Optimization over Nonnegative Polynomials: Algorithms and Applications

Amir Ali Ahmadi
Princeton, ORFE

INFORMS Optimization Society Conference (Tutorial Talk)
Princeton University
March 17, 2016

Optimization over nonnegative polynomials

Defn. A polynomial $p(x) := p(x_1, \dots, x_n)$ is nonnegative if $p(x) \geq 0, \forall x \in \mathbb{R}^n$.

Ex. Decide if the following polynomial is nonnegative:

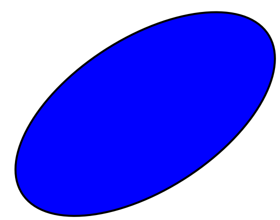
$$\begin{aligned} p(x) \;=\; & x_1^4 - 6x_1^3x_2 + 2x_1^3x_3 + 6x_1^2x_3^2 + 9x_1^2x_2^2 - 6x_1^2x_2x_3 \\ & -14x_1x_2x_3^2 + 4x_1x_3^3 + 5x_3^4 - 7x_2^2x_3^2 + 16x_2^4 \end{aligned}$$

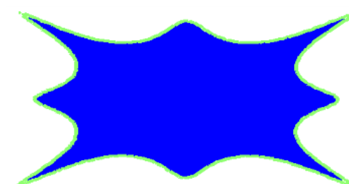
Basic semialgebraic set:

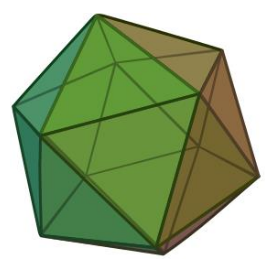
$$\{x \in \mathbb{R}^n |\ f_i(x) \geq 0, h_i(x) = 0\}$$

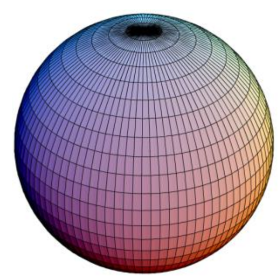
Ex. $2x_1 + 5x_1^2x_2 - x_3 \geq 0$

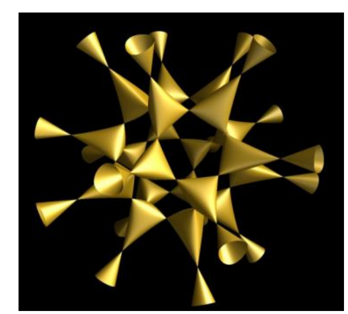
$5 - x_1^3 + 2x_1x_3 = 0$











Ubiquitous in computational mathematics!

Why would you want to do this?!

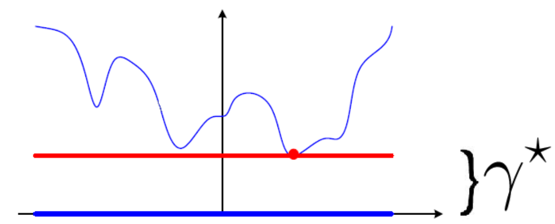
- Let’s start with three broad application areas…

1. Polynomial Optimization

$$\min_{x} p(x)$$
$$f_i(x) \leq 0$$
$$h_i(x) = 0$$

Equivalent formulation:

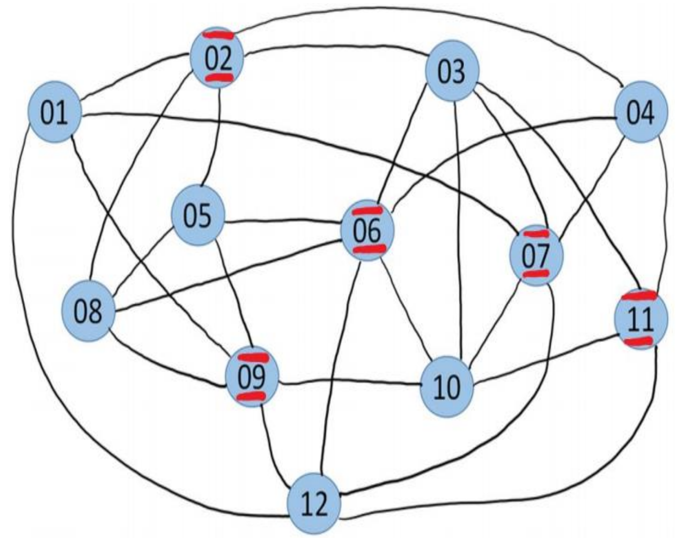
$$\max_{\gamma} \gamma$$
$$p(x) - \gamma \geq 0$$
$$\forall x \in \{f_i(x) \leq 0,\ h_i(x) = 0\}$$



- **Many applications:**
- Combinatorial optimization [next slide]
- Game theory [next slide]
- Option pricing with moment information [Merton], [Boyle], [Lo], [Bertsimas et al.],…
- The optimal power flow (OPF) problem [Lavaei, Low et al.], [Bienstock et al.]…
- Sensor network localization [Ye et al.],…
- Low-rank matrix completion, nonnegative matrix factorization, …

1.1. Optimality certificates in discrete optimization

- Finding the **independent set number** of a graph
- $5 \leq \alpha(G)$
- How to prove you cannot do better?



- One can show:

$$\alpha(G) \leq k$$

if and only if

$$-2k \sum_{(i,j)\in\bar{E}} x_i x_j y_i y_j - (1-k)\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right)$$

is nonnegative.

- Similar algebraic formulations for other combinatorial optimization problems…

1.1. Infeasibility certificates in discrete optimization

- **PARTITION**

- **Input:** A list of positive integers $a_1, \dots, a_n$.

- **Question:** Can you split them into to bags such that the sum in one equals the sum in the other?

$\{5,2,1,6,3,8,5,4,1,1,10\}$

- Note that the YES answer is easy to certify.

- How would you certify a NO answer?

1.1. Infeasibility certificates in discrete optimization

- **PARTITION**

- **Input:** A list of positive integers $a_1, \ldots, a_n$.

- **Question:** Can you split them into to bags such that the sum in one equals the sum in the other?

$$[a_1, a_2, \ldots, a_n]$$

$$\exists ?\ x_i \in \{-1, 1\} \text{ s.t. } \sum_{i=1}^{n} x_i a_i = 0$$

Infeasible iff $$\sum_{i=1}^{n} (x_i^2 - 1)^2 + (\sum_{i=1}^{n} x_i a_i)^2 > 0$$

1.2 Payoff bounds on Nash equilibria

A =

4	11	-5	6
9	-5	10	8
2	2	4	2
8	-7	0	9

B =

4	-3	5	5
-17	7	-11	0
-3	6	-1	-1
9	4	11	0

$$\Delta_n := \{x \mid x \geq 0, \sum_{i=1}^{n} x_i = 1\} \qquad (x,y) \text{ Nash if } \begin{cases} x^T A y \geq z^T A y & \forall z \in \Delta_n \\ x^T B y \geq x^T B z & \forall z \in \Delta_n \end{cases}$$

Upper bounds on player 1&2's payoffs under any Nash equilibrium:

$$\min \{\gamma : x^T A y \leq \gamma \text{ on } K\}, \quad \min\{\gamma : x^T B y \leq \gamma \text{ on } K\}$$

