5. Sum of Squares

- Polynomial nonnegativity
- Sum of squares (SOS) decompositions
- Computing SOS using semidefinite programming
- Liftings
- Dual side: moments
- Applications
    - Global optimization
    - Optimizing in parameter space
    - Lyapunov functions
    - Density functions and control synthesis

Polynomial Nonnegativity

Before dealing with *systems* of polynomial inequalities, we study the simplest nontrivial problem: one inequality.

Given $f(x_1, \ldots, x_n)$ (of even degree), is it globally nonnegative?

$$f(x_1, x_2, \ldots, x_n) \geq 0, \quad \forall x \in \mathbb{R}^n$$

- For quadratic polynomials $(d = 2)$, very easy. Essentially, checking if a matrix is PSD.
- The problem is *NP-hard* when $d \geq 4$.
- Problem is decidable, algorithms exist (more later). Very powerful, but *bad complexity properties*.
- *Many* applications. We'll see a few...

Sum of Squares Decomposition

A “simple” sufficient condition: a sum of squares (SOS) decomposition:

$$f(x) = \sum_i g_i^2(x), \qquad g_i \in \mathbb{R}[x]$$

If $f(x)$ can be written as above, then $f(x) \geq 0$.
A purely syntactic, easily verifiable certificate.

Always a sufficient condition for nonnegativity.
In some cases (univariate, quadratic, etc.), also necessary.
But in general, SOS is *not* equivalent to nonnegativity.

However, a very good thing: we can compute this efficiently using SDP.

Sum of Squares and SDP

Consider a polynomial $f(x_1, \ldots, x_n)$ of degree $2d$.

Let z be a vector with all monomials of degree less than or equal to d.

The number of components of z is $\binom{n+d}{d}$.

Then, f is SOS iff:

$$f(x) = z^T Q z, \qquad Q \succeq 0$$

- Factorize $Q = L^T L$. Then

$$f(x) = z^T L^T L z = ||Lz||^2 = \sum_i (Lz)_i^2$$

- The terms in the SOS decomposition are given by $g_i = (Lz)_i$.

- The number of squares is equal to the rank of Q.

$$f(x) = z^T Q z, \qquad Q \succeq 0$$

- Comparing terms, we obtain linear equations for the elements of Q.
- The desired matrices Q lie in the intersection of an *affine set* of matrices, and the PSD cone.
- In general, Q is *not* unique.
- Can be solved as *semidefinite program* in the standard primal form.

$$\{Q \succeq 0, \quad \mathbf{trace}\, A_i Q = b_i\}$$

Multivariate SOS Example

$$
\begin{aligned}
f(x,y) &= 2x^4 + 5y^4 - x^2y^2 + 2x^3y \\
&= \begin{bmatrix} x^2 \\ y^2 \\ xy \end{bmatrix}^T \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} x^2 \\ y^2 \\ xy \end{bmatrix} \\
&= q_{11}x^4 + q_{22}y^4 + (q_{33} + 2q_{12})x^2y^2 + 2q_{13}x^3y + 2q_{23}xy^3
\end{aligned}
$$

The existence of a PSD Q is exactly equivalent to feasibility of an SDP in the standard primal form:

$$
Q \succeq 0, \qquad \text{subject to} \begin{cases} q_{11} = 2 & q_{22} = 5 \\ 2q_{23} = 0 & 2q_{13} = 2 \\ & q_{33} + 2q_{12} = -1 \end{cases}
$$

Multivariate SOS Example (continued)

Solving numerically, we obtain a particular solution:

$$Q = \begin{bmatrix} 2 & -3 & 1 \\ -3 & 5 & 0 \\ 1 & 0 & 5 \end{bmatrix} = L^T L, \qquad L = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \end{bmatrix}$$

This Q has rank two, therefore $f(x, y)$ is the sum of two squares:

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

This representation *certifies* nonnegativity of f.

