Bounding the support of a measure from its marginal moments

Jean-Bernard Lasserre

BOUNDING THE SUPPORT OF A MEASURE FROM ITS MARGINAL MOMENTS

JEAN B. LASSERRE

Abstract. Given all moments of the marginals of a measure μ on $\mathbb{R}^n$, one provides (a) explicit bounds on its support and (b) a numerical scheme to compute the smallest box that contains the support of μ.

1. Introduction

Inverse problems in probability are ubiquitous in several important applications, and among them shape reconstruction problems. For instance, exact recovery of two-dimensional objects from finitely many moments is possible for polygons and so-called quadrature domains as shown in Golub et al. [2] and Gustafsson et al. [6], respectively. But so far, there is no inversion algorithm from moments for n-dimensional shapes. However, more recently Cuyt et al. [4] have shown how to approximately recover numerically an unknown density f defined on a compact region of $\mathbb{R}^n$, from the only knowledge of its moments. So when f is the indicator function of a compact set $A \subset \mathbb{R}^n$ one may thus recover the shape of A with good accuracy, based on moment information only. The elegant methodology developed in [4] is based on multi-dimensional homogeneous Padé approximants and uses a nice Padé *slice property*, the analogue for the moment approach of the Fourier slice theorem for the Radon transform (or projection) approach; see [4] for an illuminating discussion.

In this paper we are interested in the following inverse problem. Given an *arbitrary* finite Borel measure μ on $\mathbb{R}^n$ (not necessarily having a density with respect to the Lebesgue measure), can we compute (or at least approximate) the smallest box $\prod_{i=1}^n [a_i, b_i] \subset \mathbb{R}^n$ which contains the support of μ (not necessarily compact), from the only knowledge of its moments?

Contribution. Obviously, as we look for a box, the problem reduces to find for each $i = 1, \ldots, n$, the smallest interval $[a_i, b_i]$ (not necessarily compact) that contains the support of the marginal μ_i of μ. Of course, to bound a_i and b_i, one possibility is to compute zeros of the polynomials (p_d), $d \in \mathbb{N}$, orthogonal with respect to the measure μ_i. Indeed, for every d, the smallest (resp. largest) zero of p_d provides an upper bound on a_i (resp. a lower bound on b_i), and there is a systematic way to compute the p_d's from from the given moments of μ; see e.g. Gautschi [5, §1.2 and §2.1].

Our contribution is to provide a convergent numerical scheme for computing this smallest interval $[a_i, b_i]$, which (i) is based on the only knowledge of the moments of the marginals μ_i, $i = 1, \ldots, n$, and (ii) *avoids* computing orthogonal polynomials.

1991 *Mathematics Subject Classification.* 60B05 90C22.

Key words and phrases. inverse problems; moments of a measure; semidefinite programming.

For each i, it consists of solving 2 hierarchies (associated with each of the end points a_i and b_i) of so called *semidefinite programs*[1] in only one variable (and therefore those semidefinite programs are generalized eigenvalue problems for which even more specialized softwares exist). Importantly, we do not make any assumption on μ and in particular, μ may not have a density with respect to the Lebesgue measure as in the above cited works. In solving the two semidefinite programs at step d of the hierarchy, one provides an inner approximation $[a_d, b_d] \subset [a_i, b_i]$ such that the sequence (a_d) (resp. (b_d)), $d \in \mathbb{N}$, is monotone nonincreasing (resp. nondecreasing) and converges to a_i (resp. to b_i) as $d \to \infty$ (with possibly $a_i = -\infty$ and/or $b_i = +\infty$). Interestingly, some explicit upper (resp. lower) bounds on a_i (resp. b_i) in terms of the moments of μ_i are also available.

2. Notation and definitions

Let $\mathbb{N}$ be the set of natural numbers and denote by $\mathbf{x} = (x_1, \ldots, x_n)$ a vector of $\mathbb{R}^n$ whereas x will denote a scalar. Let $\mathbb{R}[x]$ be the ring of real univariate polynomials in the single variable x, and denote by $\mathbb{R}[x]_d$ the vector space of polynomials of degree at most d. Let $\Sigma[x] \subset \mathbb{R}[x]$ be the set of polynomials that are sums of squares (s.o.s.), and $\Sigma[x]_d$ its subspace of s.o.s. polynomials of degree at most $2d$. In the canonical basis (x^k), $k = 0, \ldots, d$, of $\mathbb{R}[x]_d$, a polynomial $f \in \mathbb{R}[x]_d$ is written

$$x \mapsto f(x) \,=\, \sum_{k=0}^{d} f_k \, x^k,$$

for some vector $\mathbf{f} = (f_k) \in \mathbb{R}^{d+1}$.