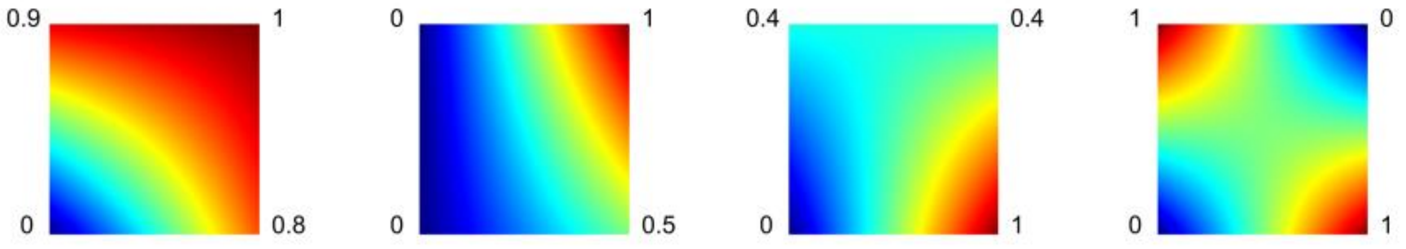
$$K := \{(x,y) \mid \sum x_i = 1, \ x \geq 0, \ \sum y_i = 1, \ y \geq 0, \ x^T A y \geq e_i^T A y, \ i=1,\dots,n, \ x^T B y \geq x^T B e_i, \ i=1,\dots,n\}$$

Set of Nash equilibria

See Session TC02 (Bowl 1, 1:30 PM)

2: Statistics and Machine Learning

- **Shape-constrained regression; e.g., *monotone regression***



Monotonic (a) | Monotonic (b) | Not Monotonic (c) | Not Monotonic (d)

(From [Gupta et al.,'15])

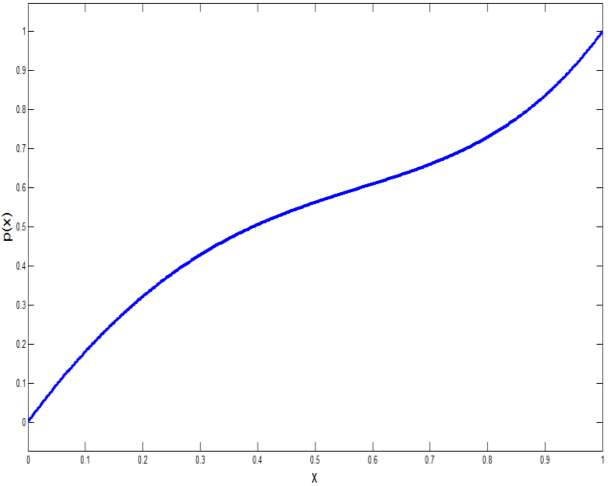
- How to parameterize a polynomial $p(x_1, x_2)$ to enforce monotonicity over $[0,1]^2$?

- Need its partial derivatives to be nonnegative over $[0,1]^2$.
- Let's see a simple example in one variable…

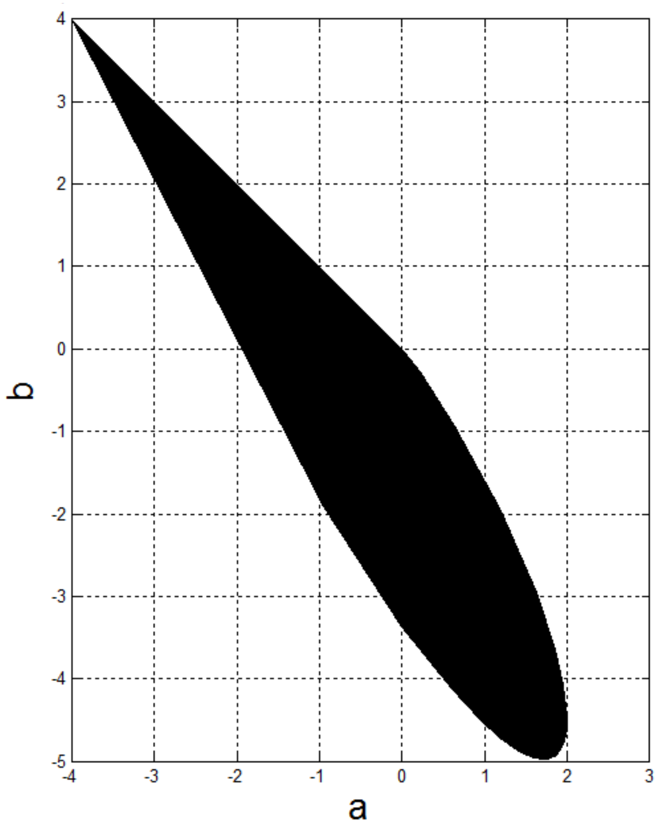
Imposing monotonicity

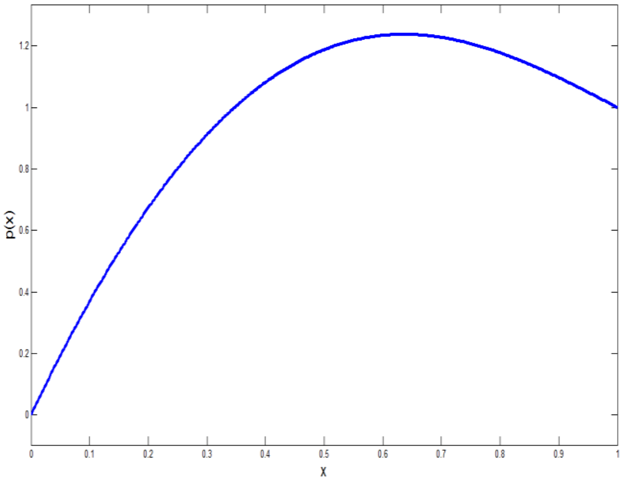
- For what values of a, b is the following polynomial monotone over $[0,1]$?

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



$a = 0, b = -2$

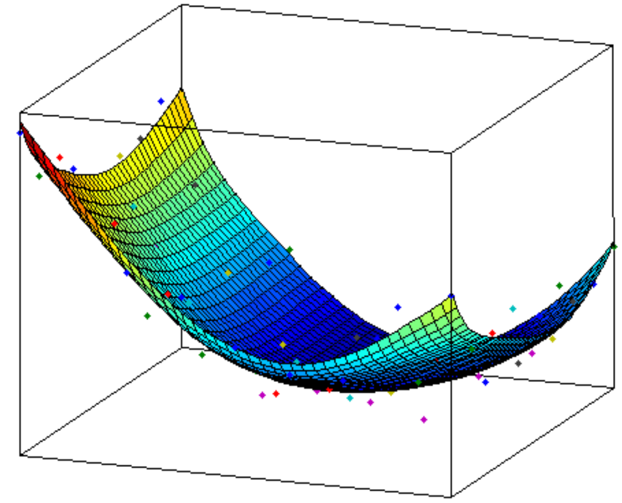




$a = -1, b = -3$

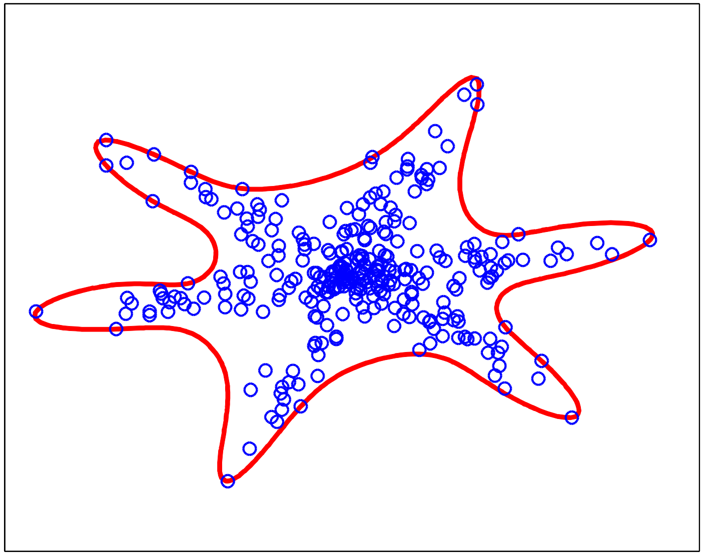
2: Statistics and Machine Learning (Ctnd.)

