Frontiers In Global Optimization, pp. 1-2
C. A. Floudas and P. M. Pardalos, Editors

Analysis of Non Convex Polynomial Programs by the Method of Moments

René J. Meziat
Departamento de Matemáticas
Universidad de los Andes
Carrera 1 No 18A-10
Bogotá, Colombia
rmeziat@uniandes.edu.co

Abstract

In this work we propose a general procedure for analyzing global minima of arbitrary mathematical programs which is based in probability measures and moments theory. We give a general characterization of global minima of arbitrary programs, and as a particular case, we characterize the global minima of unconstrained one dimensional polynomial programs by using a particular semidefinite program.

Keywords: global optimization, method of moments, semidefinite programming.

1 Introduction

The Method of Moments is a general method for treating non convex optimization problems, which is particularly well suited to cope with global optimization problems which come from several research areas in Optimization Theory like Calculus of Variations, Control Theory and Mathematical Programming. It has been successfully applied to non convex variational problems and one dimensional polynomial programs. See [10], [11], [12] and [13] for different applications of the Method of Moments. The Method of Moments takes a proper formulation in probability measures of a non convex optimization problem. Thus, when the problem can be stated in terms of polynomial expressions, we can transform the measures into algebraic moments to obtain a new convex program defined in a new set of variables that represent the moments of every measure. This procedure has been successfully employed for treating non convex variational problems when we use their generalized formulation in Young measures. See [11], [12] and [13].

The application of the moments theory to optimization problems is not new in any way, in fact it can be traced back to the works of Markov and Tchebychev, see [6] for an interesting treatment of these ideas. Recently, other authors have proposed the use of the moments theory to analyze non convex polynomial programs in global optimization, the reader should refer the works of J.B. Lasserre given in references [7] , [8] and [9], the work of Y. Nesterov in [14] and the work of N. Shor in reference [17].

The purpose of this work is to analyze the ability of the Method of Moments for treating mathematical programs given in the general form

$$\min_{t\in\Omega} f\left(t\right) \tag{1}$$

where the objective function $f\left(t\right)$ is a linear combination of simpler functions. We will illustrate the success of this analysis by studying those cases in which the objective function f is a one dimensional polynomial in some real interval Ω.

In order to properly use the Method of Moments, we must solve some particular *Problem* of *Moments* for every non convex mathematical program we are interested in. These problems are really difficult to solve, and they make part of the classical repertory of famous problems in contemporary mathematics. A short review on the classical treatment of the Problem of Moments can be found in [16]. In this paper, we describe how to solve trigonometric and algebraic moment problems in arbitrary intervals Ω on the real line, by using the classical characterizations of one dimensional positive polynomials in Ω. In this way, we will show that the solution of a particular Problem of Moments in some interval Ω, is obtained by constructing a corresponding positive semidefinite quadratic form from the parameters of the domain Ω of the polynomial program (1) . This kind of characterizations of finite sets of moments can be found in [4], [6], [14] and [16].

In order to solve a particular non convex polynomial program (1) , we can apply the Method of Moments by transforming it into an equivalent convex program. This new problem has a linear objective function, which is defined by the coefficients of the objective function f and a convex feasible set, defined by the positivity of the quadratic form that characterizes a finite set of moments. In this way, we obtain a *semidefinite relaxation* of the original non convex program. The reader can find a good introduction for semidefinite programming in [1] and [3]. Many authors call these relaxations *LMI relaxations* as they are derived by using a single linear matrix inequality. See [7], [8] and [9].

We should strongly emphasize that this paper is not meant to be a pioneer paper on semidefinite relaxations of polynomial programs as several accounts on the subject have already been published i.e. [7],[8],[9],[14] and [17]. Nevertheless, our objective here is to present a different approach to the subject by stressing the general span of the Method of Moments as it is described in Section 2, where we obtain a general characterization of the global minima of arbitrary non convex functions defined on arbitrary non convex domains. See [10] for a complete analysis of this situation. On the other hand, the approach followed in this paper relies on the analysis of the duality between a particular moment cone and the corresponding cone of positive functions as it was proposed since Tchebychev times and stated in [6], which allows us to apply the theory of the Method of Moments for characterizing the global minima of a particular non convex polynomial program. It is very important noticing here that other works on this subject use the duality of the theory of semidefinite programming in order to characterize the global optimum of one particular polynomial program, see [8]. It is also very important stress here that these ideas can be extended to higher dimensions where we must cope with constrained polynomial programs involving several variables, this important goal has been recently attained by using Putinar's characterization of positive polynomials on semialgebraic sets in R^n. The reader can find a complete account of these results in [7],[8] and [9].

The present paper is organized as follows. In Section 2 we will review the general theory behind the Method of Moments, and we will deduce the most general results for global optimization of non convex mathematical programs. In Section 3, we will explain how to solve the classical one dimensional truncated moment problems. Indeed, we will see how to solve the Hamburger's Moment Problem, the Trigonometric Moment Problem, the Stieltjes Moment Problem and the Hausdorff's Moment Problem by reducing all of them to one of the classical characterizations of one dimensional positive polynomials as has been proposed by Curto and Fialkow in [4], by Krein and Nudelman in [6] and by Nesterov in [14]. Then we will apply these results for analyzing one dimensional non convex polynomial programs. Finally, we will give some comments and remarks in Section 4.

2 General Theory of the Method of Moments

In this paper we are concerned with the search of all global minima of a continuous function $f : \Omega \to R$, defined in some arbitrary closed set $\Omega \subset R^n$. This problem is particularly difficult when no assumptions about the convexity of f or Ω are available. Nonetheless, progress in this kind of problems can be achieved by using convex analysis and measure theory.

2.1 Basic Results

We will see here how we can analyze global optimization problems when its objective function f is expressed as a linear combination of simpler functions. This is the general idea of many applications of the Method of Moments. We start by using a powerful result proposed in [8] for global optimization of polynomials and also proposed in [12] for Young measures relaxation of non convex variational problems.

