Overview

- Mathematical and computational theory, and applications to combinatorial, non-convex and nonlinear problems
  - Semidefinite programming
  - Real algebraic geometry
  - Duality and certificates

Topics

1. Convexity and duality
2. Quadratically constrained quadratic programming
3. From duality to algebra
4. Algebra and geometry
5. Sums of squares and semidefinite programming
6. Polynomials and duality; the Positivstellensatz
Discrete Problems: LQR with Binary Inputs

- linear discrete-time system \( x(t + 1) = Ax(t) + Bu(t) \) on interval \( t = 0, \ldots, N \)

- objective is to minimize the quadratic tracking error

\[
\sum_{t=0}^{N-1} (x(t) - r(t))^T Q (x(t) - r(t))
\]

- using binary inputs

\( u_i(t) \in \{-1, 1\} \) for all \( i = 1, \ldots, m \), and \( t = 0, \ldots, N - 1 \)
Nonlinear Problems: Lyapunov Stability
Graph problems

Graph problems appear in many areas: MAX-CUT, independent set, cliques, etc.

MAX CUT partitioning

- Partition the nodes of a graph in two disjoint sets, maximizing the number of edges between sets.
- Practical applications (circuit layout, etc.)
- NP-complete.

How to compute bounds, or exact solutions, for this kind of problems?
**polynomial programming**

A familiar problem

minimize \( f_0(x) \)

subject to \( f_i(x) \leq 0 \) for all \( i = 1, \ldots, m \)

\( h_i(x) = 0 \) for all \( i = 1, \ldots, p \)

objective, inequality and equality constraint functions are all *polynomials*
**polynomial nonnegativity**

first, consider the case of one inequality; given $f \in \mathbb{R}[x_1, \ldots, x_n]$

\[
\text{does there exist } x \in \mathbb{R}^n \text{ such that } f(x) < 0
\]

- if not, then $f$ is globally non-negative

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

and $f$ is called *positive semidefinite* or *PSD*

- the problem is *NP-hard*, but decidable

- many applications
certificates

the problem

\[ \text{does there exist } x \in \mathbb{R}^n \text{ such that } f(x) < 0 \]

- answer yes is easy to verify; exhibit \( x \) such that \( f(x) < 0 \)
- answer no is hard; we need a certificate or a witness
  i.e, a proof that there is no feasible point
**Sum of Squares Decomposition**

if there are polynomials $g_1, \ldots, g_s \in \mathbb{R}[x_1, \ldots, x_n]$ such that

$$f(x) = \sum_{i=1}^{s} g_i^2(x)$$

then $f$ is nonnegative

an easily checkable certificate, called a *sum-of-squares (SOS)* decomposition

- how do we find the $g_i$
- when does such a certificate exist?
example

we can write any polynomial as a \textit{quadratic function of monomials}

\[ f = 4x^4 + 4x^3y - 7x^2y^2 - 2xy^3 + 10y^4 \]

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

\[ = z^T Q(\lambda) z \]

which holds for all \( \lambda \in \mathbb{R} \)

if for some \( \lambda \) we have \( Q(\lambda) \succeq 0 \), then we can factorize \( Q(\lambda) \)
example, continued

e.g., with $\lambda = 6$, we have

$$Q(\lambda) = \begin{bmatrix} 4 & 2 & -6 \\ 2 & 5 & -1 \\ -6 & -1 & 10 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix}$$

so

$$f = \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}$$

$$= \left\| \begin{bmatrix} 2xy + y^2 \\ 2x^2 + xy - 3y^2 \end{bmatrix} \right\|^2$$

$$= (2xy + y^2)^2 + (2x^2 + xy - 3y^2)^2$$

which is an SOS decomposition
sum of squares and semidefinite programming

suppose \( f \in \mathbb{R}[x_1, \ldots, x_n] \), of degree \( 2d \)

let \( z \) be a vector of all monomials of degree less than or equal to \( d \)

\( f \) is SOS if and only if there exists \( Q \) such that

\[
Q \succeq 0 \\
f = z^T Q z
\]

- this is an SDP in standard primal form
- the number of components of \( z \) is \( \binom{n+d}{d} \)
- comparing terms gives affine constraints on the elements of \( Q \)
sum of squares and semidefinite programming

if \( Q \) is a feasible point of the SDP, then to construct the SOS representation factorize \( Q = V V^T \), and write \( V = [v_1 \ldots v_r] \), so that

\[
 f = z^T V V^T z \\
= \|V^T z\|^2 \\
= \sum_{i=1}^{r} (v_i^T z)^2
\]

- one can factorize using e.g., Cholesky or eigenvalue decomposition
- the number of squares \( r \) equals the rank of \( Q \)
example

\[ f = 2x^4 + 2x^3y - x^2y^2 + 5y^4 \]

\[
\begin{align*}
&= \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix} \\
&= q_{11}x^4 + 2q_{12}x^3y + (q_{22} + 2q_{13})x^2y^2 + 2q_{23}xy^3 + q_{33}y^4
\]

so \( f \) is SOS if and only if there exists \( Q \) satisfying the SDP

\[
Q \succeq 0 \quad q_{11} = 2 \quad 2q_{12} = 2 \\
2q_{12} + q_{22} = -1 \quad 2q_{23} = 0 \\
q_{33} = 5
\]
convexity

