Sum of Squares Optimization and Its Applications

Amir Ali Ahmadi

Princeton University

Dept. of Operations Research and Financial Engineering (ORFE)

ORF 363



Optimization over nonnegative polynomials

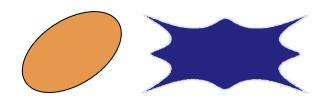
Definition by example: How to pick c_1 , c_2 , c_3 so to make

$$p(x_1, x_2) = c_1 x_1^4 - 6x_1^3 x_2 - 4x_1^3 + c_2 x_1^2 x_2^2 + 10x_1^2 + 12x_1 x_2^2 + c_3 x_2^4$$

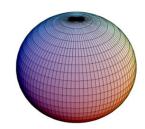
nonnegative over a given basic semialgebraic set?

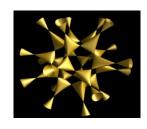
Basic semialgebraic set:
$$\{x \in \mathbb{R}^n | g_i(x) \ge 0, h_j(x) = 0\}$$

Ex:
$$x_1^3 - 2x_1x_2^4 \ge 0$$
 $x_1^4 + 3x_1x_2 - x_2^6 \ge 0$









- -This problem is fundamental to many areas of applied/computational mathematics.
- -It is the problem that "SOS optimization" is designed to solve.

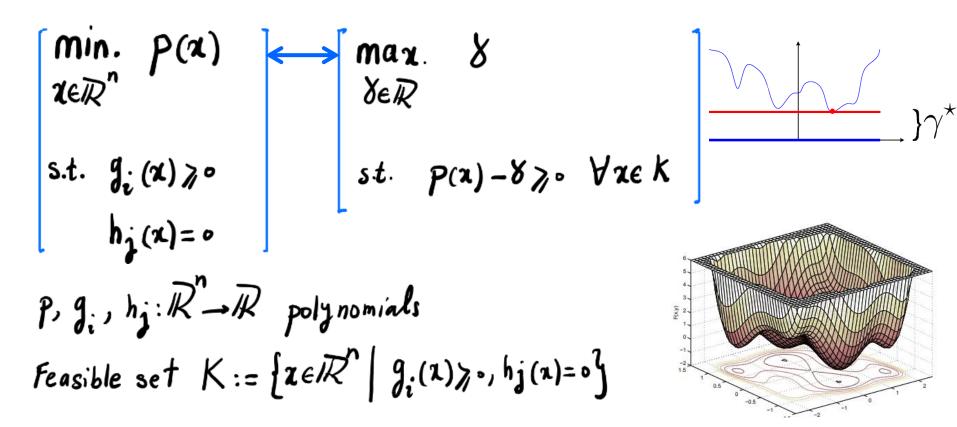


Why would you want to do this?!

Let's start with five application domains...



1. Polynomial optimization



- ■Many applications: finding equilibria in games, the optimal power flow problem, low-rank matrix factorization, dictionary learning, training of deep nets with polynomial activation function, sparse regression with nonconvex regularizes, etc.
- •Intractable in general (includes all NP-complete problem)

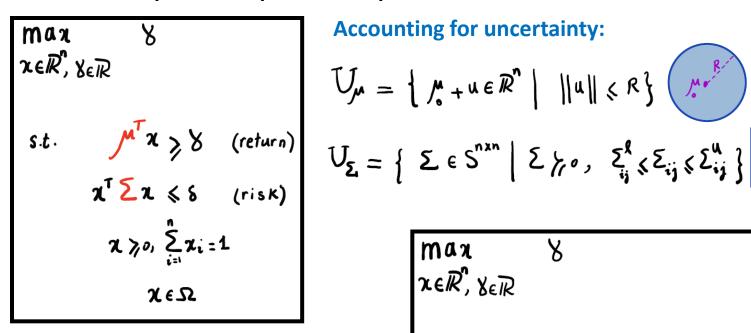


2. Optimization under input uncertainty

How to make optimal decisions when input to optimization problem is uncertain/noisy?

Example: The Markowitz portfolio optimization problem

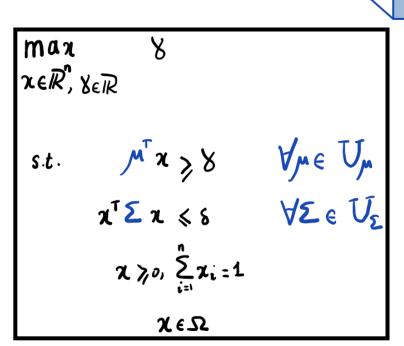




Accounting for uncertainty:

$$U_{m} = \{ \mathcal{N}_{+} u \in \mathbb{R}^{n} \mid ||u|| \leq R \}$$

$$U_{\Sigma} = \left\{ \sum_{i} S^{nxn} \mid \sum_{i,j} \sigma_{i,j} \sum_{i,j} \left(\sum_{i,j} \left(\sum_{i,j} \sum_{i,j} \left(\sum_{i,j} \sum_{i,j} \left(\sum_{i,j} \sum_{i,$$



 $\mu \in \mathbb{R}^n$: mean vector $\Sigma \in \mathbb{S}^{n \times n}$: covariance of the returns matrix of the returns

In practice estimated from past data/ML model. Optimal portfolio sensitive to this input.



