

Representations of Positive Polynomials and SDP-relaxations exploiting sparsity

Jean B. Lasserre¹

¹LAAS-CNRS and Institute of Mathematics, Toulouse, France

Banff workshop: **Positive Polynomials**, October 2006

Outline

- Problem statement
- Specialized **Positivstellensatz**: part I: (X, Y, Z)
- Specialized **Positivstellensatz**: part II: **general case**
- Sparse SDP-relaxations for global polynomial optimization
- The \mathbf{K} -moment problem with density

Outline

- Problem statement
- Specialized **Positivstellensatz**: part I: (X, Y, Z)
- Specialized **Positivstellensatz**: part II: **general case**
- Sparse SDP-relaxations for global polynomial optimization
- The \mathbf{K} -moment problem with density

Outline

- Problem statement
- Specialized **Positivstellensatz**: part I: (X, Y, Z)
- Specialized **Positivstellensatz**: part II: general case
- Sparse SDP-relaxations for global polynomial optimization
- The \mathbf{K} -moment problem with density

Outline

- Problem statement
- Specialized **Positivstellensatz**: part I: (X, Y, Z)
- Specialized **Positivstellensatz**: part II: **general case**
- Sparse SDP-relaxations for global polynomial optimization
- The K -moment problem with density

Outline

- Problem statement
- Specialized **Positivstellensatz**: part I: (X, Y, Z)
- Specialized **Positivstellensatz**: part II: **general case**
- Sparse SDP-relaxations for global polynomial optimization
- The K -moment problem with density

Outline

- Problem statement
- Specialized **Positivstellensatz**: part I: (X, Y, Z)
- Specialized **Positivstellensatz**: part II: **general case**
- Sparse SDP-relaxations for global polynomial optimization
- The \mathbf{K} -moment problem with density

Problem statement

Let $X = (X_1, \dots, X_n)$, $Y = (Y_1, \dots, Y_m)$ and $Z = (Z_1, \dots, Z_p)$.

$$\mathbf{K} := \{(x, y, z) \in \mathbb{R}^{n+m+p} : (x, y) \in \mathbf{K}_{xy}; \quad (y, z) \in \mathbf{K}_{yz}\}$$

$$\mathbf{K}_{xy} := \{(x, y) \in \mathbb{R}^{n+m} : g_j(x, y) \geq 0, \quad j \in J_{xy}\}$$

$$\mathbf{K}_{yz} := \{(y, z) \in \mathbb{R}^{m+p} : h_j(y, z) \geq 0, \quad j \in J_{yz}\}$$

and let $f \in \mathbb{R}[X, Y] + \mathbb{R}[Y, Z]$.

- So in the definition of \mathbf{K} and f there is **NO COUPLING** of variables X and Z .
- Can we have a **specialized representation** of $f > 0$ on \mathbf{K} that preserves this coupling pattern?
- If yes, can we extend to **more than two** subsets of variables (X, Y) and (Y, Z) ?

Problem statement

Let $X = (X_1, \dots, X_n)$, $Y = (Y_1, \dots, Y_m)$ and $Z = (Z_1, \dots, Z_p)$.

$$\mathbf{K} := \{(x, y, z) \in \mathbb{R}^{n+m+p} : (x, y) \in \mathbf{K}_{xy}; \quad (y, z) \in \mathbf{K}_{yz}\}$$

$$\mathbf{K}_{xy} := \{(x, y) \in \mathbb{R}^{n+m} : g_j(x, y) \geq 0, \quad j \in J_{xy}\}$$

$$\mathbf{K}_{yz} := \{(y, z) \in \mathbb{R}^{m+p} : h_j(y, z) \geq 0, \quad j \in J_{yz}\}$$

and let $f \in \mathbb{R}[X, Y] + \mathbb{R}[Y, Z]$.

- So in the definition of \mathbf{K} and f there is **NO COUPLING** of variables X and Z .
- Can we have a **specialized representation** of $f > 0$ on \mathbf{K} that preserves this coupling pattern?
- If yes, can we extend to **more than two** subsets of variables (X, Y) and (Y, Z) ?

Problem statement

Let $X = (X_1, \dots, X_n)$, $Y = (Y_1, \dots, Y_m)$ and $Z = (Z_1, \dots, Z_p)$.

$$\mathbf{K} := \{(x, y, z) \in \mathbb{R}^{n+m+p} : (x, y) \in \mathbf{K}_{xy}; \quad (y, z) \in \mathbf{K}_{yz}\}$$

$$\mathbf{K}_{xy} := \{(x, y) \in \mathbb{R}^{n+m} : g_j(x, y) \geq 0, \quad j \in J_{xy}\}$$

$$\mathbf{K}_{yz} := \{(y, z) \in \mathbb{R}^{m+p} : h_j(y, z) \geq 0, \quad j \in J_{yz}\}$$

and let $f \in \mathbb{R}[X, Y] + \mathbb{R}[Y, Z]$.

