Squared Functional Systems and Optimization Problems

Yu. Nesterov [*]

March 4, 1998

Abstract

In this paper we give an explicit description of the cones of polynomials representable as a sum of squared functions. We prove that such cones can be always seen as a linear image of the cone of positive semidefinite matrices. As a consequence of the result, we get a description of the cones of univariate polynomials, which are non-negative on a ray and on an interval, and a description of non-negative trigonometric polynomials. We discuss some applications of the results to multi-variate polynomials.

1 Introduction

Recent development of the polynomial time interior-point methods in Nonlinear Programming was essentially based on the notion of self-concordant barrier [6]. In order to apply such schemes to a convex optimization problem, we need to point out a special computable barrier function for its feasible set. In many cases, when the feasible set of the problem is defined by a finite number of convex inequality constraints, the corresponding self-concordant barriers can be constructed rather easily. However, in some important applications we can meet convex sets, which can be seen as a non-trivial intersection of infinitely many linear half-spaces. A good example of such a set delivers the cone of coefficients of univariate polynomials, which are non-negative on some segment of the real axis. A standard way of treating problems with that type of constraints consists in introducing a large enough but finite number of linear inequalities which provides us with an acceptable approximation of the feasible set (see [1] and [9] for a description of the technique and a collection of application's examples). In this paper we show that such sets can be represented as linear images of the cone of positive semi-definite matrices. Therefore the corresponding optimization problems can be solved by applying the powerful modern schemes of semidefinite programming [7, 8].

[*]CORE, Catholic University of Louvain, 34 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium; e-mail: nesterov@core.ucl.ac.be. This work was supported in a part by grant N 96-01-00293 of Russian Fond of Fundamental Research and the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. Submitted to "High Performance Optimization", Kluwer, 1998.

It is interesting that some results of this paper can be obtained from the theory of Tchebyshev systems [2]. This theory, originated from the classical results by Tchebyshev and Markov, considers the generalized polynomials of the following form:

$$p(x) = \sum_{i=1}^{n} p^{(i)} v^{(i)}(x), \quad x \in \Delta,$$

where the functional system $\{v^{(i)}(x)\}_{i=1}^{n}$ satisfies the Tchebyshev property. That is that no polynomial $p(x)$, $x \in \Delta$, can have more than $n+1$ distinct roots except $p(x) \equiv 0$. In this theory the object of the main interest is the following *moment cone*:

$$\mathcal{M} = \{c = (c^{(1)}, \ldots, c^{(n)}) : \ c^{(i)} = \int\limits_{\Delta} v^{(i)}(x) d\sigma(x), \ i = 1, \ldots, n\},$$

where $\sigma(x)$ traverses the set of all nondecreasing right continuous functions of bounded variation. At the same time it is proved (Theorem 9.1 [2]) that the cone $\mathcal{P}$ dual to $\mathcal{M}$:

$$\mathcal{P} = \{p = (p^{(1)}, \ldots, p^{(n)}) : \ \sum_{i=1}^{n} p^{(i)} c^{(i)} \geq 0 \ \forall c \in \mathcal{M}\},$$

is comprised by the polynomials $p(x)$ which are non-negative on Δ. Note that now, using the standard technique of conic duality (e.g. [6]), it is very easy to get an explicit description of such a cone $\mathcal{P}$ provided that we know the structure of the cone $\mathcal{M}$. However, it seems that such representations were never presented in a direct form.

At the same time in the literature devoted to the Tchebyshev systems we can find an interesting description of many moment cones. Note that the most of them, if not all, are given in terms of linear matrix inequalities:

$$\mathcal{M} = \{c \in R^n : \ A_i(c) \succeq 0, \ i = 1, \ldots, k\}, \tag{1}$$

where the matrices A_i depend linearly on c and $A \succeq 0$ means that the matrix A is positive semidefinite. Such representations exist for the moment problems related to the standard power functions and to the trigonometric polynomials, for different variants of Nevanlinna-Pick problems, etc. And this is not a coincidence: we will see that the description of the cone $\mathcal{M}$ in the form (1) arise always when the inclusion $p \in \mathcal{P}$ is equivalent to a possibility to represent the function $p(x)$ as a sum of squares (or weighted squares).

This observation suggests a different way for deriving some classical results. Indeed, let $\mathcal{S} = \{u^{(1)}(x), \ldots, u^{(m)}(x)\}$, $x \in \Delta$, be an arbitrary system of linearly independent functions. Define the finite-dimensional functional subspace

$$\mathcal{F}(\mathcal{S}) = \{q(x) = \sum_{k=1}^{m} q^{(k)} u^{(k)}(x), \ q = (q^{(1)}, \ldots, q^{(m)}) \in R^m\}.$$

Let us try to characterise the following convex cone:

$$K = \{p(x) = \sum_{i=1}^{N} q_i^2(x), \ q_i(x) \in \mathcal{F}(\mathcal{S}), \ i = 1, \ldots, N\}.$$

We will see that the description of the cone K depends only on the properties of the *squared functional system*

$$\mathcal{S}^2 = \{v_{ij}(x) = u^{(i)}(x)u^{(j)}(x), \; i,j = 1,\ldots,m\}.$$

Note that this approach have the following advantages.

- We can reproduce the classical results using the theorems on representation of non-negative polynomials as a sum of squares. Note that these theorems were obtained independently on the theory of Tchebyshev systems.
- We address directly the object of our interest, the cone of non-negative polynomials. The description of the corresponding moment cones can be easily derived by the tools of conic duality.
- We can work with much wider family of convex cones than it is possible in the framework of the Tchebyshev theory. Recall ([4]) that we can define a Tchebyshev system on an abstract connected compact set Δ only if there exists a homeomorphism between Δ and a one-dimensional interval or a circle. Thus the theory of Tchebyshev systems works mainly with functions of one variable. In our situation we have no direct restrictions on the structure of the set Δ. Therefore we can work also with multi-variate functions.

We should prevent a reader from considering our results as a revision of the theory of Tchebyshev systems. This theory definitely represents one of the most beautiful examples of a comprehensive applied mathematical theory. However, the style, tools and goals of this classical theory are oriented on the questions, which can be answered in a closed form. Our paper can be seen as an attempt to look again at the same objects having in mind the needs and abilities of numerical methods.

The paper is organized as follows. In Section 2 we prove the representation theorems for the cones formed by sums (or weighted sums) of squared functions. In Section 3 we give a description of the cones of non-negative uni-variate polynomials. In the last Section 4 we discuss some applications of the results in nonlinear optimization (Sections 4.1–4.4), combinatorics (Sections 4.5, 4.6) and theory of polynomials (Section 4.7).

In this paper we use the following notation. We denote $\langle p, v\rangle$ the inner product of two real vectors p and v. The meaning of this notation is clear from the spaces of the arguments. If p and v are some vectors from R^n then

$$\langle p, v\rangle = \sum_{i=1}^{n} p^{(i)} v^{(i)}.$$

Notation $p(x)$ is used for the polynomial $p(x) = \langle p, v(x)\rangle$, where $v(x)$ is some functional system. We use this notation only when the corresponding functional system is well determined by the context. If A and B are $(n \times m)$-matrices (notation $A,\, B \in R^{n\times m}$) then

$$\langle A, B\rangle = \sum_{i=1}^{n}\sum_{k=1}^{m} A^{(ik)} B^{(ik)}.$$

Notation $A \succeq 0$ means that the matrix A is positive semidefinite; $A \succeq B$ if $A - B \succeq 0$. The sign $\succeq$ indicates that the both sides of this inequality are symmetric matrices. Notation I_n is used for a unit matrix in the corresponding n-dimensional space.

Notation specific for the complex variables is introduced in Section 2.3.

2 Representation theorems

2.1 Sum of squares

Let $\mathcal{S} = \{u^{(1)}(x), \dots, u^{(m)}(x)\}$, $x \in \Delta$, be an arbitrary system of linearly independent functions. Define a finite-dimensional functional subspace

$$\mathcal{F}(\mathcal{S}) = \{q(x) = \sum_{k=1}^{m} q^{(k)} u^{(k)}(x),\ q = (q^{(1)}, \dots, q^{(m)}) \in R^m\}.$$

In this section we are interested in the description of the following finite-dimensional cone:

$$K = \{p(x) = \sum_{i=1}^{N} q_i^2(x),\ q_i(x) \in \mathcal{F}(\mathcal{S}),\ i = 1 \dots N\}.$$

In this definition N is fixed and large enough, say $N \geq m$.