3. Statistics and machine learning

Shape-constrained regression; e.g., monotone and/or convex regression

Shape constraints act as regularizer, improve test performance, make model more interpretable and trustworthy

Example 1: Shape constraints natural in most applications

Zestimate 2 \$514,690

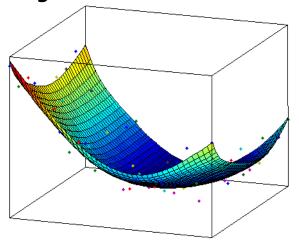




5 beds · 4 baths · 2,623 sqft







Monotonicity of a polynomial $p(x_1, ..., x_n)$ with respect to feature $j: \frac{\partial p(x)}{\partial x_i} \ge 0, \forall x \in B$

Monotonicity of a polynomial
$$p(x_1, ..., x_n)$$
 with respect to feature $j: \frac{1 + x_j}{\partial x_j} \ge 0, \forall x \in \mathbb{R}$
Example 2: "ML for fast real-time convex optimization"

$$g(b) \coloneqq \min_{x \in \mathbb{R}^n} f_0(x)$$
s.t. $f_i(x) \le b_i \ i = 1, ..., m$
 $x \in \Omega$

 f_0, \dots, f_m convex functions, Ω a convex set.

Goal: learn g(b) offline from training set; evaluate it online very fast

$$g: \mathbb{R}^m \to \mathbb{R}$$
 is

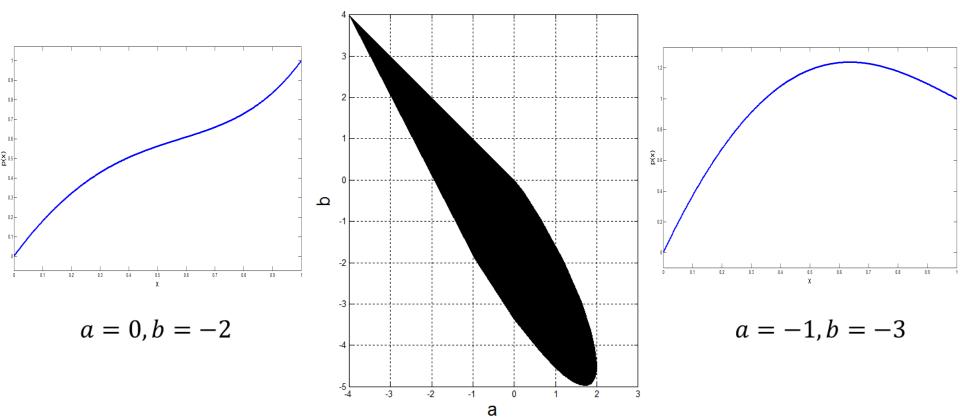
- convex
- nonincreasing w.r.t. all arguments



Imposing monotonicity

• For what values of a, b is the following polynomial monotone over [0,1]?

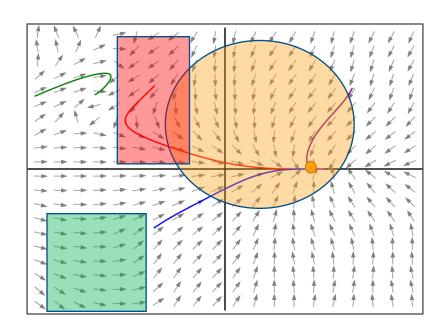
$$p(x) = x^4 + ax^3 + bx^2 - (a+b)x$$





4. Certifying properties of dynamical systems

$$\dot{x} = f(x)$$



Questions about properties of dynamical systems (e.g., stability, safety) Lyapunov theory

Search for functions satisfying nonnegativity constraints Semialgebraic parametrization