2.1.1 Theorem

Let $P(\Omega)$ be the set of all regular Borel probability measures supported in a closed set $\Omega \subset R^n$, and let f be a bounded from below continuous function $f : \Omega \to R$, then

$$\inf_{\mu \in P(\Omega)} \langle f, \mu \rangle = \inf_{t \in \Omega} f(t). \tag{2}$$

Proof

By elementary integration we have that $m \leq \langle f, \mu \rangle$ for every probability measure μ, where $m = \inf_{t \in \Omega} f(t)$. On the other hand, it is clear that $\langle f, \delta_{t_n} \rangle \to m$, where $\{t_n\} \subset \Omega$ is a minimizing sequence for the function f, and δ_t is the Dirac's measure supported in the point t.

2.1.2 Theorem

Let G be the set of all global minima of the function f in Ω, then

$$\langle f, \mu^* \rangle = \inf_{\mu \in P(\Omega)} \langle f, \mu \rangle \tag{3}$$

if and only if the support of μ^* is contained in G.

Proof

It is easy to see that $\langle f, \mu^* \rangle = m$ when $\mu^* \in P(G)$. We will verify that $\mu^* \in P(G)$ when $\langle f, \mu^* \rangle = m$. Let us assume that $t \in supp(\mu) \cap G^c$, then $f(t) > m$. Since f is continuous, there exists a neighborhood U of t such that $f \geq \gamma > m$ in U. On the other hand, $\mu(U) > 0$ because $t \in supp(\mu)$. So, $\langle f, \mu \rangle = \int_U f d\mu + \int_{U^c} f d\mu \geq \gamma\mu(U) + m\mu(U^c) > m$.

From these results, it follows that we should use the generalized optimization problem in measures

$$\min_{\mu \in P(\Omega)} \langle f, \mu \rangle \tag{4}$$

as an alternative formulation of the global optimization problem

$$\min_{t \in \Omega} f(t). \tag{5}$$

Thus, we obtain a new optimization problem with two significant features in optimization, namely, a linear objective function: $\mu \to \langle f, \mu \rangle$ and a convex feasible set: $P(\Omega)$. In addition, this formulation includes all the information about the solution of the standard global optimization problem (5). Indeed, the set $P(G)$, which is composed of all probability measures supported in G, is the solution set for the generalized problem (4).

When the objective function f can be expressed by a linear combination of simpler functions

$$f = \sum_{i=1}^{k} c_i \psi_i \tag{6}$$

where $\{\psi_1, \ldots, \psi_k\}$ is a basis of continuous functions in Ω, then every integral in (4) can be expressed by an elementary dot product in R^k

$$\langle f, \mu \rangle = \sum_{i=1}^{k} c_i x_i = c \cdot x \tag{7}$$

whose factors are the coefficients vector c from the linear combination (6) and the moment vector x whose entries

$$x_i = \int \psi_i d\mu, \quad for \quad i = 1, \ldots, k \tag{8}$$

are the moments of the measure μ with respect to the basis functions $\{\psi_1, \ldots, \psi_k\}$. Every moment vector x may also be expressed by integration in the following manner:

$$x = \langle T, \mu \rangle = \int T d\mu, \tag{9}$$

where T is the non linear transformation $T : \Omega \to R^k$ defined by the expression $T(t) = (\psi_1(t), \ldots, \psi_k(t))$. For convenience, henceforth we assume that application T is one to one.

With the help of the transformation T, we can easily define the set

$$V = \{\langle T, \mu \rangle, \mu \in P(\Omega)\} \tag{10}$$

which consists of all moment vectors of probability measures supported in Ω. Since the application

$$\mu \to \langle T, \mu \rangle : P(\Omega) \to R^k \tag{11}$$

is linear, we immediately observe that V is a convex set in R^k. In this way, we can represent every measure $\mu \in P(\Omega)$ by its respective moment vector $x \in V$. By using this representation, we can transform the generalized optimization problem (4) into the equivalent convex program

$$\min_{x \in V} c \cdot x \tag{12}$$

whose solution set is the convex set

$$W = \{\langle T, \mu \rangle, \mu \in P(G)\} \tag{13}$$

where G is the set of all global minima of the function f in Ω.

2.2 Interplay with Convex Analysis

It is remarkable that problem (12) is a convex mathematical program which encloses the information about the global minima of the objective function f in Ω. In fact, it gives a non trivial characterization of the global minima of the function f in the set Ω. Before we show the solution to the global optimization problem (5) given by the convex program (12), we have to introduce additional results that link convex analysis and measure theory.

Since the image $T(\Omega)$ is contained in the Euclidean space R^k, its convex envelope can be expressed as

$$co(T(\Omega)) = \{\langle T, \mu \rangle, \mu \in Q(\Omega)\} \tag{14}$$

where $Q(\Omega)$ is the family of all finitely supported probability measures in Ω. This means that $co(T(\Omega))$ is just the set of moment vectors of finitely supported measures in Ω. Then, $co(T(\Omega)) \subset V$.

The theory of the Method of Moments uses a convenient interplay between measure theory and convex analysis for describing the moment vector set V. In fact, from Carathedory's Theorem we know that every point of the convex hull $co(T(\Omega))$ may be represented by a convex combination with less than $k+2$ terms. On the other hand, every regular Borel probability measure can be approximated by discrete probability distributions whose moment vectors belong to the convex hull $co(T(\Omega))$. In this way, we obtain a method for approximating every point in V by convex combinations with a fixed number of terms. This is the reason that explains why the Method of Moments works particularly well in many practical applications. The following theorems confirm this remark.

2.2.1 Theorem

The convex hull $co(T(\Omega))$ is dense in the moment vectors set V, so $\overline{V} = \overline{co(T(\Omega))}$.

Proof

For a μ-measurable positive function f, its integral $\langle f, \mu \rangle$ is usually defined as the supremum of all integrals $\langle s, \mu \rangle$, where every s is a *simple function* satisfying $s \leq f$. A simple function has the form $s = \sum_{i=1}^{j} c_i \chi_{A_i}$ where every A_i is a Borel set with characteristic function χ_{A_i}. The integral of s with respect to the measure μ is defined by the elementary sum $\langle s, \mu \rangle = \sum_{i=1}^{j} c_i \mu (A_i)$. Since f is continuous, the definition of the integral $\langle f, \mu \rangle$ does not change if we consider simple functions with the form $s = \sum_{i=1}^{j} f(t_i) \chi_{A_i}$, where $t_i \in A_i$ for every index i. Thus we have $\langle s, \mu \rangle = \sum_{i=1}^{j} f(t_i) \mu (A_i) = \left\langle f, \sum_{i=1}^{j} \mu (A_i) \delta_{t_i} \right\rangle$. Finally, we extend this conclusion to every basis function ψ_i.

