Moments, Sums of Squares and Semidefinite Programming

Jean B. Lasserre

LAAS-CNRS and Institute of Mathematics, Toulouse, France

IMA Tutorial: Algebraic Algorithms in Optimization, January 2007

- the Generalized Problem of Moments (GPM)
- Some applications
- Duality between moments and nonnegative polynomials
- SDP-relaxations for the basic GPM.

- the Generalized Problem of Moments (GPM)
- Some applications
- Duality between moments and nonnegative polynomials
- SDP-relaxations for the basic GPM.

- the Generalized Problem of Moments (GPM)
- Some applications
- Duality between moments and nonnegative polynomials
- SDP-relaxations for the basic GPM.

- the Generalized Problem of Moments (GPM)
- Some applications
- Duality between moments and nonnegative polynomials
- SDP-relaxations for the basic GPM.

- the Generalized Problem of Moments (GPM)
- Some applications
- Duality between moments and nonnegative polynomials
- SDP-relaxations for the basic GPM.

The generalized problem of moments (GPM)

The GPM is the convex optimization problem:

$$\min_{\mu\,\in\, M(\mathbf{K})} \left\{ \int f_0\, d\mu \quad \Big| \quad \int f_j\, d\mu \overset{=}{\geq} b_j, \quad j=1,\ldots,p \right\}$$

with $\mathbf{K} \subseteq \mathbb{R}^n$ and $M(\mathbf{K})$ a convex set of finite Borel measures on $\mathbf{K}$. We even consider the more general GPM

$$\min_{\mu_i\,\in\, M(\mathbf{K}_i)} \left\{ \sum_{i\in I} \int f_{oi}\, d\mu_i \quad \Big| \quad \sum_{i\in I} \int f_{ji}\, d\mu_i \overset{=}{\geq} b_j, \quad j \in J \right\}$$

where for all $i \in I$, $\mathbf{K}_i \subseteq \mathbb{R}^{n_i}$ and $M(\mathbf{K}_i)$ is a convex set of finite Borel measures on $\mathbf{K}_i$.

The index set I is finite whereas J may be countable.

• The GPM has great modelling power, in various fields. Global Optimization (continuous, discrete), Control (Robust and optimal control), Nonlinear Equations, Probability and Statistics, Performance Evaluation (in e.g. Mathematical finance, Markov chains), Inverse Problems (cristallography, tomography), Numerical multivariate Integration, etc ...

• The GPM is a useful theoretical tool to prove existence and characterization of optimal solutions.

• **BUT** ... in full generality .... GPM is unsolvable numerically.

• The GPM has great modelling power, in various fields. Global Optimization (continuous, discrete), Control (Robust and optimal control), Nonlinear Equations, Probability and Statistics, Performance Evaluation (in e.g. Mathematical finance, Markov chains), Inverse Problems (cristallography, tomography), Numerical multivariate Integration, etc ...

• The GPM is a useful theoretical tool to prove existence and characterization of optimal solutions.

• **BUT** ... in full generality .... GPM is unsolvable numerically.

• The GPM has great modelling power, in various fields. Global Optimization (continuous, discrete), Control (Robust and optimal control), Nonlinear Equations, Probability and Statistics, Performance Evaluation (in e.g. Mathematical finance, Markov chains), Inverse Problems (cristallography, tomography), Numerical multivariate Integration, etc ...

• The GPM is a useful theoretical tool to prove existence and characterization of optimal solutions.

• **BUT** ... in full generality .... GPM is unsolvable numerically.

HOWEVER ... if the $\mathbf{K}_i$, $(\subset \mathbb{R}^{n_i})$ are basic semi-algebraic sets and the f_{ij} are polynomials (or even piecewise polynomials), then ...

by using results of real algebraic geometry, one may now define efficient numerical approximation chemes, based on Semidefinite Programming (SDP).

PROBLEM 1: Probability

Let $\mathbf{K} \subseteq \mathbb{R}^n$, $S \subset \mathbf{K}$ be Borel subsets, and $\Gamma \subset \mathbb{N}^n$.

