# Moment matrices, radical ideals and optimization

## Revisiting two theorems of Curto and Fialkow

Solution of the truncated complex moment problem for flat data. *Memoirs of the AMS* (119) 568, 1996

The truncated complex $K$-moment problem. *Transactions of the AMS* (352), 2000.

**Monique Laurent**

**CWI, Amsterdam**

## The moment problem

Given a sequence $y = (y_\alpha)_{\alpha \in \mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$
Does there exist a nonnegative measure $\mu$ on $\mathbb{R}^n$ with *moments:*

$$\boxed{\int x^\alpha d\mu(x) = y_\alpha \quad (\alpha \in \mathbb{N}^n)}$$

$\mu$ is then called a *representing measure* for $y$

## Variations of the problem:

- **The $F$-moment problem:** Ask for a measure $\mu$ supported by a given subset $F \subseteq \mathbb{R}^n$

- **The truncated moment problem:** We are given a *truncated* sequence $y = (y_\alpha)_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = \sum_i \alpha_i \leq t}}$

- **The complex moment problem**

# Moment matrices

$$y = (y_\alpha)_{\alpha\in\mathbb{N}^n} \qquad\qquad M(y) = (y_{\alpha+\beta})_{\alpha,\beta\in\mathbb{N}^n}$$

As usual, identify a polynomial $p(x) = \sum_\alpha p_\alpha x^\alpha$ with its sequence of coefficients $p = (p_\alpha)_\alpha$

**Fact:** If $y$ has a representing measure, then

$$p^T M(y) p = \int p(x)^2 d\mu(x) \geq 0$$

for any polynomial $p(x)$

**Therefore:**
(1) $M(y) \succeq 0$
(2) $supp(\mu) \subseteq \bigcap_{p\in \mathrm{Ker} M(y)} Zeros(p)$
(3) rank $M(y) \leq |supp(\mu)|$

**Fact:** If $\mu$ has a finite support, i.e.,

$$\mu = \sum_{i=1}^r \lambda_i \delta_{x_i} \;\text{ where } \lambda_i > 0,\; x_i \in \mathbb{R}^n$$

*Dirac measure* at $x_i$

then

$$y = \sum_{i=1}^r \lambda_i \zeta_{x_i}$$

$$M(y) = \sum_{i=1}^r \lambda_i \zeta_{x_i} \zeta_{x_i}^T$$

setting $\zeta_x := (x^\alpha)_{\alpha\in\mathbb{N}^n}$

## Moment sequences and positive polynomials

$\mathcal{M}$= cone of $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ with a representing measure
$\mathcal{M}^+$= cone of $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ with $M(y) \succeq 0$

$\mathcal{P}$= cone of polynomials nonnegative on $\mathbb{R}^n$
$\Sigma$= cone of sums of squares of polynomials

$$\mathcal{M} \subseteq \mathcal{M}^+, \quad \Sigma \subseteq \mathcal{P}$$

$$y \in \mathcal{M} \overset{\text{[Haviland]}}{\Longleftrightarrow} y^T p \geq 0 \ \forall p \in \mathcal{P}$$

$$y \in \mathcal{M}^+ \Longleftrightarrow y^T p \geq 0 \ \forall p \in \Sigma$$

$$\begin{array}{ccc} \mathcal{M} & \longleftrightarrow & \mathcal{P} \\ & \text{dual cones} & \\ \mathcal{M}^+ & \longleftrightarrow & \Sigma \end{array}$$

- $n = 1 : \mathcal{P} = \Sigma \Longrightarrow \mathcal{M} = \mathcal{M}^+$ [Hamburger's theorem]
- $n \geq 2 : \Sigma \subset \mathcal{P} \Longrightarrow \mathcal{M}^+ \subset \mathcal{M}$

**Example:** $x_1^2 x_2^2 (x_1^2 + x_2^2 - 1) + 1 \in \mathcal{P} \setminus \Sigma$

**Hilbert's theorem:** The pairs $(n, d)$ for which every degree $d$ polynomial in $n$ variables nonnegative on $\mathbb{R}^n$ is a sum of squares of polynomials are:
(1) $n = 1$, $d \geq 1$
(2) $n \geq 1$, $d = 2$
(3) $n = 2$, $d = 4$

# Positive semidefiniteness of the moment matrix is sometimes sufficient for the existence of a representing measure

**Theorem:** [Berg, Christensen, Ressel 1976]
If $M(y) \succeq 0$ and $y$ is bounded, then $y$ has a (unique) representing measure.

