Some applications of moments and SDP-relaxations

J.B. Lasserre

LAAS-CNRS,

Semidefinite Programming and Robust Optimization

IMA, Minneapolis, March 2003

- global optimization (with polynomials)
- systems of polynomial equations
- bounds on measures with moment conditions
- performance evaluation

GLOBAL OPTIMIZATION

Consider the optimization problem

$$\mathbf{P} \mapsto \mathbf{p}^* := \min\{\mathbf{g_0(x)} \,|\, \mathbf{g_i(x)} \geq \mathbf{0},\ \mathbf{i} = \mathbf{1}, \ldots \mathbf{m}\},$$

where $\mathbf{g_i(x)} : \mathbf{R^n} \to \mathbf{R}$ are all **real-valued polynomials**. Let

$$\mathbf{K} := \{\mathbf{x} \in \mathbf{R^n} \,|\, \mathbf{g_i(x)} \geq \mathbf{0},\ \mathbf{i} = \mathbf{1}, \ldots \mathbf{m}\},$$

be the feasible set, and let

$$1, x_1, x_2, \ldots x_n, x_1^2, x_1 x_2, \ldots, x_n^r,$$

be a basis (of dimension $s(r) = \binom{n+r}{n}$) of the vector space of real-valued polynomials of degree at most r, and write

$$p(x) = \sum_{|\alpha| \leq r} p_\alpha \, x^\alpha = \sum_{|\alpha| \leq r} p_\alpha \left[x_1^{\alpha_1} \ldots x_n^{\alpha_n} \right],$$

with $|\alpha| = \sum_{i=1}^n \alpha_i$, and $p = \{p_\alpha\} \in \mathbf{R}^{s(r)}$ its vector of coefficients.

Two dual points of view

I. $\mathbf{p}^*$ **global minimum** $\Leftrightarrow \mathbf{g_0(x) - p^* \geq 0 \quad \forall x \in K}$**, i.e.,**

$g_0(x) - p^*$ **is a nonnegative polynomial on** K.

$\Rightarrow$ Characterize these polynomials ....

$\longrightarrow$ (real) algebraic geometry (Representation results by Schmüdgen, Putinar, Prestel and Jacobi, etc ...)

SDP representation : Shor, Nesterov, Lasserre, Parrilo, ...

II. But we also have

$$(*) \quad \mathbf{p}^* = \min_\mu \{ \int \mathbf{g_0}(\mathbf{x})\, \mu(\mathbf{dx}) \,|\, \mu \in \mathcal{P}(\mathbf{K}) \},$$

where $\mathcal{P}(\mathbf{K})$ is the space of probability measures with support contained in $\mathbf{K}$. Indeed,

$$\int g_0(x)\, \mu(dx) \,\geq p^*, \, \forall \mu \in \mathcal{P}(\mathbf{K}),$$

and with $\mu := \delta_{x^*}$ at a global minimizer x^*,

$$\int g_0(x)\, \delta_{x^*}(dx) \,=\, p^*.$$

Observe that both I and II are valid for global optima only

$$\min_\mu \{ \langle g_0, \mu \rangle \,|\, \langle 1\!\mathrm{l}_{\mathbf{K}}, \mu \rangle \,=\, 1;\ \langle 1\!\mathrm{l}_{\mathbf{K}^c}, \mu \rangle \,=\, 0;\ \mu \,\geq\, 0 \}.$$

$\Rightarrow$ characterize these measures μ ..... (functional analysis)

The dual linear program of (*) is

$$\max_{\gamma,\lambda}\{\gamma \quad | \quad \lambda 1\!\!\mathrm{l}_{\mathbf{K}^c} + \gamma 1\!\!\mathrm{l}_{\mathbf{K}} \leq g_0(x), \quad \forall x \in \mathbf{R}^n\},$$

or, equivalently, $\max_{\gamma}\{\gamma \quad | \quad g_0(x) - \gamma \geq 0 \quad \text{on } \mathbf{K}\}$.

I. The point of view of moments

$$\int g_0(x)\,\mu(dx) \,=\, \sum_\alpha (g_0)_\alpha \int x^\alpha\,\mu(dx) \,=\, \sum_\alpha (g_0)_\alpha y_\alpha,$$

with $y_\alpha \rightarrow$ α-moment of μ. Hence, (*) reads

$$\left\{\begin{array}{l} \min_y \sum_\alpha (g_0)_\alpha y_\alpha \\ \\ y_\alpha \,=\, \int x^\alpha\,\mu(dx) \quad \forall\alpha \quad \textbf{for some probability } \mu \text{ on } \mathbf{K} \end{array}\right.$$

Hence, **translate the condition**

there is some probability μ on $\mathbf{K}$ such that

$$y_\alpha \,=\, \int x^\alpha\,\mu(dx), \qquad \forall \alpha \leq \alpha_0,$$

into a condition on $\mathbf{y}$, to obtain a finite-dimensional optimization pb on y. This is the **$\mathbf{K}$-moment problem**, which dates back to Hausdorff, Markov, Stieltjes, Hamburger, etc ...

II. The point of view of positive polynomials

=Hilbert's 17th problem on the representation of positive polynomials. In the one dimensional case,

$$p(x) \geq 0 \Leftrightarrow p(x) = \sum_{k=1}^{s} q_k(x)^2.$$

Not true anymore in R^n ....

