Optimization of Polynomials

Matthew M. Peet
Arizona State University
Thanks to S. Lall and P. Parrilo for guidance and supporting material

Lecture 03: Optimization of Polynomials

Overview

In this lecture, we will show how the LMI framework can be expanded dramatically to other forms of control problems.

1. Positivity of Polynomials
    1.1 Sum-of-Squares
2. Positivity of Polynomials on Semialgebraic sets
    2.1 Inference and Cones
    2.2 Positivstellensatz
3. Applications
    3.1 Nonlinear Analysis
    3.2 Robust Analysis and Synthesis
    3.3 Global optimization

Question: How do we Quantify the problem?

Consider: A System of Nonlinear Ordinary Differential Equations

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

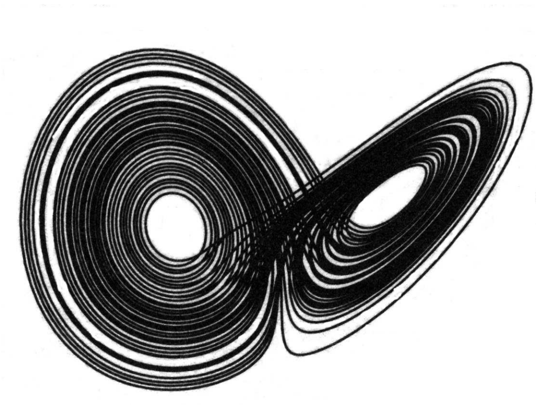
Problem 1: Exponential Decay Rate
Given a specific polynomial $f : \mathbb{R}^n \to \mathbb{R}^n$ and region $X \subset \mathbb{R}^n$,
find K and γ such that

$$\|x(t)\| \leq Ke^{-\gamma t}\|x(0)\|$$

for any $x(0) \in X$.

Stability Measure 2: Invariant Regions

Long-Range Weather Forecasting and the Lorentz Attractor



A model of atmospheric convection analyzed by E.N. Lorenz, 1963.

$$\dot{x} = \sigma(y - x) \qquad \dot{y} = rx - y - xz \qquad \dot{z} = xy - bz$$

Problem 2: Show that all trajectories converge to a set X.

An oscillating circuit model:

$$\dot{y} = -x - (x^2 - 1)y$$
$$\dot{x} = y$$

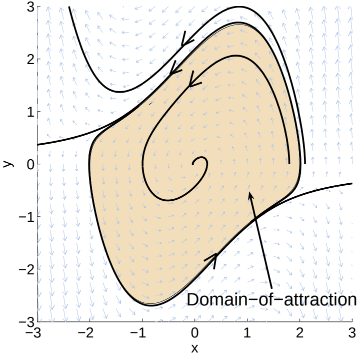


Figure : The van der Pol oscillator in reverse

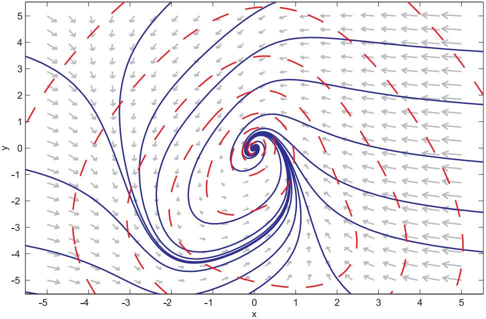
Theorem 1 (Poincaré-Bendixson).

Invariant sets in $\mathbb{R}^2$ always contain a limit cycle or fixed point.

Consider

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

with $x(0) \in X$.



Theorem 2 (Lyapunov Stability).

Suppose there exists a continuous V and $\alpha, \beta, \gamma > 0$ where

$$\beta \|x\|^2 \leq V(x) \leq \alpha \|x\|^2$$
$$\dot{V}(x) = \nabla V(x)^T f(x) \leq -\gamma \|x\|^2$$

for all $x \in X$. Then any sub-level set of V in X is a **Domain of Attraction**.

The Stability Problem is Convex

Convex Optimization of Functions: Variables $V \in \mathcal{C}[\mathbb{R}^n]$ and $\gamma \in \mathbb{R}$

$$\max_{V,\gamma} \ \gamma$$

subject to

$$V(x) - x^T x \geq 0 \qquad \forall x$$
$$\nabla V(x)^T f(x) + \gamma x^T x \leq 0 \qquad \forall x$$

The problem is *finite-dimensional* if $V(x)$ is *polynomial* of bounded degree.

Convex Optimization of Polynomials: Variables $c \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$

$$\max_{c,\gamma} \ \gamma$$

subject to

$$c^T Z(x) - x^T x \geq 0 \qquad \forall x$$
$$c^T \nabla Z(x) f(x) + \gamma x^T x \leq 0 \qquad \forall x$$

- $Z(x)$ is a fixed vector of monomial bases.

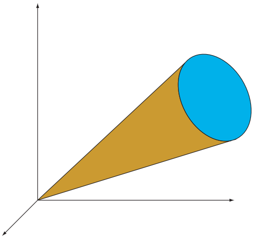
Is Nonlinear Stability Analysis Tractable or Intractable?

The Answer lies in Convex Optimization

Problem:

$$\max_{x} \; bx$$

$$\text{subject to} \quad Ax \in C$$



The problem is *convex optimization* if

- C is a convex cone.
- b and A are affine.

Computational Tractability: Convex Optimization over C is tractable if

- The set membership test for $y \in C$ is in P (polynomial-time verifiable).
- The variable x is a finite dimensional vector (e.g. $\mathbb{R}^n$).

Linear Matrix Inequality (LMI) Format:

$$\text{Find } P:$$
$$\sum_i A_i P B_i + Q > 0$$

Key Concept: System Performance is captured by Lyapunov Functions.

Lemma 3 (KYP Lemma).

For a state-space system
$$\hat{G}(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right],$$
the following are equivalent.

- $\|\hat{G}\|_{H_\infty} \le \gamma$.
- *There exists a* $P > 0$ *such that*
$$\begin{bmatrix} A^T P + PA & PB \\ B^T P & -\gamma I \end{bmatrix} + \frac{1}{\gamma}\begin{bmatrix} C^T \\ D^T \end{bmatrix}\begin{bmatrix} C & D \end{bmatrix} < 0$$

The Lyapunov $V(x) = x^T P x$ proves that $\|Gu\|_{L_2} \le \gamma \|u\|_{L_2}$.

LMI's and Non-Quadratic Lyapunov Functions

Examples of problems with explicit LMI formulations:

- H_∞-optimal control
- H_2-optimal Control (LQG/LQR)
- Quadratic Stability of systems with bounded and polytopic uncertainty

The key is that **ANY** quadratic Lyapunov function can be represented as

$$V(x) = x^T P x$$

Positive matrices can also parameterize non-quadratic Lyapunov functions:

$$V(x) = Z(x)^T P Z(x)$$

is positive if $P > 0$. ($Z(x)$ can be any vector of functions)

- Such a function is called Sum-of-Squares (SOS), denoted $V \in \Sigma_s$.

Question: Can **ANY** Lyapunov function be represented this way?

Polynomial Lyapunov Functions: A Converse Result

Theorem 4 (Peet, TAC 2009).

- *Suppose* $\dot{x}(t) = f(x(t))$ *is exponentially stable for* $\|x(0)\| \leq r$.
- *Suppose* $D^\alpha f$ *is continuous for* $\|\alpha\|_\infty \leq 3$.

Then there exists a Lyapunov function $V : \mathbb{R}^n \to \mathbb{R}$ *such that*

- V *is exponentially decreasing on* $\|x\| \leq r$.
- V *is a* **polynomial**.

Question: Can we make this Quantitative?

A Bound on the Complexity of Lyapunov Functions

Theorem 5 (Peet and Papachristodoulou, TAC, 2012).

- *Suppose* $\|x(t)\| \le K\|x(0)\|e^{-\lambda t}$ *for* $\|x(0)\| \le r$.
- *Suppose* f **is polynomial** *and* $\|\nabla f(x)\| \le L$ *on* $\|x\| \le r$.

Then there exists a **SOS polynomial** $V \in \Sigma_s$ *such that*

- V *is exponentially decreasing on* $\|x\| \le r$.
- *The degree of* V *is less than*

$$\textit{degree}(V) \le 2q^{2(Nk-1)} \cong 2q^{2c_1\frac{L}{\lambda}}$$

where q *is the degree of the vector field,* f.

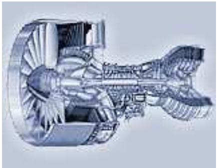
Conclusion: We can assume **ANY** Lyapunov function is of form

$$V(x) = Z_d(x)PZ_d(x)$$

where $P > 0$ and $Z_d(x)$ is the vector of monomials of degree $d \le q^{2c_1\frac{L}{\lambda}}$.

An Example of Global Stability Analysis

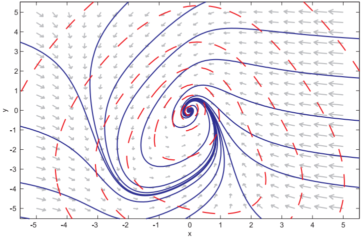
A controlled model of a jet engine (Derived from Moore-Greitzer).



