Moments and Positive Polynomials for Optimization II: LP- VERSUS SDP-relaxations

Jean B. Lasserre

LAAS-CNRS and Institute of Mathematics, Toulouse, France

EECI Course: February 2016
I. LP-representation of positive polynomials

II. Dual side: The K-moment problem

- LP-relaxations
- LP- VERSUS SDP-relaxations
I. LP-representation of positive polynomials
II: Dual side: The K-moment problem

- LP-relaxations
- LP- VERSUS SDP-relaxations
I. LP-representation of positive polynomials
II. Dual side: The K-moment problem

- LP-relaxations
- LP- VERSUS SDP-relaxations
Recall the Global Optimization problem $P$:

$$f^* := \min \{ f(x) \mid g_j(x) \geq 0, \, j = 1, \ldots, m \},$$

where $f$ and $g_j$ are all POLYNOMIALS, and let

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

be the feasible set (a compact basic semi-algebraic set)
I. LP-representation of positive polynomials

II: Dual side: The K-moment problem

Putinar Positivstellensatz

Assumption 1:
For some $M > 0$, the quadratic polynomial $M - \|X\|^2$ belongs to the quadratic module $Q(g_1, \ldots, g_m)$.

Theorem (Putinar-Jacobi-Prestel)
Let $K$ be compact and Assumption 1 hold. Then

\[ f \in \mathbb{R}[X] \text{ and } f > 0 \text{ on } K \] \Rightarrow \ f \in Q(g_1, \ldots, g_m), \ i.e.,

\[ f(x) = \sigma_0(x) + \sum_{j=1}^{m} \sigma_j(x) g_j(x), \quad \forall x \in \mathbb{R}^n \]

for some s.o.s. polynomials $\{\sigma_j\}_{j=0}^{m}$.
Assumption 1:
For some $M > 0$, the quadratic polynomial $M - \|X\|^2$ belongs to the quadratic module $Q(g_1, \ldots, g_m)$.

Theorem (Putinar-Jacobi-Prestel)

Let $K$ be compact and Assumption 1 hold. Then

\[
[f \in \mathbb{R}[X] \text{ and } f > 0 \text{ on } K] \implies f \in Q(g_1, \ldots, g_m), \ i.e.,
\]

\[
f(x) = \sigma_0(x) + \sum_{j=1}^{m} \sigma_j(x) g_j(x), \quad \forall x \in \mathbb{R}^n
\]

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

Jean B. Lasserre

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

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

• If \( x \in K \Rightarrow \|x\| \leq M \) for some (known) \( M \), then it suffices to add the redundant quadratic constraint \( M^2 - \|X\|^2 \geq 0 \), in the definition of \( K \).
• If one fixes an a priori bound on the degree of the s.o.s. polynomials \( \{\sigma_j\} \), checking \( f \in Q(g_1, \ldots, g_m) \) reduces to solving a SDP!!

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

• If \( x \in K \Rightarrow \|x\| \leq M \) for some (known) \( M \), then it suffices to add the redundant quadratic constraint \( M^2 - \|X\|^2 \geq 0 \), in the definition of \( K \).
• If one fixes an a priori bound on the degree of the s.o.s. polynomials \( \{ \sigma_j \} \), checking \( f \in Q(g_1, \ldots, g_m) \) reduces to solving a SDP!!

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

• If \( x \in K \Rightarrow \|x\| \leq M \) for some (known) \( M \), then it suffices to add the redundant quadratic constraint \( M^2 - \|X\|^2 \geq 0 \), in the definition of \( K \).
• If one fixes an a priori bound on the degree of the s.o.s. polynomials \( \{\sigma_j\} \), checking \( f \in Q(g_1, \ldots, g_m) \) reduces to solving a SDP!!

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

• If \( x \in K \Rightarrow \|x\| \leq M \) for some (known) \( M \), then it suffices to add the redundant quadratic constraint \( M^2 - \|X\|^2 \geq 0 \), in the definition of \( K \).
Assumption I:

With \( g_0 = 1 \), the family \( \{g_0, \ldots, g_m\} \) generates the algebra \( \mathbb{R}[x] \), that is, \( \mathbb{R}[x_1, \ldots, x_n] = \mathbb{R}[g_0, \ldots, g_m] \).

