Polynomial optimization and sum-of-squares relaxations

- Non-convex problems
- Illustration of SDP relaxations: The maxcut problem
- Indefinite quadratic optimization
- Lagrange relaxation
- Improving relaxations with sum-of-squares
- Three typical examples from systems and control
Static Output Feedback Stabilization

Determine a feedback gain $K$ which renders

$$\dot{x} = (A + BKC)x$$

asymptotically stable.

Equivalently, determine $K$ and $X$ such that

$$X \succ 0 \quad \text{and} \quad (A + BKC)^T X + X(A + BKC) \prec 0.$$

- This is no LMI problem since the constraint is defined with a quadratic (bilinear) function of the involved variables.

- Is a polynomial semi-definite program with matrix variables and matrix-valued constraints. What to do?

Let’s start somewhat more modestly ...
Outline

- Non-convex problems
- Illustration of SDP relaxations: The maxcut problem
- Indefinite quadratic optimization
- Lagrange relaxation
- Improving relaxations with sum-of-squares
- Three typical examples from systems and control
Example: Maxcut Problem

Given an undirected graph with \( n \) nodes and nonnegative weights \( a_{jk} = a_{kj} \geq 0 \) on the arc joining node \( j \) with \( k \), find a partition \( S \cup T \) of the nodes such that the total weight of arcs linking \( S \) with \( T \) is maximal.

Identify a cut with a vector \( x \in \{-1, 1\}^n \) as follows:

\[
x_j = 1 \text{ if } j \in S \quad \text{and} \quad x_j = -1 \text{ if } j \in T.
\]

The total weight of arcs linking \( S \) with \( T \) is

\[
\frac{1}{2} \sum_{j,k=1}^{n} a_{jk}(1 - x_j x_k) = \frac{1}{2}(c - x^T A x) \quad \text{with} \quad c = \sum_{j,k=1}^{n} a_{jk}.
\]
Example: Maxcut Problem

The maxcut problem amounts to solving

\[
\text{minimize} \quad x^T Ax - c \\
\text{subject to} \quad x_j^2 = 1, \quad j = 1, \ldots, n
\]

This is an **indefinite** quadratic optimization problem (non-convex).

Devise fast algorithms to **approximate** optimal value \(v_{\text{opt}}\):

- Compute a **guaranteed lower bound** \(v_{\text{rel}}\)
- Try to **estimate accuracy** of approximation
Example: Maxcut Problem

Let us observe that

\[ x_j^2 = 1 \quad \text{iff} \quad \text{diag}(x x^T) = I \]

and recall the following simple facts:

\[ x^T A x = \text{trace}(x x^T A), \quad x x^T \succeq 0, \quad \text{rank}(x x^T) = 1. \]

The maxcut problem can be formulated as

\[
\begin{align*}
\text{minimize} & \quad \text{trace}(x x^T A) - c \\
\text{subject to} & \quad \text{diag}(x x^T) = I, \quad x x^T \succeq 0, \quad \text{rank}(x x^T) = 1
\end{align*}
\]

We can substitute \( x x^T \) by a full matrix \( M \), since the rank constraint forces \( M \) to be of the form \( x x^T \) for some vector \( x \).
Example: Maxcut Problem

The maxcut problem can be formulated as

\[
\begin{align*}
\text{minimize} & \quad \text{trace}(MA) - c \\
\text{subject to} & \quad \text{diag}(M) = I, \quad M \succeq 0, \quad \text{rank}(M) = 1
\end{align*}
\]

Let us now relax the constraints, by just dropping \(\text{rank}(M) = 1\).
Example: Maxcut Problem

The rank-relaxation of the maxcut problem is

\[
\begin{align*}
\text{minimize} & \quad \text{trace}(MA) - c \\
\text{subject to} & \quad \text{diag}(M) = I, \ M \succeq 0, \ \text{rank}(M) = 1
\end{align*}
\]

The following facts are obvious:

- The relaxation is a semi-definite program
- It defines a lower bound on the original value: \( v_{\text{rel}} \leq v_{\text{opt}} \)
- If it has a solution \( M \) of rank one, the relaxation gap vanishes

Starting from a solution \( M \), one can devise a randomization technique which leads to good upper bounds on \( v_{\text{opt}} \) for estimation of quality.
Example: Maxcut Problem

Determine a Cholesky-factorization

\[ M = (m_1 \cdots m_n)(m_1 \cdots m_n)^T \]

with vectors \( m_1, \ldots, m_n \).