2.2.2 Theorem

When the domain Ω is compact, so are the convex hulls $co\,(T\,(\Omega))$ and $co\,(T\,(G))$.

Proof

Convex hulls of compact sets are compact.

2.2.3 Theorem

The convex hull $co\,(T\,(\Omega))$ and the set V of moment vectors are not necessarily closed.

Proof

By taking $\Omega = R$, $\psi_1 (t) = t$ and $\psi_2 (t) = e^{-t^2}$, we found that $co\,(T\,(\Omega))$ fails to be closed. In order to verify that V does not need to be closed, we will exhibit an example suggested by Pedregal in [11]. Take the probability measure $\mu = \lambda \delta_{t_1} + (1 - \lambda) \delta_{t_2}$ where

$$t_1 = -\left(\frac{1-\lambda}{\lambda}\right)^{\frac{1}{4}}, \quad t_2 = \left(\frac{\lambda}{1-\lambda}\right)^{\frac{1}{4}}, \quad 0 < \lambda < 1. \tag{15}$$

Notice that $\int (1, t, t^2, t^3, t^4)\, d\mu \rightarrow (1,0,0,0,1)$ when $\lambda \rightarrow 1$, which means that the five algebraic moments of μ converge to the vector $(1,0,0,0,1)$ when $\lambda \rightarrow 1$. However, it is easy to see that there is no positive measure on the real line with values $(1,0,0,0,1)$ as its five first algebraic moments.

2.2.4 Theorem

Let $Q_{k+1} (\Omega)$ be the set of all probability measures in Ω supported in $k+1$ points at most, then $co\,(T\,(\Omega)) = \{\langle T, \mu \rangle, \mu \in Q_{k+1} (\Omega)\}$ when T is a one to one application.

Proof

Apply Caratheodory´s theorem from Convex Analysis.

2.2.5 Theorem

Let $\mu^\varepsilon \in Q\,(\Omega)$ for every $\varepsilon > 0$ and let $y = \lim_{\varepsilon \rightarrow 0} \langle T, \mu^\varepsilon \rangle$. If y does not belong to $co\,(T\,(\Omega))$, then there exists an unbounded subsequence $\{t^{\varepsilon_n}\} \subset \Omega$ such that every t^{ε_n} belongs to the support of the measure μ^{ε_n}.

Proof

From Theorem 2.2.4, we can express every moment vector $x^\varepsilon = \langle T, \mu^\varepsilon \rangle$ by using a convex combination with at most $k+1$ terms:

$$x^{\varepsilon} = \lambda_1^{\varepsilon} T\left(t_1^{\varepsilon}\right) + \cdots + \lambda_{k+1}^{\varepsilon} T\left(t_{k+1}^{\varepsilon}\right) \tag{16}$$

where points $t_1^{\varepsilon}, \ldots, t_{k+1}^{\varepsilon}$ form the support of the measure μ^{ε}. If every family $\{t_i^{\varepsilon} : \varepsilon > 0\}$ is bounded, then there exists a converging sub sequence x^{ε_n} such that

$$y = \lim_{n\to\infty} x^{\varepsilon_n} = \overline{\lambda}_1 T\left(\bar{t}_1\right) + \cdots + \overline{\lambda}_{k+1} T\left(\bar{t}_{k+1}\right) = \langle T, \overline{\mu} \rangle \ , \tag{17}$$

where $\lim_{n\to\infty} \lambda_i^{\varepsilon_n} = \overline{\lambda}_i,\ \lim_{n\to\infty} t_i^{\varepsilon_n} = \bar{t}_i,$ for every $i = 1, \ldots, k+1$. Thus, the entries of the vector $y \in R^k$ are the moments of the finitely supported measure $\overline{\mu} = \sum_{i=1}^{k+1} \overline{\lambda}_i \delta_{\bar{t}_i}$ and this result contradicts the assumptions about y.

Although Theorem 2.2.2 states that the moment vector set V is closed in the particular cases where the feasible set Ω is bounded, Theorem 2.2.3 claims that convex set V does not need to be closed in general, so we should replace program (12) by the extended program

$$\min_{x\in\overline{V}} c \cdot x \, . \tag{18}$$

It is remarkable that any solution for this program is linked with a minimizing sequence of the function f. Let us assume that x^* is a solution to the extended program (18), and $x^* \in co\left(T\left(\Omega\right)\right)$, then $x^* = \langle T, \mu^* \rangle$ being μ^* a finitely supported measure in G. In this way, we can obtain a finite set of global minima of the function f by analyzing the support of μ^*. In the opposite case when $x^* \notin co\left(T\left(\Omega\right)\right)$, we can obtain a minimizing sequence for f by using a family of finitely supported measures μ^{ε} whose moments $x^{\varepsilon} = \langle T, \mu^{\varepsilon} \rangle$ approach x^* in R^k.

2.2.6 Theorem

If x^* is a solution to the extended program (18) in $co\left(T\left(\Omega\right)\right)$, then $x^* \in co\left(T\left(G\right)\right)$.

Proof

Since $c \cdot x^* = m \left(\equiv \inf_{\Omega} f\right)$ for $x^* = \langle T, \mu^* \rangle$ with $\mu^* \in Q\left(\Omega\right)$, we have $\langle f, \mu^* \rangle = m$. From Theorem 2.1.2, the support of measure μ^* consists of finitely many points in G.

2.2.7 Theorem

Let $\{x^{\varepsilon} = \langle T, \mu^{\varepsilon} \rangle : \varepsilon > 0\}$ be a family of moments in $co\left(T\left(\Omega\right)\right)$ such that $x^{\varepsilon} \to x^*$, where x^* is a solution to the extended program (18). Then, there exists a minimizing sequence $\{t^{\varepsilon_n}\} \subset \Omega$ for the function f, such that every term t^{ε_n} belongs to the support of the measure μ^{ε_n}.