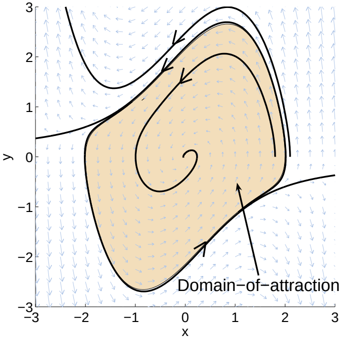
$$\dot{x} = -y - \frac{3}{2}x^2 - \frac{1}{2}x^3$$
$$\dot{y} = 3x - y$$

This is feasible with

$$V(x) = 4.5819x^2 - 1.5786xy + 1.7834y^2 - 0.12739x^3 + 2.5189x^2y - 0.34069xy^2 + 0.61188y^3 + 0.47537x^4 - 0.052424x^3y + 0.44289x^2y^2 + 0.090723y^4$$



Problem: Local Stability Analysis



Problem: Find a polynomial V such that

$$V(x) \geq \alpha \|x\|^2 \qquad \text{for } x \in X,$$

and

$$\nabla V(x)^T f(x) \leq -\gamma \|x\|^2 \qquad \text{for } x \in X,$$

where

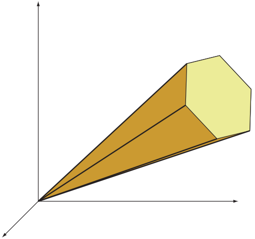
$$X := \left\{ x : \begin{matrix} p_i(x) \geq 0 & i = 1, \ldots, k \\ q_j(x) = 0 & j = 1, \ldots, m \end{matrix} \right\}$$

Optimization of Polynomials

Problem:

$$\max\ b^T x$$

$$\text{subject to } A_0(y) + \sum_i^n x_i A_i(y) \succeq 0 \quad \forall y$$



The A_i are matrices of polynomials in y. e.g. Using multi-index notation,

$$A_i(y) = \sum_\alpha A_{i,\alpha}\, y^\alpha$$

Computationally Intractable

The problem: "Is $p(x) \geq 0$ for all $x \in \mathbb{R}^n$?" (i.e. "$p \in \mathbb{R}^+[x]$?") is NP-hard.

Optimization of Polynomials:

Examples

In fact, many important problems can be cast as optimization of polynomials.

- Global Stability of Nonlinear Systems

$$f(y)^T \nabla p(y) < 0$$
$$p(y) > 0$$

- Matrix Copositivity

$$y^T M y - g(y)^T y \geq 0$$
$$g(y) \geq 0$$

- Integer Programming

$$\max \gamma$$
$$p_0(y)(\gamma - f(y)) - (\gamma - f(y))^2 + \sum_{i=1}^{n} p_i(y)(y_i^2 - 1) \geq 0$$
$$p_0(y) \geq 0$$

- Structured Singular Value (μ)

Global Positivity

Definition 6.

A Polynomial, f, is called Positive SemiDefinite (PSD) if

$$f(x) \geq 0 \quad \text{for all } x \in \mathbb{R}^n$$

Problem: How to prove that $f(x) \geq 0$ for all x?
More generally, the **Primary** problem is

$$\gamma^* = \max_{\gamma} \gamma : \qquad \gamma \leq f(x) \quad \text{for all } x \in \mathbb{R}^n$$

Then if $\gamma^* \geq 0$, f is Positive SemiDefinite.
An **Alternative** problem is

$$\min_{(\sigma,x)} \sigma : \qquad f(x) \leq \sigma \quad \text{for some } x \in \mathbb{R}^n$$

These are **Strong Alternatives**.

- They are also both NP-hard.
- If $\gamma^* = \sigma^* \geq 0$, the function f is PSD.

Global Positivity Certificates (Proofs and Counterexamples)

It is easy to identify a primal **Certificate of Infeasibility** (A Proof that f is NOT PSD). i.e. To show that

$$f(x) \geq 0 \qquad \text{for all } x \in \mathbb{R}^n$$

is FALSE, we need only find a point x with $f(x) < 0$.

It is much harder to identify a **Certificate of Feasibility** (A Proof that f is PSD).

Question: How does one prove that $f(x)$ is positive semidefinite?

What Kind of Functions do we Know are PSD?

- Any squared function is positive.
- The sum of squared forms is PSD
- The product of squared forms is PSD
- The ratio of squared forms is PSD

But is any PSD polynomial the sum, product, or ratio of squared polynomials?

- An old Question....

Definition 7.

A polynomial, $p(x) \in \mathbb{R}[x]$ is a **Sum-of-Squares (SOS)**, denoted $p \in \Sigma_s$ if there exist polynomials $g_i(x) \in \mathbb{R}[x]$ such that

$$p(x) = \sum_i^k g_i(x)^2.$$

David Hilbert created a famous list of 23 then-unsolved mathematical problems in 1900.

- Only 10 have been fully resolved.
- The 17th problem has been resolved.

"Given a multivariate polynomial that takes only non-negative values over the reals, can it be represented as a sum of squares of rational functions?"   -D. Hilbert, 1900

Hilbert's 17th was resolved in the affirmative by E. Artin in 1927.

- Any PSD polynomial is the sum, product and ratio of squared polynomials.
- If $p(x) \geq 0$ for all $x \in \mathbb{R}^n$, then

$$p(x) = \frac{g(x)}{h(x)}$$

  where $g, h \in \Sigma_s$.
- If p is positive definite, then we can assume $h(x) = (\sum_i x_i^2)^d$ for some d. That is,

$$(x_1^2 + \cdots + x_n^2)^d p(x) \in \Sigma_s$$

- If we can't find a SOS representation (certificate) for $p(x)$, we can try $(\sum_i x_i^2)^d p(x)$ for higher powers of d.

Of course this doesn't answer the question of how we find SOS representations.

Polynomial Representation - Linear

First consider the question of representation of polynomials.

- The set of polynomials is an infinite-dimensional vector space.
- The set of polynomials of degree d or less is a finite-dimensional subspace.
  - The monomials are a simple basis for the space of polynomials

Definition 8.

Define $Z_d(x)$ to be the vector of monomial bases of degree d or less.

e.g., if $x \in \mathbb{R}^2$, then

$$Z_2(x_1, x_2)^T = \begin{bmatrix} 1 & x_1 & x_2 & x_1 x_2 & x_1^2 & x_2^2 \end{bmatrix}$$

and

$$Z_4(x_1)^T = \begin{bmatrix} 1 & x_1 & x_1^2 & x_2^3 & x_1^4 \end{bmatrix}$$

Linear Representation

- Any polynomial of degree d can be represented with a vector $c \in \mathbb{R}^m$

$$p(x) = c^T Z_d(x)$$

- This representation is unique.

$$2x_1^2 + 6x_1x_2 + 4x_2 + 1 = \begin{bmatrix} 1 & 0 & 4 & 6 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & x_1 & x_2 & x_1 x_2 & x_1^2 & x_2^2 \end{bmatrix}^T$$

Polynomial Representation - Quadratic

Quadratic Representation

- Alternatively, a polynomial of degree d can be represented b a matrix $M \in \mathbb{S}^m$
$$p(x) = Z_d(x)^T M Z_d(x)$$
- However, now the problem may be under-determined

$$\begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} M_1 & M_2 & M_3 \\ M_2 & M_4 & M_5 \\ M_3 & M_5 & M_6 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}$$
$$= M_1 x^4 + 2M_2 x^3 y + (2M_3 + M_4) x^2 y^2 + 2M_5 x y^3 + M_6 y^4$$

Thus, there are infinitely many quadratic representations of p. For the polynomial
$$f(x) = 4x^4 + 4x^3 y - 7x^2 y^2 - 2xy^3 + 10y^4,$$
we can use the alternative solution
$$4x^4 + 4x^3 y - 7x^2 y^2 - 2xy^3 + 10y^4$$
$$= M_1 x^4 + 2M_2 x^3 y + (2M_3 + M_4) x^2 y^2 + 2M_5 x y^3 + M_6 y^4$$

Polynomial Representation - Quadratic

For the polynomial

$$f(x) = 4x^4 + 4x^3y - 7x^2y^2 - 2xy^3 + 10y^4,$$

we require

$$\begin{aligned} &4x^4 + 4x^3y - 7x^2y^2 - 2xy^3 + 10y^4 \\ &\qquad = M_1x^4 + 2M_2x^3y + (2M_3 + M_4)x^2y^2 + 2M_5xy^3 + M_6y^4 \end{aligned}$$

Constraint Format:

$$M_1 = 4; \quad 2M_2 = 4; \quad 2M_3 + M_4 = -7; \quad 2M_5 = -2; \quad 10 = M_6.$$

An underdetermined system of linear equations (6 variables, 5 equations).

- This yields a family of quadratic representations, parameterized by λ as

$$4x^4 + 4x^3y - 7x^2y^2 - 2xy^3 + 10y^4 = \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 4 & 2 & -\lambda \\ 2 & -7+2\lambda & -1 \\ -\lambda & -1 & 10 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}$$

  which holds for any $\lambda \in \mathbb{R}$

The Quadratic Representation is important in the case where the matrix is positive semidefinite.

Suppose: $p(x) = Z_d(x)^T Q Z_d(x)$ where $Q > 0$.

- Any positive semidefinite matrix, $Q \geq 0$ has a square root $Q = PP^T$

Hence

$$p(x) = Z_d(x)^T Q Z_d(x) = Z_d(x)^T P P^T Z_d(x).$$

Which yields

$$p(x) = \sum_i \left(\sum_j P_{i,j} Z_{d,j}(x) \right)^2$$

which makes $p \in \Sigma_s$ an SOS polynomial.

Moreover: Any SOS polynomial has a quadratic rep. with a PSD matrix.

Suppose: $p(x) = \sum_i g_i(x)^2$ is degree $2d$ (g_i are degree d).

- Each $g_i(x)$ has a linear representation in the monomials.
$$g_i(x) = c_i^T Z_d(x)$$
- Hence
$$p(x) = \sum_i g_i(x)^2 = \sum_i Z_d(x) c_i c_i^T Z_d(x) = Z_d(x) \left(\sum_i c_i c_i^T \right) Z_d(x)$$
- Each matrix $c_i c_i^T \geq 0$. Hence $Q = \sum_i c_i c_i^T \geq 0$.
- We conclude that if $p \in \Sigma_s$, there is a $Q \geq 0$ with $p(x) = Z_d(x) Q Z_d(x)$.

Lemma 9.

Suppose M is polynomial of degree $2d$. $M \in \Sigma_s$ if and only if there exists some $Q \succeq 0$ such that
$$M(x) = Z_d(x)^T Q Z_d(x).$$

Sum-of-Squares

Thus we can express the search for a SOS certificate of positivity as an LMI.

