The moment-SOS approach II: Some Applications

Jean B. Lasserre

LAAS-CNRS and Institute of Mathematics, Toulouse, France

Master 2 Optimization: Paris-Orsay, 2016

A preliminary result

- Let $\mathbf{B} \subset \mathbb{R}^n$ be a simple set like e.g., a Box $[-1, 1]^n$ or an ellipsoid,
- $\mathbf{K} := \{ (x, y) : \; x \in \mathbf{B}; \; g_j(x, y) \geq 0, \; j = 1, \ldots, m \}$ and $h : \mathbb{R}^n \times \mathbb{R}^p :\to \mathbb{R}$.

Goal: Approximate the function $F : \mathbf{B} \to \mathbb{R}$

$$x \mapsto F(x) := \sup_y \{ h(x, y) : \; (x, y) \in \mathbf{K} \}, \qquad x \in \mathbf{B}.$$

... by POLYNOMIALS $J_k \in \mathbb{R}[x]$ of increasing degree ... and with some guarantee of convergence.

A preliminary result

- Let $\mathbf{B} \subset \mathbb{R}^n$ be a simple set like e.g., a Box $[-1, 1]^n$ or an ellipsoid,
- $\mathbf{K} := \{ (x, y) : \; x \in \mathbf{B};\; g_j(x, y) \,\geq\, 0,\; j = 1, \ldots, m \}$ and $h : \mathbb{R}^n \times \mathbb{R}^p :\to \mathbb{R}$.

Goal: Approximate the function $F : \mathbf{B} \to \mathbb{R}$

$$x \mapsto F(x) \;:=\; \sup_{y} \{ h(x, y) :\; (x, y) \,\in\, \mathbf{K} \}, \qquad x \in \mathbf{B}.$$

... by POLYNOMIALS $J_k \in \mathbb{R}[x]$ of increasing degree ... and with some guarantee of convergence.

Fix $k \in \mathbb{N}$ and consider the optimization problem:

$$\inf_{J \in \mathbb{R}[x]_k} \left\{ \int_{\mathbf{B}} (J(x) - F(x))\, dx : J(x) \geq F(x), \quad \forall x \in \mathbf{B} \right\}$$

Or, equivalently:

$$\inf_{J \in \mathbb{R}[x]_k} \{ \|J - F\|_1 : p(x) \geq F(x), \quad \forall x \in \mathbf{B} \}$$

where $\|\cdot\|_1$ is the L_1-norm

$$h \mapsto \|h\|_1 := \int_{\mathbf{B}} |h(x)|\, dx, \quad h \in L_1(\mathbf{B}).$$

Fix $k \in \mathbb{N}$ and consider the optimization problem:

$$\inf_{J \in \mathbb{R}[x]_k} \left\{ \int_{\mathbf{B}} (J(x) - F(x))\, dx : J(x) \geq F(x), \quad \forall x \in \mathbf{B} \right\}$$

Or, equivalently:

$$\inf_{J \in \mathbb{R}[x]_k} \{ \|J - F\|_1 : p(x) \geq F(x), \quad \forall x \in \mathbf{B} \}$$

where $\|\cdot\|_1$ is the L_1-norm

$$h \mapsto \|h\|_1 := \int_{\mathbf{B}} |h(x)|\, dx, \quad h \in L_1(\mathbf{B}).$$

To fix ideas, let $\mathbf{B} := [-1, 1]^n = \{x : x_i^2 \leq 1, i = 1, \ldots, n\}$ and set $g_0(x) = 1$ and $g_{m+i}(x) = 1 - x_i^2$, $i = 1, \ldots, n$.

A simple strategy:

Replace the difficult constraint $J(x) - F(x)$, $\forall x \in \mathbf{B}$ with the SOS-based positivity certificate

$$J(x) - h(x, y) = \sum_{j=0}^{m+n} \sigma_j(x, y)\, g_j(x, y),$$

for some SOS polynomials (σ_j) such that $\deg(\sigma_j g_j) \leq 2k$.

Indeed

$$J(x) \geq h(x, y), \quad \forall (x, y) \in \mathbf{K}; x \in \mathbf{B} \Rightarrow J(x) \geq F(x), \quad \forall x \in \mathbf{B}.$$

To fix ideas, let $\mathbf{B} := [-1,1]^n = \{x : x_i^2 \leq 1, i = 1, \ldots, n\}$ and set $g_0(x) = 1$ and $g_{m+i}(x) = 1 - x_i^2$, $i = 1, \ldots, n$.

A simple strategy:

Replace the difficult constraint $J(x) - F(x)$, $\forall x \in \mathbf{B}$ with the SOS-based positivity certificate

$$J(x) - h(x, y) = \sum_{j=0}^{m+n} \sigma_j(x, y)\, g_j(x, y),$$

for some SOS polynomials (σ_j) such that $\deg(\sigma_j g_j) \leq 2k$.

Indeed

$$J(x) \geq h(x, y), \quad \forall (x, y) \in \mathbf{K}; x \in \mathbf{B} \Rightarrow J(x) \geq F(x), \quad \forall x \in \mathbf{B}.$$

Then solve:

$$\inf_{J,\sigma_j} \left\{ \int_{\mathbf{B}} (J(x) - F(x))\, dx : \right.$$

$$\text{s.t.} \quad J(x) - h(x,y) = \sum_{j=0}^{m+n} \sigma_j(x,y)\, g_j(x,y), \quad \forall (x,y) \in \mathbb{R}^{n+p}$$

$$\left. J \in \mathbb{R}[\mathbf{x}]_k;\ \sigma_j \text{ SOS};\ \deg(\sigma_j\, g_j) \leq 2k \right\}.$$

Theorem

☞ *This is an SDP and it has an optimal solution* $J_k^* \in \mathbb{R}[x]_k$. *Moreover,* $\|J_k^* - f\|_1 \to 0$ *as* $k \to \infty$.

Then solve:

$$\inf_{J,\sigma_j} \left\{ \int_{\mathbf{B}} (J(x) - F(x))\, dx : \right.$$

$$\text{s.t.} \quad J(x) - h(x,y) = \sum_{j=0}^{m+n} \sigma_j(x,y)\, g_j(x,y), \quad \forall (x,y) \in \mathbb{R}^{n+p}$$

$$\left. J \in \mathbb{R}[\mathbf{x}]_k;\ \sigma_j \text{ SOS}; \deg(\sigma_j\, g_j) \leq 2k \right\}.$$

Theorem

☞ *This is an SDP and it has an optimal solution* $J_k^* \in \mathbb{R}[x]_k$. *Moreover,* $\|J_k^* - f\|_1 \to 0$ *as* $k \to \infty$.

Write $J(x) = \sum_{\alpha \in \mathbb{N}^n} J_\alpha \, x^\alpha$.

I. ☞ The criterion $\int_{\mathbf{B}} (J(x) - f(x))\, dx$ reads:

$$\sum_{\alpha \in \mathbb{N}^n} J_\alpha \underbrace{\int_{\mathbf{B}} x^\alpha \, dx}_{=\gamma_\alpha \text{ known}} - \underbrace{\int_B F(\mathbf{x})\, dx}_{\text{an (unknown) constant } c} = \sum_{\alpha \in \mathbb{N}^n} J_\alpha \, \gamma_\alpha - c,$$

is LINEAR in the unknown coefficients J_α of $J \in \mathbb{R}[x]_k$.

II. ☞ The constraint

$$J(x) - h(x, y) = \sum_{j=0}^{m+n} \sigma_j(x, y)\, g_j(x, y),$$

reduces to

- LINEAR constraints on the coefficients J_α and $\sigma_{j\alpha}$,
- + semidefinite constraints to state that the σ_j's are SOS.

I. Approximation of sets with quantifiers

Let $f \in \mathbb{R}[x, y]$ and let $\mathbf{K} \subset \mathbb{R}^n \times \mathbb{R}^p$ be the semi-algebraic set:

$$\mathbf{K} := \{(x, y) : \quad g_j(x, y) \geq 0, \quad j = 1, \ldots, m\},$$

and let $\mathbf{B} \subset \mathbb{R}^n$ be the Euclidean unit ball or the box $[-1, 1]^n$.

Suppose that one wants to approximate the set:

$$R_f := \{x \in \mathbf{B} : \; f(x, y) \leq 0 \text{ for all } y \text{ such that } (x, y) \in \mathbf{K}\}$$

as closely as desired by a sequence of sets of the form:

$$\Theta_k := \{\mathbf{x} \in \mathbf{B} : \; J_k(x) \leq 0\,\}$$

for some polynomials J_k.