The sets of PSD and SOS polynomials are a convex cones; i.e.,

\[ f, g \text{ PSD} \implies \lambda f + \mu g \text{ is PSD for all } \lambda, \mu \geq 0 \]

Let \( P_{n,d} \) be the set of PSD polynomials of degree \( \leq d \)
Let \( \Sigma_{n,d} \) be the set of SOS polynomials of degree \( \leq d \)

- both \( P_{n,d} \) and \( \Sigma_{n,d} \) are convex cones in \( \mathbb{R}^N \) where \( N = \binom{n+d}{d} \)
- we know \( \Sigma_{n,d} \subset P_{n,d} \), and testing if \( f \in P_{n,d} \) is NP-hard
- but testing if \( f \in \Sigma_{n,d} \) is an SDP (but a large one)
polynomials in one variable

if \( f \in \mathbb{R}[x] \), then \( f \) is SOS if and only if \( f \) is PSD

general example

all real roots must have even multiplicity, and highest coeff. is positive

\[
f = x^6 - 10x^5 + 51x^4 - 166x^3 + 342x^2 - 400x + 200
\]

\[
= (x - 2)^2(x - (2 + i))(x - (2 - i))(x - (1 + 3i))(x - (1 - 3i))
\]

now reorder complex conjugate roots

\[
= (x - 2)^2(x - (2 + i))(x - (1 + 3i))(x - (2 - i))(x - (1 - 3i))
\]

\[
= (x - 2)^2((x^2 - 3x - 1) - i(4x - 7))((x^2 - 3x - 1) + i(4x - 7))
\]

\[
= (x - 2)^2((x^2 - 3x - 1)^2 + (4x - 7)^2)
\]

so every PSD scalar polynomial is the sum of one or two squares
quadratic polynomials

A quadratic polynomial in $n$ variables is PSD if and only if it is SOS because it is PSD if and only if

\[ f = x^T Q x \]

where $Q \geq 0$

and it is SOS if and only if

\[ f = \sum_{i} (v_i^T x)^2 \]
\[ = x^T \left( \sum_i v_i v_i^T \right) x \]
some background

In 1888, Hilbert showed that PSD=SOS if and only if

- $d = 2$, i.e., quadratic polynomials
- $n = 1$, i.e., univariate polynomials
- $d = 4, n = 2$, i.e., quartic polynomials in two variables

\[
\begin{array}{c|cccc}
\mathcal{d} \setminus \mathcal{n} & 2 & 4 & 6 & 8 \\
\hline
1 & \text{yes} & \text{yes} & \text{yes} & \text{yes} \\
2 & \text{yes} & \text{yes} & \text{no} & \text{no} \\
3 & \text{yes} & \text{no} & \text{no} & \text{no} \\
4 & \text{yes} & \text{no} & \text{no} & \text{no} \\
\end{array}
\]

- in general $f$ is PSD does not imply $f$ is SOS
some background

- 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.

\[ M(x, y) = x^2 y^4 + x^4 y^2 + 1 - 3x^2 y^2 \]

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

\[
(x^2 + y^2 + 1) M(x, y) = (x^2 y - y)^2 + (xy^2 - x)^2 + (x^2 y^2 - 1)^2 + \\
\quad + \frac{1}{4}(xy^3 - x^3 y)^2 + \frac{3}{4}(x^3 y + x^3 y - 2xy)^2
\]
The Univariate Case:

\[ f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_{2d} x^{2d} \]

\[
= \begin{bmatrix}
1 \\
x \\
\vdots \\
x^d
\end{bmatrix}^T \begin{bmatrix}
q_{00} & q_{01} & \cdots & q_{0d} \\
q_{01} & q_{11} & \cdots & q_{1d} \\
\vdots & \vdots & \ddots & \vdots \\
q_{0d} & q_{1d} & \cdots & q_{dd}
\end{bmatrix} \begin{bmatrix}
1 \\
x \\
\vdots \\
x^d
\end{bmatrix}
\]

\[
= \sum_{i=0}^{d} \left( \sum_{j+k=i} q_{jk} \right) x^i
\]

- 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.
About SOS/SDP

- The resulting SDP problem is polynomially sized (in $n$, for fixed $d$).
- 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 $\alpha, \beta$ 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 $\gamma$ s.t. $f(x, y) - \gamma$ is SOS
- A semidefinite program (convex!).
- If exact, can recover optimal solution.
- Surprisingly effective.

Solving, the maximum $\gamma$ is -1.0316. Exact value.
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:

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

The matrix \( Q \) should be PSD and satisfy the affine constraints.
The feasible set is given by:

\[
\left\{ (\alpha, \beta) \mid \exists \lambda \text{ s.t. } \begin{bmatrix} 1 & -\frac{1}{2} \alpha & \beta - \lambda \\ -\frac{1}{2} \alpha & 2\lambda & \frac{1}{2}(\alpha + 3\beta) \\ \beta - \lambda & \frac{1}{2}(\alpha + 3\beta) & 1 \end{bmatrix} \succeq 0 \right\}
\]
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 $\dot{x} = f(x)$,

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

- For linear systems $\dot{x} = Ax$, quadratic Lyapunov functions $V(x) = x^T Px$
  
  \[
  P > 0, \quad A^T P + PA < 0.
  \]

- 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:

\[
\begin{align*}
\dot{x} &= -y - \frac{3}{2}x^2 - \frac{1}{2}x^3 \\
y &= 3x - y
\end{align*}
\]

