Journal of Mathematical Sciences, Vol. 116, No. 3, 2003

THE METHOD OF MOMENTS IN GLOBAL OPTIMIZATION

R. Meziat UDC 517.977

In this paper, we present a general development, called the *method of moments,* which allows one to analyze global optimization problems in which the objective function is given as a linear combination of simpler functions. The essential feature of the method is its ability to provide nontrivial necessary and sufficient conditions for global minima. To illustrate the success of this approach, we exhibit an alternative way of finding global minima of polynomials by solving semidefinite programs.

1. Introduction

A major goal in global optimization is to find the global minima of a function f defined on a subset Ω of the Euclidean space $\mathbb{R}^v$. One approach to this problem comes from convex analysis, since we can use the convex envelope of the function f in order to locate its global minima. It is well known that every convex combination of points in Ω can be described as a discrete probability distribution μ supported in Ω such that every integral

$$\int_\Omega f(s)\,d\mu(s)$$

represents one point over the convex envelope of the function f. For this reason, we study the relaxed problem

$$\min_\mu \int_\Omega f(s)\,d\mu(s) \tag{1.1}$$

in order to find the global minima of the objective function f in Ω. For a thorough treatise on convex analysis, see Rockafellar's classical text [15].

It is true that relaxed problem (1.1) contains information about all the global minima of the function f in Ω. However, it cannot be solved easily in practice: consider the huge difficulty of describing all possible convex combinations of points in Ω. This is precisely the point that we will try to clarify here. We show how to transform problem (1.1) in order to make it more manageable. This can be done when we can express the objective function f as the composition of two functions:

$$f = g \circ F, \tag{1.2}$$

where g is linear and F is continuous. The analysis of this kind of situation is what we call the *method of moments*, since it is possible to use form (1.2) to express relaxed problem (1.1) as a programming problem with such nice properties as convexity and linearity. We will explore this general idea in Sec. 2.

Since the function g is a linear functional, every integral in (1.1) takes the form

$$\int_\Omega f(s)\,d\mu(s) = \int_\Omega g \circ F(s)\,d\mu(s) = g\left(\int_\Omega F(s)\,d\mu(s)\right).$$

Thus, we must study only the convex hull of the image of the set $\Omega \subset \mathbb{R}^v$ under the transformation $F : \Omega \to \mathbb{R}^k$. It is easy to see that

$$\operatorname{co}(F(\Omega)) = \left\{\int_\Omega F(s)\,d\mu(s) : \mu \text{ is a probability distribution}\right\},$$

Translated from Itogi Nauki i Tekhniki, Seriya Sovremennaya Matematika i Ee Prilozheniya. Tematicheskie Obzory. Vol. 94, Optimization and Related Topics–3, 2001.

where co stands for convex hull. If we were able to properly describe the convex set $\mathrm{co}(F(\Omega))$, then we would only need to study the convex problem

$$\min \{g(x) : x \in \mathrm{co}(F(\Omega))\}.$$

Assuming that transformation $F : \Omega \to \mathbb{R}^k$ is defined as

$$F(t) = (\psi_1(t), \ldots, \psi_k(t)),$$

where $\{\psi_1, \ldots, \psi_k\}$ is a linear independent set of scalar functions, the convex hull $\mathrm{co}(F(\Omega))$ is precisely the set of vectors $(m_1, \ldots, m_k)$ composed of the generalized moments

$$m_i = \int_\Omega \psi_i(s)\, d\mu(s) \quad \forall\, i = 1, \ldots, k,$$

of every probability measure μ with respect to the function set $\{\psi_1, \ldots, \psi_k\}$. From this standpoint, to describe the convex set $\mathrm{co}(F(\Omega))$, we need to solve the problem of moments for the basis $\{\psi_1, \ldots, \psi_k\}$. This is a classical problem, which is well known for a wide class of standard bases, such as the algebraic system of integer exponents and the trigonometric system of complex exponents. In this work, we show how the global minima of different kinds of functions can be analyzed by using known results on the problem of moments (see [17] for an interesting review of the most important topics within the problem of moments). Modern views on this matter can be found in [3, 12]. In addition, Shor's work on positive polynomials [16, Chap. 9] stretches the span of methods proposed here.

In Sec. 3, we describe some relevant results on the theory of moments, which are used in Sec. 4 for the analysis of global minima of polynomials. Indeed, as an illustration of the general method, we consider the special case in which the function f is a polynomial and hence it can be expressed as a linear combination of simpler functions:

$$f = \sum_i c_i\, \psi_i, \tag{1.3}$$

where the function basis $\{\psi_i\}$ can be the algebraic system $\psi_i = t^i$ or the trigonometric system $\psi_i = e^{jit}$. Thus, every integral in (1.1) can be expressed as

$$\int_\Omega f(s)\, d\mu(s) = \sum_{i=1}^k c_i\, m_i,$$

where $m_1, \ldots, m_k$ are the moments of the measure μ. Therefore, relaxed problem (1.1) can be posed as the programming problem

$$\min_{m_i} \sum_{i=1}^k c_i\, m_i, \tag{1.4}$$

where the variables $m_1, \ldots, m_k$ must be restricted to be the moments of some probability measure. Since relaxed problem (1.1) is equivalent to the global minimization of the function f in Ω, we hope that programming problem (1.4) also yields the global minima for f. This idea is developed in Sec. 4 for general algebraic polynomials, general trigonometric polynomials, and bidimensional two-degree polynomials.

In this paper, we clarify how the solutions for programming problem (1.4) can provide the global minima for general objective function (1.3). Indeed, we show nontrivial necessary and sufficient conditions to characterize the global minima of (1.3) which involve the solutions of optimization problem (1.4). To attain this goal, we must analyze the equivalence between relaxed problem (1.1) and programming problem (1.4). As a particular case, we can characterize the global minima of polynomials by solving convex optimization problems like (1.4) whose variables $m_1, \ldots, m_k$ are the entries of certain positive semidefinite matrices.

These kind of problems are called *semidefinite programming* problems since they are convex problems whose constraints are given by linear matrix inequalities. These problems arose in control theory, and many standard optimization problems can be recast as semidefinite problems. During the last decade, extensive research activity on the theory, applications, and numerical methods for these kinds of optimization problems took place. In addition, there exist many algorithms and routines to solve them (see [6, 7, 13, 16]).

2. The General Theory of The Method of Moments

Let $f : \Omega \to \mathbb{R}$ be a continuous function defined on a subset Ω of the Euclidean space $\mathbb{R}^v$. Our goal is to characterize the *global minima* of the function f in Ω. This means that we search for points t_0 such that

$$f(t_0) \le f(t) \quad \forall\, t \in \Omega.$$

In order to find these points, we can change the analysis of the optimization problem

$$\min_{t\in\Omega} f(t) \tag{2.1}$$

and study the relaxed problem

$$\min_{\mu\in \Pr(\Omega)} \int_\Omega f(s)\, d\mu(s), \tag{2.2}$$

where $\Pr(\Omega)$ represents the family of all probability measures supported in Ω. Next, we present some facts about the equivalence between problems (2.1) and (2.2).

If c is a lower bound for f in Ω, then

$$c \le \int_\Omega f(s)\, d\mu(s)$$

for every probability measure μ supported in Ω. Taking the Dirac measure $v = \delta_t$, we obtain

$$f(t) = \int_\Omega f(s)\, dv(s)$$

for every $t \in \Omega$. Therefore, we have proved that

$$\inf_{t\in\Omega} f(t) = \inf_{\mu\in \Pr(\Omega)} \int_\Omega f(s)\, d\mu(s). \tag{2.3}$$

This result implies the following proposition.

Proposition 1. *If $\mathcal{G}$ is the set of global minima for f in Ω, then every probability measure supported in $\mathcal{G}$ solves relaxed problem* (2.2)*. Conversely, if the probability measure μ^* satisfies the condition*

$$\int_\Omega f(s)\, d\mu^*(s) = \min_{\mu\in \Pr(\Omega)} \int_\Omega f(s)\, d\mu(s), \tag{2.4}$$

then every point in the support of μ^* is a global minimum for f in Ω.

Proof. It is easy to see that

$$f(s) = \min_{t\in\Omega} f(t) \quad \forall\, s \in \mathcal{G},$$

hence

$$\int_\Omega f(s)\, d\mu^*(s) = \int_{\mathcal{G}} f(s)\, d\mu^*(s) = \min_{t\in\Omega} f(t) \int_\Omega d\mu^*(s) = \min_{\mu\in \Pr(\Omega)} \int_\Omega f(s)\, d\mu(s)$$

for every probability measure μ^* supported in $\mathcal{G}$. On the other hand, we have

$$f(s) \ge \min_{t\in\Omega} f(t) \quad \forall\, s \in \Omega;$$

therefore,

$$\int_\Omega f(s)\, d\mu(s) \ge \min_{t\in\Omega} f(t)$$

for every probability measure μ whose support is contained in Ω. If there exists some point s_0 in the support of μ such that

$$f(s_0) > \min_{t\in\Omega} f(t),$$

then

$$\int_\Omega f(s)\, d\mu(s) > \min_{t\in\Omega} f(t)$$

and hence μ could not satisfy (2.4). □

These results show that there exists a theoretical equivalence between minimization problem (2.1) and relaxed problem (2.2). Now we endeavor to make such an equivalence explicit and useful.