Representation of polynomials, positive on $\mathbf{K} := \{\mathbf{x} \in \mathbf{R}^\mathbf{n} \,|\, \mathbf{g_k(x)} \geq \mathbf{0},\ \mathbf{k} = \mathbf{1}, \ldots \mathbf{m}\}$

Theorem [Schmüdgen, Putinar, Jacobi, Prestel]

Assume there is a polynomial $u(x) : \mathbf{R}^n \to \mathbf{R}$ *such that*

$$u(x) = q(x) + \sum_{k=1}^{m} g_k(x) v(x),$$

for some polynomials $q(x), v(x)$*, all sums of squares, and such that* $\{u(x) \geq 0\}$ *is compact. Then:*

Every polynomial $p(x) > 0$ *on* $\mathbf{K}$ *has the representation:*

$$(*) \qquad \mathbf{p(x)} = \sum_{\mathbf{j=1}}^{\mathbf{r_0}} \mathbf{q_j(x)^2} + \sum_{\mathbf{k=1}}^{\mathbf{m}} \mathbf{g_k(x)} \, [\sum_{\mathbf{j=1}}^{\mathbf{r_k}} \mathbf{t_{kj}(x)^2}],$$

for some (finite) family of polynomials $\{q_j(x)\}$, $\{t_{kj}(x)\}$.

For instance, the representation (*) holds whenever

- $\{g_k(x) \geq 0\}$ is compact for some k,

- when all the $g_k(x)$ are linear and $\mathbf{K}$ is compact,

In practice, one may also add the redundant constraint $M - \sum_i x_i^2 \geq 0$ for M large enough.

This is also the case when one has integrality constraints $x_i^2 = x$ for all i. $\Rightarrow$ **very general result!**

SDP-relaxations

Moment matrix. With $\alpha \in \mathbf{N}^n$, and $\mathbf{y}_{\alpha_1,\ldots,\alpha_n} \rightsquigarrow \int \mathbf{x}_1^{\alpha_1} \cdots \mathbf{x}_n^{\alpha_n} \, \mathbf{d}\mu$

$$M_2(y) = \left[\begin{array}{c|cc|ccc} 1 & y_{1,0} & y_{0,1} & y_{2,0} & y_{1,1} & y_{0,2} \\ \hline y_{1,0} & y_{2,0} & y_{1,1} & y_{3,0} & y_{2,1} & y_{1,2} \\ y_{0,1} & y_{1,1} & y_{0,2} & y_{2,1} & y_{1,2} & y_{0,3} \\ \hline y_{2,0} & y_{3,0} & y_{2,1} & y_{4,0} & y_{3,1} & y_{2,2} \\ y_{1,1} & y_{2,1} & y_{1,2} & y_{3,1} & y_{2,2} & y_{1,3} \\ y_{0,2} & y_{1,2} & y_{0,3} & y_{2,2} & y_{1,3} & y_{0,4} \end{array} \right]$$

In general, if $M_r(y)(i,1) = y_\alpha$ and $M_r(y)(1,j) = y_\beta$ then

$$M_r(y)(i,j) = y_{\alpha+\beta} = y_{\alpha_1+\beta_1,\ldots,\alpha_n+\beta_n}$$

Localizing matrix.

Given a polynomial $\theta : \mathbf{R}^n \to \mathbf{R}$ of degree w, with coefficient vector $\theta \in \mathbf{R}^{s(w)}$, let $M_r(\theta y)$ be the **localizing matrix**

$$M_r(\theta y)(i,j) \ := \sum_\alpha \theta_\alpha y_{\{\alpha(i,j)+\alpha\}}.$$

For instance, with $x \mapsto \theta(x) = 1 - x_1^2 - x_2^2$, $M_2(\theta y) =$

$$\left[\begin{array}{ccc} 1 - y_{20} - y_{02}, & y_{10} - y_{30} - y_{12}, & y_{01} - y_{21} - y_{03} \\ y_{10} - y_{30} - y_{12}, & y_{20} - y_{40} - y_{22}, & y_{11} - y_{21} - y_{12} \\ y_{01} - y_{21} - y_{03}, & y_{11} - y_{21} - y_{12}, & y_{02} - y_{22} - y_{04} \end{array}\right].$$

If $M_r(y)(i,j) = y_\beta$ then $\mathbf{M_r}(\theta\mathbf{y})(\mathbf{i},\mathbf{j}) = \sum_\alpha \theta_\alpha \mathbf{y}_{\beta+\alpha}$ that is,

$$M_r(\theta y)(i,j) \rightsquigarrow \int x^\beta \, \theta(x) \, \mu(dx)$$

If $(1, y)$ is the vector of moments up to order $2r$ of some probability measure μ on the Borel sets of $\mathbf{R}^n$, then for every polynomial $q(x) : \mathbf{R}^n \rightarrow \mathbf{R}$ of degree at most r,

$$\langle q, M_r(y)q\rangle = \int q(x)^2 \, \mu(dx) \geq 0,$$

so that $M_r(y) \succeq 0$. Similarly,

$$\langle q, M_r(\theta y)q\rangle = \int \theta(x)q(x)^2 \, \mu(dx) \geq 0,$$

and thus $M_r(\theta y) \succeq 0$ whenever μ is supported on $\{\theta(x) \geq 0\}$.

The $\mathbf{K}$**-moment problem** identifies those vectors $\mathbf{y}$ with $M_r(y) \succeq 0$ that are **moments of a measure** μ **with support contained in** $\mathbf{K}$.