Finding an upper bound (if possible optimal) on $\mathrm{Prob}\,(\mathbf{X} \in S)$, the probability that a $\mathbf{K}$-valued random variable $\mathbf{X} \in S$, given some of its moments $\gamma = \{\gamma_\alpha\}$, $\alpha \in \Gamma \subset \mathbb{N}^n$ ....

.... is equivalent to solving:

$$\sup_{\mu \in M(\mathbf{K})} \; \{\, \mu(S) \quad | \quad \int X^\alpha \, d\mu \,=\, \gamma_\alpha, \quad \alpha \in \Gamma \}$$

- $M(\mathbf{K})$ is the (convex) set of **probability measures** on $\mathbf{K} \subseteq \mathbb{R}^n$.
- $f_\alpha \equiv X^\alpha$, $\alpha \in \Gamma$ (polynomial); $f_0 = \mathrm{I}_S$ (piecewise polynomial)

PROBLEM 1: Probability

Let $\mathbf{K} \subseteq \mathbb{R}^n$, $S \subset \mathbf{K}$ be Borel subsets, and $\Gamma \subset \mathbb{N}^n$.

Finding an upper bound (if possible optimal) on $\mathrm{Prob}\,(\mathbf{X} \in S)$, the probability that a $\mathbf{K}$-valued random variable $\mathbf{X} \in S$, given some of its moments $\gamma = \{\gamma_\alpha\}$, $\alpha \in \Gamma \subset \mathbb{N}^n$ ....

.... is equivalent to solving:

$$\sup_{\mu \in M(\mathbf{K})} \left\{ \mu(S) \;\middle|\; \int X^\alpha \, d\mu = \gamma_\alpha, \quad \alpha \in \Gamma \right\}$$

- $M(\mathbf{K})$ is the (convex) set of **probability measures** on $\mathbf{K} \subseteq \mathbb{R}^n$.
- $f_\alpha \equiv X^\alpha$, $\alpha \in \Gamma$ (polynomial); $f_0 = \mathrm{I}_S$ (piecewise polynomial)

PROBLEM 1: Probability

Let $\mathbf{K} \subseteq \mathbb{R}^n$, $S \subset \mathbf{K}$ be Borel subsets, and $\Gamma \subset \mathbb{N}^n$.

Finding an upper bound (if possible optimal) on $\mathrm{Prob}\,(\mathbf{X} \in S)$, the probability that a $\mathbf{K}$-valued random variable $\mathbf{X} \in S$, given some of its moments $\gamma = \{\gamma_\alpha\}$, $\alpha \in \Gamma \subset \mathbb{N}^n$ ....

.... is equivalent to solving:

$$\sup_{\mu \in M(\mathbf{K})} \left\{ \mu(S) \quad \Big| \quad \int X^\alpha \, d\mu \, = \, \gamma_\alpha, \quad \alpha \in \Gamma \right\}$$

- $M(\mathbf{K})$ is the (convex) set of **probability measures** on $\mathbf{K} \subseteq \mathbb{R}^n$.
- $f_\alpha \equiv X^\alpha$, $\alpha \in \Gamma$ (polynomial); $f_0 = \mathrm{I}_S$ (piecewise polynomial)

Under no arbitrage, the price of an European Call Option with strike k, is given by $E[(\mathbf{X}-k)^+]$ where E is the expectation operator w.r.t. the distribution of the underlying asset $\mathbf{X}$.