2.1. Linear function decomposition. Assume that the objective function $f : \Omega \to R$ can be decomposed as

$$f = g \circ F, \tag{2.5}$$

where the function g is linear and the function F is continuous. Thus, every integral in relaxed problem (2.2) satisfies

$$\int_\Omega f(s)\, d\mu(s) = \int_\Omega g(F(s))\, d\mu(s) = g\left(\int_\Omega F(s)\, d\mu(s)\right).$$

On the other hand, it is easy to see that every integral

$$\int_\Omega F(s)\, d\mu(s)$$

represents a point within the convex hull of the image of the set Ω under the function F; therefore, we have

$$\inf_{t\in\Omega} f(t) = \inf\{g(x) : x \in \mathrm{co}(F(\Omega))\}. \tag{2.6}$$

Next, we explore the possibility for the new convex optimization problem

$$\min\{g(x) : x \in \mathrm{co}(F(\Omega))\} \tag{2.7}$$

to represent the original global optimization problem (2.1). It is very important to note that problem (2.1) can be a highly nonconvex problem, whereas proposed problem (2.7) has a linear objective function and a convex feasible set.

Proposition 2. *Assume that x_0 solves the optimization problem*

$$\min\{g(x) : x \in \mathrm{co}(F(\Omega))\}.$$

If there exist values $\lambda_i \geq 0$ and points $t_i \in \Omega$ such that

$$x_0 = \lambda_1 F(t_1) + \cdots + \lambda_r F(t_r)$$

with

$$\sum_{i=1}^{r} \lambda_i = 1,$$

then all t_i points are global minima for f in Ω.

Proof. Let μ^* be the probability measure defined as

$$\mu^* = \sum_{i=1}^{r} \lambda_i \delta_{t_i},$$

whose support is the set $\{t_1, \ldots, t_r\}$. Then we have

$$\int_\Omega f(t)\, d\mu^*(t) = g\left(\int_\Omega F(t)\, d\mu^*(t)\right) = g(x_0) = \inf_{t\in\Omega} f(t).$$

By (2.3), we conclude that μ^* solves (2.2), and, by Proposition 1, we conclude that points $t_1, \dots, t_r$ are global minima for f in Ω. $\square$

Proposition 3. *Let x_0 be an extreme point of the solution set of the optimization problem*

$$\min\{g(x) : x \in \mathrm{co}(F(\Omega)). \tag{2.8}$$

Then there exists (at least) one point $t_0 \in \Omega$ such that

$$F(t_0) = x_0,$$

which is a global minimum for the function f in Ω.

Proof. Since $x_0 \in \mathrm{co}(F(\Omega))$, there exist r points $t_i \in \Omega$ such that

$$x_0 = \sum_{i=1}^{r} \lambda_i F(t_i), \tag{2.9}$$

where $\lambda_i \geq 0$ and $\sum_{i=1}^{r} \lambda_i = 1$. By Proposition 2, we know that every t_i is a global minimum for f in Ω; therefore, every $x_i = F(t_i)$ is a solution for (2.8) and hence (2.9) shows that x_0 is not an extreme point of the solution set of (2.8) whenever $r > 1$. $\square$

Remark. The previous results provide general nontrivial necessary and sufficient conditions for points t to be global minima of the function f over the set Ω.

Corollary 1. *A point $t_0 \in \Omega$ is a global minimum for the function f in Ω if and only if it solves the nonlinear equation system*

$$F(t_0) = x_0,$$

where x_0 is an extreme point of the solution set for optimization problem (2.8).

Proof. This follows from Proposition 3. $\square$

Corollary 2. *The solutions of the nonlinear equations in λ_i and t_i variables, given by the expression*

$$\lambda_1 F(t_1) + \cdots + \lambda_r F(t_r) = x_0 \tag{2.10}$$

subject to the constraints

$$\lambda_i \geq 0, \quad \sum_{i=1}^{r} \lambda_i = 1,$$

where x_0 is a solution for optimization problem (2.8), provide all the global minima for f in Ω.

Proof. Since $F(\Omega) \subset \mathbb{R}^k$, it suffices to take $r = k + 1$ in order to satisfy Eq. (2.10). Next, we apply Proposition 2 to conclude that t_i are global minima for f in Ω. Reciprocally, if $t_1, \dots, t_r$ are global minima for f in Ω, it is easy to show that $x_0 = \sum_{i=1}^{r} \lambda_i F(t_i)$, where $\lambda_i \geq 0$ and $\sum_{i=1}^{r} \lambda_i = 1$, is a minimum for $g(x)$ in $\mathrm{co}(F(\Omega))$. $\square$

2.2. Explicit description by moments. So far we have seen that we can characterize the global minima of the objective function f by solving the optimization problem

$$\min\{g(x) : x \in \mathrm{co}(F(\Omega))\}. \tag{2.11}$$

It is very important to note that problem (2.11) has the following features, which are very useful in optimization:

(1) problem (2.11) has linear objective function g;
(2) problem (2.11) has convex feasible set $\mathrm{co}(F(\Omega))$.

For these reasons, we would like to describe the convex feasible set $\mathrm{co}(F(\Omega))$ in order to solve optimization problem (2.11) by using standard optimization techniques.

Linear-function decomposition (2.5) takes the explicit form

$$f(t) = \sum_{i=1}^{k} c_i \psi_i(t), \tag{2.12}$$

where

$$F(t) = (\psi_1(t), \ldots, \psi_k(t))$$

is a continuous transformation defined in $\Omega \subset \mathbb{R}^v$ with values in $\mathbb{R}^k$, and the linear functional g is defined by the relation

$$g(x_1, \ldots, x_k) = \sum_{i=1}^{k} c_i\, x_i$$

in $\mathbb{R}^k$. It is easy to see that the convex hull $\mathrm{co}(F(\Omega))$ of the image $F(\Omega)$ of the set Ω under the transformation F is composed of all vectors

$$(m_1, \ldots, m_k) \in \mathbb{R}^k$$

such that

$$m_i = \int_\Omega \psi_i(s)\, d\mu(s), \quad i = 1, \ldots, k,$$

where μ is some probability measure supported in Ω. Thus, the values $m_1, \ldots, m_k$ are precisely the generalized moments of the measure μ with respect to the function basis

$$\{\psi_1, \ldots, \psi_k\}.$$

The characterization of the values $m_1, \ldots, m_k$ as the moments of some measure μ is an open question in contemporary mathematics. This difficult task is called the *problem of moments*. Now, at this stage, the connection between the problem of moments and the analysis of global extremes of linear combinations such as (2.12) should be clear. Indeed, if we could find some convenient requirements for characterizing the values $m_1, \ldots, m_k$ as moments, then we could use them in order to pose optimization problem (2.11) as the programming problem

$$\min\left\{ \sum_{i=1}^{k} c_i\, m_i : \text{values } m_1, \ldots, m_k \text{ are the moments of some } \mu \in \mathrm{Pr}(\Omega) \right\}.$$

This proposal has been successful for treating some families of polynomial global optimization problems. In the next section, we review some essential facts about the problem of moments.

2.3. Trivial example. To illustrate the ability of the method to cope with nonconvex global optimization problems, we consider the step function

$$f = \sum_{i=1}^{k} c_i \psi_i,$$

where

$$\psi_i(t) = \begin{cases} 1 & \text{if } t \in \Omega_i, \\ 0 & \text{if } t \notin \Omega_i \end{cases}$$

is the *characteristic* function of the set Ω_i and the sets Ω_i form a partition of the set Ω in k disjoint subsets. Note that the function f may be extremely nonconvex, in fact, it may be as bizarre as you can imagine. Next, we see how the method of moments provides the global minima of the function f.

It is very easy to see that the values $m_1, \ldots, m_k$ are the moments of a probability measure, with respect to the family of characteristic functions $\{\psi_1, \ldots, \psi_k\}$, if and only if $m_i \geq 0$ for $i = 1, \ldots, k$ and

$$\sum_{i=1}^{k} m_i = 1.$$

Using this simple characterization, we can pose the global optimization problem as the convex program

$$\min_{m_i} \sum_{i=1}^{k} c_i\, m_i \quad \text{s.t.} \quad m_i \geq 0 \quad \text{and} \quad \sum_{i=1}^{k} m_i = 1,$$

whose solution is trivially

$$m_i^* = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j \end{cases}$$

for $i = 1, \ldots, k$, where

$$c_j = \min\{c_1, \ldots, c_k\}.$$

Therefore, the global minima for f are points $t \in \Omega$ satisfying the conditions

$$\psi_j(t) = 1, \quad \psi_i(t) = 0$$

for every $i \neq j$. Those are exactly the points t in Ω_j as we have found from a simple analysis of the step function f.