Dual theory in algebraic geometry = representation of polynomials, positive on a semi-algebraic set $\mathbf{K}$

Introduce the family $\{\mathbf{Q}_i\}$ of SDP-relaxations

$$\mathbf{Q}_i \begin{cases} \min_y \sum_\alpha (g_0)_\alpha y_\alpha & \\ M_i(y) & \succeq 0 \\ M_{i-v_k}(g_k y) & \succeq 0, \quad k = 1, \ldots m. \end{cases}$$

and the family $\{\mathbf{Q}_i^*\}$ of their dual

$$\mathbf{Q}_i^* \begin{cases} \max_{X, Z_1, \ldots Z_m \succeq 0} -X(1,1) - \sum_{k=1}^m g_k(0) Z_k(1,1) \\ \text{s.t.} \quad \langle X, B_\alpha \rangle + \sum_{k=1}^m \langle Z_k, C_\alpha^k \rangle = (g_0)_\alpha, \ \forall \alpha. \end{cases}$$

where we write

$$M_i(y) = \sum_\alpha y_\alpha B_\alpha; \quad M_{i-v_k}(g_k y) = \sum_\alpha y_\alpha C_\alpha^k, \ k = 1, \ldots m$$

SDP-RELAXATIONS' INTERPRETATION :

From the primal : $y_\alpha \to \int x^\alpha d\mu;\ \mu(\mathbf{K}) = 1$

Primal SDP-relaxation = K-moment problem.

From the dual : one has $\max \mathbf{Q}_i^* \le p^*$ and

$$p(x) - \max \mathbf{Q}_i^* = \sum_{j=1}^{s} q_j(x)^2 + \sum_{k=1}^{m} g_k(x) \left[\sum_{l=1}^{s_k} q_{kl}(x)^2 \right],$$

with degree$(q_i) \le i$ and degree$(q_{kl}) \le i - v_k$.

Dual SDP-relaxation $\longleftrightarrow$ Schmüdgen-Putinar representation of $\mathbf{p(x) - p^*}$

Theorem 1. *Assume that there exists a polynomial* $u(x) : \mathbf{R}^n \rightarrow \mathbf{R}$ *with Putinar's representation*

$$u(x) = q_0(x) + \sum_{j=1}^{m} g_i(x) q_i(x), \tag{1}$$

with $\{q_j\}$ *sums of squares, and with* $\{u(x) \geq 0\}$ *compact. Then:*

$$\inf \mathbf{Q_i} \uparrow \mathbf{p}^* = \min \mathbf{P}.$$

In addition, if $p(x) - p^*$ *has Putinar's representation (1) for polynomials* $\{q_j\}$ *of degree at most* $2N$*, then*

$$\mathbf{p}^* = \min \mathbf{Q_i}, \quad \forall \mathbf{i} \geq \mathbf{N},$$

and for every optimal solution x^* *of* $\mathbf{P}$*, the vector*

$$\mathbf{y}^* := (\mathbf{x_1^*}, \ldots, \mathbf{x_n^*}, \ldots (\mathbf{x_1^*})^{\mathbf{2i}}, \ldots (\mathbf{x_n^*})^{\mathbf{2i}})$$

is an optimal solution of $\mathbf{Q}_i$.

Software **GLOPTIPOLY** for polynomial programming:
D. Henrion and J.B. Lasserre, ACM Trans. Math. Soft.
(to appear). Free access at http://www.laas.fr/~henrion

Continuous optimization problems:
CPU times in seconds and LMI relaxation orders
required to reach global optima

pb	n	m	d	M	N	CPU	LMI
[FP2.2]	5	11	2	461	7987	11.8	3
[FP2.3]	6	13	2	209	1421	1.86	2
[FP2.4]	13	35	2	2379	17885	1012	2
[FP2.5]	6	15	2	209	1519	1.58	2
[FP2.6]	10	31	2	1000	8107	67.7	2
[FP2.7]	10	25	2	1000	7381	75.3	2
[FP2.8]	20	10	2	-	-	-	dim
[FP2.9]	24	10	2	-	-	-	dim
[FP2.10]	10	11	2	1000	5632	45.3	2
[FP3.2]	8	22	2	3002	71775	3032	3
[FP3.3]	5	16	2	125	1017	1.20	2
[FP3.4]	6	16	2	209	1568	1.50	2
[FP3.5]	3	8	2	164	4425	2.42	4
[FP4.2]	1	2	6	6	34	0.17	3
[FP4.3]	1	2	50	50	1926	0.94	25
[FP4.4]	1	2	5	6	34	0.25	3
[FP4.5]	1	2	4	4	17	0.14	2
[FP4.6]	2	2	6	27	172	0.41	3
[FP4.7]	1	2	6	6	34	0.20	3
[FP4.8]	1	2	4	4	17	0.16	2

SYSTEMS OF POLYNOMIAL EQUATIONS

$$\mathbf{S} \quad \rightarrow \quad g_1(x_1,\ldots,x_n) = 0; \ldots; \; g_n(x_1,\ldots,x_n) = 0.$$

where $g_i \in \mathbf{R}[x_1,\ldots,x_n]$ for all $i = 1,\ldots,n$.

