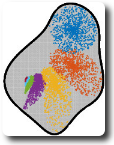
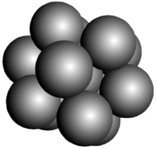
The Moment-Sums of Squares Hierarchy for Polynomial Optimization

Victor Magron, CNRS-LAAS

21 October 2019

TU Chemnitz





Outline

- Semidefinite Optimization

Outline

- Semidefinite Optimization
- Representation of positive polynomials

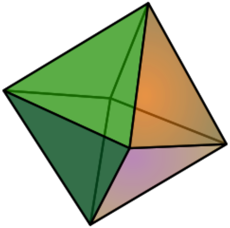
Outline

- Semidefinite Optimization
- Representation of positive polynomials
- The $\mathbf{K}$-moment problem

What is Semidefinite Optimization?

- Linear Programming (LP):

$$\min_{\mathbf{z}} \quad \mathbf{c}^\top \mathbf{z}$$
$$\text{s.t.} \quad \mathbf{A}\,\mathbf{z} \geqslant \mathbf{d} \; .$$



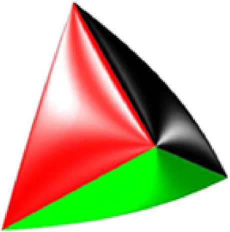
Polyhedron

  - Linear cost $\mathbf{c}$
  - Linear inequalities "$\sum_i A_{ij} z_j \geqslant d_i$"

What is Semidefinite Optimization?

- Semidefinite Programming (SDP):

$$\mathbf{P} : \min_{\mathbf{z}} \quad \mathbf{c}^\top \mathbf{z}$$
$$\text{s.t.} \quad \sum_i \mathbf{F}_i \, z_i \succcurlyeq \mathbf{F}_0 \ .$$



Spectrahedron

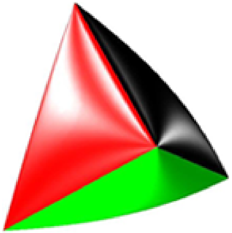
  - Linear cost $\mathbf{c}$
  - Symmetric matrices $\mathbf{F}_0, \mathbf{F}_i$
  - Linear matrix inequalities "$\mathbf{F} \succcurlyeq 0$" ($\mathbf{F}$ has nonnegative eigenvalues)

What is Semidefinite Optimization?

- Semidefinite Programming (SDP):

$$\mathbf{P} : \min_{\mathbf{z}} \quad \mathbf{c}^\top \mathbf{z}$$
$$\text{s.t.} \quad \sum_{i=1}^{n} \mathbf{F}_i \, z_i \succcurlyeq \mathbf{F}_0 \; , \quad \mathbf{A}\,\mathbf{z} = \mathbf{d} \; .$$

  - Linear cost $\mathbf{c}$
  - Symmetric matrices $\mathbf{F}_0, \mathbf{F}_i$
  - Linear matrix inequalities "$\mathbf{F} \succcurlyeq 0$" ($\mathbf{F}$ has nonnegative eigenvalues)



Spectrahedron

Example

$$\mathbf{P}: \quad \min_{\mathbf{z}} \quad \left\{ z_1 + z_2 \ : \right. \\ \text{s.t.} \quad \left. \begin{bmatrix} 3 + 2z_1 + z_2 & z_1 - 5 \\ z_1 - 5 & z_1 - 2z_2 \end{bmatrix} \succcurlyeq 0 \right\}$$

or, equivalently

$$\mathbf{P}: \min_{\mathbf{z}} \quad \left\{ z_1 + z_2 \ : \right. \\ \text{s.t.} \quad \left. \begin{bmatrix} 3 & -5 \\ -5 & 0 \end{bmatrix} + z_1 \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} + z_2 \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \succcurlyeq 0 \right\}$$

Its DUAL is the convex optimization problem:

$$\mathbf{P}^* : \max \{ \langle \mathbf{F}_0 , \mathbf{Y} \rangle \mid \quad \mathbf{Y} \succcurlyeq 0; \langle \mathbf{F}_i, \mathbf{Y} \rangle = c_i, \quad i = 1, \ldots, n \}$$

Example (continued)

The dual of

$$\mathbf{P}: \quad \min_{\mathbf{z}} \quad \Big\{ z_1 + z_2 \ : \quad \text{s.t.} \quad \begin{bmatrix} 3 & -5 \\ -5 & 0 \end{bmatrix} + z_1 \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} + z_2 \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \succcurlyeq 0 \Big\}$$

is the semidefinite program

$$\mathbf{P}^* : \quad \max_{\mathbf{Y} \succcurlyeq 0} \left\{ \langle \mathbf{Y}, - \begin{bmatrix} 3 & -5 \\ -5 & 0 \end{bmatrix} \rangle : \langle \mathbf{Y}, \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \rangle = 1; \langle \mathbf{Y}, \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix} \rangle = 1 \right\}$$

or, equivalently

$$\mathbf{P}^* : \quad \max \left\{ -3y_1 + 10y_2 \ : \ 2y_1 + 2y_2 + y_3 = 1; y_1 - 2y_3 = 1; \mathbf{Y} = \begin{bmatrix} y_1 & y_2 \\ y_2 & y_3 \end{bmatrix} \succcurlyeq 0 \right\}$$

$\mathbf{P}$ and its dual $\mathbf{P}^*$ are **convex** problems that are solvable in polynomial time to arbitrary precision $\epsilon > 0$.

$\mathbf{P}$ and its dual $\mathbf{P}^*$ are **convex** problems that are solvable in polynomial time to arbitrary precision $\epsilon > 0$.
= generalization to the convex cone $\mathcal{S}_m^+$ ($X \succcurlyeq 0$) of Linear Programming on the convex polyhedral cone $\mathbb{R}_+^m$ ($x \geqslant 0$).

Indeed, with DIAGONAL matrices

Semidefinite programming = Linear Programming!