For any \textbf{unit} vector \( u \) determine a feasibly point for maxcut as

\[ x_j = 1 \text{ if } u^T m_j \geq 0 \quad \text{and} \quad x_j = -1 \text{ if } u^T m_j < 0. \]

Choose \( u \) randomly on the unit sphere and evaluate the best cost.

On the basis of this randomization strategy, an ingenious argument by \textbf{Goemans & Williamson (95)} allows to show:

\[ 0.87856 \ldots \leq \frac{v_{\text{opt}}}{v_{\text{rel}}} \leq 1 \]

Such guarantees on the relaxation gap are usually out of reach!
Outline

- Non-convex problems
- Illustration of SDP relaxations: The maxcut problem
- Indefinite quadratic optimization
- Lagrange relaxation
- Improving relaxations with sum-of-squares
- Three typical examples from systems and control
Indefinite Quadratic Optimization

A general **quadratic function** \( f : \mathbb{R}^n \to \mathbb{R} \) can be represented as

\[
f(x) = q + 2s^T x + x^T R x = \begin{pmatrix} 1 \\ x \end{pmatrix}^T \begin{pmatrix} q & s^T \\ s & R \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix}^T P \begin{pmatrix} 1 \\ x \end{pmatrix}
\]

where \( P \) can be taken **symmetric**.

Surprisingly many problems can be transformed into the following **non-convex quadratic optimization** problem with value \( v_{\text{opt}} \):

\[
\inf \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} \quad \text{subject to} \quad \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix} \leq 0, \quad j = 1, \ldots, k
\]

Convex problem is easy: \( R_0 \succeq 0, R_1 \succeq 0, \ldots, R_k \succeq 0 \).
Sample Example: Robust Stability

If $G$ or $K$ strictly proper and systems are described in state-space:

**Affine dependence of $G$ on parameters** ...

**Affine dependence of controlled system on parameters**
Sample Example: Quadratic Stability

Closed-loop system with two parametric uncertainties:

\[
\dot{\xi} = (A + \delta_1 A_1 + \delta_2 A_2)\xi \quad \text{where} \quad |\delta_1| \leq 1, \quad |\delta_2| \leq 1.
\]

Robust stability guaranteed if there exists \( X > 0 \) such that

\[
(A + \delta_1 A_1 + \delta_2 A_2)^T X + X (A + \delta_1 A_1 + \delta_2 A_2) < 0 \quad \text{for all} \quad \delta_1, \delta_2.
\]

Polytopic Approach

Solve LMI at four generators.

Implies validity in whole uncertainty region.

Warning: \# generators grows exponentially with \# parameters

General problem is actually NP-hard (Nemirovski 93)
Suppose $A_1$, $A_2$ have rank one. Factorize $A_1 = B_1 C_1$, $A_2 = B_2 C_2$.

Uncertain system is $\dot{\xi} = (A + B_1 \delta_1 C_1 + B_2 \delta_2 C_2) \xi$.

Can be also described as

$$\dot{\xi} = A \xi + B_1 w_1 + B_2 w_2$$

$$z_1 = C_1 \xi, \quad z_2 = C_2 \xi$$

$$w_1 = \delta_1 z_1, \quad w_2 = \delta_2 z_2$$

Classical technique: Separate uncertainty from nominal system.