Moment matrix. Given a infinite sequence $\mathbf{y} := (y_k)$, $k \in \mathbb{N}$, indexed in the canonical basis (x^k) of $\mathbb{R}[x]$, let $\mathbf{H}_d(\mathbf{y}) \in \mathbb{R}^{(d+1)\times(d+1)}$ be the real Hankel matrix defined by:

$$\mathbf{H}_d(\mathbf{y})(i,j) \,=\, y_{i+j-2}, \qquad \forall\, i,j \le d. \tag{2.1}$$

The matrix $\mathbf{H}_d(\mathbf{y})$ is called the *moment* matrix associated with the sequence $\mathbf{y}$ (see e.g. Curto and Fialkow [3] and Lasserre [7]). If $\mathbf{y}$ has a *representing measure* μ (i.e., if there exists a finite Borel measure μ such that $y_k = \int \mathbf{x}^k d\mu$ for every $k \in \mathbb{N}$) then

$$\langle \mathbf{f}, \mathbf{H}_d(\mathbf{y})\mathbf{f} \rangle \,=\, \int f(x)^2 \, d\mu(x) \,\ge\, 0, \qquad \forall f \in \mathbb{R}[x]_d, \tag{2.2}$$

so that $\mathbf{H}_d(\mathbf{y}) \succeq 0$, where for a real symmetric matrix $\mathbf{A}$, the notation $\mathbf{A} \succeq 0$ (resp. $\mathbf{A} \succ 0$) stands for $\mathbf{A}$ is positive semidefinite (resp. positive definite).

Localizing matrix. Similarly, given $\theta \in \mathbb{R}[x]_s$ with vector of coefficients (θ_k), let $\mathbf{H}_d(\theta\,\mathbf{y}) \in \mathbb{R}^{(d+1)\times(d+1)}$ be the real symmetric matrix defined by:

$$\mathbf{H}_d(\theta\,\mathbf{y})(i,j) \,:=\, \sum_{k=0}^{s} \theta_k y_{i+j+k-2}, \qquad \forall\, i,j \le d. \tag{2.3}$$

The matrix $\mathbf{H}_d(\theta\,\mathbf{y})$ is called the *localizing* matrix associated with the sequence $\mathbf{y}$ and the polynomial θ (see again Lasserre [7]). Notice that the localizing matrix with respect to the constant polynomial $\theta \equiv 1$ is the moment matrix $\mathbf{H}_d(\mathbf{y})$ in

[1] A semidefinite program is a convex optimization problem that can be solved (up to arbitrary but fixed precision) in time polynomial in the input size of the problem, and for which efficient public softwares are available; see e.g. [9]

(2.1). If $\mathbf{y}$ has a representing measure μ with support contained in the level set $\{x \in \mathbb{R} : \theta(x) \geq 0\}$, then

$$\langle \mathbf{f}, \mathbf{H}_d(\theta\,\mathbf{y})\mathbf{f}\rangle \;=\; \int f(x)^2\theta(x)\,d\mu(x) \;\geq\; 0 \qquad \forall\, f \in \mathbb{R}[x]_d, \tag{2.4}$$

so that $\mathbf{H}_d(\theta\,\mathbf{y}) \succeq 0$.

Finally, for a finite Borel measure μ, denote its support by $\operatorname{supp}\mu$, that is, $\operatorname{supp}\mu$ is the smallest closed set B such that $\mu(B^c) = 0$ (where B^c denotes the complement of B).

3. Main result

We may and will restrict to the one-dimensional case because if μ is a finite Borel measure on $\mathbb{R}^n$, and if we look for a box $\mathbf{B} := \prod_{i=1}^n [a_i, b_i]$ such that $\operatorname{supp}\mu \subseteq \mathbf{B}$, then we have the following result. For every $i = 1, \ldots, n$, let μ_i be the marginal of μ with respect to the variable x_i.