Hence, finding an (**optimal**) upper bound on the price of a European Call Option with strike k, given the first p moments $\{\gamma_j\}$, reduces to solving:

$$\sup_{\mu \in M(\mathbf{K})} \left\{ \int (X-k)^+ \, d\mu \quad \middle| \quad \int X^j \, d\mu = \gamma_j, \quad j=1,\ldots,p \right\}$$

with $\mathbf{K} = \mathbb{R}_+$, and $M(\mathbf{K})$ the set of probability measures on $\mathbf{K}$.
$f_j \equiv X^j$ (polynomials), and $f_0 \equiv (X-k)^+$ (piecewise polynomial)

PROBLEM 3: Global Optimization

With $\mathbf{K} \subseteq \mathbb{R}^n$ and $p, q \in \mathbb{R}[X]$ consider the optimization problem:

$$\rho^* := \inf_x \left\{ \frac{p(x)}{q(x)} \quad \middle| \quad x \in \mathbf{K} \right\}$$

where ρ^* is the GLOBAL minimum.

Finding ρ^* is equivalent to solving

$$\inf_{\mu \in M(\mathbf{K})} \quad \left\{ \int p \, d\mu \ : \ \int q \, d\mu = 1 \right\}$$

with $M(\mathbf{K})$ being the set of finite measures on $\mathbf{K}$.

Important particular case :

Solving Polynomial Equations

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

with $g_j \in \mathbb{R}[X_1, \ldots, X_n]$ for all $j = 1, \ldots, m$.

Finding a solution $x^* \in \mathbf{K}$ that minimizes $f(X)$ on $\mathbf{K}$ is equivalent to solving

$$\inf_{\mu \in M(\mathbf{K})} \int f \, d\mu$$

with $M(\mathbf{K})$ being the set of probability measures on $\mathbf{K}$.

PROBLEM 4: Convex envelope

Let $\mathbf{K} \subseteq \mathbb{R}^n$, $f : \mathbf{K} \to \mathbb{R}$ ($= +\infty$ outside $\mathbf{K}$) and with x fixed, consider the optimization problem

$$\widehat{f}(x) \;:=\; \inf_{\mu \in M(\mathbf{K})} \left\{ \int f\, d\mu \;\middle|\; \int X_j\, d\mu \;=\; x_j, \quad j = 1, \ldots, n \right\}$$

with $M(\mathbf{K})$ being the set of probability measures on $\mathbf{K}$.

- $\widehat{f}$ is convex and is the *convex envelope* of f, defined on the convex hull $\mathrm{co}\,(\mathbf{K})$ of $\mathbf{K}$.

Let $\mathbf{K}_j \subset \mathbb{R}^{n_j}$, $j = 1, \ldots, p$, and $\mathbf{K} := \mathbf{K}_1 \times \mathbf{K}_2 \cdots \times \mathbf{K}_p \subset \mathbb{R}^n$, and with natural projections $\pi_j : \mathbf{K} :\to \mathbf{K}_j$, $j = 1, \ldots, p$.

Let ν_j be a given Borel measure on $\mathbf{K}_j$, $j = 1, \ldots, p$,

For a measure μ on $\mathbf{K}$, denote $\pi_j \, \mu$ its marginal on $\mathbf{K}_j$, i.e.

$$\pi_j \, \mu(B) \; := \; \mu\,(\pi_j^{-1}(B)) \; = \; \mu\,(\{x \in \mathbf{K} \; : \; \pi_j \, x \in B\}), \quad B \in \mathcal{B}(\mathbf{K}_j)$$

Consider the optimization problem:

$$\inf_{\mu \in M(\mathbf{K})} \left\{ \int f \, d\mu \quad | \quad \pi_j \, \mu \; = \; \nu_j, \quad j = 1, \ldots, p \right\}$$

with $M(\mathbf{K})$ being the set of finite Borel measures on $\mathbf{K}$.

Generalization of the famous Monge-Kantorovich transportation problem, with many other interesting applications, particularly in Probability.

•• If $\mathbf{K}_j$ is compact then the constraint on marginal

$$\pi_j \, \mu \,=\, \nu_j$$

is equivalent to the countably many linear equalities

$$\int X^\alpha \, d\mu \quad = \quad \int X^\alpha \, d\nu_j, \quad \forall \alpha \in \mathbb{N}^{n_j}$$

between moments of μ and ν_j ...
because the space of polynomials is dense (for the sup-norm) in the space $C(\mathbf{K}_j)$ of continuous functions on $\mathbf{K}_j$.