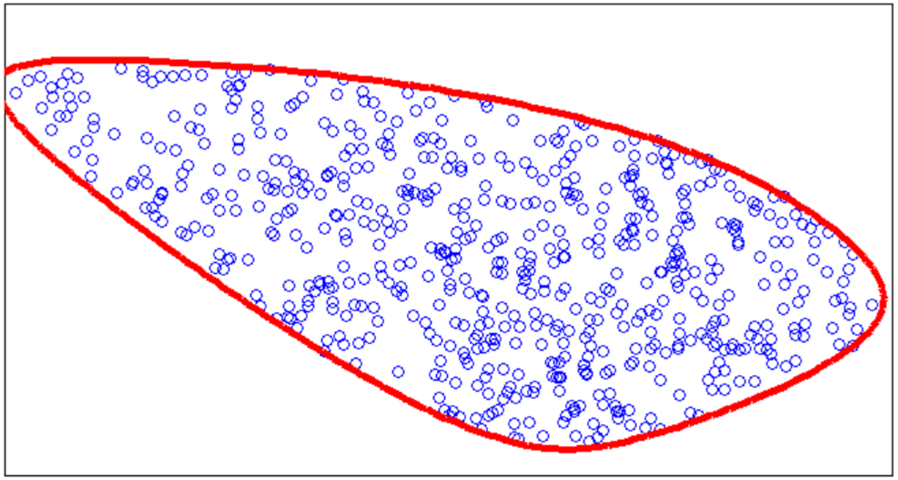
- **Convex regression**



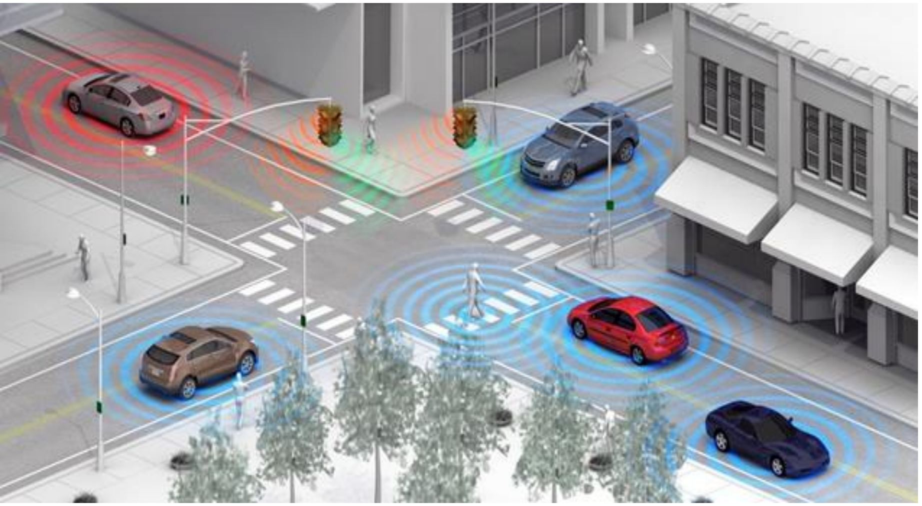
$$p(x) \text{ convex} \Leftrightarrow y^T \nabla^2 p(x) y \text{ nonnegative}$$

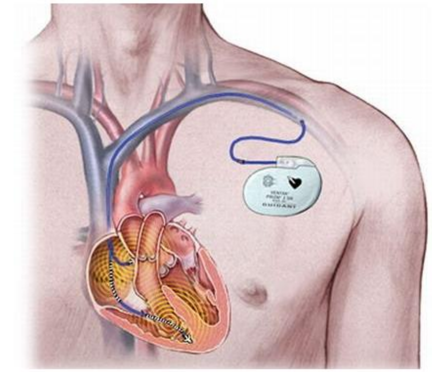
- **Clustering with semialgebraic sets**

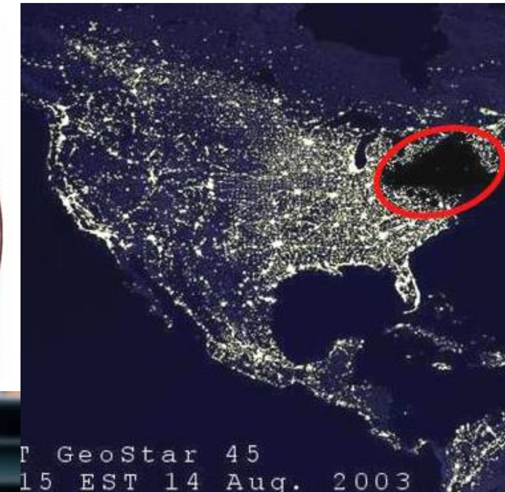


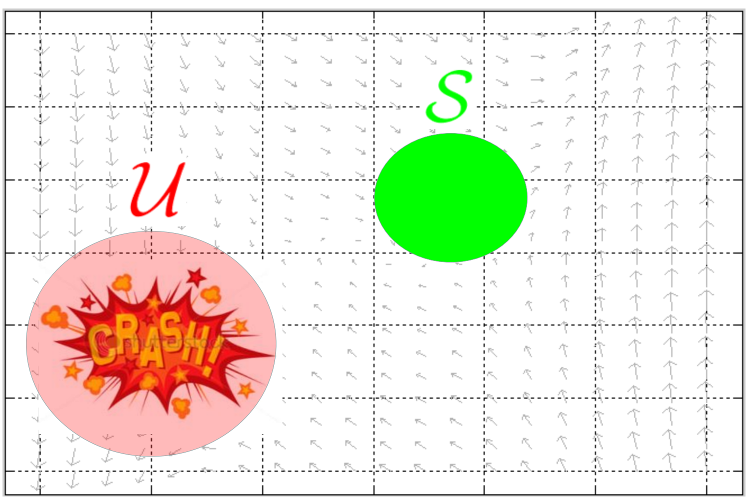


3: Safety verification of dynamical systems









$\mathcal{S}$: set that requires safety verification (where system currently operating)

$\mathcal{U}$: unsafe (or forbidden) set e.g., physical collision, variable overflow, blackout, financial crisis, …

Goal: "formal certificates" of safety

3.1: Lyapunov Barrier Certificates

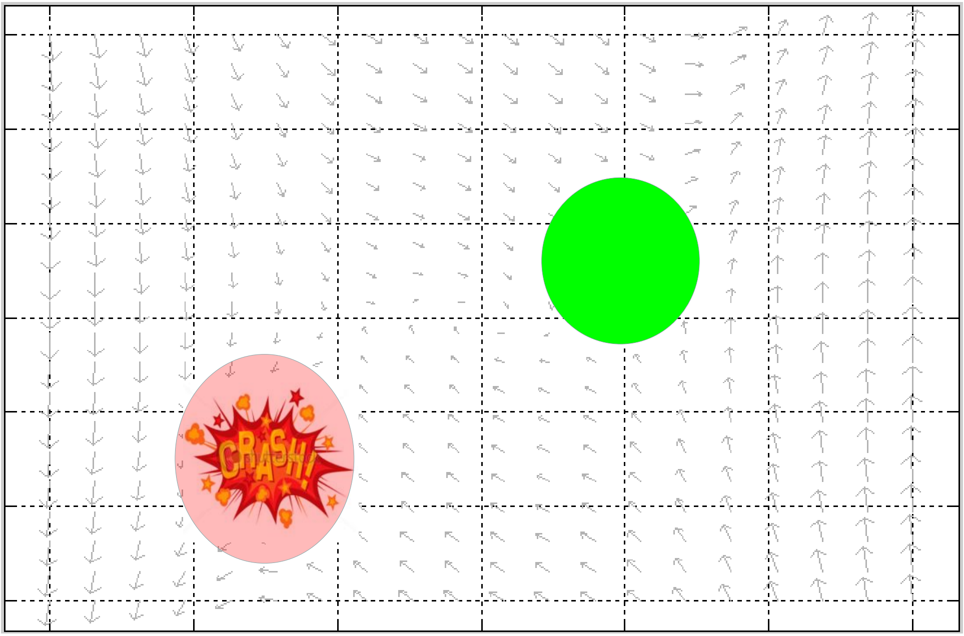
$$\dot{x} = f(x)$$

(vector valued polynomial)