The polynomial **ideal** $I = \langle g_1,\ldots,g_n\rangle \subset \mathbf{R}[x_1,\ldots,x_n]$, generated by the family $\{g_i\}$ is assumed to be zero-dimensional, i.e. the algebraic variety

$$V_{\mathbf{C}}(I) := \{z \in \mathbf{C}^n \,|\, g_k(z) = 0 \quad k = 1,\ldots,n\} \quad \textbf{is finite}$$

Define

$$V_{\mathbf{R}}(I) := \{x \in \mathbf{R}^n \,|\, g_k(x) = 0 \quad k = 1,\ldots,n\}$$

be the set of **real zeros** of the system $\mathbf{S}$.

I. Numerical solution :

Solve the SDP-relaxations

$$\mathbf{Q}_r \to \max_y \{\sum_\alpha f_\alpha y_\alpha \mid M_r(y) \succeq 0;\ M_{r-v_k}(g_k y) = 0 \quad k = 1, \ldots, n\}$$

for some arbitrary polynomial criterion f.

When I is zero-dimensional $\mathbf{Q}_r$ is exact for some r.

Systems of polynomial equations:
CPU times in seconds and SDP-relaxation orders
required to extract at least one solution

problem	n	m	d	M	N	CPU	LMI	sol
boon	6	6	4	3002	52864	1220	4	8
bifur	3	3	9	454	8717	8.20	5	2
brown	5	5	5	461	4061	6.27	3	1
butcher	7	7	4	6434	120156	-	4	mem
camera1s	6	6	2	209	952	1.33	2	2
caprasse	4	4	4	209	1285	0.58	3	2
cassou	4	4	8	4844	280151	-	8	mem
chemequ	5	5	3	461	3661	9.48	3	1
chemequs	5	5	3	124	486	6.73	2	1
cohn2	4	4	6	209	1229	0.48	3	1
cohn3	4	4	6	209	1229	0.55	3	1
comb3000	10	10	3	1000	4951	24.6	2	1
conform1	3	3	4	83	430	0.22	3	2
conform2	3	3	4	83	430	0.19	3	2
conform3	3	3	4	285	3766	3.89	5	4
conform4	3	3	4	454	8946	12.2	6	2
cpdm5	5	5	3	125	446	0.24	2	1
d1	12	12	3	-	-	-	3	dim
des18_3	8	8	3	12869	303945	-	4	mem
des22_24	10	10	2	1000	5016	77.2	1	1

problem	n	m	d	M	N	CPU	LMI	sol
discret3	8	8	2	44	89	0.31	1	1
eco5	5	5	3	461	3661	5.98	3	1
eco6	6	6	3	923	7980	57.4	3	1
eco7	7	7	3	1715	15921	256	3	1
eco8	8	8	3	3002	29565	1310	3	1
fourbar	4	4	4	69	229	0.16	2	1
geneig	6	6	3	923	7602	33.2	3	1
heart	8	8	4	3002	31545	1532	3	2
i1	10	10	3	1000	4366	44.1	2	1
ipp	8	8	2	494	2385	6.42	2	1
katsura5	6	6	2	209	952	0.74	2	1
kinema	9	9	2	714	3520	26.4	2	1
kin1	12	12	3	-	-	-	3	dim
ku10	10	10	2	1000	5016	72.5	2	1
lorentz	4	4	2	209	1705	0.64	2	2
manocha	2	2	8	90	826	1.27	6	1
noon3	3	3	3	83	430	0.22	3	1
noon4	4	4	3	209	1285	0.65	3	1
noon5	5	5	3	461	3241	4.48	3	1

problem	n	m	d	M	N	CPU	LMI	sol
proddeco	4	4	4	69	229	0.11	2	1
puma	8	8	2	3002	35505	1136	3	4
quadfor2	4	4	4	209	1495	0.75	3	2
quadgrid	5	5	5	461	3641	10.52	3	1
rabmo	9	9	5	5004	51703	-	3	mem
rbpl	6	6	3	923	7602	36.9	3	1
redeco5	5	5	2	20	41	0.16	1	1
redeco6	6	6	2	27	55	0.13	1	1
redeco7	7	7	2	35	71	0.14	1	1
redeco8	8	8	2	44	89	0.13	1	1
rediff3	3	3	2	9	19	0.09	1	1
reimer5	5	5	6	6187	264516	-	6	mem
rose	3	3	9	679	16681	79.5	7	2
s9_1	8	8	2	494	2385	5.45	2	1
sendra	2	2	7	65	453	0.34	5	1
solotarev	4	4	3	69	257	0.24	2	1
stewart1	9	9	2	714	3520	20.4	2	2
stewart2	12	10	2	1819	9191	372	2	1
trinks	6	6	3	209	925	0.78	2	1
virasoro	8	8	2	44	89	0.16	1	1

II. Characterization of zeros

Problems : Give conditions on the coefficients of $\{g_i\}$ to ensure that

- $V_{\mathbf{C}}(I) \equiv V_{\mathbf{R}}(I)$, i.e., $\mathbf{S}$ has only real zeros

- $V_{\mathbf{C}}(I) \equiv V_{\mathbf{R}}(I)$ and $V_{\mathbf{R}}(I) \subseteq \mathbf{K}$ for some specified semi-algebraic set $\mathbf{K} \subset \mathbf{R}^n$.

- $V_{\mathbf{C}}(I) \subseteq \mathbf{K}$ for some specified semi-algebraic set $\mathbf{K} \subset \mathbf{C}^n$.

Illustration on the one-dimensional case.

Let $g \in \mathbf{R}[x]$ with $x \mapsto g(x) = x^{n+1} + a_n x^n + \ldots + a_0$.