Consider the *squared* functional system

$$\mathcal{S}^2 = \{v_{ik}(x) = u^{(i)}(x) \cdot u^{(k)}(x),\ i, k = 1, \dots, m\}.$$

And let the components of the vector function $v(x) = (v^{(1)}(x), \dots, v^{(n)}(x))^T$ be the elements of some basis of a finite dimensional functional space which covers $\mathcal{S}^2$. Define the vector coefficients $\lambda_{ik} \in R^n$ as follows:

$$u^{(i)}(x) \cdot u^{(k)}(x) = \langle \lambda_{ik}, v(x) \rangle \quad \forall x \in \Delta.$$

Then we can introduce a matrix valued linear operator $\Lambda(v):\ R^n \to R^{m \times m}$, such that

$$(\Lambda(v))^{(ik)} = \langle \lambda_{ik}, v \rangle, \quad i, k = 1 \dots m.$$

Note that for $u(x) = (u^{(1)}(x), \dots, u^{(m)}(x))^T$ we have

$$u(x) u(x)^T \equiv \Lambda(v(x)), \quad x \in \Delta. \tag{2}$$

The adjoint linear operator $\Lambda^*(Y)$, $Y \in R^{m \times m}$, is defined as follows:

$$\langle Y, \Lambda(v) \rangle \equiv \langle \Lambda^*(Y), v \rangle \quad \forall Y \in R^{m \times m},\ v \in R^n.$$

Theorem 1 *1). The function $p(x) = \langle p, v(x) \rangle$, $p \in R^n$, belongs to K if and only if there exists a positive semidefinite $(m \times m)$-matrix Y such that $p = \Lambda^*(Y)$:*

$$K = \{p \in R^n :\ p = \Lambda^*(Y),\ Y \succeq 0\}. \tag{3}$$

This cone is closed convex and pointed.

$$\operatorname{rint} K = \{p \in R^n :\ p = \Lambda^*(Y),\ Y \succ 0\}. \tag{4}$$

2). $K^* = \{c \in R^n :\ \Lambda(c) \succeq 0\}$, $\operatorname{int} K^* \neq \emptyset$.

3). If $v(x)$ is a minimal system for $\mathcal{S}^2$ then K^* is a pointed cone and $\operatorname{int} K \neq \emptyset$.

4). Any $p \in K$ can be represented as a sum of at most m squares:

$$p(x) = \sum_{i=1}^{k} q_i^2(x), \quad q_i(x) \in \mathcal{F}(\mathcal{S}), i = 1, \dots, k \leq m.$$

The inclusion $p \in \operatorname{rint} K$ holds if and only if there exists such a representation with linearly independent q_i and $k = m$.

Proof:
1). If $p(x) = \langle p, v(x) \rangle$ with $p = \Lambda^*(Y)$ and $Y \succeq 0$, then for any $x \in \Delta$ we have:

$$p(x) = \langle \Lambda^*(Y), v(x) \rangle = \langle Y, \Lambda(v(x)) \rangle = \langle Y, u(x)u(x)^T \rangle = \langle Yu(x), u(x) \rangle \geq 0 \qquad (5)$$

since $Y \succeq 0$. On the other hand, if $p(x) \in K$ then there exists a system of coefficients $q_i \in R^m$, $i = 1 \dots N$, such that

$$\begin{aligned} p(x) &= \sum_{i=1}^{N} \langle q_i, u(x) \rangle^2 = \langle \sum_{i=1}^{N} q_i q_i^T, u(x)u(x)^T \rangle \\ &= \langle \sum_{i=1}^{N} q_i q_i^T, \Lambda(v(x)) \rangle = \langle \Lambda^* \left(\sum_{i=1}^{N} q_i q_i^T \right), v(x) \rangle. \end{aligned}$$

Thus, we can take $Y = \sum_{i=1}^{N} q_i q_i^T \succeq 0$ and $p = \Lambda^*(Y)$.

Note that K is closed and convex in view of (3). It is pointed since the system $\mathcal{S}$ is linearly independent.

2). The form of the dual cone K^* can be derived as follows:

$$\begin{aligned} c \in K^* \quad &\Leftrightarrow \quad \langle p, c \rangle \geq 0 \qquad \forall p \in K \\ &\Leftrightarrow \quad \langle \Lambda^*(Y), c \rangle \geq 0 \quad \forall Y \succeq 0 \\ &\Leftrightarrow \quad \langle Y, \Lambda(c) \rangle \geq 0 \quad \forall Y \succeq 0 \\ &\Leftrightarrow \quad \Lambda(c) \succeq 0. \end{aligned}$$

Since the primal cone is pointed, the interior of K^* is non-empty.

3). In order to prove $\operatorname{int} K \neq \emptyset$ we need to guarantee that the cone K^* is pointed. In view of Item 2, this is equivalent to non-degeneracy of the linear operator $\Lambda(c)$.

Assume that there exist $c \neq 0$ such that $\Lambda(c) = 0$. Without loss of generality we can think that $c^{(1)} = 1$. Then in view of (2) we have:

$$u(x)u(x)^T = \Lambda(v(x)) = \Lambda(v(x) - v^{(1)}(x)c).$$

Thus, the system $\mathcal{S}^2$ is covered by the functional system $\tilde{v}(x) = v(x) - v^{(1)}(x)c$. However, $\tilde{v}^{(1)}(x) \equiv 0$. This contradicts to our assumption on minimal degree of the system $v(x)$.

4). Note that any $Y \succeq 0 \in R^{m \times m}$ can be written as $Y = \sum_{i=1}^{k} q_i q_i^T$ with $q_i \in R^m$ and $k \leq m$. Therefore, in view of (5) we have:

$$p(x) = \sum_{i=1}^{k} \langle q_i, u(x) \rangle^2 = \sum_{i=1}^{k} q_i^2(x). \qquad \Box$$

In what follows we call $x_0 \in \Delta$ the *proper root* of a polynomial $p(x_0)$ if $p(x_0) = 0$ and $u(x_0) \neq 0$. Note that if the functional system $\mathcal{S}$ contains a constant function, then any root of the polynomial $p(x) \in K$ is proper.

Corollary 1 *1). If $p \in \operatorname{rint} K$ then $p(x) > 0$ for all $x \in \Delta$ such that $u(x) \neq 0$. If $p(x)$ has a proper root, then p belongs to the relative boundary of the cone K.*

2). Let $\gamma > 0 \in R^m$. Then $p(x) = \sum\limits_{i=1}^{m} \gamma^{(i)}[u^{(i)}(x)]^2 \in \operatorname{rint} K$.

Proof:
Indeed if x_0 is a proper root of $p(x) \in K$, then for any $Y \succeq 0$ such that $p = \Lambda^*(Y)$ we have:

$$0 = p(x_0) = \langle p, v(x_0) \rangle = \langle Y u(x_0), u(x_0) \rangle.$$

Since $u(x_0) \neq 0$ we conclude that Y cannot be positive semidefinite. Therefore in view of Item 4 of Theorem 1 p belongs to the relative boundary of the cone K. The first statement of this item is equivalent to the second one.

The second item of the coorolary follows from (4) with $Y = \operatorname{diag}\{\gamma\}$. □

We have seen that the structure of the domain Δ is not involved explicitly in the statement of Theorem 1. The cone K is completely described in terms of the linear operator $\Lambda(\cdot)$.

Note that for a particular function $p(x) = \sum\limits_{i=1}^{n} p^{(i)} \phi_i(x)$ we can check now the possiblity to represent it as a sum of squares. In order to do that, we need to fix an initial functional system $\mathcal{S}$ and an upper level functional system $v(x)$, which covers all $\phi_i(x)$ and $\mathcal{S}^2$. Of course, we have a lot of freedom in defining these objects. However, each concrete choice results in a convex feasibility problem $p \in K$.

Let us discuss some consequences of the representation (3). This relation implies that the cone K can be seen as a linear image of the cone of positive semidefinite matrices. Hence, it can be equipped with a self-concordant barrier ([6]):

$$F_K(p) = \min_Y \{-\ln \det Y : \; p = \Lambda^*(Y), \; Y \succeq 0 \in R^{m \times m}\}.$$

In view of Proposition 5.1.5 [6] the value of parameter of this barrier ν_K is equal to m. Recall that this value describes the complexity of a convex set for polynomial-time interior point schemes. Of course, the above representation of the barrier $F_K(p)$ is implicit. But we can avoid this trouble by considering the matrix Y as a part of our decision variables. Namely, we can consider the cone

$$\hat{K} = \{(p, Y) : \; p = \Lambda^*(Y), \; Y \succeq 0\}.$$

Note that this cone is pointed and the function $\hat{F}(p, Y) = -\ln \det Y$ is an m-self-concordant barrier for $\hat{K}$.

The barrier description of the dual cone is even simpler. Indeed, from Proposition 2.3.1 [6] we see that the function

$$F_{K^*}(c) = -\ln \det \Lambda(c)$$

is an m-self-concordant barrier for the cone K^*. This function is well defined since in view of Theorem 1 the interior of the cone K^* is always nonempty. In the case when the system $v(x)$ is minimal this function has a non-degenerate Hessian at any feasible point (see Theorem 2.1.1 (ii) [6]).

It is wellknown, that the gradient of the normal barriers establish one-to-one correspondence between interior points of the primal and dual cones ([6], Section 4). In our case this relation can be written in the following form.