3. The Problem of Moments

The problem of moments has challenged the minds of some of the most famous mathematicians since the 19th century. In fact, renowned scientists as Tchebyshev, Stieltjes, Markov, Hausdorff, M. Riesz, and many others were involved in the study of different forms of this old problem, and they developed rich theories around it and its connections with other fields such as mechanics and probability (see [17] for a complete review on the problem of moments). We begin by providing a general description of the problem of moments.

Given a set of functions $\psi_1, \ldots, \psi_k$ defined in $\Omega \subset \mathbb{R}^v$ and values $m_1, \ldots, m_k$, the problem of moments consists of determining a positive measure μ such that

$$m_i = \int_{\Omega} \psi_i(s)\, d\mu(s) \quad \forall\, i = 1, \ldots, k, \tag{3.1}$$

whenever it is possible. Thus, the problem of moments also includes the search for requirements in order to characterize the values $m_1, \ldots, m_k$ as a set of moments. Depending on the function basis $\psi_1, \ldots, \psi_k$ and the set Ω, the problem of moments can take different forms. For example, if $\psi_i(t) = t^i$ is the algebraic system and $\Omega = \mathbb{R}$, we refer to the problem of moments as the *Hamburger moment problem*. We also review the following classical cases: the *Stieltjes moment problem*, in which $\psi_i(t) = t^i$ and $\Omega = [0, \infty)$, the *Hausdorff moment problem*, where $\psi_i(t) = t^i$ and $\Omega = [a, b]$, and the *Toeplitz* (*trigonometric*) *moment problem*, in which $\psi_i(t) = e^{ijt}$ and $\Omega = [-\pi, \pi)$. In addition, we explore the particular case of the two-dimensional, two-degree algebraic moment problem, where

$$\psi_{ij}(x, y) = x^i y^j, \quad 0 \leq i + j \leq 2.$$

It is easy to see that vectors $(m_1, \ldots, m_k)$ satisfying (3.1) form a convex cone $\mathcal{M}$ in $\mathbb{R}^k$. Therefore, we can see that the convex cone

$$\mathcal{P} = \left\{ (c_1, \ldots, c_k) \in R^k : \sum_{i=1}^{k} c_i\, \psi_i(t) \geq 0 \quad \forall\, t \in \Omega \right\},$$

composed of all positive functions in the linear span of the basis $\psi_1, \ldots, \psi_k$, is the convex dual of $\mathcal{M}$. This observation explains why the classical approach to studying the problem of moments (3.1) involves the algebraic analysis of the positive functions in the cone $\mathcal{P}$. Indeed, the same David Hilbert proposed, at the beginning of the 20th century, the study of positive polynomials in the seventeenth problem of his famous collection of mathematical problems for the new century (see [1, 12, 16, 17]).

As an illustration, we consider the family $\mathcal{P}$ of all $2n$-degree positive polynomials. It is well known that a positive polynomial

$$q(x) = \sum_{i=0}^{2n} c_i\, x^i$$

can be expressed as the sum of two squares of polynomials, i.e., this means that there exist two polynomials A and B such that

$$q(x) = A(x)^2 + B(x)^2$$

whenever $q(x) \geq 0$. Nesterov [12] devised a general frame, using standard linear algebra tools, to properly describe the dual of convex cones $\mathcal{P}$ composed of positive functions which admit a representation as a finite sum of squares. For example, since it is obvious that A and B polynomials belong to the span of the algebraic system $\mathcal{U} = \{1, x, \dots, x^n\}$, we need (at least) the basis

$$\mathcal{V} = \{1, x, \dots, x^{2n}\}$$

to describe all possible products of functions in the linear span of $\mathcal{U}$. Next, we define the linear transformation

$$\Lambda : R^{2n+1} \to M_{n+1} : a \to (a_{k+l})_{k,l=0}^{n},$$

where M_{n+1} represents the family of all $(n+1) \times (n+1)$ matrices. The values of the operator Λ are given by the projections on the basis $\mathcal{V}$ of every product of two functions in the basis $\mathcal{U}$, i.e.,

$$x^k \cdot x^l = x^{k+l} = \sum_{i=0}^{2n} \delta_{k+l,i}\, x^i \quad \forall\, k, l = 0, 1, \dots, n;$$

therefore, in general,

$$\Lambda\, a\,(k, l) = \sum_{i=0}^{2n} \delta_{k+l,i}\, a_i = a_{k+l} \quad \forall\, a \in \mathbb{R}^{2n+1}.$$

In [12, Theorem 17.1], it was shown that vector a belongs to the dual of the convex cone $\mathcal{P}$ if and only if $\Lambda\, a$ is a positive semidefinite matrix. Hence, we conclude that any vector a belongs to the closure of the moments cone

$$\mathcal{M} = \left\{ (m_0, \dots, m_{2n}) : m_i = \int_{\mathbb{R}} x^i d\mu(x) \quad \forall i = 0, \dots, 2n \right\}$$

if and only if the Hankel matrix

$$\Lambda\, a = (a_{k+l})_{k,l=0}^{n}$$

is positive semidefinite. Recall, from convex analysis, that the second convex dual of a convex cone $\mathcal{W}$ is the closure of $\mathcal{W}$ [15].

The previous example shows the importance of the description of positive polynomials as a finite sum of squares of polynomials for the problem of moments. However, Hilbert did show that it is not possible to express every positive polynomial in several variables as a finite sum of squares of polynomials. This characterization is possible only for some special classes of positive polynomials. This point is carefully analyzed in Shor's work on positive polynomials [16, Chap. 9]. In the remainder of this section, we review some important results an the problem of moments. In the next section, we see how to apply them to find the global minima of polynomial curves.

3.1. Hamburger moment problem. If $\Omega = \mathbb{R}$ and $\psi_i(t) = t^i$ for $i = 0, 1, \ldots, 2n$, the problem of moments is called the *Hamburger moment problem.* It is well known that there exists a positive measure with algebraic moments $m_0, \ldots, m_{2n}$ if and only if the values m_i form a positive semidefinite Hankel matrix

$$H = \begin{pmatrix} m_0 & m_1 & \cdots & m_n \\ m_1 & m_2 & \cdots & m_{n+1} \\ \cdots & \cdots & \cdots & \cdots \\ m_n & m_{n+1} & \cdots & m_{2n} \end{pmatrix}.$$

I warn that this sentence applies only in a limit sense; for example, consider the matrix

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

which is a positive semidefinite Hankel matrix, but it is not possible to find a positive measure with algebraic moments

$$m_0 = 1, \quad m_1 = 0, \quad m_2 = 0, \quad m_3 = 0, \quad m_4 = 1.$$

In other words, the previous sentence does characterize the closure of the convex set $\mathrm{co}(F(R))$, where

$$F(t) = (1, t, t^2, \ldots, t^{2n}) \quad \forall\, t \in \mathbb{R}.$$

Below, we list some useful results on the Hamburger moment problem that we will need later.

Lemma 1 ([1])**.** *If $m_0, \ldots, m_{2n}$ are the algebraic moments of a positive measure, then the Hankel matrix $H = (m_{i+j})_{i,j=0}^n$ is positive semidefinite.*

Lemma 2 ([1])**.** *If $H = (m_{i+j})_{i,j=0}^n$ is a positive definite Hankel matrix, then there exist many infinite positive measures whose algebraic moments are $m_0, \ldots, m_{2n}$.*

Lemma 3 ([9])**.** *Let $H = (m_{i+j})_{i,j=0}^n$ be a positive semidefinite Hankel matrix. If H is not positive definite, then there exists a unique positive measure whose algebraic moments of orders from 0 to $2n-1$ are $m_0, \ldots, m_{2n-1}$ and such that its $2n$-order moment does not exceed m_{2n}.*

Remark. The previous lemma is a short version of Fisher's theorem [16].

Definition. Given a Hankel matrix $H = (m_{i+j})_{i,j=0}^n$, we define the *degree* of H as $n+1$ if H is nonsingular; otherwise, we define the degree of H as the smallest integer r such that $0 \le r \le n$ and

$$D_k = |m_{i+j}|_{i,j=0}^k = 0$$

for every $k = r, \ldots, n$. In general, $\deg(H) \le \mathrm{rank}(H)$. If H is nonsingular, then $\deg(H) = \mathrm{rank}(H) = n+1$.

Proposition 4. *Let H be a singular positive semidefinite Hankel matrix with degree r. Then $D_k > 0$ for every $k = 0, \ldots, r-1$ and $D_k = 0$ for every $k = r, \ldots, n$.*

Remark ([8])**.** This result follows from a thorough analysis of the structure of Hankel matrices.