Polynomial inequalities



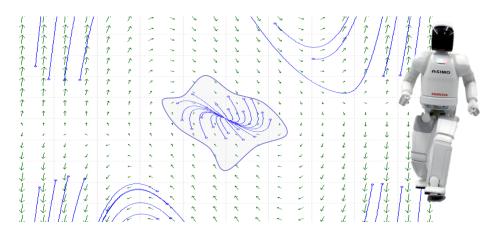


Example: certifying stability

$$\dot{x} = f(x) \quad f: \mathbb{R}^n \to \mathbb{R}^n$$

Ex.
$$\dot{x}_1 = -x_2 + \frac{3}{2}x_1^2 - \frac{1}{2}x_1^3x_2$$
 $\dot{x}_2 = 3x_1 - x_1x_2$

Locally asymptotic stability (LAS) of equilibrium points



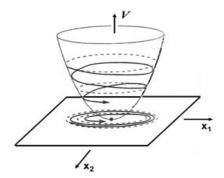
Lyapunov's theorem (and its converse):

The origin is LAS if and only if there exists a C^1 function $V: \mathbb{R}^n \to \mathbb{R}$ that vanishes at the origin and a scalar $\beta > 0$ such that

$$V(x) > 0$$

 $V(x) \le \beta \Rightarrow \dot{V}(x) = \nabla V(x)^T f(x) < 0$









(If $\dot{V}(x) < 0$ everywhere, then globally stable.)

Example: certifying collision avoidance

$$\dot{x} = f(x)$$

(vector valued polynomial)

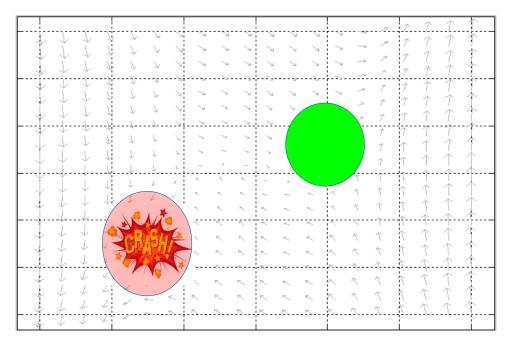


signal in the state of the stat



unsafe (or forbidden) set

(both sets basic semialgebraic)



Safety guaranteed if we find a "Lyapunov function" such that:

$$B(\mathcal{S}) < 0$$

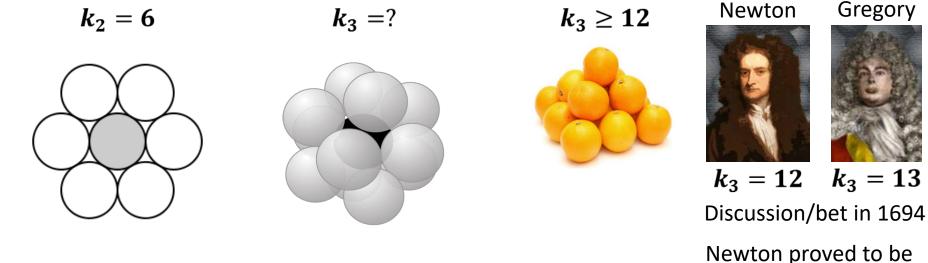
$$B(\mathcal{U}) > 0$$

$$\dot{B} = \langle \nabla B(x), f(x) \rangle \le 0$$

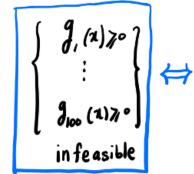


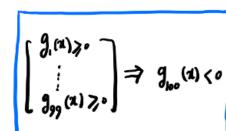
5. Automated theorem proving in geometry

Kissing number in dimension n: largest number of n-dimensional non-overlapping spheres that can simultaneously touch (or "kiss") a common unit sphere.



13 spheres impossible iff the following system is *infeasible*:





correct in 1953!



Outline

Global nonnegativity

- Sum of squares (SOS) and semidefinite programming
- Two applications
- Hilbert's 17th problem

Nonnegativity over a region

- Putinar's Positivstellensat
- Two applications
- Recap and further reading



How would you prove nonnegativity?

Ex. Decide if the following polynomial is nonnegative:

$$p(x) = x_1^4 - 6x_1^3x_2 + 2x_1^3x_3 + 6x_1^2x_3^2 + 9x_1^2x_2^2 - 6x_1^2x_2x_3 - 14x_1x_2x_3^2 + 4x_1x_3^3 + 5x_3^4 - 7x_2^2x_3^2 + 16x_2^4$$

- Not so easy! (In fact, NP-hard for degree ≥ 4)
- ■But what if I told you:

$$p(x) = (x_1^2 - 3x_1x_2 + x_1x_3 + 2x_3^2)^2 + (x_1x_3 - x_2x_3)^2 + (4x_2^2 - x_3^2)^2.$$

Natural questions:

- •Q1: Is it any easier to test for a sum of squares (SOS) decomposition?
- •Q2: Is every nonnegative polynomial SOS?