PROBLEM 6: Deterministic Optimal Control

$$
\begin{aligned}
j^* := \quad \min_{\mathbf{u}} \quad & \int_0^T h(s, x(s), u(s))\, ds + H(x(T)) \\
\dot{x}(s) \quad = \quad & f(s,\, x(s), u(s)), \quad s \in [0, T] \\
(x(s), u(s)) \quad \in \quad & X \times U, \quad s \in [0, T) \\
x(T) \quad \in \quad & X_T,
\end{aligned}
\tag{1}
$$

and with initial condition $x(0) = x_0 \in X$, and
- $X, X_T \subset \mathbb{R}^n$ and $U \subset \mathbb{R}^m$ are basic semi-algebraic sets.
- $h, f \in \mathbb{R}[t, x, u]$, $H \in \mathbb{R}[x]$

The primal SDP-relaxation corresponds to a weak formulation of the original problem in terms of occupation measures .....

whereas the dual SDP-relaxation corresponds to a weak formulation of the original problem in terms of subsolutions of the Hamilton-Jacobi-Bellman optimality conditions.

More in Didier Henrion's talk .... next week!

Duality

With $M(\mathbf{K})$ the space of Borel prob. measures on $\mathbf{K}$, the GPM

$$\min_{\mu \in M(\mathbf{K})} \left\{ \int f_0 \, d\mu \quad \middle| \quad \int f_j \, d\mu \, = \, b_j, \quad j = 1, \ldots, p \right\}$$

is the **infinite-dimensional LP**

$$\min_{\mu \in \mathcal{M}} \{ \langle f_0, \mu \rangle \mid \quad \langle f_j, \mu \rangle = b_j, \, j = 1, \ldots, p; \quad \langle 1, \mu \rangle = 1; \; \mu \geq 0 \}$$

where $\mathcal{M}$ is the **vector space** of **finite** *signed* **Borel measures** on $\mathbf{K}$. The **dual** LP reads:

$$\max_{\lambda \in \mathbb{R}^p, \gamma \in \mathbb{R}} \left\{ \quad \gamma \quad \middle| \quad f_0 - \sum_{j=1}^{p} \lambda_j \, (f_j - b_j) \, \geq \, \gamma \quad \text{on } \mathbf{K} \right\}$$

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

• to handle $\int f_j \, d\mu$, and have

relatively **simple** and **tractable** **characterizations** of :

• **measures** μ with support contained in **K**, ... or

• **functions** (e.g. $f_0 - \sum_{j=1}^{p} \lambda_j (f_j - b_j)$) nonnegative on **K**.

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

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

• to handle $\int f_j \, d\mu$, and have

relatively simple and tractable characterizations of :

• **measures** μ with support contained in **K**, ... or

• **functions** (e.g. $f_0 - \sum_{j=1}^{p} \lambda_j (f_j - b_j)$) nonnegative on **K**.

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

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

• to handle $\int f_j \, d\mu$, and have

relatively simple and tractable characterizations of :

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

• **functions** (e.g. $f_0 - \sum_{j=1}^{p} \lambda_j \, (f_j - b_j)$) nonnegative on $\mathbf{K}$.

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

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

• to handle $\int f_j \, d\mu$, and have

relatively simple and tractable characterizations of :

• **measures** μ with support contained in **K**, ... or

• **functions** (e.g. $f_0 - \sum_{j=1}^{p} \lambda_j \, (f_j - b_j)$) nonnegative on **K**.

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

A first good news ...

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

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

with $\{g_j\} \subset \mathbb{R}[X_1, \ldots, X_n]$ ...

Powerful results of real algebraic geometry and on the moment problem, provide conditions for :

• a finite Borel **measure** μ to be supported on $\mathbf{K}$ (i.e., $\mu(\mathbf{K}^c) = 0$)

• a **polynomial** f to be > 0 on $\mathbf{K}$.