- So in the definition of \mathbf{K} and f there is **NO COUPLING** of variables X and Z .
- Can we have a **specialized representation** of $f > 0$ on \mathbf{K} that preserves this coupling pattern?
- If yes, can we extend to **more than two** subsets of variables (X, Y) and (Y, Z) ?

A specialized representation : part I

Let $P(g) \subset \mathbb{R}[X, Y]$ and $Q(g) \subset \mathbb{R}[X, Y]$ be the **preordering** and **quadratic module** generated by $(g_j)_{j \in J_{xy}}$.

That is:

$$P(g) := \left\{ \sum_{J \subset J_{xy}} \sigma_J \prod_{j \in J} g_j : \sigma_J \in \Sigma_{xy}^2 \right\}$$

$$Q(g) := \left\{ \sigma_0 + \sum_{j \in J_{xy}} \sigma_j g_j : \sigma_0, \sigma_j \in \Sigma_{xy}^2 \right\}$$

Similar definitions for $P(h), Q(h) \subset \mathbb{R}[Y, Z]$.

Theorem : Las (2006)

Let $f \in \mathbb{R}[X, Y] + \mathbb{R}[Y, Z]$, and let \mathbf{K} be compact.

(a) If f is positive on \mathbf{K} then $f \in P(g) + P(h)$.

(b) If $N - \|(X, Y)\|^2 \in Q(g)$ and/or $N - \|(Y, Z)\|^2 \in Q(h)$ for some scalar N , and if f is positive on \mathbf{K} , THEN

in (a) one may replace $P(g)$ with $Q(g)$ and/or $P(h)$ with $Q(h)$.

(a) is a specialized version of Schmüdgen's Positivstellensatz with no coupling between X and Z in the representation of f .

(b) is a specialized version of Putinar, and Jacobi & Prestel's Positivstellensatz.

Theorem : Las (2006)

Let $f \in \mathbb{R}[X, Y] + \mathbb{R}[Y, Z]$, and let \mathbf{K} be compact.

(a) If f is positive on \mathbf{K} then $f \in P(g) + P(h)$.

(b) If $N - \|(X, Y)\|^2 \in Q(g)$ and/or $N - \|(Y, Z)\|^2 \in Q(h)$ for some scalar N , and if f is positive on \mathbf{K} , THEN

in (a) one may replace $P(g)$ with $Q(g)$ and/or $P(h)$ with $Q(h)$.

(a) is a specialized version of Schmüdgen's Positivstellensatz with no coupling between X and Z in the representation of f .

(b) is a specialized version of Putinar, and Jacobi & Prestel's Positivstellensatz.

Theorem : Las (2006)

Let $f \in \mathbb{R}[X, Y] + \mathbb{R}[Y, Z]$, and let \mathbf{K} be compact.

(a) If f is positive on \mathbf{K} then $f \in P(g) + P(h)$.

(b) If $N - \|(X, Y)\|^2 \in Q(g)$ and/or $N - \|(Y, Z)\|^2 \in Q(h)$ for some scalar N , and if f is positive on \mathbf{K} , THEN

in (a) one may replace $P(g)$ with $Q(g)$ and/or $P(h)$ with $Q(h)$.

(a) is a specialized version of Schmüdgen's Positivstellensatz with no coupling between X and Z in the representation of f .

(b) is a specialized version of Putinar, and Jacobi & Prestel's Positivstellensatz.

Theorem : Las (2006)

Let $f \in \mathbb{R}[X, Y] + \mathbb{R}[Y, Z]$, and let \mathbf{K} be compact.

(a) If f is positive on \mathbf{K} then $f \in P(g) + P(h)$.

(b) If $N - \|(X, Y)\|^2 \in Q(g)$ and/or $N - \|(Y, Z)\|^2 \in Q(h)$ for some scalar N , and if f is positive on \mathbf{K} , THEN

in (a) one may replace $P(g)$ with $Q(g)$ and/or $P(h)$ with $Q(h)$.

(a) is a specialized version of Schmüdgen's Positivstellensatz with no coupling between X and Z in the representation of f .

(b) is a specialized version of Putinar, and Jacobi & Prestel's Positivstellensatz.

(c) Assume that the g_j 's and h_k 's are **normalized**, i.e.,

$$0 \leq g_j \leq 1 \text{ on } \mathbf{K}_{xy}; \quad 0 \leq h_k \leq 1 \text{ on } \mathbf{K}_{yz} \quad \forall j \in J_{xy}, k \in J_{yz}$$

and that the family $\{0, 1, (g_j)_{j \in J_{xy}}\}$ (resp. $\{0, 1, (h_k)_{k \in J_{yz}}\}$) generates the algebra $\mathbb{R}[X, Y]$ (resp. $\mathbb{R}[Y, Z]$).