Lemma 4 ([3])**.** *Let $m_0, \ldots, m_{2n}$ be the first $2n+1$ algebraic moments of a positive measure μ. The measure μ is supported in $r \le n$ points if and only if the degree of the Hankel matrix $H = (m_{i+j})_{i,j=0}^n$ is r. Hence, the support of μ contains more than n points if and only if H is nonsingular.*

3.2. Trigonometric moment problem. Let $\Omega = [-\pi, \pi)$ and $\psi_i(t) = e^{jit}$ for $i = -n, \dots, n$, where j stands for the imaginary unit. In this case, the problem of moments is known as the *trigonometric moment problem.* There is a well-known result in probability called the *Bochner theorem*, which establishes that $m_{-n}, \dots, m_n$ are the trigonometric moments of a positive measure supported in $[-\pi, \pi)$ if and only if they form a positive semidefinite Toeplitz matrix

$$T = \begin{pmatrix} m_0 & m_{-1} & \cdots & m_{-n} \\ m_1 & m_0 & \cdots & m_{-n+1} \\ \cdots & \cdots & \cdots & \cdots \\ m_n & m_{n-1} & \cdots & m_0 \end{pmatrix}.$$

Some known results on the trigonometric moment problem are listed below. In the remainder of this section, we assume that m_i are complex numbers satisfying the Hermitian condition

$$\bar{m}_i = m_{-i} \quad \forall\, i = 0, \dots, n. \tag{3.2}$$

Bochner Theorem ([2])**.** *The values $m_{-n}, \dots, m_n$ are the trigonometric moments of a positive measure supported in $[-\pi, \pi)$ if and only if the Toeplitz matrix $T = (m_{i-j})_{i,j=0}^n$ is positive semidefinite.*

Lemma 5 ([3])**.** *Let $T = (m_{i-j})_{i,j=0}^n$ be a positive semidefinite Toeplitz matrix with rank $r \le n$. Then there exists a unique r-point supported measure whose trigonometric moments are $m_{-n}, \dots, m_n$.*

3.3. Stieltjes moment problem. In the case where $\Omega = [0, \infty)$ and $\psi_i(t) = t^i$ is the algebraic basis with $i = 0, \dots, n$, we refer to the moment problem as the *Stieltjes moment problem.* Now we list some known classical results on the theory of this problem.

Lemma 6 ([14])**.** *If the matrices*

$$H = (m_{i+j})_{i,j=0}^n, \quad K = (m_{i+j+1})_{i,j=0}^{n-1} \tag{3.3}$$

are both positive definite, then $m_0, \dots, m_{2n}$ are the algebraic moments of a positive measure supported in $[0, \infty)$. On the other hand, if $m_0, \dots, m_{2n}$ are the algebraic moments of some positive measure supported in $[0, \infty)$, then matrices H and K are positive semidefinite.

Lemma 7 ([14])**.** *If the matrices*

$$H = (m_{i+j})_{i,j=0}^n, \quad K = (m_{i+j+1})_{i,j=0}^{n}$$

are both positive definite, then $m_0, \dots, m_{2n}, m_{2n+1}$ are the algebraic moments of some positive measure supported in $[0, \infty)$. On the other hand, if $m_0, \dots, m_{2n+1}$ are the algebraic moments of some positive measure supported in $[0, \infty)$, then matrices H and K are positive semidefinite.

Lemma 8 ([14])**.** *If the matrices*

$$H = (m_{i+j})_{i,j=0}^n, \quad K = (m_{i+j+1})_{i,j=0}^{n-1}$$

are positive semidefinite, then there exists a positive measure supported in $[0, \infty)$ whose first $2n$ algebraic moments are $m_0, \dots, m_{2n-1}$ and such that its $2n$-order algebraic moment does not exceed m_{2n}.

Remark ([3])**.** The representation of $m_0, \dots, m_{2n}$ as the algebraic moments of some positive measure supported in $[0, \infty)$ is unique if and only if the Hankel matrix H is singular.

Lemma 9 ([14])**.** *If the matrices*

$$H = (m_{i+j})_{i,j=0}^n, \quad K = (m_{i+j+1})_{i,j=0}^{n}$$

are positive semidefinite, then there exists a positive measure supported in $[0, \infty)$ whose first $2n+1$ algebraic moments are $m_0, \dots, m_{2n}$ and such that its $(2n+1)$-order algebraic moment does not exceed m_{2n+1}.

Remark ([3]). The representation of $m_0, \dots, m_{2n+1}$ as the algebraic moments of some positive measure supported in $[0, \infty)$ is unique if and only if either the Hankel matrix H is singular or H is nonsingular and the matrix K is singular.

3.4. Hausdorff moment problem. The problem of moments is called the *Hausdorff moment problem* if $\Omega = [a, b]$ and the function basis is the algebraic system $\psi_i(t) = t^i$.

Lemma 10 ([14]). *The values $m_0, \dots, m_{2n}$ are the algebraic moments of a positive measure supported on the finite interval $[a, b]$ if and only if the matrices*

$$H = (m_{i+j})_{i,j=0}^{n}, \quad K = \Big((a+b)\, m_{i+j+1} - ab\, m_{i+j} - m_{i+j+2}\Big)_{i,j=0}^{n-1}$$

are positive semidefinite.

Lemma 11 ([14]). *The values $m_0, \dots, m_{2n+1}$ are the algebraic moments of a positive measure supported on the finite interval $[a, b]$ if and only if the matrices*

$$L = (m_{i+j+1} - a\, m_{i+j})_{i,j=0}^{n}, \quad M = (b\, m_{i+j} - m_{i+j+1})_{i,j=0}^{n}$$

are positive semidefinite.

Remark ([3]). If the Hankel matrix H is singular, then the representation of $m_0, \dots, m_{2n}$ as the moments of a positive measure supported in $[a, b]$ is unique. The same applies to an even set of $m_0, \dots, m_{2n+1}$.

3.5. Quadratic moment problem. In this kind of moments problem, we consider the family of functions

$$\psi_{i,j}(x, y) = \bar{z}^i z^j \quad \text{with} \quad 0 \leq i + j \leq 2, \tag{3.4}$$

where the real variables x and y are identified with the real and imaginary parts of the complex variable z. A linear combination

$$f(x, y) = \sum_{i,j} a_{ij} \bar{z}^i z^j$$

of the functions $\psi_{i,j}$ is real if and only if the coefficients a_{ij} are Hermitian, i.e.,

$$\bar{a}_{ij} = a_{ji}.$$

Lemma 12 ([4]). *The complex values m_{ij} are the moments of some positive measure supported in the complex plane z, with respect to basis* (3.4), *if and only if the matrix*

$$M = \begin{pmatrix} m_{00} & m_{01} & m_{10} \\ m_{10} & m_{11} & m_{20} \\ m_{01} & m_{02} & m_{11} \end{pmatrix}$$

is positive semidefinite.

It is easy to see that we can express every two-dimensional, two-degree polynomial as a linear combination of complex functions in the basis

$$\Big\{ \bar{z}^i z^j : 0 \leq i + j \leq 2 \Big\};$$

therefore, we can use this result of moments theory to study the global minima of polynomial expressions as

$$\sum_{0 \leq i+j \leq 2} c_{ij} x^i y^j \tag{3.5}$$

with real coefficients c_{ij}.

Indeed, we can see that

$$\begin{aligned} z+\bar{z} &= 2x, \\ z-\bar{z} &= 2iy, \\ z^2-\bar{z}^2 &= 4ixy, \\ 2\,z\,\bar{z}+z^2+\bar{z}^2 &= 4x^2, \\ 2\,z\,\bar{z}-z^2-\bar{z}^2 &= 4y^2. \end{aligned}$$

Therefore,

$$\begin{aligned} 1 &= \psi_{0,0}, \\ x &= \frac{1}{2}\psi_{0,1}+\frac{1}{2}\psi_{1,0}, \\ y &= -\frac{i}{2}\psi_{0,1}+\frac{i}{2}\psi_{1,0}, \\ xy &= \frac{i}{4}\psi_{2,0}-\frac{i}{4}\psi_{0,2}, \\ x^2 &= \frac{1}{2}\psi_{1,1}+\frac{1}{4}\psi_{0,1}+\frac{1}{4}\psi_{1,0}, \\ y^2 &= \frac{1}{2}\psi_{1,1}-\frac{1}{4}\psi_{0,1}-\frac{1}{4}\psi_{1,0}. \end{aligned}$$

Thus, every two-dimensional, two-degree real polynomial (3.5) can be expressed as a (Hermitian) linear combination of functions $\psi_{i,j}$ of basis (3.4).