Theorem 2 *Let the system $v(x)$ be minimal for K. Then, a polynomial p belongs to the interior of the cone K if and only if there exists $c \in \operatorname{int} K^*$ such that $p = \Lambda^*([\Lambda(c)]^{-1})$. Such c is uniquely defined and*

$$p(x) = \langle [\Lambda(c)]^{-1} u(x), u(x) \rangle. \tag{6}$$

Proof:
Since for $F_{K^*}(c) = -\ln \det \Lambda(c)$ we have

$$F'_{K^*}(c) = -\Lambda^*([\Lambda(c)]^{-1}),$$

the first statement of the theorem and uniqueness of c follows from the standard conic duality theory (see [6], Section 4). The representation (6) follows from (2) and (3):

$$\begin{aligned} p(x) &= \langle p, v(x) \rangle = \langle \Lambda^*([\Lambda(c)]^{-1}), v(x) \rangle = \langle [\Lambda(c)]^{-1}, \Lambda(v(x)) \rangle \\ &= \langle [\Lambda(c)]^{-1}, u(x)u(x)^T \rangle = \langle [\Lambda(c)]^{-1} u(x), u(x) \rangle. \end{aligned}$$

□

To conclude this section, let us prove a lower bound for the value of the parameter of a self-concordant barrier for the cone K.

Lemma 1 *Let $u(x)$ be non-degenerate. Assume that there exist k polynomials $q_i(x) \in \mathcal{F}(\mathcal{S})$ such that any $k-1$ of them has a common proper root but the polynomial $p(x) = \sum\limits_{i=1}^{k} q_i^2(x)$ belongs to the relative interior of the cone K. Then the value of the parameter of any self-concordant barrier for K cannot be less than k.*

Proof:
Note that all $q_i^2(x) \in K$ and all $p_i(x) = p(x) - q_i^2(x)$, $i = 1, \ldots, k$ belong to the relative boundary of the cone K (Corollary 1). On the other hand, $0 = p(x) - \sum\limits_{i=1}^{k} q_i^2(x) \in K$. Therefore the statement of the theorem follows from Lemma 3 in Appendix. □

We will see that Lemma 1 is convenient to get lower bounds for the parameters of the barriers for the cones related to univariate functional systems. In more complicated situations such bounds can be obtained directly from Lemma 3.

Example 1 Let $\Delta = R^n$ and the functional system $\mathcal{S}$ is defined as follows:

$$\mathcal{S} = \{u^{(ij)}(x) = x^{(i)} x^{(j)},\ i = 1, \ldots n,\ j = 1, \ldots i\}.$$

Consider the following cone:

$$K_{n,4} = \sum_{k=1}^{N} q_k^2(x), \quad q_k^2(x) \in \mathcal{F}(\mathcal{S}),\ k = 1, \ldots, N.$$

Thus, $K_{n,4}$ is a cone of homogeneous polynomials of n variables of degree four, which can be represented as a sum of squares of n-dimensional quadratic forms.

Let us compute a lower bound for the value of the parameter of an arbitrary self-concordant barrier for this cone. Let us fix some arbitray scalar $\beta > 0$. Consider the polynomial

$$\begin{aligned} p(x) &= \sum_{i=1}^{n} [u^{(ii)}(x)]^2 + \gamma \sum_{i=1}^{n} \sum_{j=1}^{i-1} [u^{ij}(x)]^2 = \sum_{i=1}^{n} [x^{(i)}]^4 + \gamma \sum_{i=1}^{n} \sum_{j=1}^{i-1} [x^{(i)}]^2 [x^{(j)}]^2 \\ &= (1 - \tfrac{\gamma}{2}) \sum_{i=1}^{n} [x^{(i)}]^4 + \tfrac{\gamma}{2} \left(\sum_{i=1}^{n} [x^{(i)}]^2 \right)^2. \end{aligned}$$

Note that $p(x) \in \operatorname{rint} K_{n,4}$ in view of Corollary 1. On the other hand, the polynomials $p_{ij}(x) = [x^{(i)}]^2 [x^{(j)}]^2$ are the recession directions of the cone $K_{n,4}$. Note that the polynomial

$$p(x) - p_{ii}(x) = \left(1 - \frac{\gamma}{2}\right) \sum_{i=1}^{n} [x^{(i)}]^4 + \frac{\gamma}{2} \left(\sum_{i=1}^{n} [x^{(i)}]^2 \right)^2 - [x^{(i)}]^4$$

vanishes at $x = e_i$, where $e_i \in R^n$ is the ith coordinate vector of this space. Therefore in view of Corollary 1 we can take in Lemma 3 $\beta_{ii} = 1$. At the same time, the polynomial

$$p(x) - p_{ij}(x) = \left(1 - \frac{\gamma}{2}\right) \sum_{i=1}^{n} [x^{(i)}]^4 + \frac{\gamma}{2} \left(\sum_{i=1}^{n} [x^{(i)}]^2 \right)^2 - (\gamma + 2) [x^{(i)}]^2 [x^{(j)}]^2$$

vanishes at $x = e_i + e_j$. Therefore we can take $\beta_{ij} = (\gamma + 2)$ for $i \neq j$. Finally, in view of definition of $p(x)$, we can take $\alpha_{ii} = 1$ and $\alpha_{ij} = \gamma$. Thus, in view of Lemma 3 we get the following bound on the parameter of a self-concordant barrier for $K_{n,4}$:

$$\nu_{n,4} \geq n + \tfrac{1}{2} n(n-1) \frac{\gamma}{\gamma + 2}.$$

Since γ is an arbitrary positive value, we conclude that $\nu_{n,4} \geq \frac{n(n+1)}{2}$. Note that in our example $m = \frac{n(n+1)}{2}$. Thus, the semidefinite representation (3) provides us with an optimal barrier for $K_{n,4}$. □

2.2 Sum of weighted squares

Note that the above results can be used in more complicated situations. Indeed, let $\mathcal{S}$ be defined as in the previous section. Consider the following convex cone:

$$K(\bar{q}) = \{ p(x) = \bar{q}(x) \sum_{i=1}^{N} q_i^2(x), \; q_i(x) \in \mathcal{F}(\mathcal{S}), \; i = 1 \ldots N \},$$

where $\bar{q}(x)$, $\bar{q} \neq 0$, is a *fixed non-negative* function:

$$\bar{q}(x) \geq 0 \quad \forall x \in \Delta.$$

In order to describe $K(\bar{q})$ we need only to change the initial functional system. Let us define

$$\bar{\mathcal{S}} = \{\bar{u}^{(1)} = \sqrt{\bar{q}(x)}u^{(1)}(x), \ldots, \bar{u}^{(m)} = \sqrt{\bar{q}(x)}u^{(m)}(x)\}.$$

Since the elements of the system $\mathcal{S}$ are linearly independent, the same is true for the elements of the system $\bar{\mathcal{S}}$. On the other hand, it is clear that

$$K(\bar{q}) = \{p(x) = \sum_{i=1}^{N} q_i^2(x),\ q_i(x) \in \mathcal{F}(\bar{\mathcal{S}}),\ i = 1 \ldots N\},$$

so we can get its description from Theorem 1.

In this section we give an explicit description of the cone, which is formed as a sum of weighted squares. Assume that we have several functional systems $\mathcal{S}_k$, $k = 1, \ldots, l$, which define the corresponding functional subspaces $\mathcal{F}(\mathcal{S}_k)$. Consider the following convex cone:

$$K(\bar{q}_1, \ldots, \bar{q}_l) = \{p(x) = \sum_{k=1}^{l} \bar{q}_k(x) \sum_{i=1}^{N} q_{ik}^2(x),\ q_{ik}(x) \in \mathcal{F}(\mathcal{S}_k),\ i = 1 \ldots N,\ k = 1, \ldots, l\},$$

where $\bar{q}_k(x)$ are some fixed non-zero functions, which are non-negative on Δ. It is clear that

$$K(\bar{q}_1, \ldots, \bar{q}_l) = \sum_{k=1}^{l} K(\bar{q}_k). \tag{7}$$

Therefore we can describe this cone using Theorem 1. In order to do that, we need to define the operators $\Lambda_k(v)$ which relate the systems $\bar{\mathcal{S}}_k^2$ with some upper-level functional system $v(x)$:

$$\bar{q}_k(x) u_k(x) u_k(x)^T = \Lambda_k(v(x)), \quad k = 1, \ldots, l.$$

The only assumption we need here is that $v(x)$ is large enough to represent any function from $\bar{\mathcal{S}}_k^2$, $k = 1, \ldots, l$.

Theorem 3 *1). The cone $K(\bar{q}_1, \ldots, \bar{q}_l)$ can be represented as follows:*

$$K(\bar{q}_1, \ldots, \bar{q}_l) = \{p \in R^n : \quad p = \sum_{k=1}^{l} \Lambda_k^*(Y_k), \quad Y_k \succeq 0 \in R^{m_k \times m_k},\ k = 1, \ldots, l\}, \tag{8}$$

where m_k is the number of elements of the system $\mathcal{S}_k$. This cone is closed, convex and pointed.

2). For the dual cone we have the following representation:

$$K^*(\bar{q}_1, \ldots, \bar{q}_l) = \{c \in R^n : \Lambda_k(c) \succeq 0 \in R^{m_k \times m_k},\ k = 1, \ldots, l\}.$$

We always have $\operatorname{int} K^*(\bar{q}_1, \ldots, \bar{q}_l) \neq \emptyset$.

3). If $v(x)$ is a minimal system for $\bigcup_{k=1}^{l} \bar{\mathcal{S}}_k^2$ then $K^*(\bar{q}_1, \ldots, \bar{q}_l)$ is a pointed cone and $\operatorname{int} K(\bar{q}_1, \ldots, \bar{q}_l) \neq \emptyset$.

