12. Positivstellensatz

- Basic semialgebraic sets
- Semialgebraic sets
- Tarski-Seidenberg and quantifier elimination
- Feasibility of semialgebraic sets
- Real fields and inequalities
- The real Nullstellensatz
- The Positivstellensatz
- Example: Farkas lemma
- Hierarchy of certificates
- Boolean minimization and the S-procedure
- Exploiting structure
Basic Semialgebraic Sets

The \textit{basic (closed) semialgebraic set} defined by polynomials $f_1, \ldots, f_m$ is

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

Examples

- The nonnegative orthant in $\mathbb{R}^n$
- The cone of positive semidefinite matrices
- Feasible set of an SDP; polyhedra and spectrahedra

Properties

- If $S_1, S_2$ are basic closed semialgebraic sets, then so is $S_1 \cap S_2$; i.e., the class is closed under intersection
- Not closed under union or projection
Semialgebraic Sets

Given the basic semialgebraic sets, we may generate other sets by set theoretic operations; unions, intersections and complements.

A set generated by a finite sequence of these operations on basic semialgebraic sets is called a \textit{semialgebraic set}.

Some examples:

- The set
  \[ S = \left\{ x \in \mathbb{R}^n \mid f(x) \neq 0 \right\} \]
  is semialgebraic, where \( \neq \) denotes \(<, \leq, =, \neq\).

- In particular every real variety is semialgebraic.

- We can also generate the semialgebraic sets via Boolean logical operations applied to polynomial equations and inequalities
Semialgebraic Sets

Every semialgebraic set may be represented as either

- an intersection of unions

\[ S = \bigcap_{i=1}^{m} \bigcup_{j=1}^{p_i} \left\{ x \in \mathbb{R}^n \mid \text{sign } f_{ij}(x) = a_{ij} \right\} \text{ where } a_{ij} \in \{-1, 0, 1\} \]

- a finite union of sets of the form

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

- in \( \mathbb{R} \), a finite union of points and open intervals

Every \textit{closed} semialgebraic set is a finite union of basic closed semialgebraic sets; i.e., sets of the form

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

- If $S_1, S_2$ are semialgebraic, so is $S_1 \cup S_2$ and $S_1 \cap S_2$
- The projection of a semialgebraic set is semialgebraic
- The closure and interior of a semialgebraic sets are both semialgebraic
- Some examples:

Sets that are not Semialgebraic

Some sets are not semialgebraic; for example

- the graph $\{ (x, y) \in \mathbb{R}^2 \mid y = e^x \}$
- the infinite staircase $\{ (x, y) \in \mathbb{R}^2 \mid y = \lfloor x \rfloor \}$
- the infinite grid $\mathbb{Z}^n$
Tarski-Seidenberg and Quantifier Elimination

Tarski-Seidenberg theorem: if $S \subseteq \mathbb{R}^{n+p}$ is semialgebraic, then so are

- $\{ x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^p \ (x, y) \in S \}$ (closure under projection)
- $\{ x \in \mathbb{R}^n \mid \forall y \in \mathbb{R}^p \ (x, y) \in S \}$ (complements and projections)

i.e., quantifiers do not add any expressive power

*Cylindrical algebraic decomposition* (CAD) may be used to compute the semialgebraic set resulting from quantifier elimination
Feasibility of Semialgebraic Sets

Suppose $S$ is a semialgebraic set; we’d like to solve the feasibility problem

Is $S$ non-empty?

More specifically, suppose we have a semialgebraic set represented by polynomial inequalities and equations

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

- 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
Certificates So Far

- **The Nullstellensatz:** a necessary and sufficient condition for feasibility of *complex* varieties

\[
\left\{ x \in \mathbb{C}^n \mid h_i(x) = 0 \ \forall i \right\} = \emptyset \iff -1 \in \text{ideal}\{h_1, \ldots, h_m\}
\]

- **Valid inequalities:** a sufficient condition for infeasibility of *real basic* semialgebraic sets

\[
\left\{ x \in \mathbb{R}^n \mid f_i(x) \geq 0 \ \forall i \right\} = \emptyset \iff -1 \in \text{cone}\{f_1, \ldots, f_m\}
\]

- **Linear Programming:** necessary and sufficient conditions via duality for *real linear* equations and inequalities
Certificates So Far