Recall that $\mathbf{B} = \{x : 1 - x_i^2 \geq 0\}$ and set $g_{m+i}(x) := 1 - x_i^2$, $i = 1, \ldots, n$. With $g_0 = 1$ and with $\mathbf{K} \subset \mathbb{R}^n \times \mathbb{R}^p$ and $k \in \mathbb{N}$, let

$$Q_k(g) = \left\{ \sum_{j=0}^{m+n} \sigma_j(x,y)\, g_j(x,y) : \quad \sigma_j \in \Sigma[x,y], \deg(\sigma_j\, g_j) \leq 2k \right\}$$

Let $x \mapsto F(x) := \max\{f(x,y) : (x,y) \in \mathbf{K}\}$, and

for every integer k consider the optimization problem:

$$\rho_k = \min_{J \in \mathbb{R}[x]_k} \left\{ \int_{\mathbf{B}} (J - F)\, dx \; : \; J(x) - f(x,y) \in Q_k(g); \; x \in \mathbf{B} \right\}$$

☞ Remember that $J(x) - f(x,y) \in Q_k(g)$ implies:

$$J(x) \geq f(x,y), \quad \forall (x,y) \in \mathbf{K} \quad \Rightarrow J(x) \geq F(x) \quad \forall x \in \mathbf{B}.$$

Recall that $\mathbf{B} = \{x : 1 - x_i^2 \geq 0\}$ and set $g_{m+i}(x) := 1 - x_i^2$, $i = 1, \ldots, n$. With $g_0 = 1$ and with $\mathbf{K} \subset \mathbb{R}^n \times \mathbb{R}^p$ and $k \in \mathbb{N}$, let

$$Q_k(g) = \left\{ \sum_{j=0}^{m+n} \sigma_j(x,y)\, g_j(x,y) : \quad \sigma_j \in \Sigma[x,y], \deg(\sigma_j\, g_j) \leq 2k \right\}$$

Let $x \mapsto F(x) := \max\,\{f(x,y) : (x,y) \in \mathbf{K}\}$, and

for every integer k consider the optimization problem:

$$\rho_k = \min_{J \in \mathbb{R}[x]_k} \left\{ \int_{\mathbf{B}} (J - F)\, dx \; : \; J(x) - f(x,y) \in Q_k(g);\; x \in \mathbf{B} \right\}$$

☞ Remember that $J(x) - f(x,y) \in Q_k(g)$ implies:

$$J(x) \geq f(x,y), \quad \forall (x,y) \in \mathbf{K} \quad \Rightarrow J(x) \geq F(x) \quad \forall x \in \mathbf{B}.$$

Recall that $\mathbf{B} = \{x : 1 - x_i^2 \geq 0\}$ and set $g_{m+i}(x) := 1 - x_i^2$, $i = 1, \ldots, n$. With $g_0 = 1$ and with $\mathbf{K} \subset \mathbb{R}^n \times \mathbb{R}^p$ and $k \in \mathbb{N}$, let

$$Q_k(g) = \left\{ \sum_{j=0}^{m+n} \sigma_j(x,y)\, g_j(x,y) : \quad \sigma_j \in \Sigma[x,y], \deg(\sigma_j\, g_j) \leq 2k \right\}$$

Let $x \mapsto F(x) := \max\{f(x,y) : (x,y) \in \mathbf{K}\}$, and

for every integer k consider the optimization problem:

$$\rho_k = \min_{J \in \mathbb{R}[x]_k} \left\{ \int_{\mathbf{B}} (J - F)\, dx \ : \ J(x) - f(x,y) \in Q_k(g);\ x \in \mathbf{B} \right\}$$

☞ Remember that $J(x) - f(x,y) \in Q_k(g)$ implies:

$$J(x) \geq f(x,y), \quad \forall (x,y) \in \mathbf{K} \quad \Rightarrow J(x) \geq F(x) \quad \forall x \in \mathbf{B}.$$

1. The criterion

$$\int_{\mathbf{B}} (J - F)\, dx = \underbrace{\int_{\mathbf{B}} -F\, dx}_{\text{unknown but constant}} + \sum_{\alpha} J_\alpha \underbrace{\int_{\mathbf{B}} \mathbf{x}^\alpha\, dx}_{\text{easy to compute}}$$

is LINEAR in the coefficients J_α of the unknown polynomial $J \in \mathbb{R}[\mathbf{x}]_k$!

2. The constraint

$$J(x) - f(x, y) = \sum_{j=0}^{m} \sigma_j(x, y)\, g_j(x, y)$$

is just LINEAR CONSTRAINTS + LMIs!

1. The criterion

$$\int_{\mathbf{B}} (J - F)\, dx = \underbrace{\int_{\mathbf{B}} -F\, dx}_{\text{unknown but constant}} + \sum_{\alpha} J_{\alpha} \underbrace{\int_{\mathbf{B}} \mathbf{x}^{\alpha}\, dx}_{\text{easy to compute}}$$

is LINEAR in the coefficients J_α of the unknown polynomial $J \in \mathbb{R}[\mathbf{x}]_k$!

2. The constraint

$$J(x) - f(x,y) = \sum_{j=0}^{m} \sigma_j(x,y)\, g_j(x,y)$$

is just LINEAR CONSTRAINTS + LMIs!

Hence, the optimization problem

$$\rho_k = \min_{J \in \mathbb{R}[x]_k} \left\{ \int_{\mathbf{B}} (J - F)\, dx \; : \; J(\mathbf{x}) - f(\mathbf{x}, y) \in Q_k(g) \right\}$$

IS AN SDP! Moreover, it has an optimal solution $J_k^* \in \mathbb{R}[x]_k$!

- Alternatively, if one uses LP-based positivity certificates for $J(\mathbf{x}) - f(\mathbf{x}, y)$, one ends up with solving an LP!

From the definition of J_k^*, the sublevel sets

$$\Theta_k := \{x \in \mathbf{B} : J_k^*(x) \leq 0\} \subset R_f, \quad k \in \mathbb{N},$$

provide a nested sequence of INNNER approximations of R_f.

Hence, the optimization problem

$$\rho_k = \min_{J \in \mathbb{R}[x]_k} \left\{ \int_{\mathbf{B}} (J - F)\, dx \;:\; J(\mathbf{x}) - f(\mathbf{x}, y) \in Q_k(g) \right\}$$

IS AN SDP! Moreover, it has an optimal solution $J_k^* \in \mathbb{R}[x]_k$!

- Alternatively, if one uses LP-based positivity certificates for $J(\mathbf{x}) - f(\mathbf{x}, y)$, one ends up with solving an LP!

From the definition of J_k^*, the sublevel sets

$$\Theta_k := \{x \in \mathbf{B} : J_k^*(x) \leq 0\} \subset R_f, \quad k \in \mathbb{N},$$

provide a nested sequence of INNNER approximations of R_f.

Hence, the optimization problem

$$\rho_k = \min_{J\in\mathbb{R}[x]_k} \left\{ \int_{\mathbf{B}} (J - F)\,dx \;:\; J(\mathbf{x}) - f(\mathbf{x}, y) \in Q_k(g) \right\}$$

IS AN SDP! Moreover, it has an optimal solution $J_k^* \in \mathbb{R}[x]_k$!

• Alternatively, if one uses LP-based positivity certificates for $J(\mathbf{x}) - f(\mathbf{x}, y)$, one ends up with solving an LP!

From the definition of J_k^*, the sublevel sets

$$\Theta_k := \{x \in \mathbf{B} : J_k^*(x) \le 0\} \subset R_f, \quad k \in \mathbb{N},$$

provide a nested sequence of INNNER approximations of R_f.

Theorem (Lass)

(Strong) convergence in $L_1(\mathbf{B})$-norm takes place, that is:

$$\lim_{k\to\infty} \int_{\mathbf{B}} |\, J_k^* - F \,|\, dx \;=\; 0$$

and, if in addition the set $\{x \in \mathbf{B} : \; F(x) = 0\}$ has Lebesgue measure zero, then

$$\lim_{k\to\infty} \; \mathrm{VOL}(R_f \setminus \Theta_k) \;=\; 0$$

Theorem (Lass)

(Strong) convergence in $L_1(\mathbf{B})$-norm takes place, that is:

$$\lim_{k\to\infty} \int_{\mathbf{B}} |\, J_k^* - F \,|\, dx \;=\; 0$$

and, if in addition the set $\{x \in \mathbf{B} : \; F(x) = 0\}$ has Lebesgue measure zero, then

$$\lim_{k\to\infty} \;\mathrm{VOL}(R_f \setminus \Theta_k) \;=\; 0$$

Ex: Polynomial Matrix Inequalities: (with D. Henrion)

Let $x \mapsto \mathbf{A}(x) \in \mathbb{R}^{p\times p}$ where $\mathbf{A}(x)$ is the matrix-polynomial

$$x \mapsto \mathbf{A}(x) = \sum_{\alpha\in\mathbb{N}^n} \mathbf{A}_\alpha \, x^\alpha \quad \left(= \sum_{\alpha\in\mathbb{N}^n} \mathbf{A}_\alpha \, x_1^{\alpha_1} \cdots x_n^{\alpha_n} \right).$$

for finitely many real symmetric matrices $(\mathbf{A}_\alpha)$, $\alpha \in \mathbb{N}^n$.

... and suppose one wants to approximate the set

$$R_{\mathbf{A}} := \{x \in \mathbf{B} \,:\, \mathbf{A}(x) \succeq 0\} = \{x \,:\, \lambda_{\min}(\mathbf{A}(x)) \geq 0\}.$$

Then:

$$R_{\mathbf{A}} = \left\{ x \in \mathbf{B} \,:\, \underbrace{y^T \mathbf{A}(x) y}_{f(x,y)} \geq 0, \quad \forall\, y \text{ s.t. } \|y\|^2 = 1 \right\}$$

Ex: Polynomial Matrix Inequalities: (with D. Henrion)

Let $x \mapsto \mathbf{A}(x) \in \mathbb{R}^{p\times p}$ where $\mathbf{A}(x)$ is the matrix-polynomial

$$x \mapsto \mathbf{A}(x) = \sum_{\alpha\in\mathbb{N}^n} \mathbf{A}_\alpha\, x^\alpha \quad \left(= \sum_{\alpha\in\mathbb{N}^n} \mathbf{A}_\alpha\, x_1^{\alpha_1}\cdots x_n^{\alpha_n}\right).$$

for finitely many real symmetric matrices $(\mathbf{A}_\alpha)$, $\alpha \in \mathbb{N}^n$.

