

A moment approach to analyze zeros of polynomial equations

J.B. Lasserre

LAAS-CNRS, Toulouse, France

MSRI workshop: April 2004

- SDP for solving systems of polynomial equations
- A moment approach to analyze the zero set of polynomial equations
- The univariate case
- The multivariate case

SYSTEM OF POLYNOMIAL EQUATIONS

$$S \rightarrow g_1(x_1, \dots, x_n) = 0; \dots; g_n(x_1, \dots, x_n) = 0.$$

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

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

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

Define

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

be the set of **real zeros** of the system S.

Efficient **Symbolic software** packages exist, especially if S is in **triangular form** (e.g., Aubry, Lazard, Maza, Rouillier)

Numerical solution via SDP-relaxations

Moment matrix. With $\alpha \in \mathbb{N}^n$, and $y_{\alpha_1, \dots, \alpha_n} \rightsquigarrow \int x_1^{\alpha_1} \cdots x_n^{\alpha_n} d\mu$

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

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

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

Localizing matrix.

Given a polynomial $\theta : \mathbf{R}^n \rightarrow \mathbf{R}$ 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 $x \mapsto \theta(x) = 1 - x_1^2 - x_2^2$, $M_2(\theta y) =$

$$\begin{bmatrix} 1 - y_{20} - y_{02}, & y_{10} - y_{30} - y_{12}, & y_{01} - y_{21} - y_{03} \\ y_{10} - y_{30} - y_{12}, & y_{20} - y_{40} - y_{22}, & y_{11} - y_{21} - y_{12} \\ y_{01} - y_{21} - y_{03}, & y_{11} - y_{21} - y_{12}, & y_{02} - y_{22} - y_{04} \end{bmatrix}.$$

If $M_r(y)(i, j) = y_{\beta}$ then $M_r(\theta y)(i, j) = \sum_{\alpha} \theta_{\alpha} y_{\beta + \alpha}$ 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 **K-moment problem** identifies those vectors y with $M_r(y) \succeq 0$ that are **moments of a measure μ with support contained in K** .

Dual theory in algebraic geometry = representation of polynomials, positive on a semi-algebraic set K

With $f \in \mathbf{R}[x]$, introduce the family $\{Q_i\}$ of **SDP-relaxations**

$$Q_i \begin{cases} \min_y \sum_{\alpha} f_{\alpha} y_{\alpha} \\ M_i(y) \succeq 0 \\ M_{i-v_k}(g_k y) = 0, \quad k = 1, \dots, n. \end{cases}$$

and the family $\{Q_i^*\}$ of their **dual**

$$Q_i^* \begin{cases} \max_{X \succeq 0, Z_1, \dots, Z_n} -X(1, 1) - \sum_{k=1}^m g_k(0) Z_k(1, 1) \\ \text{s.t. } \langle X, B_{\alpha} \rangle + \sum_{k=1}^m \langle Z_k, C_{\alpha}^k \rangle = f_{\alpha}, \quad \forall \alpha \neq 0 \end{cases}$$

where we write

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

I. Numerical solution via SDP-relaxations

Let $f \in \mathbf{R}[x]$ be arbitrary, fixed.

Theorem: Let I be a zero-dimensional radical ideal. Then :

(a) $\max \mathbf{Q}_{r_0}^* = \min \mathbf{Q}_{r_0} = \min \{f(x) \mid x \in \mathbf{S}\}$, for some r_0 , and all $r \geq r_0$.

(b) $f - f^* = q_0 + \sum_{j=1}^n q_j g_j$ for some s.o.s. polynomials $\{q_j\}_{j=1}^n$.

(c) Every optimal solution $y^* = \{y_\alpha^*\}$ of \mathbf{Q}_r is the vector of moments of some probability measure supported on the real zeros of G .

So, when I is a radical zero-dimensional ideal, \mathbf{Q}_r is exact for all $r \geq r_0$. (Lasserre (grid case), Laurent, Parrilo for the general case)

Solving Systems of polynomial equations with GLOPTIPOLY
software:

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., 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 subset $\mathbf{K} \subset \mathbf{C}^n$ (or semialgebraic set \mathbf{K}' of \mathbf{R}^{2n}).

The univariate case.

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

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 $[u, v] \subset \mathbf{R}$.

Define the **Newton sums** (counting multiplicities)

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

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 $K := [u, v]$!!

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

$$\begin{bmatrix} 1 & s_1 & \cdots & s_n \\ s_1 & s_2 & \cdots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & \cdots & s_{2n} \end{bmatrix}; \quad \text{and} \quad \begin{bmatrix} s_1 & s_2 & \cdots & s_{n+1} \\ s_2 & s_3 & \cdots & s_{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n+1} & s_{n+2} & \cdots & s_{2n+1} \end{bmatrix}.$$

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

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

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

$$v.H(n, s) \succeq M(n, s) \succeq u.H(n, s)$$

The (univariate) complex case

The complex moment matrix : Consider the basis of monomials

$$1, z, \bar{z}, z^2, z\bar{z}, \bar{z}^2, \dots, z^n, z^{n-1}\bar{z}, \dots, z\bar{z}^{n-1}, \bar{z}^n, \dots$$

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

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

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

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

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

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

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

Let $z \mapsto p(z) := a_0 + a_1z + \dots + a_nz^n + z^{n+1} \in \mathbf{R}[z]$ be a polynomial with real coefficients, and let

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

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 : $\mathbf{K} = \{z \in \mathbf{C} \mid z + \bar{z} \leq 0\}$ for stability of linear systems, which gave the Routh-Hurwitz and Liénard criteria involving determinants, linear in the coefficients of p .