Try a generic 4th order polynomial Lyapunov function.

\[
V(x, y) = \sum_{0 \leq j+k \leq 4} c_{jk} x^j y^k
\]

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!
Lyapunov Example

After solving the SDPs, we obtain a Lyapunov function.

\[
V = 4.5819x^2 - 1.5786xy + 1.7834y^2 - 0.12739x^3 + 2.5189x^2y - 0.34069xy^2 \\
+ 0.61188y^3 + 0.47537x^4 - 0.052424x^3y + 0.44289x^2y^2 + 0.0000018868xy^3 + 0.090723y^4
\]
**Lyapunov Example**

Find a Lyapunov function for

\[
\begin{align*}
\dot{x} &= -x + (1 + x) y \\
\dot{y} &= -(1 + x) x.
\end{align*}
\]

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:

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

The matrices are positive definite, so 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.
Example

minimize \[ x_1 x_2 \]
subject to \[ x_1 \geq 0 \]
\[ x_2 \geq 0 \]
\[ x_1^2 + x_2^2 \leq 1 \]

- The objective is not convex.
- The Lagrange dual function is

\[ g(\lambda) = \inf_x \left( x_1 x_2 - \lambda_1 x_1 - \lambda_2 x_2 + \lambda_3 (x_1^2 + x_2^2 - 1) \right) \]

\[ = \begin{cases} -\lambda_3 - \frac{1}{2} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}^T \begin{bmatrix} 2\lambda_3 & 1 \\ 1 & 2\lambda_3 \end{bmatrix}^{-1} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} & \text{if } \lambda_3 > \frac{1}{2} \\ -\infty & \text{otherwise, except bdry} \end{cases} \]
Example, continued

The dual problem is

\[
\begin{align*}
\text{maximize} & \quad g(\lambda) \\
\text{subject to} & \quad \lambda_1 \geq 0 \\
& \quad \lambda_2 \geq 0 \\
& \quad \lambda_3 \geq \frac{1}{2}
\end{align*}
\]

- By symmetry, if the maximum \( g(\lambda) \) is attained, then \( \lambda_1 = \lambda_2 \) at optimality
- The optimal \( g(\lambda^*) = -\frac{1}{2} \) at \( \lambda^* = (0, 0, \frac{1}{2}) \)
- Here we see an example of a duality gap; the primal optimal is strictly greater than the dual optimal
Example, continued

It turns out that, using the Schur complement, the dual problem may be written as

\[
\begin{align*}
\text{maximize} & \quad \gamma \\
\text{subject to} & \quad \begin{bmatrix}
-2\gamma - 2\lambda_3 & \lambda_1 & \lambda_2 \\
\lambda_1 & 2\lambda_3 & 1 \\
\lambda_2 & 1 & 2\lambda_3
\end{bmatrix} > 0 \\
\lambda_1 & > 0 \\
\lambda_2 & > 0
\end{align*}
\]

We’ll see a systematic way to convert a dual problem to an SDP, whenever the objective and constraint functions are polynomials.
The Dual is Not Intrinsic

- The dual problem, and its corresponding optimal value, are *not properties of the primal feasible set and objective function alone.*

- Instead, they depend on the particular equations and inequalities used.

To construct equivalent primal optimization problems with different duals:

- replace the objective $f_0(x)$ by $h(f_0(x))$ where $h$ is increasing

- introduce new variables and associated constraints, e.g.

\[
\begin{align*}
\text{minimize} & \quad (x_1 - x_2)^2 + (x_2 - x_3)^2 \\
\text{is replaced by} & \\
\text{minimize} & \quad (x_1 - x_2)^2 + (x_4 - x_3)^2 \\
\text{subject to} & \quad x_2 = x_4
\end{align*}
\]

- add redundant constraints
**Example**

Adding the redundant constraint $x_1 x_2 \geq 0$ to the previous example gives

\[
\begin{align*}
\text{minimize} & \quad x_1 x_2 \\
\text{subject to} & \quad x_1 \geq 0 \\
& \quad x_2 \geq 0 \\
& \quad x_1^2 + x_2^2 \leq 1 \\
& \quad x_1 x_2 \geq 0
\end{align*}
\]

Clearly, this has the same primal feasible set and same optimal value as before.
Example Continued

The Lagrange dual function is

\[
g(\lambda) = \inf_x \left( x_1 x_2 - \lambda_1 x_1 - \lambda_2 x_2 + \lambda_3 (x_1^2 + x_2^2 - 1) - \lambda_4 x_1 x_2 \right)
\]

\[
= \begin{cases} 
-\lambda_3 - \frac{1}{2} \left[ \begin{array}{c} \lambda_1 \\ \lambda_2 \\
\end{array} \right]^T \left[ \begin{array}{cc} 2\lambda_3 & 1 - \lambda_4 \\ 1 - \lambda_4 & 2\lambda_3 \\
\end{array} \right]^{-1} \left[ \begin{array}{c} \lambda_1 \\ \lambda_2 \\
\end{array} \right] & \text{if } 2\lambda_3 > 1 - \lambda_4 \\
-\infty & \text{otherwise, except bdry}
\end{cases}
\]

- Again, this problem may also be written as an SDP. The optimal value is \( g(\lambda^*) = 0 \) at \( \lambda^* = (0, 0, 0, 1) \)
- Adding redundant constraints makes the dual bound **tighter**
- This always happens! Such redundant constraints are called **valid inequalities**.
Constructing Valid Inequalities

The function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is called a valid inequality if

\[
f(x) \geq 0 \quad \text{for all feasible } x
\]

Given a set of inequality constraints, we can generate others as follows.