Lemma 3.1. *Let $\mathbf{B} \subset \mathbb{R}^n$ be the box $\prod_{i=1}^n [a_i, b_i]$ with possibly $a_i = -\infty$ and/or $b_i = +\infty$. Then $\operatorname{supp}\mu \subseteq \mathbf{B}$ if and only if $\operatorname{supp}\mu_i \subseteq [a_i, b_i]$ for every $i = 1, \ldots, n$.*

Proof. For every $i = 1, \ldots, n$, let $A_i \subset \mathbb{R}^n$ be the Borel set $\{\mathbf{x} \in \mathbb{R}^n : x_i \in \mathbb{R} \setminus [a_i, b_i]\}$. Then $0 = \mu(A_i) = \mu_i(\mathbb{R} \setminus [a_i, b_i])$, shows that $\operatorname{supp}\mu_i \subseteq [a_i, b_i]$, $i = 1, \ldots, n$. Conversely, if $\mu_i \subseteq [a_i, b_i]$, $i = 1, \ldots, n$, then $0 = \mu_i(\mathbb{R} \setminus [a_i, b_i]) = \mu(A_i)$, $i = 1, \ldots, n$. But since $\mathbf{B}^c = \cup_{i=1}^n A_i$ we also obtain $\mu(\mathbf{B}^c) = 0$. □

So Lemma 3.1 tells us that it is enough to consider separate conditions for the marginals μ_i, $i = 1, \ldots, n$. Therefore, all we need to know is the sequence of moments

$$\mathbf{y}_k^i := \int x_i^k\, d\mu(\mathbf{x}) \;=\; \int x^k\, d\mu_i(x), \qquad k = 0, 1, \ldots$$

of the marginal μ_i of μ, for every $i = 1, \ldots, n$.

Hence we now consider the one-dimensional case. For a real number a, let $\theta_a \in \mathbb{R}[x]$ be the polynomial $x \mapsto \theta_a(x) = (x - a)$. Recall that the support of a finite Borel measure μ on $\mathbb{R}$ (denoted $\operatorname{supp}\mu$) is the smallest closed set B such that $\mu(\mathbb{R} \setminus B) = 0$. For instance is μ is supported on $(a, b] \cup [c, d) \cup \{e\}$, with $a < b < c < d < e$ then $\operatorname{supp}\mu$ is the closed set $[a, b] \cup [c, d] \cup \{e\}$ and is contained in the interval $[a, e]$.

3.1. **Bounds on $\operatorname{supp}\mu$.** We first derive bounds on scalars a and b that satisfy $\operatorname{supp}\mu \subseteq (-\infty, a]$ and/or $\operatorname{supp}\mu \subseteq [b, +\infty)$.

Proposition 3.2. *Let μ be a finite and non trivial Borel measure on the real line $\mathbb{R}$, with associated sequence of moments $\mathbf{y} = (y_k)$, $k \in \mathbb{N}$, all finite. Then:*

(i) $\operatorname{supp}\mu \subseteq [a, +\infty)$ *if and only if*

$$\mathbf{H}_d(\theta_a\,\mathbf{y}) \succeq 0, \qquad \forall d \in \mathbb{N}. \tag{3.1}$$

(ii) $\operatorname{supp}\mu \subseteq [-\infty, b]$ *if and only if*

$$\mathbf{H}_d(-\theta_b\,\mathbf{y}) \succeq 0, \qquad \forall d \in \mathbb{N}. \tag{3.2}$$

For each fixed $d \in \mathbb{N}$, the condition $\mathbf{H}_d(\theta_a\,\mathbf{y}) \succeq 0$ (resp. $\mathbf{H}_d(-\theta_b\,\mathbf{y}) \succeq 0$) determines a basic semi-algebraic set of the form $\{(a, \mathbf{y}) : p_{dk}(a, \mathbf{y}) \geq 0, k = 0, \ldots d - 1\}$

(resp. $\{(a,\mathbf{y}) : (-1)^{d-k}p_{dk}(b,\mathbf{y}) \geq 0,\, k = 0,\ldots d-1\}$*) for some polynomials* $(p_{dk}) \subset \mathbb{R}[x,\mathbf{y}]$.