... and suppose one wants to approximate the set

$$R_{\mathbf{A}} := \{x \in \mathbf{B} : \mathbf{A}(x) \succeq 0\} = \{x : \lambda_{\min}(\mathbf{A}(x)) \geq 0\}.$$

Then:

$$R_{\mathbf{A}} = \left\{ x \in \mathbf{B} : \underbrace{y^T\mathbf{A}(x)y}_{f(x,y)} \geq 0, \quad \forall\, y \text{ s.t. } \|y\|^2 = 1 \right\}$$

Let **B** be the unit disk $\{\mathbf{x} : \|\mathbf{x}\| \leq 1\}$ and let:

$$R_{\mathbf{A}} := \left\{ \mathbf{x} \in \mathbf{B} \;:\; \mathbf{A}(\mathbf{x}) \left(= \begin{bmatrix} 1 - 16x_1x_2 & x_1 \\ x_1 & 1 - x_1^2 - x_2^2 \end{bmatrix} \right) \succeq 0 \right\}$$

Then by solving relatively simple semidefinite programs, one may approximate $R_{\mathbf{A}}$ with sublevel sets of the form:

$$\Theta_k := \{x \in \mathbf{B} : J_k^*(x) \geq 0\}$$

for some polynomial J_k^* of degree $k = 2, 4, \ldots$ and with

$$\mathrm{VOL}\,(R_A \setminus \Theta_k) \to 0 \qquad \text{as } k \to \infty.$$

Illustrative example (continued)

Let $\mathbf{B}$ be the unit disk $\{\mathbf{x} : \|\mathbf{x}\| \leq 1\}$ and let:

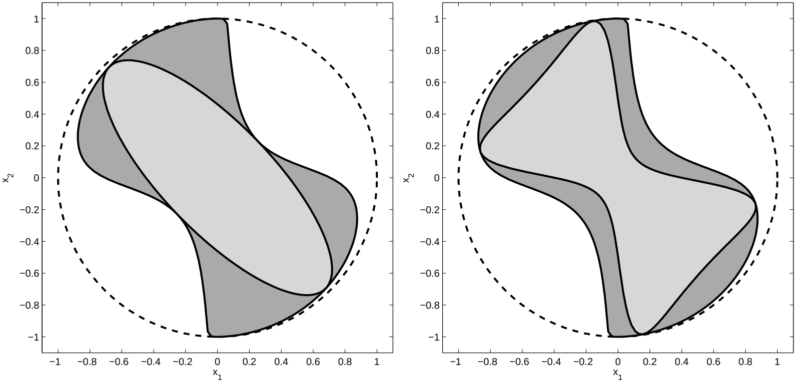
$$R_{\mathbf{A}} := \left\{ \mathbf{x} \in \mathbf{B} \;:\; \mathbf{A}(\mathbf{x}) \left(= \begin{bmatrix} 1 - 16x_1x_2 & x_1 \\ x_1 & 1 - x_1^2 - x_2^2 \end{bmatrix} \right) \succeq 0 \right\}$$

Then by solving relatively simple semidefinite programs, one may approximate $R_{\mathbf{A}}$ with sublevel sets of the form:

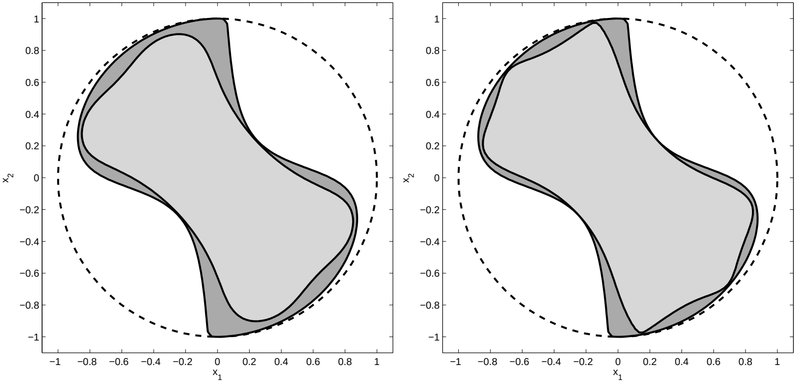
$$\Theta_k := \{ x \in \mathbf{B} : J_k^*(x) \geq 0 \}$$

for some polynomial J_k^* of degree $k = 2, 4, \ldots$ and with

$$\mathrm{VOL}\,(R_A \setminus \Theta_k) \to 0 \qquad \text{as } k \to \infty.$$



Θ_2 (left) and Θ_4 (right) inner approximations (light gray) of (dark gray) embedded in unit disk **B** (dashed).



Θ_6 (left) and Θ_8 (right) inner approximations (light gray) of (dark gray) embedded in unit disk $\mathbf{B}$ (dashed).

II. Convex Underestimators of Polynomials

A typical MINLP is of the form:

$$\mathbf{P}: \quad \inf_{\mathbf{x}} \{ f(\mathbf{x}) : \mathbf{x} \in \mathbf{K} \cap \mathbf{B};\ x_i \in \{0,1\},\ \forall i \in I \},$$

where

- $\mathbf{K} := \{\mathbf{x} : g_j(\mathbf{x}) \leq 0,\ j = 1, \ldots, m\}$ for some $(g_j) \subset \mathbb{R}[\mathbf{x}]$,
- f is a (low) degree-d non convex polynomial ($d \leq 4$).
- $\mathbf{B}$ is a box, e.g., $[0,1]^n$

hence very difficult to solve!

A typical MINLP is of the form:

$$\mathbf{P}: \quad \inf_{\mathbf{x}} \{ f(\mathbf{x}) : \mathbf{x} \in \mathbf{K} \cap \mathbf{B};\ x_i \in \{0,1\},\ \forall i \in I \},$$

where

- $\mathbf{K} := \{\mathbf{x} : g_j(\mathbf{x}) \leq 0,\ j = 1, \ldots, m\}$ for some $(g_j) \subset \mathbb{R}[\mathbf{x}]$,
- f is a (low) degree-d non convex polynomial ($d \leq 4$).
- $\mathbf{B}$ is a box, e.g., $[0,1]^n$

hence very difficult to solve!

• In the context of large scale MINLP the most efficient & popular strategy is to use BRANCH & BOUND combined with efficient LOWER BOUNDING techniques used at each node of the search tree.

• Typically, f is a sum $\sum_k f_k$ where each f_k “sees" only very few variables (say $3, 4$). The same observation is true for each g_j in the constraints:

Hence a very appealing idea is to pre-compute CONVEX UNDERESTIMATORS $\widehat{f_k} \leq f_k$ and $\widehat{g_j} \leq g_j$ for each non convex f_k and each non convex g_j, independently and separately!

$\rightarrow$ hence potentially many BUT LOW-DIMENSIONAL problems.

• In the context of large scale MINLP the most efficient & popular strategy is to use BRANCH & BOUND combined with efficient LOWER BOUNDING techniques used at each node of the search tree.

• Typically, f is a sum $\sum_k f_k$ where each f_k "sees" only very few variables (say $3, 4$). The same observation is true for each g_j in the constraints:

Hence a very appealing idea is to pre-compute CONVEX UNDERESTIMATORS $\widehat{f_k} \leq f_k$ and $\widehat{g_j} \leq g_j$ for each non convex f_k and each non convex g_j, independently and separately!

$\rightarrow$ hence potentially many BUT LOW-DIMENSIONAL problems.

• In the context of large scale MINLP the most efficient & popular strategy is to use BRANCH & BOUND combined with efficient LOWER BOUNDING techniques used at each node of the search tree.

• Typically, f is a sum $\sum_k f_k$ where each f_k "sees" only very few variables (say $3, 4$). The same observation is true for each g_j in the constraints:

Hence a very appealing idea is to pre-compute CONVEX UNDERESTIMATORS $\widehat{f_k} \leq f_k$ and $\widehat{g_j} \leq g_j$ for each non convex f_k and each non convex g_j, independently and separately!

$\rightarrow$ hence potentially many BUT LOW-DIMENSIONAL problems.

• Then at each node of the search tree of the B & B tree one computes a LOWER BOUND by solving convex optimization problems the form:

$$\widehat{\mathbf{P}}: \quad \inf_{\mathbf{x}} \{ \sum_k \widehat{f_k}(\mathbf{x}) : \widehat{g}_j(\mathbf{x}) \leq 0, \; j = 1, \ldots, m; \; \mathbf{x} \in \mathbf{B} \}$$

where some of the integer variables are fixed at 0 or 1.

and then one explores the search tree.

Hence one has to solve the generic problem

Compute a "tight" convex polynomial underestimator $p \leq f$ of a non convex polynomial f on a box $\mathbf{B} \subset \mathbb{R}^n$.

Message:

"Good" CONVEX POLYNOMIAL UNDERESTIMATORS can be computed efficiently!

Hence one has to solve the generic problem

Compute a "tight" convex polynomial underestimator $p \leq f$ of a non convex polynomial f on a box $\mathbf{B} \subset \mathbb{R}^n$.

Message:

“Good" CONVEX POLYNOMIAL UNDERESTIMATORS can be computed efficiently!

I: Characterizing convex polynomial underestimators

1. $p(\mathbf{x}) \leq f(\mathbf{x})$ for every $\mathbf{x} \in \mathbf{B}$.
2. p convex on $\mathbf{B} \rightarrow \nabla^2 p(\mathbf{x}) \succeq 0$ for all $\mathbf{x} \in \mathbf{B}$,

$$\Longleftrightarrow \quad \mathbf{u}^T \nabla^2 p(\mathbf{x})\, \mathbf{u} \geq 0, \; \forall (\mathbf{x}, \mathbf{u}) \in \mathbf{B} \times \mathbf{U},$$

where $\mathbf{U} := \{\mathbf{u} : \|\mathbf{u}\|^2 \leq 1\}$.