Using SOSTOOLS: `[Q,Z]=findsos(2*x^4+5*y^4-x^2*y^2+2*x^3*y)`

Some Background

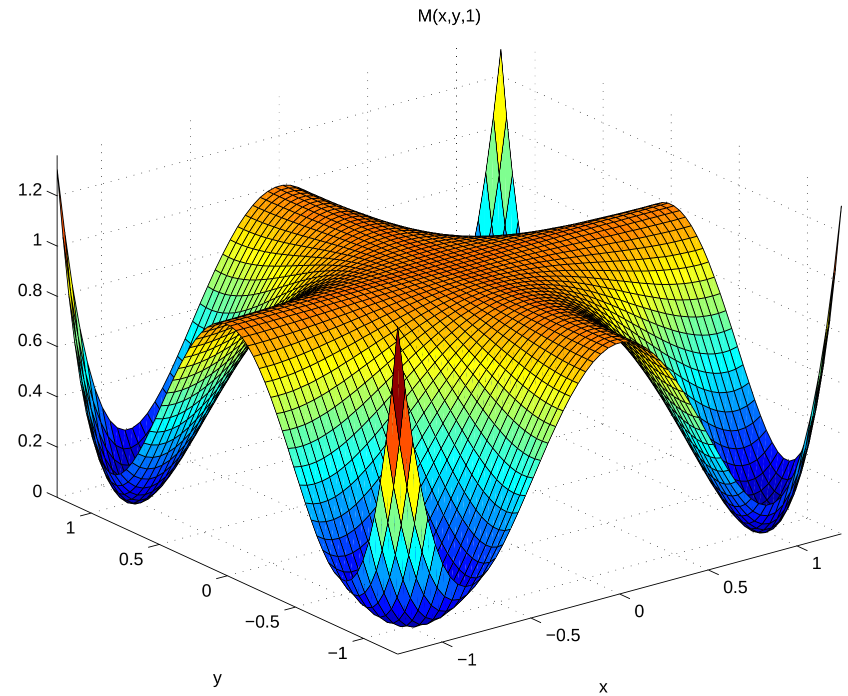
- In 1888, Hilbert showed that PSD=SOS if and only if
  - $d = 2$. Quadratic polynomials. SOS decomposition follows from Cholesky, square root, or eigenvalue decomposition.
  - $n = 1$. Univariate polynomials.
  - $d = 4, n = 2$. Quartic polynomials in two variables.
- Connections with Hilbert's 17th problem, solved by Artin: every PSD polynomial is a SOS of *rational functions*.
- If f is not SOS, then can try with gf, for some g.
  - For fixed f, can optimize over g too
  - Otherwise, can use a "universal" construction of Pólya-Reznick.

More about this later.

The Motzkin Polynomial

A positive semidefinite polynomial, that is *not* a sum of squares.

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



- Nonnegativity follows from the arithmetic-geometric inequality applied to $(x^2y^4, x^4y^2, 1)$
- Introduce a nonnegative factor $x^2 + y^2 + 1$
- Solving the SDPs we obtain the decomposition:

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

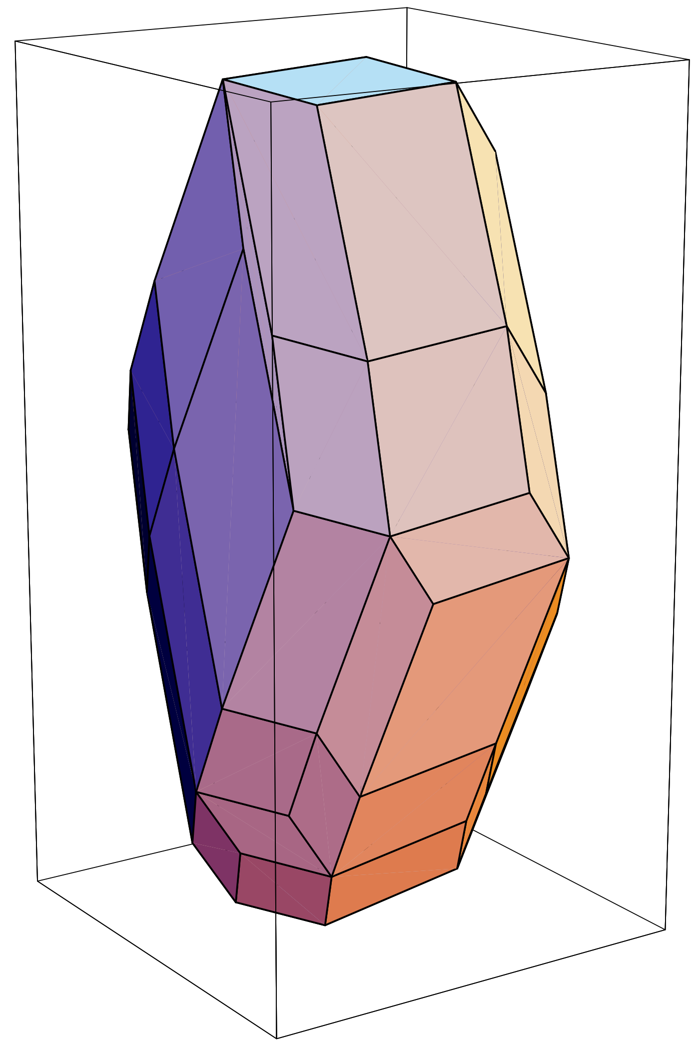
The Univariate Case:

$$
\begin{aligned}
f(x) &= a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_{2d} x^{2d} \\
&= \begin{bmatrix} 1 \\ x \\ \vdots \\ x^d \end{bmatrix}^T \begin{bmatrix} q_{00} & q_{01} & \cdots & q_{0d} \\ q_{01} & q_{11} & \cdots & q_{1d} \\ \vdots & \vdots & \ddots & \vdots \\ q_{00} & q_{1d} & \cdots & q_{dd} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ \vdots \\ x^d \end{bmatrix} \\
&= \sum_{i=0}^{d} \left(\sum_{j+k=i} q_{jk} \right) x^i
\end{aligned}
$$

- In the univariate case, the SOS condition is exactly equivalent to non-negativity.
- The matrices A_i in the SDP have a Hankel structure. This can be exploited for efficient computation.