With $\mathbf{y}$ *and* d *fixed, the condition (3.1) (resp. (3.2)) yields an upper bound* $a \leq \underline{a}_d$ *(resp. a lower bound* $b \geq \overline{b}_d$*), and the sequence* $(\underline{a}_d)$ *(resp.* $(\overline{b}_d)$*),* $d \in \mathbb{N}$*, is monotone nonincreasing (resp. nondecreasing). In particular,*

$$(3.3)\quad a \;\leq\; \min\left[\frac{y_1}{y_0}, \frac{y_1+y_3}{y_0+y_2}\right] \quad \textit{and} \quad a \leq \min\left[\frac{y_1}{y_0}, \frac{y_3}{y_2}, \frac{y_1+y_3}{y_0+y_2}\right] \;\; \textit{if } y_2 \neq 0$$

$$(3.4)\quad b \;\geq\; \max\left[\frac{y_1}{y_0}, \frac{y_1+y_3}{y_0+y_2}\right] \quad \textit{and} \quad b \geq \max\left[\frac{y_1}{y_0}, \frac{y_3}{y_2}, \frac{y_1+y_3}{y_0+y_2}\right] \;\; \textit{if } y_2 \neq 0,$$

as well as

$$(3.5)\qquad a \;\leq\; \frac{y_0y_3 - y_1y_2 - \sqrt{(y_0y_3-y_1y_2)^2 - 4(y_0y_2-y_1^2)(y_1y_3-y_2^2)}}{2(y_0y_2-y_1^2)}$$

$$(3.6)\qquad b \;\geq\; \frac{y_0y_3 - y_1y_2 + \sqrt{(y_0y_3-y_1y_2)^2 - 4(y_0y_2-y_1^2)(y_1y_3-y_2^2)}}{2(y_0y_2-y_1^2)}$$

if $(y_0y_3-y_1y_2)^2 \geq 4(y_0y_2-y_1^2)(y_1y_3-y_2^2)$.

Proof. (3.1) and (3.2) is well-known and can be found in e.g. Lasserre [7, Theorem 3.2]. Next, write the characteristic polynomial $t \mapsto c(t)$ of $\mathbf{H}_d(\theta_a,\mathbf{y})$ in the form

$$c(t)\;(=\det(tI - \mathbf{H}_d(\theta_a\,\mathbf{y}))) \;=\; t^d + \sum_{k=0}^{d-1}(-1)^{d-k}p_{dk}(a,\mathbf{y})\,t^k, \qquad t \in \mathbb{R},$$

for some polynomials $(p_{dk}) \subset \mathbb{R}[a,\mathbf{y}]$, and where p_{dk} is of degree k in the variable a. Then $\mathbf{H}_d(a\,\mathbf{y}) \succeq 0$, if and only if c has all its roots nonnegative, which in turn, by Descarte's rule, happens if and only if $p_{dk}(a,\mathbf{y}) \geq 0$, for every $k = 0,\ldots,d-1$. In fact $(a,\mathbf{y})$ belong to the closure of a convex connected component of $\{(a,\mathbf{y}) : p_{d0}(a,\mathbf{y})\,(=\det H_d(\theta_a\,\mathbf{y})) > 0\}$. A similar argument applies for (3.2) with now the characteristic polynomial

$$\tilde{c}(t) \;=\; \det(tI - \mathbf{H}_d(-\theta_b\,\mathbf{y})) \;=\; (-1)^d c(-t) \;=\; t^d + \sum_{k=0}^{d-1} p_{dk}(b,\mathbf{y})\,t^k, \qquad t \in \mathbb{R}.$$

So with $\mathbf{y}$ and d fixed, $a \mapsto p_{dk}(a,\mathbf{y})$ is a univariate polynomial for every k, and so the conditions (3.1) provide a bound of the form $a \leq \underline{a}_d$ for some $\underline{a}_d$ since if a satisfies (3.1) then so does $a' \leq a$. Similarly, the conditions (3.2) provide a bound of the form $b \geq \overline{b}_d$ since if b satisfies (3.2) then so does $b' \geq b$. The scalar $\underline{a}_d$ (resp. $\overline{b}_d$) may be taken as the smallest (resp. largest) root of the polynomial $x \mapsto p_{d0}(x,\mathbf{y})$; bounds in terms of the coefficients of $p_{d0}(x,\mathbf{y})$ are available in the literature.