Hence we have the two "Positivity constraints"

$$\begin{aligned} f(\mathbf{x}) - p(\mathbf{x}) &\geq 0, \quad \forall\, \mathbf{x} \in \mathbf{B} \\ \mathbf{u}^T \nabla^2 p(\mathbf{x})\, \mathbf{u} &\geq 0, \quad \forall (\mathbf{x}, \mathbf{u}) \in \mathbf{B} \times \mathbf{U}. \end{aligned}$$

I: Characterizing convex polynomial underestimators

1. $p(\mathbf{x}) \leq f(\mathbf{x})$ for every $\mathbf{x} \in \mathbf{B}$.
2. p convex on $\mathbf{B} \rightarrow \nabla^2 p(\mathbf{x}) \succeq 0$ for all $\mathbf{x} \in \mathbf{B}$,

$$\Longleftrightarrow \quad \mathbf{u}^T \nabla^2 p(\mathbf{x})\, \mathbf{u} \geq 0, \; \forall (\mathbf{x}, \mathbf{u}) \in \mathbf{B} \times \mathbf{U},$$

where $\mathbf{U} := \{ \mathbf{u} : \|\mathbf{u}\|^2 \leq 1 \}$.

Hence we have the two "Positivity constraints"

$$\begin{aligned} f(\mathbf{x}) - p(\mathbf{x}) &\geq 0, \quad \forall \mathbf{x} \in \mathbf{B} \\ \mathbf{u}^T \nabla^2 p(\mathbf{x})\, \mathbf{u} &\geq 0, \quad \forall (\mathbf{x}, \mathbf{u}) \in \mathbf{B} \times \mathbf{U}. \end{aligned}$$

I: Characterizing convex polynomial underestimators

1. $p(\mathbf{x}) \leq f(\mathbf{x})$ for every $\mathbf{x} \in \mathbf{B}$.
2. p convex on $\mathbf{B} \to \nabla^2 p(\mathbf{x}) \succeq 0$ for all $\mathbf{x} \in \mathbf{B}$,

$$\iff \quad \mathbf{u}^T \nabla^2 p(\mathbf{x})\,\mathbf{u} \,\geq\, 0,\ \forall (\mathbf{x}, \mathbf{u}) \in \mathbf{B} \times \mathbf{U},$$

where $\mathbf{U} := \{\,\mathbf{u} : \|\mathbf{u}\|^2 \leq 1\}$.

Hence we have the two "Positivity constraints"

$$\begin{aligned} f(\mathbf{x}) - p(\mathbf{x}) &\geq 0, \quad \forall\, \mathbf{x} \in \mathbf{B} \\ \mathbf{u}^T \nabla^2 p(\mathbf{x})\,\mathbf{u} &\geq 0, \quad \forall (\mathbf{x}, \mathbf{u}) \in \mathbf{B} \times \mathbf{U}. \end{aligned}$$

II: Characterizing "tightness"

One possibility is to evaluate the L_1-norm $\displaystyle\int_{\mathbf{B}} |f(\mathbf{x}) - p(\mathbf{x})|\, d\mathbf{x}$

$$\rightarrow \int_{\mathbf{B}} (f(\mathbf{x}) - p(\mathbf{x})\, d\mathbf{x} = \underbrace{\int_{\mathbf{B}} f(\mathbf{x}) d\mathbf{x}}_{constant} - \underbrace{\int_{\mathbf{B}} p(\mathbf{x})\, d\mathbf{x}}_{\text{linear in } p!}$$

Indeed, writing $p(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha\, \mathbf{x}^\alpha$,

$$\int_{\mathbf{B}} p(\mathbf{x})\, d\mathbf{x} = \sum_{\alpha \in \mathbb{N}^n} p_\alpha \underbrace{\int_{\mathbf{B}} \mathbf{x}^\alpha\, d\mathbf{x}}_{\gamma_\alpha},$$

where γ_α is known (and easy to compute)!

II: Characterizing "tightness"

One possibility is to evaluate the L_1-norm $\int_{\mathbf{B}} |f(\mathbf{x}) - p(\mathbf{x})|\, d\mathbf{x}$

$$\rightarrow \int_{\mathbf{B}} (f(\mathbf{x}) - p(\mathbf{x})\, d\mathbf{x} = \underbrace{\int_{\mathbf{B}} f(\mathbf{x}) d\mathbf{x}}_{\textit{constant}} - \underbrace{\int_{\mathbf{B}} p(\mathbf{x})\, d\mathbf{x}}_{\text{linear in } p!}$$

Indeed, writing $p(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha\, \mathbf{x}^\alpha$,

$$\int_{\mathbf{B}} p(\mathbf{x})\, d\mathbf{x} = \sum_{\alpha \in \mathbb{N}^n} p_\alpha \underbrace{\int_{\mathbf{B}} \mathbf{x}^\alpha\, d\mathbf{x}}_{\gamma_\alpha},$$

where γ_α is known (and easy to compute)!

II: Characterizing "tightness"

One possibility is to evaluate the L_1-norm $\int_{\mathbf{B}} |f(\mathbf{x}) - p(\mathbf{x})| \, d\mathbf{x}$

$$\rightarrow \int_{\mathbf{B}} (f(\mathbf{x}) - p(\mathbf{x}) \, d\mathbf{x} = \underbrace{\int_{\mathbf{B}} f(\mathbf{x}) d\mathbf{x}}_{\textit{constant}} - \underbrace{\int_{\mathbf{B}} p(\mathbf{x}) \, d\mathbf{x}}_{\text{linear in } p!}$$

Indeed, writing $p(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha \, \mathbf{x}^\alpha$,

$$\int_{\mathbf{B}} p(\mathbf{x}) \, d\mathbf{x} = \sum_{\alpha \in \mathbb{N}^n} p_\alpha \underbrace{\int_{\mathbf{B}} \mathbf{x}^\alpha \, d\mathbf{x}}_{\gamma_\alpha},$$

where γ_α is known (and easy to compute)!

Hence computing the best degree-d convex polynomial underestimator of f reduces to solve the CONVEX optimization problem:

$$\mathbf{P}: \quad \rho = \inf_{p \in \mathbb{R}[\mathbf{x}]_d} \sum_{\alpha \in \mathbb{N}^n_d} p_\alpha \, \gamma_\alpha$$

$$\text{s.t.} \quad f(\mathbf{x}) - p(\mathbf{x}) \geq 0, \; \forall \mathbf{x} \in \mathbf{B}$$

$$\mathbf{u}^T \nabla^2 p(\mathbf{x}) \, \mathbf{u} \geq 0, \; \forall (\mathbf{x}, \mathbf{u}) \in \mathbf{B} \times \mathbf{U}.$$

☞ which has an optimal solution $p^* \in \mathbb{R}[\mathbf{x}]_d$

Hence we replace **P** with the HIERARCHY of SEMIDEFINITE PROGRAMS

$$\rho_\ell = \inf_{p \in \mathbb{R}[\mathbf{x}]_d} \sum_{\alpha \in \mathbb{N}^n_d} p_\alpha \, \gamma_\alpha$$

$$\text{s.t.} \quad f - p = \sigma_0(\mathbf{x}) + \sum_{j=1}^m \sigma_j(\mathbf{x}) g_j(\mathbf{x}) + \sigma_{m+1}(M - \|\mathbf{x}\|^2)$$

$$\mathbf{u}^T \nabla^2 p(\mathbf{x}) \, \mathbf{u} \; = \psi_0(\mathbf{x}, \mathbf{u}) + \sum_{j=1}^m \psi_j(\mathbf{x}, \mathbf{u}) g_j(\mathbf{x})$$

$$+ \psi_{m+1}(\mathbf{x}, \mathbf{u})(M - \|\mathbf{x}\|^2) + \psi_{m+2}(\mathbf{x}, \mathbf{u})(1 - \|\mathbf{u}\|^2);$$

$$\deg(\sigma_j, \psi_j) \; \leq \; 2\ell,$$

parametrized by ℓ.

$\rightarrow$ and optimal solution $p^*_\ell \in \mathbb{R}[\mathbf{x}]_d$.

Hence we replace **P** with the HIERARCHY of SEMIDEFINITE PROGRAMS

$$\rho_\ell = \inf_{p\in\mathbb{R}[\mathbf{x}]_d} \quad \sum_{\alpha\in\mathbb{N}^n_d} p_\alpha\,\gamma_\alpha$$

$$\text{s.t.} \quad f - p = \sigma_0(\mathbf{x}) + \sum_{j=1}^m \sigma_j(\mathbf{x}) g_j(\mathbf{x}) + \sigma_{m+1}(M - \|\mathbf{x}\|^2)$$

$$\mathbf{u}^T\nabla^2 p(\mathbf{x})\,\mathbf{u} \;=\psi_0(\mathbf{x},\mathbf{u}) + \sum_{j=1}^m \psi_j(\mathbf{x},\mathbf{u}) g_j(\mathbf{x})$$

$$+\psi_{m+1}(\mathbf{x},\mathbf{u})(M - \|\mathbf{x}\|^2) + \psi_{m+2}(\mathbf{x},\mathbf{u})(1 - \|\mathbf{u}\|^2);$$

$$\deg(\sigma_j,\psi_j) \;\leq\; 2\ell,$$

parametrized by ℓ.

$\to$ and optimal solution $p^*_\ell \in \mathbb{R}[\mathbf{x}]_d$.

Theorem (Lass & T. Phan Thanh (JOGO 2013))

$\rho_\ell \to \rho$ *and* $p^*_\ell \to p^* \in \mathbb{R}[\mathbf{x}]_d$, *as* $\ell \to \infty$

$\to$ Provides the best results in the comparison:

Guzman, Y. A; Hasan, M. M. F.; Floudas, C. A: *Computational Comparison of Convex Underestimators for Use in a Branch-and-Bound Global Optimization Framework*, Optimization in Science and Engineering; Springer, 2014; pp 229-246.