$\mathcal{S}$: needs safety verification

$\mathcal{U}$: unsafe (or forbidden) set

(both sets semialgebraic)

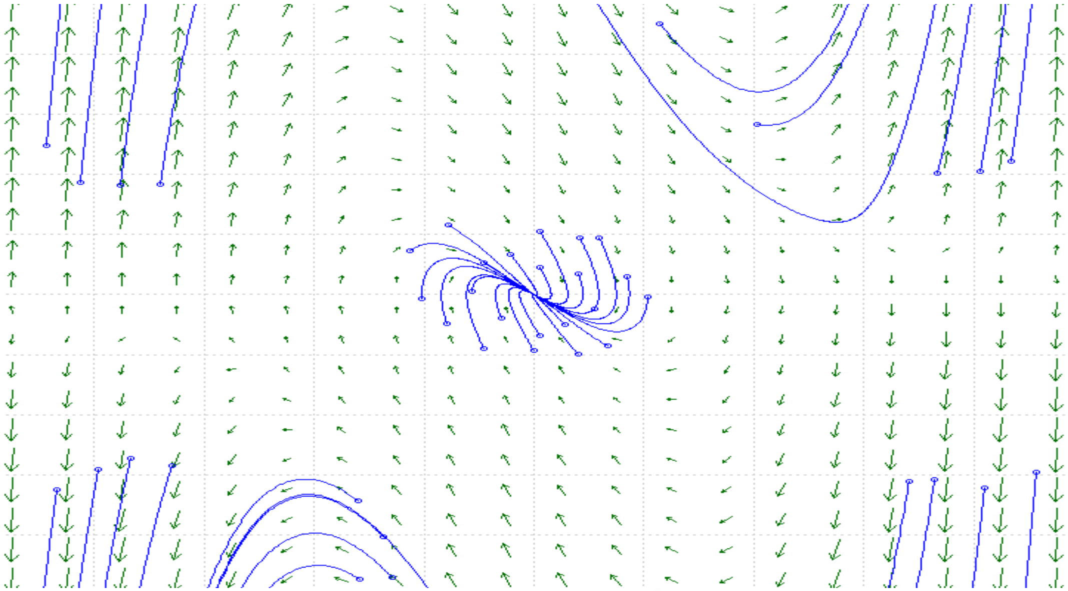


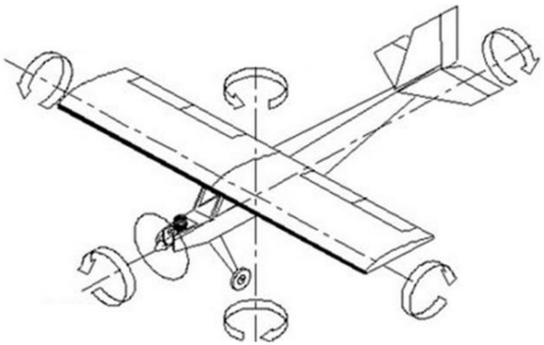
Safety assured if we find a "Lyapunov function" such that:

$$B(\mathcal{S}) < 0 \qquad \dot{B} = \langle \nabla B(x), f(x) \rangle \leq 0$$
$$B(\mathcal{U}) > 0$$

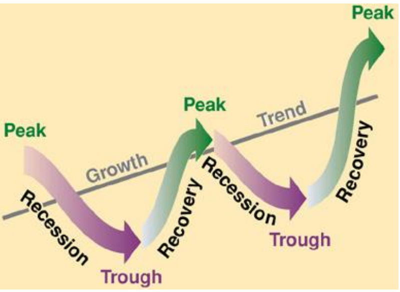
3.2: Stability of dynamical systems

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

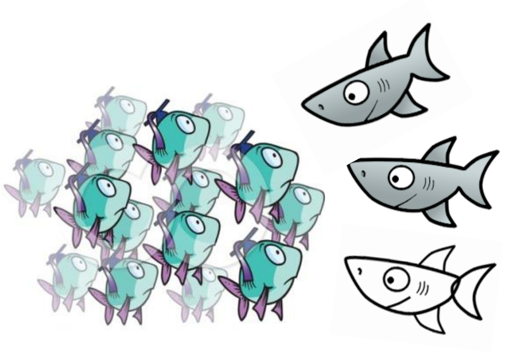




Control



Dynamics of prices



Equilibrium populations

Spread of epidemics

Robotics

Lyapunov's theorem for asymptotic stability

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

Existence of a (Lyapunov) function

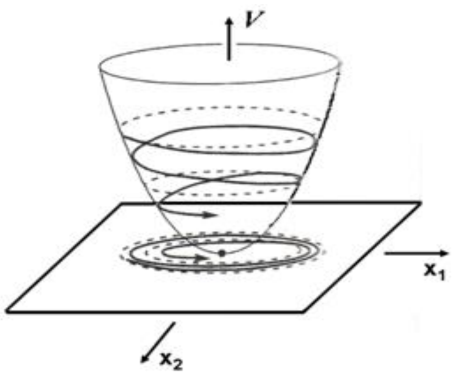
$$V(x) : \mathbb{R}^n \to \mathbb{R}$$

$$\dot{V}(x) = \langle \frac{\partial V}{\partial x}, f(x) \rangle$$

such that

$$V(x) > 0,$$

$$V(x) \leq \beta \Rightarrow \dot{V}(x) < 0$$



implies $\{x | \, V(x) \leq \beta\}$ is in the region of attraction (ROA).

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

How would you prove nonnegativity?

Ex. Decide if the following polynomial is nonnegative:

$$\begin{aligned} p(x) &= x_1^4 - 6x_1^3x_2 + 2x_1^3x_3 + 6x_1^2x_3^2 + 9x_1^2x_2^2 - 6x_1^2x_2x_3 \\ &\quad -14x_1x_2x_3^2 + 4x_1x_3^3 + 5x_3^4 - 7x_2^2x_3^2 + 16x_2^4 \end{aligned}$$

- Not so easy! (In fact, **NP-hard for degree ≥ 4**) (Have seen the proof already!)
- But what if I told you:

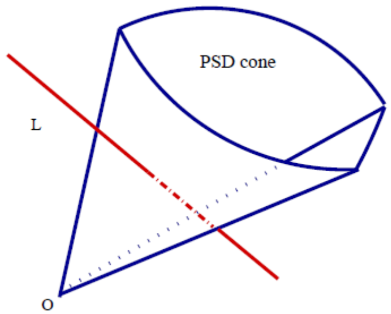
$$\begin{aligned} p(x) &= (x_1^2 - 3x_1x_2 + x_1x_3 + 2x_3^2)^2 + (x_1x_3 - x_2x_3)^2 \\ &\quad +(4x_2^2 - x_3^2)^2. \end{aligned}$$

Natural questions:

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

Sum of Squares and Semidefinite Programming

[Lasserre], [Nesterov], [Parrilo]



Q. Is it any easier to decide sos?

- Yes! Can be reduced to a **semidefinite program (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]

- Can also efficiently **search and optimize** over sos polynomials
  - This enables numerous applications…

SOS→SDP

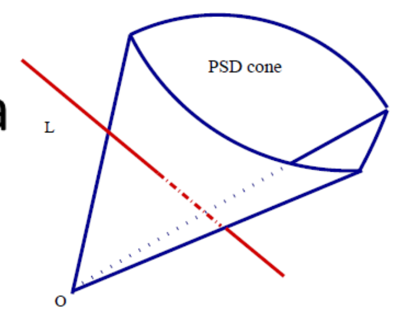
Thm: A polynomial ***p(x)*** of degree ***2d*** is sos if and only if there exists a matrix Q such that

$$Q \succcurlyeq 0,$$
$$p(x) = z(x)^T Q z(x),$$

where

$$z = [1, x_1, x_2, \ldots, x_n, x_1 x_2, \ldots, x_n^d]^T$$

The set of such matrices Q forms the feasible set of a semidefinite program.