$\mathbf{P}$ and its dual $\mathbf{P}^*$ are **convex** problems that are solvable in polynomial time to arbitrary precision $\epsilon > 0$.
= generalization to the convex cone $\mathcal{S}_m^+$ $(X \succcurlyeq 0)$ of Linear Programming on the convex polyhedral cone $\mathbb{R}_+^m$ $(x \geqslant 0)$.

Indeed, with DIAGONAL matrices

Semidefinite programming = Linear Programming!

Several academic SDP software packages exist, (e.g. MATLAB "LMI toolbox", SeduMi, SDPT3, ...). However, so far, size limitation is more severe than for LP software packages.
Pioneer contributions by A. Nemirovsky, Y. Nesterov, N.Z. Shor, B.D. Yudin,...

Consider the polynomial optimization problem:

$$\mathbf{P}: \quad f^* = \min\{ f(\mathbf{x}) : \quad g_j(\mathbf{x}) \geqslant 0, \, j = 1, \ldots, m \}$$

for some polynomials f, $g_j \in \mathbb{R}[\mathbf{x}]$.

Consider the polynomial optimization problem:

$$\mathbf{P}: \quad f^* = \min\{ f(\mathbf{x}) : \quad g_j(\mathbf{x}) \geqslant 0,\ j = 1, \ldots, m \}$$

for some polynomials f, $g_j \in \mathbb{R}[\mathbf{x}]$.

Why Polynomial Optimization?

After all ... $\mathbf{P}$ is just a particular case of Non Linear Programming (NLP)!

True!

... if one is interested with a LOCAL optimum only!!

True!

... if one is interested with a LOCAL optimum only!!

When searching for a local minimum ...

Optimality conditions and descent algorithms use basic tools from REAL and CONVEX analysis and linear algebra

True!

... if one is interested with a LOCAL optimum only!!

When searching for a local minimum ...

Optimality conditions and descent algorithms use basic tools from REAL and CONVEX analysis and linear algebra

The focus is on how to improve f by looking at a NEIGHBORHOOD of a nominal point $\mathbf{x} \in \mathbf{K}$, i.e., LOCALLY AROUND $\mathbf{x} \in \mathbf{K}$, and in general, no GLOBAL property of $\mathbf{x} \in \mathbf{K}$ can be inferred.

True!

... if one is interested with a LOCAL optimum only!!

When searching for a local minimum ...

Optimality conditions and descent algorithms use basic tools from REAL and CONVEX analysis and linear algebra

The focus is on how to improve f by looking at a NEIGHBORHOOD of a nominal point $\mathbf{x} \in \mathbf{K}$, i.e., LOCALLY AROUND $\mathbf{x} \in \mathbf{K}$, and in general, no GLOBAL property of $\mathbf{x} \in \mathbf{K}$ can be inferred.

The fact that f and g_j are POLYNOMIALS does not help much!

BUT for GLOBAL Optimization

... the picture is different!

BUT for GLOBAL Optimization

... the picture is different!

Remember that for the GLOBAL minimum f^*:

$$f^* = \sup \{ \lambda : f(\mathbf{x}) - \lambda \geqslant 0 \quad \forall \mathbf{x} \in \mathbf{K} \}$$

(Not true for a local minimum!))

BUT for GLOBAL Optimization

... the picture is different!

Remember that for the GLOBAL minimum f^*:

$$f^* = \sup \{ \lambda : f(\mathbf{x}) - \lambda \geqslant 0 \quad \forall \mathbf{x} \in \mathbf{K} \}$$

(Not true for a local minimum!))

and so to compute f^* ...
one needs to handle EFFICIENTLY the difficult constraint

$$f(\mathbf{x}) - \lambda \geqslant 0 \quad \forall \mathbf{x} \in \mathbf{K}$$

i.e. one needs
TRACTABLE CERTIFICATES of POSITIVITY on $\mathbf{K}$
for the polynomial $\mathbf{x} \mapsto f(\mathbf{x}) - \lambda$!

Global Optimization

Consider the GLOBAL optimization problem

$$(\mathbf{P}) \quad f^* := \min_x \{ f(\mathbf{x}) \mid \quad g_j(\mathbf{x}) \geqslant 0, j = 1, \ldots, m\}$$

where f, $g_j(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ are all **real-valued functions.** Let

$$\mathbf{K} := \{ \mathbf{x} \in \mathbb{R}^n \mid \quad g_j(\mathbf{x}) \geqslant 0, j = 1, \ldots, m\}$$

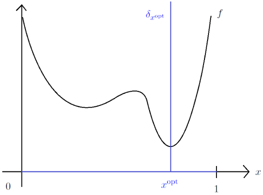
be the feasible set.

Two dual points of view

I: THE PRIMAL SIDE

f^* is a **global minimum if and only if**

$$(*) \quad f^* = \min_{\mu \in M(\mathbf{K})} \left\{ \int_{\mathbf{K}} f \, d\mu \mid \quad \mu(\mathbf{K}) = 1 \right\}$$



i.e. one optimizes over the set $\mathcal{P}(\mathbf{K}) \subset M(\mathbf{K})$ of probability measures with support contained in $\mathbf{K}$.

$$f \geqslant f^* \Rightarrow \int_{\mathbf{K}} f \, d\mu \geqslant \int_{\mathbf{K}} f^* \, d\mu = f^*, \forall \mu \in \mathcal{P}(\mathbf{K})$$

On the other hand, with $\mathbf{x} \in \mathbf{K}$ and $\mu := \delta_{\mathbf{x}}$, $\int f \, d\mu = f(\mathbf{x})$

Two dual points of view (continued)

... But also: II. THE DUAL SIDE

f^* is a **global minimum** **if and only if:**

$$f^* = \sup_{\lambda} \{\lambda : f(\mathbf{x}) - \lambda \geqslant 0 \quad \forall \mathbf{x} \in \mathbf{K}\}$$

and in particular, $\mathbf{x} \mapsto f(\mathbf{x}) - f^*$ is nonnegative on $\mathbf{K}$.