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:

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

- $M(\mathbf{K})$ 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)

III. Applications in 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:

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

- $M(\mathbf{K})$ 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)

III. Applications in 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:

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

- $M(\mathbf{K})$ 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)

Notice that writing

$$\mu = \nu + \varphi, \qquad \text{with}$$

- φ supported on S, and
- ν supported on $\mathbf{K} \setminus S$,

equivalently, one has to solve

$$\mathbf{P}: \quad \sup_{\nu,\varphi} \left\{ \int_S 1\, d\varphi \quad \Big| \quad \int_{\mathbf{K}\setminus S} X^\alpha\, d\nu + \int_S X^\alpha\, d\varphi = \gamma_\alpha, \quad \alpha \in \Gamma \right\}$$

(since we maximize $\varphi(S)$), one may take ν supported on $\mathbf{K}$.)

Assume that $\Gamma \subset \mathbb{N}^n_d$. Then the dual of **P** reads:

$$\mathbf{P}^*: \quad \inf_{p_\alpha} \left\{ \sum_{\alpha \in \Gamma} p_\alpha \, \gamma_\alpha : \quad p \geq 1 \text{ on } S; \quad p \geq 0 \text{ on } \mathbf{K} \right\}$$

where $p \in \mathbb{R}[\mathbf{x}]_d$ is a polynomial

$$\mathbf{x} \mapsto p(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n_d} p_\alpha \, \mathbf{x}^\alpha; \qquad p_\alpha = 0 \quad \forall \alpha \in \mathbb{N}^n_d \setminus \Gamma.$$

Let K and $S \subset K$ be compact semi-algebraic sets:

$$\begin{aligned} \mathbf{K} &= \{\mathbf{x} \in \mathbb{R}^n : \quad h_k(\mathbf{x}) \geq 0, \quad k = 1, \ldots, p\} \\ S &= \{\mathbf{x} \in \mathbb{R}^n : \quad g_j(\mathbf{x}) \geq 0, \quad j = 1, \ldots, m\} \end{aligned}$$

Assume that $\Gamma \subset \mathbb{N}^n_d$. Then the dual of **P** reads:

$$\mathbf{P}^* : \quad \inf_{p_\alpha} \left\{ \sum_{\alpha \in \Gamma} p_\alpha \, \gamma_\alpha : \quad p \geq 1 \text{ on } S; \quad p \geq 0 \text{ on } \mathbf{K} \right\}$$

where $p \in \mathbb{R}[\mathbf{x}]_d$ is a polynomial

$$\mathbf{x} \mapsto p(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n_d} p_\alpha \, \mathbf{x}^\alpha; \qquad p_\alpha = 0 \quad \forall \alpha \in \mathbb{N}^n_d \setminus \Gamma.$$

Let K and $S \subset K$ be compact semi-algebraic sets:

$$\begin{aligned} \mathbf{K} &= \{\mathbf{x} \in \mathbb{R}^n : \quad h_k(\mathbf{x}) \geq 0, \quad k = 1, \ldots, p\} \\ S &= \{\mathbf{x} \in \mathbb{R}^n : \quad g_j(\mathbf{x}) \geq 0, \quad j = 1, \ldots, m\} \end{aligned}$$

Again the basic idea of the moment-SOS approach is to replace the positivity constraints

$$p - 1 \geq 0 \text{ on } S; \quad p \geq 0 \text{ on } \mathbf{K}$$

with the SOS-positivity certificates

$$p(\mathbf{x}) - 1 = \sigma_0(\mathbf{x}) + \sum_{k=1}^{p} \sigma_k(\mathbf{x})\, h_k(\mathbf{x})$$

$$p(\mathbf{x}) = \psi_0(\mathbf{x}) + \sum_{j=1}^{m} \psi_j(\mathbf{x})\, g_j(\mathbf{x})$$

for some SOS polynomial $(\sigma_k) \subset \mathbb{R}[\mathbf{x}]_t$ and $(\psi_j) \subset \mathbb{R}[\mathbf{x}]_t$,

☞ and let t increase.

Again the basic idea of the moment-SOS approach is to replace the positivity constraints

$$p-1\geq 0 \text{ on } S;\quad p\geq 0 \text{ on } \mathbf{K}$$

with the SOS-positivity certificates

$$p(\mathbf{x})-1 \;=\; \sigma_0(\mathbf{x})+\sum_{k=1}^{p}\sigma_k(\mathbf{x})\,h_k(\mathbf{x})$$

$$p(\mathbf{x}) \;=\; \psi_0(\mathbf{x})+\sum_{j=1}^{m}\psi_j(\mathbf{x})\,g_j(\mathbf{x})$$

for some SOS polynomial $(\sigma_k)\subset\mathbb{R}[\mathbf{x}]_t$ and $(\psi_j)\subset\mathbb{R}[\mathbf{x}]_t$,

☞ and let t increase.

Again the basic idea of the moment-SOS approach is to replace the positivity constraints

$$p - 1 \geq 0 \text{ on } S; \quad p \geq 0 \text{ on } \mathbf{K}$$

with the SOS-positivity certificates

$$\begin{aligned} p(\mathbf{x}) - 1 &= \sigma_0(\mathbf{x}) + \sum_{k=1}^{p} \sigma_k(\mathbf{x})\, h_k(\mathbf{x}) \\ p(\mathbf{x}) &= \psi_0(\mathbf{x}) + \sum_{j=1}^{m} \psi_j(\mathbf{x})\, g_j(\mathbf{x}) \end{aligned}$$

for some SOS polynomial $(\sigma_k) \subset \mathbb{R}[\mathbf{x}]_t$ and $(\psi_j) \subset \mathbb{R}[\mathbf{x}]_t$,

☞ and let t increase.

☞ One ends up in solving the hierarchy of semidefinite programs indexed by $t \in \mathbb{N}$:

$$\rho_t = \min_{p \in \mathbb{R}[\mathbf{x}]_d} \sum_{\alpha \in \Gamma} p_\alpha \, \gamma_\alpha : \qquad \text{subject to:}$$

$$\begin{aligned} p_\alpha &= 0, \quad \forall \alpha \in \mathbb{N}^n_d \setminus \Gamma \\ p(\mathbf{x}) - 1 &= \sigma_0(\mathbf{x}) + \sum_{k=1}^{p} \sigma_k(\mathbf{x}) \, h_k(\mathbf{x}) \\ p(\mathbf{x}) &= \psi_0(\mathbf{x}) + \sum_{j=1}^{m} \psi_j(\mathbf{x}) \, g_j(\mathbf{x}) \end{aligned}$$

$$\sigma_k, \psi_j \quad SOS; \qquad \deg(\sigma_k \, h_k) \leq 2t; \quad \deg(\psi_j \, g_j) \leq 2t;$$

The volume of a basic-semi algebraic set

Let $S \subset \mathbb{R}^n$ be a compact basic semi-algebraic set. Let $\mathbf{K}$ be a BOX $[0, a]^n$ containing S and let:

$$\gamma_\alpha = \int_{\mathbf{K}} X^\alpha \, dx = \frac{a^{n+|\alpha|}}{\prod_{k=1}^n (1+\alpha_k)!}, \quad \forall \alpha \in \mathbb{N}^n$$

Theorem

The (Lebesgue) volume of the set S is obtained as:

$$\sup_{\nu, \varphi} \left\{ \int_S 1 \, d\varphi \quad | \quad \int_S X^\alpha \, d\varphi + \int_{\mathbf{K}} X^\alpha \, d\nu = \gamma_\alpha, \quad \alpha \in \mathbb{N}^n \right\}$$

The volume of a basic-semi algebraic set

Let $S \subset \mathbb{R}^n$ be a compact basic semi-algebraic set. Let $\mathbf{K}$ be a BOX $[0, a]^n$ containing S and let:

$$\gamma_\alpha = \int_{\mathbf{K}} X^\alpha \, dx = \frac{a^{n+|\alpha|}}{\prod_{k=1}^n (1+\alpha_k)!}, \quad \forall \alpha \in \mathbb{N}^n$$

Theorem

The (Lebesgue) volume of the set S is obtained as:

$$\sup_{\nu, \varphi} \left\{ \int_S 1 \, d\varphi \quad | \quad \int_S X^\alpha \, d\varphi + \int_{\mathbf{K}} X^\alpha \, d\nu = \gamma_\alpha, \quad \alpha \in \mathbb{N}^n \right\}$$

PROBLEM 4: Measures with given marginals

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:

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

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

PROBLEM 4: Measures with given marginals

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 $\pi_j : \mathbf{K} :\rightarrow \mathbf{K}_j$, $j = 1, \ldots, p$.

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:

$$f^* \, = \, \inf_{\mu \, \in M(\mathbf{K})} \, \{ \int_{\mathbf{K}} f \, d\mu \quad | \quad \pi_j \, \mu \, = \, \nu_j, \quad j = 1, \ldots, p \, \}$$

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 marginals

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

is equivalent to the countably many linear equalities

$$\int_{\mathbf{K}} X^\alpha \, d\mu \quad = \quad \int_{\mathbf{K}_j} X^\alpha \, d\nu_j, \quad \forall \alpha \in \mathbb{N}^{n_j}$$

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 $\mathbf{K}_j$.