Example

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

$$p(x) = z^TQz \qquad z = (x_1^2, x_1x_2, x_2^2, x_1x_3, x_2x_3, x_3^2)^T$$

$$Q = \begin{pmatrix} 1 & -3 & 0 & 1 & 0 & 2 \\ -3 & 9 & 0 & -3 & 0 & -6 \\ 0 & 0 & 16 & 0 & 0 & -4 \\ 1 & -3 & 0 & 2 & -1 & 2 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 2 & -6 & 4 & 2 & 0 & 5 \end{pmatrix}$$

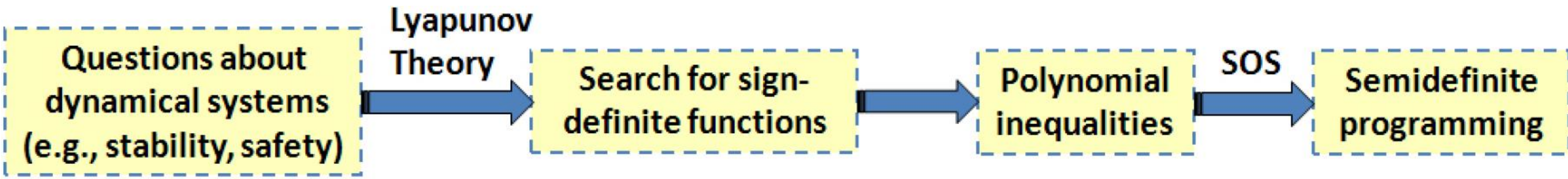
$$Q = \sum_{i=1}^3 a_ia_i^T$$

$$a_1 = (1, -3, 0, 1, 0, 2)^T, \; a_2 = (0, 0, 0, 1, -1, 0)^T, \; a_3 = (0, 0, 4, 0, 0, -1)^T$$

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

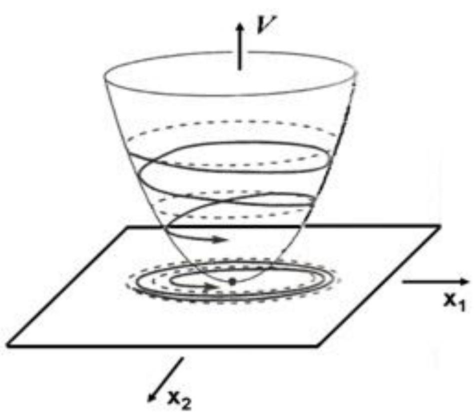
Fully automated in YALMIP, SOSTOOLS, SPOTLESS, GloptiPoly, …

Automated search for Lyapunov functions



Ex. Lyapunov's stability theorem.

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



Lyapunov function

$$V(x) : \mathbb{R}^n \to \mathbb{R}$$

$$\dot{V}(x) = \langle \frac{\partial V}{\partial x}, f(x) \rangle$$

$$\begin{array}{ll} V(x) & \text{sos} \\ -\dot{V}(x) & \text{sos} \end{array} \Rightarrow \begin{array}{l} V(x) > 0 \\ -\dot{V}(x) > 0 \end{array} \Rightarrow \text{GAS}$$

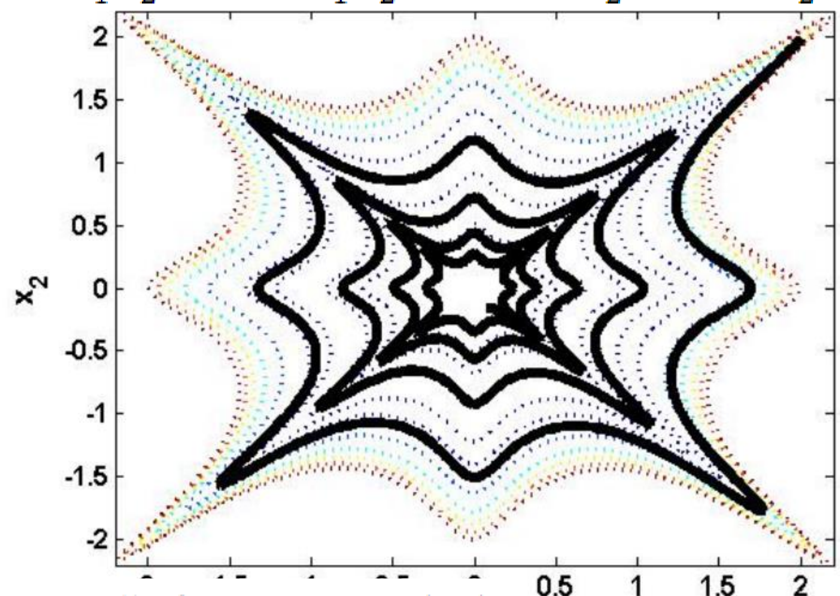
(similar local version)

Global stability

$$\begin{array}{ll} V(x) & \text{sos} \\ -\dot{V}(x) & \text{sos} \end{array} \Rightarrow \begin{array}{l} V(x) > 0 \\ -\dot{V}(x) > 0 \end{array} \Rightarrow \text{GAS}$$

Example.

$$\begin{aligned} \dot{x}_1 &= -0.15x_1^7 + 200x_1^6x_2 - 10.5x_1^5x_2^2 - 807x_1^4x_2^3 + 14x_1^3x_2^4 + 600x_1^2x_2^5 - 3.5x_1x_2^6 + 9x_2^7 \\ \dot{x}_2 &= -9x_1^7 - 3.5x_1^6x_2 - 600x_1^5x_2^2 + 14x_1^4x_2^3 + 807x_1^3x_2^4 - 10.5x_1^2x_2^5 - 200x_1x_2^6 - 0.15x_2^7 \end{aligned}$$



Output of SDP solver:

$$\begin{aligned} V &= 0.02x_1^8 + 0.015x_1^7x_2 + 1.743x_1^6x_2^2 - 0.106x_1^5x_2^3 - 3.517x_1^4x_2^4 \\ &\quad +0.106x_1^3x_2^5 + 1.743x_1^2x_2^6 - 0.015x_1x_2^7 + 0.02x_2^8. \end{aligned}$$

Hilbert's 1888 Paper

Q. SOS $\overset{?}{\Leftarrow}$ Nonnegativity

Polynomials

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

Forms (homog. polynomials)

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

Homogenization:

$$\begin{cases} p_h(x,y) = y^d p(x/y) \\ p(x) = p_h(x,1) \end{cases}$$

$$p(x) = x_1^2 x_2^2 + x_1 x_2 + 4$$

$$p_h(x,y) = x_1^2 x_2^2 + x_1 x_2 y^2 + 4y^4$$

Motzkin (1967):

$$M(x_1, x_2, x_3) = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 x_3^2 + x_3^6$$

Robinson (1973):

$$\begin{aligned} R(x_1, x_2, x_3, x_4) \quad = \quad & x_1^2(x_1 - x_4)^2 + x_2^2(x_2 - x_4)^2 + x_3^2(x_3 - x_4)^2 \\ & + 2x_1 x_2 x_3 (x_1 + x_2 + x_3 - 2x_4) \end{aligned}$$

The Motzkin polynomial

$$M(x_1, x_2, x_3) = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 x_3^2 + x_3^6$$

- How to show it's nonnegative but not sos?!