(i) If \( f_1 \) and \( f_2 \) define valid inequalities, then so does \( h(x) = f_1(x) + f_2(x) \)

(ii) If \( f_1 \) and \( f_2 \) define valid inequalities, then so does \( h(x) = f_1(x) f_2(x) \)

(iii) For any \( f \), the function \( h(x) = f(x)^2 \) defines a valid inequality

Now we can use algebra to generate valid inequalities.
Valid Inequalities and Cones

- The set of polynomial functions on $\mathbb{R}^n$ with real coefficients is denoted $\mathbb{R}[x_1, \ldots, x_n]$

- Computationally, they are easy to parametrize. We will consider polynomial constraint functions.

A set of polynomials $P \subset \mathbb{R}[x_1, \ldots, x_n]$ is called a cone if

(i) $f_1 \in P$ and $f_2 \in P$ implies $f_1 f_2 \in P$

(ii) $f_1 \in P$ and $f_2 \in P$ implies $f_1 + f_2 \in P$

(iii) $f \in \mathbb{R}[x_1, \ldots, x_n]$ implies $f^2 \in P$

It is called a proper cone if $-1 \notin P$

By applying the above rules to the inequality constraint functions, we can generate a cone of valid inequalities
Algebraic Geometry

- There is a correspondence between the geometric object (the feasible subset of $\mathbb{R}^n$) and the algebraic object (the cone of valid inequalities).
- This is a dual relationship; we’ll see more of this later.
- The dual problem is constructed from the cone.
- For equality constraints, there is another algebraic object; the ideal generated by the equality constraints.
- For optimization, we need to look both at the geometric objects (for the primal) and the algebraic objects (for the dual problem).
Cones

- For $S \subset \mathbb{R}^n$, the cone defined by $S$ is

$$C(S) = \left\{ f \in \mathbb{R}[x_1, \ldots, x_n] \mid f(x) \geq 0 \text{ for all } x \in S \right\}$$

- If $P_1$ and $P_2$ are cones, then so is $P_1 \cap P_2$

- A polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ is called a sum-of-squares (SOS) if

$$f(x) = \sum_{i=1}^{r} s_i(x)^2$$

for some polynomials $s_1, \ldots, s_r$ and some $r \geq 0$. The set of SOS polynomials $\Sigma$ is a cone.

- Every cone contains $\Sigma$. 

---

Sum of Squares

S. Lall, Stanford 2003.11.12.04
Cones

The set $\text{monoid}\{f_1, \ldots, f_m\} \subset \mathbb{R}[x_1, \ldots, x_n]$ is the set of all finite products of polynomials $f_i$, together with 1.

The smallest cone containing the polynomials $f_1, \ldots, f_m$ is

$$\text{cone}\{f_1, \ldots, f_m\} = \left\{ \sum_{i=1}^{r} s_i g_i \mid s_0, \ldots, s_r \in \Sigma, \quad g_i \in \text{monoid}\{f_1, \ldots, f_m\} \right\}$$

$\text{cone}\{f_1, \ldots, f_m\}$ is called the \textit{cone generated by} $f_1, \ldots, f_m$
Explicit Parametrization of the Cone

- If $f_1, \ldots, f_m$ are valid inequalities, then so is every polynomial in $\text{cone}\{f_1, \ldots, f_m\}$

- The polynomial $h$ is an element of $\text{cone}\{f_1, \ldots, f_m\}$ if and only if

$$h(x) = s_0(x) + \sum_{i=1}^{m} s_i(x)f_i(x) + \sum_{i\neq j} r_{ij}(x)f_i(x)f_j(x) + \ldots$$

where the $s_i$ and $r_{ij}$ are sums-of-squares.
An Algebraic Approach to Duality

Suppose $f_1, \ldots, f_m$ are polynomials, and consider the feasibility problem:

$$\text{does there exist } x \in \mathbb{R}^n \text{ such that } f_i(x) \geq 0 \text{ for all } i = 1, \ldots, m$$

Every polynomial in $\text{cone}\{f_1, \ldots, f_m\}$ is non-negative on the feasible set.

So if there is a polynomial $q \in \text{cone}\{f_1, \ldots, f_m\}$ which satisfies

$$q(x) \leq -\varepsilon < 0 \quad \text{for all } x \in \mathbb{R}^n$$

then the primal problem is infeasible.
Example

Let’s look at the feasibility version of the previous problem. Given \( t \in \mathbb{R} \), does there exist \( x \in \mathbb{R}^2 \) such that

\[
\begin{align*}
x_1 x_2 & \leq t \\
x_1^2 + x_2^2 & \leq 1 \\
x_1 & \geq 0 \\
x_2 & \geq 0
\end{align*}
\]

Equivalently, is the set \( S \) nonempty, where

\[
S = \left\{ x \in \mathbb{R}^n \mid f_i(x) \geq 0 \text{ for all } i = 1, \ldots, m \right\}
\]

where

\[
\begin{align*}
f_1(x) &= t - x_1 x_2 \\
f_2(x) &= 1 - x_1^2 - x_2^2 \\
f_3(x) &= x_1 \\
f_4(x) &= x_2
\end{align*}
\]
Example Continued

Let \( q(x) = f_1(x) + \frac{1}{2}f_2(x) \). Then clearly \( q \in \text{cone}\{f_1, f_2, f_3, f_4\} \) and

\[
q(x) = t - x_1x_2 + \frac{1}{2}(1 - x_1^2 - x_2^2)
\]

\[
= t + \frac{1}{2} - \frac{1}{2}(x_1 + x_2)^2
\]

\[
\leq t + \frac{1}{2}
\]

So for any \( t < -\frac{1}{2} \), the primal problem is infeasible.