A General Method: Liftings

Consider this polytope in $\mathbb{R}^3$ (a zonotope).
It has 56 facets, and 58 vertices.



Optimizing a linear function over this set, requires a *linear program* with 56 constraints (one per face).

However, this polyhedron is a three-dimensional *projection* of the 8-dimensional hypercube $\{x \in \mathbb{R}^8, -1 \leq x_i \leq 1\}$.

Therefore, by using additional variables, we can solve the same problem, by using an LP with *only 16 constraints*.

Liftings

By going to higher dimensional representations, things may become easier:

- "Complicated" sets can be the projection of much simpler ones.
- A polyhedron in $\mathbb{R}^n$ with a "small" number of faces can project to a lower dimensional space with *exponentially* many faces.
- Basic semialgebraic sets can project into non-basic semialgebraic sets.

An essential technique in integer programming.

Advantages: compact representations, avoiding "case distinctions," etc.

Example

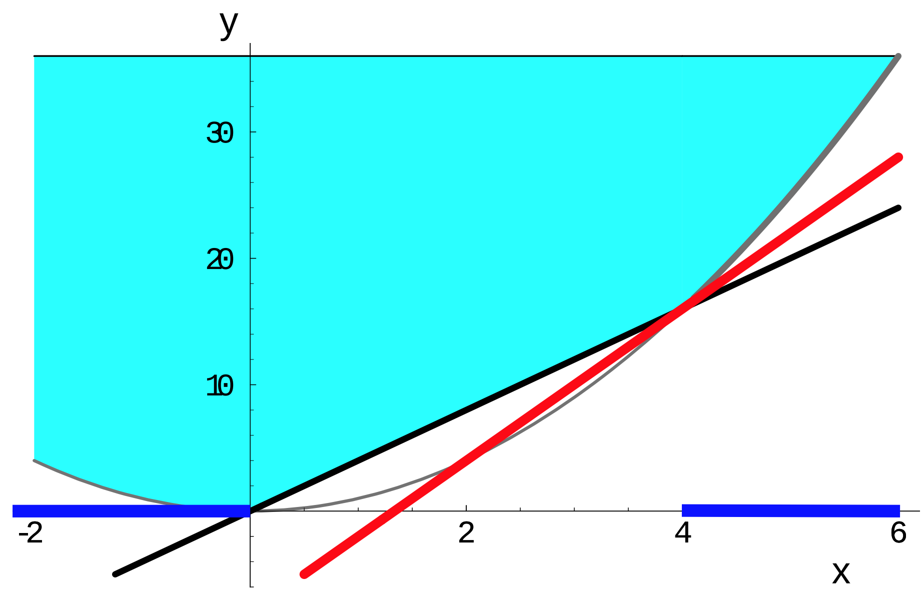
$$\begin{array}{rl} \text{minimize} & (x-3)^2 \\ \text{subject to} & x\,(x-4) \geq 0 \end{array}$$

The feasible set is $[-\infty, 0] \cup [4, \infty]$. *Not* convex, or even connected.

Consider the lifting $L : \mathbb{R} \to \mathbb{R}^2$, with $L(x) = (x, x^2) =: (x, y)$.

Rewrite the problem in terms of the lifted variables.

- For every lifted point, $\begin{bmatrix} 1 & x \\ x & y \end{bmatrix} \succeq 0$.
- Constraint becomes: $y - 4x \geq 0$
- Objective is now: $y - 6x + 9$



We "get around" nonconvexity: interior points are now on the *boundary*.

The Dual Side of SOS: Moment Sequences

The SDP dual of the SOS construction gives efficient *semidefinite liftings*.

For the univariate case: $L : \mathbb{R} \to \mathbb{S}^{d+1}$, with

$$L(x) = \begin{bmatrix} 1 & x & \dots & x^d \\ x & x^2 & \dots & x^{d+1} \\ \vdots & \vdots & \ddots & \vdots \\ x^d & x^{d+1} & \dots & x^{2d} \end{bmatrix}$$

The matrices $L(x)$ are Hankel, positive semidefinite, and rank one.

The convex hull $\mathbf{co}\, L(x)$ therefore contains only PSD Hankel matrices.

$$\mathbf{Hankel}(w) := \begin{bmatrix} 1 & w_1 & \dots & w_d \\ w_1 & w_2 & \dots & w_{d+1} \\ \vdots & \vdots & \ddots & \vdots \\ w_d & w_{d+1} & \dots & w_{2d} \end{bmatrix}$$

(in fact, in the univariate case *every* PSD Hankel is in the convex hull)

SOS Dual (continued)

For nonnegativity, want to rule out the existence of x with $f(x) < 0$.