Take the numerical example

$$4x^4 + 4x^3y - 7x^2y^2 - 2xy^3 + 10y^4$$

The question of an SOS representation is equivalent to

$$\text{Find} \quad M = \begin{bmatrix} M_1 & M_2 & M_3 \\ M_2 & M_4 & M_5 \\ M_3 & M_5 & M_6 \end{bmatrix} \geq 0 \quad \text{such that}$$

$$M_1 = 4; \quad 2M_2 = 4; \quad 2M_3 + M_4 = -7; \quad 2M_5 = -2; \quad M_6 = 10.$$

In fact, this is feasible for

$$M = \begin{bmatrix} 4 & 2 & -6 \\ 2 & 5 & -1 \\ -6 & -1 & 10 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ 2 & 1 & -3 \end{bmatrix}$$

Sum-of-Squares

We can use this solution to construct an SOS certificate of positivity.

$$\begin{aligned}
4x^4 + 4x^3y - 7x^2y^2 - 2xy^3 + 10y^4 &= \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 4 & 2 & -6 \\ 2 & 5 & -1 \\ -6 & -1 & 10 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix} \\
&= \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ 2 & 1 & -3 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix} \\
&= \begin{bmatrix} 2xy + y^2 \\ 2x^2 + xy + 3y^2 \end{bmatrix}^T \begin{bmatrix} 2xy + y^2 \\ 2x^2 + xy + 3y^2 \end{bmatrix} \\
&= (2xy + y^2)^2 + (2x^2 + xy + 3y^2)^2
\end{aligned}$$

SOS Programming:

Numerical Example

This also works for matrix-valued polynomials.

$$M(y,z) = \begin{bmatrix} (y^2+1)z^2 & yz \\ yz & y^4+y^2-2y+1 \end{bmatrix}$$

$$\begin{aligned}
\begin{bmatrix} (y^2+1)z^2 & yz \\ yz & y^4+y^2-2y+1 \end{bmatrix} &= \begin{bmatrix} z & 0 \\ yz & 0 \\ 0 & 1 \\ 0 & y \\ 0 & y^2 \end{bmatrix}^T \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & -1 & -1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z & 0 \\ yz & 0 \\ 0 & 1 \\ 0 & y \\ 0 & y^2 \end{bmatrix} \\
&= \begin{bmatrix} z & 0 \\ yz & 0 \\ 0 & 1 \\ 0 & y \\ 0 & y^2 \end{bmatrix}^T \begin{bmatrix} 0 & 1 & 1 & -1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 0 & 1 & 1 & -1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z & 0 \\ yz & 0 \\ 0 & 1 \\ 0 & y \\ 0 & y^2 \end{bmatrix} \\
&= \begin{bmatrix} yz & 1-y \\ z & y^2 \end{bmatrix}^T \begin{bmatrix} yz & 1-y \\ z & y^2 \end{bmatrix} \in \Sigma_s
\end{aligned}$$

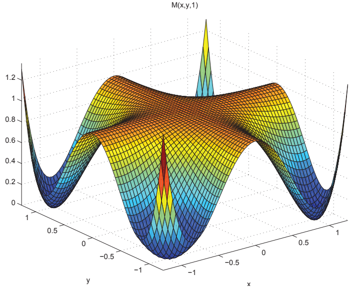
Problems with SOS

Unfortunately, a Sum-of-Squares representation is not necessary for positivity.

- Artin included ratios of squares.

Counterexample: The Motzkin Polynomial

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



However, $(x^2+y^2+1)M(x,y)$ is a Sum-of-Squares.

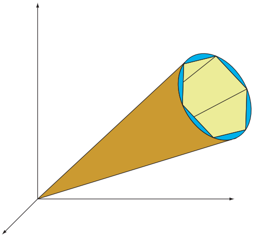
$$\begin{aligned}(x^2+y^2+1)M(x,y) &= (x^2y-y)^2 + (xy^3-x)^2 + (x^2y^2-1)^2 \\ &+ \frac{1}{4}(xy^3-x^3y)^2 + \frac{3}{4}(xy^3+x^3y-2xy)^2\end{aligned}$$

Sum-of-Squares (SOS) Programming

Problem:

$$\max\ b^T x$$

$$\text{subject to} \quad A_0(y) + \sum_i^n x_i A_i(y) \in \Sigma_s$$



Definition 10.

$\Sigma_s \subset \mathbb{R}^+[x]$ is the cone of *sum-of-squares* matrices. If $S \in \Sigma_s$, then for some $G_i \in \mathbb{R}[x]$,

$$S(y) = \sum_{i=1}^{r} G_i(y)^T G_i(y)$$

Computationally Tractable: $S \in \Sigma_s$ is an SDP constraint.

Global Stability Analysis

At this point, we can express the problem of global stability of a polynomial vector field as an LMI.

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

$$\max_{c,\gamma} \ \gamma$$

subject to

$$V(x) - \|x\|^2 = c^T Z(x) - x^T x \in \Sigma_s$$
$$\dot{V}(x) + \gamma\|x\|^2 = c^T \nabla Z(x) f(x) + \gamma x^T x \in \Sigma_s$$

If this is feasible with $\gamma > 0$, then

$$V(x) \geq \|x\|^2 \qquad \text{and} \qquad \dot{V}(x) \leq -\gamma\|x\|^2$$

which implies the system is globally exponentially stable.

- The process can be automated (See SOSTOOLS)
- Creates a set of equality constraints between c and a PSD matrix,

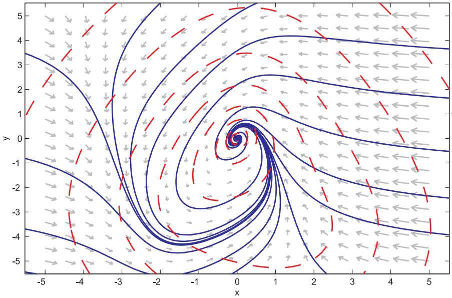
An Example of Global Stability Analysis

A controlled model of a jet engine.

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

This is feasible with

$$V(x) = 4.5819x^2 - 1.5786xy + 1.7834y^2 - 0.12739x^3 + 2.5189x^2y - 0.34069xy^2$$
$$+ 0.61188y^3 + 0.47537x^4 - 0.052424x^3y + 0.44289x^2y^2 + 0.090723y^4$$



Problems with SOS

The problem is that most nonlinear stability problems are **local**.

- Global stability requires a unique equilibrium.
- Very few nonlinear systems are globally stable.

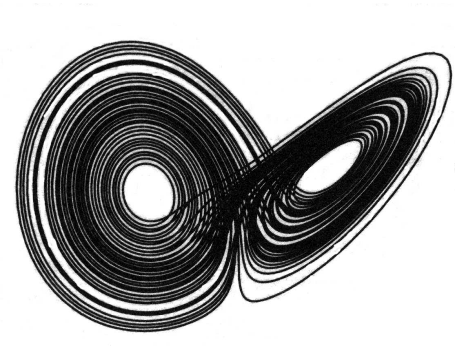


Figure : The Lorentz Attractor

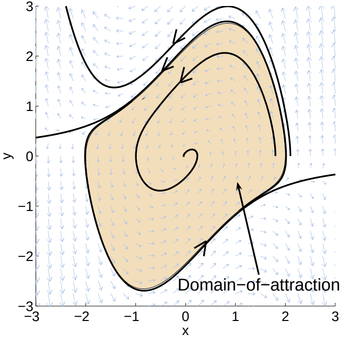


Figure : The van der Pol oscillator in reverse

Local Positivity

A more interesting question is the question of local positivity.
Question: Is $y(x) \geq 0$ for $x \in X$, where $X \subset \mathbb{R}^n$.

Examples:

- Matrix Copositivity:

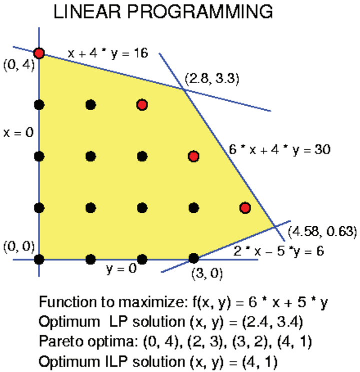
$$y^T M y \geq 0 \qquad \text{for all } y \geq 0$$

- Integer Programming (Upper bounds)

$$\min \gamma$$
$$\gamma \geq f_i(y)$$
$$\text{for all } y \in \{-1, 1\}^n \text{ and } i = 1, \cdots, k$$

- Local Lyapunov Stability

$$V(x) \geq \|x\|^2 \qquad \text{for all } \|x\| \leq 1$$
$$\nabla V(x)^T f(x) \leq 0 \qquad \text{for all } \|x\| \leq 1$$



All these sets are **Semialgebraic**.

Semialgebraic Sets

The first step is to unify the representation of sets.

Definition 11.

A set $X \subset \mathbb{R}^n$ is **Semialgebraic** if it can be represented using polynomial equality and inequality constraints.

$$X := \left\{ x : \begin{matrix} p_i(x) \geq 0 & i = 1, \ldots, k \\ q_j(x) = 0 & j = 1, \ldots, m \end{matrix} \right\}$$

If there are only equality constraints, the set is **Algebraic**.

Note: A semialgebraic set can also include $\neq$ and $<$.

Integer Programming

$$\{-1, 1\}^n = \{y \in \mathbb{R}^n \, : \, y_i^2 - 1 = 0\}$$

Local Lyapunov Stability

$$\{x \, : \, \|x\| \leq 1\} = \{x \, : \, 1 - x^T x \geq 0\}$$

Not that the *representation* of a semialgebraic set is **NOT UNIQUE**.

- Some representations are better than others...

Feasibility of a Semialgebraic Set

Positivity of f on X is equivalent to a positive solution of either the Primary or Alternative problems.