This corresponds to Lagrange multipliers \((1, \frac{1}{2})\) for the thm. of alternatives.

Alternatively, this is a proof by contradiction.

- If there exists \( x \) such that \( f_i(x) \geq 0 \) for \( i = 1, \ldots, 4 \) then we must also have \( q(x) \geq 0 \), since \( q \in \text{cone}\{f_1, \ldots, f_4\} \)
- But we proved that \( q \) is negative if \( t < -\frac{1}{2} \)
Example Continued

We can also do better by using other functions in the cone. Try

\[ q(x) = f_1(x) + f_3(x)f_4(x) = t \]

giving the stronger result that for any \( t < 0 \) the inequalities are infeasible. Again, this corresponds to Lagrange multipliers \((1, 1)\)

- In both of these examples, we found \( q \) in the cone which was globally negative. We can view \( q \) as the Lagrangian function evaluated at a particular value of \( \lambda \)

- The Lagrange multiplier procedure is searching over a particular subset of functions in the cone; those which are generated by linear combinations of the original constraints.

- By searching over more functions in the cone we can do better
Normalization

In the above example, we have

\[ q(x) = t + \frac{1}{2} - \frac{1}{2}(x_1 + x_2)^2 \]

We can also show that \(-1 \in \text{cone}\{f_1, \ldots, f_4\}\), which gives a very simple proof of primal infeasibility.

Because, for \( t < -\frac{1}{2} \), we have

\[ -1 = a_0 q(x) + a_1 (x_1 + x_2)^2 \]

and by construction \( q \) is in the cone, and \((x_1 + x_2)^2\) is a sum of squares.

Here \( a_0 \) and \( a_1 \) are positive constants

\[ a_0 = \frac{-2}{2t + 1} \quad a_1 = \frac{-1}{2t + 1} \]
An Algebraic Dual Problem

Suppose $f_1, \ldots, f_m$ are polynomials. The primal feasibility problem is

\[
\text{does there exist } x \in \mathbb{R}^n \text{ such that } f_i(x) \geq 0 \quad \text{for all } i = 1, \ldots, m
\]

The dual feasibility problem is

\[
\text{Is it true that } -1 \in \text{cone}\{f_1, \ldots, f_m\}
\]

If the dual problem is feasible, then the primal problem is infeasible.

In fact, a result called the Positivstellensatz implies that strong duality holds here.
Interpretation: Searching the Cone

- Lagrange duality is searching over *linear combinations* with nonnegative coefficients

\[ \lambda_1 f_1 + \cdots + \lambda_m f_m \]

to find a globally negative function as a certificate.

- The above algebraic procedure is searching over *conic combinations*

\[ s_0(x) + \sum_{i=1}^{m} s_i(x) f_i(x) + \sum_{i \neq j} r_{ij}(x) f_i(x) f_j(x) + \ldots \]

where the \( s_i \) and \( r_{ij} \) are *sums-of-squares*
Interpretation: Formal Proof

We can also view this as a type of *formal proof*:

- View $f_1, \ldots, f_m$ are *predicates*, with $f_i(x) \geq 0$ meaning that $x$ satisfies $f_i$.

- Then $\text{cone}\{f_1, \ldots, f_m\}$ consists of predicates which are *logical consequences* of $f_1, \ldots, f_m$.

- If we find $-1$ in the cone, then we have a proof by contradiction.

Our objective is to *automatically search* the cone for negative functions; i.e., proofs of infeasibility.
Example: Linear Inequalities

Does there exist \( x \in \mathbb{R}^n \) such that

\[
Ax \geq 0
\]
\[
c^T x \leq -1
\]

Write \( A \) in terms of its rows \( A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} \),

then we have inequality constraints defined by linear polynomials

\[
f_i(x) = a_i^T x \quad \text{for } i = 1, \ldots, m
\]
\[
f_{m+1}(x) = -1 - c^T x
\]
Example: Linear Inequalities

We’ll search over functions $q \in \text{cone}\{f_1, \ldots, f_{m+1}\}$ of the form

$$q(x) = \sum_{i=1}^{m} \lambda_{i} f_{i}(x) + \mu f_{m+1}(x)$$

Then the algebraic form of the dual is:

$$\text{does there exist } \lambda_{i} \geq 0, \mu \geq 0 \text{ such that } q(x) = -1 \text{ for all } x$$

if the dual is feasible, then the primal problem is infeasible
**Example: Linear Inequalities**

The above dual condition is

\[ \lambda^T A x + \mu(-1 - c^T x) = -1 \quad \text{for all } x \]

which holds if and only if \( A^T \lambda = \mu c \) and \( \mu = 1 \).

So we can state the duality result as follows.