Then $f > 0$ on $\mathbf{K} \Rightarrow f \in C(g, 1 - g) + C(h, 1 - h)$.

where $C(g, 1 - g)$ is the cone generated by $(g_j, (1 - g_j))_{j \in J_{xy}}$

$$C(g, 1 - g) := \left\{ \sum_{\alpha, \beta \in \mathbb{N}^{|J_{xy}|}} c_{\alpha\beta} g^\alpha (1 - g)^\beta : c_{\alpha\beta} \geq 0 \right\}$$

(c) is a **specialized Positivstellensatz** inspired from Krivine's representation.

(c) Assume that the g_j 's and h_k 's are **normalized**, i.e.,

$$0 \leq g_j \leq 1 \text{ on } \mathbf{K}_{xy}; \quad 0 \leq h_k \leq 1 \text{ on } \mathbf{K}_{yz} \quad \forall j \in J_{xy}, k \in J_{yz}$$

and that the family $\{0, 1, (g_j)_{j \in J_{xy}}\}$ (resp. $\{0, 1, (h_k)_{k \in J_{yz}}\}$) generates the algebra $\mathbb{R}[X, Y]$ (resp. $\mathbb{R}[Y, Z]$).

Then $f > 0$ on $\mathbf{K} \Rightarrow f \in C(g, 1 - g) + C(h, 1 - h)$.

where $C(g, 1 - g)$ is the cone generated by $(g_j, (1 - g_j))_{j \in J_{xy}}$

$$C(g, 1 - g) := \left\{ \sum_{\alpha, \beta \in \mathbb{N}^{|J_{xy}|}} c_{\alpha\beta} g^\alpha (1 - g)^\beta : c_{\alpha\beta} \geq 0 \right\}$$

(c) is a **specialized Positivstellensatz** inspired from **Krivine's** representation.

Fact:

Let ψ be a probability measure on $X \times Y$, a *cartesian product* of Borel spaces. Then one may **disintegrate** ψ into

- a **stochastic kernel** $\psi_{X|Y}(dx | y)$ on X given Y , and
- a probability measure $\psi_Y(dy)$ on Y , and so

$$\psi(A \times B) = \int_B \psi_{X \times Y}(A | y) \psi_Y(dy), \quad A \in \mathcal{B}(X), B \in \mathcal{B}(Y)$$

ψ_Y is the **marginal** of ψ on Y whereas $\psi_{X \times Y}$ is the **conditional probability**.

$\psi_{X|Y}(dx | y)$ is a stochastic kernel on X given Y if

- for every fixed $y \in Y$, $\psi_{X|Y}(dx | y)$ is a p.m. on X
- for every fixed $A \in \mathcal{B}(X)$, the function $y \mapsto \psi_{X|Y}(A | y)$ is a bounded Borel measurable on Y .

Then the proof relies on the following crucial property:

Let $\nu(d(x, y))$ and $\varphi(d(y, z))$ be probability measures with same marginal on y , i.e.,

$$\nu_y(dy) \equiv \varphi_y(dy) =: \psi(dy),$$

then one may construct $\mu(d(x, y, z))$ such that

$$\mu_{xy}(d(x, y)) \equiv \nu(d(x, y)); \quad \mu_{yz}(d(y, z)) \equiv \varphi(d(y, z)).$$

In fact μ is defined via:

$$\mu(A \times B \times C) = \int_B \nu_{x|y}(A|y) \varphi_{z|y}(C|y) \psi(dy)$$

Then the proof relies on the following crucial property:

Let $\nu(d(x, y))$ and $\varphi(d(y, z))$ be probability measures with same marginal on y , i.e.,

$$\nu_y(dy) \equiv \varphi_y(dy) =: \psi(dy),$$

then one may construct $\mu(d(x, y, z))$ such that

$$\mu_{xy}(d(x, y)) \equiv \nu(d(x, y)); \quad \mu_{yz}(d(y, z)) \equiv \varphi(d(y, z)).$$

In fact μ is defined via:

$$\mu(A \times B \times C) = \int_B \nu_{x|y}(A|y) \varphi_{z|y}(C|y) \psi(dy)$$

Observe that apart from the usual compactness assumption on K and on its representation, **no other assumption** on the **coupling pattern** of variables is needed. That is:

If there is **no coupling** between X and Z , then a **specialized Positivstellensatz** with no coupling between X and Z **also exists!**

.... Yields **specialized SDP-relaxations** for polynomial optimization with significant **computational savings** ... More later

Observe that apart from the usual compactness assumption on K and on its representation, **no other assumption** on the **coupling pattern** of variables is needed. That is:

If there is **no coupling** between X and Z , then a **specialized Positivstellensatz** with no coupling between X and Z **also exists!**

.... Yields **specialized SDP-relaxations** for polynomial optimization with significant **computational savings** ... More later

part II: general case

Assume that $\mathbf{K} \subset \mathbb{R}^n$ is compact and one knows some scalar M such that $M - \|X\|^2 > 0$ on \mathbf{K} .