**Theorem A:** [Curto-Fialkow 1996]
If $M(y) \succeq 0$ and $M(y)$ has finite rank $r$, then $y$ has a (unique) representing measure.

Moreover, $supp(\mu) = \bigcap_{p \in \text{Ker} M(y)} Zeros(p)$ has cardinality $r$.

**C-F proof:** based on operator theory [spectral theorem + Riesz representation theorem]

**New (elementary) proof:** use Hilbert's Nullstellensatz and the fact that Ker $M(y)$ is a *radical ideal.*

**Proof of:** $M(y) \succeq 0$, rank $M(y) = r \Longrightarrow$
$y$ has a $r$-atomic representing measure

**Lemma 1:** $(fg)^T M(y) h = f^T M(y)(gh)$ for any polynomials $f, g, h$

**Corollary 1:** $I :=$ Ker $M(y)$ is a radical ideal in $\mathbb{R}[x_1, \ldots, x_n]$.

**Lemma 2:** Let $\mathcal{B}$ be a set of monomials indexing a maximal nonsingular principal submatrix of $M(y)$. Then, $\mathcal{B}$ is a basis of $\mathbb{R}[x_1, \ldots, x_n]/I$.

**Corollary 2:** As $I$ is radical, its variety:

$$V(I) = \{x \in \mathbb{C}^n \mid p(x) = 0 \ \forall p \in I\}$$

has cardinality $|V(I)| = r$ and, by the Nullstellensatz,

$$p \in I \Longleftrightarrow p(v) = 0 \ \forall v \in V(I)$$

$$\boxed{\dim \ \mathbb{R}[x_1, \ldots, x_n]/I = |V(I)| = r}$$

**Our objective:** Show that $y$ has a representing measure supported by $V(I)$.

## Proving that $\boxed{V(I) \subseteq \mathbb{R}^n}$

Let $p_v \in \mathbb{C}[x_1, \ldots, x_n]$ be *interpolation polynomials* at $v \in V(I)$; i.e., $p_v(v') = 1$ if $v = v'$ and 0 otherwise.
Let $Z$ be the matrix with columns $\zeta_v$ ($v \in V(I)$) and let $\tilde{Z}$ be the matrix with rows $p_v$ ($v \in V(I)$).

$$Z = \begin{array}{c} \quad S \quad\; T \quad\; \overline{T} \\ \begin{pmatrix} \\ A & B & \overline{B} \\ \\ \end{pmatrix} \end{array}, \quad \tilde{Z} = \begin{array}{c} S \\ T \\ \overline{T} \end{array} \begin{pmatrix} C \\ D \\ \overline{D} \end{pmatrix}$$

setting $V(I) = S \cup T \cup \overline{T}$, $S := V(I) \cap \mathbb{R}^n$, $\overline{T} = \{\overline{v} \mid v \in T\}$

**Lemma 3:** $\boxed{M(y) = Z \mathrm{diag}(\tilde{Z} y) Z^T}$

**Corollary 3:** $\boxed{M(y) \succeq 0 \Longrightarrow T = \emptyset}$

**Proof:**

$$M(y) = \begin{pmatrix} \\ A & B & \overline{B} \\ \\ \end{pmatrix} \begin{pmatrix} a & & \\ & b & \\ & & \overline{b} \end{pmatrix} \begin{pmatrix} A^T \\ B^T \\ \overline{B}^T \end{pmatrix}$$

$$= \underbrace{A \mathrm{diag}(a) A^T}_{A_+ A_+^T - A_- A_-^T} + \underbrace{B \mathrm{diag}(b) B^T + \overline{B \mathrm{diag}(b) B^T}}_{EE^T - FF^T}$$

$$= (A_+ A_+^T + EE^T) - (A_- A_-^T + FF^T)$$

As $A_+, A_-, E, F$ are real matrices,

$$\begin{aligned} |S| + 2|T| = \mathrm{rank}\ M(y) &\leq \mathrm{rank}(A_+ A_+^T + EE^T) \\ &\leq |\{v \in S \mid a_v > 0\}| + |\{v \in T \mid b_v \neq 0\}| \\ &\leq |S| + |T| \end{aligned}$$

**Lemma 4:** $$\boxed{M(y) = \sum_{v \in V} [p_v^T M(y) p_v] \; \zeta_v \zeta_v^T}$$

**Proof:** Denote by $N$ the RHS matrix.
As $\{p_v \mid v \in V\}$ is a basis of $\mathbb{R}[x_1, \ldots, x_n]/I$, it suffices to verify that

$$p_v^T M(y) p_{v'} = p_v^T N p_{v'} \text{ for all } v, v' \in V.$$

- Obvious if $v = v'$.
- If $v \neq v'$, $p_v^T N p_{v'} = 0$ and $p_v^T M(y) p_{v'} = 1^T M(y)(p_v p_{v'}) = 0$, since $p_v p_{v'} \in I$.