The **Primary** problem is

$$\max_{\gamma} \gamma :$$
$$\gamma < f(x) \quad \text{for } \textit{all } x \in X$$

The **Alternative** problem is

$$\min_{\sigma,x} \sigma :$$
$$f(x) \le \sigma \quad \text{for } \textit{some } x \in X$$

Define $\qquad S_\gamma = \{x : \gamma \ge f(x), x \in X\}$ (A Semialgebraic Set)

The **Primary** problem is

$$\max \gamma :$$
$$S_\gamma = \emptyset$$

The **Alternative** problem is

$$\min \sigma :$$
$$S_\sigma \ne \emptyset$$

The feasibility problem is the difficult part.

Positivity and Feasibility

Positivity and Feasibility are the same problem...:

Lemma 12.

$f(x) \geq 0$ *for all* $x \in \mathbb{R}^n$ *if and only if*

$$Y = \{x \,:\, -f(x) > 0\} = \emptyset$$

Now, let

$$X := \left\{ x : \begin{matrix} p_i(x) \geq 0 & i = 1, \ldots, k \\ q_j(x) = 0 & j = 1, \ldots, m \end{matrix} \right\}$$

$f(x) \geq 0$ *for all* $x \in X$ *if and only if*

$$Y = \{x \,:\, -f(x) > 0,\, p_i(x) \geq 0,\, q_j(x) = 0\} = \emptyset$$

Problem: How to test if there exists an $y \in Y$ (Feasibility of Y)???

Feasibility of a Semialgebraic Set

Solving the feasibility problem yields a solution to the optimization problem.
Convex Optimization: For convex functions f_i,

$$\min\ f_0(x) :$$
$$\text{subject to} \quad f_i(x) \le 0, \quad h_j(x) = 0$$

Geometric Representation: Let $S := \{x \in \mathbb{R}^n \,:\, f_i(x) \le 0, \quad h_j(x) = 0\}$

$$\min\ f_0(x) :$$
$$\text{subject to} \quad x \in S$$

Feasibility Representation: For γ, define

$$S_\gamma := \{x \in \mathbb{R}^n \,:\, f_0(x) \le \gamma,\, f_i(x) \le 0,\, h_j(x) = 0\}$$

$$\min\ \gamma :$$
$$\text{subject to } S_\gamma \ne \emptyset$$

Given an efficient test for feasibility (Is $S_\gamma = \emptyset$?), **Bisection** will solve the Optimization Problem.

Optimization Problem:

$$\gamma^* = \min_{\gamma} \; \gamma :$$
$$\text{subject to } S_\gamma \neq \emptyset$$

Bisection Algorithm:

1 Initialize feasible $\gamma_u = b$

2 Initialize infeasible $\gamma_l = a$

3 Set $\gamma = \frac{\gamma_u + \gamma_l}{2}$

5 If S_γ feasible, set $\gamma_u = \frac{\gamma_u + \gamma_l}{2}$

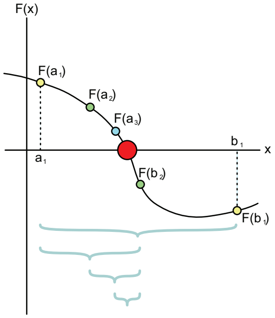
4 If S_γ infeasible, set $\gamma_l = \frac{\gamma_u + \gamma_l}{2}$

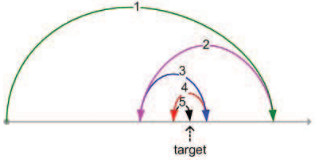
6 $k = k + 1$

7 Goto 3

Then $\gamma^* \in [\gamma_l, \gamma_u]$ and $|\gamma_u - \gamma_l| \leq \frac{b-a}{2^k}$.

Bisection with oracle also solves the Primary Problem. ($\max \gamma : \; S_\gamma = \emptyset$)





A Problem of Representation and Inference

As with SOS and PSD Functions, Testing Feasibility is a Question of Representation.

Consider how to represent a semialgebraic set:
Example: A representation of the interval $X = [a, b]$.

- A first order representation:

$$\{x \in \mathbb{R} : x - a \geq 0,\, b - x \geq 0\}$$

- A quadratic representation:

$$\{x \in \mathbb{R} : (x - a)(b - x) \geq 0\}$$

- We can add arbitrary polynomials which are PSD on X to the representation.

$$\{x \in \mathbb{R} : (x - a)(b - x) \geq 0,\, x - a \geq 0\}$$
$$\{x \in \mathbb{R} : (x^2 + 1)(x - a)(b - x) \geq 0\}$$
$$\{x \in \mathbb{R} : (x - a)(b - x) \geq 0,\, (x^2 + 1)(x - a)(b - x) \geq 0,\, (x - a)(b - x) \geq 0\}$$

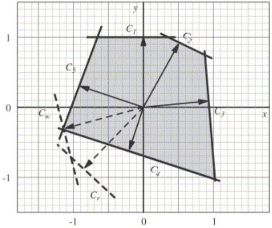
A Problem of Representation and Inference

Why are all these representations valid?

- We are adding redundant constraints to the set.
- $x - a \geq 0$ and $b - x \geq 0$ for $x \in [a, b]$ implies

$$(x - a)(b - x) \geq 0.$$

- $x^2 + 1$ is SOS, so is obviously positive on $x \in [a, b]$.



How are we creating these redundant constraints?

- **Logical Inference**
- Using existing polynomials which are positive on X to create new ones.

Big Questions:

- Can ANY polynomial which is positive on $[a, b]$ can be constructed this way?
- Given f, can we use inference to prove that $f(x) \geq 0$ for any $x \in X = [a, b]$?

The Cone of Inference

Definition 13.

Given a semialgebraic set S, a function f is called a **valid inequality** on S if

$$f(x) \geq 0 \qquad \text{for all } x \in S$$

Question: How to construct valid inequalities?

- Closed under addition: If f_1 and f_2 are valid, then $h(x) = f_1(x) + f_2(x)$ is valid
- Closed under multiplication: If f_1 and f_2 are valid, then $h(x) = f_1(x)f_2(x)$ is valid
- Contains all Squares: $h(x) = g(x)^2$ is valid for ANY polynomial g.

A set of inferences constructed in such a manner is called a cone.

The Cone of Inference

Definition 14.

The set of polynomials $C \subset \mathbb{R}[x]$ is called a **Cone** if

- $f_1 \in C$ and $f_2 \in C$ implies $f_1 + f_2 \in C$.
- $f_1 \in C$ and $f_2 \in C$ implies $f_1 f_2 \in C$.
- $\Sigma_s \subset C$.

Note: this is **NOT** the same definition as in optimization.

The Cone of Inference

The set of inferences is a cone

Definition 15.

For any set, S, the cone $C(S)$ is the set of polynomials PSD on S

$$C(S) := \{f \in \mathbb{R}[x] \,:\, f(x) \geq 0 \text{ for all } x \in S\}$$

The big question: how to test $f \in C(S)$???

Corollary 16.

$f(x) \geq 0$ for all $x \in S$ if and only if $f \in C(S)$

The Monoid

Suppose S is a semialgebraic set and define its *monoid*.

Definition 17.

For given polynomials $\{f_i\} \subset \mathbb{R}[x]$, we define $\texttt{monoid}(\{f_i\})$ as the set of all products of the f_i

$$\texttt{monoid}(\{f_i\}) := \{h \in \mathbb{R}[x] : h(x) = \prod f_1^{a_1}(x) f_2^{a_k}(x) \cdots f_k^{a_2}(x),\, a \in \mathbb{N}^k\}$$

- $1 \in \texttt{monoid}(\{f_i\})$
- $\texttt{monoid}(\{f_i\})$ is a subset of the cone defined by the f_i.
- The monoid does not include arbitrary sums of squares

The Cone of Inference

If we combine $\texttt{monoid}(\{f_i\})$ with Σ_s, we get $\texttt{cone}(\{f_i\})$.

Definition 18.

For given polynomials $\{f_i\} \subset \mathbb{R}[x]$, we define $\texttt{cone}(\{f_i\})$ as

$$\texttt{cone}(\{f_i\}) := \{h \in \mathbb{R}[x] : h = \sum s_i g_i,\, g_i \in \texttt{monoid}(\{f_i\}),\, s_i \in \Sigma_s\}$$

If

$$S := \{x \in \mathbb{R}^n \,:\, f_i(x) \geq 0,\, i = 1 \cdots, k\}$$

$\texttt{cone}(\{f_i\}) \subset C(S)$ is an approximation to $C(S)$.

- The key is that it is possible to test whether $f \in \texttt{cone}(\{f_i\}) \subset C(S)$!!!
    - Sort of... (need a degree bound)
    - Use e.g. SOSTOOLS

More on Inference

Corollary 19.

$h \in \textit{cone}(\{f_i\}) \subset C(S)$ if and only if there exist $s_i, r_{ij}, \cdots \in \Sigma_s$ such that

$$h(x) = s_0 + \sum_i s_i f_i + \sum_{i \neq j} r_{ij} f_i f_j + \sum_{i \neq j \neq k} r_{ijk} f_i f_j f_k + \cdots$$

Note we must include all possible combinations of the f_i

- A finite number of variables s_i, r_{ij}.
- $s_i, r_{ij} \in \Sigma_s$ is an SDP constraint.
- The equality constraint acts on the coefficients of f, s_i, r_{ij}.

This gives a sufficient condition for $h(x) \geq 0$ for all $x \in S$.

- Can be tested using, e.g. SOSTOOLS

Numerical Example

Example: To show that $h(x) = 5x - 9x^2 + 5x^3 - x^4$ is PSD on the interval $[0, 1] = \{x \in \mathbb{R}^n : x(1-x) \ge 0\}$, we use $f_1(x) = x(1-x)$. This yields the constraint

$$h(x) = s_0(x) + x(1-x)s_1(x)$$

We find $s_0(x) = 0$, $s_1(x) = (2-x)^2 + 1$ so that

$$5x - 9x^2 + 5x^3 - x^4 = 0 + ((2-x)^2 + 1)x(1-x)$$

Which is a certificate of non-negativity of h on $S = [0, 1]$

Note: the original representation of S matters:

- If we had used $S = \{x \in \mathbb{R} : x \ge 0,\ 1 - x \ge 0\}$, then we would have had 4 SOS variables

$$h(x) = s_0(x) + xs_1(x) + (1-x)s_2(x) + x(1-x)s_3(x)$$

The complexity can be *decreased* through judicious choice of representation.