Finally (3.3)-(3.6) are obtained with $d = 1$ in which case (3.1) and (3.2) read:

$$\begin{bmatrix} y_1 - ay_0 & y_2 - ay_1 \\ y_2 - ay_1 & y_3 - ay_2 \end{bmatrix} \succeq 0; \qquad \begin{bmatrix} by_0 - y_1 & by_1 - y_2 \\ by_1 - y_2 & by_2 - y_3 \end{bmatrix} \succeq 0.$$

□

3.2. **Computing the smallest interval $[a,b] \supseteq \operatorname{supp}\mu$.**

Theorem 3.2 provides bounds (some of them explicit) in terms of bounds on the largest (or smallest) root of some univariate polynomial whose coefficients are polynomials in $\mathbf{y}$. But one may also get numerical sharp bounds via solving the following sequence of semidefinite programs, indexed by d:

$$a_d = \max_a \{\, a \,:\, \mathbf{H}_d(\theta_a\,\mathbf{y}) \succeq 0 \,\} \tag{3.7}$$

$$b_d = \min_b \{\, b \,:\, \mathbf{H}_d(\theta_b\,\mathbf{y}) \succeq 0 \,\}, \tag{3.8}$$

where $\mathbf{H}_d(\theta_*\,\mathbf{y})$ is the localizing matrix associated with the polynomial $\theta_* \in \mathbb{R}[x]$. Observe that (3.7) and (3.8) are semidefinite programs with only one variable! For instance for $d=1$, (3.7) reads

$$a_1 = \max_a \left\{ a \,:\, \begin{bmatrix} y_1 - ay_0 & y_2 - ay_1 \\ y_2 - ay_1 & y_3 - ay_2 \end{bmatrix} \succeq 0 \right\},$$

whereas (3.8) reads

$$b_1 = \min_b \left\{ b \,:\, \begin{bmatrix} by_0 - y_1 & by_1 - y_2 \\ by_1 - y_2 & by_2 - y_3 \end{bmatrix} \succeq 0 \right\}$$

For more details on semidefinite programming the interested reader is referred to e.g. [9]. And we obtain:

Theorem 3.3. *Let μ be a finite Borel measure with all moments $\mathbf{y} = (y_k)$ finite. Then $\operatorname{supp}\mu \subseteq [a^*, b^*]$, with possibly $a^* = -\infty$ and/or $b^* = +\infty$, and where:*

(i) a_d is an optimal solution of (3.7) for all $d \in \mathbb{N}$, and the sequence (a_d), $d \in \mathbb{N}$, is monotone nonincreasing with $a_d \downarrow a^*$ as $d \to \infty$.

(ii) b_d is an optimal solution of (3.8) for all $d \in \mathbb{N}$, and and the sequence (b_d), $d \in \mathbb{N}$, is monotone nondecreasing with $b_d \uparrow a^*$ as $d \to \infty$.

(iii) a^* (resp. b^*) is the largest (resp. smallest) scalar such that $\operatorname{supp}\mu \subseteq [a^*, b^*]$.

Proof. We prove the statements for (i) and (iii) only because similar arguments hold for (ii). We first prove that (3.7) has always a feasible solution. If $\operatorname{supp}\mu \subset [a, +\infty)$ for some $a > -\infty$, then a is obviously feasible for the semidefinite program (3.7), for every $d \in \mathbb{N}$. If there is no such a, consider the finite sequence of moments $\mathbf{y}_d = (y_0, \ldots, y_{2d+1})$. By Tchakaloff's theorem (see e.g. [1, 8], [7, Theorem B.12]) there exists a measure ν supported on finitely many points $x_0 \le x_1 \le \cdots \le x_t$, with $t \le 2d+2$ (hence $\operatorname{supp}\nu = \cup_{i=0}^t \{x_i\} \subset [x_0, +\infty)$), and with same moments as μ, up to degree $2d+1$. Hence in view of what precedes, x_0 is feasible for (3.7). Next, as every feasible solution is bounded above by y_1/y_0 and as we maximize, it follows that (3.7) has an optimal solution a_d for every $d \in \mathbb{N}$.

Next, observe that $a_d \le a_k$ whenever $d \ge k$ because the feasible set of (3.7) for d is contained in that for k and every feasible solution is bounded above by y_1/y_0. Therefore the sequence (a_d), $d \in \mathbb{N}$, is monotone nonincreasing and thus, converges to a^* with possibly $a^* = -\infty$.