Proof:

The proof of this theorem is very close to that of Theorem 1. The representation (8) follows from (7) and Theorem 1. In view of (8) this cone is convex and closed. It is pointed since any system $\mathcal{S}_k$ is formed by linearly independent functions. Hence, we get that the interior of the dual cone $K^*(\bar{q}_1,\ldots,\bar{q}_l)$ is non-empty. Its analytic representation can be obtained as follows:

$$
\begin{aligned}
c \in K^*(\bar{q}_1,\ldots,\bar{q}_l) \quad & \Leftrightarrow \quad \langle p, c\rangle \geq 0 && \forall p \in K(\bar{q}_1,\ldots,\bar{q}_l),\\
& \Leftrightarrow \quad \langle \sum_{k=1}^{l} \Lambda_k^*(Y_k), c\rangle \geq 0 && \forall Y_k \succeq 0 \in R^{m_k \times m_k},\ k=1,\ldots,l,\\
& \Leftrightarrow \quad \langle Y_k, \Lambda_k(c)\rangle \geq 0 && \forall Y_k \succeq 0 \in R^{m_k \times m_k},\ k=1,\ldots,l,\\
& \Leftrightarrow \quad \Lambda_k(c) \succeq 0 \in R^{m_k \times m_k}, \quad k=1,\ldots,l.
\end{aligned}
$$

Finally, the cone $K^*(\bar{q}_1,\ldots,\bar{q}_l)$ contains a line if and only if there exists a vector $c \in R^n$ such that $\Lambda_k(c) = 0$ for all $k = 1,\ldots,l$. However, in this case we can reduce the size of the functional system $v(x)$ (see the proof of Item 3 of Theorem 1). □

In the case of minimal $v(x)$ the cone $K^*(\bar{q}_1,\ldots,\bar{q}_l)$ can be equipped with a non-degenerate self-concordant barrier

$$F_{K^*(\bar{q}_1,\ldots,\bar{q}_l)}(c) = -\sum_{k=1}^{l} \ln\det \Lambda_k(c).$$

Using this barrier we can prove the following duality relation.

Theorem 4 *Let the system $v(x)$ be minimal for $K(\bar{q}_1,\ldots,\bar{q}_l)$. Then, a polynomial p is an interior point of the cone $K(\bar{q}_1,\ldots,\bar{q}_l)$ if and only if there exists $c \in \operatorname{int} K^*(\bar{q}_1,\ldots,\bar{q}_l)$ such that*

$$p = \sum_{k=1}^{l} \Lambda_k^*([\Lambda_k(c)]^{-1}).$$

Such c is uniquely defined and

$$p(x) = \sum_{k=1}^{l} \bar{q}_k(x)\langle [\Lambda_k(c)]^{-1} u_k(x), u_k(x)\rangle. \tag{9}$$

The proof of this statement is similar to that of Theorem 2.

2.3 Functional systems with complex values

In this section we use the following notation. For a complex number $x = a + jb \in C$, where $j = \sqrt{-1}$, we denote $\bar{x} = a - jb$. Then $| x |^2 = x\bar{x}$. For two vectors x, $y \in C^n$ we denote

$$\langle x, y\rangle_C = \sum_{i=1}^{n} x^{(i)}\bar{y}^{(i)}.$$

Thus, $\| x \|_C = \langle x, x\rangle_C^{1/2}$. For a complex matrix $A \in C^{n\times n}$ we write $A \succeq 0$ if A is Hermitian and positive semidefinite:

$$\langle Ax, x\rangle_C \in R, \quad \langle Ax, x\rangle_C \geq 0, \quad \forall x \in C^n.$$

For a matrix $A \in C^{m\times n}$ we denote $A^{\#}$ its complex adjoint:

$$\langle Ax, y\rangle_C = \langle x, A^{\#}y\rangle_C \quad \forall x, y \in C^n.$$

Clearly, $A^{\#} = \bar{A}^T$. We denote by $H^{m\times m}$ the linear space of Hermitian $(m \times m)$-matrices. For $A \in H^{m\times m}$ we have $A = A^{\#}$. Finally, the functions $\mathrm{Re}\,(\cdot)$ and $\mathrm{Im}\,(\cdot)$ extract the real and imaginary parts of the argument:

$$x \equiv \mathrm{Re}\,(x) + j\mathrm{Im}\,(x).$$

For x and $y \in C^n$ we denote $\langle x, y\rangle_H$ their Hermitian inner product:

$$\langle x, y\rangle_H = \langle \mathrm{Re}\,(x), \mathrm{Re}\,(y)\rangle + \langle \mathrm{Im}\,(x), \mathrm{Im}\,(y)\rangle = \tfrac{1}{2}[\langle x, y\rangle_C + \langle \bar{x}, \bar{y}\rangle_C] = \mathrm{Re}\,(\langle x, y\rangle_C).$$

Note that for two Hermitian matrices X and Y we have:

$$\langle X, Y\rangle_C = \tfrac{1}{2}[\langle X, Y\rangle_C + \langle X^T, Y^T\rangle_C] = \tfrac{1}{2}[\langle X, Y\rangle_C + \langle \bar{X}, \bar{Y}\rangle_C] = \langle X, Y\rangle_H.$$

Let $\mathcal{W} = \{w^{(1)}(x), \dots, w^{(m)}(x)\}$, $x \in \Delta$, be an arbitrary system of linearly independent complex-valued functions. Define a finite-dimensional functional subspace

$$\mathcal{F}_C(\mathcal{W}) = \{q(x) = \sum_{k=1}^{m} q^{(k)} w^{(k)}(x),\ q = (q^{(1)}, \dots, q^{(m)}) \in C^m\}.$$

Consider the following finite-dimensional cone:

$$K = \{p(x) = \sum_{i=1}^{N} \mid q_i(x) \mid^2,\ q_i(x) \in \mathcal{F}_C(\mathcal{W}),\ i = 1 \dots N\}.$$

Again, in this definition N is fixed and large enough, say $N \geq m$.

Consider the *squared* functional system

$$\mathcal{W}^2 = \{v_{ik}(x) = w^{(i)}(x) \cdot \bar{w}^{(k)}(x),\ i, k = 1, \dots, m\}.$$

And let the components of the vector function $v(x) \in C^n$ span a finite dimensional functional space which covers the system $\mathcal{W}^2$. Define a matrix valued linear operator $\Lambda(v) :\ C^n \to C^{m\times m}$ as follows:

$$w(x)w(x)^{\#} = \Lambda(v(x)), \quad x \in \Delta, \tag{10}$$

where $w(x) = (w^{(1)}(x), \dots, w^{(m)}(x))$. Since $w(x)w(x)^{\#}$ is a Hermitian matrix and the operator $\Lambda(v) = \sum_{i=1}^{n} \Lambda_i v^{(i)}$ is linear, we have also the following representation:

$$w(x)w(x)^{\#} = (\Lambda(v(x)))^{\#} = \left(\sum_{i=1}^{n} \Lambda_i v^{(i)}(x)\right)^{\#} = \sum_{i=1}^{n} \Lambda_i^{\#} \bar{v}^{(i)}(x). \tag{11}$$

Adding (10) and (11) we get a symmetric expression:

$$w(x)w(x)^{\#} = \Lambda_H(v(x)), \quad x \in \Delta, \tag{12}$$

where $\Lambda_H(v) = \frac{1}{2}[\sum_{i=1}^{n} \Lambda_i v^{(i)} + \Lambda_i^{\#} \bar{v}^{(i)}]$. If for some i the matrix Λ_i is Hermitian, the complex part of the function $v_i(x)$ is not involved in the above representation. Therefore without loss of generality we can assume that for such i the function $v_i(x)$ has only real values. Thus, we can fix the following structure of the operator $\Lambda_H(v)$:

$$\begin{aligned} \Lambda_H(v(x)) &= w(x)w(x)^{\#}, \quad x \in \Delta, \\ \Lambda_H(v): & \quad E \to H^{m \times m}, \end{aligned} \tag{13}$$

where E is a direct product of some real and some complex multidimensional spaces.