Stengle's Positivstellensatz

We have two big questions

- How close an approximation is $\texttt{cone}(\{f_i\}) \subset C(S)$ to $C(S)$?
  - Cannot always be exact since not every positive polynomial is SOS.
- Can we reduce the complexity?

Both these questions are answered by *Positivstellensatz* Results. Recall

$$S := \{x \in \mathbb{R}^n \,:\, f_i(x) \geq 0,\, i = 1 \cdots, k\}$$

Theorem 20 (Stengle's Positivstellensatz).

$S = \emptyset$ if and only if $-1 \in \textit{cone}(\{f_i\})$. That is, $S = \emptyset$ if and only if there exist $s_i, r_{ij}, \cdots \in \Sigma_s$ such that

$$-1 = s_0 + \sum_i s_i f_i + \sum_{i \neq j} r_{ij} f_i f_j + \sum_{i \neq j \neq k} r_{ijk} f_i f_j f_k + \cdots$$

Note that this is not exactly what we were asking.

- We would prefer to know whether $h \in \texttt{cone}(\{f_i\})$
- Difference is important for reasons of convexity.

Stengle's Positivstellensatz

Problem: We want to know whether $f(x) > 0$ for all $x \in \{x : g_i(x) \geq 0\}$.

Corollary 21.

$f(x) > 0$ for all $x \in \{x : g_i(x) \geq 0\}$ if and only if there exist $s_i, q_{ij}, r_{ij}, \cdots \in \Sigma_s$ such that

$$f\left(s_{-1} + \sum_i q_i g_i + \sum_{i \neq j} q_{ij} g_i g_j + \sum_{i \neq j \neq k} q_{ijk} g_i g_j g_k + \cdots\right)$$
$$= 1 + s_0 + \sum_i s_i g_i + \sum_{i \neq j} r_{ij} g_i g_j + \sum_{i \neq j \neq k} r_{ijk} g_i g_j g_k + \cdots$$

Proof.

$f(x) > 0$ for all $x \in \{x : g_i(x) \geq 0\}$ is equivalent to infeasibility of

$$S := \{x : -f(x) \geq 0,\ g_i(x) \geq 0\}$$

By applying the Positivstellensatz, we obtain the conditions. □

Stengle's Positivstellensatz

$$f\left(s_{-1}+\sum_i q_i g_i+\sum_{i\neq j} q_{ij} g_i g_j+\sum_{i\neq j\neq k} q_{ijk} g_i g_j g_k+\cdots\right)$$
$$=1+s_0+\sum_i s_i g_i+\sum_{i\neq j} r_{ij} g_i g_j+\sum_{i\neq j\neq k} r_{ijk} g_i g_j g_k+\cdots$$

The Good:

- Now possible to test whether $f(x)>0$ for $x\in S$
- No special conditions on S or f.

The Bad:

- Condition is bilinear in the f and s_i, q_{ij}.
  - Need a specialized solver.
- Strict Positivity
  - $f(x)>\epsilon$ for some $\epsilon>0$ and all $x\in S$
  - Actually, Stengle's Positivstellensatz has a weak form.
- Lots of variables.

Stengle's "Weak" Positivstellensatz

$$S := \{x \in \mathbb{R}^n : f_i(x) \geq 0,\ i = 1, \ldots, n_K\}.$$

Theorem 22.

Let S be given as above and let I denote the set of subsets of $\{0, \ldots, n_K\}$. Then $f_0(x) \geq 0$ for all $x \in S$ if and only if there exist $s \in \Sigma_s$, $s_J \in \Sigma_s$ for $J \in I$ and $k \in \mathbb{Z}^+$ such that

$$s(x) + \sum_{J \in I} s_J(x) \prod_{i \in J} f_i(x) + f_0(x)^{2k} = 0$$

Condition:

$$f_0 \left(s_{-1} + \sum_i q_i g_i + \sum_{i \neq j} q_{ij} g_i g_j + \sum_{i \neq j \neq k} q_{ijk} g_i g_j g_k + \cdots \right)$$
$$= f_0(x)^{2k} + s_0 + \sum_i s_i g_i + \sum_{i \neq j} r_{ij} g_i g_j + \sum_{i \neq j \neq k} r_{ijk} g_i g_j g_k + \cdots$$

Schmudgen's Positivstellensatz

If the set is compact, then the problem can be convexified.

Theorem 23.

Suppose that $S = \{x : g_i(x) \geq 0\}$ is compact. If $f(x) > 0$ for all $x \in S$, then there exist $s_i, r_{ij}, \cdots \in \Sigma_s$ such that

$$f = s_0 + \sum_i s_i g_i + \sum_{i \neq j} r_{ij} g_i g_j + \sum_{i \neq j \neq k} r_{ijk} g_i g_j g_k + \cdots$$

Note that Schmudgen's Positivstellensatz is essentially the same as Stengle's except for a single term.

- Now we can include both f and s_i, r_{ij} as variables.
- Reduces the number of variables substantially.

The complexity is still high (Lots of SOS multipliers).

Putinar's Positivstellensatz

If the semialgebraic set is P-Compact, then we can improve the situation further.

Definition 24.

We say that $f_i \in \mathbb{R}[x]$ for $i = 1, \ldots, n_K$ define a **P-compact** set K_f, if there exist $h \in \mathbb{R}[x]$ and $s_i \in \Sigma_s$ for $i = 0, \ldots, n_K$ such that the level set $\{x \in \mathbb{R}^n : h(x) \geq 0\}$ is compact and such that the following holds.

$$h(x) - \sum_{i=1}^{n_K} s_i(x) f_i(x) \in \Sigma_s$$

The condition that a region be P-compact may be difficult to verify. However, some important special cases include:

- Any region K_f such that all the f_i are linear.
- Any region K_f defined by f_i such that there exists some i for which the level set $\{x : f_i(x) \geq 0\}$ is compact.

Putinar's Positivstellensatz

P-Compact is not hard to satisfy.

Corollary 25.

Any compact set can be made P-compact by inclusion of a redundant constraint of the form $f_i(x) = \beta - x^T x$ *for sufficiently large* β*.*

Thus P-Compact is a property of the *representation* and not the set.

Example: The interval $[a, b]$.

- Not Obviously P-Compact:

$$\{x \in \mathbb{R} : x^2 - a^2 \geq 0,\, b - x \geq 0\}$$

- P-Compact:

$$\{x \in \mathbb{R} : (x - a)(b - x) \geq 0\}$$

If S is P-Compact, Putinar's Positivstellensatz dramatically reduces the complexity

Theorem 26 (Putinar's Positivstellesatz).

Suppose that $S = \{x : g_i(x) \geq 0\}$ *is P-Compact. If* $f(x) > 0$ *for all* $x \in S$*, then there exist* $s_i \in \Sigma_s$ *such that*

$$f = s_0 + \sum_i s_i g_i$$

A single multiplier for each constraint.

Relationship to the S-Procedure

A Classical LMI.
The S-procedure asks the question:

- Is $z^T F z \geq 0$ for all $z \in \{x : x^T G x \geq 0\}$?

Corollary 27 (S-Procedure).

$z^T F z \geq 0$ for all $z \in \{x : x^T G x \geq 0\}$ if there exists a $\tau \geq 0$ such that $F - \tau G \succeq 0$.

The S-procedure is **Necessary** if $\{x : x^T G x > 0\} \neq \emptyset$.

A Special Case of Putinar's Positivstellensatz when

- $f(x) = x^T F x$
- $g(x) = x^T G x$
- $s(x) = \tau$
- we replace ≥ 0 with $\in \Sigma_s$

Robust Analysis and Synthesis

Recall the problem of parametric uncertainty.

$$\dot{x}(t) = A(\alpha)x(t) + B(\alpha)u(t)$$
$$y(t) = C(\alpha)x(t) + D(\alpha)u(t)$$

where $\alpha \in X$ where X is a semialgebraic set.

$$\text{Let} \quad X := \left\{ x : \begin{matrix} p_i(x) \geq 0 & i = 1, \ldots, k \\ q_j(x) = 0 & j = 1, \ldots, m \end{matrix} \right\}$$

Almost any analysis or synthesis condition expressed as an LMI can be made robust using SOS and Positivstellensatz results.

Robust Analysis and Synthesis

Theorem 28.

$\dot{x}(t) = A(\alpha)x(t)$ is stable for all $\alpha \in X$ if there exist $\epsilon > 0$, polynomials T_i and SOS polynomials $P, R, S_i \in \Sigma_s$ such that

$$P(\alpha) - \epsilon I = R(\alpha)$$

$$-\left(A(\alpha)^T P(\alpha) + P(\alpha)A(\alpha)\right) = S_0(\alpha) + \sum_i S_i(\alpha)p_i(\alpha) + \sum_j T_i(\alpha)q_j(\alpha)$$

The difference from the S-procedure is that

- The Lyapunov function is parameter-dependent (Not Quadratic Stability).
- The multipliers are parameter-dependent.

Polynomial Programming

PS results can be applied to the problem of polynomial programming.

Polynomial Optimization

$$c_1^* = \max_y \; d^T y$$

subject to $f_i(y) \ge 0 \quad i = 1, \cdots, k$

This problem is **NOT** convex.