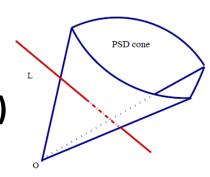


Sum of squares and semidefinite programming

[Lasserre], [Nesterov], [Parrilo]

Q1: Is it any easier to decide SOS?

■Yes! Can be reduced to a semidefinite program (SDP)



- Can also efficiently search and optimize over SOS polynomials
- As we will see, this latter property is very important in applications...

Semidefinite programming (SDP)

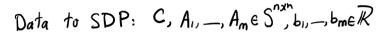
- A broad generalization of linear programs
- Can be solved to arbitrary accuracy in polynomial time (e.g., using interior point algorithms) [Nesterov, Nemirovski], [Alizadeh]

min. Tr
$$(CX)$$
XeS^{n×n}
st. Tr $(AiX)=bi$ $i=1,...,m$

$$X > 0$$
L'psd"

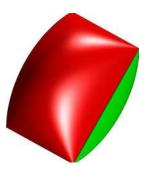
Notes:
$$Tr(CX) = \sum_{i,j} C_{ij} X_{ij}$$

Eigenvalues of X are >0.

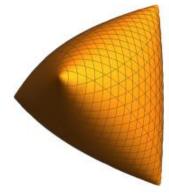


Feasible set called a "spectrahedron":











SOS→SDP

Thm:

A polynomial p of degree 2d is SOS if and only if $\exists Q \geqslant 0$ such that

$$p(x) = z(x)^T Q z(x)$$

where $z = \begin{bmatrix} 1, x_1, \dots, x_n, x_1 x_2, \dots, x_n^d \end{bmatrix}^T$ is the vector of monomials of degree up to d.

(It follows that checking membership or optimizing a linear function over the set of SOS polynomials is an SDP)

$$Q_{\gamma_0} \Rightarrow Q = V^T V \Rightarrow p(x) = Z^T(x) V^T V Z(x) = \left\| V Z(x) \right\|^2 = \sum_{i=1}^{r} \left(U_i^T Z(x) \right)^2.$$

$$\exists \, \sigma_{i,i} - \sigma_{i} \in \mathbb{R}^{\binom{n+d}{d}} \text{ s.t. } p(x) = \sum_{i=1}^{r} \left(\sigma_{i}^{T} z(x)\right)^{2} = \sum_{i=1}^{r} \left(z^{T}(x) \, \sigma_{i}^{T}\right) \left(\sigma_{i}^{T} z(x)\right) = z^{T}(x) \left(\sum_{i=1}^{r} \sigma_{i}^{T} \sigma_{i}^{T}\right) z(x).$$



Example

$$P(x) = 10 x^4 - 2 x^3 - 7 x^2 + 4x + 4$$

Is p SOS?

$$P(x) = \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ z^2 \end{bmatrix}$$

$$\forall x$$

SDP feasibility problem

Find Q yo s.t.

$$q_{33} = 10$$
, $2q_{23} = -2$
 $q_{22} + 2q_{13} = -7$
 $2q_{12} = 4$, $q_{11} = 4$

$$Q = \begin{bmatrix} 4 & 2 & -6 \\ 2 & 5 & -1 \\ -6 & -1 & 10 \end{bmatrix}$$

$$Q = \begin{bmatrix} 4 & 2 & -6 \\ 2 & 5 & -1 \\ -6 & -1 & 10 \end{bmatrix} \stackrel{eg.}{=} V^{T} V$$

$$V^{T} V$$

$$\Rightarrow p(x) = Z^{T}(x) V^{T} V + 2(x) = \left\| V + 2(x) \right\|^{2} = \left\| \begin{bmatrix} 0 & 2 & 1 \\ 2 & 1 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^{2} \end{bmatrix} \right\|^{2}$$



$$P(x) = (2x + x^2)^2 + (2 + x - 3x^2)^2$$

Let's revisit two of our applications!



Optimization over nonnegative polynomials



Sum of squares (SOS) programming



Semidefinite programming (SDP)





1) Nonconvex unconstrained minimization

Find:
$$p! = \inf \left(\frac{4x^2 - \frac{21}{10}x^4 + \frac{1}{3}x^6 + xy - 4y^2 + 4y^4 + x^2y}{(x,y)eR^2} \right)$$

$$P_{sos} := \sup_{\delta eR} \delta$$

$$s.t. \quad P(x,y) - \delta \quad SOS$$

$$P_{sos} \leq p^*$$

```
solvertime: 0.6 (s)
p=4*x^2-2.1*x^4+(1/3)*x^6+1*x*y-4*y^2+4*y^4+x^2*y;
solvesos (sos (p-gam), -gam, [], [gam])
                                                              p sos =
p sos=double(gam)
                                                                -2.921560950963582
[inf,z,Q]=solvesos(p-p sos);
sdisplay(z{1})
                                                             p at xstar =
                                       xstar =
[v,d]=eig(double(Q{1}));
```

zxstar=v(:,1)/v(1,1);xstar=[zxstar(3);zxstar(2)] p_at_xstar=replace(p,[x,y],[xstar(1),xstar(2)])