4. Analysis of Global Minima of Polynomials

In this section, we show how the theory of moments can help us to analyze global optimization problems if the objective function is an algebraic or trigonometric polynomial. Indeed, we give a new characterization of the global minima for some classes of polynomials. The essential point that we want to emphasize is that we can obtain information about the global minima of nonlinear, nonconvex polynomials from the solution of some equivalent convex program. Moreover, such a convex program belongs to one particular class of convex programs that have been thoroughly studied in recent years, *semidefinite programs.*

A semidefinite program is an optimization problem in which the objective function is a linear combination

$$c^t x = c_1x_1+\cdots+c_kx_k,$$

whose variables are constrained by a set of *linear matrix inequalities.* A linear matrix inequality is a constraint expressed in the form

$$A_0+x_1A_1+\cdots+x_kA_k \geq 0, \tag{4.1}$$

where $A_0,\dots,A_k$ is a given set of symmetric matrices and the symbol $\geq$ stands for the positive semidefinite condition imposed on the left-hand side matrix in (4.1).

It is easy to see that these programs are convex. Indeed, they have a linear objective function and a convex feasible set. However, it is very important to note that their feasible set is a convex cone. In this sense, semidefinite programming is a generalization of the classical linear programming in which the feasible set is the positive cone $\mathbb{R}^n_+$. Currently, there is active research on the theory and algorithms for semidefinite programming. Some important concepts and tools have been developed around it. For instance, consider the tremendous development on *interior point methods* and the introduction of *self-concordant barrier functions.* It is not our purpose to explain the essentials of semidefinite programming here; for an introductory review on this subject, see [18].

4.1. Algebraic case. The first situation that we consider here is the analysis of the global minima of polynomials defined on the real line. We show that it is possible to determine the global minima for algebraic polynomials in the form

$$f(t) = \sum_{i=0}^{2n} c_i\, t^i \quad \text{with} \quad c_{2n} > 0 \tag{4.2}$$

from the solution of one particular semidefinite program. Thus, the following theorem establishes that the global minima of the optimization problem

$$\min_{t \in R} f(t) \tag{4.3}$$

are equivalent to the minima of the programming problem

$$\min_{m_i} \sum_{i=0}^{2n} c_i\, m_i, \tag{4.4}$$

where the variables $m_0, \ldots, m_{2n}$ are restricted to be the entries on a positive semidefinite Hankel matrix $H = (m_{i+j})_{i,j=0}^{n}$ with $m_0 = 1$. It is easy to see that the optimization problem given in (4.4) is a semidefinite program because of the symmetry on the Hankel matrix

$$H = \begin{pmatrix} m_0 & m_1 & m_2 & \cdots & m_n \\ m_1 & m_2 & \cdots\cdots & & m_{n+1} \\ m_2 & \cdots\cdots & & & m_{n+2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ m_n & m_{n+1} & m_{n+2} & \cdots & m_{2n} \end{pmatrix}.$$

Theorem 1. *If μ^* is a probability measure supported in the set of global minima of the polynomial $f(t)$ given in* (4.2)*, then the algebraic moments of μ^* solve program* (4.4)*. On the other hand, if $m_1^*, \ldots, m_{2n}^*$ solve program* (4.4)*, then there exists a unique probability measure μ^* supported in the set of global minima of the polynomial $f(t)$ whose algebraic moments are $m_1^*, \ldots, m_{2n}^*$.*

Proof. Let μ^* be a convex combination of points in the set of global minima of $f(t)$ and let $m_1^*, \ldots, m_{2n}^*$ be its algebraic moments. Observe that we are considering μ^* as a probability measure supported in the real line. Given an arbitrary set of values $m_1, \ldots, m_{2n}$, which form a positive semidefinite Hankel matrix $H = (m_{i+j})_{i,j=0}^{n}$ with $m_0 = 1$, we can find a measure μ such that its first $2n$ algebraic moments are $1, m_1, \ldots, m_{2n-1}$ and such that its $2n$-order moment does not exceed m_{2n}. That is the meaning of Lemma 3. Therefore, we have

$$\sum_{k=0}^{2n} c_k m_k^* = \int_R f(s)\, d\mu^*(s) \le \int_R f(s)\, d\mu(s) \le \sum_{k=0}^{2n} c_k m_k, \tag{4.5}$$

where the first inequality is implied by Proposition 1 and the assumption on the support of μ^*. The second inequality in (4.5) is implied by the positivity of the higher coefficient c_{2n} of the polynomial $f(t)$. In this way, we have shown that values $m_1^*, \ldots, m_{2n}^*$ solve semidefinite program (4.4).

Conversely, if $m_1^*, \ldots, m_{2n}^*$ solve semidefinite program (4.4), then we can find a unique probability measure $\bar{\mu}$ whose first $2n$ algebraic moments are $1, m_1^*, \ldots, m_{2n-1}^*$ and such that its $2n$-order moment does not exceed m_{2n}^*. This is true by Lemma 3 and the fact that the Hankel matrix $H^* = (m_{i+j}^*)_{i,j=0}^{n}$ cannot be positive definite. Indeed, if H^* is positive, then all the determinants

$$D_k = |m_{i+j}^*|_{i,j=0}^{k}$$

are positive for $k = 0, \ldots, n$. Thus, $m_1^*, \ldots, m_n^*$ do not belong to the boundary of the feasible set; therefore, they cannot be the solution for semidefinite program (4.4). Now we see that the $2n$-order moment of the measure $\bar{\mu}$ is m_{2n}^*. If this is not the case, then the moments of measure $\bar{\mu}$ are a better choice for problem (4.4). This contradicts the assumption about the values $m_1^*, \dots, m_{2n}^*$. Finally, given any probability measure μ supported in the real line, its algebraic moments $m_1, \dots, m_{2n}$ form a positive semidefinite Hankel matrix $H = (m_{i+j})_{i,j=0}^n$ with $m_0 = 1$. Thus,

$$\int_R f(s) d\bar{\mu}(s) = \sum_{k=0}^{2n} c_k m_k^* \le \sum_{k=0}^{2n} c_k m_k = \int_R f(s) d\mu(s),$$

where the inequality follows from the assumption about the values $m_1^*, \dots, m_{2n}^*$. Proposition 1 implies that the measure $\bar{\mu}$ is supported on the set of global minima of the polynomial $f(t)$ in (4.2). □

Remark. The previous theorem also establishes the equivalence between the relaxed problem

$$\min_{\mu \in \Pr(R)} \int_R f(s)\, d\mu(s) \tag{4.6}$$

and semidefinite program (4.4) in the following sense.

Corollary 3. *The algebraic moments of every solution μ^* for* (4.6) *solve* (4.4). *For every solution $m_1^*, \dots, m_{2n}^*$ of problem* (4.4)*, there exists a unique probability measure μ^* with algebraic moments $m_1^*, \dots, m_{2n}^*$, which solves* (4.6).

Corollary 4. *Let $m_1^*, \dots, m_{2n}^*$ be a solution of problem* (4.4) *and r be the degree of the Hankel matrix $H^* = (m_{i+j}^*)_{i,j=0}^n$ with $m_0^* = 1$. Then polynomial* (4.2) *has at least r global minima.*

Proof. Let μ^* be the unique measure with moments $1, m_1^*, \dots, m_{2n}^*$. Then, by Lemma 4, we can see that μ^* has r supporting points. Therefore, $f(t)$ has at least r global minima. □

Corollary 5. *If polynomial $f(t)$ in* (4.2) *has k global minima, then the set of solutions for semidefinite program* (4.4) *is a k-simplex in $\mathbb{R}^{2n}$.*

Proof. Assume that $\{t_1, \dots, t_k\}$ is the set of global minima of polynomial $f(t)$ in (4.2). Let $m_1^*, \dots, m_{2n}^*$ be a solution of semidefinite program (4.4). Then there exists a unique probability measure μ^* whose algebraic moments are $m_1^*, \dots, m_{2n}^*$. Moreover, the support of μ^* is composed of some points in $\{t_1, \dots, t_k\}$; therefore, the moments $m_1^*, \dots, m_{2n}^*$ can be described as a convex combination of the moments of the Dirac measures $\delta_{t_1}, \dots, \delta_{t_k}$. Now we must show that the moments of the Dirac measures $\delta_{t_1}, \dots, \delta_{t_k}$ are linearly independent vectors in $\mathbb{R}^{2n}$. Since $k \le n < 2n$, it suffices to consider the Vandermonde matrix

$$V = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_k \\ t_1^2 & t_2^2 & \cdots & t_k^2 \\ \cdots & \cdots & \cdots & \cdots \\ t_1^{k-1} & t_2^{k-1} & \cdots & t_k^{k-1} \end{pmatrix}$$

whose determinant equals $\prod_{i<j} (t_i - t_j)$. □

Corollary 6. *Assume that $\{t_1, \dots, t_k\}$ is the set of global minima of the polynomial $f(t)$ given in* (4.2)*. Then the moments of the Dirac measures $\delta_{t_1}, \dots, \delta_{t_k}$ are the extreme points of the solution set for semidefinite program* (4.4).