M nonnegative

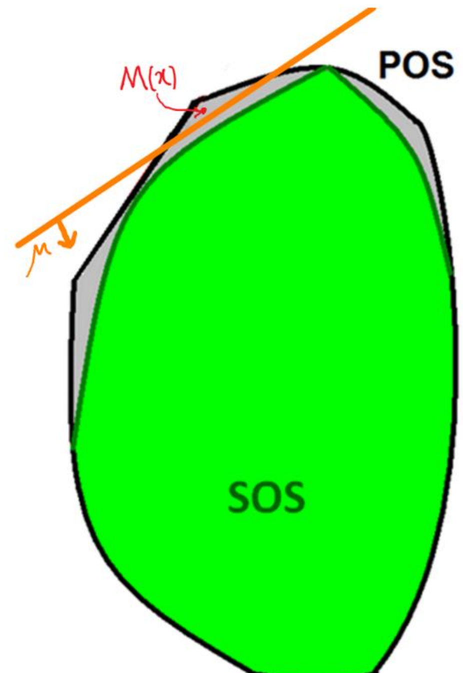
- $\frac{1}{3} M(x) = \frac{x_1^4 x_2^2 + x_1^2 x_2^4 + x_3^6}{3} - x_1^2 x_2^2 x_3^2$ AMGM inequality

- $M(x)(x_1^2 + x_2^2 + x_3^2)$ sos

M not sos

- $M(x) = \sum q_i^2(x)$ (q_i cubic)

- Separating hyperplane



$\langle \mu, \vec{M} \rangle < 0$

$\langle \mu, \vec{P} \rangle \geq 0 \quad \forall p$ sos

The good news

- In relatively small dimensions and degrees, it seems difficult to construct nonnegative polynomials that are not sos
- Especially true if additional structure is required
- For example, the following is OPEN:

Construct a *convex*, nonnegative polynomial that is not sos

- Empirical evidence from various domains over the last decade:

SOS is a very powerful relaxation.

Hilbert's 17th Problem (1900)

Q. p **nonnegative** $\overset{?}{\Rightarrow}$ $p = \sum_i \left(\frac{g_i}{q_i}\right)^2$

- Artin (1927): **Yes!**

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

Positivstellensatz: a complete algebraic proof system

- Let's motivate it with a toy example:

Consider the task of proving the statement:

$$\forall a, b, c, x, \;\; ax^2 + bx + c = 0 \Rightarrow b^2 - 4ac \geq 0$$

Short algebraic proof (certificate):

$$b^2 - 4ac = (2ax + b)^2 - 4a(ax^2 + bx + c)$$

- The Positivstellensatz vastly generalizes what happened here:
  - Algebraic certificates of infeasibility of any system of polynomial inequalities (or algebraic implications among them)
  - **Automated** proof system (via semidefinite programming)

Positivstellensatz: a generalization of Farkas lemma

Farkas lemma (1902):

$Ax = b$ and $x \geq 0$ is infeasible

$$\Updownarrow$$

There exists a y such that $y^T A \geq 0$ and $y^T b < 0$.

(The S-lemma is also a theorem of this type for quadratics)

Positivstellensatz

Stengle (1974):

The basic semialgebraic set

$$K := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, i = 1, \ldots, m, h_i(x) = 0, i = 1, \ldots, k\}$$

is empty

$$\Updownarrow$$

there exist polynomials $t_1, \ldots, t_k$ *and sum of squares polynomials* $s_0, s_1, \ldots, s_m, s_{12}, s_{13}, \ldots, s_{m-1m}, s_{123}, \ldots, s_{m-2m-1m}, \ldots, s_{12\ldots m}$ *such that*

$$\begin{aligned} -1 \;=\; & \sum_{i=1}^{k} t_i(x)h_i(x) + s_0(x) + \sum_{\{i\}} s_i(x)g_i(x) \\ + \; & \sum_{\{i,j\}} s_{ij}(x)g_i(x)g_j(x) + \sum_{\{i,j,k\}} s_{ijk}(x)g_i(x)g_j(x)g_k(x) \\ + \; & \cdots + s_{ijk\ldots m}(x)g_i(x)g_j(x)g_k(x)\ldots g_m(x). \end{aligned}$$

- Comments:
  - Hilbert's 17th problem is a straightforward corollary
  - Other versions due to Shmudgen and Putinar (can look simpler)

Parrilo/Lasserre SDP hierarchies

Recall POP:

$$\begin{array}{ll} \text{minimize} & p(x) \\ \text{subject to} & x \in K := \{x \in \mathbb{R}^n \mid g_i(x) \geq 0, h_i(x) = 0\} \end{array}$$

Idea:

obtain the largest lower bound by finding the largest γ for which the set $\{x \in K, p(x) \leq \gamma\}$ is empty.

certify this emptiness by finding Positivstellensatz certificates.

$$\begin{aligned} -1 \;=\; & \sum_{i=1}^{k} t_i(x)h_i(x) + s_0(x) + \sum_{\{i\}} s_i(x)g_i(x) \\ +\; & \sum_{\{i,j\}} s_{ij}(x)g_i(x)g_j(x) + \sum_{\{i,j,k\}} s_{ijk}(x)g_i(x)g_j(x)g_k(x) \\ +\; & \cdots + s_{ijk\ldots m}(x)g_i(x)g_j(x)g_k(x)\ldots g_m(x) \end{aligned}$$

In level l of the hierarchy, degree of the polynomials t_i and the sos polynomails s_i is bounded by l.

Comments:

- Each fixed level of the hierarchy is an SDP of polynomial size
- Originally, Parrilo's hierarchy is based on Stengle's Psatz, whereas Lasserre's is based on Putinar's Psatz