Conditions on the $\{a_i\}$ for the n zeros $\{x(j)\} \subset \mathbf{C}$ (counting multiplicities) of g to be all real and contained in the interval $[A, B] \subset \mathbf{R}$.

Define the Newton sums (counting multiplicities)

$$s_k := \frac{1}{n} \sum_{j=1}^{n} x(j)^k \qquad k = 0, 1, \ldots,$$

The s_k's are the **moments** of the **probability measure**

$$\mu := \frac{1}{n} \sum_{j=1}^{n} \delta_{x(j)} \quad \text{counting multiplicities}$$

Hence, write that μ is supported on $\mathbf{K} := [A, B]$!!

Let $M(n,s)$ and $H(n,s)$ be the respective Hankel matrices

$$\begin{bmatrix} 1 & s_1 & \cdots & s_n \\ s_1 & s_2 & \cdots & s_{n+1} \\ & & & \\ s_n & s_{n+1} & \cdots & s_{2n} \end{bmatrix}; \quad \text{and} \begin{bmatrix} s_1 & s_2 & \cdots & s_{n+1} \\ s_2 & s_2 & \cdots & s_{n+2} \\ & & & \\ s_{n+1} & s_{n+2} & \cdots & s_{2n+1} \end{bmatrix}.$$

Theorem : [Lasserre (J. Alg. Comb. (2002))]

(i) All the zeros of g are real **iff** $M(n,s) \succeq 0$.
(and rank($M_n(s,y)$)) are distinct.)

(ii) All the zeros of g are real and contained in $[A,B]$ **iff :**

$$BM(n,s) \succeq H(n,s) \succeq AM(n,s)$$

The complex case

The complex moment matrix : Consider the basis of monomials

$$1, z, \overline{z}, z^2, z\overline{z}, \overline{z}^2, \ldots, z^n, z^{n-1}\overline{z}, \ldots z\overline{z}^{n-1}, \overline{z}^n, \ldots$$

of the complex polynomials $q \in \mathbf{C}[z, \overline{z}]$, that is,

$$z \mapsto q(z, \overline{z}) = \sum q_{ij} \overline{z}^i z^j \quad \text{for finitely many } q_{ij}.$$

For a measure μ on $\mathbf{C}$ let $y_{ij} := \int \overline{z}^i z^j \, d\mu$ for all $i, j = 0, 1, \ldots$, and let $M_r(y)$ be its moment matrix, i.e.,

$$[M_r(i,1) = y_{pq} \text{ and } M_r(1,j) = y_{vw}] \Rightarrow \quad M_r(i,j) = y_{p+v,q+w}.$$

Then, $\forall\, f \in \mathbf{C}[z, \overline{z}]$ with degree $\leq r$, and coefficient vector f,

$$\langle \mathbf{f}, \mathbf{M_r f} \rangle = \int \overline{f}(\mathbf{z})\, \mathbf{f}(\mathbf{z})\, \mu(\mathbf{dz}) = \int |\mathbf{f}|^2 \, \mathbf{d}\mu,$$

so that $M_r(y)$ is Hermitian and positive semidefinite ($M_r(y) \succeq 0$)

Let $p(z) = a_0 + a_1 z + \ldots a_n z^n + z^{n+1} \in \mathbf{R}[z]$ be a polynomial with real coefficients, and let

$$\mathbf{K} := \{z \in \mathbf{C} \mid \quad g_k(z, \overline{z}) \geq 0; \quad k = 1, \ldots, m\}$$

be a given (nonnecessarily compact) semialgebraic set of $\mathbf{C}$.

Problem: Under which condition on the coefficients $\{a_j\}$ do we have all the zeros of p contained in $\mathbf{K}$?

Example : $\mathbf{K} = \{z \in \mathbf{C} \mid \quad z + \overline{z} \leq 0\}$ for stability of linear systems.

Let $\{z(k)\}_{k=1}^{n} \subset \mathbf{C}$ be the n zeros of p (counting their multiplicity), and define the probability measure μ on $\mathbf{C}$

$$\mu := \frac{1}{n} \sum_{k=1}^{n} \delta_{z(k)}; \quad y_{ij}^{*} = \int \overline{z}^{i} z^{j} \, d\mu = y_{ji}^{*}.$$

Hence, $M_r(y^*) \succeq 0$ is a real symmetric matrix.

$$y^*_{0j} = y^*_{j0} := \int z^j \, d\mu = s_j \quad j_{th} \text{ Newton sum } = f_j(a)$$

The Newton sums s_j's are known and (easy to compute) functions of the a_j's, but not the y^*_{ij}.

Also, because $p(z) = 0$, we have

$$z^p = \sum_{k=0}^{n} \beta_k(p) z^k; \qquad \overline{z}^p = \sum_{k=0}^{n} \beta_k(p) \overline{z}^k$$

for some $\{\beta_k(p)\} \subset \mathbf{R}$. Hence, for all p, q we have

$$(*) \quad y^*_{pq} = \sum_{0 \le i,j \le n} \gamma_{ij}(p,q) \, y^*_{ij} \qquad \text{for some } \{\gamma_{ij}(p,q)\} \subset \mathbf{R}.$$

and the $\gamma_{ij}(p,q)$'s are easy to compute.