**Corollary 4:** The measure $\mu := \sum_{v \in V} [p_v^T M(y) p_v] \delta_v$ is a representing measure for $y$.

## The Flat Extension Theorem for truncated sequences

**Theorem B:** [Curto-Fialkow 1996]
Let $y = (y_\alpha)_{|\alpha| \le 2t}$ such that $M_t(y) \succeq 0$ and

$$\text{rank } M_t(y) = \text{rank } M_{t-1}(y)$$

[$M_t(y)$ is a *flat extension* of $M_{t-1}(y)$].
Then, $M_t(y)$ has a flat extension $M_{t+1}(y)$.

$$\Longrightarrow M_t(y) \text{ has a flat extension } M(y) \succeq 0$$

$$\Longrightarrow y \text{ has a representing measure}$$

**Application:** [Curto-Fialkow 1998]
A truncated sequence $y = (y_\alpha)_{|\alpha| \le 2t}$ has a representing measure with finite support
$\Longleftrightarrow M_t(y) \succeq 0$ has a positive extension $M_{t+k}(y)$ (for some $k \ge 1$) which in turn has a flat extension $M_{t+k+1}(y)$.

## The $F$-moment problem

When does $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ have a representing measure supported by the basic closed semi-algebraic set:

$$F = \{x \in \mathbb{R}^n \mid h_1(x) \geq 0, \ldots, h_m(x) \geq 0\}?$$

**Necessary conditions:** $M(y) \succeq 0$, $M(h_j y) \succeq 0$ $(\forall j)$

**The conditions are sufficient** when $F$ is compact and satisfies Putinar's condition.
In the truncated case, the conditions are sufficient when some flat extension condition is satisfied.

**Theorem C:** [Curto-Fialkow 2000]
Set $d := \max_j \lceil \deg(h_j)/2 \rceil$ and let $y = (y_\alpha)_{|\alpha| \leq 2t}$.
If $M_t(y) \succeq 0$ has a flat extension $M_{t+d}(y)$ and $M_t(h_j y) \succeq 0$ $(\forall j)$, then $y$ has a representing measure supported by $F$.

**Proof:** By Theorems A + B, $y$ has a representing measure supported by $V(I)$.

Remains to show that $V(I) \subseteq F$.

For this, choose interpolation polynomials $p_v$ at $v \in V(I)$ having degree at most $t$. Then,

$$0 \leq p_v^T M_t(h_j y) p_v = \int p_v(x)^2 h_j(x) d\mu(x)$$

$$\Longrightarrow h_j(v) \geq 0 \quad \forall j, \qquad \text{i.e., } v \in F$$

## Application to optimization

$$\begin{aligned} p^* &= \min p(x) \text{ s.t. } \overbrace{h_1(x) \geq 0, \ldots, h_m(x) \geq 0}^{x \in F} \\ &= \min p^T y \text{ s.t. } y \text{ has a repr. measure supported by } F \\ &= \max \rho \text{ s.t. } p(x) - \rho \geq 0 \text{ on } F. \end{aligned}$$

**Two dual bounds:** $\boxed{\rho_t^* \leq p_t^* \leq p^*}$

$$p_t^* := \min p^T y \text{ s.t. } M_t(y) \succeq 0,\ M_{t-d_j}(h_j y) \succeq 0\ \forall j, y_0 = 1$$

$$\begin{aligned} \rho_t^* := \max \rho \text{ s.t. } \quad & p(x) - \rho = u_0 + \sum_j u_j h_j \\ & u_0, u_j \in \Sigma,\ \deg(u_0), \deg(u_j h_j) \leq 2t \end{aligned}$$

• [Lasserre 2001] **Asymptotic convergence** to $p^*$ when $F$ is compact and satisfies Putinar's condition
• [Las 01][Lau 02] **Finite convergence** in the presence of equations defining a radical zero-dimensional ideal.