Is Linear Fractional Representation (LFR) of uncertain system.
Sample Example: Towards Quadratic Problem

Original inequality:

\[ x^T (A + B_1 \delta_1 C_1 + B_2 \delta_2 C_2)^T X x + x^T X (A + B_1 \delta_1 C_1 + B_2 \delta_2 C_2) x < 0 \]

With auxiliary variables:

\[(Ax + B_1 w_1 + B_2 w_2)^T X x + x^T X (Ax + B_1 w_1 + B_2 w_2) < 0\]

\[ z_1 = C_1 x, \quad z_2 = C_2 x, \quad w_1 = \delta_1 z_1, \quad w_2 = \delta_2 z_2 \]

Elementary observation:

\[ w_j = \delta_j z_j \quad \text{for some} \quad |\delta_j| \leq 1 \iff w_j^2 \leq z_j^2. \]

Allows to eliminate \( \delta_1 \) and \( \delta_2 \) ...
Sample Example: Robust Stability

We infer that

\[(A + \delta_1 A_1 + \delta_2 A_2)^T X + X (A + \delta_1 A_1 + \delta_2 A_2) \prec 0 \text{ for all } \delta_1, \delta_2\]

if and only if

\[(Ax + B_1 w_1 + B_2 w_2)^T X x + x^T X (Ax + B_1 w_1 + B_2 w_2) < 0\]

for all \(x, w_1, w_2\) and \(z_1, z_2\) satisfying

\[z_1 = C_1 x, \quad z_2 = C_2 x, \quad w_1^2 \leq z_1^2, \quad w_2^2 \leq z_2^2, \quad x^T x \geq 1.\]

Need to check whether maximum of quadratic function over quadratic inequality constraints is negative.

Many many other questions addressed from this point of view in SIAM book by **Boyd, El Ghaoui, Feron, Balakrishnan (94).**
Outline

- Non-convex problems
- Illustration of SDP relaxations: The maxcut problem
- Indefinite quadratic optimization
- Lagrange relaxation
- Improving relaxations with sum-of-squares
- Three typical examples from systems and control
QP: Problem

Quadratic optimization problem with optimal value $v_{\text{opt}}$:

$$\inf \left( \begin{array}{c} 1 \\ x \end{array} \right)^T P_0 \left( \begin{array}{c} 1 \\ x \end{array} \right) \text{ subject to } \left( \begin{array}{c} 1 \\ x \end{array} \right)^T P_j \left( \begin{array}{c} 1 \\ x \end{array} \right) \leq 0, \quad j = 1, \ldots, k$$

Just by definition of infimum: $v_{\text{opt}}$ is the largest $v$ for which

$$\left( \begin{array}{c} 1 \\ x \end{array} \right)^T P_0 \left( \begin{array}{c} 1 \\ x \end{array} \right) \geq v \quad \text{for all } x \text{ with } \left( \begin{array}{c} 1 \\ x \end{array} \right)^T P_j \left( \begin{array}{c} 1 \\ x \end{array} \right) \leq 0, \quad j = 1, \ldots, k$$
QP: Towards LMIs

\[
\begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} - v \geq 0 \text{ for all } x \text{ with } \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix} \leq 0, \ j = 1, \ldots, k
\]

\[
\uparrow
\]

There exist \( s_1 \geq 0, \ldots, s_k \geq 0 \) with

\[
\begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} + \sum_{j=1}^{k} s_j \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix} - v \geq 0 \text{ for all } x
\]

Sufficient condition involves \textbf{Lagrangian} of optimization problem.
QP: Towards LMIs

\[
\begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} - v \geq 0 \quad \text{for all } x \quad \text{with} \quad \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix} \leq 0, \ j = 1, \ldots, k
\]

\[
\uparrow
\]

There exist \( s_1 \geq 0, \ldots, s_k \geq 0 \) with

\[
\begin{pmatrix} 1 \\ x \end{pmatrix}^T \left( P_0 + \sum_{j=1}^k s_j P_j - \begin{pmatrix} v \\ 0 \\ 0 \end{pmatrix} \right) \begin{pmatrix} 1 \\ x \end{pmatrix} \geq 0 \quad \text{for all } x
\]

\[
\updownarrow
\]

There exist \( s_1 \geq 0, \ldots, s_k \geq 0 \) with

\[
P_0 + \sum_{j=1}^k s_j P_j - \begin{pmatrix} v \\ 0 \\ 0 \end{pmatrix} \succeq 0
\]
QP: Lagrange Relaxation

Quadratic optimization problem with optimal value $v_{\text{opt}}$:

$$\inf \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} \text{ subject to } \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix} \leq 0, \ j = 1, \ldots, k$$

The largest $v$ for which there exist $s_1 \geq 0, \ldots, s_k \geq 0$ with

$$P_0 + \sum_{j=1}^{k} s_j P_j - \begin{pmatrix} v & 0 \\ 0 & 0 \end{pmatrix} \succeq 0$$

is a lower bound on $v_{\text{opt}}$ and denoted as $v_{\text{rel}}$.

