"Global Optimization with Polynomials"

Geoffrey Schiebinger, Stephen Kemmerling

Math 301, 2010/2011

March 16, 2011
"Global Optimization with Polynomials and the problem of moments", by Jean B. Lasserre (2001)

Goal: Solve \( \min_{x \in K} p(x) \), \( p \) arbitrary polynomial,
\( K = \bigcap \{x|g_i(x) \geq 0\} \), \( g_i \) arbitrary Polynomials.

Result: Possible as Sequence of SDPs, approaching the solution.
Outline

- The Moment Problem. Equivalence.
- SDP relaxation. Exactness.
- General unconstrained case ($p$ not S.O.S., $K = \mathbb{R}^n$).
- Constrained case.
- Detecting Optimality.
- Generalizations/Applications.
- Conclusions.
The Problem of Moments I

1. Given a polynomial $p : \mathbb{R}^n \to \mathbb{R}$, consider $\mathbb{P} \mapsto p^* := \min_{x \in \mathbb{R}^n} p(x)$.
2. Moment Formulation: $\mathcal{P} \mapsto p^* := \min_{\mu \in \mathcal{P}(\mathbb{R}^n)} \int p(x) d\mu$.
3. Assumption: Minimizer always exists.

Theorem: $\mathbb{P}$ and $\mathcal{P}$ are equivalent. Specifically

(a) $\inf \mathbb{P} = \inf \mathcal{P}$
(b) $x^* = \text{argmin} \mathbb{P} \Rightarrow \mu^* = \delta_{x^*} = \text{argmin} \mathcal{P}$
(c) $\mu^* = \text{argmin} \mathcal{P} \Rightarrow p(x) = \min \mathbb{P}, \mu^* - a.e.$
Theorem: $\mathbb{P}$ and $\mathcal{P}$ are equivalent. Specifically

(a) $\inf \mathbb{P} = \inf \mathcal{P}$

Proof. We have $p(x) = \int p \, d\delta_x$, thus $\inf \mathcal{P} \leq \inf \mathbb{P}$. Conversely let $p^* := \inf \mathbb{P}$. Then, since $p(x) \geq p^* \ \forall x$, we have $\inf \mathcal{P} = \inf_\mu \int p \, d\mu \geq p^* = \inf \mathbb{P}$. 
The Problem of Moments II

Theorem: $\mathbb{P}$ and $\mathcal{P}$ are equivalent. Specifically

(a) $\inf \mathbb{P} = \inf \mathcal{P}$

Proof. We have $p(x) = \int p \, d\delta_x$, thus $\inf \mathcal{P} \leq \inf \mathbb{P}$.

Conversely let $p^* := \inf \mathbb{P}$. Then, since $p(x) \geq p^* \forall x$, we have

$\inf \mathcal{P} = \inf_{\mu} \int p \, d\mu \geq p^* = \inf \mathbb{P}$. 

(b) $x^* = \arg\min \mathbb{P} \Rightarrow \mu^* = \delta_{x^*} = \arg\min \mathcal{P}$

Proof. Immediate from $p(x^*) \leq p(x) \forall x$. 

The Problem of Moments II

Theorem: $\mathbb{P}$ and $\mathcal{P}$ are equivalent. Specifically

(a) $\inf \mathbb{P} = \inf \mathcal{P}$

*Proof.* We have $p(x) = \int p \, d\delta_x$, thus $\inf \mathcal{P} \leq \inf \mathbb{P}$. Conversely, let $p^* := \inf \mathbb{P}$. Then, since $p(x) \geq p^* \forall x$, we have $\inf \mathcal{P} = \inf_{\mu} \int p \, d\mu \geq p^* = \inf \mathbb{P}$.

(b) $x^* = \text{argmin} \mathbb{P} \Rightarrow \mu^* = \delta_{x^*} = \text{argmin} \mathcal{P}$

*Proof.* Immediate from $p(x^*) \leq p(x) \forall x$.