**Stopping criterion:** [Henrion-Lasserre 2003]
Let $y$ be an optimum solution to $p_t^*$.
If rank $M_t(y) =$ rank $M_{t-d}(y)$, then $p_t^* = p^*$.

**Proof:** By Theorem C, $y$ has a representing measure $\mu = \sum_{v \in V(I)} \lambda_v \delta_v$ supported by $V(I) \subseteq F$.
Then, $p_t^* = p^T y = \sum_v \lambda_v p(v) \geq p^* \Longrightarrow p_t^* = p^*$.

**Extracting a global minimizer:** Compute the points in $V(I)$, which are global minimizers of $p(x)$ over $F$.

## Alternative elementary proof for Theorem A using the spectral theorem for commutative symmetric matrices

(inspired by [Freedman-Lovász-Schrijver])

Assume $M(y) \succeq 0$, rank $M(y) = r$, set $I = \text{Ker } M(y)$.

• Equip the algebra $\mathcal{A} := \mathbb{R}[x_1, \ldots, x_n]/I$ with the inner product:

$$\langle p, q \rangle = p^T M(y) q$$

• W.r.t. an orthomal basis, the *multiplication operator*:

$$\begin{aligned} L_f : \mathcal{A} &\to \mathcal{A} \\ g &\mapsto fg \end{aligned}$$

has a symmetric matrix $M_f$

• The matrices $M_{x_1}, \ldots, M_{x_n}$ pairwise commute
$\Longrightarrow$ They have a common system $\{p_1, \ldots, p_r\}$ of *real* eigenvectors, forming a basis of $\mathcal{A}$.

• $p_i p_j = 0$ $(i \neq j)$, $p_i^2 = p_i$ [system of idempotents of $\mathcal{A}$]
$\langle 1, p_i \rangle = \langle 1, p_i^2 \rangle = \langle p_i, p_i \rangle \geq 0$

• Write $x_i = \sum_{\ell=1}^{r} \beta_\ell^{(i)} p_\ell$ and set $z_\ell := (\beta_\ell^{(1)}, \ldots, \beta_\ell^{(n)})$. Then,

$$x_1^{\alpha_1} \cdots x_n^{\alpha_n} = \sum_{\ell=1}^{r} (\beta_\ell^{(1)})^{\alpha_1} \cdots (\beta_\ell^{(n)})^{\alpha_n} p_\ell = \sum_{\ell=1}^{r} z_\ell^\alpha p_\ell$$

Therefore,

$$y_\alpha = \langle 1, x^\alpha \rangle = \sum_{\ell=1}^{r} \langle 1, p_\ell \rangle z_\ell^\alpha$$

Hence, $\mu := \sum_{\ell=1}^{r} \langle 1, p_\ell \rangle \delta_{z_\ell}$ is a representing measure for $y$.

# Some bibliography


C. Berg, J.P.R. Christensen, and P. Ressel. Positive definite functions on Abelian semigroups. *Math. Ann.* **223** (1976), 253–272.

R.E. Curto and L.A. Fialkow. Solution of the truncated complex moment problem for flat data. *Mem. Amer. Math. Soc.* vol. 568, Amer. Math. Soc., Providence, RI, 1996.

R.E. Curto and L.A. Fialkow. Flat extensions of positive moment matrices: recursively generated relations. *Mem. Amer. Math. Soc.* vol. 648, Amer. Math. Soc., Providence, RI, 1998.

R.E. Curto and L.A. Fialkow. The truncated complex $K$-moment problem. *Trans. Amer. Math. Soc.* **352** (2000), 2825–2855.

M. Freedman, L. Lovász, and A. Schrijver. Reflection positivity, rank connectivity and homomorphism of graphs. Preprint, 2004.

D. Henrion and J.-B. Lasserre. Detecting global optimality and extracting solutions in GloptiPoly. Preprint, 2003.

J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* **11** (2001), 796–817.

J.B. Lasserre. Polynomials nonnegative on a grid and discrete representations. *Trans. Amer. Math. Soc.* **354** (2001), 631–649.

M. Laurent. Revisiting two theorems of Curto and Fialkow on moment matrices. Preprint, January 2004. To appear in the Proceedings of the AMS.

M. Laurent. Semidefinite representations for finite varieties. Preprint, 2002.