Optimization of Polynomials:

$$c_2^* = \max_x \; b^T x$$

subject to $A_0(y) + \sum_i^n x_i A_i(y) \succeq 0 \quad \forall y$

This Problem is convex (use SOS)

Optimization of polynomials can be used to find the maximum value c_1^*.

$$c_1^* = \min_\gamma \; \gamma$$

subject to $\gamma \ge d^T y \quad \forall y \in \{y \,:\, f_i(y) \ge 0,\, i = 1, \cdots, k\}$

Polynomial Programming

Optimization of polynomials can be used to find the maximum value c_1^*.

$$c_1^* = \min_{\gamma} \; \gamma$$
$$\text{subject to } \gamma \ge d^T y \quad \forall y \in \{y \,:\, f_i(y) \ge 0,\, i = 1, \cdots, k\}$$

Reformulate as

$$c_1^* = \min_{\gamma} \; \gamma \qquad \text{subject to}$$
$$\gamma - d^T y = s_0(y) + \sum_i s_i(y) f_i(y)$$
$$s_i \in \Sigma_s$$

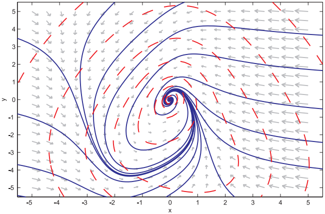
Which is easily solved using SOSTOOLS.

Return to Lyapunov Stability

We can now recast the search for a Lyapunov function.

Let

$$X := \left\{ x : p_i(x) \geq 0 \quad i = 1, \ldots, k \right\}$$



Theorem 29.

Suppose there exists a polynomial v, a constant $\epsilon > 0$, and sum-of-squares polynomials s_0, s_i, t_0, t_i such that

$$v(x) - \sum_i s_i(x) p_i(s) - s_0(s) - \epsilon\, x^T x = 0$$

$$-\nabla v(x)^T f(x) - \sum_i t_i(x) p_i(s) - t_0(x) - \epsilon\, x^T x = 0$$

Then the system is exponentially stable on any $Y_\gamma := \{x : v(x) \leq \gamma\}$ where $Y_\gamma \subset X$.

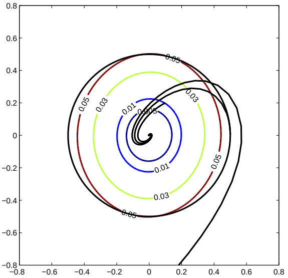
Note: Find the largest Y_γ via bisection.

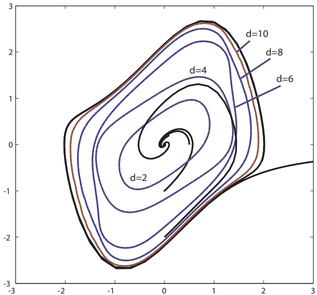
Van-der-Pol Oscillator

$$\dot{x}(t) = -y(t)$$
$$\dot{y}(t) = -\mu(1 - x(t)^2)y(t) + x(t)$$

Procedure:

1. Use Bisection to find the largest ball on which you can find a Lyapunov function.
2. Use Bisection to find the largest level set of that Lyapunov function on which you can find a Lyapunov function. Repeat





Computing the Attractive Region

Advective approaches also work well.

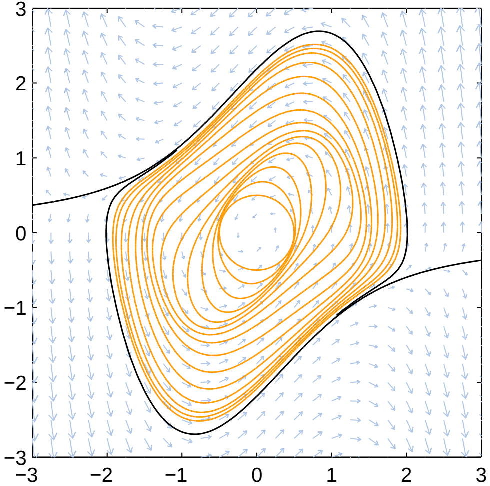


Figure : The van der Pol Oscillator using advection [T.-C. Wang]

Computing the Invariant Region

Numerical Results using the Advective Approach

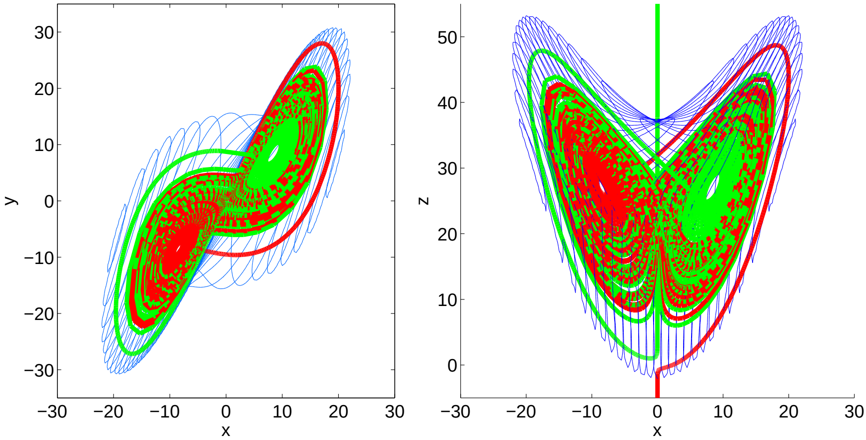


Figure : The Lorentz Attractor using advection [T.C. Wang]

Equality Constraints

Before we conclude, it is worth considering the question of equality constraints:
Integer Programming (Upper bounds)

$$\min \gamma$$
$$\gamma \geq f_i(y)$$
$$\text{for all } y \in \{-1, 1\}^n \text{ and } i = 1, \cdots, k$$

where

$$\{-1, 1\}^n = \{y \in \mathbb{R}^n \,:\, y_i^2 - 1 = 0\}$$

Incompressible Navier-Stokes

$$\frac{\partial v}{\partial t} = -v \cdot \nabla v + \frac{1}{\rho} \left(\nabla \cdot T + f - \nabla p\right)$$

Requires conservation of mass

$$\nabla v = 0$$

Which imposes an algebraic constraint on the state.
Question: How to test whether

$$f(x) \geq 0 \qquad \text{for all } x \in \{x \,:\, g(x) = 0\}?$$

Inference

The approach is the same as for inequality constraints.

- If $g_1(x) = 0$ for $x \in S$ and $g_2(x) = 0$ for $x \in S$, then

$$g_1(x) + g_2(x) = 0 \qquad \text{for all } x \in S$$

- If $g(x) = 0$ for $x \in S$, then for ANY polynomial $h(x)$,

$$g(x)h(x) = 0 \qquad \text{for all } x \in S$$

Let

$$S = \{x : f_i(x) = 0,\, i = 1, \cdots k\}$$

Definition 30.

We say f is a **valid equality** on S if $f(x) = 0$ for $x \in S$

Inference

Definition 31.

A set of polynomials, I is called an **ideal** if

- $f_1 + f_2 \in I$ for all $f_1, f_2 \in I$.
- $fg \in I$ for all $f \in I$ and $g \in \mathbb{R}[x]$.

Define ideal(S) to be the set of valid equalities on S.

Definition 32.

For given polynomials $\{f_i\} \subset \mathbb{R}[x]$, we define $\texttt{ideal}(\{f_i\})$ as

$$\texttt{ideal}(\{f_i\}) := \{h \in \mathbb{R}[x] : h = \sum f_i g_i,\, g_i \in \mathbb{R}[x]\}$$

$\texttt{ideal}(\{f_i\})$ gives a set of valid equalities on

$$S = \{x : f_i(x) = 0,\, i = 1, \cdots k\}$$

$\texttt{ideal}(\{f_i\})$ is the smallest ideal containing $\{f_i\}$.

- How closely does $\texttt{ideal}(\{f_i\})$ approximate ideal(S)?

The Nullstellensatz

Theorem 33 (Hilbert's Nullstellensatz).

For $f_i \in \mathbb{C}[x]$, let

$$S = \{x \in \mathbb{C}^n : f_i(x) = 0,\, i = 1, \cdots k\}$$

Then $S = \emptyset$ if and only if $1 \in \texttt{ideal}(\{f_i\})$

If $1 \in \texttt{ideal}(\{f_i\})$, then feasibility of the set would imply $1 = 0$.

- Predates Positivstellensatz results.
- Not valid over the reals

This gives an algorithmic approach to testing feasibility: Find g_i such that

$$-1 = \sum f_i g_i,\, g_i \in \mathbb{C}[x]\}$$

This involves solving linear equations (no optimization).

Positivstellensatz with Equality

Fortunately, the equality constraint are included in later versions of the Positivstellensatz.

Theorem 34 (Putinar's Positivstellesatz).

Suppose that $S = \{x : g_i(x) \geq 0,\, h_i(x) = 0\}$ *is P-Compact. If* $f(x) > 0$ *for all* $x \in S$*, then there exist* $s_i \in \Sigma_s$ *and* $t_i \in \mathbb{R}[x]$ *such that*

$$f = s_0 + \sum_i s_i g_i + \sum_j t_j h_j$$

Theorem 35 (Schmüdgen's Positivstellesatz).

Suppose that $S = \{x : g_i(x) \geq 0,\, h_i(x) = 0\}$ *is compact. If* $f(x) > 0$ *for all* $x \in S$*, then there exist* $s_i, r_{ij}, \cdots \in \Sigma_s$ *and* $t_i \in \mathbb{R}[x]$ *such that*

$$f = 1 + \sum_j t_j h_j + s_0 + \sum_i s_i g_i + \sum_{i \neq j} r_{ij} g_i g_j + \sum_{i \neq j \neq k} r_{ijk} g_i g_j g_k + \cdots$$

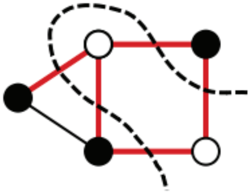


Figure : Division of a set of nodes to maximize the weighted cost of separation

Goal: Assign each node i an index $x_i = -1$ or $x_j = 1$ to maximize overall cost.

