate polynomials SOS **not equivalent** to $p(x_1, \ldots, x_m) \geq 0$

(**Motzkin polynomial**)

$M(x, y) := x^4y^2 + x^2y^4 + 1 - 3x^2y^2$ nonneg but not SOS.

SOS and SDP

Theorem Let $p(x)$ be of degree $2d$. Then, $p(x) \geq 0$ (or SOS) if and only if there exists a $Q \in \mathcal{S}_+^{d+1}$ that satisfies $p(x) = z^T Q z$, where $z = [1, x, \ldots, x^d]^T$.

SOS and SDP

Theorem Let $p(x)$ be of degree $2d$. Then, $p(x) \geq 0$ (or SOS) if and only if there exists a $Q \in \mathcal{S}_+^{d+1}$ that satisfies $p(x) = z^T Q z$, where $z = [1, x, \ldots, x^d]^T$.

- If $p(x) \geq 0$, then we have $p(x) = \sum_i^m q_i^2(x)$

SOS and SDP

Theorem Let $p(x)$ be of degree $2d$. Then, $p(x) \geq 0$ (or SOS) if and only if there exists a $Q \in \mathcal{S}_+^{d+1}$ that satisfies $p(x) = z^T Q z$, where $z = [1, x, \ldots, x^d]^T$.

- If $p(x) \geq 0$, then we have $p(x) = \sum_i^m q_i^2(x)$
- Obviously, degree of any q_i at most d

SOS and SDP

Theorem Let $p(x)$ be of degree $2d$. Then, $p(x) \geq 0$ (or SOS) if and only if there exists a $Q \in \mathcal{S}_+^{d+1}$ that satisfies $p(x) = z^T Q z$, where $z = [1, x, \ldots, x^d]^T$.

- If $p(x) \geq 0$, then we have $p(x) = \sum_i^m q_i^2(x)$
- Obviously, degree of any q_i at most d
- Write a vector of polynomials

$$\begin{bmatrix} q_1(x) \\ q_2(x) \\ \vdots \\ q_m(x) \end{bmatrix} = V \begin{bmatrix} x^0 \\ x^1 \\ \vdots \\ x^d \end{bmatrix}$$

where row i of $V \in \mathbb{R}^{m \times (d+1)}$ contains coefficients of the q_i.

SOS and SDP

- Denote $[x]_d := [x^0, x^1, \ldots, x^d]^T$

SOS and SDP

- Denote $[x]_d := [x^0, x^1, \ldots, x^d]^T$
- Then, since $q = V[x]_d$, we have $\sum_i q_i^2(x) = (V[x]_d)^T(V[x]_d)$ which is nothing but $[x]_d^T Q[x]_d$, where $Q = V^T V \succeq 0$.

SOS and SDP

- Denote $[x]_d := [x^0, x^1, \dots, x^d]^T$
- Then, since $q = V[x]_d$, we have $\sum_i q_i^2(x) = (V[x]_d)^T(V[x]_d)$ which is nothing but $[x]_d^T Q[x]_d$, where $Q = V^T V \succeq 0$.
- Conversely, if there is a Q such that $p(x) = [x]_d^T Q[x]_d$, just take Cholesky factorization $Q = R^T R$, to obtain SOS decomp. of p

SOS and SDP

- Denote $[x]_d := [x^0, x^1, \dots, x^d]^T$
- Then, since $q = V[x]_d$, we have $\sum_i q_i^2(x) = (V[x]_d)^T(V[x]_d)$ which is nothing but $[x]_d^T Q [x]_d$, where $Q = V^T V \succeq 0$.
- Conversely, if there is a Q such that $p(x) = [x]_d^T Q [x]_d$, just take Cholesky factorization $Q = R^T R$, to obtain SOS decomp. of p
- If we are given p, how to find SOS decomp / matrix Q?

Remark: N. Z. Shor (inventor of subgradient method), seems to be first to establish connection between SOS decompositions and convexity.