As one may expect, the conditions are *dual* to each other ....

A first good news ...

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

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

with $\{g_j\} \subset \mathbb{R}[X_1, \ldots, X_n]$ ...

Powerful results of real algebraic geometry and on the moment problem, provide conditions for :

• a finite Borel **measure** μ to be supported on $\mathbf{K}$ (i.e., $\mu(\mathbf{K}^c) = 0$)

• a **polynomial** f to be > 0 on $\mathbf{K}$.

As one may expect, the conditions are *dual* to each other ....

A first good news ...

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

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

with $\{g_j\} \subset \mathbb{R}[X_1, \ldots, X_n]$ ...

Powerful results of real algebraic geometry and on the moment problem, provide conditions for :

• a finite Borel **measure** μ to be supported on $\mathbf{K}$ (i.e., $\mu(\mathbf{K}^c) = 0$)

• a **polynomial** f to be > 0 on $\mathbf{K}$.

As one may expect, the conditions are *dual* to each other ....

A second good news ...(continued)

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

• the **moments** $y_\alpha := \int X^\alpha \, d\mu$, $\alpha \in \mathbb{N}^n$, of μ with **support** in **K**

• the **coefficients** $\{q_{j\alpha}\}$ of sum of squares (s.o.s.) **polynomials** $\{q_j\}_{j=0}^m \subset \mathbb{R}[x]$, in e.g. Putinar's s.o.s. representation

$$f = q_0 + \sum_{j=1}^m q_j \, g_j, \qquad \text{if } f > 0 \text{ on } \mathbf{K}.$$

• Linear Inequalities instead of LMIs are also available .. but less efficient and ill-behaved ... despite so far, LP software packages are more powerful than SDP packages!!

A second good news ...(continued)

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

• the **moments** $y_\alpha := \int X^\alpha \, d\mu$, $\alpha \in \mathbb{N}^n$, of μ with **support** in **K**

• the **coefficients** $\{q_{j\alpha}\}$ of sum of squares (s.o.s.) **polynomials** $\{q_j\}_{j=0}^m \subset \mathbb{R}[x]$, in e.g. Putinar's s.o.s. representation

$$f = q_0 + \sum_{j=1}^m q_j \, g_j, \qquad \text{if } f > 0 \text{ on } \mathbf{K}.$$

• Linear Inequalities instead of LMIs are also available .. but less efficient and ill-behaved ... despite so far, LP software packages are more powerful than SDP packages!!

A second good news ...(**continued**)

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

• the **moments** $y_\alpha := \int X^\alpha \, d\mu$, $\alpha \in \mathbb{N}^n$, of μ with **support** in **K**

• the **coefficients** $\{q_{j\alpha}\}$ of sum of squares (s.o.s.) **polynomials** $\{q_j\}_{j=0}^m \subset \mathbb{R}[x]$, in e.g. Putinar's s.o.s. representation

$$f = q_0 + \sum_{j=1}^m q_j \, g_j, \qquad \text{if } f > 0 \text{ on } \mathbf{K}.$$

• Linear Inequalities instead of LMIs are also available .. but less efficient and ill-behaved ... despite so far, LP software packages are more powerful than SDP packages!!

A second good news ...(continued)

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

- the **moments** $y_\alpha := \int X^\alpha \, d\mu$, $\alpha \in \mathbb{N}^n$, of μ with **support** in **K**
- the **coefficients** $\{q_{j\alpha}\}$ of sum of squares (s.o.s.) **polynomials** $\{q_j\}_{j=0}^m \subset \mathbb{R}[x]$, in e.g. Putinar's s.o.s. representation

$$f \;=\; q_0 + \sum_{j=1}^m q_j \, g_j, \qquad \text{if } f > 0 \text{ on } \mathbf{K}.$$

- Linear Inequalities instead of LMIs are also available .. but less efficient and ill-behaved ... despite so far, LP software packages are more powerful than SDP packages!!