Proof

From Theorem 2.2.4, we can write every moment vector x^{ε} by using the convex combination

$$x^{\varepsilon} = \langle T, \mu^{\varepsilon} \rangle = \lambda_1^{\varepsilon} T\left(t_1^{\varepsilon}\right) + \cdots + \lambda_{k+1}^{\varepsilon} T\left(t_{k+1}^{\varepsilon}\right) \tag{19}$$

where every measure $\mu^{\varepsilon} \in Q_{k+1}\left(\Omega\right)$. Since x^* is a solution for program (18), then

$$c \cdot x^{\varepsilon} = \langle f, \mu^{\varepsilon} \rangle = \lambda_1^{\varepsilon} f\left(t_1^{\varepsilon}\right) + \cdots + \lambda_{k+1}^{\varepsilon} f\left(t_{k+1}^{\varepsilon}\right) \to m \left(\equiv \inf_{\Omega} f\right). \tag{20}$$

Let us assume that the family $\{t_i^{\varepsilon} : \varepsilon > 0, i = 1, \ldots, k+1\}$ does not contain any minimizing sequences of the function f. Then $f\left(t_i^{\varepsilon}\right) > m + \gamma$ with $\gamma > 0$, which prevents $\langle f, \mu^{\varepsilon} \rangle$ from converging to m.

2.3 Global Optimization

The success of the theory of the Method of Moments relies on the important fact that it provides an alternative characterization of the global minima of the function f.

2.3.1 Theorem

Let us assume that f does not have any unbounded minimizing sequence. If $x^* \in R^k$ is a solution for the extended program (18), then there exist finitely many points $t_1, \ldots, t_\rho \in G$ and positive values $\lambda_1, \ldots, \lambda_\rho$ such that

$$x^* = \lambda_1 T(t_1) + \cdots + \lambda_\rho T(t_\rho) \quad \textit{and } 1 = \lambda_1 + \cdots + \lambda_\rho \tag{21}$$

where ρ may be chosen to be less than $k+2$.

Proof

Since f has no unbounded minimizing sequence, Theorem 2.2.7 implies that $x^* \in co(T(\Omega))$. Hence, $x^* = \langle T, \mu^* \rangle$ where μ^* is finitely supported. From Theorem 2.1.2, we can verify that every point in the support of μ^* is a global minimum of the objective function f. Thus we also have that $\mu^* \in P(G)$ because $\langle f, \mu^* \rangle = c \cdot x^* = m \, (\equiv \inf_\Omega f)$. Finally, from Theorem 2.2.4 we have $x^* = \langle T, \overline{\mu} \rangle$ where $\overline{\mu}$ is supported in $k+1$ points at the most. And applying Theorem 2.1.2 again, we conclude that $\overline{\mu}$ is supported in G.

2.3.2 Corollary

If finitely many points $t_1, \ldots, t_\rho \in \Omega$ satisfy (21), then every t_i is a global minimum of f in Ω.

Proof

By taking the measure $\mu^* = \sum_{i=1}^{\rho} \lambda_i \delta_{t_i}$ we found that $\langle f, \mu^* \rangle = c \cdot x^* = m$, then the support of μ^* is contained in G because of Theorem 2.1.2.

2.3.3 Theorem

Let us assume that f does not have any unbounded minimizing sequence. If $x^* \in R^k$ is an extreme point of the solution set of the convex program (18), then there exists a global minimum $t \in \Omega$, of the objective function f, satisfying the set of k non linear equations

$$x^* = T(t). \tag{22}$$

Proof

If $\rho > 1$ in (21), then x^* could not be an extreme point of the solution set of program (18).

2.3.4 Corollary

Every point $t \in \Omega$ satisfying the set of k non linear equations (22) is a global minimum of the objective function f.

Proof

Take $\mu^* = \delta_t$ and observe that $f(t) = \langle f, \mu^* \rangle = m$, then $t \in G$ because t is the support of μ^* and Theorem 2.1.2.

2.3.5 Theorem

For arbitrary points $t_1, \ldots, t_\rho \in G$ and positive values $\lambda_1, \ldots, \lambda_\rho$ satisfying $\lambda_1 + \cdots + \lambda_\rho = 1$, the point $x^* \in R^k$ defined by

$$x^* = \lambda_1 T(t_1) + \cdots + \lambda_\rho T(t_\rho) \tag{23}$$

is a solution of the extended program (18).

Proof

Take $c \cdot x^* = \lambda_1 f(t_1) + \cdots + \lambda_\rho f(t_\rho) = m\,(\equiv \inf_\Omega f)$.

From these results we conclude that a necessary and sufficient condition for a finite number of points $t_1, \ldots, t_\rho \in \Omega$ be a set of global minima of f , is that they satisfy the k non linear equations

$$x^* = \lambda_1 T(t_1) + \cdots + \lambda_\rho T(t_\rho) \tag{24}$$

for some positive $\lambda_1, \ldots, \lambda_\rho$ with $\lambda_1 + \cdots + \lambda_\rho = 1$, where $x^* \in R^k$ is a moment vector that solves the extended program (18). In order to estimate a particular set of global minima of the function f, we must solve equations (24) for a particular solution of the program (18). However, this question is equivalent to looking for a finitely supported measure μ^* whose moments (with respect to the basis $\psi_1, \ldots, \psi_k$) are the optimal values $x_1^*, \ldots, x_k^*$. Notice that we are mostly interested in finding the support of μ^* rather than determining the measure μ^*. The answer to this question comes again from the Problem of Moments where it is clarified how to recover a measure from its moments. Then, in order to apply the Method of Moments on specific problems, we need a proper characterization of the closure of the set V of all moment vectors and a practical method for recovering every finitely supported measure from its moments. The reader can find a recent survey on one dimensional truncated moment problems in [4].

3 One Dimensional Moment Problems

The Problem of Moments consists in determining the conditions which guarantee that values $x_1, \ldots, x_k$ are the moments of an arbitrary positive measure μ with respect to a particular basis of functions $\psi_1, \ldots, \psi_k$, which is defined in some domain Ω. The solution of the Problem of Moments also should provide techniques for recovering the measure μ from the sequence of moments $x_1, \ldots, x_k$. This is a classical problem in modern mathematics in which great mathematicians have been involved since the nineteenth century. For a classical introductory review on the Problem of Moments see [16]. Here we solve several truncated moment problems using a powerful tool from convex analysis. To attain this task, we will use the classical duality between the moment cone and the cone of the corresponding positive functions. This tool is introduced in [6] and can be traced back to the seminal works of Markov and Tchebychev. The reader should also see [14] for a different perspective on the links between moment theory and global optimization with linear algebra tools.