- The cost if x_i and x_j do not share the same index is w_{ij}.
- The cost if they share an index is 0
- The weight $w_{i,j}$ are given.
- Thus the total cost is

$$\frac{1}{2}\sum_{i,j} w_{i,j}(1 - x_i x_j)$$

MAX-CUT

The optimization problem is the integer program:

$$\max_{x_i^2=1} \frac{1}{2} \sum_{i,j} w_{i,j}(1 - x_i x_j)$$

The MAX-CUT problem can be reformulated as

$$\begin{aligned} &\min \gamma : \\ &\gamma \geq \max_{x_i^2=1} \frac{1}{2} \sum_{i,j} w_{i,j}(1 - x_i x_j) \quad \text{for all} \quad x \in \{x \, : \, x_i^2 = 1\} \end{aligned}$$

We can compute a bound on the max cost using the Nullstellensatz

$$\begin{aligned} &\min_{p_i \in \mathbb{R}[x],\, s_0 \in \Sigma_s} \gamma : \\ &\gamma - \frac{1}{2} \sum_{i,j} w_{i,j}(1 - x_i x_j) + \sum_i p_i(x)(x_i^2 - 1) = s_0(x) \end{aligned}$$

MAX-CUT

Consider the MAX-CUT problem with 5 nodes

$$w_{12} = w_{23} = w_{45} = w_{15} = .5 \quad \text{and} \quad w_{14} = w_{24} = w_{25} = w_{34} = 0$$

where $w_{ij} = w_{ji}$. The objective function is

$$f(x) = 2.5 - .5x_1x_2 - .5x_2x_3 - .5x_3x_4 - .5x_4x_5 - .5x_1x_5$$

We use SOSTOOLS and bisection on γ to solve

$$\min_{p_i \in \mathbb{R}[x],\, s_0 \in \Sigma_s} \gamma :$$

$$\gamma - f(x) + \sum_i p_i(x)(x_i^2 - 1) = s_0(x)$$

We achieve a least upper bound of $\gamma = 4$.
However!

- we don't know if the optimization problem achieves this objective.
- Even if it did, we could not recover the values of $x_i \in [-1, 1]$.

MAX-CUT

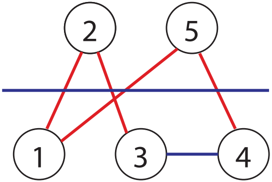


Figure : A Proposed Cut

Upper bounds can be used to VERIFY optimality of a cut.
We Propose the Cut

- $x_1 = x_3 = x_4 = 1$
- $x_2 = x_5 = -1$

This cut has objective value

$$f(x) = 2.5 - .5x_1x_2 - .5x_2x_3 - .5x_3x_4 - .5x_4x_5 - .5x_1x_5 = 4$$

Thus verifying that the cut is optimal.

MAX-CUT code

```
 pvar x1 x2 x3 x4 x5;
vartable = [x1; x2; x3; x4; x5];
prog = sosprogram(vartable);

gamma = 4;
f = 2.5 - .5*x1*x2 - .5*x2*x3 - .5*x3*x4 - .5*x4*x5 - .5*x5*x1;

bc1 = x1^2 - 1 ;
bc2 = x2^2 - 1 ;
bc3 = x3^2 - 1 ;
bc4 = x4^2 - 1 ;
bc5 = x5^2 - 1 ;

for i = 1:5
[prog, p{1+i}] = sospolyvar(prog,Z);
end;

expr = (gamma-f)+p{1}*bc1+p{2}*bc2+p{3}*bc3+p{4}*bc4+p{5}*bc5;

prog = sosineq(prog,expr);
prog = sossolve(prog);
```

The Structured Singular Value

For the case of structured parametric uncertainty, we define the structured singular value.

$$\mathbf{\Delta} = \{\Delta = \operatorname{diag}(\delta_1 I_{n1}, \cdots, \delta_s I_{ns} \,:\, \delta_i \in \mathbb{R}\}$$

- δ_i represent unknown parameters.

Definition 36.

Given system $M \in \mathcal{L}(L_2)$ and set $\mathbf{\Delta}$ as above, we define the **Structured Singular Value** of $(M, \mathbf{\Delta})$ as

$$\mu(M, \mathbf{\Delta}) = \frac{1}{\inf_{\substack{\Delta \in \mathbf{\Delta} \\ I - M\Delta \text{ is singular}}} \|\Delta\|}$$

The fundamental inequality we have is $\mathbf{\Delta}_\gamma = \{\operatorname{diag}(\delta_i), : \sum_i \delta_i^2 \leq \gamma\}$. We want to find the largest γ such that $I - M\Delta$ is stable for all $\Delta \in \mathbf{\Delta}_\gamma$

The Structured Singular Value, μ

The system

$$\dot{x}(t) = A_0 x(t) + M p(t), \qquad p(t) = \Delta(t) q(t),$$
$$q(t) = N x(t) + Q p(t), \qquad \Delta \in \mathbf{\Delta}$$

is stable if there exists a $P(\delta) \in \Sigma_s$ such that

$$\dot{V} = x^T P(\delta)(A_0 x + Mp) + (A_0 x + Mp)^T P(\delta) x < \epsilon x^T x$$

for all x, p, δ such that

$$(x, p, \delta) \in \left\{ x, p, \delta \, : \, p = \operatorname{diag}(\delta_i)(Nx + Qp), \, \sum_i \delta_i^2 \leq \gamma \right\}$$

Proposition 1 (Lower Bound for μ).

$\mu \geq \gamma$ if there exist polynomial $h \in \mathbb{R}[x, p, \delta]$ and $s_i \in \Sigma_s$ such that

$$x^T P(\delta)(A_0 x + Mp) + (A_0 x + Mp)^T P(\delta) x - \epsilon x^T x$$
$$= -s_0(x, p, \delta) - (\gamma - \sum_i \delta_i^2) s_1(x, p, \delta) - (p - \operatorname{diag}(\delta_i)(Nx + Qp)) h(x, p, \delta)$$

Variations - Polya's Formulation

Recall that Hilbert's 17th was resolved in the affirmative by E. Artin in 1927.

- Any PSD polynomial p is the sum, product and ratio of squared polynomials.
$$p(x) = \frac{g(x)}{h(x)}$$
  where $g, h \in \Sigma_s$.

It was later shown by Habricht that if p is strictly positive, then we may assume $h(x) = (\sum_i x_i^2)^d$ for some d. That is,

$$(x_1^2 + \cdots + x_n^2)^d p(x) \in \Sigma_s$$

Question: Given properties of p, may we assume a structure for h?

Yes: Polya was able to show that if $p(x)$ has the structure

$$p(x) = \tilde{p}(x_i^2, \cdots, x_n^2),$$

then we may assume that s is a sum of squared monomials (prima facie SOS).

$$s(x) = \sum_{\alpha \in \mathbb{N}^n} (c_\alpha x^\alpha)^2$$

where $x^\alpha = \prod_i x_i^{\alpha_i}$.

Variations - Polya's Formulation

Now suppose that we are interested in polynomials defined over the positive orthant:

$$X := \{x \,:\, x_i \geq 0,\, i = 1, \cdots\}$$

A polynomial $f(x_1, \cdots, x_n) > 0$ for all $x \in X$ if and only if $f(x_1^2, \cdots, x_n^2) \geq 0$ for all $x \in \mathbb{R}^n$.
Thus, by Polya's result if $f(x_1, \cdots, x_n) > 0$ for all $x \in X$, then

$$(\sum_i x_i^2)^{d_p} f(x_1^2, \cdots, x_n^2) = \sum_{\alpha \in \mathbb{N}^n} (c_\alpha x^\alpha)^2$$

for some $d_p > 0$.
Now making the substitution $x_i^2 \to y_i$, we have the condition

Theorem 37.

If $f(x_1, \cdots, x_n) > 0$ for all $x \in X$ then there exist $c_\alpha \geq 0$ and $d_p \geq 0$ such that

$$\left(\sum_i y_i\right)^{d_p} f(y_1, \cdots, y_n) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha|_1 \leq d + d_p}} c_\alpha y^\alpha$$

Variations - Polya's Formulation

Now suppose the polynomial, p, is homogeneous. Then $p(x) > 0$ for all $X/0$ if and only if $p(x) > 0$ for $x \in \Delta := \{x : \sum_i x_i = 1\}$ where Δ is the unit simplex.

Consider the uncertainty set $\Delta := \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 1,\ x_i \geq 0\}$.

Theorem 38 (Polya's Theorem).

Let $F(x)$ be a real homogeneous polynomial which is positive on Δ. Then for a sufficiently large $d \in \mathbb{N}$,

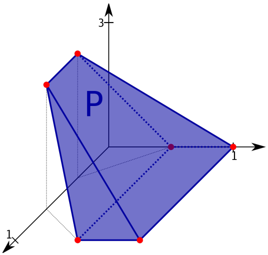
$$(x_1 + x_2 + \cdots + x_n)^d F(x)$$

has all its coefficients strictly positive.

The algorithmic nature was noted by Polya himself:

> *"The theorem gives a systematic process for deciding whether a given form F is strictly positive for positive x. We multiply repeatedly by $\sum x$, and, if the form is positive, we shall sooner or later obtain a form with positive coefficients."* -G. Pólya, 1934

Variations - Polya's Formulation



For example, if we have a finite number of operating points A_i, and want to ensure performance for all combinations of these points.

$$\dot{x}(t) = Ax(t) \quad \text{where} \quad A \in \left\{ \sum_i A_i \mu_i \; : \; \mu_i \geq 0, \; \sum_i \mu_i = 1 \right\}$$

This is equivalent to the existence of a polynomial P such that $P(\mu) > 0$ for all $\mu \in \Delta$ and such that

$$A(\mu)^T P(\mu) + P(\mu) A(\mu) < 0 \quad \text{for all} \quad \mu \in \Delta$$

$$\text{where} \quad A(\mu) = \sum_i A_i \mu_i$$

Variations - Polya's Formulation

A more challenging case is if $A(\alpha)$ is *nonlinear* in some parameters, α.
Simple Example: Angle of attack (α)

$$\dot{\alpha}(t) = -\frac{\rho v^2 c_\alpha(\alpha, M)}{2I}\alpha(t)$$

The time-varying parameters are:

- velocity, v and Mach number, M (M depends on Reynolds #);
- density of air, ρ;
- Also, we sometimes treat α itself as an uncertain parameter.