Putinar-Jacobi-Prestel's Positivstellensatz

$$Q(g) := \{f \in \mathbb{R}[X] : \quad f = f_0 + \sum_{j=1}^m f_j\, g_j, \quad f_j \text{ s.o.s.}\},$$

is the quadratic module generated by the g_j's.

• $f \in Q(g)$ is an obvious **certificate of nonnegativity** of f on $\mathbf{K}$.

Assumption 1: There exists some $u \in Q(g)$ such that the level set $\{x \in \mathbb{R}^n \mid u(x) \geq 0\}$ is compact.

Theorem (Putinar): Let $\mathbf{K}$ compact and Assumption 1 hold.
Then $[\,\mathbf{f} \in \mathbb{R}[\mathbf{X}]$ and $\mathbf{f} > \mathbf{0}$ on $\mathbf{K}\,] \Rightarrow \mathbf{f} \in \mathbf{Q}(\mathbf{g}_1, \ldots, \mathbf{g}_m)$.

Putinar-Jacobi-Prestel's Positivstellensatz

$$Q(g) := \{f \in \mathbb{R}[X] : \quad f = f_0 + \sum_{j=1}^{m} f_j\, g_j, \quad f_j \text{ s.o.s.}\},$$

is the quadratic module generated by the g_j's.

• $f \in Q(g)$ is an obvious **certificate of nonnegativity** of f on **K**.

Assumption 1: There exists some $u \in Q(g)$ such that the level set $\{x \in \mathbb{R}^n \mid u(x) \geq 0\}$ is compact.

Theorem (Putinar): Let **K** compact and Assumption 1 hold. Then $[\,\mathbf{f} \in \mathbb{R}[\mathbf{X}]$ and $\mathbf{f} > \mathbf{0}$ on $\mathbf{K}\,] \Rightarrow \mathbf{f} \in \mathbf{Q}(\mathbf{g_1}, \ldots, \mathbf{g_m})$.

• **With an apriori bound on the degree of the s.o.s. polynomials $\{f_j\}$, checking $f \in Q(g)$ is solving a 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 $\{ x \mid g_j(x) \geq 0\}$ is compact for some $j \in \{1, \ldots, m\}$.

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

• **With an apriori bound on the degree of the s.o.s. polynomials $\{f_j\}$, checking $f \in Q(g)$ is 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 $j \in \{1, \ldots, m\}$.

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

• **With an apriori bound on the degree of the s.o.s. polynomials $\{f_j\}$, checking $f \in Q(g)$ is solving a 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 $\{\, x \mid g_j(x) \geq 0\}$ is compact for some $j \in \{1, \ldots, m\}$.

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

• With an apriori bound on the degree of the s.o.s. polynomials $\{f_j\}$, checking $f \in Q(g)$ is solving a 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 $\{\, x \mid g_j(x) \geq 0\}$ is compact for some $j \in \{1, \ldots, m\}$.

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

Putinar versus Karush-Kuhn-Tucker

Let $f^* := \min_x \{ f(x) \; : \; g_j(x) \geq 0, \; j = 1, \ldots, m\}$. Then:

$$\text{KKT}: \quad \nabla \left[f(x^*) - \sum_{j=1}^m \lambda_j \, g_j(x^*) \right] = 0. \quad \lambda_j \, g_j(x^*) = 0; \quad \lambda_j \geq 0$$

Convex case: $\Rightarrow x^*$ is a global minimum of $L = f - \sum_j \lambda_j \, g_j$ and

$$f - f^* - \sum_{j=1}^m \lambda_j \, g_j \geq 0 \quad \text{on } \mathbb{R}^n$$

In general : x^* stationary point of L or **local** minimizer only!!

Putinar: $f - f^* \geq 0$ on **K**. $\qquad f - f^* - \sum_{j=1}^m h_j \, g_j$ is s.o.s.!