Let $I = \{1, \dots, n\} = \bigcup_{i=1}^p I_p$ Let $\mathbb{R}[X(I_k)] = \mathbb{R}[X_i \mid i \in I_k]$.

Sparsity pattern:

- $f \in \mathbb{R}[X(I_1)] + \dots + \mathbb{R}[X(I_p)]$.
- $\forall j = 1, \dots, m, \quad g_j \in \mathbb{R}[X(I_{k(j)})]$ for some $k(j) \in \{1, \dots, p\}$.

In other words ... a monomial X^α appears in f or g_j if and only if $\text{support}(\alpha) := \{i \mid \alpha_i \neq 0\} \subseteq I_k$ for some k .

part II: general case

Assume that $\mathbf{K} \subset \mathbb{R}^n$ is compact and one knows some scalar M such that $M - \|X\|^2 > 0$ on \mathbf{K} .

Let $I = \{1, \dots, n\} = \bigcup_{i=1}^p I_p$ Let $\mathbb{R}[X(I_k)] = \mathbb{R}[X_i \mid i \in I_k]$.

Sparsity pattern:

- $f \in \mathbb{R}[X(I_1)] + \dots + \mathbb{R}[X(I_p)]$.
- $\forall j = 1, \dots, m, \quad g_j \in \mathbb{R}[X(I_{k(j)})]$ for some $k(j) \in \{1, \dots, p\}$.

In other words ... a monomial X^α appears in f or g_j if and only if $\text{support}(\alpha) := \{i \mid \alpha_i \neq 0\} \subseteq I_k$ for some k .

A sparse Positivstellensatz

- Add the p redundant quadratic constraints

$$g_{m+k}(X) = M - \|X(l_k)\|^2 \geq 0, \quad k = 1, \dots, p$$

in the definition of the constraint set \mathbf{K} . ($g_{m+k} \in \mathbb{R}[X(l_k)]$.)

Let $\mathcal{G}_k := \{g_j : g_j \in \mathbb{R}[X(l_k)]\}$ $k = 1, \dots, p$.

Theorem : Las (2006)

Let $f \in \sum_{k=1}^p \mathbb{R}[X(l_k)]$, let $\mathbf{K} \subset \mathbb{R}^n$ be compact, and assume that

$$\forall k = 2, \dots, p, \quad l_k \cap \left(\bigcup_{l=1}^{k-1} l_l \right) \subseteq l_s, \quad \text{for some } s \leq k-1$$

Then f positive on $\mathbf{K} \Rightarrow f \in \sum_{k=1}^p Q(\mathcal{G}_k)$

A sparse Positivstellensatz

- Add the p redundant quadratic constraints

$$g_{m+k}(X) = M - \|X(l_k)\|^2 \geq 0, \quad k = 1, \dots, p$$

in the definition of the constraint set \mathbf{K} . ($g_{m+k} \in \mathbb{R}[X(l_k)]$.)

Let $\mathcal{G}_k := \{g_j : g_j \in \mathbb{R}[X(l_k)]\}$ $k = 1, \dots, p$.

Theorem : Las (2006)

Let $f \in \sum_{k=1}^p \mathbb{R}[X(l_k)]$, let $\mathbf{K} \subset \mathbb{R}^n$ be compact, and assume that

$$\forall k = 2, \dots, p, \quad l_k \cap \left(\bigcup_{l=1}^{k-1} l_l \right) \subseteq l_s, \quad \text{for some } s \leq k-1$$

Then f positive on $\mathbf{K} \Rightarrow f \in \sum_{k=1}^p Q(\mathcal{G}_k)$

A sparse Positivstellensatz

- Add the p redundant quadratic constraints

$$g_{m+k}(X) = M - \|X(l_k)\|^2 \geq 0, \quad k = 1, \dots, p$$

in the definition of the constraint set \mathbf{K} . $(g_{m+k} \in \mathbb{R}[X(l_k)].)$

Let $\mathcal{G}_k := \{g_j : g_j \in \mathbb{R}[X(l_k)]\}$ $k = 1, \dots, p$.