Two dual points of view (continued)

... But also: II. THE DUAL SIDE

f^* is a **global minimum if and only if:**

$$f^* = \sup_{\lambda} \{\lambda : f(\mathbf{x}) - \lambda \geqslant 0 \quad \forall \mathbf{x} \in \mathbf{K}\}$$

and in particular, $\mathbf{x} \mapsto f(\mathbf{x}) - f^*$ is nonnegative on $\mathbf{K}$.

Observe that these two characterizations are proper to the global optimum f^*, and are NOT valid for a local minimum!

BUT ... this is just LP-Duality

The optimization problem

$$f^* = \min_{\mu \in M(\mathbf{K})} \left\{ \int_{\mathbf{K}} f \, d\mu \quad | \quad \mu(\mathbf{K}) = 1 \right\}$$

is the **infinite-dimensional** LP

$$f^* = \min_{\mu \in \mathcal{M}(\mathbf{K})} \{ \langle f, \mu \rangle \mid \quad \langle 1, \mu \rangle = 1; \; \mu \geqslant 0 \}$$

where :

- $\mathcal{M}(\mathbf{K})$ is the **space** of **finite** *signed* **Borel measures** on $\mathbf{K}$, and
- $\langle \cdot, \cdot \rangle$ is the duality bracket between $\mathcal{C}(\mathbf{K})$ and $\mathcal{M}(\mathbf{K})$:

$$\langle f, \mu \rangle = \int_{\mathbf{K}} f \, d\mu, \qquad \forall \, f \in \mathcal{C}(\mathbf{K}), \; \mu \in \mathcal{M}(\mathbf{K})$$

As in the finite dimensional case ...., the **dual** LP reads:

$$f^* = \max_{\gamma \in \mathbb{R}} \{ \quad \gamma \quad | \quad f - \gamma\, 1 \in \mathcal{C}(\mathbf{K})_+ \}$$

As in the finite dimensional case ...., the **dual** LP reads:

$$f^* = \max_{\gamma \in \mathbb{R}} \{ \quad \gamma \quad | \quad f - \gamma \mathbf{1} \in \mathcal{C}(\mathbf{K})_+ \}$$

or, equivalently:

$$f^* = \max_{\gamma \in \mathbb{R}} \{ \quad \gamma \quad | \quad f(\mathbf{x}) - \gamma \geqslant 0 \quad \forall \mathbf{x} \in \mathbf{K} \}$$

☞

Hence, to solve (or at least approximate) either LP, one needs:

relatively simple and tractable characterizations of:

☞

Hence, to solve (or at least approximate) either LP, one needs:

relatively simple and tractable characterizations of:

- **measures** μ with support contained in $\mathbf{K}$, ... or

☞

Hence, to solve (or at least approximate) either LP, one needs:

relatively simple and tractable characterizations of:

- **measures** μ with support contained in $\mathbf{K}$, ... or
- **functions** (e.g. $f - \gamma$) nonnegative on $\mathbf{K}$.

☞

Hence, to solve (or at least approximate) either LP, one needs:

relatively simple and tractable characterizations of:

- **measures** μ with support contained in $\mathbf{K}$, ... or
- **functions** (e.g. $f - \gamma$) nonnegative on $\mathbf{K}$.

☞

Hence, to solve (or at least approximate) either LP, one needs:

relatively simple and tractable characterizations of:

- **measures** μ with support contained in $\mathbf{K}$, ... or
- **functions** (e.g. $f - \gamma$) nonnegative on $\mathbf{K}$.

☞ Not possible in general .... BUT ...

Concerning the primal LP, NOTICE that if f is a polynomial, i.e.,

$$f(\mathbf{x}) = \sum_{\alpha} f_{\alpha}\, \mathbf{x}^{\alpha} \quad \left(= \sum_{\alpha} f_{\alpha}\, x_1^{\alpha_1} \cdots x_1^{\alpha_n} \right)$$

then $$\int_{\mathbf{K}} f\, d\mu = \sum_{\alpha} f_{\alpha} \underbrace{\int_{\mathbf{K}} \mathbf{x}^{\alpha}\, d\mu}_{= y_{\alpha}} = \sum_{\alpha} f_{\alpha}\, y_{\alpha},$$

and so the primal LP reads:

$$f^* = \min \left\{ \sum_{\alpha} f_{\alpha}\, y_{\alpha} \, : \, y_0 = 1; \quad y \in \Delta \right\},$$

where $\Delta = \{ y = (y_{\alpha}) \, : \, \exists\, \mu \in \mathcal{M}(\mathbf{K})$ s.t. $y_{\alpha} = \int_{\mathbf{K}} \mathbf{x}^{\alpha}\, d\mu, \, \forall \alpha \}$.

Concerning the primal LP, NOTICE that if f is a polynomial, i.e.,

$$f(\mathbf{x}) = \sum_\alpha f_\alpha \, \mathbf{x}^\alpha \quad \left(= \sum_\alpha f_\alpha \, x_1^{\alpha_1} \cdots x_1^{\alpha_n} \right)$$

then $$\int_{\mathbf{K}} f \, d\mu = \sum_\alpha f_\alpha \underbrace{\int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu}_{= y_\alpha} = \sum_\alpha f_\alpha \, y_\alpha,$$

and so the primal LP reads:

$$f^* = \min \left\{ \sum_\alpha f_\alpha \, y_\alpha \ : \ y_0 = 1; \quad y \in \Delta \right\},$$

where $\Delta = \{ y = (y_\alpha) \ : \ \exists \, \mu \in \mathcal{M}(\mathbf{K})$ s.t. $y_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu, \ \forall \alpha \}$.

$\rightarrow$ a FINITE-DIMENSIONAL CONVEX optimization problem!!

SIMILARLY for the dual LP, NOTICE that if f is a polynomial of degree d, then the constraint of the dual LP

$$f(\mathbf{x}) - \gamma \geqslant 0 \quad \forall \mathbf{x} \in \mathbf{K}$$

is the same as stating that the polynomial $f - \gamma 1$ belongs to the FINITE-DIMENSIONAL convex cone

$$\Theta_d = \{g \in \mathbb{R}[\mathbf{x}]_d : g \geqslant 0 \text{ on } \mathbf{K}\}.$$

and so the dual LP reads:

$$f^* = \max \{\gamma : f - \gamma 1 \in \Theta_d\},$$

$\to$ a FINITE-DIMENSIONAL CONVEX optimization problem!!