In the lifted variables, we can look at:

$$\left\{ \mathbf{Hankel}(w) \succeq 0, \quad \sum_i a_i w_i < 0 \right\}$$

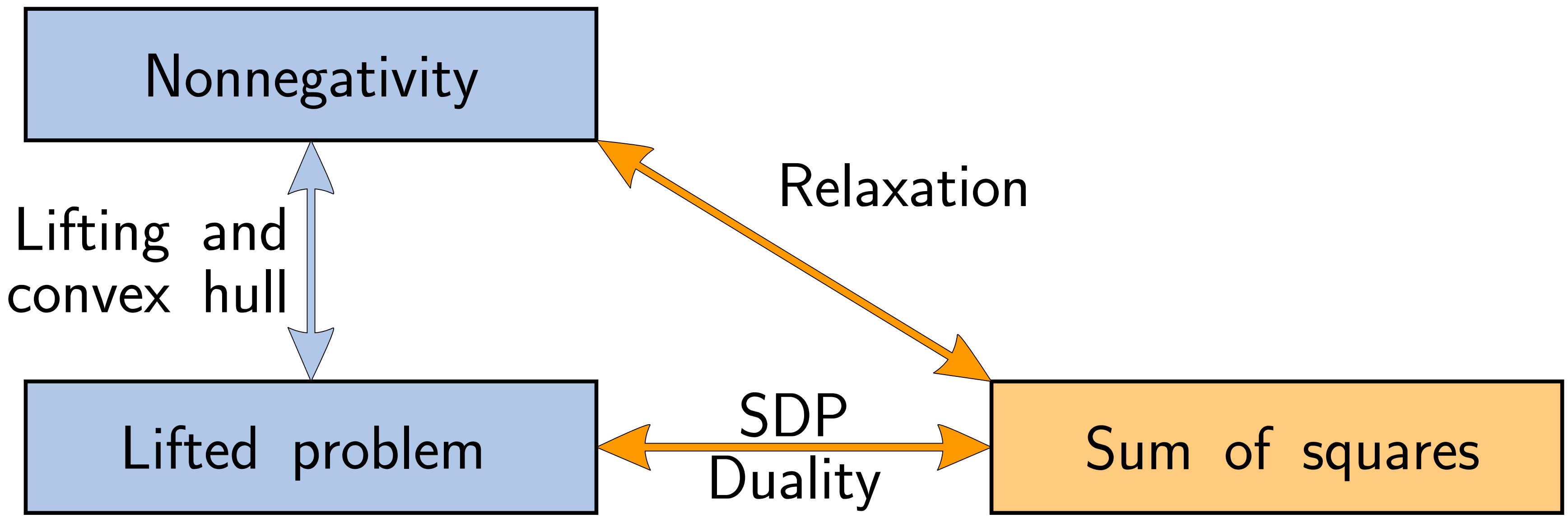
This is *exactly* the SDP dual of the univariate SOS construction.

$$\left\{ Q \succeq 0, \quad \sum_{j+k=i} q_{jk} = a_i \right\}$$

- If the first problem is feasible, there is always a w such that $\mathbf{Hankel}(w)$ is *rank one*. It corresponds directly to the lifting of a primal point.

Direct extensions to the multivariate case. Though in general, PSD $\neq$ SOS.

A General Scheme



- Lifting corresponds to a classical problem of *moments*.
- The solution to the lifted problem *may* suggest candidate points where the polynomial is negative.
- The sums of squares *certify* or *prove* polynomial nonnegativity.

We'll be generalizing this. . .

About SOS/SDP

- The resulting SDP problem is polynomially sized (in n).
- By properly choosing the monomials, we can exploit structure (sparsity, symmetries, ideal structure).
- An important feature: the problem is still a SDP *if the coefficients of F are variable*, and the dependence is affine.
- Can optimize over SOS polynomials in affinely described families.

  For instance, if we have $p(x) = p_0(x) + \alpha p_1(x) + \beta p_2(x)$, we can "easily" find values of α, β for which $p(x)$ is SOS.

This fact will be *crucial* in everything that follows...

Global Optimization

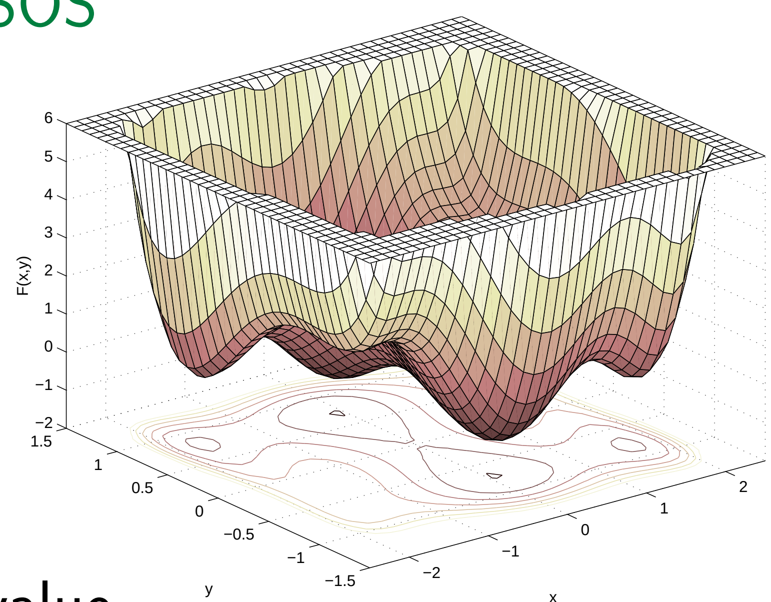
Consider the problem

$$\min_{x,y} f(x,y)$$

with

$$f(x,y) := 4x^2 - \frac{21}{10}x^4 + \frac{1}{3}x^6 + xy - 4y^2 + 4y^4$$

- Not convex. Many local minima. NP-hard.
- Find the largest γ s.t. $f(x,y) - \gamma$ is SOS
- Essentially due to Shor (1987).
- A semidefinite program (convex!).
- If exact, can recover optimal solution.
- *Surprisingly* effective.