If $a^* = -\infty$ then there is no a such that $\operatorname{supp}\mu \subseteq [a, +\infty)$ because we would have $a_d \ge a$ for all d. Next, consider the case $a^* > -\infty$, and let $d \in \mathbb{N}$ be fixed. Using $\mathbf{H}_d(\theta_{a_d}\,\mathbf{y}) \succeq 0$ and the continuity of $a \mapsto \mathbf{H}_d(\theta_a\,\mathbf{y})$, one obtains $\mathbf{H}_d(\theta_{a^*}\,\mathbf{y}) \succeq 0$. As d fixed was arbitrary, we then obtain $\mathbf{H}_d(\theta_{a^*}\,\mathbf{y}) \succeq 0$ for every $d \in \mathbb{N}$. But then by [7, Theorem 3.2], μ is supported on the set $\{x \,:\, \theta_{a^*}(x) \ge 0\}$, which shows that $\operatorname{supp}\mu \subseteq [a^*, +\infty)$.

Concerning (iii), if $a^* > -\infty$ then a^* is the largest a such that $\operatorname{supp}\mu \subseteq [a, +\infty)$ because if $\operatorname{supp}\mu \subseteq [a, +\infty)$ then a is feasible for (3.7), for every $d \in \mathbb{N}$; therefore, $a \leq a_d$ for every d, which in turn implies $a \leq a^*$. □

Next, in the case where μ is known to have compact support one may even consider the following single hierarchy of semidefinite programs

$$\rho_d = \min_{b,a} \{\, b - a \,:\, \mathbf{H}_d(\theta_a\,\mathbf{y}),\ \mathbf{H}_d(-\theta_b\,\mathbf{y}) \succeq 0\,\}, \tag{3.9}$$

indexed by d, and with now two variables a and b. We obtain the following result of which the proof is omitted.

Corollary 3.4. *Assume that μ has compact support. Then:*

(a) *The semidefinite program (3.9) has an optimal solution (a_d, b_d) for every $d \in \mathbb{N}$.*

(b) *Let (a_d, b_d), $d \in \mathbb{N}$, be a sequence of optimal solutions of (3.9). As $d \to \infty$, $(a_d, b_d) \to (a^*, b^*)$ and $\operatorname{supp}\mu \subseteq [a^*, b^*]$. Moreover, $[a^*, b^*]$ is the smallest interval which contains $\operatorname{supp}\mu$ and if the support of μ is an interval then $\operatorname{supp}\mu = [a^*, b^*]$.*

3.3. **Duality.** We interpret the dual of the semidefinite program (3.7). Let $\mathcal{S}_d \subset \mathbb{R}^{(d+1)\times(d+1)}$ be the cone of real symmetric matrices. The dual of (3.7) is the semidefinite program:

$$\begin{aligned} a_d^* = \min_{\mathbf{Z}\in\mathcal{S}_d}\ & \langle \mathbf{H}_d(x\,\mathbf{y}), \mathbf{Z}\rangle \\ \text{s.t.}\ & \langle \mathbf{H}_d(\mathbf{y}), \mathbf{Z}\rangle \,=\, 1\,;\ \mathbf{Z} \succeq 0. \end{aligned} \tag{3.10}$$

If $\mathbf{Z} \succeq 0$ is feasible, using the singular value decomposition of $\mathbf{Z}$ one may write $\mathbf{Z} = \sum_i \mathbf{f}_i\mathbf{f}_i^T$ for some vectors $(\mathbf{f}_i) \subset \mathbb{R}^{d+1}$, and so

$$1 = \langle \mathbf{H}_d(\mathbf{y}), \mathbf{Z}\rangle \,=\, \sum_i \langle \mathbf{f}_i, \mathbf{H}_d(\mathbf{y})\mathbf{f}_i\rangle = \sum_i \int f_i^2\, d\mu \,=\, \int \sigma\, d\mu,$$

with $\sigma = \sum_i f_i^2 \in \Sigma[x]_d$, and similarly

$$\langle \mathbf{H}_d(x\,\mathbf{y}), \mathbf{Z}\rangle \,=\, \sum_i \langle \mathbf{f}_i, \mathbf{H}_d(x\,\mathbf{y})\mathbf{f}_i\rangle = \sum_i \int x f_i(x)^2 d\mu(x) \,=\, \int x\sigma(x) d\mu(x).$$