- Is called Lagrange relaxation of the original QP. (Shor 87)
- It’s the S-procedure in control. (Lur’e & Postnikov 44)
Sample Example: A Remarkable Result

Quadratic stability for polytopic uncertain systems: Exists $X \succ 0$ with

$$
\left( A + \sum_{k=1}^{p} \delta_k A_k \right)^T X + X \left( A + \sum_{k=1}^{p} \delta_k A_k \right) \prec 0 \text{ for all } |\delta_k| \leq 1.
$$

- Lagrange relaxation leads to sufficient LMI condition.
- If sufficient condition fails, quadratic stability might still be true.

Suppose $A_k$ have rank one. If sufficient condition from Lagrange relaxation fails, then **quadratic stability fails** for the larger set

$$
|\delta_1| \leq \frac{\pi}{2}, \ldots, |\delta_p| \leq \frac{\pi}{2}.
$$

Ben-Tal & Nemirovski (02)
QP: Lagrange Relaxation

• Suppose that there is only one constraint, which satisfies

\[
\begin{pmatrix}
1 \\
x_0
\end{pmatrix}^T P_1 \begin{pmatrix}
1 \\
x_0
\end{pmatrix} < 0 \text{ for some point } x_0.
\]

Then it is guaranteed that \( v_{\text{opt}} = v_{\text{rel}} \). (S-procedure is lossless)

• A similar result holds for two constraints if the problem is complex. This is the reason why the \( D \)-scalings \( \mu \)-upper bound relaxation is exact for three complex uncertainty blocks.

• There is no hope for generic estimations on the gap \( v_{\text{opt}} - v_{\text{rel}} \).

Can we check whether \( v_{\text{opt}} = v_{\text{rel}} \) for particular problem instance?

Can we systematically reduce \( v_{\text{opt}} - v_{\text{rel}} \) (improve relaxation)?
Reminder: Linear SDP Duality

If \( x \) satisfies \( F_0 + x_1 F_1 + \cdots + x_n F_n \preceq 0 \) and \( M \succeq 0 \) then

\[
\mathbf{c}^T \mathbf{x} \geq \sum_{j=1}^{n} c_j x_j + \text{trace}(M[F_0 + x_1 F_1 + \cdots + x_n F_n]) = \sum_{j=1}^{n} (c_j + \text{trace}(M F_j)) x_j + \text{trace}(M F_0).
\]

Infimization on both sides leads to

\[
\inf_{F_0 + x_1 F_1 + \cdots + x_n F_n \preceq 0} \mathbf{c}^T \mathbf{x} \geq \sup_{M \succeq 0, c_j + \text{trace}(M F_j) = 0, j=1,\ldots,n} \text{trace}(M F_0).
\]

- The original **primal** SDP in variables \( x_1, \ldots, x_n \) has been related to its **dual** in matrix variable \( M \).
- Inequality is called **weak duality**. Observe analogy to relaxation!
Reminder: Linear SDP Duality

Weak duality:
\[
\inf_{F_0 + x_1 F_1 + \cdots + x_n F_n \preceq 0} c^T x \quad \geq \quad \sup_{M \succeq 0, c_j + \text{trace}(M F_j) = 0, j = 1, \ldots, n} \text{trace}(M F_0).
\]

Lagrange Duality Theorem
If primal constraint strictly feasible then equality holds with \( \max \).