3.1 General Theory

Let us define

$$M = \left\{ x \in R^k : x_i = \int \psi_i d\mu, i = 1, \ldots k, \mu \textit{ positive mesure in } \Omega \right\} \tag{25}$$

as the set of moment vectors of all positive measures supported in Ω. We can easily see that M is a convex cone in R^k. We also define P as follows:

$$P = \left\{ c \in R^k : \sum_{i=1}^{k} c_i \psi_i(t) \geq 0, \forall t \in \Omega \right\}, \tag{26}$$

where the vectors c in R^k determine non-negative functions in Ω. It is also easy to check out that P is a closed convex cone in R^k. The usual way for solving moment problems is to analyze the cone P, since its dual is exactly the closure of the moment cone M.

3.1.1 Theorem

The dual of the cone P is the closure of the cone M.

Proof

For arbitrary vectors $c \in P$ and $x \in M$, we have

$$c \cdot x = \int \left(\sum_{i=1}^{k} c_i \psi_i \right) d\mu \geq 0. \tag{27}$$

Thus, $P \subset M^*$, and $\overline{M} \subset P^*$. If there exists a point $t_0 \in \Omega$ such that

$$\sum_{i=1}^{k} c_i \psi_i(t_0) < 0, \tag{28}$$

then $\sum_{i=1}^{k} c_i x_i^0 < 0$, where x^0 is the moment vector of the Dirac measure δ_{t_0}. Thus, $M^* \subset P$, and $P^* \subset \overline{M}$.

By using the duality statement of Theorem 3.1.1, we can find the answer to many classical truncated moment problems, provided we can properly characterize the corresponding family of positive functions P. This procedure has been carried out in [4], [6] and [14].

3.2 Hamburger's Moment Problem

When the function basis is the algebraic system $1, t, \ldots, t^{2r}$, and the domain Ω is the real line, the moment problem is referred to as the Hamburger's Moment Problem. It is well known, from elementary algebra, that every positive polynomial $\sum_{i=0}^{2r} c_i t^i$ on the real line can be expressed as the sum of two squares, that is

$$\sum_{i=0}^{2r} c_i t^i = \left(\sum_{i=0}^{r} a_i t^i \right)^2 + \left(\sum_{i=0}^{r} b_i t^i \right)^2. \tag{29}$$

However, it will be useful to express (29) by using quadratic forms, so we claim that every positive polynomial $\sum_{i=0}^{2r} c_i t^i$ on the real line can be expressed in the following form:

$$\sum_{i=0}^{2r} c_i t^i = \sum_{i=0}^{r} \sum_{j=0}^{r} a_i a_j t^{i+j} + \sum_{i=0}^{r} \sum_{j=0}^{r} b_i b_j t^{i+j}. \tag{30}$$

For solving the classical Hamburger's Moment Problem, we only need to apply previous duality statement and the decomposition (30) for positive polynomials. If values $x_0, \ldots, x_{2r}$ are the algebraic moments of a positive measure supported on the real line, then $x \in P^*$, and thus

$$\sum_{i=0}^{2r} c_i x_i \geq 0 \tag{31}$$

for the coefficients c of every positive polynomial $\sum_{i=0}^{2r} c_i t^i$. In particular, for arbitrary values $a_0, \ldots, a_r$ we have

$$\sum_{i=0}^{r} \sum_{j=0}^{r} a_i a_j x_{i+j} \geq 0 \tag{32}$$

due to $\left(\sum_{i=0}^{r} a_i t^i\right)^2 \geq 0$. Thus, we conclude that a necessary condition for a vector $x \in R^{2k+1}$ to be a moment vector, is that its components form a positive semidefinite Hankel matrix $H = (x_{i+j})_{i,j=0}^{r}$.

On the other hand, assuming that the entries of a vector $x \in R^{2r+1}$ compose a positive semidefinite Hankel matrix $H = (x_{i+j})_{i,j=0}^{r}$, we can see from expression (30) that

$$\sum_{i=0}^{2r} c_i x_i = \sum_{i=0}^{r} \sum_{j=0}^{r} a_i a_j x_{i+j} + \sum_{i=0}^{r} \sum_{j=0}^{r} b_i b_j x_{i+j} \geq 0 \tag{33}$$

for the coefficients c of every positive polynomial $\sum_{i=0}^{2r} c_i t^i$ on the real line. Therefore, $x \in P^* = \overline{M}$ and we conclude that x is a moment vector or at least it is a limit point of a sequence of moment vectors. This procedure may be applied to obtain the characterization of the moment vectors for other bases and domains.

3.3 Trigonometric Moment Problem

From the Riesz-Fejer Theorem in complex analysis, we know that every positive trigonometric polynomial $\sum_{i=-r}^{r} c_i e^{ijt}$ can be expressed as

$$\sum_{i=-r}^{r} c_i e^{ijt} = \left| \sum_{k=0}^{r} a_k e^{kjt} \right|^2 = \sum_{l=0}^{r} \sum_{k=0}^{r} a_k \overline{a_l} e^{j(k-l)t}. \tag{34}$$

By using the quadratic form (34) and the arguments explained above, we easily solve the Trigonometric Moment Problem.

The closure of the cone M of all moment vectors of positive measures supported in the unitary circumference S^1, with respect to the trigonometric system $e^{-rjt}, \ldots, e^{rjt}$, is the set of all vectors $x \in C^{2r+1}$ whose entries form a positive semidefinite Toeplitz matrix $T = (x_{k-l})_{k,l=0}^{r}$.

3.4 Stieltjes and Hausdorff's Moment Problem

The Stieltjes Moment Problem arises when we consider the algebraic system $1, t, \ldots, t^r$ on the semi-axis $\Omega = [0, \infty)$ of the real line. If we restrict the domain to a bounded interval $\Omega = [a, b]$, we obtain the Hausdorff's Moment Problem.