So define the moment matrix $M_r(y)$ (or $M_r(y,s)$)

$$\text{Ex}: \; M_2(y) = M_2(y,s) = \begin{bmatrix} 1 & s_1 & s_1 & s_2 & y_{11} & s_2 \\ s_1 & y_{11} & s_2 & y_{12} & y_{12} & s_3 \\ s_1 & s_2 & y_{11} & s_3 & y_{12} & y_{12} \\ s_2 & y_{12} & s_3 & y_{22} & y_{13} & s_4 \\ y_{11} & y_{12} & y_{12} & y_{13} & y_{22} & y_{13} \\ s_2 & s_3 & y_{12} & s_4 & y_{13} & y_{22} \end{bmatrix}$$

with $\mathbf{y}$ **unknown** in lieu of $\mathbf{y}^*$, and using (*)

$$y_{pq} = \sum_{0 \le i,j \le n} \gamma_{ij}(p,q)\, y_{ij}$$

For $z \mapsto g_k(z,\overline{z}) = \sum_{u,v} g_k(u,v)\overline{z}^u\, z^v$, define the usual localizing matrices

$$M_r(y,s)(i,j) = y_{pq} \Rightarrow \quad M_{g_k,r}(y,s)(i,j) = \sum g_k(u,v)\, y_{p+u,q+v}$$

for all $k = 1,\ldots,m$.

Theorem : All the zeros of p are contained in $\mathbf{K}$ if and only if

$$M_n(y,s) \succeq 0; \quad M_{g_k,n}(y,s) \succeq 0, \qquad k = 1,\ldots,m$$

for some vector $\{y_{ij}\}$. (In which case $y = y^*$ is unique)

The proof uses a nice result of Curto and Fialkow on flat positive extensions of moment matrices.

Multivariable case $\mathbb{C}^n$

Consider the system $\mathbf{S}$ of polynomial equations

$$h_1(x_1) = 0;\ h_2(x_1, x_2) = 0; \ldots;\ h_n(x_1, x_2, \ldots, x_n) = 0$$

in **triangular form**, where :

$$h_k(x) = h_{k1}(x_1, \ldots, x_{k-1})\, x_k^{r_k} + h_{k2}(x_1, \ldots, x_k); \quad k = 2, \ldots, n$$

and $h_{k1}(x_1, \ldots, x_{k-1}) \neq 0$ whenever $h_i(x) = 0, k = 1, \ldots k-1$.

(i) Every system of polynomial equations associated with a zero-dimensional ideal is a finite union of such triangular systems

(ii) Symbolic computation packages can obtain this form.

Let $\{z(k)\}_{k=1}^{t} \subset \mathbf{C}^n$ be the t complex zeros of the system $\mathbf{S}$. Then, if one defines the *generalized* Newton sums

$$s_\alpha := \int z_1^{\alpha_1} \cdots z_n^{\alpha_n} \, d\mu, \quad \text{with } \mu := \frac{1}{t} \sum_{k=1}^{t} \delta_{z(k)}$$

(**) One may compute $\{s_\alpha\}$ recursively as rational fractions of the coefficients of the polynomials $\{h_k\}$ that define $\mathbf{S}$.

Similarly, let

$$y^*_{\alpha\beta} = \int \overline{z}^{\alpha} \, z^{\beta} \, \mu(dz) \qquad \alpha, \beta \in \mathbf{N}^n.$$

As in the one-dimensional case,

$$(*) \qquad y^*_{\eta\delta} = \sum_{\alpha_i, \beta_i < r_i \, \forall i} \gamma_{\alpha\beta}(\eta, \delta) \, y^*_{\alpha\beta}$$

Let $\mathbf{K} := \{z \in \mathbf{C}^n \mid \quad g_k(z, \overline{z}) \geq 0,\ k = 1, \ldots, m\} \subset \mathbf{C}^n$ be given.

One defines the multivariable analogues of the moment matrix $M_r(y, s)$ and localizing matrices $M_{g_k,r}(y, s)$, with $\mathbf{y}$ unknown in lieu of $\mathbf{y}^*$

Theorem : All the zeros of $\mathbf{S}$ are contained in $\mathbf{K}$ if and only if

$$M_n(y, s) \succeq 0; \quad M_{g_k,n}(y, s) \succeq 0, \qquad k = 1, \ldots, m$$

for some vector $\{y_{ij}\}$. (In which case $y = y^*$ is unique)

Again, we use Curto and Fialkow's result on flat positive extensions of moment matrices.

BOUNDS ON MEASURES WITH MOMENT CONDITIONS

[Lasserre (2002)], Annals Appl. Prob.

Let $\Gamma \subset \mathbf{N}^n$ and let $\{\gamma_\alpha\}_{\alpha\in\Gamma}$ be a finite sequence of scalars.
Problem: Given a semi-algebraic set

$$\mathbf{K} := \{x \in \mathbf{R}^n \,|\, g_i(x) \geq 0, \quad i = 1, \ldots, m\}$$

we want to find, or approximate

$$\rho^* := \sup_{\mu\in\mathcal{P}(\mathbf{R^n})} \{\mu(\mathbf{K}) \,|\, \int \mathbf{x}^\alpha \, \mathbf{d}\mu = \gamma_\alpha \quad \alpha \in \Gamma\}$$

Applications in Probability, Finance, Queuing, ...

Work motivated by nice results from **Bertsimas and Popescu** (1999)-(2002) for $\mathbf{K}$ convex, and first-and second order moment conditions.