First good news ...

When:

- $\mathbf{K} \subset \mathbb{R}^n$ is the compact semi-algebraic set

$$\mathbf{K} := \{ \mathbf{x} \in \mathbb{R}^n \mid \quad g_j(\mathbf{x}) \geqslant 0, \quad j = 1, \ldots, m \}$$

with $\{g_j\} \subset \mathbb{R}[x_1, \ldots, x_n]$ ...

First good news ...

When:

- $\mathbf{K} \subset \mathbb{R}^n$ is the compact semi-algebraic set

$$\mathbf{K} := \{ \mathbf{x} \in \mathbb{R}^n \mid \quad g_j(\mathbf{x}) \geqslant 0, \quad j = 1, \dots, m \}$$

  with $\{g_j\} \subset \mathbb{R}[x_1, \dots, x_n]$ ...

- and f is a POLYNOMIAL

POWERFUL results of real algebraic geometry provide

- a CHARACTERIZATION of **polynomials** POSITIVE on $\mathbf{K}$.

POWERFUL results of real algebraic geometry provide

- a CHARACTERIZATION of **polynomials** POSITIVE on $\mathbf{K}$.

  $\rightarrow$ which is what we need to solve the dual LP!

POWERFUL results of real algebraic geometry provide

- a CHARACTERIZATION of **polynomials** POSITIVE on $\mathbf{K}$.

  $\rightarrow$ which is what we need to solve the dual LP!

- a CHARACTERIZATION of real sequences $y = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, such that

$$y_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu, \qquad \forall \alpha \in \mathbb{N}^n,$$

  for some Borel Measure μ on $\mathbf{K}$.

POWERFUL results of real algebraic geometry provide

- a CHARACTERIZATION of **polynomials** POSITIVE on $\mathbf{K}$.

  $\to$ which is what we need to solve the dual LP!

- a CHARACTERIZATION of real sequences $y = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, such that

$$y_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu, \qquad \forall \alpha \in \mathbb{N}^n,$$

  for some Borel Measure μ on $\mathbf{K}$.

  $\to$ which is what we need to solve the primal LP!

Second good news ...

In both cases ... these conditions can translate into **Linear Matrix Inequalities** (**LMI**) on :

Second good news ...

In both cases ... these conditions can translate into **Linear Matrix Inequalities** (**LMI**) on :

- The **moments** $y_\alpha := \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu$ of μ (**Primal side**)

Second good news ...

In both cases ... these conditions can translate into **Linear Matrix Inequalities** (**LMI**) on :

- The **moments** $y_\alpha := \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu$ of μ (**Primal side**)
- The **coefficients** of some sum of squares (s.o.s.) **polynomials** $\{q_j\}_{j=0}^m \subset \mathbb{R}[\mathbf{x}]$, for the representation of a polynomial positive on $\mathbf{K}$.

Second good news ...

In both cases ... these conditions can translate into **Linear Matrix Inequalities** (**LMI**) on :

- The **moments** $y_\alpha := \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu$ of μ (**Primal side**)
- The **coefficients** of some sum of squares (s.o.s.) **polynomials** $\{q_j\}_{j=0}^m \subset \mathbb{R}[\mathbf{x}]$, for the representation of a polynomial positive on $\mathbf{K}$.

$\rightarrow$ PRACTICAL COMPUTATION is possible!

Other and (not s.o.s. based) representations of positive polynomials are available (Krivine, Handelman, Vasilescu).

† They lead to Linear Inequalities instead of LMIs and so

... to LP-relaxations instead of SDP-relaxations

.. but less efficient and ill-behaved ... despite so far, LP software packages are more powerful than SDP packages!!

Putinar's Positivstellensatze

Let $\mathbf{K} \subset \mathbb{R}^n$ be the basic semi-algebraic set:

$$\mathbf{K} := \{ x \in \mathbb{R}^n \, : \, g_j(\mathbf{x}) \geqslant 0, \quad j = 1, \ldots, m \}$$

for some polynomials $\{g_j\} \subset \mathbb{R}[\mathbf{x}]$.

and with g_0 being the constant polynomial 1,

define the quadratic module

$$\mathcal{Q}(g_1, \ldots, g_m) = \{g \in \mathbb{R}[X] : g = \sum_{j=0}^{m} \sigma_j \, g_j\},$$

where the $(\sigma_j)_{j=0}^{m}$ are s.o.s. polynomials.

• $f \in \mathcal{Q}(g_1, \ldots, g_m)$ is also an obvious **certificate of nonnegativity** on $\mathbf{K}$.

- $f \in \mathcal{Q}(g_1, \ldots, g_m)$ is also an obvious **certificate of nonnegativity** on $\mathbf{K}$.

- It requires $m+1$ s.o.s. weights σ_j.

Assumption 1:

For some $M > 0$, the quadratic polynomial $M - \|\mathbf{x}\|^2$ belongs to the quadratic module $\mathcal{Q}(g_1, \ldots, g_m)$

Assumption 1:

For some $M > 0$, the quadratic polynomial $M - \|\mathbf{x}\|^2$ belongs to the quadratic module $\mathcal{Q}(g_1, \ldots, g_m)$

Theorem (Putinar-Jacobi-Prestel)

Let $\mathbf{K}$ be compact and Assumption 1 hold. Then

$$[\, f \in \mathbb{R}[\mathbf{x}] \quad \textit{and} \quad f > 0 \textit{ on } \mathbf{K} \,] \Rightarrow f \in \mathcal{Q}(g_1, \ldots, g_m), \;\; i.e.,$$

$$f(\mathbf{x}) \; = \; \sigma_0(\mathbf{x}) + \sum_{j=1}^{m} \sigma_j(\mathbf{x}) \, g_j(\mathbf{x}), \qquad \forall x \in \mathbb{R}^n$$

for some s.o.s. polynomials $\{\sigma_j\}_{j=0}^{m}$.