Theorem : Las (2006)

Let $f \in \sum_{k=1}^p \mathbb{R}[X(l_k)]$, let $\mathbf{K} \subset \mathbb{R}^n$ be compact, and assume that

$$\forall k = 2, \dots, p, \quad l_k \cap \left(\bigcup_{l=1}^{k-1} l_l \right) \subseteq l_s, \quad \text{for some } s \leq k-1$$

Then f positive on $\mathbf{K} \Rightarrow f \in \sum_{k=1}^p Q(\mathcal{G}_k)$.

That is

$$f = \sum_{k=1}^p \left(\sigma_{0k} + \sum_{g_j \in \mathcal{G}_k} \sigma_{jk} g_j \right), \quad \sigma_{jk} \in \mathbb{R}[X(I_k)]^2$$

The hypothesis

$$\forall k = 2, \dots, p, \quad I_k \cap \left(\bigcup_{l=1}^{k-1} I_l \right) \subseteq I_s \quad \text{for some } s \leq k-1$$

is the **Running Intersection Property (RIP)** in graph theory.

Used in Bayesian networks for efficient **global update** of information from **local modifications** at some nodes of the network...

That is

$$f = \sum_{k=1}^p \left(\sigma_{0k} + \sum_{g_j \in \mathcal{G}_k} \sigma_{jk} g_j \right), \quad \sigma_{jk} \in \mathbb{R}[X(I_k)]^2$$

The hypothesis

$$\forall k = 2, \dots, p, \quad I_k \cap \left(\bigcup_{l=1}^{k-1} I_l \right) \subseteq I_s \quad \text{for some } s \leq k-1$$

is the **Running Intersection Property (RIP)** in graph theory.

Used in Bayesian networks for efficient **global update** of information from **local modifications** at some nodes of the network...

That is

$$f = \sum_{k=1}^p \left(\sigma_{0k} + \sum_{g_j \in \mathcal{G}_k} \sigma_{jk} g_j \right), \quad \sigma_{jk} \in \mathbb{R}[X(I_k)]^2$$

The hypothesis

$$\forall k = 2, \dots, p, \quad I_k \cap \left(\bigcup_{l=1}^{k-1} I_l \right) \subseteq I_s \quad \text{for some } s \leq k-1$$

is the **Running Intersection Property (RIP)** in graph theory.

Used in Bayesian networks for efficient **global update** of information from **local modifications** at some nodes of the network...

The proof relies on the following crucial property: Let $\nu_k(dX_i : i \in I_k)$, $k = 1, \dots, p$ be **probability measures** with **consistent marginals**, i.e., whenever $I_k \cap I_j \neq \emptyset$..

$$\nu_{k_j}(dX_i : i \in I_k \cap I_j) \equiv \nu_{j_k}(dX_i : i \in I_k \cap I_j) =: \nu_{jk}$$

Then, under **RIP**, one may construct from the local ν_k 's ... a **global prob. measure** $\mu(dX_i : i = 1, \dots, n)$ on \mathbb{R}^n , such that

$$\mu_{x_i}(dX_i : i \in I_k) \equiv \nu_k, \quad k = 1, \dots, p$$

The proof relies on the following crucial property: Let $\nu_k(dX_i : i \in I_k)$, $k = 1, \dots, p$ be **probability measures** with **consistent marginals**, i.e., whenever $I_k \cap I_j \neq \emptyset$..

$$\nu_{kj}(dX_i : i \in I_k \cap I_j) \equiv \nu_{jk}(dX_i : i \in I_k \cap I_j) =: \nu_{jk}$$

Then, under **RIP**, one may construct from the local ν_k 's ... a **global prob. measure** $\mu(dX_i : i = 1, \dots, n)$ on \mathbb{R}^n , such that

$$\mu_{X_i}(dX_i : i \in I_k) \equiv \nu_k, \quad k = 1, \dots, p$$

SPARSE SDP-relaxations for global optimization

For the global polynomial optimization problem

$$f^* = \min_x \{ f(x) : g_j(x) \geq 0, \quad j = 1, \dots, m \}$$

REPLACE the standard hierarchy of SDP-relaxations

$$Q_r^* \left\{ \begin{array}{l} \max_{\lambda, q_0, \dots, q_m} \lambda \\ f - \lambda = q_0 + \sum_{j=1}^m q_j g_j \\ q_j \text{ s.o.s.} \in \mathbb{R}[X] \quad \deg(q_j g_j) \leq 2r, \quad \forall j = 0, \dots, m \end{array} \right.$$