Write $\mu = \varphi + \psi$ with $\varphi(\mathbf{K}^c) = 0$ and let $\{y_\alpha, z_\alpha\}$ be the respective moments of φ, ψ, so that

$$\int \mathbf{x}^\alpha \mathbf{d}\mu = \mathbf{y}_\alpha + \mathbf{z}_\alpha \quad \alpha \in \mathbf{N}^\mathbf{n}.$$

$$\rho^* = \begin{cases} & \sup y_0 \\ \text{s.t.} & y_\alpha + z_\alpha \quad = \gamma_\alpha \quad \alpha \in \Gamma \\ & \varphi(\mathbf{K}^c) \quad = 0 \end{cases}$$

→ **SDP-relaxations**

$$\mathbf{Q}_r \begin{cases} & \sup y_0 \\ \text{s.t.} & M_r(y), M_r(z) \succeq 0 \\ & M_{r-v_i}(g_i y) \succeq 0 \quad i = 1, \ldots, m \\ & y_\alpha + z_\alpha = \gamma_\alpha \quad \alpha \in \Gamma \end{cases}$$

Theorem: (a) As $r \to \infty$

$$\sup \mathbf{Q}_r \downarrow \delta \geq \rho^*.$$

(b) If in addition, $\mathbf{Q}_r$ is solvable and an optimal solution y^*, z^* satisfies

$$\operatorname{rank} M_r(y^*) = \operatorname{rank} M_{r-d}(y^*); \quad \operatorname{rank} M_r(z^*) = \operatorname{rank} M_{r-1}(z^*),$$

then

$$\max \mathbf{Q}_r = \rho^* = \max_{\mu \in \mathcal{P}(\mathbf{R}^n)} \{\mu(\mathbf{K}) \mid \int x^\alpha \, d\mu = \gamma_\alpha, \quad \alpha \in \Gamma.$$

$\longrightarrow$ the bound $\max \mathbf{Q}_r$ is tight!

Performance evaluation: Invariant probability measures of Markov chains

Let $(X, \mathcal{B}, Q)$ be a time-homogeneous Markov chain (MC) $\Phi_\bullet = (\Phi_0, \Phi_1, \ldots)$ with **state space** X, and **t.p.f.** Q.

- $Q(x, .)$ is a **prob. measure** on $\mathcal{B}$ for all $x \in X$.

- $x \mapsto Q(x, B)$ is **measurable** for all $B \in \mathcal{B}$.

$$\text{Prob}(\Phi_t \in B \,|\, \Phi_{t-1} = x) = Q(x, B) \quad B \in \mathcal{B},\, x \in X.$$

$\mu \in \mathcal{P}(X)$ is an **invariant prob. measure** for the MC $\Phi_\bullet$ iff $\mu Q = \mu$, i.e.,

$$\mu(B) = \int_X Q(x, B)\, \mu(dx) \qquad \forall\, B \in \mathcal{B}.$$

If μ is unique and $f \in L_1(\mu)$ then :

$$\lim_{\mathbf{T}\to\infty} \mathbf{E_x} \frac{1}{T} \sum_{\mathbf{t=0}}^{\mathbf{T-1}} \mathbf{f}(\mathbf{\Phi_t}) = \int_{\mathbf{X}} \mathbf{f}\, \mathbf{d}\mu \quad \mu\text{–a.e.,}$$

and for μ-a.a. $x \in X$,

$$\lim_{\mathbf{T}\to\infty} \frac{1}{T} \sum_{\mathbf{t=0}}^{\mathbf{T-1}} \mathbf{f}(\mathbf{\Phi_t}) = \int_{\mathbf{X}} \mathbf{f}\, \mathbf{d}\mu \quad \mathbf{P_x}\text{–a.s}$$

Simulation gives only an **estimator** of $\int f\, d\mu$.

Remark: $X = \mathbf{R}^n$. If $Qf \in \mathbf{R}[x_1, \ldots, x_n]$ when $f \in \mathbf{R}[x_1, \ldots, x_n]$ then

$$y_\alpha = \int x^\alpha\, d\mu = \int (Qx^\alpha)\, d\mu = \langle A_\alpha, y \rangle$$

Invariance of $\mu \Rightarrow$ linear relations between its moments!

SDP-relaxations: Let $f = \sum_\alpha f_\alpha x^\alpha \in \mathbf{R}[x]$, and

$$\mathbf{Q}_r \left\{ \begin{array}{lll} & \max_y \text{ (or } \min_y) & \sum_\alpha f_\alpha y_\alpha \\ \text{s.t.} & M_r(y) & \succeq 0 \\ & y_\alpha - \langle A_\alpha, y \rangle & = 0, \quad \forall\, |\alpha| \leq r \end{array} \right.$$

$$\text{Then} \quad \min \mathbf{Q}_r \leq \int_X f\, d\mu \leq \max \mathbf{Q}_r \qquad \forall\, r$$

so that

$$\sup_r \min \mathbf{Q}_r = \underline{\rho} \leq \int_X f\, d\mu \leq \overline{\rho} = \inf_r \max \mathbf{Q}_r$$

If μ is the **unique inv. prob.** measure then one gets **upper and lower bounds** on $\int f\, d\mu$. More generally :

$$\sup_{\mathbf{r}} \min \mathbf{Q_r} = \underline{\rho} \leq \inf_{\mu\mathbf{Q}=\mu} \int_{\mathbf{X}} \mathbf{f}\, \mathbf{d}\mu \leq \sup_{\mu\mathbf{Q}=\mu} \int_{\mathbf{X}} \mathbf{f}\, \mathbf{d}\mu \leq \overline{\rho} = \inf_{\mathbf{r}} \max \mathbf{Q_r}$$

LP-relaxations: Stockbridge, Helmes

$$\text{LP}_r \begin{cases} & \max_y \text{ (or } \min_y) \sum_\alpha f_\alpha y_\alpha \\ \text{s.t.} & \text{Hausdorff (linear) moment conditions} \\ & y_\alpha - \langle A_\alpha, y \rangle = 0, \quad \forall\, |\alpha| \leq r \end{cases}$$

SDP better than LP-relaxations which are ill-posed because of binomial coefficients in the Hausdorff moment conditions. But SDP software packages are still limited.