• **If one fixes an a priori bound on the degree of the s.o.s. polynomials** $\{\sigma_j\}$, checking $f \in \mathcal{Q}(g_1, \ldots, g_m)$ reduces to solving an SDP!!

• **If one fixes an a priori bound on the degree of the s.o.s. polynomials** $\{\sigma_j\}$, checking $f \in \mathcal{Q}(g_1, \ldots, g_m)$ reduces to solving an SDP!!

• **Moreover, Assumption 1 holds true if e.g. :**

- all the g_j's are linear (hence $\mathbf{K}$ is a polytope), or if
- the set $\{ \mathbf{x} \mid g_j(\mathbf{x}) \geqslant 0\}$ is compact for some $j \in \{1, \ldots, m\}$

• **If one fixes an a priori bound on the degree of the s.o.s. polynomials** $\{\sigma_j\}$, checking $f \in \mathcal{Q}(g_1, \ldots, g_m)$ reduces to solving an SDP!!

• **Moreover, Assumption 1 holds true if e.g. :**

- all the g_j's are linear (hence $\mathbf{K}$ is a polytope), or if
- the set $\{ \mathbf{x} \mid g_j(\mathbf{x}) \geqslant 0\}$ is compact for some $j \in \{1, \ldots, m\}$

• **If** $\mathbf{x} \in \mathbf{K} \Rightarrow \|\mathbf{x}\| \leqslant M$ **for some (known)** M, then it suffices to add the redundant quadratic constraint $M^2 - \|\mathbf{x}\|^2 \geqslant 0$, in the definition of $\mathbf{K}$

• **If one fixes an a priori bound on the degree of the s.o.s. polynomials** $\{\sigma_j\}$, checking $f \in \mathcal{Q}(g_1, \ldots, g_m)$ reduces to solving an SDP!!

• **Moreover, Assumption 1 holds true if e.g. :**
- all the g_j's are linear (hence $\mathbf{K}$ is a polytope), or if
- the set $\{ \mathbf{x} \mid g_j(\mathbf{x}) \geqslant 0\}$ is compact for some $j \in \{1, \ldots, m\}$

• **If** $\mathbf{x} \in \mathbf{K} \Rightarrow \|\mathbf{x}\| \leqslant M$ **for some (known)** M, then it suffices to add the redundant quadratic constraint $M^2 - \|\mathbf{x}\|^2 \geqslant 0$, in the definition of $\mathbf{K}$

A representation in terms of s.o.s. (sums of squares) is interesting BECAUSE checking whether some given polynomial $f \in \mathbb{R}[\mathbf{x}]$ is s.o.s. reduces to solving an SDP ... that one may solve efficiently to arbitrary precision, in time polynomial in the input size!

A representation in terms of s.o.s. (sums of squares) is interesting BECAUSE checking whether some given polynomial $f \in \mathbb{R}[\mathbf{x}]$ is s.o.s. reduces to solving an SDP ... that one may solve efficiently to arbitrary precision, in time polynomial in the input size!

Indeed, let

$$v_d(\mathbf{x}) = (\mathbf{x}^\alpha), \qquad |\alpha| := \sum_i \alpha_i \leqslant d$$

be a basis of $\mathbb{R}[\mathbf{x}]_d$ (polynomials of degree at most d)

Let $f \in \mathbb{R}[\mathbf{x}]_{2d}$ be a s.o.s. polynomial, that is, $f = \sum_{k=1}^s q_k(\mathbf{x})^2$, for some polynomials $\{q_k\}_{k=1}^s \subset \mathbb{R}[\mathbf{x}]_d$.

Denote also $q_k = \{q_{k\alpha}\}_{\alpha \in \mathbb{N}^n}$, the vector of coefficients of the polynomial q_k, in the basis $v_d(\mathbf{x})$, that is,

$$q_k(\mathbf{x}) = \langle q_k, v_d(\mathbf{x}) \rangle = \sum_{|\alpha| \leqslant r} q_{k\alpha} \mathbf{x}^\alpha$$

and define the real symmetric matrix $Q := \sum_{k=1}^s q_k q_k^T \succcurlyeq 0$.

$$\langle v_d(\mathbf{x}), Q\, v_d(\mathbf{x}) \rangle = \sum_{k=1}^s \langle q_k, v_d(\mathbf{x}) \rangle^2 = \sum_{k=1}^s q_k(\mathbf{x})^2 = f(\mathbf{x})$$

Denote also $q_k = \{q_{k\alpha}\}_{\alpha \in \mathbb{N}^n}$, the vector of coefficients of the polynomial q_k, in the basis $v_d(\mathbf{x})$, that is,

$$q_k(\mathbf{x}) = \langle q_k, v_d(\mathbf{x}) \rangle = \sum_{|\alpha| \leqslant r} q_{k\alpha} \mathbf{x}^\alpha$$

and define the real symmetric matrix $Q := \sum_{k=1}^s q_k q_k^T \succcurlyeq 0$.

$$\langle v_d(\mathbf{x}), Q\, v_d(\mathbf{x}) \rangle = \sum_{k=1}^s \langle q_k, v_d(\mathbf{x}) \rangle^2 = \sum_{k=1}^s q_k(\mathbf{x})^2 = f(\mathbf{x})$$

Conversely, let $Q \succcurlyeq 0$ be a real $s(d) \times s(d)$ positive semidefinite symmetric matrix ($s(d)$ is the dimension of the vector space $\mathbb{R}[\mathbf{x}]_d$). As $Q \succcurlyeq 0$, write $Q = \sum_{k=1}^s q_k q_k^T$, so that

$$f(\mathbf{x}) = \langle v_d(\mathbf{x}), Q\, v_d(\mathbf{x}) \rangle = \sum_{k=1}^s \langle q_k, v_d(\mathbf{x}) \rangle^2 = \sum_{k=1}^s q_k(\mathbf{x})^2$$

is s.o.s.