Solving, the maximum γ is -1.0316. Exact value.

Why Does This Work?

Three *independent* facts, theoretical and experimental:

- The existence of efficient algorithms for SDP.
- The size of the SDPs grows much slower than the Bézout number μ.
  - A bound on the number of (complex) critical points.
  - A reasonable estimate of complexity.
  - The bad news: $\mu = (2d-1)^n$ (for dense polynomials).
  - Almost all (exact) algebraic techniques scale as μ.
- The lower bound f^{SOS} very often coincides with f^*.
  (Why? what does *often* mean?)

SOS provides *short proofs*, even though they're not guaranteed to exist.

Coefficient Space

Let $f_{\alpha\beta}(x) = x^4 + (\alpha + 3\beta)x^3 + 2\beta x^2 - \alpha x + 1$.

What is the set of values of $(\alpha, \beta) \in \mathbb{R}^2$ for which $f_{\alpha\beta}$ is PSD? SOS?

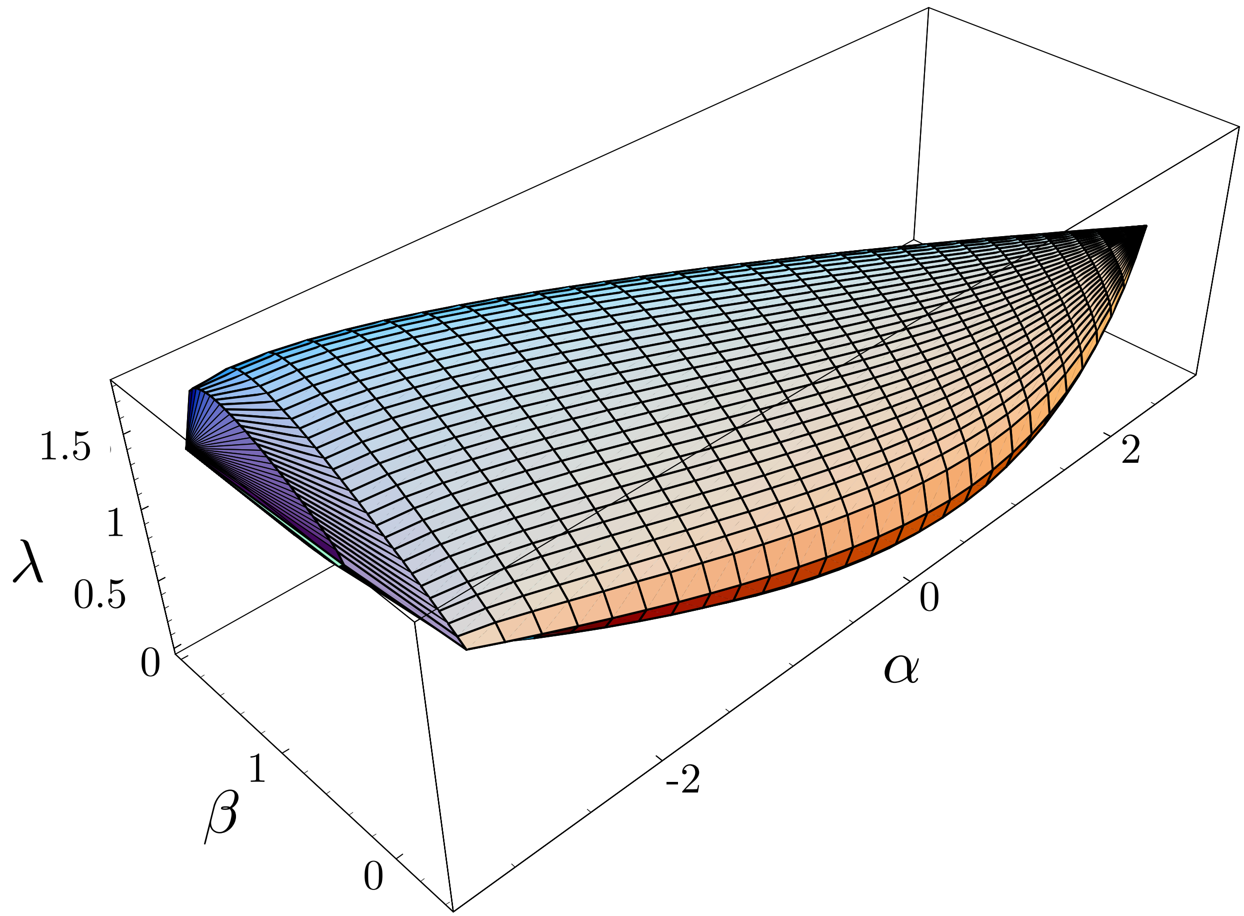
To find a SOS decomposition:

$$\begin{aligned} f_{\alpha,\beta}(x) &= 1 - \alpha x + 2\beta x^2 + (\alpha + 3\beta)x^3 + x^4 \\ &= \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix}^T \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} \\ &= q_{11} + 2q_{12}x + (q_{22} + 2q_{13})x^2 + 2q_{23}x^3 + q_{33}x^4 \end{aligned}$$