Now we can define the adjoint operator $\Lambda_H^*(\cdot) : H^{m\times m} \to E$ as follows:

$$\langle Y, \Lambda_H(v)\rangle_H = \langle \Lambda_H^*(Y), v\rangle_H, \quad \forall Y \in H^{m\times m},\ v \in E.$$

Let us find an explicit form of $\Lambda_H^*(\cdot)$. Note that

$$\begin{aligned} \langle Y, \Lambda_H(v)\rangle_H &= \frac{1}{2}\sum_{i=1}^{n} \langle Y, \Lambda_i v^{(i)} + \Lambda_i^{\#}\bar{v}^{(i)}\rangle_H \\ &= \frac{1}{2}\sum_{i=1}^{n} \langle Y, \Lambda_i v^{(i)} + \Lambda_i^{\#}\bar{v}^{(i)}\rangle_C \\ &= \frac{1}{2}\sum_{i=1}^{n} [\bar{v}^{(i)}\langle Y, \Lambda_i\rangle_C + v^{(i)}\langle Y, \Lambda_i^{\#}\rangle_C] \\ &= \frac{1}{2}\sum_{i=1}^{n} [\bar{v}^{(i)}\langle Y, \Lambda_i\rangle_C + v^{(i)}\langle \bar{Y}, \bar{\Lambda}_i\rangle_C] \\ &= \frac{1}{2}\sum_{i=1}^{n} [\bar{v}^{(i)}\langle Y, \Lambda_i\rangle_C + v^{(i)}\overline{\langle Y, \Lambda_i\rangle_C}] \\ &= \frac{1}{2}[\langle p, v\rangle_C + \langle \bar{p}, \bar{v}\rangle_C] = \langle p, v\rangle_H \end{aligned}$$

with the vector p defined as follows:

$$p^{(i)} = \langle Y, \Lambda_i\rangle_C, \quad i = 1, \dots, n. \tag{14}$$

For such a vector we write $p = \Lambda^*(Y)$. Note that Hermitian Λ_i generates real $p^{(i)}$.

Now we can prove the following theorem.

Theorem 5 *1). The function $p(x) = \langle p, v(x)\rangle_H$, $p \in E$, belongs to K if and only if there exists a Hermitian positive semidefinite $(m \times m)$-matrix Y such that $p = \Lambda_H^*(Y)$:*

$$K = \{p \in E : \ p = \Lambda_H^*(Y),\ Y \succeq 0 \in H^{m\times m}\}. \tag{15}$$

This cone is closed convex and pointed.

$$\operatorname{rint} K = \{p \in E : \ p = \Lambda^*(Y),\ Y \succ 0 \in H^{m\times m}\}. \tag{16}$$

2). Its dual cone K^* defined with respect to the inner produch $\langle \cdot, \cdot \rangle_H$ can be represented as follows:

$$K^* = \{c \in E : \; \Lambda_H(c) \succeq 0 \in H^{m \times m}\}, \quad \mathrm{int}\, K^* \neq \emptyset. \tag{17}$$

3). If the size of the system $v(x)$ in the representation (13) cannot be reduced, then K^* is a pointed cone and $\mathrm{int}\, K \neq \emptyset$.

4). Any $p \in K$ can be represented as a sum of at most m squares:

$$p(x) = \sum_{i=1}^{k} \mid q_i(x) \mid^2, \quad q_i(x) \in \mathcal{F}_C(\mathcal{W}), i = 1, \ldots, k \leq m.$$

The inclusion $p \in \mathrm{rint}\, K$ holds if and only if there exists such a representation with linearly independent q_i and $k = m$.

Proof:

1). The cone K is pointed since the system $\mathcal{W}$ consists of linearly independant functions. It is convex and closed in view of representation (15), which we are going to prove now.

Let $p(x) \in K$. Then there exists an $(m \times N)$-matix Q such that

$$p(x) = \langle Q^{\#} w(x), Q^{\#} w(x) \rangle_C.$$

Therefore in view of (14) we have:

$$\begin{aligned} p(x) &= \langle QQ^{\#} w(x), w(x) \rangle_C = \langle QQ^{\#}, w(x) w^{\#}(x) \rangle_C = \langle QQ^{\#}, w(x) w^{\#}(x) \rangle_H \\ &= \langle QQ^{\#}, \Lambda_H(v(x)) \rangle_H = \langle \Lambda_H^*(QQ^{\#}), v(x) \rangle_H. \end{aligned}$$

Thus, any $p(x) \in K$ can be representes as $p(x) = \langle p, v(x) \rangle_H$, where $p = \Lambda_H^*(Y)$ and $Y = QQ^{\#} \succeq 0 \in H^{m \times m}$. On the other hand, if $p(x) = \langle p, v(x) \rangle_H$ with $p = \Lambda_H^*(Y)$ and $Y \succeq 0 \in H^{m \times m}$, we have the following:

$$\begin{aligned} p(x) &= \langle \Lambda_H^*(Y), v(x) \rangle_H = \langle Y, \Lambda_H(v(x)) \rangle_H \\ &= \langle Y, \Lambda_H(v(x)) \rangle_C = \langle Y, w(x) w(x)^{\#} \rangle_C = \langle Y w(x), w(x) \rangle_C. \end{aligned}$$

Since Y is Hermitian and positive semidefinite, we can represent it in the form $Y = \sum_{i=1}^{k} q_i q_i^{\#}$ with $q_i \in C^n$, $i = 1, \ldots, k \leq m$. Therefore

$$p(x) = \sum_{i=1}^{k} \langle q_i q_i^{\#} w(x), w(x) \rangle_C = \sum_{i=1}^{k} \mid q_i^{\#} w(x) \mid^2 .$$

Thus, in this case $p(x) \in K$.

2). Let us derive the representation (17).

$$\begin{aligned} c \in K^* \quad &\Leftrightarrow \quad \langle p, c \rangle_H \geq 0 \qquad \forall p \in K \\ &\Leftrightarrow \quad \langle \Lambda_H^*(Y), c \rangle_H \geq 0 \qquad \forall Y \succeq 0 \in H^{m \times m} \\ &\Leftrightarrow \quad \langle Y, \Lambda_H(c) \rangle_H \geq 0 \qquad \forall Y \succeq 0 \in H^{m \times m} \\ &\Leftrightarrow \quad \Lambda(c) \succeq 0 \in H^{m \times m}. \end{aligned}$$

The proofs of the rest statements of the theorem are similar to those of Theorem 1. □

Same as in the previous sections, we can establish some duality relations between the interior of the cones K and K^*. form.

Theorem 6 *Let the system $v(x)$ be minimal for K. Then, a polynomial p belongs to the interior of the cone K if and only if there exists $c \in \operatorname{int} K^*$ such that $p = \Lambda_H^*([\Lambda_H(c)]^{-1})$. Such c is uniquely defined and*

$$p(x) = \langle [\Lambda(c)]^{-1} u(x), u(x) \rangle_C. \tag{18}$$

Proof:
Using the technique of Section 5.4.5, [6], it can be proved that the function $F(X) = -\ln \det X$, $X \in H^{m \times m}$, is an m-self-concordant barrier for the cone of Hermitian positive semidefinite matrices. Define $F_{K^*}(c) = -\ln \det \Lambda_H(c)$, $c \in E$. Let us compute its gradient with respect to Hermitian inner product. The first differential of this function can be written in the following form:

$$-DF_{K^*}(c)[h] = \langle [\Lambda_H(c)]^{-1}, \Lambda_H(h) \rangle_H = \langle \Lambda_H^*([\Lambda_H(c)]^{-1}), h \rangle_H, \quad h \in E.$$

Thus, $-F'_{K^*}(c) = \Lambda_H^*([\Lambda_H(c)]^{-1})$ and the first statement of the theorem follows from the standard conic duality.

Let us prove the representation (18). In view of (12) we have

$$\begin{aligned} p(x) &= \langle p, v(x) \rangle_H = \langle \Lambda_H^*([\Lambda_H(c)]^{-1}), v(x) \rangle_H = \langle [\Lambda_H(c)]^{-1}, \Lambda_H(v(x)) \rangle_H \\ &= \langle [\Lambda_H(c)]^{-1}, u(x) u(x)^{\#} \rangle_H = \langle [\Lambda(c)]^{-1} u(x), u(x) \rangle_C. \end{aligned}$$

□

3 Cones of non-negative polynomials

In this section we describe the cones of univariate non-negative polynomials. We can get this description using the results of Section 2 and the classical results related to representability of such polynomials as sums of squares.

3.1 Polynomials on an infinite interval

Consider the vector function $v(t) = (1, t, t^2, \dots, t^{2n}) \in R^{2n+1}$, $t \in R$. We are interested in the description of the following convex cone:

$$K_\infty = \{ p \in R^{2n+1} : \ \langle p, v(t) \rangle \geq 0 \ \forall t \in R \}.$$

Denote $u(t) = (1, t, \dots, t^n) \in R^{n+1}$. In accordance with Markov-Lukacs theorem [5, 3], $p(t) \equiv \langle p, v(t) \rangle$ belongs to K_∞ if and only if it can be represented as follows:

$$p(t) = \langle q_1, u(t) \rangle^2 + \langle q_2, u(t) \rangle^2$$

with some $q_1,q_2 \in R^{n+1}$. Thus, we can get a description of the cone K_∞ from Theorem 1. To this end we need to define a linear operator Λ such that

$$u(t)u(t)^T = \Lambda(v(t)).$$

Clearly, we can take $\Lambda(v) = \sum_{i=1}^{2n+1} v^{(i)} H_{n,i}$ with the matrices $H_{n,i} \in R^{(n+1)\times(n+1)}$ defined as follows:

$$H_{n,i}^{(kl)} = \begin{cases} 1, & \text{if } k+l = i+1, \\ 0, & \text{otherwise.} \end{cases} \tag{19}$$

Thus, $\Lambda(v)$ is a Hankel matrix. Note that the operator $\Lambda^*(Y) : R^{(n+1)\times(n+1)} \to R^{2n+1}$ is defined as follows:

$$\Lambda^*(Y)^{(i)} = \langle Y, H_{n,i} \rangle, \quad i = 1, \ldots, 2n+1.$$

Using Theorem 1 we get the following result.