Next, write the matrix $v_d(\mathbf{x})\, v_d(\mathbf{x})^T$ as:

$$v_d(\mathbf{x})\, v_d(\mathbf{x})^T \;=\; \sum_{\alpha \in \mathbb{N}^n_{2d}} \mathbf{B}_\alpha\, \mathbf{x}^\alpha,$$

so that checking whether

$$f(\mathbf{x}) = \langle v_d(\mathbf{x}), Q\, v_d(\mathbf{x}) \rangle \;=\; \langle Q, v_d(\mathbf{x})\, v_d(\mathbf{x})^T \rangle,$$

for some $Q \succcurlyeq 0$ reduces to checking the LMI

$$\begin{cases} \langle \mathbf{B}_\alpha, Q \rangle &=& f_\alpha, \quad \alpha \in \mathbb{N}^n,\ |\alpha| \leqslant 2d \\ Q &\succcurlyeq& 0 \end{cases}.$$

Example

Let $t \mapsto f(t) = 6 + 4t + 9t^2 - 4t^3 + 6t^4$. Is f an SOS? Do we have

$$f(t) = \begin{bmatrix} 1 \\ t \\ t^2 \end{bmatrix}^T \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \begin{bmatrix} 1 \\ t \\ t^2 \end{bmatrix}$$

Example

Let $t \mapsto f(t) = 6 + 4t + 9t^2 - 4t^3 + 6t^4$. Is f an SOS? Do we have

$$f(t) = \begin{bmatrix} 1 \\ t \\ t^2 \end{bmatrix}^T \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \begin{bmatrix} 1 \\ t \\ t^2 \end{bmatrix}$$

We must have:

$$a = 6\,;\, 2\,b = 4;\, d + 2\,c = 9;\, 2\,e = -4;\, f = 6.$$

And so we must find a scalar c such that

$$Q = \begin{bmatrix} 6 & 2 & c \\ 2 & 9 - 2c & -2 \\ c & -2 & 6 \end{bmatrix} \succcurlyeq 0.$$

With $c = -4$ we have

$$Q = \begin{bmatrix} 6 & 2 & -4 \\ 2 & 17 & -2 \\ -4 & -2 & 6 \end{bmatrix} \succcurlyeq 0.$$

and

$$Q = 2 \begin{bmatrix} \sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{bmatrix}' + 9 \begin{bmatrix} 2/3 \\ -1/3 \\ -2/3 \end{bmatrix} \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}'$$

$$+18 \begin{bmatrix} 1/\sqrt{(}18) \\ 4/\sqrt{(}18) \\ -1/\sqrt{(}18) \end{bmatrix} \begin{bmatrix} 1/\sqrt{(}18) \\ 4/\sqrt{(}18) \\ -1/\sqrt{(}18) \end{bmatrix}'$$

and so

$$f(t) = (1 + t^2)^2 + (2 - t - 2t^2)^2 + (1 + 4t - t^2)^2$$

II. DUAL side: The K-moment problem

Let $\{\mathbf{x}^\alpha\}$ be a canonical basis for $\mathbb{R}[\mathbf{x}]$, and let $y := \{y_\alpha\}$ be a given sequence indexed in that basis.

The K-moment problem

Given $\mathbf{K} \subset \mathbb{R}^n$, **does there exist a measure** μ on $\mathbf{K}$, such that

$$y_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha \, d\mu, \qquad \forall \alpha \in \mathbb{N}^n \quad ?$$

(where $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$).

Given $y = \{y_\alpha\}$, let $L_y : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$, be the linear functional

$$f \left(= \sum_\alpha f_\alpha \, \mathbf{x}^\alpha\right) \quad \mapsto \quad L_y(f) := \sum_{\alpha \in \mathbb{N}^n} f_\alpha \, y_\alpha .$$

Moment matrix $M_d(y)$

with rows and columns also indexed in the basis $\{\mathbf{x}^\alpha\}$.

$$M_d(y)(\alpha, \beta) := L_y(\mathbf{x}^{\alpha+\beta}) = y_{\alpha+\beta}, \quad \alpha, \beta \in \mathbb{N}^n, \quad |\alpha|, |\beta| \leqslant d.$$

For instance in $\mathbb{R}^2$: $\quad M_1(y) = \begin{bmatrix} \overbrace{y_{00}}^{1} & | & \overbrace{y_{10}}^{X_1} & \overbrace{y_{01}}^{X_2} \\ - & & - & - \\ y_{10} & | & y_{20} & y_{11} \\ y_{01} & | & y_{11} & y_{02} \end{bmatrix}$

Importantly ...

$$M_d(y) \succcurlyeq 0 \iff L_y(h^2) \geqslant 0, \qquad \forall h \in \mathbb{R}[X]_d$$

Localizing matrix

The "Localizing matrix" $M_d(\theta y)$ w.r.t. a polynomial $\theta \in \mathbb{R}[X]$
with $\mathbf{x} \mapsto \theta(\mathbf{x}) = \sum_\gamma \theta_\gamma \, \mathbf{x}^\gamma$, has its rows and columns also indexed in the basis $\{X^\alpha\}$ of $\mathbb{R}[\mathbf{x}]_d$, and with entries:

$$\begin{aligned} M_d(\theta\, y)(\alpha, \beta) &= L_y(\theta\, \mathbf{x}^{\alpha+\beta}) \\ &= \sum_{\gamma \in \mathbb{N}^n} \theta_\gamma \, y_{\alpha+\beta+\gamma}, \qquad \begin{cases} \alpha, \beta \in \mathbb{N}^n \\ |\alpha|, |\beta| \leqslant d. \end{cases} \end{aligned}$$

For instance, in $\mathbb{R}^2$, and with $X \mapsto \theta(\mathbf{x}) := 1 - x_1^2 - x_2^2$,

$$M_1(\theta\, y) = \begin{bmatrix} \overbrace{y_{00} - y_{20} - y_{02}}^{1} & \overbrace{y_{10} - y_{30} - y_{12}}^{x_1} & \overbrace{y_{01} - y_{21} - y_{03}}^{x_2} \\ y_{10} - y_{30} - y_{12} & y_{20} - y_{40} - y_{22} & y_{11} - y_{21} - y_{12} \\ y_{01} - y_{21} - y_{03} & y_{11} - y_{21} - y_{12} & y_{02} - y_{22} - y_{04} \end{bmatrix}$$

Importantly ...

$$M_d(\theta\, y) \succcurlyeq 0 \quad \Longleftrightarrow \quad L_y(h^2\, \theta) \geqslant 0, \qquad \forall h \in \mathbb{R}[\mathbf{x}]_d$$

Putinar's dual conditions

Again $\mathbf{K} := \{ \mathbf{x} \in \mathbb{R}^n \mid \quad g_j(\mathbf{x}) \geqslant 0, j = 1, \ldots, m\}$.