Lyapunov Barrier Certificates

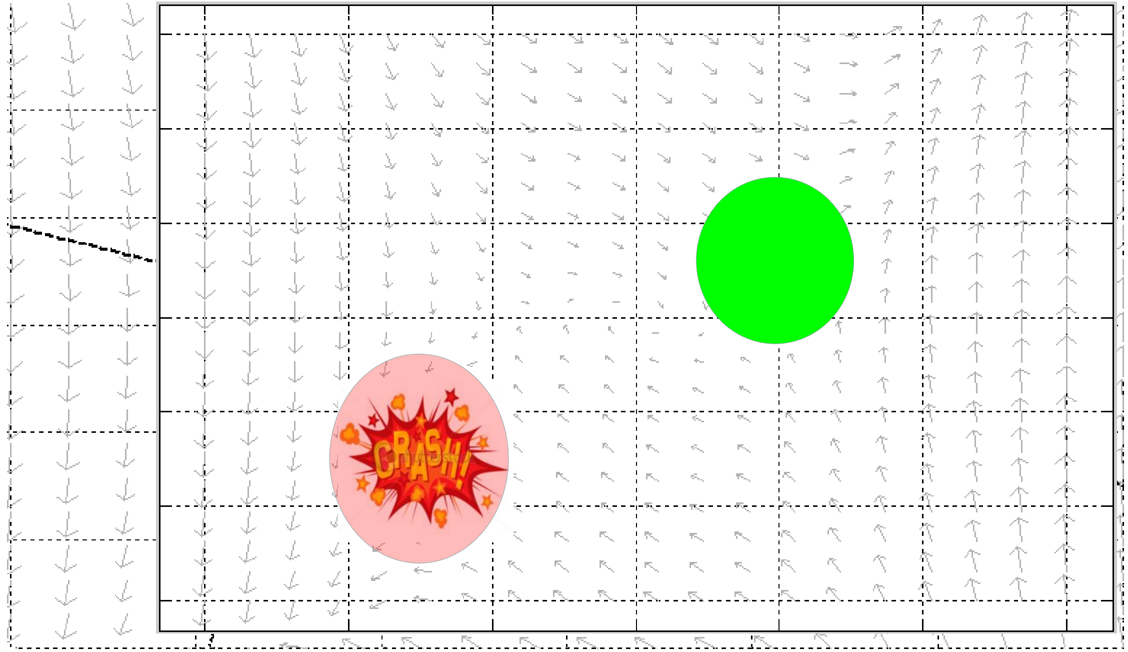
$$\dot{x} = f(x)$$

(vector valued polynomial)

$\mathcal{S}$: needs safety verification

$\mathcal{U}$: unsafe (or forbidden) set

(both sets semialgebraic)



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

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

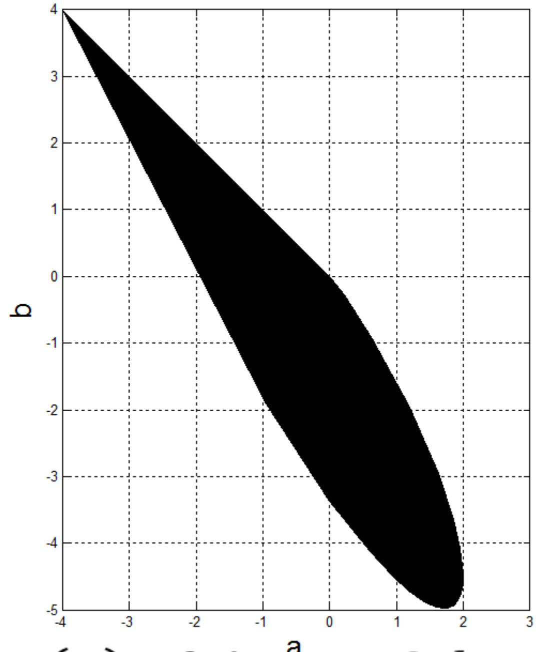
$$\dot{B} = \langle \nabla B(x), f(x) \rangle \leq 0$$

(extends to uncertain or hybrid systems, stochastic differential equations and probabilistic guarantees [Prajna et al.])

How did I plot this?

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

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

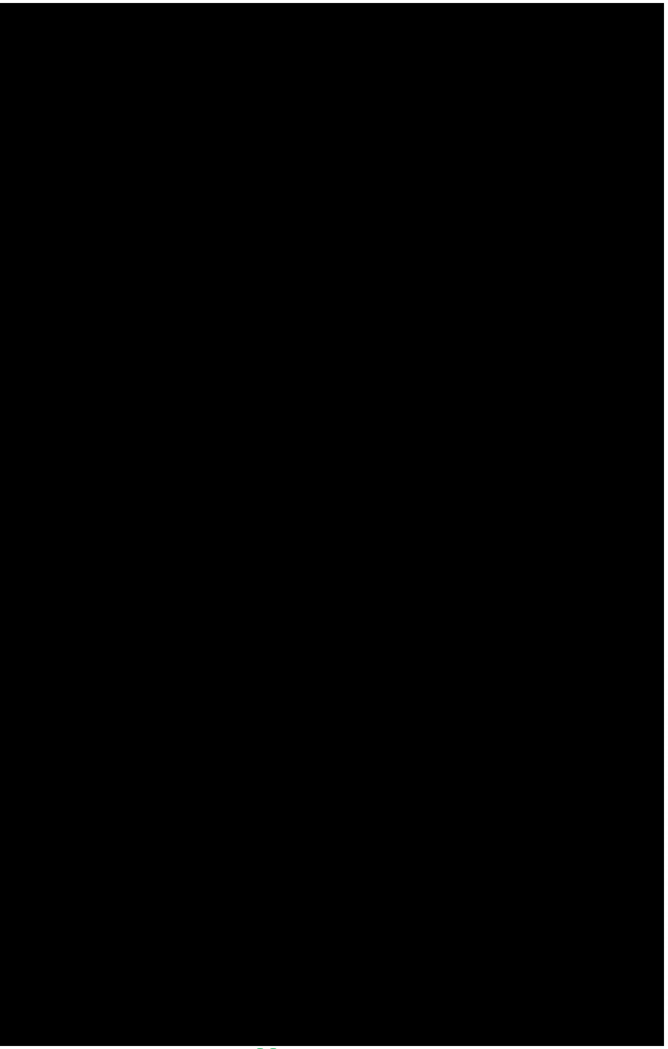


Theorem. A polynomial $p(x)$ of degree $2d$ is monotone on $[0,1]$ if and only if

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

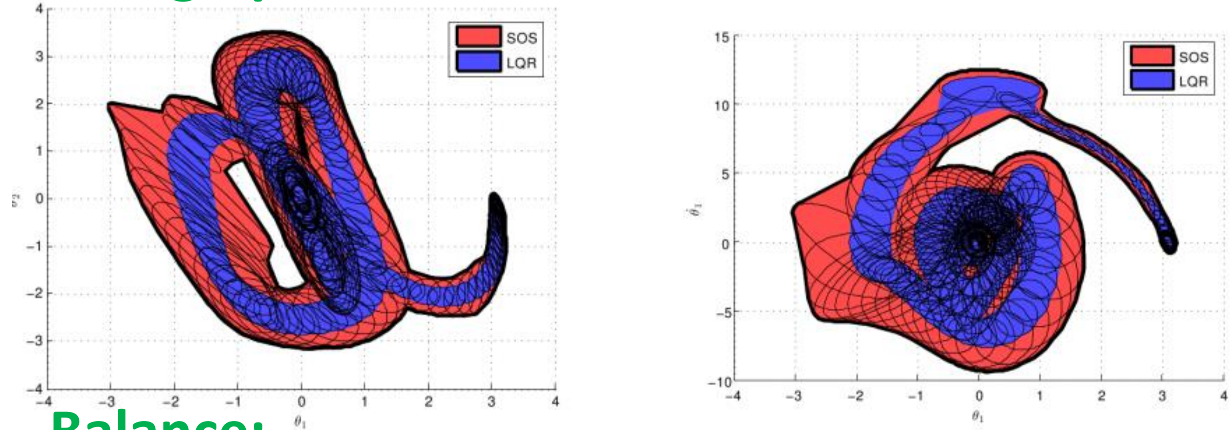
where $s_1(x)$ and $s_2(x)$ are some SOS polynomials of degree $2d-2$.