1.832996144755618

-0.922931478421273

-2.921559422066406

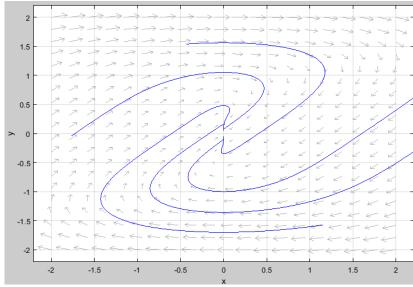
2) Automated proof of global asymptotic stability



To: Amir Ali Ahmadi

Hi Amir Ali,

I hope life and career are going well.



I have a question that I assume might take little more than 5-10 of your time but please feel free to let me know if it would actually take more.

Today in class we got into an interesting discussion with students about what a strict Lyapunov function would be for the system

.edu> on behalf of

$$dx/dt = -x + y^3$$
$$dy/dt = -x$$

A non-strict L.f.. is easy, $V = x^2/2 + y^4/4$, with $dV/dt = -x^2$. One could then deduce g.a.s. by a Barbashin-Krasovskii/Lasalle argument, but that's not satisfactory.

I started constructing a strict one in real time and it quickly got out of hand, necessitating higher and higher powers and many cross terms. I inevitably thought of you and your (and Pablo's) SOS program that would spit out a good strict V within seconds.

If you can plug in this system and let me know what comes out, I'd appreciate it, and my 40-50 students in class would learn a few things (complexity of Lyapunov functions, automated options for finding them, etc.).

Best regards,



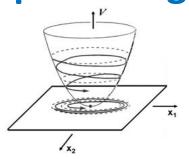
Automated proof of global asymptotic stability

$$\begin{cases} \dot{x} = -x + y^3 \\ \dot{y} = -x \end{cases}$$

[V,c,m]=polynomial([x;y],4,2);

sdpvar x y
xdot=-x+y^3;

ydot=-x;



>> sdisplay(clean(double(c)'*m,1e-3))

Find
$$V(x,y)$$
 of degree 4 s.t.

$$V(x,y) = -\langle \nabla V(x,y), (-x+y^3) \rangle = -\langle \nabla V(x,y), (-x+y^3) \rangle = -\langle \nabla V(x,y), (-x+y^3) \rangle$$

```
Vdot=jacobian(V,[x,y])*[xdot;ydot];

FF=[sos(V),sos(-Vdot)]
solvesos(FF,[],[],[c])
```

$$V(x,y) = x^2 - \frac{1}{3}xy + \frac{1}{6}y^2 + \frac{1}{2}y^4$$

1.00000084865*x^2-0.333330248293*x*y+0.166665124147*y^2+0.500118639025*y^4

$$= \left(\chi - \frac{1}{6}y\right)^2 + \frac{5}{36}y^2 + \frac{1}{2}y^4 \qquad \text{(hence positive definite)}$$

$$\frac{1}{36} (\chi^2 + y^2) + \frac{1}{2} y^4$$
 (hence radially unbounded)

$$\sqrt{(x,y)} = -\frac{5}{3} x^2 - \frac{1}{3} y^4$$
 (hence negative definite)

Hilbert's 1888 Paper

Q2: SOS $\stackrel{?}{\Leftarrow}$ Nonnegativity

n,d	2	4	≥6
1	yes	yes	yes
2	yes	yes	no
3	yes	no	no
≥4	yes	no	no



Motzkin (1967):

$$M(x_1,x_2) = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 + 1$$

Robinson (1973):

$$R(x_1, x_2, x_3) = x_1^2(x_1 - 1)^2 + x_2^2(x_2 - 1)^2 + x_3^2(x_3 - 1)^2$$



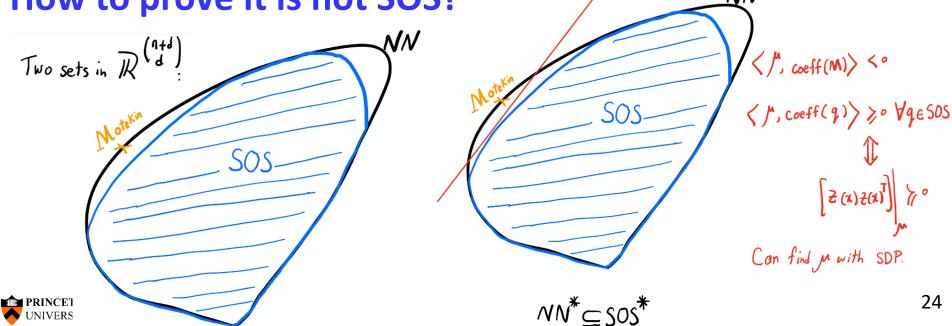
The Motzkin polynomial

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

How to prove it is nonnegative?