Assumption 1:

For some $M > 0$, the quadratic polynomial

$M - \|\mathbf{x}\|^2$ is in the quadratic module $\mathcal{Q}(g_1, \ldots, g_m)$

Putinar's dual conditions

Again $\mathbf{K} := \{ \mathbf{x} \in \mathbb{R}^n \mid g_j(\mathbf{x}) \geqslant 0, j = 1, \ldots, m\}$.

Assumption 1:

For some $M > 0$, the quadratic polynomial $M - \|\mathbf{x}\|^2$ is in the quadratic module $\mathcal{Q}(g_1, \ldots, g_m)$

Theorem (Putinar: dual side)

Let $\mathbf{K}$ be compact, and Assumption 1 hold. Then $y = \{y_\alpha\}$ has a representing measure μ on $\mathbf{K}$ if and only if

$$(**) \quad L_y(f^2) \geqslant 0; \quad L_y(f^2 g_j) \geqslant 0, \quad \forall j = 1, \ldots, m; \quad \forall f \in \mathbb{R}[\mathbf{x}].$$

Checking whether (**) holds for all $f \in \mathbb{R}[\mathbf{x}]$ with degree $\leqslant d$

reduces to checking whether $M_d(y) \succcurlyeq 0$ and $M_d(g_j\, y) \succcurlyeq 0$, for all $j = 1, \ldots, m$!

$\rightarrow m + 1$ LMI conditions to verify!

A Hierarchy of SDP Relaxations

Recall the PRIMAL LP

$$f^* = \min_{\mu \in \mathcal{M}(\mathbf{K})} \left\{ \int_{\mathbf{K}} f \, d\mu \; : \; \mu(\mathbf{K}) = 1 \right\}$$

where $\mathcal{M}(\mathbf{K})$ is the space of Borel measures on $\mathbf{K}$

A Hierarchy of SDP Relaxations

Let $\deg g_j = 2v_j$ or $2v_j - 1$. The SDP-relaxation $\mathbf{Q}_d$, $d \in \mathbb{N}$, reads:

$$\mathbf{Q}_d \begin{cases} \rho_d = \min\limits_{y} \quad L_y(f) & (\rightarrow \text{think of } \int f \, d\mu) \\ \text{s.t.} \quad \underbrace{M_{d-v_j}(g_j\, y) \succcurlyeq 0, \quad j = 0, \ldots m}_{\text{necessary conditions for } y_\alpha = \int \mathbf{x}^\alpha d\mu} & . \\ L_y(1) \ = \ 1 & (\rightarrow \mu(\mathbf{K}) = 1) \end{cases}$$

A Hierarchy of SDP Relaxations

... whose dual is the SDP

$$\mathbf{Q}_d^* \begin{cases} \rho_d^* = \max\limits_{\lambda, \{\sigma_j\}} \quad \lambda \\ \qquad \text{s.t.} \quad f - \lambda = \sum\limits_{k=0}^{m} \sigma_k \, g_k \\ \qquad \qquad \{\sigma_k\} \text{ are s.o.s.}; \quad \deg \sigma_0, \deg \sigma_k \, g_k \leqslant 2d \end{cases}$$

A Hierarchy of SDP Relaxations

Recall that $\mathbf{K} \subset \mathbb{R}^n$ is the basic semi-algebraic set

$$\mathbf{K} := \{ \ x \in \mathbb{R}^n \mid \quad g_j(\mathbf{x}) \geqslant 0, \, j = 1, \ldots, m \}.$$

Assumption 1:

For some $M > 0$, $M - \|\mathbf{x}\|^2$ is in $\mathcal{Q}(g_1, \ldots, g_m)$

A Hierarchy of SDP Relaxations

Recall that $\mathbf{K} \subset \mathbb{R}^n$ is the basic semi-algebraic set

$$\mathbf{K} := \{ \ x \in \mathbb{R}^n \mid \quad g_j(\mathbf{x}) \geqslant 0, \ j = 1, \ldots, m \}.$$

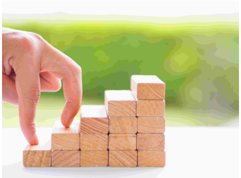
Assumption 1:

For some $M > 0$, $M - \|\mathbf{x}\|^2$ is in $\mathcal{Q}(g_1, \ldots, g_m)$

Theorem (Lasserre 01)

Let $\mathbf{K}$ be compact, and let Assumption 1 holds. Then:

- $\rho_d^* \leqslant \rho_d \leqslant f^*$ *for all* d *and* $\rho_d^*, \rho_d \uparrow f^*$ *as* $d \to \infty$*.*
- *If* $\operatorname{int} \mathbf{K} \neq \emptyset$*, then* $\rho_d = \rho_d^*$ *and the "sup" is attained.*



Global optimality check & extracting solutions

Exactness of a particular SDP-relaxation

Let y be an optimal solution of $\mathbf{Q}_d$ and let $2v \geqslant \max_j \deg g_j$. If

$$\operatorname{rank} M_d(y) = \operatorname{rank} M_{d-v}(y) \; (=: \; s)$$

then $\rho_d = f^*$

and one may extract s GLOBAL MINIMIZERS

Imperial College Press Optimization Series Vol. 1

Moments, Positive Polynomials and Their Applications

Many important problems in global optimization, algebra, probability and statistics, applied mathematics, control theory, financial mathematics, inverse problems, etc. can be modeled as a particular instance of the *Generalized Moment Problem* (GMP).