The matrix Q should be PSD and satisfy the affine constraints.

The feasible set is given by:

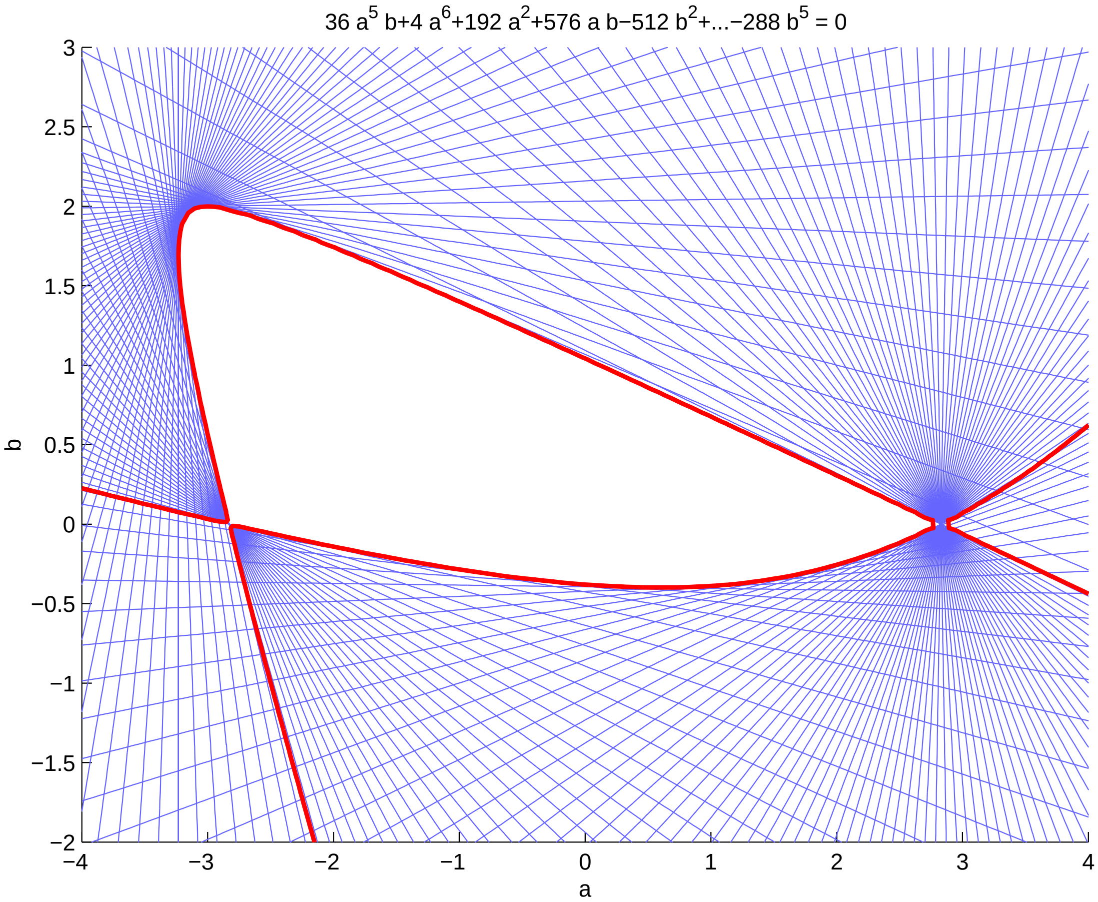
$$\left\{ (\alpha, \beta) \quad \Big| \quad \exists \lambda \text{ s.t. } \begin{bmatrix} 1 & -\frac{1}{2}\alpha & \beta - \lambda \\ -\frac{1}{2}\alpha & 2\lambda & \frac{1}{2}(\alpha + 3\beta) \\ \beta - \lambda & \frac{1}{2}(\alpha + 3\beta) & 1 \end{bmatrix} \succeq 0 \right\}$$



What is the set of values of $(\alpha, \beta) \in \mathbb{R}^2$ for which $f_{\alpha\beta}$ PSD? SOS?

Recall: in the univariate case PSD=SOS, so here the sets are the same.

- Convex and semialgebraic.
- It is the projection of a spectrahedron in $\mathbb{R}^3$.
- We can easily test membership, or even optimize over it!