Assumption II:

Recall that \( K \) is compact. Hence we also assume with no loss of generality (but possibly after scaling) that for every \( j = 1, \ldots, m \):

\[
0 \leq g_j(x) \leq 1 \quad \forall x \in K.
\]
Krivine-Handelman-Vasilescu Positivstellensatz

Assumption I:
With $g_0 = 1$, the family $\{g_0, \ldots, g_m\}$ generates the algebra $\mathbb{R}[x]$, that is, $\mathbb{R}[x_1, \ldots, x_n] = \mathbb{R}[g_0, \ldots, g_m]$.

Assumption II:
Recall that $K$ is compact. Hence we also assume with no loss of generality (but possibly after scaling) that for every $j = 1, \ldots, m$:

$$0 \leq g_j(x) \leq 1 \quad \forall x \in K.$$
Notation: for $\alpha, \beta \in \mathbb{N}^m$, let

$$g(x)^\alpha = g_1(x)^{\alpha_1} \cdots g_m(x)^{\alpha_m}$$

$$\quad \quad \quad (1 - g(x))^\beta = (1 - g_1(x))^{\beta_1} \cdots (1 - g_m(x))^{\beta_m}$$

Theorem (Krivine,Vasilescu Positivstellensatz)

Let Assumption I and Assumption II hold:

If $f \in \mathbb{R}[x_1, \ldots, x_m]$ is POSITIVE on $K$ then

$$f(x) = \sum_{\alpha,\beta \in \mathbb{N}^m} c_{\alpha\beta} g(x)^\alpha (1 - g(x))^\beta, \quad \forall x \in \mathbb{R}^n,$$

for finitely many positive coefficients $(c_{\alpha\beta})$. 
Notation: for \( \alpha, \beta \in \mathbb{N}^m \), let

\[
\begin{align*}
g(x)^\alpha &= g_1(x)^{\alpha_1} \cdots g_m(x)^{\alpha_m} \\
(1 - g(x))^\beta &= (1 - g_1(x))^{\beta_1} \cdots (1 - g_m(x))^{\beta_m}
\end{align*}
\]

**Theorem (Krivine, Vasilescu Positivstellensatz)**

Let Assumption I and Assumption II hold:

If \( f \in \mathbb{R}[x_1, \ldots, x_m] \) is **POSITIVE** on \( K \) then

\[
f(x) = \sum_{\alpha, \beta \in \mathbb{N}^m} c_{\alpha \beta} g(x)^\alpha (1 - g(x))^\beta, \quad \forall x \in \mathbb{R}^n,
\]

for **finitely many positive coefficients** \( (c_{\alpha \beta}) \).
Testing whether

\[ f(x) = \sum_{\alpha,\beta \in \mathbb{N}^m} c_{\alpha\beta} \, g(x)^\alpha \left(1 - g(x)\right)^\beta, \quad \forall \, x \in \mathbb{R}^n, \]

for finitely many positive coefficients \((c_{\alpha\beta})\) and with

\[ \sum_i \alpha_i + \beta_i \leq d \]

... reduces to solving a LP!.

Indeed, recall that \(f(x) = \sum_{\gamma} f_{\gamma} \, x^\gamma\). So expand

\[ \sum_{\alpha,\beta \in \mathbb{N}^m} c_{\alpha\beta} \, g(x)^\alpha \left(1 - g(x)\right)^\beta = \sum_{\gamma \in \mathbb{N}^n} \theta_{\gamma}(c) \, x^\gamma \]

and state that

\[ f_{\gamma} = \theta_{\gamma}(c), \quad \forall \gamma \in \mathbb{N}^n; \quad c \geq 0. \quad \rightarrow \quad \text{a linear system!} \]
Testing whether

\[ f(x) = \sum_{\alpha, \beta \in \mathbb{N}^m} c_{\alpha \beta} g(x)^\alpha (1 - g(x))^\beta, \quad \forall x \in \mathbb{R}^n, \]

for finitely many positive coefficients \((c_{\alpha \beta})\) and with
\[ \sum_i \alpha_i + \beta_i \leq d \]

... reduces to solving a LP!.

Indeed, recall that \(f(x) = \sum_{\gamma} f_{\gamma} x^\gamma\). So expand

\[ \sum_{\alpha, \beta \in \mathbb{N}^m} c_{\alpha \beta} g(x)^\alpha (1 - g(x))^\beta = \sum_{\gamma \in \mathbb{N}^n} \theta_{\gamma}(c) x^\gamma \]

and state that

\[ f_{\gamma} = \theta_{\gamma}(c), \quad \forall \gamma \in \mathbb{N}_{2d}^n; \quad c \geq 0. \rightarrow \text{a linear system!} \]
Testing whether

\[ f(x) = \sum_{\alpha, \beta \in \mathbb{N}^m} c_{\alpha \beta} g(x)^\alpha (1 - g(x))^\beta, \quad \forall x \in \mathbb{R}^n, \]

for finitely many positive coefficients \((c_{\alpha \beta})\) and with \(\sum_i \alpha_i + \beta_i \leq d\)

... reduces to solving a LP!.