**Farkas Lemma**

If there exists \( \lambda \in \mathbb{R}^m \) such that

\[ A^T \lambda = c \quad \text{and} \quad \lambda \geq 0 \]

then there does not exist \( x \in \mathbb{R}^n \) such that

\[ A x \geq 0 \quad \text{and} \quad c^T x \leq -1 \]
Farkas Lemma

Farkas Lemma states that the following are strong alternatives

(i) there exists $\lambda \in \mathbb{R}^m$ such that $A^T \lambda = c$ and $\lambda \geq 0$
(ii) there exists $x \in \mathbb{R}^n$ such that $Ax \geq 0$ and $c^T x < 0$

Since this is just Lagrangian duality, there is a geometric interpretation

(i) $c$ is in the convex cone

$$\{ A^T \lambda \mid \lambda \geq 0 \}$$

(ii) $x$ defines the hyperplane

$$\{ y \in \mathbb{R}^n \mid y^T x = 0 \}$$

which separates $c$ from the cone
Optimization Problems

Let’s return to optimization problems instead of feasibility problems.

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{subject to} & \quad f_i(x) \geq 0 \quad \text{for all } i = 1, \ldots, m
\end{align*}
\]

The corresponding feasibility problem is

\[
\begin{align*}
t - f_0(x) \geq 0 \\
f_i(x) \geq 0 \quad \text{for all } i = 1, \ldots, m
\end{align*}
\]

One simple dual is to find polynomials \( s_i \) and \( r_{ij} \) such that

\[
t - f_0(x) + \sum_{i=1}^{m} s_i(x) f_i(x) + \sum_{i \neq j} r_{ij}(x) f_i(x) f_j(x) + \ldots
\]

is globally negative, where the \( s_i \) and \( r_{ij} \) are \textit{sums-of-squares}
Optimization Problems

We can combine this with a maximization over $t$

\[
\begin{align*}
\text{maximize} & \quad t \\
\text{subject to} & \quad t - f_0(x) + \sum_{i=1}^{m} s_i(x) f_i(x) + \\
& \quad \sum_{i=1}^{m} \sum_{j=1}^{m} r_{ij}(x) f_i(x) f_j(x) \leq 0 \text{ for all } x
\end{align*}
\]

$s_i, r_{ij}$ are sums-of-squares

- The variables here are (coefficients of) the polynomials $s_i, r_i$
- We will see later how to approach this kind of problem using semidefinite programming
equality constraints

consider the feasibility problem

does there exist $x \in \mathbb{R}^n$ such that
\[
f_i(x) = 0 \quad \text{for all } i = 1, \ldots, m
\]

the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is called a **valid equality constraint** if

\[
f(x) = 0 \quad \text{for all feasible } x
\]

given a set of equality constraints, we can generate others as follows

(i) if $f_1$ and $f_2$ are valid equalities, then so is $f_1 + f_2$

(ii) for any $h \in \mathbb{R}[x_1, \ldots, x_n]$, if $f$ is a valid equality, then so is $hf$

using these will make the dual bound **tighter**
ideals and valid equality constraints

A set of polynomials \( I \subset \mathbb{R}[x_1, \ldots, x_n] \) is called an ideal if

(i) \( f_1 + f_2 \in I \) for all \( f_1, f_2 \in I \)

(ii) \( fh \in I \) for all \( f \in I \) and \( h \in \mathbb{R}[x_1, \ldots, x_n] \)

- Given \( f_1, \ldots, f_m \), we can generate an ideal of valid equalities by repeatedly applying these rules.

- This gives the ideal generated by \( f_1, \ldots, f_m \),

\[
\text{ideal}\{f_1, \ldots, f_m\} = \left\{ \sum_{i=1}^{m} h_i f_i \mid h_i \in \mathbb{R}[x_1, \ldots, x_n] \right\}
\]

Written \( \text{ideal}\{f_1, \ldots, f_m\} \), or sometimes \( \langle f_1, \ldots, f_m \rangle \).
**generators of an ideal**

- every polynomial in $\text{ideal}\{f_1, \ldots, f_m\}$ is a valid equality.

- $\text{ideal}\{f_1, \ldots, f_m\}$ is the smallest ideal containing $f_1, \ldots, f_m$.

- the polynomials $f_1, \ldots, f_m$ are called the *generators*, or a *basis*, of the ideal.

**properties of ideals**

- if $I_1$ and $I_2$ are ideals, then so is $I_1 \cap I_2$

- an ideal generated by one polynomial is called a *principal ideal*
Feasibility of Semialgebraic Sets

Suppose $S$ is a *semialgebraic set* represented by polynomial inequalities and equations

$$S = \left\{ x \in \mathbb{R}^n \mid f_i(x) \geq 0, \ h_j(x) = 0 \text{ for all } i = 1, \ldots, m, \ j = 1, \ldots, p \right\}$$

we’d like to solve the feasibility problem

Is $S$ non-empty?

- Important, non-trivial result: the feasibility problem is *decidable*.
- But NP-hard (even for a single polynomial, as we have seen)
- We would like to *certify* infeasibility
The Real Nullstellensatz

Recall $\Sigma$ is the cone of polynomials representable as sums of squares.