Strong duality due to convexity and Slater’s constraint qualification:
\[
\inf_{F_0 + x_1 F_1 + \cdots + x_n F_n \preceq 0} c^T x \quad = \quad \max_{M \succeq 0, c_j + \text{trace}(M F_j) = 0, j = 1, \ldots, n} \text{trace}(M F_0).
\]

Can routinely dualize SDPs ... and compute dual optimal solution.
QP: Interpretation of Dual of Lagrange Relaxation

Quadratic program:

\[
\begin{align*}
\inf & \quad \left( \begin{array}{c} 1 \\ x \end{array} \right)^T P_0 \left( \begin{array}{c} 1 \\ x \end{array} \right) \\
\text{subject to} & \quad \left( \begin{array}{c} 1 \\ x \end{array} \right)^T P_j \left( \begin{array}{c} 1 \\ x \end{array} \right) \leq 0, & j = 1, \ldots, k
\end{align*}
\]

Lagrange Relaxation: Maximize \( v \) over \( s_1 \geq 0, \ldots, s_k \geq 0 \) and

\[
P_0 + \sum_{j=1}^k s_j P_j - \begin{pmatrix} v & 0 \\ 0 & 0 \end{pmatrix} \succeq 0.
\]

Dual of Relaxation: Minimize \( \text{trace}(MP_0) \) over

\[
\text{trace}(MP_1) \leq 0, \ldots, \text{trace}(MP_n) \leq 0, \quad M_{11} = 1.
\]

This is rank relaxation! If optimal \( M \) has rank one then \( v_{\text{opt}} = v_{\text{rel}} \)!
QP: Summary

**Quadratic program** with optimal value $v_{\text{opt}}$:

$$
\inf \left( \begin{array}{c} 1 \\ x \end{array} \right)^T P_0 \left( \begin{array}{c} 1 \\ x \end{array} \right) \quad \text{subject to} \quad \left( \begin{array}{c} 1 \\ x \end{array} \right)^T P_j \left( \begin{array}{c} 1 \\ x \end{array} \right) \leq 0, \quad j = 1, \ldots, k
$$

- Manifold applications in control.
- Constructed Lagrange relaxation for computing lower bound $v_{\text{rel}}$.
  
  Was based on elementary weak duality argument.
- Can routinely compute dual optimal solution $M$ of relaxation.

  If $M$ has rank one then the relaxation gap vanishes: $v_{\text{opt}} = v_{\text{rel}}$.

We then say that the relaxation is **exact**.
Outline

• Non-convex problems
• Illustration of SDP relaxations: The maxcut problem
• Indefinite quadratic optimization
• Lagrange relaxation
• Improving relaxations with sum-of-squares
• Three typical examples from systems and control
QP: How to Improve Relaxation?

\[
\begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} - v \geq 0 \quad \text{for all } x \text{ with } \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix} \leq 0, \quad j = 1, \ldots, k
\]

\[\uparrow\]

Exist \( s_1 \geq 0, \ldots, s_k \geq 0 \) with

\[
\begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} - v + \sum_{j=1}^{k} s_j \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix} \geq 0 \quad \text{for all } x
\]

Conservative to use real numbers \( s_1 \geq 0, \ldots, s_k \geq 0 \).

Try instead nonnegative quadratic functions:

\[s_j(x) = \begin{pmatrix} 1 \\ x \end{pmatrix}^T S_j \begin{pmatrix} 1 \\ x \end{pmatrix} \quad \text{with } S_j \succeq 0.\]
Geometric Interpretation

**Convex**: Separation by *linear* functionals

**Non-convex**: Separation by *polynomial* functionals
QP: How to Improve Relaxation?

\[
\begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} - v \geq 0 \quad \text{for all } x \text{ with } \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix} \leq 0, \; j = 1, \ldots, k
\]

\[ \uparrow \]

Exist \( S_1 \succeq 0, \ldots, S_k \succeq 0 \) with

\[
\begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} - v + \sum_{j=1}^{k} \left( \begin{pmatrix} 1 \\ x \end{pmatrix}^T S_j \begin{pmatrix} 1 \\ x \end{pmatrix} \right) \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix} \geq 0 \; \forall \; x
\]