3.4.1 Solution to the Stieltjes Problem - Even Case

The closure of the cone M of all moment vectors of positive measures supported in the semiaxis $[0,\infty)$, with respect to the algebraic system $1, t \ldots, t^{2r}$, is the set of all vectors $x \in R^{2r+1}$ whose entries form two positive semidefinite Hankel matrices given by

$$H_1 = (x_{i+j})_{i,j=0}^{r} \quad H_2 = (x_{i+j+1})_{i,j=0}^{r-1}. \tag{35}$$

Proof

Since we can express an arbitrary, even degree, non negative polynomial $\sum_{i=0}^{2r} c_i t^i$ on the semiaxis $[0,\infty)$ in the form

$$\sum_{i=0}^{2r} c_i t^i = \left(\sum_{i=0}^{r} a_i t^i\right)^2 + t\left(\sum_{i=0}^{r-1} b_i t^i\right)^2 \tag{36}$$

see [6]. This expression may be rewritten by the following couple of quadratic forms

$$\sum_{i=0}^{2r} c_i t^i = \sum_{i=0}^{r}\sum_{j=0}^{r} a_i a_j t^{i+j} + \sum_{i=0}^{r-1}\sum_{j=0}^{r-1} b_i b_j t^{i+j+1}. \tag{37}$$

Then, we can repeat the arguments used in the proof of Hamburger´s Moment Problem.

3.4.2 Solution to the Stieltjes Problem - Odd Case

The closure of the cone M of all moment vectors of positive measures supported in the semiaxis $[0,\infty)$, with respect to the algebraic system $1, t \ldots, t^{2r+1}$, is the set of all vectors $x \in R^{2r+2}$ whose entries form two positive semidefinite Hankel matrices with the following form

$$H_1 = (x_{i+j})_{i,j=0}^{r} \quad H_2 = (x_{i+j+1})_{i,j=0}^{r}. \tag{38}$$

Proof

For odd degree, positive polynomials in $[0,\infty)$, we have the analogous expression

$$\sum_{i=0}^{2r+1} c_i t^i = \left(\sum_{i=0}^{r} a_i t^i\right)^2 + t\left(\sum_{i=0}^{r} b_i t^i\right)^2 \tag{39}$$

which can be written as the sum of two quadratic forms:

$$\sum_{i=0}^{2r+1} c_i t^i = \sum_{i=0}^{r}\sum_{j=0}^{r} a_i a_j t^{i+j} + \sum_{i=0}^{r}\sum_{j=0}^{r} b_i b_j t^{i+j+1}. \tag{40}$$

3.4.3 Solution to the Hausdorff's Problem - Even Case

The closure of the cone M of all moment vectors of positive measures supported in the bounded interval $[\kappa_1, \kappa_2]$, with respect to the algebraic system $1, t \ldots, t^{2r}$, is the set of all *vectors $x \in R^{2r+1}$ whose entries make positive semidefinite the following symmetric matrices:*

$$H_1 = (x_{i+j})_{i,j=0}^{r} \quad H_2 = ((\kappa_1 + \kappa_2)\, x_{i+j+1} - \kappa_1 \kappa_2 x_{i+j} - x_{i+j+2})_{i,j=0}^{r-1}. \tag{41}$$

Proof

From Markov-Luckas Theorem [6], we can express every even degree positive polynomial on the bounded interval $\Omega = [\kappa_1, \kappa_2]$ as

$$\sum_{i=0}^{2r} c_i t^i = \left(\sum_{i=0}^{r} a_i t^i\right)^2 + (t - \kappa_1)(\kappa_2 - t)\left(\sum_{i=0}^{r-1} b_i t^i\right)^2. \tag{42}$$

Then we can write this expression by using the following quadratic forms:

$$\sum_{i=0}^{r}\sum_{j=0}^{r} a_i a_j t^{i+j} + \sum_{i=0}^{r-1}\sum_{j=0}^{r-1} b_i b_j \left((\kappa_1 + \kappa_2)\, t^{i+j+1} - \kappa_1 \kappa_2 t^{i+j} - t^{i+j+2}\right) \tag{43}$$

and repeat the arguments for the Hamburger´s Moment Problems.

3.4.4 Solution to the Hausdorff's Problem - Odd Case

The closure of the cone M of all moment vectors of positive measures supported in the bounded interval $[\kappa_1, \kappa_2]$, with respect to the algebraic system $1, t \ldots, t^{2r+1}$, is the set of all vectors $x \in R^{2r+2}$ whose entries form two positive semidefinite symmetric matrices given by

$$H_1 = (x_{i+j+1} - \kappa_1 x_{i+j})_{i,j=0}^{r} \quad H_2 = (\kappa_2 x_{i+j} - x_{i+j+1})_{i,j=0}^{r}. \tag{44}$$

Proof

Again, from Markov-Luckas Theorem, we can express every odd degree positive polynomial on the bounded interval $\Omega = [\kappa_1, \kappa_2]$ as

$$\sum_{i=0}^{2r+1} c_i t^i = (t - \kappa_1)\left(\sum_{i=0}^{r} a_i t^i\right)^2 + (\kappa_2 - t)\left(\sum_{i=0}^{r} b_i t^i\right)^2. \tag{45}$$

Hence we have the quadratic form expression

$$\sum_{i=0}^{r}\sum_{j=0}^{r} a_i a_j \left(t^{i+j+1} - \kappa_1 t^{i+j}\right) + \sum_{i=0}^{r}\sum_{j=0}^{r} b_i b_j \left(\kappa_2 t^{i+j} - t^{i+j+1}\right). \tag{46}$$

At this stage we should note the important fact that every one dimensional moment problem was solved by using a particular set of quadratic forms coming from the classical characterizations of positive polynomials in intervals.

3.5 Measure Recovery

The second question behind a particular moment problem is about the construction of a measure μ from a set of values $x_1, \ldots, x_k$ which are supposed to be the moments of μ. Once again, this is a very difficult problem in modern mathematics. However, for one dimensional algebraic and trigonometric moment problems we have the right answer. In Section 5 of [10] the reader can find the proper methods for obtaining a finitely supported measure μ from a finite sequence of one dimensional moments. These methods come from the characterization of one dimensional truncated moments provided by Curto and Fialkow in [4].