Proof. Given the Dirac measure $\mu^* = \delta_{t_i}$, where t_i is a global minimum for $f(t)$, we know, by Proposition 1, that μ^* solves relaxed problem (4.6). If the algebraic moments $t_i, t_i^2, \dots, t_i^{2n}$ of μ^* can be expressed as a convex combination of the other moments of the Dirac measures $\delta_{t_1}, \dots, \delta_{i-1}, \delta_{i+1} \dots, \delta_{t_k}$, that would imply that the Vandermonde matrix V is singular. Since points $t_1, \dots, t_k$ are distinct, the Vandermonde matrix V is nonsingular. Thus, we conclude that moments $t_i, t_i^2, \dots, t_i^{2n}$ form an extreme point of the solution set for semidefinite program (4.4). Note that we have used implicitly the fact that $k \leq n < 2n$. $\square$

Corollary 7. *Let* $m_1^*, \ldots, m_{2n}^*$ *be an extreme point of the solution set for semidefinite program* (4.4). *Then* m_1^* *is a global minimum of the polynomial* $f(t)$ *in* (4.2).

Proof. This follows from Proposition 3. $\square$

Corollary 8. *If polynomial* (4.2) *has only one global minimum and* $m_1^*, \ldots, m_{2n}^*$ *solve* (4.4), *then* m_1^* *is the global minimum for* f.

Proof. In this case, $m_1^*, \ldots, m_{2n}^*$ is an extreme point of the solution set for (4.4). $\square$

4.2. Trigonometric case. Now we analyze the global minima for curves given by trigonometric polynomials of the form

$$f(t) = \sum_{i=-n}^{n} c_i \, e^{jit}. \tag{4.7}$$

In these cases, the polynomial f is real if and only if the complex coefficients c_i are Hermitian, i.e.,

$$\bar{c}_i = c_{-i}.$$

Since every complex exponential function e^{jit} is 2π-periodic, it is obvious that our analysis refers to the fundamental interval $-\pi \leq t < \pi$. The next theorem shows how we can reduce this global optimization problem to a semidefinite program. The method is very similar to the polynomial case studied above. We see that the global minima for the optimization problem $\min\limits_{-\pi<t\leq\pi} f(t)$ are equivalent to the minima of the semidefinite program

$$\min_{m_i} \sum_{i=-n}^{n} c_i \, m_i, \tag{4.8}$$

where the variables m_i are restricted to be the entries on a positive semidefinite Toeplitz matrix $T = (m_{i-j})_{i,j=0}^{n}$ with $m_0 = 1$.

Since variables $m_{-n}, \ldots, m_n$ are complex, it is not completely clear why optimization problem (4.8) is a semidefinite program. To clarify this point, we must take the real and imaginary parts of the values $m_{-n}, \ldots, m_n$ as a new set of real variables. Then objective function (4.8) can be expressed as a linear combination of the new set of variables and the new feasible set becomes a convex cone in some Euclidean space.

Theorem 2. *Let* μ^* *be a probability measure supported in the interval* $[-\pi, \pi)$ *whose supporting points are global minima for polynomial* (4.7). *Then the set of trigonometric moments* $m_{-n}^*, \ldots, m_n^*$ *of the measure* μ^* *is a solution for optimization problem* (4.8). *On the other hand, if* $m_{-n}^*, \ldots, m_n^*$ *solve problem* (4.8), *then there exists a unique probability measure* μ^* *with trigonometric moments* $m_{-n}^*, \ldots, m_n^*$, *whose supporting points are global minima of polynomial* (4.7) *in the interval* $[-\pi, \pi)$.

Proof. Assume that μ^* is a probability measure supported in some subset of the set of global minima of the polynomial $f(t)$ given in (4.7). We also assume that μ^* is supported in the interval $[-\pi, \pi)$. Given any set of values $m_{-n}, \ldots, m_n$ that forms a semidefinite Toeplitz matrix $T = (m_{i-j})_{i,j=0}^{n}$ with $m_0 = 1$, we know, by the Bochner theorem, that it is possible to find a probability measure μ supported in $[-\pi, \pi)$

whose trigonometric moments are $m_{-n}, \dots, m_n$. If $m^*_{-n}, \dots, m^*_n$ are the trigonometric moments of the measure μ^*, we have

$$\sum_{-n}^{n} c_i m_i^* = \int_{-\pi}^{\pi} f(s)\, d\mu^*(s) \leq \int_{-\pi}^{\pi} f(s)\, d\mu(s) = \sum_{-n}^{n} c_i m_i, \tag{4.9}$$

where the inequality in (4.9) follows from Proposition 1. Finally, we conclude that values $m^*_{-n}, \dots, m^*_n$ form a solution for program (4.8).

Conversely, if $m^*_{-n}, \dots, m^*_n$ is a solution for program (4.8), it is easy to show that the Toeplitz matrix $T^* = (m^*_{i-j})^n_{i,j=0}$ is singular. Let r be the rank of matrix T^*. By Lemma 5, we know that there exists a unique probability measure μ^* with trigonometric moments $m^*_{-n}, \dots, m^*_n$, which is supported in r points located on the interval $[-\pi, \pi)$. Given any arbitrary probability measure μ supported in $[-\pi, \pi)$, its trigonometric moments $m_{-n}, \dots, m_n$ form a positive semidefinite Toeplitz matrix $T = (m_{i-j})^n_{i,j=0}$ with $m_0 = 1$. Since the values $m^*_{-n}, \dots, m^*_n$ solve (4.8), we have

$$\int_{-\pi}^{\pi} f(s) d\mu^*(s) = \sum_{-n}^{n} c_i m_i^* \leq \sum_{-n}^{n} c_i m_i = \int_{-\pi}^{\pi} f(s) d\mu(s).$$

Thus, measure μ^* solves the relaxed problem

$$\min_{\mu \in \Pr([-\pi,\pi))} \int_{-\pi}^{\pi} f(s)\, d\mu(s) \tag{4.10}$$

and, by Proposition 1, we conclude that the supporting points of μ^* are global minima of the trigonometric polynomial $f(t)$ in $[-\pi, \pi)$. $\square$

Remark. The previous theorem also establishes the equivalence between relaxed problem (4.10) and optimization problem (4.8).

Corollary 9. *Let μ^* be a solution for relaxed problem* (4.10)*. Then the trigonometric moments of μ^* solve* (4.8)*. On the other hand, if the values $m^*_{-n}, \dots, m^*_n$ solve* (4.8)*, then there exists a unique probability measure μ^* with trigonometric moments $m^*_{-n}, \dots, m^*_n$, which solves* (4.10)*.*

Corollary 10. *Let $m^*_{-n}, \dots, m^*_n$ be a solution for* (4.8) *and r be the rank of the Toeplitz matrix $T^* = (m^*_{i-j})^n_{i,j=0}$. Then the polynomial f given in* (4.7) *has at least r global minima.*

Proof. By Lemma 5, we know that there exists a unique measure μ^* supported in r points whose trigonometric moments are the values $m^*_{-n}, \dots, m^*_n$. By Theorem 2, we know that the supporting points of μ^* are global minima for the polynomial $f(t)$. $\square$

Corollary 11. *If the trigonometric polynomial f given in* (4.7) *has k global minima, then the set of solutions of programming problem* (4.8) *is a k-simplex in $\mathbb{R}^{2n}$.*

Proof. We use the new set of $2n$ real variables $x_1, \dots, x_n;\ y_1, \dots, y_n$, where

$$x_i = \mathrm{Re}(m_i), \quad y_i = \mathrm{Im}(m_i)$$

for $i = 1, \dots, n$. Assume that $\{t_1, \dots, t_k\}$ is the set of global minima for the polynomial $f(t)$ in $[-\pi, \pi)$. Given a solution $m^*_{-n}, \dots, m^*_n$ for problem (4.8), there exists a unique measure μ^* with trigonometric moments $m^*_{-n}, \dots, m^*_n$, whose supporting points are global minima for $f(t)$ in $[-\pi, \pi)$. Therefore, the values $m^*_{-n}, \dots, m^*_n$ can be expressed as convex combinations of the trigonometric moments of the Dirac measures $\delta_{t_1}, \dots, \delta_{t_k}$. Hence the real vector

$$\left(x_1^*, \dots, x_n^*;\ y_1^*, \dots, y_n^*\right)$$

can be expressed as a convex combination of k $2n$-dimensional real vectors

$$\begin{aligned}&\Big(\cos(t_1),\dots,\cos(nt_1);\ \sin(t_1),\dots,\sin(nt_1)\Big),\ \dots\\&\Big(\cos(t_k),\dots,\cos(nt_k);\sin(t_k),\dots,\sin(nt_k)\Big).\end{aligned} \tag{4.11}$$

Next, we see that the set of vectors (4.11) is linearly independent. We can identify the real vectors (4.11) in $\mathbb{R}^{2n}$ with the set of complex vectors in $\mathbb{C}^n$ given as

$$\Big(e^{jt_1},\dots,e^{jnt_1}\Big),\ \dots,\ \Big(e^{jt_k},\dots,e^{jnt_k}\Big).$$

Taking the Vandermonde-type matrix

$$\bar{V}=\begin{pmatrix} z_1 & z_1^2 & \cdots & z_1^k\\ z_2 & z_2^2 & \cdots & z_2^k\\ \dots & \dots & \dots & \dots\\ z_k & z_k^2 & \cdots & z_k^k\end{pmatrix},$$

where $z_1=e^{jt_1},\dots,z_k=e^{jt_k}$, it is easy to see that

$$\det(\bar{V})=e^{j(t_1+\cdots+t_k)}\prod_{i>j}(z_i-z_j).$$

Therefore, we conclude that the real vectors in (4.11) form a linearly independent set since points $t_1,\dots,t_k$ are distinct. □

Remark. In the previous argument, we implicitly used the fact that $k\le n$.