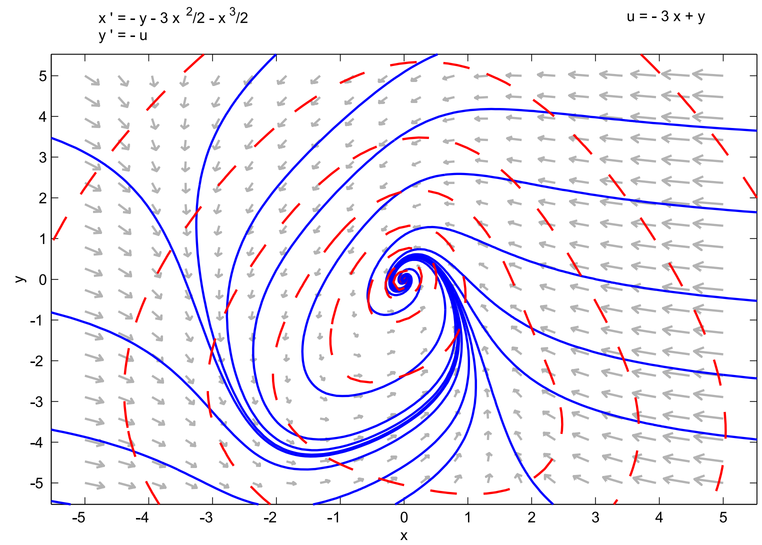
Defined by the curve: $288\beta^5 - 36\alpha^2\beta^4 + 1164\alpha\beta^4 + 1931\beta^4 - 132\alpha^3\beta^3 + 1036\alpha^2\beta^3 + 1956\alpha\beta^3 - 2592\beta^3 - 112\alpha^4\beta^2 + 432\alpha^3\beta^2 + 1192\alpha^2\beta^2 - 1728\alpha\beta^2 + 512\beta^2 - 36\alpha^5\beta + 72\alpha^4\beta + 360\alpha^3\beta - 576\alpha^2\beta - 576\alpha\beta - 4\alpha^6 + 60\alpha^4 - 192\alpha^2 - 256 = 0$

Lyapunov Stability Analysis

To prove asymptotic stability of $\dot{x} = f(x)$,

$$V(x) > 0 \quad x \neq 0$$

$$\dot{V}(x) = \left(\frac{\partial V}{\partial x}\right)^T f(x) < 0, \quad x \neq 0$$

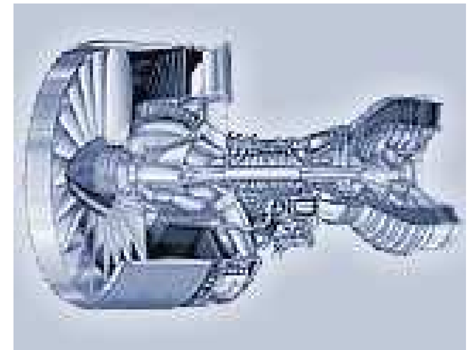


- For linear systems $\dot{x} = Ax$, quadratic Lyapunov functions $V(x) = x^T P x$

$$P > 0, \qquad A^T P + PA < 0.$$

- With an affine family of candidate polynomial V, $\dot{V}$ is also affine.
- Instead of *checking nonnegativity*, use a *SOS condition*.
- Therefore, for polynomial vector fields and Lyapunov functions, we can check the conditions using the theory described before.

Lyapunov Example



A jet engine model (derived from Moore-Greitzer), with controller:

$$\dot{x} = -y + \frac{3}{2}x^2 - \frac{1}{2}x^3$$
$$\dot{y} = 3x - y$$

Try a generic 4th order polynomial Lyapunov function.