These results are briefly recalled in the following. If we take the values $x_0, \ldots, x_{2r}$ as the algebraic moments of a positive measure μ supported on the real line, its supporting points can be estimated by finding the roots of the polynomial

$$P(t) = \begin{vmatrix} x_0 & x_1 & \cdots & x_j \\ & & \cdots & \\ x_{j-1} & x_j & \cdots & x_{2j-1} \\ 1 & t & \cdots & t^{2j} \end{vmatrix} \tag{47}$$

where j is linked with the rank of the Hankel matrix $H = (x_{i+j})_{i,j=0}^{r}$. See [10] for a similar result applying on truncated trigonometric moments problem. Observe here that we only need to know the supporting points of μ to determine the global minima of f by the Method of Moments.

3.6 One Dimensional Polynomial Programs

The solutions presented here to one dimensional moment problems allow us to apply the theory of the Method of Moments for solving mathematical programs involving one-dimensional polynomials. Since we have characterized the one dimensional algebraic moments of positive measures on the line, and because we have a practical method to estimate its supporting points from its moment sequence, then we can fruitfully apply the general theory of the Method of Moments for analyzing arbitrary, non convex, one dimensional polynomial programs. A detailed exposition of the application of the Method of Moments to one dimensional polynomial programs may be found in [10].

For instance, to estimate the global minima of a particular one-dimensional polynomial given by

$$f(t) = \sum_{i=0}^{2r} c_i t^i \tag{48}$$

we should solve the corresponding semidefinite program:

$$\min_x \sum_{i=0}^{2r} c_i x_i \quad s.t. \quad (x_{i+j})_{i,j=0}^{r} \geq 0 \textit{ and } x_0 = 1. \tag{49}$$

As the polynomial f has no unbounded minimizing sequence, we conclude that every solution of the convex program (49) provides a set of global minima of the polynomial f in R.

3.6.1 Theorem

For every solution $x^* \in R^{2r+1}$ of the semidefinite program (49), there exist finitely many points $t_1, \ldots, t_\rho \in G$ and positive values $\lambda_1, \ldots, \lambda_\rho$ satisfying the equations

$$x_j^* = \lambda_1 t_1^j + \cdots + \lambda_\rho t_\rho^j, \quad j = 0, \ldots, 2r \tag{50}$$

where G is the set of all global minima of the polynomial f given in (48). Here ρ may be chosen to be less than $2r+3$.

Proof

Apply Theorem 2.3.1 and the solution of the Hamburger´s Moment Problem.

3.6.2 Theorem

A necessary and sufficient condition for finitely many points

$$t_1, \ldots, t_\rho \in R \tag{51}$$

to be global minima of the polynomial f given in (48), is that the following $2r+1$ equations

$$x_j^* = \lambda_1 t_1^j + \cdots + \lambda_\rho t_\rho^j, \quad j = 0, \ldots, 2r \tag{52}$$

hold true for some solution $x^* \in R^{2r+1}$ of the semidefinite program (49) and positive values $\lambda_1, \ldots, \lambda_\rho$.

Proof

Apply Theorem 2.3.1, Corollary 2.3.2 and the solution of the Hamburger´s Moment Problem.

3.6.3 Corollary

If x^* is an extreme point of the solution set of program (49), then x_1^* is a global minimum in R of the polynomial f given by the expression (48).

Proof

Apply Theorem 2.3.3 and the solution of the Hamburger's Moment Problem.

Since we have obtained an explicit method for determining a finitely supported measure from a sequence of its algebraic moments, we can find the global minima of any one dimensional algebraic polynomial with the form (48). Let us assume that $x^* \in R^{2r+1}$ is a solution of the semidefinite program (49), then the roots of the polynomial

$$P^*(t) = \begin{vmatrix} x_0^* & x_1^* & \cdots & x_j^* \\ & & \cdots & \\ x_{j-1}^* & x_j^* & \cdots & x_{2j-1}^* \\ 1 & t & \cdots & t^{2j} \end{vmatrix} \tag{53}$$

are global minima of the polynomial f in (48). See the proof of this fundamental result in [10].

Each one of the classical one dimensional moment problems allows us to solve a general family of non convex one dimensional polynomial programs. The following results illustrate this statement.

3.6.4 Theorem

A necessary and sufficient condition for finitely many points

$$z_1, \ldots, z_\rho \in S^1 \equiv \{z \in C : |z| = 1\} \tag{54}$$

to be global minima of the trigonometric polynomial

$$f(z) = \sum_{i=-r}^{r} c_i z^i \tag{55}$$

is that the equations

$$x_j^* = \lambda_1 z_1^j + \cdots + \lambda_\rho z_\rho^j, j = -r, \ldots, r \tag{56}$$

hold true for some solution $x^* \in C^{2r+1}$ of the semidefinite program

$$\min_x \sum_{i=-r}^{r} c_i x_i \quad s.t. \quad (x_{k-l})_{k,l=0}^{r} \geq 0 \quad and \quad x_0 = 1 \tag{57}$$

and positive values $\lambda_1, \ldots, \lambda_\rho$.

Proof

Apply Theorem 2.3.1, Corollary 2.3.2 and the solution of the Trigonometric Moment Problem.

The procedure for recovering a finitely supported measure from its trigonometric moments is explained in [10].

We can settle similar results for global optimization of one dimensional polynomial programs defined in arbitrary intervals of the real line. By using the even cases of Stieltjes and Hausdorff's Moment Problems we obtain the following theorems. The reader can infer the analogous results for odd cases.

3.6.5 Theorem

A necessary and sufficient condition for points $t_1, \ldots, t_\rho \geq 0$ to be global minima of the even degree, algebraic polynomial f given by (48), is that the equations

$$x_j^* = \lambda_1 t_1^j + \cdots + \lambda_\rho t_\rho^j, \quad j = 0, \ldots, 2r \tag{58}$$

hold true for some solution $x^* \in R^{2r+1}$ of the semidefinite program:

$$\min_x \sum_{i=0}^{2r} c_i x_i \quad s.t. \quad (x_{k+l})_{k,l=0}^{r} \geq 0, \ (x_{k+l+1})_{k,l=0}^{r-1} \geq 0 \ and \ x_0 = 1 \tag{59}$$

and positive values $\lambda_1, \ldots, \lambda_\rho$.

Proof

Apply Theorem 2.3.1, Corollary 2.3.2 and the solution for the even case of the Stieltjes Moment Problem.