SOS and SDP

❋ **SOSTOOLS** package automatically translates between SOS polynomial and its SDP representation.

SOS and SDP

❋ **SOSTOOLS** package automatically translates between SOS polynomial and its SDP representation.
Suppose $p(x) \geq 0$, then $p(x) = [x]_d^T Q [x]_d$.

SOS and SDP

❋ **SOSTOOLS** package automatically translates between SOS polynomial and its SDP representation.
Suppose $p(x) \geq 0$, then $p(x) = [x]_d^T Q[x]_d$. We need to find Q.

SOS and SDP

❋ **SOSTOOLS** package automatically translates between SOS polynomial and its SDP representation.
Suppose $p(x) \geq 0$, then $p(x) = [x]_d^T Q[x]_d$. We need to find Q.
Expanding out the product above we have

$$\sum_{j,k=0}^{d} q_{jk} x^{j+k} = \sum_{i=0}^{2d} \Big(\sum_{j+k=i} q_{jk} \Big) x^i.$$

Since $p(x) = p_n x^n + \ldots + p_1 x + p_0$. Thus, matching coeffts

SOS and SDP

❄ **SOSTOOLS** package automatically translates between SOS polynomial and its SDP representation.
Suppose $p(x) \geq 0$, then $p(x) = [x]_d^T Q[x]_d$. We need to find Q.
Expanding out the product above we have

$$\sum_{j,k=0}^{d} q_{jk} x^{j+k} = \sum_{i=0}^{2d} \Big(\sum_{j+k=i} q_{jk} \Big) x^i.$$

Since $p(x) = p_n x^n + \ldots + p_1 x + p_0$. Thus, matching coeffts

$$p_i = \sum_{j+k=i} q_{jk}, \qquad i = 0, \ldots, 2d.$$

- These are $2d+1$ linear constraints on Q
- We also have $Q \succeq 0$

SOS and SDP

❄ **SOSTOOLS** package automatically translates between SOS polynomial and its SDP representation.
Suppose $p(x) \geq 0$, then $p(x) = [x]_d^T Q[x]_d$. We need to find Q.
Expanding out the product above we have

$$\sum_{j,k=0}^{d} q_{jk} x^{j+k} = \sum_{i=0}^{2d} \Big(\sum_{j+k=i} q_{jk} \Big) x^i.$$

Since $p(x) = p_n x^n + \ldots + p_1 x + p_0$. Thus, matching coeffts

$$p_i = \sum_{j+k=i} q_{jk}, \qquad i = 0, \ldots, 2d.$$

- These are $2d+1$ linear constraints on Q
- We also have $Q \succeq 0$
- Thus, finding feasible Q is an SDP

Mini-challenge

Exercise: Prove that for $1 \leq n \leq m$, the polynomial $p(x) = \frac{1}{2}\binom{2m}{2n}(1+x)^{2m-2n} + \frac{1}{2}q(x)$ is nonnegative, where

$$q(x) = \sum_{j=n}^{m} \binom{2m}{2j}(1-x)^{2m-2j}(-4x)^{j-n}.$$

Mini-challenge

Exercise: Prove that for $1 \le n \le m$, the polynomial $p(x) = \frac{1}{2}\binom{2m}{2n}(1+x)^{2m-2n} + \frac{1}{2}q(x)$ is nonnegative, where

$$q(x) = \sum_{j=n}^{m} \binom{2m}{2j}(1-x)^{2m-2j}(-4x)^{j-n}.$$

- Other computational tricks may be more suitable?

Mini-challenge

Exercise: Prove that for $1 \leq n \leq m$, the polynomial $p(x) = \frac{1}{2}\binom{2m}{2n}(1+x)^{2m-2n} + \frac{1}{2}q(x)$ is nonnegative, where

$$q(x) = \sum_{j=n}^{m} \binom{2m}{2j}(1-x)^{2m-2j}(-4x)^{j-n}.$$

- Other computational tricks may be more suitable?