Indeed, recall that \(f(x) = \sum_{\gamma} f_\gamma x^\gamma\). So expand

\[ \sum_{\alpha, \beta \in \mathbb{N}^m} c_{\alpha \beta} g(x)^\alpha (1 - g(x))^\beta = \sum_{\gamma \in \mathbb{N}^n} \theta_\gamma(c) x^\gamma \]

and state that

\[ f_\gamma = \theta_\gamma(c), \quad \forall \gamma \in \mathbb{N}_2^d; \quad c \geq 0. \rightarrow \text{a linear system!} \]
Let \( \{X^\alpha\} \) be a canonical basis for \( \mathbb{R}[X] \), and let \( y := \{y_\alpha\} \) be a given sequence indexed in that basis.

Recall the K-moment problem

Given \( K \subset \mathbb{R}^n \), does there exist a measure \( \mu \) on \( K \), such that

\[
y_\alpha = \int_K X^\alpha \, d\mu, \quad \forall \alpha \in \mathbb{N}^n,
\]

(where \( X^\alpha = X_1^{\alpha_1} \cdots X_n^{\alpha_n} \)).
Given $y = \{y_\alpha\}$, let $L_y : \mathbb{R}[X] \to \mathbb{R}$, be the linear functional

$$f (= \sum_{\alpha} f_\alpha X^\alpha) \mapsto L_y(f) := \sum_{\alpha \in \mathbb{N}^n} f_\alpha y_\alpha.$$ 

Moment matrix $M_d(y)$

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

$$M_d(y)(\alpha, \beta) := L_y(X^{\alpha+\beta}) = y_{\alpha+\beta}, \quad \alpha, \beta \in \mathbb{N}^n, \quad |\alpha|, |\beta| \leq d.$$
For instance in $\mathbb{R}^2$: $M_1(y) =$

\[
\begin{bmatrix}
1 & x_1 & x_2 \\
y_{00} & y_{10} & y_{01} \\
y_{10} & y_{20} & y_{11} \\
y_{01} & y_{11} & y_{02}
\end{bmatrix}
\]

Importantly . . .

$M_d(y) \succeq 0 \iff L_y(h^2) \geq 0, \quad \forall h \in \mathbb{R}[X]_d$
Localizing matrix

The “Localizing matrix” $M_d(\theta y)$ w.r.t. a polynomial $\theta \in \mathbb{R}[X]$

with $X \mapsto \theta(X) = \sum \theta_\gamma X^\gamma$, has its rows and columns also indexed in the basis $\{X^\alpha\}$ of $\mathbb{R}[X]_d$, and with entries:

$$M_d(\theta y)(\alpha, \beta) = L_y(\theta X^{\alpha+\beta})$$

$$= \sum_{\gamma \in \mathbb{N}^n} \theta_\gamma y_{\alpha+\beta+\gamma}, \quad \left\{ \begin{array}{l} \alpha, \beta \in \mathbb{N}^n \\ |\alpha|, |\beta| \leq d. \end{array} \right.$$
For instance, in $\mathbb{R}^2$, and with $X \mapsto \theta(X) := 1 - X_1^2 - X_2^2$,

$$M_1(\theta y) = \begin{bmatrix}
1 & X_1 & X_2 \\
Y_{00} - Y_{20} - Y_{02} & Y_{10} - Y_{30} - Y_{12} & Y_{01} - Y_{21} - Y_{03} \\
Y_{10} - Y_{30} - Y_{12} & Y_{20} - Y_{40} - Y_{22} & Y_{11} - Y_{21} - Y_{12} \\
Y_{01} - Y_{21} - Y_{03} & Y_{11} - Y_{21} - Y_{12} & Y_{02} - Y_{22} - Y_{04}
\end{bmatrix}.$$
Putinar’s dual conditions

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

**Assumption 1**: For some $M > 0$, the quadratic polynomial $M - \|X\|^2$ is in the quadratic module $Q(g_1, \ldots, g_m)$.

**Theorem (Putinar: dual side)**

Let $K$ be compact, and Assumption 1 hold.