Therefore, equivalently, (3.10) reads

$$a_d^* = \min_{\sigma\in\Sigma[x]_d} \left\{ \int x\, \underbrace{\sigma(x)d\mu(x)}_{d\nu(x)} \,:\, \int \underbrace{\sigma\, d\mu}_{d\nu} \,=\, 1 \right\}, \tag{3.11}$$

and $a_d^* \geq a_d$ (which is called *weak duality*). Indeed, for any two feasible solutions $a, \mathbf{Z}$ of (3.9) and (3.10) respectively, using $\mathbf{Z} \succeq 0$ and $\mathbf{H}_d(\theta_a\,\mathbf{y}) \succeq 0$, yields

$$0 \leq \langle \mathbf{Z}, \mathbf{H}_d(\theta_a\,\mathbf{y})\rangle \,=\, \int (x-a)\,\sigma(x)\,d\mu(x) \,=\, \int x\sigma(x)\,d\mu(x) - a,$$

that is, $a \leq \int x\sigma(x)d\mu(x)$. So in the dual semidefinite program (3.11), one searches for a sum of squares polynomial $\sigma \in \Sigma[x]_d$ of degree at most $2d$ (normalized to satisfy $\int \sigma d\mu = 1$), which minimizes $\int x\sigma d\mu$. Equivalently, one searches for a probability measure ν with density $\sigma \in \Sigma[x]_d$ with respect to μ, which minimizes the upper bound $\int x d\nu$ on the global minimum of x on the support of μ.

Similarly, the dual of (3.8) is the semidefinite program:

$$b_d^* = \max_{\sigma \in \Sigma[x]_d} \left\{ \int x \underbrace{\sigma(x) d\mu(x)}_{d\nu(x)} \,:\, \int \underbrace{\sigma\, d\mu}_{d\nu} = 1. \right\}. \tag{3.12}$$

Again, by weak duality, $b_d^* \le b_d$ and in (3.12) one searches for a probability measure ν with density $\sigma \in \Sigma[x]_d$ with respect to μ, which maximizes the lower bound $\int x d\nu$ on the global maximum of x on the support of μ.

Theorem 3.5. *Suppose that μ is such that $\mathbf{H}_d(\mathbf{y}) \succ 0$ for all d (for instance if μ has no atom). Then there is no duality gap between (3.7) and (3.11) (resp. (3.8) and (3.12)), i.e. $a_d = a_d^*$ (resp. $b_d = b_d^*$). In addition (3.11) (resp. (3.12)) has an optimal solution $\sigma^* \in \Sigma[x]_d$ (resp. $\psi^* \in \Sigma[x]_d$), and*

$$\int (x - a_d)\, \sigma^*(x)\, d\mu(x) \,=\, 0 \,=\, \int (b_d - x)\, \psi^*(x)\, d\mu(x). \tag{3.13}$$

Proof. From Theorem 3.3, we know that (3.7) has an optimal solution a_d. By Tchakaloff's theorem, let ν be the measure supported on the finitely many points $x_0, \ldots, x_t$ (with $t \le 2d+2$), and with same moments as μ, up to degree $2d+1$. There are positive weights λ_k, $k = 1, \ldots, t$, such that

$$\int p d\mu = \int p d\nu = \sum_{k=0}^{t} \lambda_k p(x_k) \qquad \forall\, p \in \mathbb{R}[x]_{2d+1}$$

(see e.g. [1, 8]). Hence $\mathbf{H}_d(\theta_a\, \mathbf{y}) \succ 0$ for every $a < x_0$, because for every $\mathbf{f}\,(\neq 0) \in \mathbb{R}^{d+1}$ (hence every $f\,(\neq 0) \in \mathbb{R}[x]_d$), $\langle \mathbf{f}, \mathbf{M}_d(\mathbf{y})\mathbf{f}\rangle = \sum_k \lambda_k f(x_k)^2 > 0$, and so

$$\begin{aligned} \langle \mathbf{f}, \mathbf{H}_d(\theta_a\, \mathbf{y})\mathbf{f}\rangle \;&=\; \int f(x)^2 (x-a)\, d\mu(x) \\ &=\; \int f(x)^2 (x-a)\, d\nu(x) \,=\, \sum_{k=0}^{t} f(x_k)^2 (x_k - a)\lambda_k \,>\, 0. \end{aligned}$$