Theorem 7 *The cones K_∞ and K_∞^* can be represented as follows:*

$$K_\infty = \{p \in R^{2n+1} : \ p = \Lambda^*(Y), \ Y \succeq 0 \in R^{(n+1)\times(n+1)}\},$$

$$K_\infty^* = \{c \in R^{2n+1} : \Lambda(c) \succeq 0 \in R^{(n+1)\times(n+1)}\}.$$

Both cones K_∞ and K_∞^* are convex, closed and pointed cones with non-empty interior.

In view of Theorem 2 we have the following dual representation for $p \in \operatorname{int} K_\infty$:

$$p(t) = \langle [\Lambda(c)]^{-1} u(t), u(t) \rangle,$$

where the point $c \in \operatorname{int} K_\infty^*$ is uniquely defined by the equation

$$p = \Lambda^*([\Lambda(c)]^{-1}).$$

Finaly, let us choose $(n+1)$ values $t_1 < t_2 < \ldots < t_{n+1}$ and consider the polynomials

$$q_i(t) = \prod_{k \neq i} (t - t_k), \quad i = 1, \ldots, n+1.$$

Then, using Lemma 1 we get that any self-concordant barrier for the cone K_∞ cannot have the value of the parameter less than $(n+1)$. Thus, the barrier

$$F_*(c) = -\ln \det \Lambda(c)$$

is optimal for the dual cone.

3.2 Polynomials on a semi-infinite interval

Consider the vector function $v(t) = (1, t, t^2, \ldots, t^n) \in R^{n+1}$, $t \geq 0$. Define the following convex cone:

$$K_{0,\infty} = \{p \in R^{n+1} : \ \langle p, v(t) \rangle \geq 0 \ \forall t \geq 0\}.$$

Define $n_1 = [\frac{n}{2}]$ and $n_2 = [\frac{n-1}{2}]$. Note that $2n_1 \le n$, $2n_2 \le n-1$ and $n_1 + n_2 = n-1$. Denote

$$u_1(t) = (1, t, \dots, t^{n_1}) \in R^{n_1+1}, \quad u_2(t) = (1, t, \dots, t^{n_1}) \in R^{n_2+1}.$$

In accordance with Markov-Lukacs theorem [5, 3], $p(t) \equiv \langle p, v(t) \rangle$ belongs to $K_{0,\infty}$ if and only if it has the following representation:

$$p(t) = \langle q_1, u_1(t) \rangle^2 + t \langle q_2, u_2(t) \rangle^2$$

with some $q_1 \in R^{n_1+1}$ and $q_2 \in R^{n_2+1}$. Thus, we can get a description of the cone $K_{0,\infty}$ from Theorem 3. In order to apply this theorem we need to define two linear operators Λ_1 and Λ_2 such that

$$u_1(t)u_1(t)^T = \Lambda_1(v(t)), \quad tu_2(t)u_2(t)^T = \Lambda_2(v(t)).$$

Clearly, we can take

$$\Lambda_1(v) = \sum_{i=1}^{2n_1+1} v^{(i)} H_{n_1,i}, \quad \Lambda_2(v) = \sum_{i=1}^{2n_2+1} v^{(i+1)} H_{n_2,i},$$

with the matrices $H_{(\cdot)}$ defined by (19). Then the adjoint operators are defined as follows:

$$\Lambda_1^*(Y) : R^{(n_1+1)\times(n_1+1)} \to R^{n+1}, \quad \Lambda_1^*(Y)^{(i)} = \begin{cases} \langle Y, H_{n_1,i} \rangle & 1 \le i \le 2n_1 + 1, \\ 0 & \text{otherwise}. \end{cases}$$

$$\Lambda_2^*(Y) : R^{(n_2+1)\times(n_2+1)} \to R^{n+1}, \quad \Lambda_1^*(Y)^{(i)} = \begin{cases} \langle Y, H_{n_2,i-1} \rangle & 2 \le i \le 2n_2 + 2, \\ 0 & \text{otherwise}. \end{cases}$$

Using Theorem 3 we get the following result.

Theorem 8 *The cones $K_{0,\infty}$ and $K_{0,\infty}^*$ can be represented as follows:*

$$\begin{aligned} K_{0,\infty} &= \{p \in R^{n+1} : \quad p = \Lambda_1^*(Y_1) + \Lambda_2^*(Y_2), \\ &\qquad Y_1 \succeq 0 \in R^{(n_1+1)\times(n_1+1)},\ Y_2 \succeq 0 \in R^{(n_2+1)\times(n_2+1)}\}, \\ K_{0,\infty}^* &= \{c \in R^{n+1} : \quad \Lambda_1(c) \succeq 0 \in R^{(n_1+1)\times(n_1+1)},\ \Lambda_2(c) \succeq 0 \in R^{(n_2+1)\times(n_2+1)}\}. \end{aligned}$$

Both cones $K_{0,\infty}$ and $K_{0,\infty}^*$ are convex, closed and pointed cones with non-empty interior.

Proof:
The only thing we need to check is that $\Lambda_1(c) = 0$ and $\Lambda_2(c) = 0$ implies $c = 0 \in R^{n+1}$. Indeed, the first conditions implies that $c^{(i)} = 0$ for $i = 1, \dots, 2n_1 + 1$. From the second condition we have $c^{(i)} = 0$ for $i = 2, \dots, 2n_2 + 2$. But $(2n_1 + 1) + (2n_2 + 2) = 2n + 1$. Therefore one of these integers is greater than n. □

In view of Theorem 4 we have the following dual representation for $p \in \text{int}\, K_{0,\infty}$:

$$p(t) = \langle [\Lambda_1(c)]^{-1} u_1(t), u_1(t) \rangle + t \langle [\Lambda_2(c)]^{-1} u_2(t), u_2(t) \rangle,$$

where the point $c \in \text{int}\, K_{0,\infty}^*$ is uniquely defined by the equation

$$p = \Lambda_1^*([\Lambda_1(c)]^{-1}) + \Lambda_2^*([\Lambda_2(c)]^{-1}).$$

3.3 Polynomials on a finite interval

Let us fix some interval $[a,b] \subset R$. Consider the vector function $v(t) = (1, t, t^2, \ldots, t^n) \in R^{n+1}$, $t \in [a,b]$. Define the following convex cone:

$$K_{a,b} = \{p \in R^{n+1} : \langle p, v(t) \rangle \geq 0 \ \forall t \in [a,b]\}.$$

In order to describe this cone we need to consider two cases.

3.3.1 Polynomials of even degree

Let $n = 2m$. Denote

$$u_1(t) = (1, t, \ldots, t^m) \in R^{m+1}, \quad u_2(t) = (1, t, \ldots, t^{m-1}) \in R^m.$$

In accordance with Markov-Lukacs theorem [5, 3], $p(t) \equiv \langle p, v(t) \rangle$ belongs to $K_{a,b}$ if and only if it has the following representation:

$$p(t) = \langle q_1, u_1(t) \rangle^2 + (t-a)(b-t)\langle q_2, u_2(t) \rangle^2$$

with some $q_1 \in R^{m+1}$ and $q_2 \in R^m$. Therefore we can get a description of the cone $K_{a,b}$ from Theorem 3. Let us introduce two linear operators Λ_1 and Λ_2 as follows

$$u_1(t)u_1(t)^T = \Lambda_1(v(t)), \quad (t-a)(b-t)u_2(t)u_2(t)^T = \Lambda_2(v(t)).$$

Clearly, we can take $\Lambda_1(v) = \sum\limits_{i=1}^{2m+1} v^{(i)} H_{m,i}$. In order to define the second operator note that

$$(t-a)(b-t)u_2(t)u_2(t)^T = (b+a)tu_2(t)u_2(t)^T - t^2 u_2(t)u_2(t)^T - abu_2(t)u_2(t)^T.$$

Therefore we can take

$$\begin{aligned} \Lambda_2(v) &= (b+a)\sum_{i=2}^{2m} v^{(i)} H_{m-1,i-1} - \sum_{i=3}^{2m+1} v^{(i)} H_{m-1,i-2} - ab \sum_{i=1}^{2m-1} v^{(i)} H_{m-1,i} \\ &= \sum_{i=1}^{2m-1} [(b+a)v^{(i+1)} - v^{(i+2)} - abv^{(i)}] H_{m-1,i}. \end{aligned}$$

Using Theorem 3 we get the following result.

Theorem 9 *The cones $K_{a,b}$ and $K^*_{a,b}$ can be represented as follows:*

$$\begin{aligned} K_{a,b} &= \{p \in R^{n+1} : \quad p = \Lambda_1^*(Y_1) + \Lambda_2^*(Y_2), \\ &\qquad\qquad Y_1 \succeq 0 \in R^{(m+1)\times(m+1)}, \ Y_2 \succeq 0 \in R^{m \times m}\}, \\ K^*_{a,b} &= \{c \in R^{n+1} : \quad \Lambda_1(c) \succeq 0 \in R^{(m+1)\times(m+1)}, \ \Lambda_2(c) \succeq 0 \in R^{m\times m}\}. \end{aligned}$$

Both cones $K_{a,b}$ and $K^*_{a,b}$ are convex, closed and pointed cones with non-empty interior.