Remark: We note that testing nonnegativity of multivariate polynomials (of degree 4 or higher) is NP-Hard.

Application

$\mapsto$ Global optimization of a univariate polynomial $p(x)$

Application

- $\mapsto$ Global optimization of a univariate polynomial $p(x)$
- $\mapsto$ Instead of seeking $x^* \in \operatorname{argmin} p(x)$, first attempt to find a good lower bound on optimal value $p(x^*)$

Application

- $\mapsto$ Global optimization of a univariate polynomial $p(x)$
- $\mapsto$ Instead of seeking $x^* \in \operatorname{argmin} p(x)$, first attempt to find a good lower bound on optimal value $p(x^*)$
- $\mapsto$ A number γ is a global lower bound on $p(x)$, iff

$$p(x) \geq \gamma \quad \forall x \quad \Leftrightarrow \quad p(x) - \gamma \geq 0, \ \forall x.$$

Application

- $\mapsto$ Global optimization of a univariate polynomial $p(x)$
- $\mapsto$ Instead of seeking $x^* \in \operatorname{argmin} p(x)$, first attempt to find a good lower bound on optimal value $p(x^*)$
- $\mapsto$ A number γ is a global lower bound on $p(x)$, iff

$$p(x) \geq \gamma \quad \forall x \quad \Leftrightarrow \quad p(x) - \gamma \geq 0,\ \forall x.$$

- $\mapsto$ Now optimize to get tightest bound, so

$$\max \quad \gamma \quad \text{s.t.} \quad p(x) - \gamma \text{ is SOS.}$$

- $\mapsto$ Turn this into SDP for SOS; solve SDP to obtain γ^*

Application

- $\mapsto$ Global optimization of a univariate polynomial $p(x)$
- $\mapsto$ Instead of seeking $x^* \in \operatorname{argmin} p(x)$, first attempt to find a good lower bound on optimal value $p(x^*)$
- $\mapsto$ A number γ is a global lower bound on $p(x)$, iff

$$p(x) \geq \gamma \quad \forall x \quad \Leftrightarrow \quad p(x) - \gamma \geq 0,\ \forall x.$$

- $\mapsto$ Now optimize to get tightest bound, so

$$\max \quad \gamma \quad \text{s.t.} \quad p(x) - \gamma \text{ is SOS.}$$

- $\mapsto$ Turn this into SDP for SOS; solve SDP to obtain γ^*
- $\mapsto$ Note, optimal γ^* gives global minimum of polynomial, even though p may be highly nonconvex!

Applications

- Polynomials nonnegative only over an interval

Applications

- Polynomials nonnegative only over an interval
- Minimizing ratio of two polynomials (where $q(x) > 0$)

$$\frac{p(x)}{q(x)} \geq \gamma \quad \leftrightarrow \quad p(x) - \gamma q(x) \geq 0.$$

Applications

- Polynomials nonnegative only over an interval
- Minimizing ratio of two polynomials (where $q(x) > 0$)

$$\frac{p(x)}{q(x)} \geq \gamma \quad \leftrightarrow \quad p(x) - \gamma q(x) \geq 0.$$

- Several others (in nonlinear control, etc.)

Applications

- Polynomials nonnegative only over an interval
- Minimizing ratio of two polynomials (where $q(x) > 0$)

$$\frac{p(x)}{q(x)} \geq \gamma \quad \leftrightarrow \quad p(x) - \gamma q(x) \geq 0.$$

- Several others (in nonlinear control, etc.)
- (Lower bounds for minima of multivariate polynomials)

References

♡ P. Parrilo. *Algebraic techniques and semidefinite optimization*. MIT course, 6.256.

♡ P. Parrilo's website.

♡ G. Blekherman, P. Parrilo, R. R. Thomas. *Semidefinite optimization and convex algebraic geometry* (2012).

What we did not cover?