<table>
<thead>
<tr>
<th>Degree \ Field</th>
<th>Complex</th>
<th>Real</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>Range/Kernel</td>
<td>Farkas Lemma</td>
</tr>
<tr>
<td></td>
<td>Linear Algebra</td>
<td>Linear Programming</td>
</tr>
<tr>
<td>Polynomial</td>
<td>Nullstellensatz</td>
<td>???</td>
</tr>
<tr>
<td></td>
<td>Bounded degree: LP</td>
<td>???</td>
</tr>
<tr>
<td></td>
<td>Groebner bases</td>
<td>???</td>
</tr>
</tbody>
</table>

We’d like a method to construct certificates for

- *polynomial* equations
- over the *real* field
Real Fields and Inequalities

If we can test feasibility of real equations then we can also test feasibility of real inequalities and inequations, because

- **inequalities:** there exists \( x \in \mathbb{R} \) such that \( f(x) \geq 0 \) if and only if there exists \((x, y) \in \mathbb{R}^2\) such that \( f(x) = y^2 \)

- **strict inequalities:** there exists \( x \) such that \( f(x) > 0 \) if and only if there exists \((x, y) \in \mathbb{R}^2\) such that \( y^2 f(x) = 1 \)

- **inequations:** there exists \( x \) such that \( f(x) \neq 0 \) if and only if there exists \((x, y) \in \mathbb{R}^2\) such that \( yf(x) = 1 \)

The underlying theory for real polynomials called **real algebraic geometry**
Real Varieties

The real variety defined by polynomials $h_1, \ldots, h_m \in \mathbb{R}[x_1, \ldots, x_n]$ is

$$\mathcal{V}_\mathbb{R}\{h_1, \ldots, h_m\} = \{ x \in \mathbb{R}^n \mid h_i(x) = 0 \text{ for all } i = 1, \ldots, m \}$$

We’d like to solve the feasibility problem; is $\mathcal{V}_\mathbb{R}\{h_1, \ldots, h_m\} \neq \emptyset$?

We know

- Every polynomial in ideal$\{h_1, \ldots, h_m\}$ vanishes on the feasible set.
- The (complex) Nullstellensatz:
  $$-1 \in \text{ideal}\{h_1, \ldots, h_m\} \implies \mathcal{V}_\mathbb{R}\{h_1, \ldots, h_m\} = \emptyset$$
- But this condition is not necessary over the reals
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

$$-1 = s + t_1 h_1 + \cdots + t_m h_m$$
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

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

\[
\begin{align*}
    f_i(x) &\geq 0 \quad \text{for all } i = 1, \ldots, m \\
    h_j(x) &= 0 \quad \text{for all } j = 1, \ldots, p
\end{align*}
\]

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

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

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

\[ s_1, s_2 \in \Sigma \text{ and } t \in \mathbb{R}[x, y] \]

such that

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

A certificate is given by

\[ s_1 = \frac{1}{3} + 2\left(y + \frac{3}{2}\right)^2 + 6\left(x - \frac{1}{6}\right)^2, \quad s_2 = 2, \quad t = -6. \]
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 \( s_i, r_{ij}, \ldots \in \sum \) 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 \textit{semidefinite program}
- This gives a \textit{hierarchy} of syntactically verifiable certificates
- The validity of a certificate may be easily checked; e.g., linear algebra, random sampling
- Unless NP=co-NP, the certificates cannot \textit{always} be polynomially sized.
Example: Farkas Lemma

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

\[
Ax + b \geq 0 \quad Cx + d = 0
\]

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

\[
\lambda^T A + \mu^T C = 0 \quad \lambda^T b + \mu^T d = -1 \quad \lambda \geq 0
\]
Hierarchy 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 $f_i(x) \geq 0$.

$$f(x) = s_0 + s_1 f_1 + \cdots + s_n f_n, \quad s_i \in \Sigma$$

**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 \( Q \succeq \Lambda \), \( \gamma = \text{trace} \Lambda - \varepsilon \).

The corresponding optimization problem is

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

$$(x^T x)^d(x^T M x) = s_0 + \sum_i s_i x_i + \sum_{jk} s_{jk} x_j x_k + \cdots$$
**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.