(c) $\mu^* = \text{argmin} \mathcal{P} \Rightarrow p(x) = \min \mathbb{P}, \mu^* - a.e.$

*Proof.* Let $B \subset \mathbb{R}^n$, with $\mu^*(B) > 0$ and $p(x) > p^* \forall x \in B$. Then $\int p \, d\mu^* = \int_B p \, d\mu^* + \int_{\mathbb{R}^n - B} p \, d\mu^* > p^*$. Contradiction to (a).
The Problem of Moments III

With \( p = \sum_{\alpha} p_{\alpha} x^{\alpha} : \int p(x) d\mu = \sum_{\alpha} p_{\alpha} \int x^{\alpha} d\mu = \sum_{\alpha} p_{\alpha} y_{\alpha} \)

Thus:

\[
\mathcal{P} \begin{cases} 
\min_{y} \sum_{\alpha} p_{\alpha} y_{\alpha} \\
\text{s.t. } y_{\alpha} \text{ are moments}
\end{cases}
\]

Relaxation:

\[
\mathcal{Q} \begin{cases} 
\min_{y} \sum_{\alpha} p_{\alpha} y_{\alpha} \\
\text{s.t. } M_{m}(y) \succeq 0
\end{cases}
\]

where \( M_{m}(y) \) is the moment matrix up to degree \( m \), i.e. its entries are \( y_{\alpha} \) with \( \sum_{i} \alpha_{i} \leq m \) and \( \langle p, M_{m}(y)p \rangle \geq \int p^{2} d\mu_{y} \). Slater's Condition holds for \( \mathcal{Q} \). \( \mathcal{Q} = \mathcal{P} \) if \( p(x) - p^{*} \) is S.O.S.
Theorem: Let \( p(x) = : \mathbb{R}^n \to \mathbb{R} \) be a 2m-degree polynomial of the form \( \sum_\alpha p_\alpha x^\alpha \) with global minimum \( p^* = \min \mathbb{P} \) and such that \( \|x^*\| \leq a \) for some \( a > 0 \) at some global minimizer \( x^* \).

Then as \( N \to \infty \), one has

\[
\inf Q_a^N \uparrow p^*.
\]

Here \( Q_a^N \) is the convex LMI problem:

\[
\begin{cases}
\inf_y \sum_\alpha p_\alpha y_\alpha \\
M_N(y) \succeq 0, \\
M_{N-1}(\theta y) \succeq 0.
\end{cases}
\]

and \( \theta(x) = a - \|x\|^2 \), \( M_m(\theta y)(i,j) = \sum_\alpha \theta_\alpha y_{\{\beta(i,j)+\alpha\}} \).

What really matters: \( \langle v, M_m(\theta y)v \rangle = \int \theta(x)v(x)^2 \mu_y(dx) \)
Writing $M_N(y) = \sum_{\alpha} y_{\alpha} B_{\alpha}$ for appropriate matrices $\{B_{\alpha}\}$ and $M_{N-1}(\theta y) = \sum_{\alpha} y_{\alpha} C_{\alpha}$ for appropriate matrices $\{C_{\alpha}\}$, we can express the dual

$$
(\mathbb{Q}^N_a)^* \begin{cases} 
\sup_{X, Z \succeq 0} & -X(1, 1) - a^2 Z(1, 1), \\
\langle X, B_{\alpha} \rangle + \langle Z, C_{\alpha} \rangle & = p_{\alpha}, \alpha \neq 0
\end{cases}
$$

Let $K_a = \{x : \theta(x) > 0\}$. 

---

Geoffrey Schiebinger, Stephen Kemmerling (Math 301, 2010/2011)

"Global Optimization with Polynomials"
Writing $M_N(y) = \sum_{\alpha} y_{\alpha} B_{\alpha}$ for appropriate matrices $\{B_{\alpha}\}$ and $M_{N-1}(\theta y) = \sum_{\alpha} y_{\alpha} C_{\alpha}$ for appropriate matrices $\{C_{\alpha}\}$, we can express the dual

$$(\mathbb{Q}^N_a)^* \begin{cases} \sup_{X,Z \succeq 0} -X(1,1) - a^2 Z(1,1), \\ \langle X, B_{\alpha} \rangle + \langle Z, C_{\alpha} \rangle = p_{\alpha}, \alpha \neq 0 \end{cases}$$

Let $K_a = \{x : \theta(x) > 0\}$.