Corollary 12. *Assume that $\{t_1,\dots,t_k\}$ is the set of global minima of the trigonometric polynomial $f(t)$ given in* (4.7)*. Then the trigonometric moments of the Dirac measures $\delta_{t_1},\dots,\delta_{t_k}$ are the extreme points of the solution set for semidefinite program* (4.8)*.*

Proof. We must take the real and imaginary parts of the trigonometric moments

$$m_i=\int_{-\pi}^{\pi}e^{jit}d\delta_{t_l}=e^{jit_l}$$

for $i=-n,\dots,n$ and $l=1,\dots,k$. Next, we must follow the same argument given in Corollary 11. □

Corollary 13. *If the values $m^*_{-n},\dots,m^*_n$ form an extreme point of the solution set of program* (4.8)*, then the expression*

$$t_0=\begin{cases}\ \ \arccos(\mathrm{Re}(m_1^*)) & \textit{if } \mathrm{Im}(m_1^*)\ge 0,\\ -\arccos(\mathrm{Re}(m_1^*)) & \textit{if } \mathrm{Im}(m_1^*)<0\end{cases} \tag{4.12}$$

gives one global minimum for f.

Proof. Expression (4.12) is the converse for

$$m_1^*=\int_{-\pi}^{\pi}e^{jt}\,\delta_{t_o}(t)\,dt=e^{jt_0}=\cos t_0+j\,\sin t_0.$$

The remainder of the proof follows from Proposition 3. □

4.2.1. One side bounded problems. Now we consider the optimization problem

$$\min_{t\geq 0} f(t), \tag{4.13}$$

where the objective function is a polynomial of the form

$$f(t) = \sum_{i=0}^{k} c_i\, x^i, \qquad c_k > 0. \tag{4.14}$$

In this kind of problem, we must analyze two different possibilities. The first one, when the polynomial f has even degree $k = 2n$, and the second one, when the polynomial f has odd degree $k = 2n+1$. Next, we will see that global minima for problem (4.13) can be characterized by using the solutions for a certain semidefinite program.

If $k = 2n$, the global minima for problem (4.13) are equivalent to the minima of the optimization problem

$$\min_{m_i} \sum_{i=0}^{2n} c_i\, m_i, \tag{4.15}$$

where the variables $m_0, \dots, m_{2n}$ are constrained by the positiveness on the Hankel matrices

$$H = (m_{i+j})_{i,j=0}^{n}, \quad K = (m_{i+j+1})_{i,j=0}^{n-1} \tag{4.16}$$

with $m_0 = 1$. It is easy to see that the optimization problem given in (4.15) is a semidefinite program. The next theorem shows that program (4.15) is equivalent to the finding of the global minima of polynomial (4.14) for $t \geq 0$ and $k = 2n$.

Theorem 3. *Assume that $k = 2n$. Let μ^* be a probability measure supported in $[0, \infty)$, whose supporting points are global minima of the polynomial $f(t)$ in $[0, \infty)$. Then the algebraic moments $m_1^*, \dots, m_k^*$ of measure μ^* solve optimization problem (4.15). Conversely, if $m_1^*, \dots, m_k^*$ solve problem (4.15), then there exists a unique probability measure μ^* with algebraic moments $m_1^*, \dots, m_k^*$, whose supporting points are global minima for $f(t)$ in $[0, \infty)$.*

Proof. Using Lemmas 6 and 8, our assertion follows by the same arguments given in the proof of Theorem 1. □

If $k = 2n+1$, the global minima for problem (4.13) are equivalent to the minima of the programming problem

$$\min_{m_i} \sum_{i=0}^{2n+1} c_i\, m_i, \tag{4.17}$$

where the variables $m_0, \dots, m_{2n+1}$ are constrained by the positiveness of the Hankel matrices

$$H = (m_{i+j})_{i,j=0}^{n}, \quad K = (m_{i+j+1})_{i,j=0}^{n} \tag{4.18}$$

with $m_0 = 1$. We stress that optimization problem (4.17) is a semidefinite program. The next theorem shows that it is equivalent to the finding of the global minima in $[0, \infty)$ of polynomial (4.14) if it has odd degree $k = 2n+1$.

Theorem 4. *Assume that $k = 2n+1$. Let μ^* be one probability measure supported in $[0, \infty)$, whose supporting points are global minima of the polynomial $f(t)$ in $[0, \infty)$. Then the algebraic moments $m_1^*, \dots, m_k^*$ of measure μ^* solve optimization problem (4.17). Conversely, if $m_1^*, \dots, m_k^*$ solve problem (4.17), then there exists a unique probability measure μ^* supported in $[0, \infty)$, with algebraic moments $m_1^*, \dots, m_k^*$, whose supporting points are global minima for $f(t)$ in $[0, \infty)$.*

Proof. Using Lemmas 7 and 9, our assertion follows by the same argument that we used in the proof of Theorem 1. $\square$

4.2.2. Bounded Problems. Now we consider global optimization problems

$$\min_{a\le t\le b} f(t),$$

where f is a polynomial of the form

$$f(t) = \sum_{i=0}^{k} c_i\, t^i, \qquad c_k \neq 0. \tag{4.19}$$

Again, we can characterize global minima for f in $[a, b]$ using certain equivalent semidefinite program.

The global minima for polynomial (4.19) with even degree $k = 2n$ on the interval $[a, b]$ are equivalent to the minima of the semidefinite programming problem

$$\min_{m_i} \sum_{i=0}^{2n} c_i\, m_i, \tag{4.20}$$

where the variables $m_1, \dots, m_{2n}$ are restricted to be the entries in the positive semidefinite Hankel-type matrices

$$H = (m_{i+j})_{i,j=0}^{n}, \quad K = \Big((a+b)\, m_{i+j+1} - ab\, m_{i+j} - m_{i+j+2}\Big)_{i,j=0}^{n-1}$$

with $m_0 = 1$. Now it should be obvious to the reader how to prove this equivalence by using the content of Lemma 10.

We can obtain a similar result for the case where the polynomial f in (4.19) has odd degree $k = 2n+1$. In this case, the global minima for polynomial (4.19) with even degree $k = 2n + 1$ on the interval $[a, b]$ are equivalent to the minima of the programming problem

$$\min_{m_i} \sum_{i=0}^{2n+1} c_i\, m_i, \tag{4.21}$$

where the variables $m_1, \dots, m_{2n+1}$ are restricted to be the entries in the positive semidefinite Hankel-type matrices

$$H = \Big(m_{i+j+1} - a\, m_{i+j}\Big)_{i,j=0}^{n}, \quad K = \Big(b\, m_{i+j} - m_{i+j+1}\Big)_{i,j=0}^{n}$$

with $m_0 = 1$. The reader should be able to use Lemma 11 in order to prove this equivalence.

4.2.3. Bidimensional Problems. Despite its elementary nature, the (global) minima for two-dimensional, two-degree real polynomials

$$f(x, y) = \sum_{0\le i+j\le 2} c_{ij} x^i y^j$$

can be characterized by the semidefinite program

$$\min_{m_{ij}} \sum_{0\le i+j\le 2} a_{ij} m_{ij} \quad \text{s.t.} \quad \begin{pmatrix} m_{00} & m_{01} & m_{10} \\ m_{10} & m_{11} & m_{20} \\ m_{01} & m_{02} & m_{11} \end{pmatrix} \ge 0, \tag{4.22}$$

where the coefficients a_{ij} come from the representation of polynomial f in the complex system $\bar{z}^i z^j$, i.e.,

$$f(x, y) = \sum_{0\le i+j\le 2} a_{ij} \bar{z}^i z^j. \tag{4.23}$$

In this situation, the variables m_{ij} are Hermitian complex variables; therefore, we can use their real and imaginary parts as a new set of real variables for the optimization problem. In this new setting, we conserve the essentials of semidefinite programming, i.e., a linear objective function and one particular convex cone as the feasible set. Lemma 12 allows us to characterize the global minima of the bidimensional polynomial in (4.23) from the solution set for semidefinite program (4.22).