(hence ≥ 0 on $\mathbb{R}^n$)

When **Putinar's representation** holds for the polynomial $f - f^*$ (≥ 0 on **K**)... it provides a **global optimality certificate** for f^*

... the analogue in **nonconvex polynomial optimization** of the **KKT-optimality conditions** for the general convex case ....

a **highly nontrivial extension ..!!!**

When **Putinar's representation** holds for the polynomial $f - f^*$ (≥ 0 on $\mathbf{K}$)... it provides a **global optimality certificate** for f^*

... the analogue in **nonconvex polynomial optimization** of the **KKT-optimality conditions** for the general convex case ....

a **highly nontrivial extension ..!!!**

When **Putinar's representation** holds for the polynomial $f - f^*$ (≥ 0 on **K**)... it provides a **global optimality certificate** for f^*

... the analogue in **nonconvex polynomial optimization** of the **KKT-optimality conditions** for the general convex case ....

a **highly nontrivial extension ..!!!**

Example: $1 = \min_x \{ \quad x \mid \quad x^2 = 1; \quad x \geq 1/2\}$.
KKT: One has the non convex Lagrangian

$$L = X - 1 - \lambda_1(X - 1/2) - \lambda_2\,(X^2 - 1)$$

with $\lambda_1 = 0$ and $\lambda_2 = 1/2$
Putinar: $X - 1 \geq 0$ on $\mathbf{K} = \{\mathbf{x} \mid \mathbf{x^2} = \mathbf{1};\ \mathbf{x} \geq \mathbf{1/2}\} = \{\mathbf{1}\}$ and ...

$$X - 1 - \frac{1}{3}(X - 1)^2\,(X - \frac{1}{2}) + \frac{1}{3}\,(X - \frac{5}{2})\,(X^2 - 1) = 0$$

Note that $-(X^* - 5/2)/3 = \frac{1}{2} = \lambda_2$ and $(X^* - 1)^2/3 = 0 = \lambda_1$.

The constraint $X - 1/2 \geq 0$ is **not active** at $x^* = 1$, but its MULTIPLIER $(X - 1)^2/3 \neq 0$. Its value at $x^* = 1$ is $0 = \lambda_1$!

SDP-relaxations for solving the GPM

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

Let $L_y : \mathbb{R}[X] \to \mathbb{R}$, be the linear functional

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

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

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

Given $y = \{y_\alpha\}$ indexed in the basis $\{X^\alpha\}$, introduce the **moment matrix** $M_r(y)$ with rows and columns also indexed in that basis, and defined as follows:

$$M_r(y)(\alpha,\beta) := L_y(X^{\alpha+\beta}) = y_{\alpha+\beta}, \quad \alpha,\beta \in \mathbb{N}^n, \quad |\alpha|,|\beta| \leq r.$$

For instance, and for illustration purposes, in $\mathbb{R}^2$,

$$M_1(y) = \left[\begin{array}{c|cc} y_{00} & y_{10} & y_{01} \\ - & - & - \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{array}\right]$$

Then

$$\left[L_y(f^2) \geq 0, \quad \forall f,\ \deg(f) \leq r \right] \quad \Leftrightarrow \quad M_r(y) \succeq 0$$

Similarly, given $\theta \in \mathbb{R}[X]$, $X \mapsto \theta(X) = \sum_\gamma \theta_\gamma X^\gamma$, one defines the localizing matrix $M_r(\theta\, y)$, with respect to y, θ, also indexed in the basis $\{X^\alpha\}$, by

$$M_r(\theta\, y)(\alpha,\beta) \,=\, L_y(\theta\, X^{\alpha+\beta}) \,=\, \sum_{\gamma\in\mathbb{N}^n} \theta_\gamma\, y_{\alpha+\beta+\gamma}, \quad \left\{ \begin{array}{l} \alpha,\beta \in \mathbb{N}^n \\ |\alpha|, |\beta| \leq r. \end{array} \right.$$