WITH the hierarchy of **SPARSE** SDP-relaxations

$$\mathbf{Q}_r^* \left\{ \begin{array}{l}
 \max_{\lambda, q_0, \dots, q_m} \lambda \\
 f - \lambda = q_0 + \sum_{j=1}^{m+p} q_j g_j \\
 q_j \text{ s.o.s.} \in \mathbb{R}[X(I_{k(j)})] \quad \deg(q_j g_j) \leq 2r, \quad \forall j = 1, \dots, m \\
 q_0 = \sum_{k=1}^p q_{0k}, \quad q_{0k} \text{ s.o.s.} \in \mathbb{R}[X(I_k)] \quad \deg(q_{0k}) \leq 2r
 \end{array} \right.$$

Convergence

Theorem : If the sparsity pattern $I = \bigcup_{k=1}^p I_k$ satisfies the running intersection property

$$\forall k = 2, \dots, p, \quad I_k \cap \left(\bigcup_{l=1}^{k-1} I_l \right) \subseteq I_s, \quad \text{for some } s \leq k-1$$

.... then $\min \mathbf{Q}_r \uparrow f^*$ as $r \rightarrow \infty$.

Sketch of the proof

Recall that one has

$$\begin{aligned}
 f^* &= \min_x \{ f(x) : g_j(x) \geq 0, \quad j = 1, \dots, m \} = \min_x \{ f(x) : x \in \mathbf{K} \} \\
 &= \min_{\mu} \left\{ \int f d\mu : \mu \in \mathcal{P}(\mathbf{K}) \right\}
 \end{aligned}$$

Let $\mathcal{G}_j := \{ g_k : g_k \in \mathbb{R}[X(l_j)] \} \quad j = 1, \dots, p.$

$\mathbf{K}_j := \{ x \in \mathbb{R}^{n_j} : g_k(x) \geq 0, \quad g_k \in \mathcal{G}_j \}$

$$\rho^* := \min_{\mu_1, \dots, \mu_p} \left\{ \sum_{j=1}^p \int f_j d\mu_j : \mu_j \in \mathcal{P}(\mathbf{K}_j) \right\} \leq f^*$$

Sketch of the proof

Recall that one has

$$\begin{aligned} f^* &= \min_x \{ f(x) : g_j(x) \geq 0, \quad j = 1, \dots, m \} = \min_x \{ f(x) : x \in \mathbf{K} \} \\ &= \min_{\mu} \left\{ \int f d\mu : \mu \in \mathcal{P}(\mathbf{K}) \right\} \end{aligned}$$

Let $\mathcal{G}_j := \{ g_k : g_k \in \mathbb{R}[X(l_j)] \} \quad j = 1, \dots, p.$

$$\mathbf{K}_j := \{ x \in \mathbb{R}^{n_j} : g_k(x) \geq 0, \quad g_k \in \mathcal{G}_j \}$$

$$\rho^* := \min_{\mu_1, \dots, \mu_p} \left\{ \sum_{j=1}^p \int f_j d\mu_j : \mu_j \in \mathcal{P}(\mathbf{K}_j) \right\} \leq f^*$$

In computing ρ^* ... one has replaced the **single** prob. measure μ on the **BIG** space \mathbb{R}^n with

ρ measures μ_1, \dots, μ_p on the **SMALLER** spaces $\mathbb{R}^{n_1}, \dots, \mathbb{R}^{n_p}$

So in general $\rho^* \leq f^*$... BUT under the running intersection property **RIP**

from the μ_j 's ... one may construct a **global** prob. measure μ on the whole \mathbb{R}^n , such that the marginal of μ on K_j is μ_j , for all $j = 1, \dots, p$. And therefore,

$$\rho^* = \sum_{j=1}^p \int f_j d\mu_j = \int f d\mu = f^*$$

In computing ρ^* ... one has replaced the **single** prob. measure μ on the **BIG** space \mathbb{R}^n with

ρ measures μ_1, \dots, μ_p on the **SMALLER** spaces $\mathbb{R}^{n_1}, \dots, \mathbb{R}^{n_p}$

So in general $\rho^* \leq f^*$... BUT under the running intersection property **RIP**

from the μ_j 's ... one may construct a **global** prob. measure μ on the whole \mathbb{R}^n , such that the marginal of μ on K_j is μ_j , for all $j = 1, \dots, p$. And therefore,

$$\rho^* = \sum_{j=1}^p \int f_j d\mu_j = \int f d\mu = f^*$$

STANDARD vs SPARSE SDP-relaxations

Let $\kappa := \max_{k=1,\dots,p} \text{card } I_k$.

STANDARD SDP-relaxation \mathbf{Q}_r	SPARSE SDP-relaxation \mathbf{Q}_r
$O(n^{2r})$ VARIABLES	$p O(\kappa^{2r})$ VARIABLES
$m + 1$ LMIs of size $O(n^r)$	$m + 2p$ LMIs of size $O(\kappa^r)$

Hence, **SPARSITY** yields a **big saving** if κ is **small**, as is frequently the case in applications with large scale problems.