3.6.6 Theorem

A necessary and sufficient condition for points $t_1, \ldots, t_\rho$ to be global minima in the interval $[\kappa_1, \kappa_2]$ of the even degree, algebraic polynomial f given by (48), is that the equations

$$x_j^* = \lambda_1 t_1^j + \cdots + \lambda_\rho t_\rho^j, \quad j = 0, \ldots, 2r \tag{60}$$

hold true for some solution $x^* \in R^{2r+1}$ of the semidefinite program:

$$\min_x \sum_{i=0}^{2r} c_i x_i \quad s.t. \quad (x_{k+l})_{k,l=0}^{r} \geq 0 \tag{61}$$

$$((\kappa_1 + \kappa_2)\, x_{k+l+1} - \kappa_1\kappa_2\, x_{k+l} - x_{k+l+2})_{k,l=0}^{r-1} \geq 0 \quad and \quad x_0 = 1 \tag{62}$$

and positive values $\lambda_1, \ldots, \lambda_\rho$.

Proof

Apply Corollary 2.3.2, Theorem 3.1.1 and the solution of the even case of the Hausdorff's Moment problem.

In this section, we have applied the Method of Moments for transforming a one dimensional polynomial program into an equivalent semidefinite program. The key in this procedure is to find a convenient positive semidefinite quadratic form which characterizes the algebraic moments of positive measures on intervals. See [1] for a review on semidefinite programming.

3.7 Example

Let us illustrate the theory introduced by treating the next polynomial program

$$\min_{t \in R}\ 0.2\, t^4 - t^3 + t^2 + t \tag{63}$$

which is described by a non convex curve with only one global minima. Its semidefinite relaxation takes the form:

$$\begin{array}{l} \min\ 0.2\, m_4 - m_3 + m_2 + m_1 \\ s.t. \\ \begin{pmatrix} 1 & m_1 & m_2 \\ m_1 & m_2 & m_3 \\ m_2 & m_3 & m_4 \end{pmatrix} \geq 0 \end{array} \tag{64}$$

and it can be easily formulated as a semidefinite program. By using available routines for semidefinite programming described in [5], we obtain the following answer:

$$\begin{array}{l} m_1^* = -0.3263 \\ m_2^* = 0.1065 \\ m_3^* = -0.0348 \\ m_4^* = 0.0113 \end{array} \tag{65}$$

and we conclude that the non convex polynomial in (63) has a global minima at point $t^* = -0.3263$.

4 Concluding Remarks

In this paper we provide a general characterization of global minima of arbitrary programs as has been stated in Section 2. Thus, we have seen how to treat one dimensional non convex polynomial programs by reducing them to a single semidefinite program which encloses the information about the optimal solutions of the original problem. We have actually attained this task by focussing on the particular quadratic form that solves the truncated moment problem in a given real interval. We also stress that several global minima of one dimensional polynomial programs can be estimated by using algebraic tools proposed by Curto and Fialkow in [4] . These methods have been carefully explained in [10]. For the analysis of higher dimension polynomial programs under constraints the reader should refer to [7], [8] and [9].

Acknowledgments: The author wishes to thank Dr. Jean B. Lasserre for his suggestions on bibliography about moment theory and optimization kindly provided at ACAGO conference held in Greece in 2000. This work has been supported by the following research grants: *Proyecto 1596-Fundación para la Investigación y la Tecnología* and *Proyecto Aula Cimne Uniandes.*

References

[1] Ben-Tal, A. and A. Nemirovski, Lectures on Modern Convex Optimization, MPS-SIAM, 2001.

[2] Berg, C. et al., A remark on the multidimensional moment problem, Math. Ann. 223, p.163-169, 1979.

[3] Boyd, S. et al., Linear Matrix Inequalities and Control Theory, SIAM, 1994.

[4] Curto, R. and L.A. Fialkow, Recursiveness, positivity and truncated moment problems, Houston Journal of Mathematics, vol. 17, No. 4, 1991.

[5] Gahinet, P. et al., LMI Control Toolbox User's Guide, The MathWorks Inc., 1995.

[6] Krein, M.G. and A.A. Nudel'man, The Markov Moment Problem and Extremal Problems, Translations of Mathematical Monographs, vol. 50, AMS, 1977.

[7] Lasserre, J., Semidefinite programming vs lp relaxations for polynomial programming, Mathematics of Operations Research Journal, vol. 27, No 2, pp. 347-360, 2002.

[8] Lasserre, J., Global optimization with polynomials and the problem of moments, SIAM J. Optim., vol 11, No 3, pp. 796-817, 2001.

[9] Lasserre, J., New positive semidefinite relaxations for nonconvex quadratic programs, in Advances in Convex Analysis and Global Optimization, Non Convex Optimization and Its Applications Series, vol. 54, Kluwer, 2001.

[10] Meziat, R., The method of moments in global optimization, Journal of Mathematical Sciences, vol 116, No3, pp. 3303-3324, Kluwer, 2003.

[11] Meziat, R., P. Pedregal and J.J. Egozcue, From a nonlinear, nonconvex variational problem to a linear, convex formulation, J. Appl. Math. Optm., vol 47, pp. 27-44, Springer Verlag, New York, 2003.

[12] Meziat, R., P. Pedregal and J.J. Egozcue, The method of moments for non convex variational problems, in Advances in Convex Analysis and Global Optimization, Non Convex Optimization and Its Applications Series, vol. 54, pp. 371-382, Kluwer, 2001.

[13] Meziat, R., Two dimensional non convex variational problems, to appear in the proceedings of the International Workshop in Control and Optimization, Erice, Italy, 2001.

[14] Nesterov Y., Squared functional systems and optimization problems, in : High Performance Optimization, H. Frenk, K. Roos, T. Terlaky, S. Zhang eds, Kluwer Academic Publishers, Dordrecht, 2000.

[15] Putinar, M., Positive polynomials on compact semi-algebraic sets, Indiana University Mathematics Journal, vol. 42, No. 3, pp. 969-984, 1993.

[16] Shoat, J.A. and J.D. Tamarkin, The Problem of Moments, Mathematical Surveys 1, AMS, 1943.

[17] Shor, N.Z., Nondifferentiable Optimization and Polynomial Problems, Kluwer, 1998.