Hence for each $d \in \mathbb{N}$, we may consider the truncated version

$$\rho_d = \inf_{\mu \in M(\mathbf{K})} \{ \int_{\mathbf{K}} f\, d\mu : \int_{\mathbf{K}} X^\alpha d\mu = \underbrace{\int_{\mathbf{K}_j} X^\alpha\, d\nu_j}_{=\gamma_{j\alpha} \text{ known}}, \quad \alpha \in \mathbb{N}^{n_j}_d; j = 1, \ldots, p\}$$

with dual

$$\rho_d^* = \sup_{p_j \in \mathbb{R}[X_j]_d} \{ \sum_{j=1}^{p} \underbrace{\int_{\mathbf{K}_j} p_j\, d\nu_j}_{= \sum_{\alpha \in \mathbb{N}^{n_j}_d} p_{j\alpha}\, \gamma_{j\alpha}} : f(X) - \sum_{j=1}^{p} p_j(X_j) \geq 0 \quad \text{on } \mathbf{K} \}$$

Hence for each $d \in \mathbb{N}$, we may consider the truncated version

$$\rho_d = \inf_{\mu \in M(\mathbf{K})} \{ \int_{\mathbf{K}} f \, d\mu : \int_{\mathbf{K}} X^\alpha d\mu = \underbrace{\int_{\mathbf{K}_j} X^\alpha \, d\nu_j}_{=\gamma_{j\alpha} \text{ known}}, \quad \alpha \in \mathbb{N}_d^{n_j}; j = 1, \ldots, p \}$$

with dual

$$\rho_d^* = \sup_{p_j \in \mathbb{R}[X_j]_d} \{ \sum_{j=1}^{p} \underbrace{\int_{\mathbf{K}_j} p_j \, d\nu_j}_{= \sum_{\alpha \in \mathbb{N}_d^{n_j}} p_{j\alpha} \gamma_{j\alpha}} \quad : f(X) - \sum_{j=1}^{p} p_j(X_j) \geq 0 \quad \text{on } \mathbf{K} \}$$

Then if $\mathbf{K} \subset \mathbb{R}^n$ is compact and in the form

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

one replaces the positivity-on-K constraint

$$f(X) - \sum_{j=1}^{p} p_j(X_j) \geq 0 \quad \text{on } \mathbf{K}$$

with the SOS-based Putinar's certificate

$$f(X) - \sum_{j=1}^{p} p_j(X_j) = \sigma_0(X) + \sum_{j=1}^{m} \sigma_j(X)\, g_j(X),$$

with SOS-weights $(\sigma_j) \subset \mathbb{R}[X]$.

... and solve the hierarchy of semidefinite programs

$$\theta_d = \sup_{p_j \in \mathbb{R}[X_j]_d} \left\{ \sum_{j=1}^{p} \sum_{\alpha \in \mathbb{N}_d^{n_j}} p_{j\alpha}\, \gamma_{j\alpha} : \right.$$

$$\text{s.t.} \quad f(X) - \sum_{j=1}^{p} p_j(X_j) = \sigma_0(X) + \sum_{j=1}^{m} \sigma_j(X)\, g_j(X) \left. \right\},$$

with SOS-weights (σ_j) of degree at most $2d$.

☞ One may prove that $\rho_d \to f^*$ as $d \to \infty$.

... and solve the hierarchy of semidefinite programs

$$\theta_d = \sup_{p_j \in \mathbb{R}[X_j]_d} \quad \{ \sum_{j=1}^{p} \sum_{\alpha \in \mathbb{N}^{n_j}_d} p_{j\alpha}\, \gamma_{j\alpha} :$$

$$\text{s.t.} \qquad f(X) - \sum_{j=1}^{p} p_j(X_j) \;=\; \sigma_0(X) + \sum_{j=1}^{m} \sigma_j(X)\, g_j(X) \,\},$$

with SOS-weights (σ_j) of degree at most $2d$.

☞ One may prove that $\rho_d \to f^*$ as $d \to \infty$.

IV. Gaussian measures of semi-algebraic sets

Let μ be the Gaussian measure on $\mathbb{R}^n$ with density $\mathbf{x} \mapsto \exp(-\|\mathbf{x}\|^2)$ and let $\mathbf{K} \subset \mathbb{R}^n$ be the non necessarily compact basic semi-algebraic set

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

Goal:

Approximate $\mu(\mathbf{K})$ as closely as desired

☞ Can be difficult even in small dimension $n = 2, 3$.

IV. Gaussian measures of semi-algebraic sets

Let μ be the Gaussian measure on $\mathbb{R}^n$ with density $\mathbf{x} \mapsto \exp(-\|\mathbf{x}\|^2)$ and let $\mathbf{K} \subset \mathbb{R}^n$ be the non necessarily compact basic semi-algebraic set

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

Goal:

Approximate $\mu(\mathbf{K})$ as closely as desired

☞ Can be difficult even in small dimension $n = 2, 3$.

Theorem (Lass 2015)

Let $f \in \mathbb{R}[\mathbf{x}]$ be strictly positive μ-a.e. on $\mathbf{K}$, and let $M(\mathbf{K})$ (resp. $M(\mathbb{R}^n)$) be the space of finite Borel measures on $\mathbf{K}$ (resp. $\mathbb{R}^n$). Then the optimization problem:

$$f_1^* = \sup_{\nu,\phi} \left\{ \int_{\mathbf{K}} f\, d\phi : \; \phi + \nu \;=\; \mu; \; \phi \in M(\mathbf{K}),\; \nu \in M(\mathbb{R}^n) \right\},$$

has a unique optimal solution $(\phi^*, \nu^*) = (\mu_{\mathbf{K}}, \mu - \mu_{\mathbf{K}})$ where $\mu_{\mathbf{K}}$ is the restriction of μ to $\mathbf{K}$, that is:

$$\phi^*(B) \;=\; \mu_{\mathbf{K}}(B) \;=\; \mu(\mathbf{K} \cap B), \qquad \forall B \in \mathcal{B}(\mathbb{R}^n).$$

In particular, $\phi^*(\mathbf{K}) = \mu(\mathbf{K})$, and $f^* = \mu(\mathbf{K})$ if $f = 1$.

Proof

From $\phi + \nu = \mu$ one deduces $\phi \leq \mu$ and therefore

$$f^* \leq \int_{\mathbf{K}} f d\mu = \int f\, d\mu_{\mathbf{K}}.$$

On the other hand the pair $(\phi^*, \nu^*) = (\mu_{\mathbf{K}}, \mu - \mu_{\mathbf{K}})$ is a feasible solution with associated cost

$$\int_{\mathbf{K}} f\, d\phi^* = \int f\, d\mu_{\mathbf{K}},$$

which proves the optimality of (ϕ^*, ν^*).

Uniqueness is more delicate. Assume there is another optimal solution ϕ, ν). From $\phi \leq \mu$ one deduces $\phi \ll \mu$ and so by Radon-Nykodim

$$\phi(B \cap \mathbf{K}) = \int_{B \cap \mathbf{K}} g\, d\mu \leq \int_{B \cap \mathbf{K}} d\mu, \quad \forall B \in \mathcal{B}(\mathbb{R}^n),$$

for some nonnegative measurable function g. Hence $g \leq 1$, μ-a.e. on $\mathbf{K}$.

Proof

From $\phi + \nu = \mu$ one deduces $\phi \leq \mu$ and therefore

$$f^* \leq \int_{\mathbf{K}} f d\mu = \int f \, d\mu_{\mathbf{K}}.$$

On the other hand the pair $(\phi^*, \nu^*) = (\mu_{\mathbf{K}}, \mu - \mu_{\mathbf{K}})$ is a feasible solution with associated cost

$$\int_{\mathbf{K}} f \, d\phi^* = \int f \, d\mu_{\mathbf{K}},$$

which proves the optimality of (ϕ^*, ν^*).

Uniqueness is more delicate. Assume there is another optimal solution ϕ, ν). From $\phi \leq \mu$ one deduces $\phi \ll \mu$ and so by Radon-Nykodim

$$\phi(B \cap \mathbf{K}) = \int_{B \cap \mathbf{K}} g \, d\mu \leq \int_{B \cap \mathbf{K}} d\mu, \quad \forall B \in \mathcal{B}(\mathbb{R}^n),$$

for some nonnegative measurable function g. Hence $g \leq 1$, μ-a.e. on $\mathbf{K}$.

On the other hand, by optimality of ϕ^* and ϕ,

$$f^* = \int_{\mathbf{K}} f d\mu = \int f \, d\phi^* = \int f \, d\phi$$

$$= \int_{\mathbf{K}} f \, g \, d\mu$$

which implies

$$0 = \int_{\mathbf{K}} f(1-g) \, d\mu,$$

Combining this with $f > 0$ and $g \leq 1$ μ-a.e. on $\mathbf{K}$, yields $g = 1$, μ-a.e. on $\mathbf{K}$.

☞ This yields the desired result that $\phi = \phi^*$. □

A dual view

A possible dual for the above LP is the LP:

$$\rho^* = \inf_{p \in \mathbb{R}[\mathbf{x}]} \{ \int_{\mathbf{K}} p \, d\mu : \ p \geq f \text{ on } \mathbf{K}; \ p \geq 0 \text{ on } \mathbb{R}^n \},$$

Indeed it trivially holds that $\rho^* \geq f^*$.

A tractable version is obtained by replacing:

- the "hard" positivity constraint $p - f \geq 0$ on $\mathbf{K}$, with the positivity-on-$\mathbf{K}$ certificate

$$p - f = \sigma_0 + \sum_{j=1}^{m} \sigma_j \, g_j; \quad \sigma_j \text{ is SOS for all } j \}.$$

- the "hard" positivity constraint $p \geq 0$ on $\mathbb{R}^n$ with p is SOS.

A dual view

A possible dual for the above LP is the LP:

$$\rho^* = \inf_{p\in\mathbb{R}[\mathbf{x}]} \{\int_{\mathbf{K}} p\,d\mu : \; p \geq f \text{ on } \mathbf{K};\; p \geq 0 \text{ on } \mathbb{R}^n\},$$

Indeed it trivially holds that $\rho^* \geq f^*$.