Fact: For all $p(x)$ strictly positive on $K_a$, we can write

$$p(x) = \sum_{i=1}^{r_1} q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} t_j(x)^2$$

(See, e.g. Berg (1980), "The multidimensional moment problem and semi-groups").
Proof

From $x^\ast \in K_a$, and with $y^\ast = (x_1^\ast, \ldots, (x_1^\ast)^{2N}, \ldots, (x_n^\ast)^{2N})$, it follows that $M_N(y^\ast), M_{N-1}(\theta y^\ast) \succeq 0$ so that $y^\ast$ is admissible for $Q_a^N$ and thus $\inf Q_a^N \leq p^\ast$.
Proof

- From \( x^* \in K_a \), and with \( y^* = (x_1^*, \ldots, (x_1^*)^{2N}, \ldots, (x_n^*)^{2N}) \), it follows that \( M_N(y^*), M_{N-1}(\theta y^*) \geq 0 \) so that \( y^* \) is admissible for \( Q_N^a \) and thus \( \inf Q_N^a \leq p^* \)
- Let \( \epsilon > 0. \) \( p(x) - (p^* - \epsilon) > 0, \) so \( \exists N_0 \) such that
  \[
p(x) - p^* + \epsilon = \sum_{i=1}^{r_1} q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} t_j(x)^2
\]
f for some polynomials \( q_i \) of degree at most \( N_0 \), and polynomials \( t_j \) of degree at most \( N_0 - 1 \).
Proof

- From \( x^* \in K_a \), and with \( y^* = (x_1^*, \ldots, (x_1^{2N}), \ldots, (x_n^{2N})) \), it follows that \( M_N(y^*), M_{N-1}(\theta y^*) \succeq 0 \) so that \( y^* \) is admissible for \( Q_a^N \) and thus \( \inf Q_a^N \leq p^* \).

- Let \( \epsilon > 0 \). \( p(x) - (p^* - \epsilon) > 0 \), so \( \exists N_0 \) such that

\[
p(x) - p^* + \epsilon = \sum_{i=1}^{r_1} q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} t_j(x)^2
\]

for some polynomials \( q_i \) of degree at most \( N_0 \), and polynomials \( t_j \) of degree at most \( N_0 - 1 \).

\[
X = \sum_{i=1}^{r_1} q_i q_i', \quad Z = \sum_{i=1}^{r_2} t_j t_j', \quad X, Z \succeq 0
\]
Proof

- From $x^* \in K_a$, and with $y^* = (x_1^*, \ldots, (x_1^*)^{2N}, \ldots, (x_n^*)^{2N})$, it follows that $M_N(y^*), M_{N-1}(\theta y^*) \geq 0$ so that $y^*$ is admissible for $Q_a^N$ and thus $\inf Q_a^N \leq p^*$.

- Let $\epsilon > 0$. $p(x) - (p^* - \epsilon) > 0$, so $\exists N_0$ such that

$$p(x) - p^* + \epsilon = \sum_{i=1}^{r_1} q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} t_j(x)^2$$

for some polynomials $q_i$ of degree at most $N_0$, and polynomials $t_j$ of degree at most $N_0 - 1$.

$$X = \sum_{i=1}^{r_1} q_i q_i', \quad Z = \sum_{i=1}^{r_2} t_j t_j'$$

$X, Z \succeq 0$

- Therefore $(X, Z)$ admissible for $(Q_a^{N_0})^*$ with value

$$-X(1, 1) - a^2 Z(1, 1) = p^* - \epsilon,$$

and therefore

$$p^* - \epsilon \leq \inf Q_a^{N_0} \leq p^*$$
Optimality Conditions

\[ \inf \mathbb{Q}_a^N = p^* \iff p(x) - p^* = \sum_{i=1}^{r_1} q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} t_j(x)^2, \]
with \( \deg(q_i) \leq N, \deg(t_j) \leq N - 1. \)
Optimality Conditions

- \( \inf_{Q_a^N} = p^* \text{ iff } p(x) - p^* = \sum_{i=1}^{r_1} q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} t_j(x)^2, \) 
  with \( \deg(q_i) \leq N, \deg(t_j) \leq N - 1. \)

- Practical sufficient condition: \( \text{Rank } M_N(y) = \text{Rank } M_{N-1}(y). \)

(See, e.g. Curto, Fialkow (2000): "The truncated complex K-moment problem")
Optimality Conditions

\[ \inf Q_a^N = p^* \iff p(x) - p^* = \sum_{i=1}^{r_1} q_i(x)^2 + \theta(x) \sum_{j=1}^{r_2} t_j(x)^2, \]

with \( \deg(q_i) \leq N \), \( \deg(t_j) \leq N - 1 \).

- Practical sufficient condition: \( \text{Rank } M_N(y) = \text{Rank } M_{N-1}(y) \).