$$(x^{2} + y^{2} + 1) M(x, y) = (x^{2}y - y)^{2} + (xy^{2} - x)^{2} + (x^{2}y^{2} - 1)^{2} + \frac{1}{4}(xy^{3} - x^{3}y)^{2} + \frac{3}{4}(xy^{3} + x^{3}y - 2xy)^{2}$$

How to prove it is not SOS?



Hilbert's 17th Problem (1900)

Q.
$$p$$
 nonnegative $\stackrel{?}{\Rightarrow}$ $p = \sum_{i} \left(\frac{g_i}{q_i}\right)^2$

- Artin (1927): Yes!
- Implications:
 - $p \ge 0 \Rightarrow \exists h \text{ sos } \text{such that } p.h \text{ sos }$
 - **Reznick:** (under mild conditions) can take $h = (\sum_i x_i^2)^r$
 - Certificates of nonnegativity can always be given with sos (i.e., with semidefinite programming)!
 - We'll see how the Positivstellensatz generalizes this even further

Outline of the rest of the talk...

Global nonnegativity

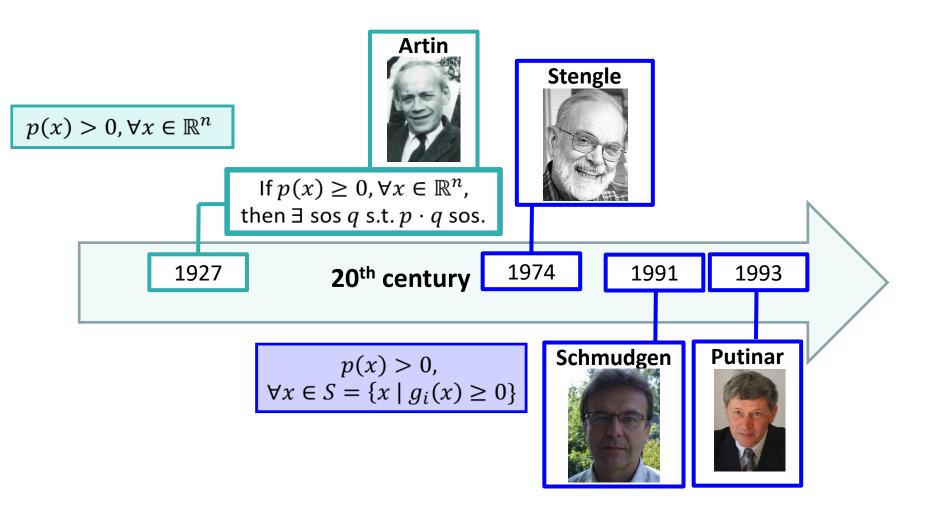
- Sum of squares (SOS) and semidefinite programming
- Two applications
- Hilbert's 17th problem

Nonnegativity over a region

- Putinar's Positivstellensatz
- Two applications
- Recap and further reading



Positivstellensatz





Putinar's Positivstellensatz (1993)

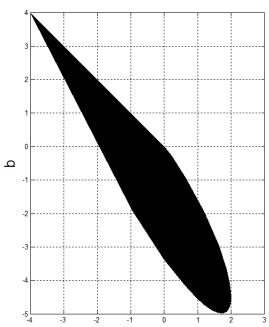
- Search for SOS multipliers can be automated via SDP!
- This means that SDP-based proofs of implications between polynomial inequalities always possible!
- Degree bounds on SOS multipliers available in special cases.



How did I plot this?

• For what values of a, b is the following polynomial monotone over [0,1]?

$$p(x) = x^4 + ax^3 + bx^2 - (a+b)x$$



Theorem. A polynomial p(x) of degree 2d is monotone on [0,1] if and only if

$$p'(x) = xs_1(x) + (1 - x)s_2(x),$$

where $s_1(x)$ and $s_2(x)$ are some SOS polynomials of degree 2d-2.