Suppose $h_1, \ldots, h_m \in \mathbb{R}[x_1, \ldots, x_n]$.

$$-1 \in \Sigma + \text{ideal}\{h_1, \ldots, h_m\} \iff \mathcal{V}_\mathbb{R}\{h_1, \ldots, h_m\} = \emptyset$$

Equivalently, there is no $x \in \mathbb{R}^n$ such that

$$h_i(x) = 0 \quad \text{for all } i = 1, \ldots, m$$

if and only if there exists $t_1, \ldots, t_m \in \mathbb{R}[x_1, \ldots, x_n]$ and $s \in \Sigma$ such that

$$-1 = s + t_1h_1 + \cdots + t_mh_m$$
Example

Suppose \( h(x) = x^2 + 1 \). Then clearly \( \mathcal{V}_\mathbb{R}\{h\} = \emptyset \)

We saw earlier that the complex Nullstellensatz cannot be used to prove emptyness of \( \mathcal{V}_\mathbb{R}\{h\} \)

But we have

\[
-1 = s + th
\]

with

\[
s(x) = x^2 \quad \text{and} \quad t(x) = -1
\]

and so the real Nullstellensatz implies \( \mathcal{V}_\mathbb{R}\{h\} = \emptyset \).

The polynomial equation \(-1 = s + th\) gives a certificate of infeasibility.
The Positivstellensatz

We now turn to feasibility for basic semialgebraic sets, with primal problem

\[
\text{Does there exist } x \in \mathbb{R}^n \text{ such that} \\
f_i(x) \geq 0 \quad \text{for all } i = 1, \ldots, m \\
h_j(x) = 0 \quad \text{for all } j = 1, \ldots, p
\]

Call the feasible set \( S \); recall

- every polynomial in \( \text{cone}\{f_1, \ldots, f_m\} \) is nonnegative on \( S \)
- every polynomial in \( \text{ideal}\{h_1, \ldots, h_p\} \) is zero on \( S \)

The Positivstellensatz (Stengle 1974)

\[
S = \emptyset \iff -1 \in \text{cone}\{f_1, \ldots, f_m\} + \text{ideal}\{h_1, \ldots, h_m\}
\]
Example

Consider the feasibility problem

\[ S = \left\{ (x, y) \in \mathbb{R}^2 \mid f(x, y) \geq 0, h(x, y) = 0 \right\} \]

where

\[ f(x, y) = x - y^2 + 3 \]
\[ h(x, y) = y + x^2 + 2 \]

By the P-satz, the primal is infeasible if and only if there exist polynomials \( s_1, s_2 \in \Sigma \) and \( t \in \mathbb{R}[x, y] \) such that

\[ -1 = s_1 + s_2 f + th \]

A certificate is given by

\[ s_1 = \frac{1}{3} + 2(y + \frac{3}{2})^2 + 6(x - \frac{1}{6})^2, \quad s_2 = 2, \quad t = -6. \]
Explicit Formulation of the Positivstellensatz

The primal problem is

\[
\begin{align*}
&\text{Does there exist } x \in \mathbb{R}^n \text{ such that } \\
&f_i(x) \geq 0 \quad \text{for all } i = 1, \ldots, m \\
&h_j(x) = 0 \quad \text{for all } j = 1, \ldots, p
\end{align*}
\]

The dual problem is

\[
\begin{align*}
&\text{Do there exist } t_i \in \mathbb{R}[x_1, \ldots, x_n] \text{ and } s_i, r_{ij}, \ldots \in \Sigma \text{ such that } \\
&-1 = \sum_i h_i t_i + s_0 + \sum_i s_i f_i + \sum_{i \neq j} r_{ij} f_i f_j + \cdots
\end{align*}
\]

These are \textit{strong alternatives}
Testing the Positivstellensatz

Do there exist \( t_i \in \mathbb{R}[x_1, \ldots, x_n] \) and \( s_i, r_{ij}, \ldots \in \Sigma \) such that

\[
-1 = \sum_i t_i h_i + s_0 + \sum_i s_i f_i + \sum_{i \neq j} r_{ij} f_i f_j + \cdots
\]

- This is a convex feasibility problem in \( t_i, s_i, r_{ij}, \ldots \)
- To solve it, we need to choose a subset of the cone to search; i.e., the maximum degree of the above polynomial; then the problem is a semidefinite program
- This gives a hierarchy of syntactically verifiable certificates
- The validity of a certificate may be easily checked; e.g., linear algebra, random sampling
- Unless \( \text{NP} = \text{co-NP} \), the certificates cannot always be polynomially sized.
**Example: Farkas Lemma**

The primal problem; does there exist \( x \in \mathbb{R}^n \) such that

\[
\begin{align*}
Ax + b & \geq 0 \\
Cx + d & = 0
\end{align*}
\]

Let \( f_i(x) = a_i^T x + b_i \), \( h_i(x) = c_i^T x + d_i \). Then this system is infeasible if and only if

\[
-1 \in \text{cone}\{f_1, \ldots, f_m\} + \text{ideal}\{h_1, \ldots, h_p\}
\]

Searching over *linear combinations*, the primal is infeasible if there exist \( \lambda \geq 0 \) and \( \mu \) such that

\[
\lambda^T (Ax + b) + \mu^T (Cx + d) = -1
\]

Equating coefficients, this is equivalent to

\[
\begin{align*}
\lambda^T A & + \mu^T C = 0 \\
\lambda^T b & + \mu^T d = -1 \\
\lambda & \geq 0
\end{align*}
\]
Hierarch of Certificates

- Interesting connections with logic, proof systems, etc.
- Failure to prove infeasibility (may) provide points in the set.
- Tons of applications: optimization, copositivity, dynamical systems, quantum mechanics...