Then a sequence $y = (y_\alpha), \ \alpha \in \mathbb{N}^n$, has a representing measure $\mu$ on $K$ if and only if

\[(**) \quad L_y(f^2) \geq 0; \quad L_y(f^2 g_j) \geq 0, \quad \forall j = 1, \ldots, m; \quad \forall f \in \mathbb{R}[X].\]
Putinar’s dual conditions

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

Assumption 1: For some $M > 0$, the quadratic polynomial $M - \|X\|^2$ is in the quadratic module $Q(g_1, \ldots, g_m)$.

Theorem (Putinar: dual side)

Let $K$ be compact, and Assumption 1 hold.

Then a sequence $y = (y_\alpha)$, $\alpha \in \mathbb{N}^n$, has a representing measure $\mu$ on $K$ if and only if

$$\text{(**) } L_y(f^2) \geq 0; \quad L_y(f^2 g_j) \geq 0, \quad \forall j = 1, \ldots, m; \quad \forall f \in \mathbb{R}[X].$$
Checking whether \( \text{(**)} \) holds for all \( f \in \mathbb{R}[X] \) with degree \( \leq d \) 

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

\[ \rightarrow m + 1 \text{ LMI conditions to verify!} \]
I. LP-representation of positive polynomials

II: Dual side: The K-moment problem

Krivine-Vasilescu: dual side

Theorem

Let $K$ be compact, and Assumption I and II hold.

Then the sequence $y = (y_\alpha), \alpha \in \mathbb{N}^n$, has a representing measure $\mu$ on $K$ if and only if

$$L_y(g^\alpha (1 - g)^\beta) \geq 0, \quad \forall \alpha, \beta \in \mathbb{N}^m.$$
With \( f \in \mathbb{R}[x] \), consider the hierarchy of LP-relaxations

\[
\rho^*_d = \max_{\lambda, c_{\alpha\beta}} \lambda \\
\left\{
\begin{array}{c}
\rho_d = \max_{\lambda, c_{\alpha\beta}} \lambda \\
f - \lambda = \sum_{\alpha, \beta \in \mathbb{N}^m} c_{\alpha\beta} g^\alpha (1 - g)^\beta \\
c_{\alpha\beta} \geq 0, \quad \forall |\alpha + \beta| \leq 2d
\end{array}
\right.
\]

and of course \( \rho_d = \rho^*_d \) for all \( d \).
Theorem

Assume that $K$ is compact and Assumption I and II hold. Then the LP-relaxations CONVERGE, that is,

$$\rho_d \uparrow f^* \quad \text{as } d \to \infty.$$ 

- The SHERALI-ADAMS RLT’s hierarchy is exactly this type of LP-relaxations.
- Its convergence for 0/1 programs was proved with ah-hoc arguments.
- In fact, the rationale behind such convergence if Krivine-Vasilescu Positivstellensatz.
I. LP-representation of positive polynomials
II: Dual side: The K-moment problem

Theorem

Assume that $K$ is compact and Assumption I and II hold. Then the LP-relaxations CONVERGE, that is,

$$\rho_d \uparrow f^* \text{ as } d \to \infty.$$  

- The SHERALI-ADAMS RLT’s hierarchy is exactly this type of LP-relaxations.
- Its convergence for 0/1 programs was proved with ah-hoc arguments.
- In fact, the rationale behind such convergence if Krivine-Vasilescu Positivstellensatz.
Theorem

Assume that $K$ is compact and Assumption I and II hold. Then the LP-relaxations CONVERGE, that is,

$$\rho_d \uparrow f^* \quad \text{as } d \to \infty.$$ 

- The SHERALI-ADAMS RLT's hierarchy is exactly this type of LP-relaxations.
- Its convergence for 0/1 programs was proved with ah-hoc arguments.
- In fact, the rationale behind such convergence if Krivine-Vasilescu Positivstellensatz.
Theorem

Assume that $K$ is compact and Assumption I and II hold. Then the LP-relaxations CONVERGE, that is,

$$\rho_d \uparrow f^* \text{ as } d \to \infty.$$

- The SHERALI-ADAMS RLT’s hierarchy is exactly this type of LP-relaxations.
- Its convergence for 0/1 programs was proved with ah-hoc arguments.
- In fact, the rationale behind such convergence if Krivine-Vasilescu Positivstellensatz.
Some remarks on LP-relaxations

1. Notice the presence of binomial coefficients in both primal and dual LP-relaxations ... which yields numerical ill-conditioning for relatively large $d$.