Let's end with a couple applications:

- Finance
- Control

Optimization over nonnegative polynomials



Sum of squares (SOS) programming

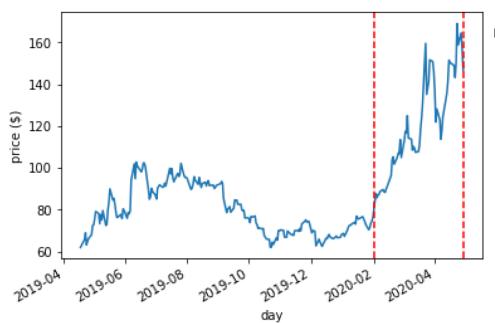


Semidefinite programming (SDP)



Distributionally robust optimization

What's the probability that Zoom's stock goes bust?



■Three months starting Feb 1, 2020

$$r_i = \frac{P_i - P_{i-1}}{P_i}, \quad i = 1, \dots, 61$$

•Empirical moments $m_k = \mathbb{E}[r^k]$:

$$m_1 = 0.0068, m_2 = 0.0034,$$

 $m_3 = 2 \times 10^{-6}, m_4 = 5 \times 10^{-5}$

- ■The distribution of r is supported on [-0.4,0.4] but is otherwise unknown
- ■What is the probability that Zoom's stock return will be below -0.1 today?
- ■Want the worst-case probability over all distributions whose first 4 moments are within 10% of those computed from data.



Sum of squares optimization can compute this probability!

$$\alpha := \inf_{q,r,s,\gamma} \quad \gamma$$
s.t. $q(x) = \sum_{k=0}^{4} q_k x^k$ is a degree-4 (univariate) polynomial,
$$r(x), s(x) \text{ are quadratic polynomials that are sos,}$$

$$q_0 + \sum_{k=1}^{4} q_k m_k' \le \gamma \ \forall m_k' \in [0.9 \ m_k, 1.1 \ m_k] \text{ for } k = 1, \dots, 4,$$

$$q(x) - (0.4^2 - x^2) \ s(x) \text{ is sos,}$$

$$q(x) - 1 - (0.4 + x)(-0.1 - x)r(x) \text{ is sos.}$$

$$\Rightarrow q(x) \ge 0 \quad \forall x \in [-0.4, 0.4]$$

$$\mathbb{P}(r \in [-0.4, -0.1]) = \mathbb{E}[1_{[-0.4, -0.1]}]$$

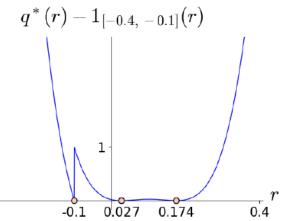
-0.4

$$\Rightarrow 1_{[-0.4,-0.1]} \le q(x) \; \forall x \in [-0.4,0.4]$$

$$\Rightarrow \mathbb{E}\left[1_{[-0.4,-0.1]}\right] \leq \mathbb{E}\left[q(x)\right] = \sum_{k=0}^{\infty} q_k m_k \leq \gamma$$

In fact, we always have

$$\mathbb{P}(r \in [-0.4, -0.1]) = \alpha$$

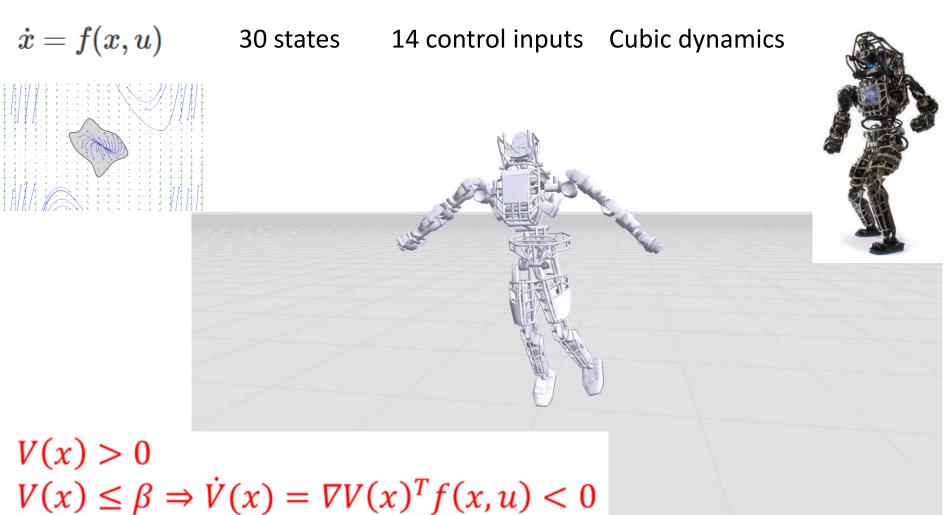


 $\mathbb{P}(r \in [-0.4, -0.1]) \le \alpha$

Optimizer terminated. Time: 0.17
alpha = 0.2073



Stabilizing a humanoid robot on one foot







Certifying collision avoidance

$$\dot{x} = f(x)$$

(vector valued polynomial)