5. Measure Recovery

The first goal of the problem of moments is to find a measure μ from the set of its moments

$$m_i = \int_\Omega \psi_i(s)\, d\mu(s)$$

with respect to the function basis $\psi_1, \dots, \psi_k$. In the previous sections, we established the relationship between the global minima of the linear combination

$$f = \sum_{i=1}^{k} c_i \psi_i, \tag{5.1}$$

the relaxed problem

$$\min_{\mu \in \Pr(\Omega)} \int_\Omega f(s)\, d\mu(s), \tag{5.2}$$

and the convex program

$$\min_{m_i} \sum_{i=1}^{k} c_i\, m_i. \tag{5.3}$$

In fact, Theorems 1–4 claim that the solution of program (5.3) corresponds to the moments of a probability measure, which solves relaxed problem (5.2). On the other hand, Proposition 1 claims that the supporting points of the solution of relaxed problem (5.2) are global minima for linear combination (5.1). Therefore, it seems natural to use the optimum moments $m_1^*, \dots, m_k^*$ of problem (5.3) to obtain the global minima of function (5.1). This suggestion works very well in the cases studied before.

5.1. Algebraic case. Given a positive semidefinite Hankel matrix $H = (m_{i+j})_{i,j=0}^{n}$ of degree $r \le n$, we know, by Lemma 3, that there exists a unique positive measure μ whose algebraic moments are $m_0, \dots, m_{2n-1}$ such that its $2n$-order moment does not exceed m_{2n}. It is well known in the theory of moments [9] that the polynomial

$$P(x) = \begin{vmatrix} m_0 & m_1 & \cdots & m_r \\ \multicolumn{4}{c}{\cdots\cdots\cdots\cdots\cdots\cdots} \\ m_{r-1} & m_r & \cdots & m_{2r-1} \\ 1 & x & & x^r \end{vmatrix}$$

has r different real roots which are the supporting points of the measure μ. Thus, we obtain the following result.

Theorem 5. *Given a polynomial*

$$f(t) = \sum_{k=1}^{2n} c_i\, t^i \tag{5.4}$$

with $c_{2n} > 0$*, assume that the values* $m_1^*, \dots, m_{2n}^*$ *solve the semidefinite program*

$$\min_{m_i} \sum_{i=1}^{2n} c_i\, m_i,$$

where the variables $m_1, \dots, m_{2n}$ form a positive semidefinite Hankel matrix $H = (m_{i+j})_{i,j=0}^n$ with $m_0 = 1$. If $m_1^*, \dots, m_{2n}^*$ form a Hankel matrix $H^* = (m_{i+j}^*)_{i,j=0}^n$ of degree r, then the roots of the polynomial

$$P^*(x) = \begin{vmatrix} 1 & m_1^* & \cdots & m_r^* \\ \cdots & \cdots & \cdots & \cdots \\ m_{r-1}^* & m_r^* & \cdots & m_{2r-1}^* \\ 1 & x & & x^r \end{vmatrix} \tag{5.5}$$

are global minima of polynomial (5.4).

It is interesting to note that the classical method of finding the local minima of a polynomial is to estimate the roots of its derivative, which in turn is also a polynomial. Here we propose that the roots of the polynomial (5.5) provide the global minima of original polynomial (5.4).

5.2. Trigonometric case. Let $T = (m_{i-j})_{i,j=0}^n$ be a positive semidefinite Toeplitz matrix with rank $r \le n$. It was shown [3] that the submatrix $T_{r-1} = (m_{i-j})_{i,j=0}^{r-1}$ is nonsingular. If we take the r-dimensional vector $(a_0, \dots, a_{r-1})$ as the solution of the linear system

$$\begin{pmatrix} m_0 & m_{-1} & \cdots & m_{-r+1} \\ m_1 & m_0 & \cdots & m_{-r} \\ \cdots & \cdots & \cdots & \cdots \\ m_{r-1} & m_{r-2} & \cdots & m_0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{r-1} \end{pmatrix} = \begin{pmatrix} m_{-r} \\ m_{-r+1} \\ \vdots \\ m_{-1} \end{pmatrix},$$

then it can be shown [3] that the set of roots of the polynomial

$$P(z) = z^r - a_0 - a_1 z - \cdots - a_{r-1} z^{r-1}$$

contains the supporting points of the unique measure supported in $[-\pi, \pi)$ with trigonometric moments $m_{-n}, \dots, m_n$. In this context, we identify the complex variable $z = e^{jt}$ with points t on the interval $-\pi \le t < \pi$. From these results, we can establish the following theorem.

Theorem 6. *Given a trigonometric polynomial*

$$f(t) = \sum_{i=-n}^{n} c_i e^{ijt}$$

with Hermitian coefficients, assume that the values $m_{-n}^*, \dots, m_n^*$ solve the program

$$\min_{m_i} \sum_{i=-n}^{n} c_i \, m_i,$$

where the variables $m_{-n}, \dots, m_n$ form a positive semidefinite Toeplitz matrix $T = (m_{i-j})_{i,j=0}^n$ with $m_0 = 1$. If $m_{-n}^*, \dots, m_n^*$ form a Toeplitz matrix $T^* = (m_{i-j}^*)_{i,j=0}^n$ with rank r, then the roots of the polynomial

$$P^*(z) = z^r - a_0^* - a_1^* z - \cdots - a_{r-1}^* z^{r-1}$$

are global minima of the trigonometric polynomial $f(t)$ on the interval $[-\pi, \pi)$, where the coefficients $a_0^*, \dots, a_{r-1}^*$ are the solution of the linear system

$$\begin{pmatrix} 1 & m_{-1}^* & \cdots & m_{-r+1}^* \\ m_1^* & 1 & \cdots & m_{-r}^* \\ \cdots & \cdots & \cdots & \cdots \\ m_{r-1}^* & m_{r-2}^* & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_0^* \\ a_1^* \\ \vdots \\ a_{r-1}^* \end{pmatrix} = \begin{pmatrix} m_{-r}^* \\ m_{-r+1}^* \\ \vdots \\ m_{-1}^* \end{pmatrix}.$$

Proof. There exists a unique probability measure supported in $[-\pi, \pi)$ with algebraic moments m_{-n}^*, …, m_n^*, which is supported in r points (see Lemma 5). On the other hand, P^* has at most r zeros. $\square$

6. Conclusion

In this paper, we presented an alternative way of studying global optimization problems from the standpoint of the theory of moments. There are many questions still to be solved completely, for example, how to extend the method to higher dimensions and how to treat nonclassical moments problems. There is work in progress to clarify these points. However, we hope these ideas will be a good source for research and applications in the growing field of global optimization.

From another point of view, the essentials of the method of moments has been successfully applied to another kind of optimization problems. Indeed, the method of moments also allows us to solve nonconvex variational problems coming from elasticity models used in materials science and mechanical engineering [10, 11].

REFERENCES

1. N. Akhiezer, *The Classical Moment Problem*, Oliver and Boyd (1965).
2. S. Bochner, *Lectures on Fourier Analysis*, Princeton (1959).
3. R. Curto and L. Fialkow, "Recursiveness, positivity, and truncated moment problems," *Houston J. Math.,* **17**, No. 4 (1991).
4. R. Curto and L. Fialkow, "Solution of the truncated complex moment problem for flat data," *Mem. Amer. Math. Soc.,* **119**, No. 568 (1996).
5. F. Gantmacher, *The Theory of Matrices*, Chelsea (1989).
6. P. Gahinet, A. Nemirovsky, A. J. Laub, and M. Chilali, *LMI Control Toolbox*, The MathWorks Inc. (1995).
7. H. Frenk et al (Eds.), *High Performance Optimization*, Kluwer (2000).
8. I. Iohvidov, *Hankel and Toeplitz Matrices and Forms*, Birkhauser (1982).
9. N. Akhiezer and M. Krein, "Some questions in the theory of moments," *Transl. Math. Monogr.*, Vol. 2, Amer. Math. Soc. (1962).
10. R. Meziat, *El método de los momentos para problemas variacionales no convexos*, Ph.D. thesis, Universidad Politécnica de Cataluña, Barcelona (2001).
11. R. Meziat, J. J. Egozcue, and P. Pedregal, "The method of moments for nonconvex variational problems," *Proc. Int. Conf. Convex Anal. Global Optimization* (N. Hadjisavvas, Ed.), Kluwer (2001).
12. Y. Nesterov, *Squared Functional Systems and Optimization Problems, Higher Performance Optimization*, Kluwer (2000).
13. Y. Nesterov and A. Nemirovsky, *Interior Point Polynomial Methods in Convex Programming*, SIAM (1994).
14. M. Krein and A. Nudel'man, "The Markov moment problem and extremal problems," *Transl. Math. Monogr.*, Vol. 59, Amer. Math. Soc. (1977).
15. R. T. Rockafellar, *Convex Analysis*, Princeton (1970).
16. N. Shor, *Nondifferentiable Optimization and Polynomial Problems*, Kluwer (2000).
17. J. Shohat and J. Tamarkin, *The Problem of Moments*, Math. Surv., Vol. 1, Amer. Math. Soc. (1943).
18. L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.,* **38**, No. 1 (1996).