Crucial implication stays true! (Weak Duality)

However, how is it possible to check \textbf{global} positivity of a \textbf{fourth degree multivariable polynomial} with LMIIs?
Example: The Idea

Consider the polynomial

\[ p(x_1, x_2) = 3x_1^4 + 2x_1^3x_2 - x_1^2x_2^2 + 5x_2^4. \]

Mimic quadratic case and write the polynomial as

\[ p(x_1, x_2) = \begin{pmatrix} x_2^2 \\ x_1^2 \\ x_1x_2 \end{pmatrix}^T \begin{pmatrix} 3 & 0 & 1 \\ 0 & 5 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix}. \]

Note that the representing matrix is not unique. Indeed we have

\[ p(x_1, x_2) = \begin{pmatrix} x_2^2 \\ x_1^2 \\ x_1x_2 \end{pmatrix}^T \begin{pmatrix} 3 & -\lambda & 1 \\ -\lambda & 5 & 0 \\ 1 & 0 & -1 + 2\lambda \end{pmatrix} \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix} \text{ for all } \lambda \in \mathbb{R}. \]

Search \( \lambda \in \mathbb{R} \) for which representing matrix is positive semidefinite.
Example: The Idea

For $\lambda_0 \approx 0.67$ have

$$\begin{pmatrix} 3 & -\lambda_0 & 1 \\ -\lambda_0 & 5 & 0 \\ 1 & 0 & -1 + 2\lambda_0 \end{pmatrix} \approx \begin{pmatrix} 3 & -\lambda_0 & 1 \\ -\lambda_0 & 5 & 0 \\ 1 & 0 & -1 + 2\lambda_0 \end{pmatrix} \begin{pmatrix} 3 & -\lambda_0 & 1 \\ -\lambda_0 & 5 & 0 \\ 1 & 0 & -1 + 2\lambda_0 \end{pmatrix} = L^T L \quad \text{with} \quad L \approx \begin{pmatrix} 1.7 & -0.4 & 0.6 \\ 0 & 2.2 & 0.1 \end{pmatrix}.$$ 

Therefore

$$p(x_1, x_2) = \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix}^T L^T L \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix} = \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix}^T \begin{pmatrix} 1.7x_1^2 - 0.4x_2^2 + 0.6x_1x_2 \\ 2.2x_2^2 + 0.1x_1x_2 \end{pmatrix} =$$

$$= (1.7x_1^2 - 0.4x_2^2 + 0.6x_1x_2)^2 + (2.2x_2^2 + 0.1x_1x_2)^2.$$ 

We have written $p$ as $p_1^2 + p_2^2$. This is called a sum-of-squares (SOS) decomposition, and it proves that $p$ is globally non-negative.
Summary: Sum-of-Squares

Suppose we are given a multivariable polynomial

\[ p(x) = \sum_{j_1=1}^{k_1} \cdots \sum_{j_n=1}^{k_n} c_{j_1,\ldots,j_n} x_1^{j_1} \cdots x_n^{j_n} \quad (k_1,\ldots,k_n \text{ even}) \]

Define the vector of all monomials up to degrees \( k_1/2, \ldots, k_n/2 \):

\[ z(x) = \begin{pmatrix} 1 & x_1 & \cdots & x_n & x_1^2 & x_1x_2 & \cdots \end{pmatrix}^T. \]

Determine all symmetric matrices \( Q_0, Q_1, \ldots, Q_d \) such that

\[ p(x) = z(x)^T [Q_0 + \lambda_1 Q_1 + \cdots + \lambda_d Q_d] z(x) \quad \text{for all} \quad \lambda_1, \ldots, \lambda_d \in \mathbb{R}. \]

The polynomial \( p \) is SOS iff there exist \( \lambda_1, \ldots, \lambda_d \in \mathbb{R} \) such that

\[ Q_0 + \lambda_1 Q_1 + \cdots + \lambda_d Q_d \succeq 0. \]

So-called Gram-matrix method. (Choi, Lam, Reznick 95)
Summary: Sum-of-Squares

If \( p \) is SOS it is globally non-negative. The converse is not true.