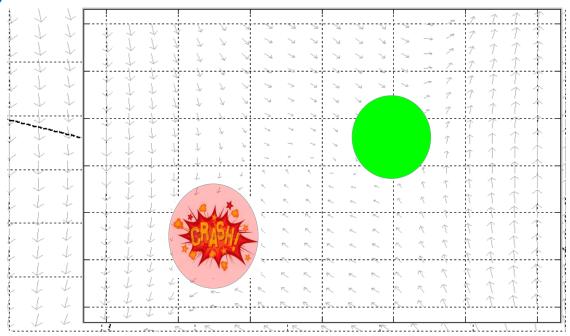


ineeds safety verification



unsafe (or forbidden) set

(both sets basic semialgebraic)



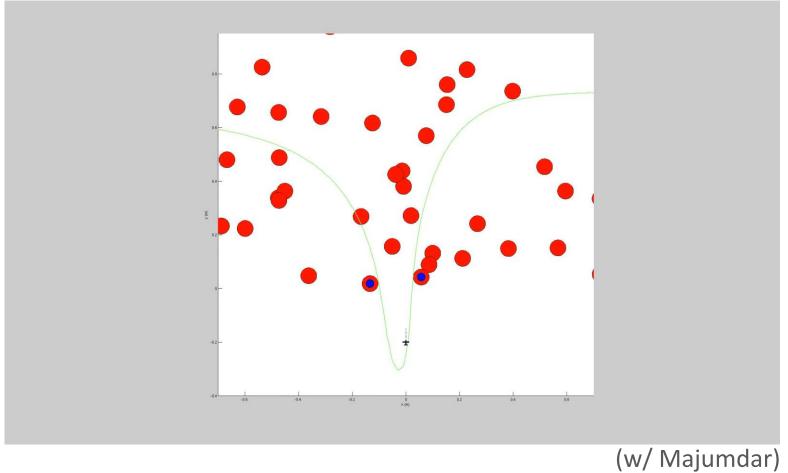
Safety guaranteed if we find a "Lyapunov function" such that:

$$B(\mathcal{S}) < 0$$

$$B(\mathcal{U}) > 0$$

$$\dot{B} = \langle \nabla B(x), f(x) \rangle \le 0$$

Real-time collision avoidance certificates



(vv) rviajarriaa

Dubins car model

Run-time: 20 ms

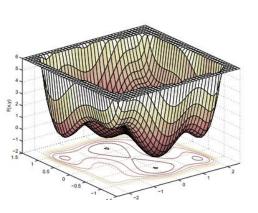


Recap: "See an inequality? Think SOS!"

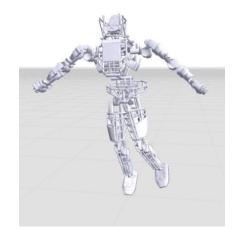
Is
$$p(x) \ge 0$$
 on $\{g_1(x) \ge 0, ..., g_m(x) \ge 0\}$?

Automated SOS-based proofs via SDP!

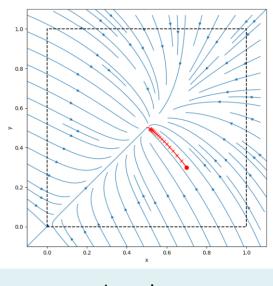
Many applications!



Optimization

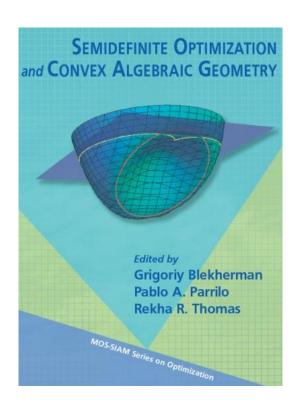


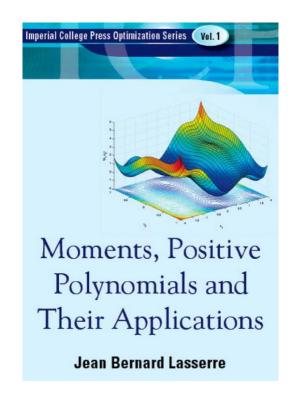
Control



Learning

Want to learn more?





SUMS OF SQUARES, MOMENT MATRICES AND OPTIMIZATION OVER POLYNOMIALS

MONIQUE LAURENT*

Applications of sums of squares

Georgina Hall