For instance, in $\mathbb{R}^2$, and with $X \mapsto \theta(X) := 1 - X_1^2 - X_2^2$,

$$M_1(\theta\, y) = \left[\begin{array}{ccc} y_{00} - y_{20} - y_{02}, & y_{10} - y_{30} - y_{12}, & y_{01} - y_{21} - y_{03} \\ y_{10} - y_{30} - y_{12}, & y_{20} - y_{40} - y_{22}, & y_{11} - y_{21} - y_{12} \\ y_{01} - y_{21} - y_{03}, & y_{11} - y_{21} - y_{12}, & y_{02} - y_{22} - y_{04} \end{array} \right].$$

Then

$$\left[L_y(f^2\, \theta) \geq 0, \quad \forall f, \deg(f) \leq r \right] \quad \Leftrightarrow \quad M_r(\theta\, y) \succeq 0$$

SDP-relaxations

$$\min_{\mu \in M(\mathbf{K})} \left\{ \int f_0 \, d\mu \quad \middle| \quad \int f_j \, d\mu \; = \; b_j, \quad j = 1, \ldots, p \right\}$$

($M(\mathbf{K})$ space of Borel prob. measures on $\mathbf{K}$, and $\{f_j\} \subset \mathbb{R}[X]$)

Let $\deg g_i = 2v_i$ or $2v_i - 1$. The SDP-relaxation $\mathbf{Q}_r$ reads:

$$\mathbf{Q}_r \begin{cases} \min\limits_{y} & L_y(f) \\ \text{s.t.} & M_r(y) \succeq 0 \\ & M_{r-v_i}(g_i\, y) \succeq 0 \quad i = 1, \ldots m. \\ & L_y(1) = 1 \\ & L_y(f_j - b_j) = 0 \quad j = 1, \ldots p. \end{cases}$$

... whose dual is the SDP

$$\mathbf{Q}_r^* \left\{ \begin{array}{ll} \displaystyle\max_{\lambda,\gamma,\{q_j\}} & \gamma \\ \text{s.t.} & \displaystyle f_0 - \sum_{j=1}^{p} \lambda_j \,(f_j - b_j) - \gamma \;=\; q_0 + \sum_{j=1}^{m} q_j \, g_j \\ & \{q_j\} \text{ are s.o.s.}; \quad \deg q_0,\ \deg q_j g_j \;\leq 2r \end{array} \right.$$

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

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

Assumption 1: There exists some $u \in Q(g_1, \ldots, g_m)$ such that the level set $\{x \in \mathbb{R}^n \,|\, u(x) \geq 0\}$ is compact.

Theorem: Let $\mathbf{K}$ be compact, and let Assumption 1 hold, and consider the basic **GPM** with optimal value ρ^*. Then :

- $\sup \mathbf{Q}_r^* \leq \inf \mathbf{Q}_r$ and $\inf \mathbf{Q}_r \uparrow \rho^*$ as $r \to \infty$
- If $\mathrm{int}\,\mathbf{K} \neq \emptyset$ and the GPM has a feasible solution with a density $\sup \mathbf{Q}_r^* = \max \mathbf{Q}_r^* = \inf \mathbf{Q}_r \uparrow \rho^*$.

Detecting global optimality and extracting solutions

• When $\mathbf{K}$ is compact, then the basic GPM has an optimal solution μ^*, with optimal value ρ^*.

• By Caratheodory theorem there exists an at most $(p+2)$-atomic probability measure φ on $\mathbf{K}$ such that

$$\int f_j\,d\varphi = \int f_j\,d\mu, \quad j = 1,\ldots,p; \qquad \int f_0\,d\varphi = \rho^*$$

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

$$\mathrm{rank}\, M_r(y) = \mathrm{rank}\, M_{r-v}(g_j\, y) = s$$

$\min \mathbf{Q}_r = \rho^*$ and one may extract a s-atomic optimal solution φ.

GloptiPoly 3 is a software package initially devoted to solving global optimization problems with polynomials.

http://www.laas.fr/∼henrion/software

with detection of optimality and extraction of solutions.