This stuff actually works! (SOS on Acrobot)

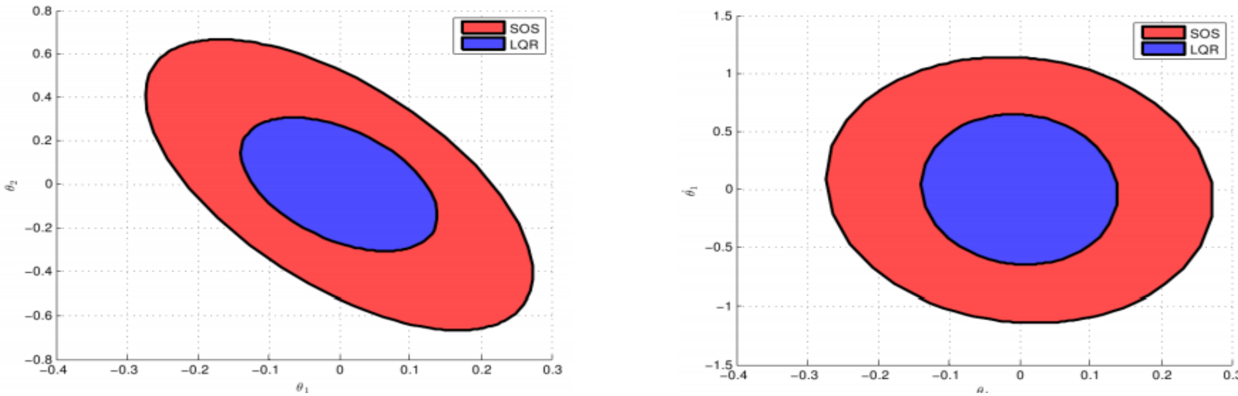


Controller designed by SOS

Swing-up:



Balance:



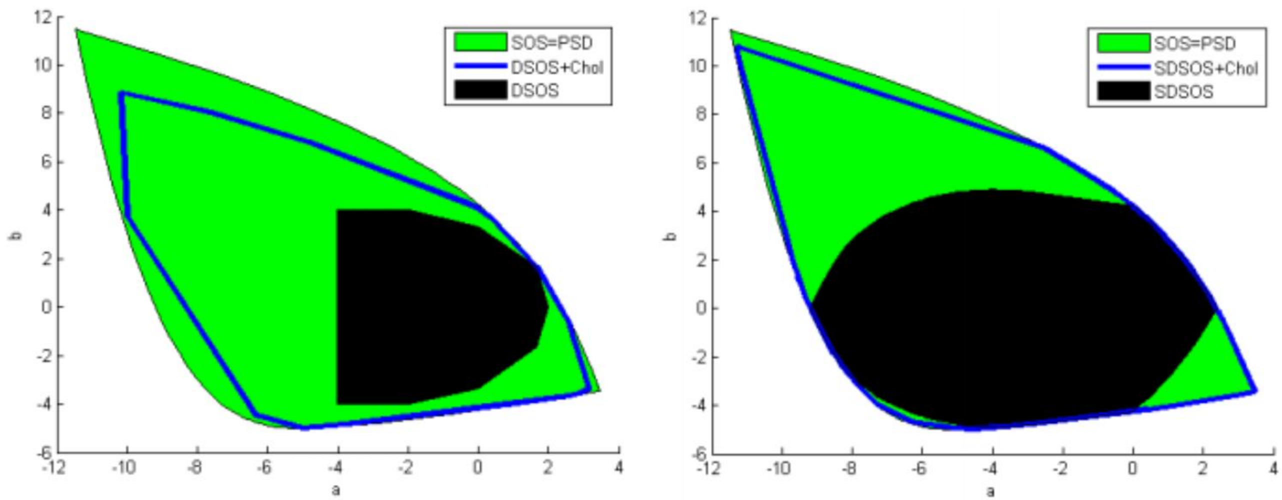
[Majumdar, AAA, Tedrake]
(Best paper award - *IEEE Conf. on Robotics and Automation*, '13)

Some recent algorithmic developments

Approximating sum of squares programs with LPs and SOCPs

(DSOS and SDSOS Optimization)

[AAA, Majumdar]
[AAA, Hall]
[AAA, Dash, Hall]
[Majumdar, AAA, Tedrake]



(a) LP inner approximations (b) SOCP inner approximations

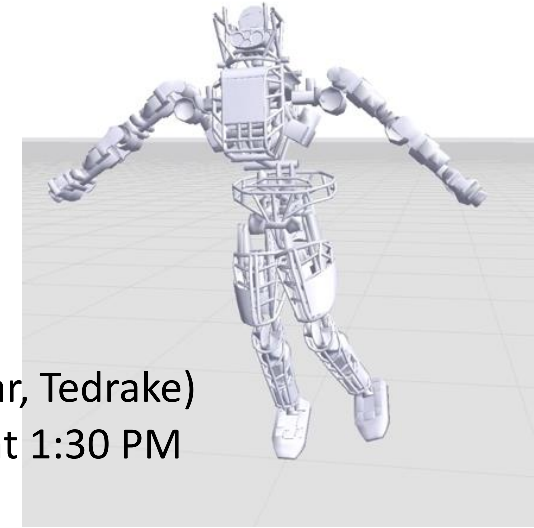
- Do we really need semidefinite programming?
- Tradeoffs between speed and accuracy of approximation?

Highlights:

- Order of magnitude speed-up in practice.
- New applications at larger scale now within reach.
- Potential for real-time algebraic optimization.

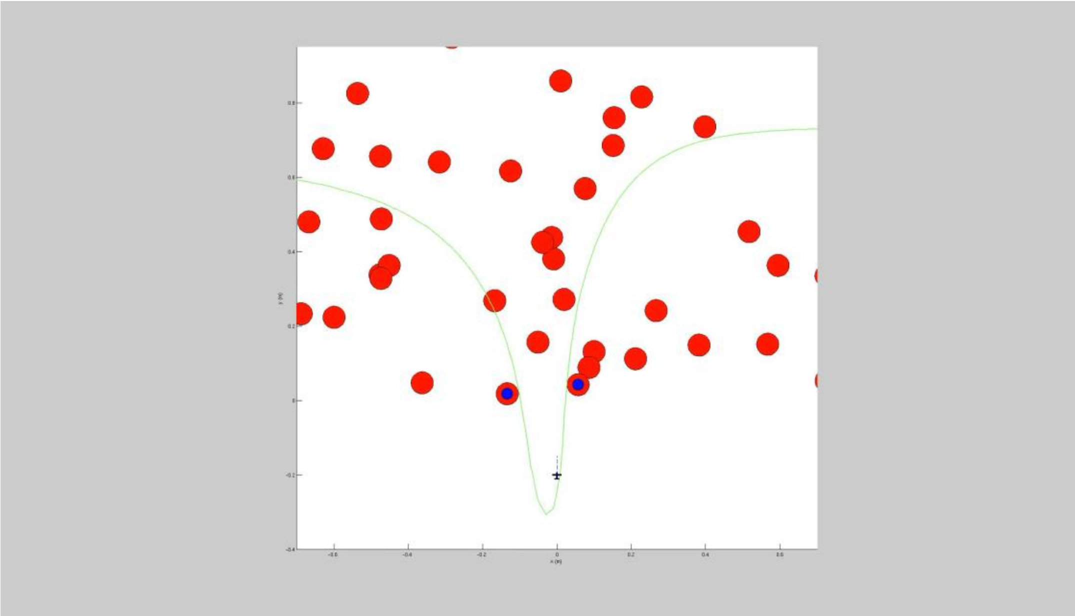
Some recent algorithmic developments

Stabilization/ collision avoidance by **SDSOS Optimization**



(w/ Majumdar, Tedrake)
TC02, Bowl 1, at 1:30 PM

(w/ Majumdar)



Want to know more?

- Next tutorial: the Lasserre hierarchy
  - Etienne de Klerk
- After lunch: follow-up session
  - TC02, Bowl 1, at 1:30 PM
- Tomorrow (FA02)
  - Javad Lavaei, polynomial optimization in energy applications
- Tomorrow (FD01)
  - Russ Tedrake's plenary talk, applications in robotics

Thank you!

aaa.princeton.edu