Proof:
Again, we need to check that $\Lambda_1(c) = 0$ and $\Lambda_2(c) = 0$ implies $c = 0 \in R^{n+1}$. However, even the first equation alone provides us with the result. $\Box$

In view of Theorem 4 we have the following dual representation for $p \in \operatorname{int} K_{a,b}$:

$$p(t) = \langle [\Lambda_1(c)]^{-1} u_1(t), u_1(t) \rangle + (t-a)(b-t)\langle [\Lambda_2(c)]^{-1} u_2(t), u_2(t) \rangle,$$

where the point $c \in \operatorname{int} K^*_{a,b}$ is uniquely defined by the equation

$$p = \Lambda_1^*([\Lambda_1(c)]^{-1}) + \Lambda_2^*([\Lambda_2(c)]^{-1}).$$

3.3.2 Polynomials of odd degree

Let $n = 2m+1$. Denote $u(t) = (1, t, \ldots, t^m) \in R^{m+1}$. In accordance with Markov-Lukacs theorem [5, 3], $p(t) \equiv \langle p, v(t) \rangle$ belongs to $K_{a,b}$ if and only if it has the following representation:

$$p(t) = (t-a)\langle q_1, u(t) \rangle^2 + (b-t)\langle q_2, u(t) \rangle^2$$

with some $q_1, q_2 \in R^{m+1}$. Let us introduce two linear operators Λ_1 and Λ_2 as follows

$$(t-a)u(t)u(t)^T = \Lambda_1(v(t)), \quad (b-t)u(t)u(t)^T = \Lambda_2(v(t)).$$

Clearly, we can take

$$\begin{aligned} \Lambda_1(v) &= \sum_{i=1}^{2m+1} (v^{(i+1)} - a v^{(i)}) H_{m,i}, \\ \Lambda_2(v) &= \sum_{i=1}^{2m+1} (b v^{(i)} - v^{(i+1)}) H_{m,i}. \end{aligned}$$

It remains to apply Theorem 3.

Theorem 10 *The cones $K_{a,b}$ and $K^*_{a,b}$ can be represented as follows:*

$$\begin{aligned} K_{a,b} &= \{p \in R^{n+1} : \quad p = \Lambda_1^*(Y_1) + \Lambda_2^*(Y_2),\ Y_1, Y_2 \succeq 0 \in R^{(m+1)\times(m+1)}\}, \\ K^*_{a,b} &= \{c \in R^{n+1} : \quad \Lambda_1(c) \succeq 0 \in R^{(m+1)\times(m+1)},\ \Lambda_2(c) \succeq 0 \in R^{(m+1)\times(m+1)}\}. \end{aligned}$$

Both cones $K_{a,b}$ and $K^*_{a,b}$ are convex, closed and pointed cones with non-empty interior.

Proof:
Let $\Lambda_1(c) = 0$ and $\Lambda_2(c) = 0$. Then $c^{(i+1)} = ac^{(i)}$ and $c^{(i+1)} = bc^{(i)}$ for $i = 1, \ldots, n-1$. Since $a \neq b$ we conclude that $c = 0$. $\Box$

In view of Theorem 4 we have the following dual representation for $p \in \operatorname{int} K_{a,b}$:

$$p(t) = (t-a)\langle [\Lambda_1(c)]^{-1} u(t), u(t) \rangle + (b-t)\langle [\Lambda_2(c)]^{-1} u(t), u(t) \rangle,$$

where the point $c \in \operatorname{int} K^*_{a,b}$ is uniquely defined by the equation

$$p = \Lambda_1^*([\Lambda_1(c)]^{-1}) + \Lambda_2^*([\Lambda_2(c)]^{-1}).$$

3.4 Trigonometric polynomials

In this and the next sections we use the notation of Section 2.3. Consider the functional system

$$w(z) = (1, z, \ldots, z^n)^T, \quad z \in \Delta = \{z = e^{jt},\ t \in [0, 2\pi]\}.$$

Define the following convex cone:

$$K_T = \{p \in C^{n+1} :\ \mathrm{Re}\,(p(z)) \equiv \langle p, w(z)\rangle_H \geq 0\ \forall z \in \Delta\}.$$

In accordance with Riesz-Fejer theorem $p \in K_T$ if and only if there exists some $q \in C^{n+1}$ such that

$$p(z) = \langle p, w(z)\rangle_H = \mid \sum_{i=1}^{n+1} q^{(i)} w^{(i)}(z) \mid^2, \quad z \in \Delta.$$

In order apply Theorem 5 we need to define the corresponding operator $\Lambda_H(v)$. Note that for $z \in \Delta$ we have

$$(w(z)w(z)^{\#})^{(kl)} = z^{k-l}, \quad k, l = 1, \ldots, n+1.$$

Therefore we can take $v(z) = (1, z, \ldots, z^n)^T$ and define

$$\Lambda_H(v) = \tfrac{1}{2} \sum_{i=1}^{n+1} [T_i v^{(i)} + T_i^T \bar{v}^{(i)}],$$

where $T_i \in R^{(n+1)\times(n+1)}$ is a basis of Toeplitz matrices:

$$T_1 = I_{n+1}, \quad T_i^{(kl)} = \begin{cases} 2, & \text{if } k - l = i - 1, \\ 0, & \text{otherwise.} \end{cases} \qquad i = 2, \ldots, n+1.$$

Thus the matrix $\Lambda_H(v)$ is Hermitian and Toeplitz. Since only T_1 is Hermitian, we take $E = R \times C^n$. Clearly,

$$w(z)w^{\#}(z) = \Lambda_H(v(z)).$$

Now we can apply Theorem 5.

Theorem 11

$$K_T = \{p \in E : p = \Lambda_H^*(Y),\ Y \succeq 0 \in H^{(n+1)\times(n+1)}\},$$

$$K_T^* = \{c \in E :\ \Lambda_H(c) \succeq 0 \in H^{(n+1)\times(n+1)}\}.$$

Both cones K_T and K_T^* are closed convex and pointed cones with non-empty interior.

Recall that $p = \Lambda_H^*(Y)$ is defined as follows: $p^{(i)} = \langle Y, T_i\rangle_C$, $i = 1, \ldots, n+1$. In view of Theorem 6 we have the following dual representation for $p \in \mathrm{int}\, K_T$:

$$\mathrm{Re}\,(p(z)) = \langle [\Lambda_H(c)]^{-1} u(z), u(z)\rangle_C,$$

where the point $c \in \mathrm{int}\, K_T^*$ is uniquely defined by the equation

$$p = \Lambda_H^*([\Lambda_H(c)]^{-1}).$$

3.5 Nevanlinna-Pick problems

Let us fix n complex values inside the unit circle:

$$z_k \in C : \mid z_k \mid < 1, \quad k = 1, \ldots, n.$$

Consider the system $w(z) = (w^{(1)}(z), \ldots w^{(n)}(z))$ defined as follows:

$$w^{(k)}(z) \quad = \tfrac{1}{z - z_k}, \ z \in \Delta = \{z = e^{jt}, \ t \in [0, 2\pi]\}, \quad k = 1, \ldots, n.$$

Then we can define the following family of rational functions:

$$\mathcal{F} = \{\rho(z) = \sum_{k=1}^{n} \rho^{(k)} w^{(k)}(z), \ \rho = (\rho^{(1)}, \ldots, \rho^{(n)}) \in C^n\}.$$

Consider the following convex cone:

$$K_{NP} = \{\rho \in C^n : \ \mathrm{Re}\,(\rho(z)) \geq 0 \ \forall z \in \Delta\}.$$

Note that any $\rho(z) \in \mathcal{F}$ can be represented as follows:

$$\rho(z) = \frac{p(z)}{d(z)},$$

where $p(z)$ is a polynomial in $z \in C$ of degree $n-1$ and $d(z) = \prod_{k=1}^{n} (z - z_k)$. On the other hand,

$$\mathrm{Re}\,(\rho(z)) = \tfrac{1}{2}(\rho(z) + \overline{\rho(z)}) = \frac{p(z)\overline{d(z)} + \overline{p(z)}d(z)}{2 \mid d(z) \mid^2} \equiv \frac{\mathrm{Re}\,(p_d(z))}{\mid d(z) \mid^2},$$

where $p_d(z)$ is a polynomial in $z \in C$ of degree $n-1$. In accordance with Riesz-Fejer theorem $\rho \in K_{NP}$ if and only if there exists some polynomial $q(z)$ of degree $n-1$ such that

$$\mathrm{Re}\,(p_d(z)) = \mid q(z) \mid^2, \quad z \in \Delta.$$

Therefore

$$\mathrm{Re}\,(\rho(z)) = \frac{\mid q(z) \mid^2}{\mid d(z) \mid^2} = \left| \frac{q(z)}{d(z)} \right|^2 = \mid \rho_1(z) \mid^2$$

for some $\rho_1(z) \in \mathcal{F}$. Thus, we can describe the cone K_{NP} using Theorem 5.