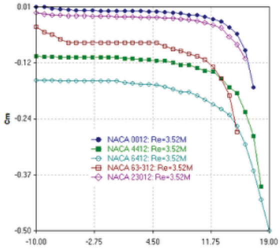


Figure : C_M vs. α and Re #

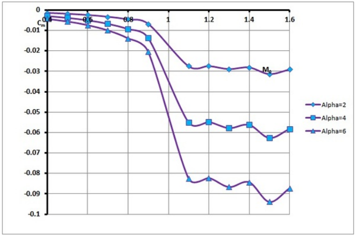


Figure : C_M vs. Mach # and α

Variations - Handelman's Formulation

Polya was not alone in looking for structure on s.

Recall Schmudgen's Positivstellensatz.

Theorem 39.

Suppose that $S = \{x : g_i(x) \geq 0\}$ is compact. If $f(x) > 0$ for all $x \in S$, then there exist $s_i, r_{ij}, \cdots \in \Sigma_s$ such that

$$f = s_0 + \sum_i s_i g_i + \sum_{i \neq j} r_{ij} g_i g_j + \sum_{i \neq j \neq k} r_{ijk} g_i g_j g_k + \cdots$$

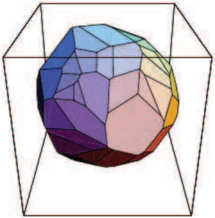
Suppose that S is a **CONVEX** polytope

$$S := \{x \in \mathbb{R}^n \, : \, a_i^T x \leq b_i, \; i = 1, \cdots \}$$

Then we may assume all the s_i are positive scalars.

Variations - Handelman's Formulation

Let $S := \{x \in \mathbb{R}^n \,:\, a_i^T \leq b_i\}$.



Theorem 40.

Suppose that $S := \{x \in \mathbb{R}^n \,:\, a_i^T x \leq b_i\}$ *is compact and convex with non-empty interior. If* $p(x) > 0$ *for all* $x \in S$*, then there exist constants* $s_i, r_{ij}, \dots > 0$ *such that*

$$p = s_0 + \sum_i s_i g_i + \sum_{i \neq j} r_{ij} g_i g_j + \sum_{i \neq j \neq k} r_{ijk} g_i g_j g_k + \cdots$$

Variations - Handelman's Formulation

Example: Consider the hypercube

$$S := \{(x,y) \,:\, -1 \le x \le 1,\ -1 \le y \le 1\}$$

Now the polytope is defined by 4 inequalities

$$g_1(x,y) = -x+1; \quad g_2(x,y) = x+1; \quad g_3(x,y) = -y+1; \quad g_4(x,y) = y+1$$

Which yields the following vector of bases

$$\begin{bmatrix} g_1 \\ \vdots \\ g_3 g_4 \end{bmatrix} = \begin{bmatrix} -x+1 \\ x+1 \\ -y+1 \\ y+1 \\ x^2-2x+1 \\ x^2+2x+1 \\ y^2-2y+1 \\ y^2+2y+1 \\ -x^2+1 \\ xy-x-y+1 \\ -xy-x+y+1 \\ -xy+x-y+1 \\ -y^2+1 \end{bmatrix}$$

Variations - Handelman's Formulation

First put the function in the linear basis

$$p(x) = -(y^2 + xy + y) + 3 = \begin{bmatrix} 3 & 0 & -1 & -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ xy \\ x^2 \\ y^2 \end{bmatrix}$$

Then convert the Handelman basis to the original basis

$$\begin{bmatrix} g_1 \\ \vdots \\ g_3 g_4 \end{bmatrix} = \begin{bmatrix} -x+1 \\ x+1 \\ -y+1 \\ y+1 \\ x^2-2x+1 \\ x^2+2x+1 \\ y^2-2y+1 \\ y^2+2y+1 \\ -x^2+1 \\ xy-x-y+1 \\ -xy-x+y+1 \\ -xy+x-y+1 \\ -y^2+1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & & & & \\ 1 & 1 & & & & \\ 1 & & -1 & & & \\ 1 & & 1 & & & \\ 1 & -2 & & & 1 & \\ 1 & 2 & & & 1 & \\ 1 & & -2 & & & 1 \\ 1 & & 2 & & & 1 \\ 1 & & & & -1 & \\ 1 & -1 & -1 & 1 & & \\ 1 & -1 & 1 & -1 & & \\ 1 & 1 & -1 & -1 & & \\ 1 & & & & & -1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ xy \\ x^2 \\ y^2 \end{bmatrix}$$

Variations - Handelman's Formulation

Now the positivity constraint becomes $c_i > 0$ and

$$p(x) = \begin{bmatrix} c_1 \\ \vdots \\ c_{13} \end{bmatrix}^T \begin{bmatrix} g_1(x) \\ \vdots \\ g_3(x)g_4(x) \end{bmatrix}.$$

Therefore, substituting the expressions of the previous slide

$$\begin{bmatrix} 3 \\ 0 \\ -1 \\ -1 \\ 0 \\ -1 \end{bmatrix}^T \begin{bmatrix} 1 \\ x \\ y \\ xy \\ x^2 \\ y^2 \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_{13} \end{bmatrix}^T \begin{bmatrix} 1 & -1 & & & & \\ 1 & 1 & & & & \\ 1 & & -1 & & & \\ 1 & & 1 & & & \\ 1 & -2 & & & 1 & \\ 1 & 2 & & & 1 & \\ 1 & & -2 & & & 1 \\ 1 & & 2 & & & 1 \\ 1 & & & & -1 & \\ 1 & -1 & -1 & 1 & & \\ 1 & -1 & 1 & -1 & & \\ 1 & 1 & -1 & -1 & & \\ 1 & & & & & -1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ xy \\ x^2 \\ y^2 \end{bmatrix}$$

Variations - Handelman's Formulation

Finally, we have that positivity of p can be expressed as the search for $c_i > 0$ such that

$$\begin{bmatrix} 1 & -1 & & & & \\ 1 & 1 & & & & \\ 1 & & -1 & & & \\ 1 & & 1 & & & \\ 1 & -2 & & & 1 & \\ 1 & 2 & & & 1 & \\ 1 & & -2 & & & 1 \\ 1 & & 2 & & & 1 \\ 1 & & & & -1 & \\ 1 & -1 & -1 & 1 & & \\ 1 & -1 & 1 & -1 & & \\ 1 & 1 & -1 & -1 & & \\ 1 & & & & & -1 \end{bmatrix}^T \begin{bmatrix} c_1 \\ \vdots \\ c_{13} \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ -1 \\ -1 \\ 0 \\ -1 \end{bmatrix}$$

Which is of the form $A^T x = b$ in variables $x > 0$.

Recall: Optimization over the positive orthant is called *Linear Programming*.

- b is determined by the coefficients of the polynomial, p
- b may itself be a variable if we are searching over positive polynomials.

Variations - Handelman's Formulation

For the polynomial

$$p(x) = -(y^2 + xy + y) + 3 = \begin{bmatrix} 3 & 0 & -1 & -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ xy \\ x^2 \\ y^2 \end{bmatrix}$$

The Linear Program is feasible with

$$x = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

This corresponds to the form

$$\begin{aligned} p(x) &= g_3(x)g_4(x) + g_2(x)g_4(x) + g_1(x) \\ &= (-y^2 + 1) + (-xy + x - y + 1) + (-x + 1) \\ &= -y^2 - xy - y + 3 \end{aligned}$$

Variations - Handelman's Formulation

Now consider the polynomial

$$p(x) = x^2 + y^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & x & y & xy & x^2 & y^2 \end{bmatrix}^T$$

Clearly, $p(x, y) \geq 0$ for all $(x, y) \in S$. However the LP is NOT feasible. Consider the point $(x, y) = (0, 0)$. Then $p(0, 0) = 0$ and

$$p(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ x \\ y \\ xy \\ x^2 \\ y^2 \end{bmatrix}_{(x,y)=0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_{13} \end{bmatrix}^T \begin{bmatrix} -x+1 \\ x+1 \\ -y+1 \\ y+1 \\ x^2-2x+1 \\ x^2+2x+1 \\ y^2-2y+1 \\ y^2+2y+1 \\ -x^2+1 \\ xy-x-y+1 \\ -xy-x+y+1 \\ -xy+x-y+1 \\ -y^2+1 \end{bmatrix}_{(x,y)=(0,0)} = \begin{bmatrix} c_1 \\ \vdots \\ c_{13} \end{bmatrix}^T \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Which implies $\sum_i c_i = 0$. Since the $c_i \geq 0$, this implies $c = 0$, which is not feasible.

Variations - Handelman's Formulation

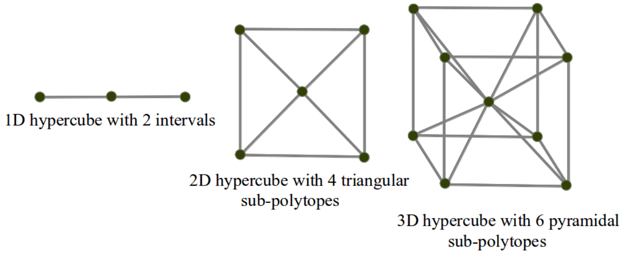
Conclusion: For many representations, the strict positivity is necessary.

- Polya's representation precludes interior-point zeros.
- Handelman's representation precludes interior-point zeros.
- Bernstein's representation precludes interior-point zeros.

In each of these cases, we may have zeros at vertices of the set.

- This makes searching for a Lyapunov function impossible.
    - Must be positive on a neighborhood of the $x = 0$ with $V(0) = 0$.

One Solution: Partition the space so that the zero point is a vertex of each set.



Conclusion

1. Optimization of Polynomials can be used to solve an broad class of problem.
    1.1 Stability Analysis of Vector Fields
    1.2 Bounds for Polynomial Programming
    1.3 Integer Programming
2. SDP, Sum-of-Squares and the Positivstellensatz allow us to optimize polynomials

Next Time:

- More applications of SOS