Some necessary and sufficient conditions provided in the pioneering work of Kalman, later extended by Mazko, Gutman and Jury, Chilali and Gahinet, ...

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}

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

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}^* .

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

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

for some $\{\beta_k(p)\} \subset \mathbf{R}$.

Hence, for all p, q we have

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

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 y **unknown** in lieu of y^* , and using (*) to have

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

Therefore, $M_r(y, s)$ contains only the $n(n+1)/2$ variables $\{y_{ij}\}$, with $1 \leq i \leq j \leq n$.

For $z \mapsto g_k(z, \bar{z}) = \sum_{u,v} g_k(u, v) \bar{z}^u z^v$, define the localizing matrices $\{M_r(g_k, y, s)\}$

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

for all $k = 1, \dots, m$, and again using (*).

Theorem : Let μ be the uniform probability measure on the zeros of p (counting multiplicities). Then:

- (a) The SDP constraint $M_{2n}(y, s) \succeq 0$ yields a unique solution y^* , the vector of moments of μ (up to order $2n$).
- (b) All the zeros of p are contained in \mathbb{K} if and only if $M_{2n}(g_k, y^*, s) \succeq 0$, for all $k = 1, \dots, m$

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

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

in **triangular form**, where :

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

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

- (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 $\{z(k)\}_{k=1}^t \subset \mathbb{C}^n$ 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

$$y_{\alpha\beta}^* = \int \bar{z}^\alpha z^\beta \mu(dz) \quad \alpha, \beta \in \mathbb{N}^n.$$

As in the one-dimensional case,

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

So, **once** the $\{y_{\alpha\beta}^*\}$ (with $\alpha_j, \beta_j < r_j$ for all $j = 1, \dots, n$) are known, then **all the $y_{\alpha\beta}^*$'s are known via (*)!**

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

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

$$(*) \quad y_{\eta\delta} = \sum_{\alpha_i, \beta_i < r_i \forall i} \gamma_{\alpha\beta}(\eta, \delta) y_{\alpha\beta}$$

so that $M_r(\mathbf{y}, s)$ and $M_{g_k, r}(\mathbf{y}, s)$ contain only the variables $\{y_{\alpha\beta}\}$ with $\alpha_i, \beta_i \leq r_i$ for all $i = 1, \dots, n$.

Theorem : Let μ be the uniform probability measure on the zeros of S (counting multiplicities). Let $r_0 := \sum_{j=1}^n r_j - 1$. Then:

(a) The SDP constraint $M_{2r_0}(y, s) \succeq 0$ yields a unique solution y^* , the vector of moments of μ (up to order $2r_0$).

(b) $\text{rank}(M_{2r_0}(y, s))$ gives the number of distinct zeros of S .

(c) Let $I = \langle g_1, \dots, g_n \rangle$. $f \in \mathbf{R}[z]$ of degree $2p$ or $2p - 1$. Then

$$f \in \sqrt{I} \Leftrightarrow M_p(\mu^*)f = 0.$$

(d) All the zeros of p are contained in \mathbf{K} if and only if $M_{2r_0}(g_k, y^*, s) \succeq 0$, for all $k = 1, \dots, m$

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

Real zeros

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

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

Let $M_r(y)$ be the **real moment matrix** with rows and columns indexed in the basis $1, x_1, \dots, x_n, x_1^2, \dots, x_n^r$.

If μ is a probability measure with support on the **real zeros** of \mathbf{S} and $y = \{y_\alpha\}$ is the vector of its moments, one has

$$(*) \quad y_\beta = \sum_{\alpha_i < r_i \forall i} \gamma_\alpha(\beta) y_\alpha$$

for some real coefficients $\{\gamma_\alpha(\beta)\}_\alpha$, easy to compute.

Replace y_β with $(*)$ in $M_r(y)$, whenever $\beta_j > r_j$, for some j . So, $M_r(y)$ has only **finitely many** unknowns $\{y_\alpha\}$, for all r .

Theorem : Let $r_0 := \sum_{j=1}^n r_j - 1$. Then:

(a) The number s_0 of distinct zeros of S is the maximum rank of the real moment matrices $M_{r_0}(y)$ which are positive semidefinite.

(b) Let $M_{r_0}(y^*) \succeq 0$ with $\text{rank}(M_{r_0}(y^*)) = s_0$. Then, $M_r(y^*) \succeq 0$ and $M_r(y^*)$ has rank s_0 for all $r \geq r_0$.

(c) Let $I = \langle g_1, \dots, g_n \rangle$, and $f \in \mathbf{R}[x]$ be of degree $\leq r$. Then

$$f \in I(V_{\mathbf{R}}(I)) \Leftrightarrow M_r(y^*)f = 0.$$

(d) All the zeros of p are contained in \mathbf{K} if and only if $M_{r_0}(g_k, y^*) \succeq 0$, for all $k = 1, \dots, m$

(**) Solving $M_{r_0}(y) \succeq 0$ with $M_{r_0}(y)$ of maximal rank is NP-hard!!

Note in passing that

S has only **real zeros** if and only if $M_{r_0}(s) \succeq 0$, where $s = \{s_\alpha\}$ is the (known) vector of Newton's sums of S .