Hence every $a < x_0$ is strictly feasible for (3.7), that is, Slater's condition[2] holds. But this implies that there is no duality gap, i.e., $a_d = a_d^*$, and in addition, the dual (3.11) has an optimal solution σ^*; see e.g. [9]. For same reasons, $b_d = b_d^*$ and (3.12) has an optimal solution. Therefore,

$$0 \,=\, \int x\sigma^*(x) d\mu(x) - a_d^* \,=\, \int (x - a_d^*)\sigma^*(x) d\mu(x) \,=\, \int (x - a_d)\sigma^*(x) d\mu(x),$$

and similarly,

$$0 \,=\, b_d^* - \int x\psi^*(x) d\mu(x) \,=\, \int (b_d^* - x)\psi^*(x) d\mu(x) \,=\, \int (b_d - x)\psi^*(x) d\mu(x),$$

which is the desired result (3.13). □

In Theorem 3.5, write $\sigma^* \in \Sigma[x]_d$ as $\sigma^* = \sum_\ell p_\ell^2$ for some polynomials $(p_\ell) \subset \mathbb{R}[x]_d$, with respective coefiicient vectors $\mathbf{p}_d \in \mathbb{R}^{d+1}$. Then by (3.13)

$$\int \sigma^*(x)(x - a_d) d\mu(x) \,=\, \sum_\ell \int p_\ell(x)^2 (x - a_d) d\mu(x) \,=\, \sum_\ell \langle \mathbf{p}_\ell, \mathbf{H}_d(\theta_{a_d}\, \mathbf{y})\mathbf{p}_\ell\rangle = 0,$$

[2]For a convex optimization problem $\min_{\mathbf{x}}\{\, f(\mathbf{x}) \,:\, g_j(\mathbf{x}) \ge 0,\, j = 1, \ldots, m\}$, Slater's condition states that there exists $\mathbf{x}_0$ such that $g_j(\mathbf{x}_0) > 0$ for every $j = 1, \ldots, m$.

so that for every ℓ, $\langle \mathbf{p}_\ell, \mathbf{H}_d(\theta_{a_d}\,\mathbf{y})\mathbf{p}_\ell\rangle = 0$ (since $\mathbf{H}_d(\theta_{a_d}\,\mathbf{y}) \succeq 0$), that is, every $\mathbf{p}_\ell$ is in the kernel of $\mathbf{H}_d(\theta_{a_d}\,\mathbf{y})$.

References

[1] C. Bayer and J. Teichmann. The proof of Tchakaloff's theorem, Proc. Amer. Math. Soc. **134** (2006), pp. 3035–3040.

[2] G.H. Golub, P. Milanfar, and J. Varah. A stable numerical method for inverting shape from moments, SIAM J. Sci. Comput. **21** (1999), pp. 1222–1243.

[3] R.E. Curto and L.A. Fialkow. Recursiveness, positivity, and truncated moment problems, Houston J. Math. **17** (1991), pp. 603–635.

[4] A. Cuyt, G.H. Golub, P. Milanfar, and B. Verdonk. Multidimensional integral inversion with application in shape reconstruction, SIAM J. Sci. Comput. **27** (2005), pp. 1058–1070.

[5] W. Gautschi. *Orthogonal Polynomials: Computation and Approximation*, Oxford University Press, 2004.

[6] B. Gustafsson, C. He, P. Milanfar, and M. Putinar. Reconstructing planar domains from their moments, Inverse Problems **16** (2000), pp. 1053–1070.

[7] J.B. Lasserre. *Moments, Positive Polynomials and Their Applications*, Imperial College Press, London, 2009.

[8] M. Putinar. A note on Tchakaloff's theorem, Proc. Amer. Math. Soc. **125** (1997), pp. 2409–2414.

[9] L. Vandenberghe and S. Boyd. Semidefinite programming, SIAM Rev. **38** (1996), pp. 49?95.

LAAS-CNRS and Institute of Mathematics, University of Toulouse, LAAS, 7 avenue du Colonel Roche, 31077 Toulouse Cédex 4,France

E-mail address: lasserre@laas.fr