A tractable version is obtained by replacing:

- the "hard" positivity constraint $p - f \geq 0$ on $\mathbf{K}$, with the positivity-on-$\mathbf{K}$ certificate

$$p - f = \sigma_0 + \sum_{j=1}^{m} \sigma_j\, g_j; \quad \sigma_j \text{ is SOS for all } j\}.$$

- the "hard" positivity constraint $p \geq 0$ on $\mathbb{R}^n$ with p is SOS.

so as to obtain the hierarchy of semidefinite approximations

indexed by $d \in \mathbb{N}$:

$$\rho_d^* = \inf_{p \in \mathbb{R}[\mathbf{x}]_d} \left\{ \int_{\mathbb{R}^n} p \, d\mu : \; p - f \; = \; \sigma_0 + \sum_{j=1}^{m} \sigma_j \, g_j; \quad p, \sigma_j \text{ all SOS } \right\}$$

where the degree of the SOS p, σ_j is bounded by $2d$.

Theorem (Lass 2015)

For every $d \in \mathbb{N}$, $\rho_d^* \geq f^*$ and $\rho_d^* \to f^*$ as $d \to \infty$.

so as to obtain the hierarchy of semidefinite approximations

indexed by $d \in \mathbb{N}$:

$$\rho_d^* = \inf_{p \in \mathbb{R}[\mathbf{x}]_d} \left\{ \int_{\mathbb{R}^n} p \, d\mu : \; p - f \; = \; \sigma_0 + \sum_{j=1}^m \sigma_j \, g_j ; \quad p, \sigma_j \text{ all SOS } \right\}$$

where the degree of the SOS p, σ_j is bounded by $2d$.

Theorem (Lass 2015)

For every $d \in \mathbb{N}$, $\rho_d^* \geq f^*$ and $\rho_d^* \to f^*$ as $d \to \infty$.

One may do the same for the complement $\mathbf{K}^c := \mathbb{R}^n \setminus \mathbf{K}$ as soon as one can write

$$\mathbf{K}^c = \bigcup_{i=1}^{p} \Omega_i; \quad \mu(\Omega_i \cap \Omega_j) = 0 \quad \forall (i,j)$$

so that $\mu(\mathbf{K}^c) = \sum_{i=1}^{p} \mu(\Omega_i)$. In doing so one obtains for each $i = 1, \ldots, p$ a sequence $(\theta_{id})_{d \in \mathbb{N}}$ such that

$$\sum_{i=1}^{p} \theta_{id} \geq \mu(\mathbf{K}^c) \quad \text{and} \quad \lim_{d\to\infty} \sum_{i=1}^{p} \theta_{id} = \mu(\mathbf{K}^c) = \mu(\mathbb{R}^n) - \mu(\mathbf{K}).$$

Theorem (Lass 2015)

With $f = 1$ one obtains $\underbrace{\mu(\mathbb{R}^n) - \sum_{j=1}^{p} \theta_{id}}_{\omega_d^*} \leq \mu(\mathbf{K}) \leq \rho_d^*$ *for all d,*

and

$$\lim_{d\to\infty} \omega_d^* = \mu(\mathbf{K}) = \lim_{d\to\infty} \rho_d^*.$$

One may do the same for the complement $\mathbf{K}^c := \mathbb{R}^n \setminus \mathbf{K}$ as soon as one can write

$$\mathbf{K}^c = \bigcup_{i=1}^{p} \Omega_i; \quad \mu(\Omega_i \cap \Omega_j) = 0 \quad \forall (i,j)$$

so that $\mu(\mathbf{K}^c) = \sum_{i=1}^{p} \mu(\Omega_i)$. In doing so one obtains for each $i = 1, \ldots, p$ a sequence $(\theta_{id})_{d \in \mathbb{N}}$ such that

$$\sum_{i=1}^{p} \theta_{id} \geq \mu(\mathbf{K}^c) \quad \text{and} \quad \lim_{d \to \infty} \sum_{i=1}^{p} \theta_{id} = \mu(\mathbf{K}^c) = \mu(\mathbb{R}^n) - \mu(\mathbf{K}).$$

Theorem (Lass 2015)

With $f = 1$ *one obtains* $\underbrace{\mu(\mathbb{R}^n) - \sum_{j=1}^{p} \theta_{id}}_{\omega_d^*} \leq \mu(\mathbf{K}) \leq \rho_d^*$ *for all* d*,*

and

$$\lim_{d \to \infty} \omega_d^* = \mu(\mathbf{K}) = \lim_{d \to \infty} \rho_d^*.$$

Examples

Let $n = 2$, and $d\mu = \exp(-\|\mathbf{x}\|^2/\sigma)\, d\mathbf{x}$ and let $\mathbf{K}$ be the non-convex quadratic

$$\mathbf{x} \mapsto \mathbf{x}^T \mathbf{A}\mathbf{x} = 0.56\, x_1^2 + 0.96\, x_1 x_2 - 1.24\, x_2^2.$$

$$\mathbf{K} = \{(x, y) : (\mathbf{x} - \mathbf{u})^T \mathbf{A}\, (\mathbf{x} - \mathbf{u}) \leq 1\} \quad \text{(non-compact)},$$

$$\text{with} \quad \mathbf{u} \;=\; (0.1, 0.5) \text{ and } (0.5, 0.1).$$

	$\mathbf{u} = (0.5, 0.1)$		
σ	ρ_9^*	ω_9^*	$100\,(\rho^*9 - \omega_9^*)/\omega_9^*$
1	2.829605	2.824718	0.17%
0.8	1.876731	1.876609	0.006%
	$\mathbf{u} = (0.1, 0.5)$		
σ	ρ_9^*	ω_9^*	$100\,(\rho_9^* - \omega_9^*)/\omega_9^*$
1	2.989832	2.986599	0.10%
0.8	1.969188	1.969103	0.004%

Examples

Let $n = 2$, and $d\mu = \exp(-\|\mathbf{x}\|^2/\sigma)\, d\mathbf{x}$ and let $\mathbf{K}$ be the non-convex quadratic

$$\mathbf{x} \mapsto \mathbf{x}^T\mathbf{A}\mathbf{x} = 0.56\, x_1^2 + 0.96\, x_1 x_2 - 1.24\, x_2^2.$$

$$\mathbf{K} = \{(x,y) : (\mathbf{x}-\mathbf{u})^T\mathbf{A}\,(\mathbf{x}-\mathbf{u}) \leq 1\} \quad \text{(non-compact)},$$

$$\text{with} \quad \mathbf{u} \; = \; (0.1, 0.5) \text{ and } (0.5, 0.1).$$

	$\mathbf{u} = (0.5, 0.1)$		
σ	ρ_9^*	ω_9^*	$100\,(\rho^*9 - \omega_9^*)/\omega_9^*$
1	2.829605	2.824718	0.17%
0.8	1.876731	1.876609	0.006%
	$\mathbf{u} = (0.1, 0.5)$		
σ	ρ_9^*	ω_9^*	$100\,(\rho_9^* - \omega_9^*)/\omega_9^*$
1	2.989832	2.986599	0.10%
0.8	1.969188	1.969103	0.004%

More details and (non-compact) examples in `arXiv:1508.06132`.

Conclusion

- Provides a sequence of converging upper and lower bounds on $\mu(\mathbf{K})$ for non necessarily compact basic semi-algebraic sets $\mathbf{K}$.
- A general methodology not set-$\mathbf{K}$-dependent.
- Also works for the exponential measure on the positive orthant $\mathbb{R}^n_+$, and in fact any measure μ provided that it satisfies Carleman's condition and one knows all its moments.

... but of course ...

With rough basic implementation and present state-of-the-art SDP solvers, one can obtain a few upper and lower bounds only and for dimension $n = 2$ or $n = 3$. For $d \geq 15$ numerical accuracy problems show up.

- ☞ Some non-trivial tricks (based on Stokes' formula) permit to improve the quality of bounds.
- ☞ Much remains to be done for a better implementation

... but of course ...

With rough basic implementation and present state-of-the-art SDP solvers, one can obtain a few upper and lower bounds only and for dimension $n = 2$ or $n = 3$. For $d \geq 15$ numerical accuracy problems show up.

- ☞ Some non-trivial tricks (based on Stokes' formula) permit to improve the quality of bounds.
- ☞ Much remains to be done for a better implementation

... but of course ...

With rough basic implementation and present state-of-the-art SDP solvers, one can obtain a few upper and lower bounds only and for dimension $n = 2$ or $n = 3$. For $d \geq 15$ numerical accuracy problems show up.

- ☞ Some non-trivial tricks (based on Stokes' formula) permit to improve the quality of bounds.
- ☞ Much remains to be done for a better implementation

V. Lebesgue decomposition in action

Given two measures μ and ν on $\mathbb{R}^n$,

one would like to approximate the Lebesgue decomposition

$$\phi + \psi \;=\; \mu; \quad \phi \ll \nu; \quad \psi \perp \nu,$$

of μ with respect to ν.

... based on the sole knowledge of the moments

$$y_\alpha \;=\; \int_{\mathbb{R}^n} \mathbf{x}^\alpha \, d\mu, \quad z_\alpha \;=\; \int_{\mathbb{R}^n} \mathbf{x}^\alpha \, d\nu, \quad \alpha \in \mathbb{N}^n.$$

of μ and ν.

V. Lebesgue decomposition in action

Given two measures μ and ν on $\mathbb{R}^n$,

one would like to approximate the Lebesgue decomposition

$$\phi + \psi \,=\, \mu; \quad \phi \ll \nu; \quad \psi \perp \nu,$$

of μ with respect to ν.

... based on the sole knowledge of the moments

$$y_\alpha \,=\, \int_{\mathbb{R}^n} \mathbf{x}^\alpha \, d\mu, \quad z_\alpha \,=\, \int_{\mathbb{R}^n} \mathbf{x}^\alpha \, d\nu, \quad \alpha \in \mathbb{N}^n.$$

of μ and ν.