- See Springer Encyclopedia on Optimization (over 4500 pages!)
- Convex relaxations of nonconvex problems in greater detail
- Algorithms (trust-region methods, cutting plane techniques, bundle methods, active-set methods, and 100s of others)
- Applications of our techniques
- Software, systems ideas techniques, implementation details
- Theory: convex analysis, geometry, probability
- Noncommutative polynomial optimization (where often we might just care for just a "feasibility" test)
- Convex optimization in inf-dimensional Hilbert, Banach spaces
- Semi-infinite and infinite programming
- Multi-stage stochastic programming, chance constraints, robust optimization, tractable approximations of hard problems
- Optimization on manifolds, on matrix manifolds
- And 100s of other things!

Thanks!

Hope you learned something new!!

Thanks!

Hope you learned something new!!

Some algebra

Ideals and cones

Given a set of multivariate polynomials $\{f_1, \ldots, f_m\}$, define

$$\textbf{ideal}(f_1, \ldots, f_m) := \left\{ f \mid f = \sum\nolimits_i t_i f_i, \quad t_i \in \mathbb{R}[x] \right\}.$$

$$\textbf{cone}(f_1, \ldots, f_m) := \left\{ g \mid g = s_0 + \sum_{\{i\}} s_i f_i + \sum_{\{i,j\}} s_{ij} f_i f_j + \ldots \right\},$$

where each term is a squarefree product of f_i, with a coefficient $s_\alpha \in \mathbb{R}[x]$ that is a sum of squares.
The sum is finite, with a total of $2^m - 1$ terms, corresponding to the nonempty subsets of $\{f_1, \ldots, f_m\}$.

Algebraic connections

Note: Every polynomial in **ideal**(f_i) vanishes in the solution set of $f_i(x) = 0$.
Note: Every element of **cone**(f_i) is nonnegative on the feasible set $f_i(x) \geq 0$.

Example $Ax = b$ is infeasible $\leftrightarrow$ there exists a μ, such that $A^T\mu = 0$ and $b^T\mu = -1$.

Theorem Hilbert's Nullstellensatz: Let $f_1(z), \dots, f_m(z)$ be polynomials in complex variables $z_1, \dots, z_n$. Then,

$$f_i(z) = 0, (i = 1, \dots, m) \quad \text{is infeasible in } \mathbb{C}^n$$
$$\Leftrightarrow \quad -1 \in \textbf{ideal}(f_1, \dots, f_m).$$

Exercise: Verify the "easy" direction of the above theorems.

Semialgebraic connections

Farkas lemma and Positivstellensatz

Theorem (Farkas lemma). $Ax + b = 0$ and $Cx + d \geq 0$ is infeasible **is equivalent to**

$$\exists \lambda \geq 0, \mu \text{ s.t. } \begin{cases} A^T\mu + C^T\lambda = 0 \\ b^T\mu + d^T\lambda = -1. \end{cases}$$

Theorem (Positivstellensatz). The system $f_i(x) = 0$ for $i = 1, \ldots, m$ and $g_i(x) \geq 0$ for $i = 1, \ldots, p$ is infeasible in $\mathbb{R}^n$ **is equivalent to**

$$\exists F(x), G(x) \in \mathbb{R}[x] \text{ s.t. } \begin{cases} F(x) + G(x) = -1 \\ F(x) \in \textbf{ideal}(f_1, \ldots, f_m) \\ G(x) \in \textbf{cone}(g_1, \ldots, g_p). \end{cases}$$

What it means?

- For every infeasible system of polynomial equations and inequalities, there exists a simple algebraic identity that directly certifies non-existence of real solutions.
- Evaluation of polynomial $F(x) + G(x)$ at any feasible point should produce a nonnegative number. But this expression is identically equal to -1, a contradiction.
- Degree of $F(x)$ and $G(x)$ can be exponential.
- These cones and ideals are always convex sets (regardless of original polynomial); similar to dual function being always concave, regardless of primal.