Tests by Kojima's group with the specialized software **SPARSEPOPT** on a sample of nontrivial and nonconvex large scale problems, are very encouraging:

For instance, they could solve problems with up to $n = 500$ and even $n = 1000$ variables with κ in the range $[3, 7]$.

For much smaller problems ... one cannot even implement the first standard SDP-relaxation!

The \mathbf{K} -moment problem with density

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

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

and let μ be a given Borel measure on \mathbf{K} with moments

$$z_\alpha = \int \mathbf{x}^\alpha d\mu, \quad \alpha \in \mathbb{N}^n.$$

Problem: Given a sequence $\mathbf{y} = (y_\alpha) \subset \mathbb{R}$, does there exist a measure φ on \mathbf{K} , such that

- $\varphi \ll \mu$ (i.e. φ is absolutely continuous w.r.t. μ)

- the density $\frac{d\varphi}{d\mu}$ is in $L_\infty(\mathbf{K}, \mu)$, that is $d\varphi = f d\mu$, for some

$f \in L_\infty(\mathbf{K}, \mu)$

The \mathbf{K} -moment problem with density

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

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

and let μ be a given Borel measure on \mathbf{K} with moments

$$z_\alpha = \int \mathbf{x}^\alpha d\mu, \quad \alpha \in \mathbb{N}^n.$$

Problem: Given a sequence $\mathbf{y} = (y_\alpha) \subset \mathbb{R}$, does there exist a measure φ on \mathbf{K} , such that

- $\varphi \ll \mu$ (i.e. φ is absolutely continuous w.r.t. μ)

- the density $\frac{d\varphi}{d\mu}$ is in $L_\infty(\mathbf{K}, \mu)$, that is $d\varphi = f d\mu$, for some

$f \in L_\infty(\mathbf{K}, \mu)$

The compact case

For a sequence $y = (y_\alpha)$, define $L_y : \mathbb{R}[X] \rightarrow \mathbb{R}$ by

$$h (= \sum_{\alpha} h_{\alpha} X^{\alpha}) \mapsto L_y(h) = \sum_{\alpha \in \mathbb{N}^n} h_{\alpha} y_{\alpha}$$

Theorem: (Lasserre): Let \mathbf{K} be compact.

(a) The sequence $y = (y_{\alpha})$ has a representing measure $\varphi \ll \mu$ with density $f \in L_{\infty}(\mathbf{K}, \mu)$ if and only if $\exists \kappa > 0$ such that

$$0 \leq L_y(h) \leq \kappa L_z(h), \quad \forall h \in P(g) \quad (= P(g_1, \dots, g_m))$$

(b) If $N - \|X\|^2 \in Q(g)$ then in (a) replace $P(g)$ with $Q(g)$.

Sketch of proof:

Only if part is straightforward. If $\varphi = \int f d\mu$ for some $f \in L_\infty(\mathbf{K}, \mu)$ and as φ is supported on \mathbf{K} , one has

$$L_y(h^2) = \int h^2 f d\mu \leq \|f\|_\infty \int h^2 d\mu = \|f\|_\infty L_z(h^2) \quad \forall h \in P(g).$$

If part. $L_y(h^2) \geq 0$ for all $h \in P(g)$ implies that φ is supported on \mathbf{K} . Next, let $\gamma := \kappa z - y$. Then

$$L_\gamma(h^2) \geq 0 \quad \forall h \in P(g) \Rightarrow \gamma_\alpha = \int_{\mathbf{K}} X^\alpha d\psi \quad \forall \alpha$$

for some measure ψ on \mathbf{K} . Next, from the definition of γ

$$\int_{\mathbf{K}} h d(\psi + \varphi) = \int_{\mathbf{K}} h \kappa d\mu \quad \forall h \in \mathbb{R}[X]$$

and so $\psi + \varphi = \kappa\mu$, which implies $\varphi \leq \mu$ and so

Sketch of proof:

Only if part is straightforward. If $\varphi = \int f d\mu$ for some $f \in L_\infty(\mathbf{K}, \mu)$ and as φ is supported on \mathbf{K} , one has

$$L_y(h^2) = \int h^2 f d\mu \leq \|f\|_\infty \int h^2 d\mu = \|f\|_\infty L_z(h^2) \quad \forall h \in P(g).$$

If part. $L_y(h^2) \geq 0$ for all $h \in P(g)$ implies that φ is supported on \mathbf{K} . Next, let $\gamma := \kappa z - y$. Then

$$L_\gamma(h^2) \geq 0 \quad \forall h \in P(g) \Rightarrow \gamma_\alpha = \int_{\mathbf{K}} X^\alpha d\psi \quad \forall \alpha$$

for some measure ψ on \mathbf{K} . Next, from the definition of γ

$$\int_{\mathbf{K}} h d(\psi + \varphi) = \int_{\mathbf{K}} h \kappa d\mu \quad \forall h \in \mathbb{R}[X]$$

and so $\psi + \varphi = \kappa\mu$, which implies $\varphi \leq \mu$ and so $\varphi \ll \mu$.