- Due to Hilbert (1888). Lead to the formulation of his 17th problem in his 1900 Address to International Congress of Mathematicians:

  Are non-negative polynomials SOS of rational functions?

  Positive answer given by Artin (1927).

- A concrete example for the gap was only given by Motzkin (1965):

  \[
  x^4 y^2 + x^2 y^4 + z^6 - 3x^2 y^2 z^2 \quad \text{AGM inequality} \quad \geq \quad 3^{\frac{3}{2}} x^6 y^6 z^6 - 3x^2 y^2 z^2 \geq 0.
  \]

- The converse is true for degree-two polynomials in an arbitrary number of variables. Was exploited in constructing rank-relaxation.
QP: How to Improve Relaxation?

\[
\begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} - v \geq 0 \text{ for all } x \text{ with } \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix} \leq 0, \ j = 1, \ldots, k
\]

\[
\uparrow
\]

Exist \( S_1 \succeq 0, \ldots, S_k \succeq 0 \) with

\[
\begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} - v + \sum_{j=1}^{k} \left( \begin{pmatrix} 1 \\ x \end{pmatrix}^T S_j \begin{pmatrix} 1 \\ x \end{pmatrix} \right) \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix} \geq 0 \ \forall \ x
\]

\[
\uparrow
\]

Exist \( S_1 \succeq 0, \ldots, S_k \succeq 0 \) such that

\[
\begin{pmatrix} 1 \\ x \end{pmatrix}^T P_0 \begin{pmatrix} 1 \\ x \end{pmatrix} - v + \sum_{j=1}^{k} \left( \begin{pmatrix} 1 \\ x \end{pmatrix}^T S_j \begin{pmatrix} 1 \\ x \end{pmatrix} \right) \begin{pmatrix} 1 \\ x \end{pmatrix}^T P_j \begin{pmatrix} 1 \\ x \end{pmatrix}
\]

is SOS.

Due to affine dependence on \( v, S_1, \ldots, S_k \) this is still an LMI!
Let $f_0(x)$ and $f_1(x), \ldots, f_k(x)$ be arbitrary polynomials. Collect in column vector $z_d(x)$ all monomials up to degree $d$.

\[
f_0(x) - v \geq 0 \quad \text{for all} \quad x \quad \text{with} \quad f_j(x) \leq 0, \; j = 1, \ldots, k
\]

\[
\exists S_1 \succeq 0, \ldots, S_k \succeq 0 \quad \text{with}
\]

\[
f_0(x) - v + \sum_{j=1}^{k} [z_d(x)^T S_j z_d(x)] f_j(x) \geq 0 \quad \text{for all} \quad x
\]

\[
\exists S_1 \succeq 0, \ldots, S_k \succeq 0 \quad \text{with}
\]

\[
f_0(x) - v + \sum_{j=1}^{k} [z_d(x)^T S_j z_d(x)] f_j(x) \quad \text{is SOS.}
\]

Again: This is still an LMI!
Polynomial Optimization: SOS Relaxation

Polynomial program with optimal value $v_{\text{opt}}$:

Minimize $f_0(x)$ subject to $f_1(x) \leq 0, \ldots, f_k(x) \leq 0$.

Computing largest $v$ for which there exist $S_1 \succeq 0, \ldots, S_k \succeq 0$ with

$$f_0(x) - v + \sum_{j=1}^{k} [z_d(x)^T S_j z_d(x)] f_j(x)$$

is SOS

amounts to solving an SDP. Denote its optimal value by $v_d$.

- Improving family of lower bound relaxations: $v_0 \leq v_1 \leq \ldots \leq v_{\text{opt}}$.

- **Surprise:** Converges under weak hypotheses: $\lim_{d \to \infty} v_d = v_{\text{opt}}$.

- Often good approximation power for $d = 0, 1, 2$ in practice.

Parrilo & Sturmfels (01), Lasserre (01)
Remarks

- Convergence if exist $R > 0$ and SOS polynomials $s_1(x), \ldots, s_k(x)$:

$$R - \|x\|^2 + s_1(x)f_1(x) + \cdots + s_k(x)f_k(x) \text{ is SOS.}$$

Is trivially true if one of the constraint functions is $\|x\|^2 - R$.