By definition of ϕ and ψ:

ϕ has a DENSITY w.r.t. ν in $L_1(\nu)$ (called the Radon-Nikodym derivative of μ w.r.t. ν). That is, there exists a nonnegative measurable function $f \in L_1(\nu)$ such that:

$$\phi(A) = \int_A f(\mathbf{x})\, d\nu(\mathbf{x}), \qquad \forall A \in \mathcal{B}(\mathbb{R}^n).$$

CLAIM: If one assumes that :

- f is in $L_\infty(\nu)$ (instead of $L_1(\nu)$), and $\|f\|_\infty < M$ for some M,
- Both moment sequences (y_α) and (z_α), $\alpha \in \mathbb{N}^n$ satisfy Carleman's condition:

$$+\infty = \sum_{k=1}^{\infty} \left(\int X_i^{2k} \, d\mu \right)^{-1/2k} = \sum_{k=1}^{\infty} \left(\int X_i^{2k} \, d\nu \right)^{-1/2k}$$

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

THEN ... one may approximate as closely as desired any fixed set of moments of ϕ and ψ.

CLAIM: If one assumes that :

- f is in $L_\infty(\nu)$ (instead of $L_1(\nu)$), and $\|f\|_\infty < M$ for some M,
- Both moment sequences (y_α) and (z_α), $\alpha \in \mathbb{N}^n$ satisfy Carleman's condition:

$$+\infty = \sum_{k=1}^{\infty} \left(\int X_i^{2k} \, d\mu \right)^{-1/2k} = \sum_{k=1}^{\infty} \left(\int X_i^{2k} \, d\nu \right)^{-1/2k}$$

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

THEN ... one may approximate as closely as desired any fixed set of moments of ϕ and ψ.

CLAIM: If one assumes that :

- f is in $L_\infty(\nu)$ (instead of $L_1(\nu)$), and $\|f\|_\infty < M$ for some M,
- Both moment sequences (y_α) and (z_α), $\alpha \in \mathbb{N}^n$ satisfy Carleman's condition:

$$+\infty = \sum_{k=1}^{\infty} \left(\int X_i^{2k}\, d\mu \right)^{-1/2k} = \sum_{k=1}^{\infty} \left(\int X_i^{2k}\, d\nu \right)^{-1/2k}$$

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

THEN ... one may approximate as closely as desired any fixed set of moments of ϕ and ψ.

A hierarchy of semidefinite approximations

Denote the moments of μ and ν by:

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

Let $\gamma > 0$ be fixed, and consider the hierarchy of semidefinite programs $\mathbf{P}_d$ indexed by $d \in \mathbb{N}$:

$$\begin{array}{rrl} \mathbf{P}_d : \quad \rho_d = & \displaystyle\sup_{y,u,v} & y_0 \\ & \text{s.t.} & y_\alpha + u_\alpha = \mu_\alpha, \quad \forall \alpha \in \mathbb{N}^n_d \\ & & y_\alpha + v_\alpha = \gamma\, \nu_\alpha, \quad \forall \alpha \in \mathbb{N}^n_d \\ & & \\ & & \mathbf{M}_d(y), \mathbf{M}_d(u), \mathbf{M}_d(v) \succeq 0 \end{array}$$

Let ϕ^* and ψ^* be the Lebesgue decomposition of μ w.r.t. ν, and let $f^* \in L_1(\nu)$ be the density of ϕ^* w.r.t. ν.

Theorem (Lass 2015)

(i) For each $d \in \mathbb{N}$, the semidefinite program has an optimal solution (y^d, u^d, v^d).

(ii) Moreover as $d \to \infty$, the triplet of sequences (y^d, u^d, v^d) converges to some triplet of sequences (y^*, u^*, v^*) where

$$y^*_\alpha = \int \mathbf{x}^\alpha \,(\gamma \wedge f^*)\, d\nu = \int \mathbf{x}^\alpha\, f^*_\gamma\, d\nu, \quad \forall \alpha \in \mathbb{N}^n.$$

with $\|f^*_\gamma\|_\infty \leq \gamma$.

(iii) So if $f^* \in L_\infty(\nu)$ with $\|f^*\|_\infty \leq \gamma$, then

$$y^*_\alpha = \int \mathbf{x}^\alpha\, d\phi, \quad \forall \alpha \in \mathbb{N}^n.$$

Let ϕ^* and ψ^* be the Lebesgue decomposition of μ w.r.t. ν, and let $f^* \in L_1(\nu)$ be the density of ϕ^* w.r.t. ν.

Theorem (Lass 2015)

(i) For each $d \in \mathbb{N}$, the semidefinite program has an optimal solution (y^d, u^d, v^d).

(ii) Moreover as $d \to \infty$, the triplet of sequences (y^d, u^d, v^d) converges to some triplet of sequences (y^*, u^*, v^*) where

$$y^*_\alpha = \int \mathbf{x}^\alpha \, (\gamma \wedge f^*) \, d\nu = \int \mathbf{x}^\alpha \, f^*_\gamma \, d\nu, \quad \forall \alpha \in \mathbb{N}^n.$$

with $\|f^*_\gamma\|_\infty \leq \gamma$.

(iii) So if $f^* \in L_\infty(\nu)$ with $\|f^*\|_\infty \leq \gamma$, then

$$y^*_\alpha = \int \mathbf{x}^\alpha \, d\phi, \quad \forall \alpha \in \mathbb{N}^n.$$

Let ϕ^* and ψ^* be the Lebesgue decomposition of μ w.r.t. ν, and let $f^* \in L_1(\nu)$ be the density of ϕ^* w.r.t. ν.

Theorem (Lass 2015)

(i) For each $d \in \mathbb{N}$, the semidefinite program has an optimal solution (y^d, u^d, v^d).

(ii) Moreover as $d \to \infty$, the triplet of sequences (y^d, u^d, v^d) converges to some triplet of sequences (y^*, u^*, v^*) where

$$y^*_\alpha = \int \mathbf{x}^\alpha\,(\gamma \wedge f^*)\,d\nu = \int \mathbf{x}^\alpha\, f^*_\gamma\,d\nu, \quad \forall \alpha \in \mathbb{N}^n.$$

with $\|f^*_\gamma\|_\infty \leq \gamma$.

(iii) So if $f^* \in L_\infty(\nu)$ with $\|f^*\|_\infty \leq \gamma$, then

$$y^*_\alpha = \int \mathbf{x}^\alpha\, d\phi, \quad \forall \alpha \in \mathbb{N}^n.$$

Let ϕ^* and ψ^* be the Lebesgue decomposition of μ w.r.t. ν, and let $f^* \in L_1(\nu)$ be the density of ϕ^* w.r.t. ν.

Theorem (Lass 2015)

(i) For each $d \in \mathbb{N}$, the semidefinite program has an optimal solution (y^d, u^d, v^d).

(ii) Moreover as $d \to \infty$, the triplet of sequences (y^d, u^d, v^d) converges to some triplet of sequences (y^*, u^*, v^*) where

$$y^*_\alpha = \int \mathbf{x}^\alpha \, (\gamma \wedge f^*) \, d\nu = \int \mathbf{x}^\alpha \, f^*_\gamma \, d\nu, \quad \forall \alpha \in \mathbb{N}^n.$$

with $\|f^*_\gamma\|_\infty \leq \gamma$.

(iii) So if $f^* \in L_\infty(\nu)$ with $\|f^*\|_\infty \leq \gamma$, then

$$y^*_\alpha = \int \mathbf{x}^\alpha \, d\phi, \quad \forall \alpha \in \mathbb{N}^n.$$

Examples

Let $n = 2$, $p \in (0, 1)$ and

- ν is the Gaussian with density $\mathbf{x} \mapsto \exp(-\|\mathbf{x}\|^2)$,
- θ is the measure uniformy distributed on the circle $\{\mathbf{x} : x_1^2 + x_2^2 = 1\}$.

Define the measure μ to be

$$\mu \;=\; p\,\nu \,+ (1-p)\,\theta,$$

so that the Lebesgue decomposition of μ w.r.t. ν is $(\phi, \psi) = (p\,\nu, (1-p)\,\theta)$.

The table below show relative error between the approximate moments $\mathbf{u} = (u_\alpha)$ of degree 2 and 4, of the singular part ψ and those of $p\,\theta$ computed with moments up to order $2d = 14$.

approx. moments	x_1^2	x_1^4	$x_1^2 x_2^2$	$L_{\mathbf{u}^d}((x_1^2 + x_2^2 - 1)^2)$
p=0.1	0.19%	0.52%	0.53%	0.001
p=0.2	3.7%	8.12%	12.14%	0.16

Same thing but now with ν being uniformly supported on the unit box $[-1, 1]^n$.

approx. moments	x_1^2	x_1^4	$x_1^2 x_2^2$	$L_{\mathbf{u}^d}((x_1^2 + x_2^2 - 1)^2)$
p=0.1	0.26%	0.93%	0.61%	0.0001
p=0.2	8.8%	10.2%	7.5%	0.08

The table below show relative error between the approximate moments $\mathbf{u} = (u_\alpha)$ of degree 2 and 4, of the singular part ψ and those of $p\,\theta$ computed with moments up to order $2d = 14$.

approx. moments	x_1^2	x_1^4	$x_1^2x_2^2$	$L_{\mathbf{u}^d}((x_1^2+x_2^2-1)^2)$
p=0.1	0.19%	0.52%	0.53%	0.001
p=0.2	3.7%	8.12%	12.14%	0.16

Same thing but now with ν being uniformly supported on the unit box $[-1,1]^n$.

approx. moments	x_1^2	x_1^4	$x_1^2x_2^2$	$L_{\mathbf{u}^d}((x_1^2+x_2^2-1)^2)$
p=0.1	0.26%	0.93%	0.61%	0.0001
p=0.2	8.8%	10.2%	7.5%	0.08