By the Radon-Nykodim theorem

$d\varphi = f d\mu$ for some nonnegative $f \in L_1(\mathbf{K}, \mu)$.

But from $\varphi \leq \kappa \mu$, one gets

$$\int_A (f - \kappa) d\mu \leq 0 \quad \forall A \in \mathcal{B}(\mathbf{K})$$

which clearly implies $f \leq \kappa$ μ -almost everywhere

... and so ... $f \in L_\infty(\mathbf{K}, \mu)$ with $\|f\|_\infty \leq \kappa$ □

For all $J \subseteq \{1, \dots, m\}$, let $g_J(X) := \prod_{j \in J} g_j(X)$, and

$$P_r(g) := \left\{ \sum_{J \subseteq \{1, \dots, m\}} \sigma_J g_J, \quad \sigma_J \text{ s.o.s.} \quad \deg \sigma_J \leq r \right\}$$

If z is given, testing the condition $0 \leq L_z(h) \leq \kappa L_y(h)$ for all $h \in P_r(g)$ is solving a SDP in the **SINGLE VARIABLE** κ .

$$\kappa_r := \min_{\kappa \geq 0} \left\{ \kappa : 0 \preceq M_r(g_J z) \preceq \kappa M_r(g_J y) \right\}$$

Notice that $\kappa_r \uparrow$ as $r \rightarrow \infty$ and so, $\lim_{r \rightarrow \infty} \kappa_r < \infty$ is a

necessary and sufficient condition for z to have a representing measure $\varphi \ll \mu$ with density in $L_\infty(\mathbf{K}, \mu)$.

The case $\mathbf{K} = \mathbb{R}^n$

Theorem: (Lass): Let $y = (y_\alpha)$ be the moment sequence of a finite Borel measure μ on \mathbb{R}^n , and assume that

$$\sum_{k=1}^{\infty} \left(L_y(X_i^{2k}) \right)^{-\frac{1}{2k}} = +\infty, \quad i = 1, \dots, n$$

Then the sequence $z = (z_\alpha)$ has a representing measure $\varphi \ll \mu$ with density $f \in L_\infty(\mathbb{R}^n, \mu)$ if $\exists \kappa_1, \kappa_2 > 0$ such that

$$\kappa_1 L_y(X_i^{2k}) \leq L_z(X_i^{2k}) \leq \kappa_2 L_y(X_i^{2k}), \quad k = 0, 1, \dots,$$

for all $i = 1, \dots, n$.

- For every $i = 1, \dots, n$, one has

$$L_z(X_i^{2k})^{-1/2k} \geq \kappa_2^{-1/2k} L_y(X_i^{2k})^{-1/2k}, \quad k = 0, 1, \dots,$$

and so

$$\sum_{k=1}^{\infty} \left(L_z(X_i^{2k}) \right)^{-\frac{1}{2k}} = +\infty, \quad i = 1, \dots, n$$

which implies that z has a *unique* representing finite Borel measure φ on \mathbb{R}^n .

- The sequence $\gamma = \kappa_2 y - z$ has a representing (possibly signed) measure $\kappa_2 \mu - \varphi$.

But for all $i = 1, \dots, n$,

$$L_\gamma(X_i^{2k}) \leq (\kappa_2 - \kappa_1) L_z(X_i^{2k}), \quad k = 0, 1, \dots$$

which implies

$$L_\gamma(X_i^{2k})^{-1/2k} \geq (\kappa_2 - \kappa_1)^{-1/2k} L_\gamma(X_i^{2k})^{-1/2k}, \quad k = 0, 1, \dots,$$

and so γ has a **unique** representing Borel measure ψ on \mathbb{R}^n .

- Next from the definition of γ ,

$$\int h d(\psi + \varphi) = \int h \kappa_2 d\mu, \quad \forall h \in \mathbb{R}[X]$$

and so $\psi + \varphi = \kappa_2 \mu$ because μ is **moment determinate** ...

Therefore, $\varphi \leq \kappa_2 \mu$ which implies $\varphi \ll \mu$.

- By the Radon-Nykodim theorem $d\varphi = \kappa_2 f d\mu$ for some $f \in L_1(\mathbb{R}^n, \mu)$, and in fact $f \in L_\infty(\mathbb{R}^n, \mu)$, with $\|f\|_\infty \leq \kappa_2$. □