Putinar (93), Schweighofer (05)

- The Lagrange dual of the SOS relaxation is the so-called Moment relaxation, the higher-degree variant of rank relaxation seen earlier.

The rank-one condition for exactness can be generalized to Moment relaxation. Can verify whether $v_{opt} = v_d$ for finite $d$, and can extract optimal solutions.

Lasserre (01), Henrion & Lasserre (05)
Software

Matlab-based tool for modeling optimization problems. Interface to various SDP solvers and easy access to SOS and moment relaxations:

- Yalmip (Löfberg)

Dedicated tools for SDP based relaxation of polynomial optimization:

- GloptiPoly (Herion, Lasserre)
  - SOSTools (Pranja, Papachristodoulou, Parrilo)

Freely available SDP code:

- SeDuMi (Sturm), SDPT3 (Toh, Todd, Tütüncü), CSDP (Borchers)
  - SDPA (Kojima), DSDP (Benson, Ye), PENSDDP (Kočvara, Stingl)
Outline

• Non-convex problems
• Illustration of SDP relaxations: The maxcut problem
• Indefinite quadratic optimization
• Lagrange relaxation
• Improving relaxations with sum-of-squares
• Three typical examples from systems and control
Example: Robust Stability of Polynomials

Test whether all following polynomials are Hurwitz:

\[ s^3 + (3 - \delta_1^2 + \delta_2) s^2 + (3 + \delta_1) s + (0.9 + \delta_1 \delta_2), \quad \delta_1 \in [-1, 1], \quad \delta_2 \in [-1, 1]. \]

Due to Routh-Hurwitz stability criterion, this amounts to testing

\[
\begin{align*}
3 - \delta_1^2 + \delta_2 &> 0 \\
(3 + \delta_1 + \delta_2)(3 + \delta_1) - (0.9 + \delta_1 \delta_2) &> 0
\end{align*}
\]

for all \( \delta_1, \delta_2 \) with \( \begin{cases} 
\delta_1^2 \leq 1 \\
\delta_2^2 \leq 1
\end{cases} \)

Yalmip-based relaxation for first positivity condition:

```matlab
sdpvar d1 d2 v; z=monolist([d1 d2],d); n=length(z);
S1=sdpvar(n); S2=sdpvar(n); L=set(S1>0)+set(S2>0);
L=L+set(sos(3+d1-d2^2-v+z^T*S1*z*(d1^2-1)+z^T*S2*z*(d2^2-1)));
solvesos(L,-v); double(v);
```
Example: Stabilization by low degree controllers

Transfer function system:

\[ G(s) = \frac{n(s)}{d(s)} \]

Transfer function controller:

\[ K(s) = \frac{x(s)}{y(s)} \]

Search \( x(s) \) and \( y(s) \) of fixed degree such that

characteristic polynomial \( n(s)x(s) + d(s)y(s) \) is Hurwitz.

The Routh-Hurwitz criterion leads to polynomial inequalities for the unknowns coefficients of \( x(s) \) and \( y(s) \).

Need to find feasible point of system of polynomial inequalities.
Example: Nonlinear Stability Analysis

Consider a differential equation defined with a polynomial function $f(x)$:

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

Suppose $f(x_0) = 0$. Then $x_0$ is an asymptotically stable equilibrium if there exists a polynomial Lyapunov function $V(x)$ such that

$$V(x) > V(x_0) \quad \text{and} \quad \partial V(x)f(x) < 0 \quad \text{for all} \quad x \neq x_0.$$ 

With monomials $b_1(x), \ldots, b_N(x)$ can parametrize $V(x)$ as

$$V(x) = c_1 b_1(x) + \cdots + c_N b_N(x).$$

Search of coefficients $c_1, \ldots, c_N$ in order to satisfy the above polynomial inequalities can be performed with SOS relaxation.
Polynomial Optimization in Control

SOS degree 0

SOS degree 1

SOS degree d

Exactness Verifiable