Ex: I. Barnsley Iterated function systems:

$$x_{t+1} = f_{\xi_t}(x_t), \quad t = 0, 1, \ldots,$$

- $X \equiv \mathbf{R}^n$, and $\xi_\bullet$ is a sequence of i.i.d. random variables with Prob$(\xi_0 = i) = p_i$, $i = 1, \ldots, N$.

- $f_i : \mathbf{R}^n \to \mathbf{R}$ is a polynomial for every $i = 1, \ldots, N$

Let μ be an invariant prob. with all moments $\{y_\alpha\}$ finite.

$$y_\alpha = \int_{\mathbf{R}^n} x^\alpha d\mu = \sum_{j=1}^{N} \int_{\mathbf{R}^n} p_j f_j(x)^\alpha \, d\mu = \langle A_\alpha, y \rangle$$

Ex : Logistic map:

$X \equiv [0,1]$ and $x_{t+1} = 4x_t(1-x_t)$.

$$y_n = \int_0^1 x^n \, d\mu = \int_0^1 [4x(1-x)]^n \, d\mu = \sum_{j=0}^{n} 4^n \binom{n}{j} y_{n+j}$$

II. Diffusions.

$$\mathbf{dX_s = b(s, X_s)\, ds + \sigma(s, X_s)\, d\, W_s, \quad X_0 = x_0,}$$

where $\{W_s\}$ is a standard one-dimensional Brownian motion.

The associated generator $\mathcal{A}$ with domain $\mathcal{D}$ satisfies

$$\mathcal{A} = \frac{\partial}{\partial t} + \mathbf{b(t, x)}\, \frac{\partial}{\partial x} + \frac{1}{2}\sigma\mathbf{(t, x)^2}\, \frac{\partial^2}{\partial x^2},$$

and

$$\mathbf{f(\tau, X_\tau) - f(t, X_t) - \int_t^\tau \mathcal{A}f(s, X_s)\, ds, \quad \forall\, t < \tau \in [0, T]}$$

is a martingale for each $f \in \mathcal{D}$.

Performance criterion :

$$J(\tau) = E_{0,x_0}\,[\int_0^\tau h(s, X_s)\, ds + g(\tau, X_\tau)\,]$$

Define the measure ν_τ and probability measure μ_τ on $[0,T]\times\mathbf{R}$

$$\nu_\tau(A) \;:=\; E_{0,x}\left[\int_0^\tau \mathrm{I}_A(s,X_s)\,ds\right]; \qquad \mu_\tau(A) \;:=\; E_{0,x}\left[\mathrm{I}_A(\tau,X_\tau)\right]$$

with respective moments $\{u_{ij}\}$ and $\{v_{ij}\}$. From

$$\int f\,d\mu_\tau \;=\; f(0,x_0)+\int \mathcal{A}f(t,x)\,d\nu_\tau(t,x),$$

and with f a polynomial t^ix^j of $\mathbf{R}[t,x]$

$$(*) \qquad \int t^ix^j\,d\mu_\tau \;=\; v_{ij} \;=\; 0^ix_0^j+\int \mathcal{A}t^ix^j\,d\nu_\tau(t,x),$$

and if $\mathcal{A}$ maps polynomials into polynomials then (*) reads

$$v_{ij} \;=\; 0^ix_0^j+\langle W_{ij},u\rangle,$$

a **linear relation between moments** of μ_τ and ν_τ!!

SDP Relaxations :

ν_τ is supported on $[0, \tau) \times \mathbf{R}$ and μ_τ is supported on $\{\tau\} \times \mathbf{R}$.

If h, g are polynomials, then one approximates J_τ by

$$\mathbf{Q}_r \rightarrow \left\{ \begin{array}{rcl} \text{max (min)} & & \sum_\alpha h_\alpha u_\alpha + g_\alpha v_\alpha \\ M_r(v) & \succeq & 0 \quad \rightarrow \quad \mu_\tau \\ \tau M_{r-1}(u) - M_{r-1}(t,u) & \succeq & 0 \quad \rightarrow \quad \nu_\tau \ (\tau - t \geq 0) \\ M_{r-1}(t,u) & \succeq & 0 \quad \rightarrow \quad \nu_\tau \ (t \geq 0) \\ v_{ij} - \langle W_{ij}, u \rangle & = & 0^i x_0^j \quad \forall\, i + j \leq 2r \end{array} \right.$$

One obtains upper and lower bounds, because

$$\sup_r \min \mathbf{Q}_r \leq \quad J_\tau \quad \leq \inf_r \max \mathbf{Q}_r$$

CONCLUSION

Theory of moments and positive polynomials have a **wide range of potential important applications ....**

SDP-relaxations for $\mathbf{K}$-moment conditions or Putinar's representation of polynomials > 0 on $\mathbf{K}$, appear to be very efficient.

but **SIZE LIMITATION** ... $\Rightarrow$

Need for efficient SDP-solvers for large-scale problems!!