  (See, e.g. Curto, Fialkow (2000): "The truncated complex K-moment problem")

- Extraction of optimal point possible with SVD of \( M_N(y) \).

  (See, e.g. Henrion, Lasserre (2005): "Detecting global optimality and extracting solutions in GloptiPoly")
Constrained Case

Consider $P_K \mapsto p^* := \min_{x \in K} p(x)$, with $K = \bigcap \{x | g_i(x) \geq 0\}$, $g_i$ arbitrary polynomials.

Assumption: $x \in K \Rightarrow ||x||^2 \leq a$ for some $a$. (Weaker Possible!)

Then, analogous to the unconstrained case, let

$$Q_N^K \left\{ \begin{array}{l}
\min_y \sum_\alpha p_\alpha y_\alpha \\
\text{s.t. } M_N(y) \succeq 0 \\
M_{N-\lceil \deg(g_i)/2 \rceil}(g_i y) \succeq 0, \; i = 1, \ldots, r.
\end{array} \right.$$ 

and we have $\inf Q_N^K \uparrow p^*_K$ as $N \to \infty$. Proof proceeds as in the unconstrained case, but using that (given assumption)

$$p(x) = \sum_{i=1}^{r_1} q_i(x)^2 + \sum_{k=1}^{r} g_k(x) \sum_{j=1}^{r_2} t_j(x)^2$$

for some $q_i$, $t_j$.

(See, e.g. Jacobi,Prestel (2000), "On Special Representations of strictly positive polynomials").
The rate of convergence in $\inf \mathcal{Q}_a^N \uparrow p^*$ is unknown.

Also, solving $\mathcal{Q}_N$ gets expensive very quickly:

$$M_2(y) = \begin{bmatrix}
1 & y_{1,0} & y_{0,1} & y_{2,0} & y_{1,1} & y_{0,2} \\
y_{1,0} & y_{2,0} & y_{1,1} & y_{3,0} & y_{2,1} & y_{1,2} \\
y_{0,1} & y_{1,1} & y_{0,2} & y_{2,1} & y_{1,2} & y_{0,3} \\
y_{2,0} & y_{3,0} & y_{2,1} & y_{4,0} & y_{3,1} & y_{2,2} \\
y_{1,1} & y_{2,1} & y_{1,2} & y_{3,1} & y_{2,2} & y_{1,3} \\
y_{0,2} & y_{1,2} & y_{0,3} & y_{2,2} & y_{1,3} & y_{0,4}
\end{bmatrix}$$

In practice, however, it works well:

GloptiPoly: Global Optimization over Polynomials with Matlab and SeDuMi

Didier Henrion, Jean-Bernard Lasserre
December, 2006

SOSTOOLS by Stephen Prajna, Antonis Papachristodoulou, Peter Seiler, Pablo A. Parrilo
Sparsity in Coefficient Vectors

- Convergent SDP-relaxations in polynomial optimization with sparsity. Lasserre, 2006
- Similar result holds: \( \inf Q_r \uparrow p^* \) as \( r \to \infty \), where
  \[ P: \inf_{x \in \mathbb{R}^n} \{ f(x) | x \in K \} \]
  \[ K := \{ x \in \mathbb{R}^n | g_j(x) \geq 0, j = 1, \ldots, m \} \]
  \( g_j \) and \( f \) depend only on \( \{ x_i | i \in I_k \} \) for some \( k \), and \( |I_k| \leq \kappa \)
- Advantage: the number of variables is \( \mathcal{O}(\kappa^2 r) \), instead of \( \mathcal{O}(n^2 r) \)
- LMI’s of size \( \mathcal{O}(\kappa^r) \) instead of \( \mathcal{O}(n^r) \).
- Significant when \( \kappa < n \)
- Necessary condition on the way the \( I_k \) are related: \( I_{k+1} \cap \bigcup_{j=1}^k I_j \subseteq I_s \)
  for some \( s \leq k \) (running intersection property).
Generalization to non-commuting variables

- Same flavor: construct a sequence of SDP’s that solve the problem of interest

Applications in quantum chemistry and quantum mechanics:

- Computing atomic and molecular ground state energies (solving Hartree Fock equations)
- Computing upper-bounds on the maximal violation of Bell inequalities.

Convergent relaxations of polynomial optimization problems with non-commuting variables
S. Pironio, M. Navascues, A. Acin 2009
Conclusions

- Very General framework.
- Quite a few interesting applications.
- Software available.
- Substantial Computational Challenges!