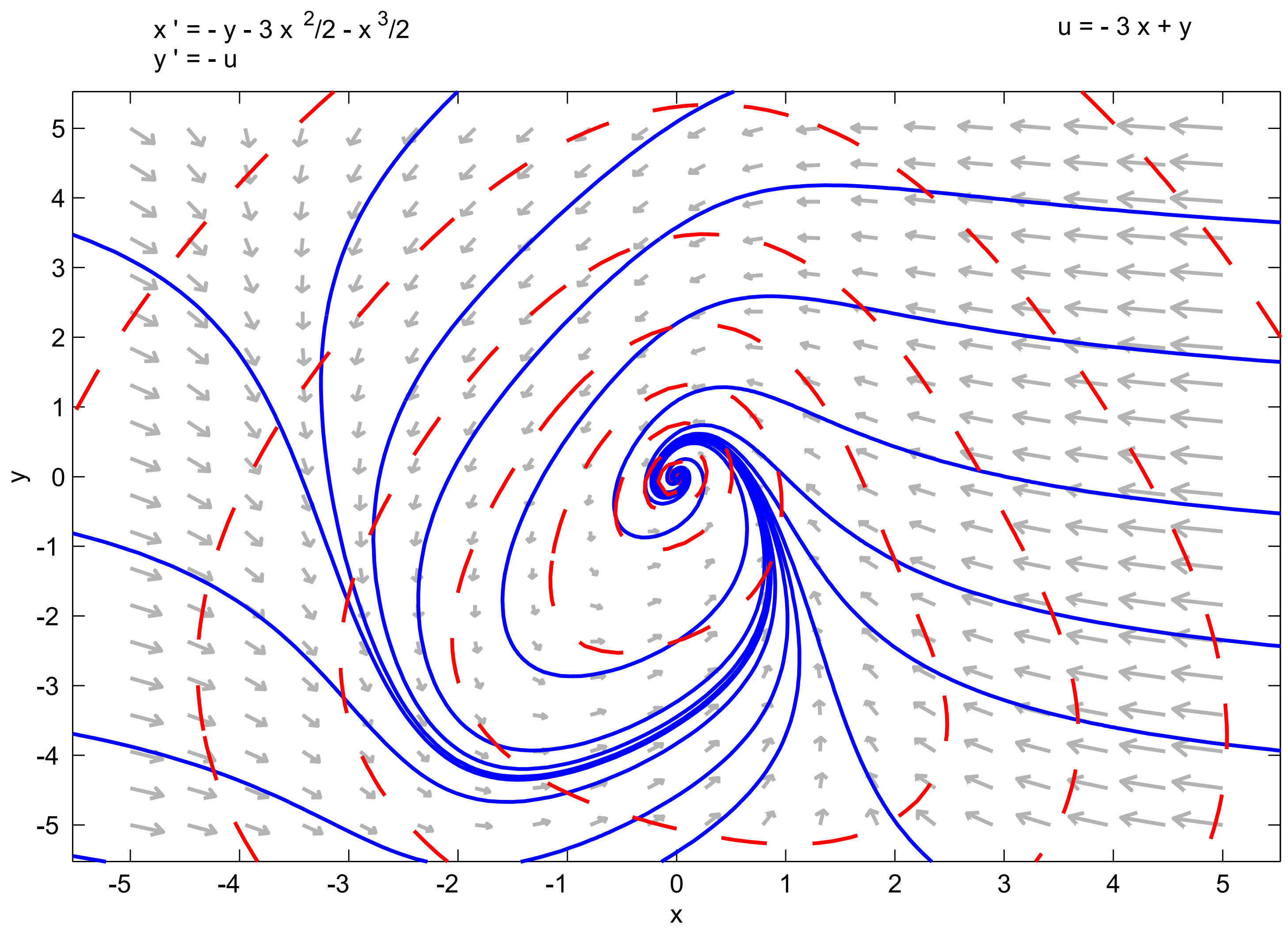
$$V(x,y) = \sum_{0 \le j+k \le 4} c_{jk} x^j y^k$$

Find a $V(x,y)$ that satisfies the conditions:

- $V(x,y)$ is SOS.
- $-\dot{V}(x,y)$ is SOS.

Both conditions are affine in the c_{jk}. Can do this directly using SOS/SDP!

After solving the SDPs, we obtain a Lyapunov function.



Lyapunov Example (2)

(M. Krstić) Find a Lyapunov function for *global asymptotic stability*:

$$\dot{x} = -x + (1+x)\,y$$
$$\dot{y} = -(1+x)\,x.$$

Using SOSTOOLS we easily find a quartic polynomial:

$$V(x,y) = 6x^2 - 2xy + 8y^2 - 2y^3 + 3x^4 + 6x^2y^2 + 3y^4.$$

Both $V(x,y)$ and $(-\dot{V}(x,y))$ are SOS:

$$V(x,y) = \begin{bmatrix} x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 6 & -1 & 0 & 0 & 0 \\ -1 & 8 & 0 & 0 & -1 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & -1 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}, \quad -\dot{V}(x,y) = \begin{bmatrix} x \\ y \\ x^2 \\ xy \end{bmatrix}^T \begin{bmatrix} 10 & 1 & -1 & 1 \\ 1 & 2 & 1 & -2 \\ -1 & 1 & 12 & 0 \\ 1 & -2 & 0 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ x^2 \\ xy \end{bmatrix}$$

The matrices are positive *definite*; this *proves* asymptotic stability.

Extensions

- Other linear differential inequalities (e.g. Hamilton-Jacobi).
- Many possible variations: nonlinear $\mathcal{H}_\infty$ analysis, parameter dependent Lyapunov functions, etc.
- Can also do local results (for instance, on compact domains).
- Polynomial and rational vector fields, or functions with an underlying algebraic structure.
- Natural extension of the LMIs for the linear case.
- Only for *analysis*. Proper *synthesis* is trickier...

Nonlinear Control Synthesis

Recently, Rantzer provided an alternative stability criterion, in some sense "dual" to the standard Lyapunov one.

$$\nabla \cdot (\rho f) > 0$$

- The *synthesis* problem is now *convex* in $(\rho, u\rho)$.

$$\nabla \cdot [\rho(f + gu)] > 0$$

- Parametrizing $(\rho, u\rho)$, can apply SOS methods.

Example:

$$\dot{x} = y - x^3 + x^2$$
$$\dot{y} = u$$

A stabilizing controller is:

$$u(x, y) = -1.22x - 0.57y - 0.129y^3$$