General Scheme
Special Cases

Many known methods can be interpreted as fragments of P-satz refutations.

- LP duality: linear inequalities, constant multipliers.
- S-procedure: quadratic inequalities, constant multipliers
- Standard SDP relaxations for QP.
- The linear representations approach for functions $f$ strictly positive on the set defined by $f_i(x) \geq 0$.

$$f(x) = s_0 + s_1f_1 + \cdots + s_nf_n, \quad s_i \in \Sigma$$

Converse Results

- Losslessness: when can we restrict a priori the class of certificates?
- Some cases are known; e.g., additional conditions such as linearity, perfect graphs, compactness, finite dimensionality, etc, can ensure specific a priori properties.
Example: Boolean Minimization

\[ x^T Q x \leq \gamma \]
\[ x_i^2 - 1 = 0 \]

A P-satz refutation holds if there is \( S \succeq 0 \) and \( \lambda \in \mathbb{R}^n, \varepsilon > 0 \) such that

\[ -\varepsilon = x^T S x + \gamma - x^T Q x + \sum_{i=1}^{n} \lambda_i (x_i^2 - 1) \]

which holds if and only if there exists a diagonal \( \Lambda \) such that \( Q \succeq \Lambda \), \( \gamma = \trace \Lambda - \varepsilon \).

The corresponding optimization problem is

\[
\begin{align*}
\text{maximize} & \quad \trace \Lambda \\
\text{subject to} & \quad Q \succeq \Lambda \\
& \quad \Lambda \text{ is diagonal}
\end{align*}
\]
**Example: S-Procedure**

The primal problem; does there exist $x \in \mathbb{R}^n$ such that

$$x^T F_1 x \geq 0$$

$$x^T F_2 x \geq 0$$

$$x^T x = 1$$

We have a P-satz refutation if there exists $\lambda_1, \lambda_2 \geq 0$, $\mu \in \mathbb{R}$ and $S \succeq 0$ such that

$$-1 = x^T S x + \lambda_1 x^T F_1 x + \lambda_2 x^T F_2 x + \mu(1 - x^T x)$$

which holds if and only if there exist $\lambda_1, \lambda_2 \geq 0$ such that

$$\lambda_1 F_1 + \lambda_2 F_2 \leq -I$$

Subject to an additional mild constraint qualification, this condition is also *necessary* for infeasibility.
Exploiting Structure

What algebraic properties of the polynomial system yield efficient computation?

- **Sparseness**: few nonzero coefficients.
  - Newton polytopes techniques
  - Complexity does not depend on the degree

- **Symmetries**: invariance under a transformation group
  - Frequent in practice. Enabling factor in applications.
  - Can reflect underlying physical symmetries, or modelling choices.
  - SOS on *invariant rings*
  - Representation theory and invariant-theoretic techniques.

- **Ideal structure**: Equality constraints.
  - SOS on *quotient rings*
  - Compute in the coordinate ring. Quotient bases (Groebner)
Example: Structured Singular Value

- Structured singular value $\mu$ and related problems: provides better upper bounds.

- $\mu$ is a measure of robustness: how big can a structured perturbation be, without losing stability.

- A standard semidefinite relaxation: the $\mu$ upper bound.
  - Morton and Doyle’s counterexample with four scalar blocks.
  - Exact value: approx. 0.8723
  - Standard $\mu$ upper bound: 1
  - New bound: 0.895
Example: Matrix Copositivity

A matrix $M \in \mathbb{R}^{n \times n}$ is **copositive** if

$$x^T M x \geq 0 \quad \forall x \in \mathbb{R}^n, x_i \geq 0.$$ 

- The set of copositive matrices is a convex closed cone, but...
- Checking copositivity is coNP-complete
- Very important in QP. Characterization of local solutions.
- The P-satz gives a family of computable SDP conditions, via:

  $$(x^T x)^d(x^T M x) = s_0 + \sum_i s_i x_i + \sum_{j,k} s_{j,k} x_j x_k + \cdots$$
Example: Geometric Inequalities

Ono’s inequality: For an acute triangle,

\[(4K)^6 \geq 27 \cdot (a^2 + b^2 - c^2)^2 \cdot (b^2 + c^2 - a^2)^2 \cdot (c^2 + a^2 - b^2)^2\]

where \(K\) and \(a, b, c\) are the area and lengths of the edges.

The inequality is true if:

\[
\begin{align*}
    t_1 &:= a^2 + b^2 - c^2 \geq 0 \\
    t_2 &:= b^2 + c^2 - a^2 \geq 0 \\
    t_3 &:= c^2 + a^2 - b^2 \geq 0
\end{align*}
\]

\[
\Rightarrow (4K)^6 \geq 27 \cdot t_1^2 \cdot t_2^2 \cdot t_3^2
\]

A simple proof: define

\[s(x, y, z) = (x^4 + x^2y^2 - 2y^4 - 2x^2z^2 + y^2z^2 + z^4)^2 + 15 \cdot (x - z)^2(x + z)^2(z^2 + x^2 - y^2)^2.\]

We have then

\[(4K)^6 - 27 \cdot t_1^2 \cdot t_2^2 \cdot t_3^2 = s(a, b, c) \cdot t_1 \cdot t_2 + s(c, a, b) \cdot t_1 \cdot t_3 + s(b, c, a) \cdot t_2 \cdot t_3\]

therefore proving the inequality.