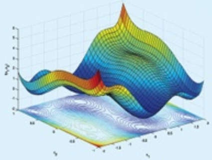
This book introduces, in a unified manual, a new general methodology to solve the GMP when its data are polynomials and basic semi-algebraic sets. This methodology combines semidefinite programming with recent results from real algebraic geometry to provide a hierarchy of semidefinite relaxations converging to the desired optimal value. Applied on appropriate cones, standard duality in convex optimization nicely expresses the duality between moments and positive polynomials.

In the second part of this invaluable volume, the methodology is particularized and described in detail for various applications, including global optimization, probability, optimal context, mathematical finance, multivariate integration, etc., and examples are provided for each particular application.

Imperial College Press
www.icpress.co.uk
ISBN-13 978-981-270-701-7
ISBN-10 981-270-701-8

Vol. 1
Moments, Positive Polynomials and Their Applications
Lasserre
ICP

Imperial College Press Optimization Series Vol. 1



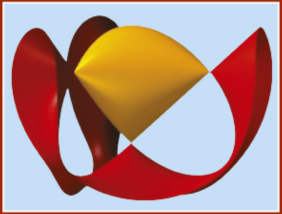
Moments, Positive Polynomials and Their Applications

Jean Bernard Lasserre

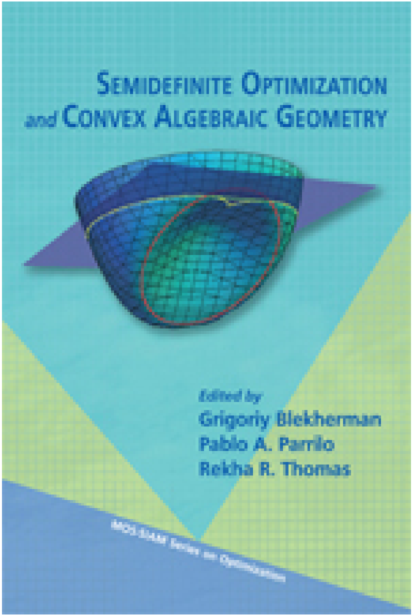
Imperial College Press

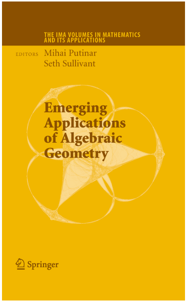
CAMBRIDGE TEXTS IN APPLIED MATHEMATICS

An Introduction to Polynomial and Semi-Algebraic Optimization

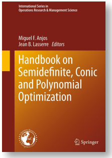


JEAN BERNARD LASSERRE


SEMIDEFINITE OPTIMIZATION
and CONVEX ALGEBRAIC GEOMETRY
Edited by
Grigoriy Blekherman
Pablo A. Parrilo
Rekha R. Thomas
MOS-SIAM Series on Optimization



See in particular the Chapter by M. Laurent



2012, 2012, XI, 957 p. 57 illus.

M.F. Anjos, Polytechnique Montréal, QC, Canada; **J.B. Lasserre**, LAAS, Toulouse Cedex 4, France (Eds.)

Handbook on Semidefinite, Conic and Polynomial Optimization

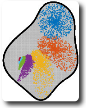
Series: International Series in Operations Research & Management Science

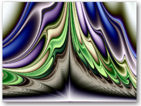
- **Summarizes research and developments of last ten years and brings field up to date**
- **Individual sections covering theory, algorithms, software and applications**
- **Editors are quite prominent in the field.**

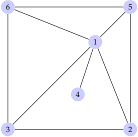
Semidefinite and conic optimization is a major and thriving research area within the optimization community. Although semidefinite optimization has been studied (under different names) since at least the 1940s, its importance grew immensely during the 1990s after polynomial-time interior-point methods for linear optimization were extended to solve semidefinite optimization problems.
Since the beginning of the 21st century, not only has research into semidefinite and conic optimization continued unabated, but also a fruitful interaction has developed with algebraic geometry through the close connections between semidefinite matrices and polynomial optimization. This has brought about important new results and led to an even higher level of research activity.
This Handbook on Semidefinite, Conic and Polynomial Optimization provides the reader with a snapshot of the state-of-the-art in the growing and mutually enriching areas of semidefinite optimization, conic optimization, and polynomial optimization. It contains a compendium of the recent research activity that has taken place in these thrilling areas, and will appeal to doctoral students, young graduates, and experienced researchers alike. The Handbook's thirty-one chapters are organized into four parts:Theory, covering significant theoretical developments as well as the interactions between conic optimization and polynomial optimization;Algorithms, documenting the directions of current algorithmic development;Software, providing an overview of the state-of-the-art;Applications, dealing with the application areas where semidefinite and conic optimization has made a significant impact in recent years.

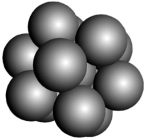
Next talks

1. Dynamical Systems $\oplus$ Roundoff errors $\oplus$ Verification



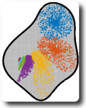


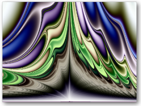


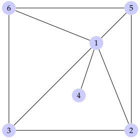


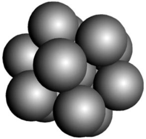
Next talks

1. Dynamical Systems ⊕ Roundoff errors ⊕ Verification

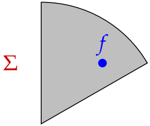






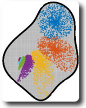
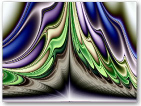
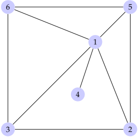


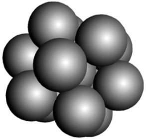
2. Exact certificates of positivity



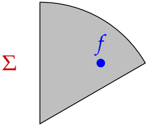
Next talks

1. Dynamical Systems ⊕ Roundoff errors ⊕ Verification



2. Exact certificates of positivity

Thank you for your attention!

homepages.laas.fr/vmagron