2. Let $x^* \in K$ be a global minimizer, and for $x \in K$, let $J(x)$ be the set of active constraints, i.e., $g_j(x) = 0$ or $1 - g_k(x) = 0$.

Then FINITE convergence CANNOT occur if there exists nonoptimal $x \in K$ with $J(x) \supseteq J(x^*)$!

→ And so ... not possible for CONVEX problems in general!

For instance, if $K$ is a Polytope then FINITE convergence is possible only if every global minimizer is a vertex of $K$!
Some remarks on LP-relaxations

1. Notice the presence of \textit{binomial coefficients} in both primal and dual LP-relaxations ... which yields \textit{numerically ill-conditioning} for relatively large $d$.

2. Let $x^* \in K$ be a global minimizer, and for $x \in K$, let $J(x)$ be the set of active constraints, i.e., $g_j(x) = 0$ or $1 - g_k(x) = 0$.

Then \textit{FINITE convergence CANNOT} occur if there exists nonoptimal $x \in K$ with $J(x) \supset J(x^*)$!

\rightarrow And so ... not possible for \textit{CONVEX} problems in general!

For instance, if $K$ is a Polytope then \textit{FINITE convergence is possible only if every global minimizer is a vertex of $K$}!
Some remarks on LP-relaxations

1. Notice the presence of binomial coefficients in both primal and dual LP-relaxations ... which yields numerically ill-conditioning for relatively large $d$.

2. Let $x^* \in K$ be a global minimizer, and for $x \in K$, let $J(x)$ be the set of active constraints, i.e., $g_j(x) = 0$ or ... $1 - g_k(x) = 0$.

Then FINITE convergence CANNOT occur if there exists nonoptimal $x \in K$ with $J(x) \supseteq J(x^*)$!

→ And so ... not possible for CONVEX problems in general!

For instance, if $K$ is a Polytope then FINITE convergence is possible only if every global minimizer is a vertex of $K$!
1. Notice the presence of binomial coefficients in both primal and dual LP-relaxations ... which yields numerical ill-conditioning for relatively large $d$.

2. Let $x^* \in K$ be a global minimizer, and for $x \in K$, let $J(x)$ be the set of active constraints, i.e., $g_j(x) = 0$ or $1 - g_k(x) = 0$.

Then finite convergence CANNOT occur if there exists nonoptimal $x \in K$ with $J(x) \supseteq J(x^*)$!

→ And so ... not possible for CONVEX problems in general!

For instance, if $K$ is a Polytope then finite convergence is possible only if every global minimizer is a vertex of $K$!
Some remarks on LP-relaxations

1. Notice the presence of binomial coefficients in both primal and dual LP-relaxations ... which yields numerically ill-conditioning for relatively large $d$.

2. Let $x^* \in K$ be a global minimizer, and for $x \in K$, let $J(x)$ be the set of active constraints, i.e., $g_j(x) = 0$ or ... $1 - g_k(x) = 0$.

Then FINITE convergence CANNOT occur if there exists nonoptimal $x \in K$ with $J(x) \supseteq J(x^*)$!

→ And so ... not possible for CONVEX problems in general!

For instance, if $K$ is a Polytope then FINITE convergence is possible only if every global minimizer is a vertex of $K$!
LP- and SDP-Relaxations with their dual

<table>
<thead>
<tr>
<th>Primal LP-relaxation</th>
<th>Primal SDP-relaxation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \min_y L_y(f) )</td>
<td>( \min_y L_y(f) )</td>
</tr>
<tr>
<td>( L_y(g^\alpha (1 - g)^\beta) \geq 0 )</td>
<td>( L_y(h^2 g_j) \geq 0, \ j = 1, \ldots, m )</td>
</tr>
<tr>
<td>( \forall \alpha, \beta \in \mathbb{N}^m, \</td>
<td>\alpha + \beta</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dual LP-relaxation</th>
<th>Dual SDP-relaxation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \max_{\lambda, {c_{\alpha \beta}}} \lambda )</td>
<td>( \max_{\lambda, {\sigma_j}} \lambda )</td>
</tr>
<tr>
<td>( f - \lambda = \sum_{\alpha, \beta \in \mathbb{N}^m} c_{\alpha \beta} g^\alpha (1 - g)^\beta )</td>
<td>( f - \lambda = \sum_{j=0}^{m} \sigma_j g_j )</td>
</tr>
<tr>
<td>( c_{\alpha \beta} &gt; 0; \</td>
<td>\alpha + \beta</td>
</tr>
</tbody>
</table>