In order to do that, we need to define the corresponding operator $\Lambda_H(v)$. Note that for $z \in \Delta$ and any k, l, $1 \leq k, l \leq n$, we have

$$\frac{z + z_k}{z - z_k} + \frac{\bar{z} + \bar{z}_l}{\bar{z} - \bar{z}_l} = 2 \frac{1 - z_k \bar{z}_l}{(z - z_k)(\bar{z} - \bar{z}_l)}.$$

Therefore

$$w^{(k)}(z)\overline{w^{(l)}(z)} = \frac{1}{2(1 - z_k \bar{z}_l)} \left(\frac{z + z_k}{z - z_k} + \frac{\bar{z} + \bar{z}_l}{\bar{z} - \bar{z}_l} \right).$$

Thus, we can choose $v(z) = (v^{(1)}(z), \ldots, v^{(n)}(z))$ as follows:

$$v^{(k)}(z) = \frac{z + z_k}{z - z_k}, \ z \in \Delta, \ k = 1, \ldots, n.$$

This choice leads to the following representation of the operator $\Lambda_H(v)$, $v \in E$:

$$\Lambda_H(v) = \frac{1}{2}\sum_{k=1}^{n}(P_k v^{(k)} + P_k^{\#}\bar{v}^{(k)}),$$

$$(P_k)^{(il)} = \begin{cases} \frac{1}{1-z_i\bar{z}_l}, & i = k, \\ 0, & \text{otherwise.} \end{cases} \qquad k, i, l = 1, \ldots, n.$$

Now we can apply Theorem 5.

Theorem 12

$$K_{NP} = \{\rho \in C^n : \rho = \Lambda_H^*(Y),\ Y \succeq 0 \in H^{n\times n}\},$$

$$K_{NP}^* = \{c \in C^n :\ \Lambda_H(c) \succeq 0 \in H^{n\times n}\}.$$

Both cones K_{NP} and K_{NP}^* are closed convex and pointed cones with non-empty interior.

Recall that $\rho = \Lambda_H^*(Y)$ is defined as follows: $\rho^{(k)} = \langle Y, P_k\rangle_C$, $k = 1, \ldots, n$.

In view of Theorem 6 we have the following dual representation for $\rho \in \operatorname{int} K_{NP}$:

$$\operatorname{Re}(\rho(z)) = \langle [\Lambda_H(c)]^{-1} u(z), u(z)\rangle_C,$$

where the point $c \in \operatorname{int} K_{NT}^*$ is uniquely defined by the equation

$$\rho = \Lambda_H^*([\Lambda_H(c)]^{-1}).$$

4 Discussion

Let us discuss some questions related to applications of the results presented in the Sections 2 and 3. Let K be a cone of polynomials, representable as a sum of squares. For univariate polynomials the inclusion $p \in K$ is equivalent to non-negativity of the function $p(t)$, $t \in \Delta$. We also know that such a cone can be equipped with a self-concordant barrier $F(p)$.

4.1 Polynomial bounds for a trajectory.

First of all note that for any fixed polynomials $l(t)$ and $u(t)$ the set of polynomials p, $l(t) \le p(t) \le u(t)$, $t \in \Delta$, is described by the pair of inclusions $p - l \in K$ and $u - p \in K$. Therefore the two-side polynomial bounds on the coordinates of a polynomial curve $(p_1(t), \ldots, p_m(t))$ can be written in terms of the cone K. Since we can point out a self-concordant barrier for the resulting set, the polynomial-time interior point schemes can be used in order to solve the corresponding optimization problem.

4.2 Minimization of univariate polynomials

Assume we need to minimize a polynomial $p(t)$ over its domain. This problem is equivalent to finding an intersection of the line $\{p+\tau e_0, \tau \in R\}$ with the boundary of the cone K (here

e_0 is the coordinate vector, which corresponds to the constant term of the polynomial). Note that this problem is convex and it can be solved by the following technique.

Let us take an initial value of τ large enough:

$$p + \tau_0 e_0 \in \operatorname{int} K.$$

Then we can iterate the process

$$\tau_{k+1} = \tau_k - \frac{1}{\langle F''(p + \tau_k e_0)e_0, e_0\rangle^{1/2}}, \quad k = 0, \dots .$$

From the general theory of self-concordant functions [6] we get the following rate of convergence for this process:

$$\tau_k - \tau^* \le \frac{(n+1)(\tau_0 - \tau^*)}{\left(1 + \frac{\gamma}{\sqrt{n+1}}\right)^k}, \tag{20}$$

where γ is a positive absolute constant.

4.3 Maximization of polynomial fractions

Let $p_1(t)$ and $p_2(t)$ be some polynomials. We are interested in the maximal τ such that

$$p_1(t) \ge \tau p_2(t)$$

for all $t \in \Delta$. In the case of $p_2 \in \operatorname{int} K$ this problem is equivalent to maximizing the fraction $p_1(t)/p_2(t)$. Thus, our problem is

$$\max\{\tau : \ p_1 - \tau p_2 \in K\}.$$

Provided by a suitable initial value of τ_0 such that $p_1 - \tau_0 p_2 \in \operatorname{int} K$, we can iterate the process

$$\tau_{k+1} = \tau_k + \frac{1}{\langle F''(p_1 - \tau_k p_2)p_2, p_2\rangle^{1/2}}, \quad k = 0, \dots .$$

The rate of convergence of this process is similar to the estimate (20). In the case when the initial value τ_0 cannot be easily found, it can be computed by a preliminary process, based on the standard interior-point technique.

4.4 Dual problems

In Section 2 we have seen that the primal cone K of polynomials representable as a sum of squares can be equipped with an implicit self-concordant barrier. Thus, any optimization problem

$$\begin{aligned} &\min \langle c, p\rangle, \\ \text{s.t} \quad & Ap = b, \\ & p \in K, \end{aligned} \tag{21}$$

can be solved by polynomial-time interior point schemes (see [6, 7, 8]). However, since the barrier for the cone K is implicit, it is reasonable to rewrite the problem (21) in an extended form:

$$\begin{aligned} & \min \langle c, p \rangle, \\ \text{s.t} \quad & Ap = b, \\ & p = \Lambda^*(Y), \\ & p \in K, \; Y \succeq 0. \end{aligned} \tag{22}$$

Then the barrier for the feasible cone of this problem is just $F(p, Y) = -\ln \det Y$.

At the same time, note that in the problem (22) we significantly increase the number of variables. Therefore, from the computational point of view the problem (22) is not very attractive. On the other hand, it is clear that we have a better alternative. Indeed, let us consider the problem dual to (21):

$$\begin{aligned} & \max \langle b, y \rangle, \\ \text{s.t} \quad & s + A^T y = c, \\ & s \in K^*. \end{aligned} \tag{23}$$

We have seen that the inclusion $s \in K^*$ is equivalent to the linear matrix inequality $\Lambda(s) \succeq 0$. Therefore the problem (23) has, in fact, the following form:

$$\begin{aligned} & \max \langle b, y \rangle, \\ \text{s.t} \quad & s + A^T y = c, \\ & \Lambda(s) \succeq 0. \end{aligned} \tag{24}$$

Note that the barrier for the feasible cone of the problem (24) is $F_*(s) = -\ln \det \Lambda(s)$. In order to apply interior-point schemes to (24) we need to compute the gradient and the Hessian of this barrier. The expressions for these objects includes the matrix $\Lambda(s)^{-1}$. In Section 3 we have seen that for univariate polynomials $\Lambda(s)$ is a kind of a Hankel or a Toeplirtz matrix. It is wellknown that, for example, a product of an inverse Toeplitz matrix with a vector can be computed by Fast Fourier Transform (FFT) in $O(n \ln^2 n)$ operations. It is very interesting to study the possibility to apply the FFT technique for computing the objects important for the interior point schemes. We may hope that this could lead to superfast optimization methods for problems related to non-negative univariate polynomials.

4.5 Combinatorial problems

It is wellknown that some NP-hard combinatorial problems can be rewritten as a problem of minimizing a multivariate polynomial. Consider, for example, the problem of finding a boolean solution $x^{(i)} = \pm 1$, $i = 1, \dots, n$, satisfying a single linear equation $\langle c, x \rangle = 0$. This problem is equivalent to the following minimization problem:

$$\min_{x \in R^n} \left[p(x) = \langle c, x \rangle^4 + n \sum_{i=1}^{n} (x^{(i)})^4 - (\sum_{i=1}^{n} (x^{(i)})^2)^2 \right]. \qquad (25)$$

The function $p(x)$ is a kind of interest because of the following properties.

- It is non-negative for any x and $p(x) = 0$.
- The set of the global solutions of the problem (25) is described as follows:

$$x^{(i)} = \pm \tau, \; i = 1, \dots, n, \quad \langle c, x \rangle = 0.$$

- This function is homogeneous in x of degree four. Therefore, it has no local minimum: $p'(x) = 0$ implies $p(x) = 0$.

At the same time, it is clear that $p(x) \in K_{n,4}$ (see Example 1 for definition of the cone). Indeed, denote

$$\begin{aligned} p_c(x) &= \langle c, x \rangle^4, \\ p_2(x) &= (\sum_{i=1}^{n} (x^{(i)})^2)^2, \\ p_4(x) &= n \sum_{i=1}^{n} (x^{(i)})^4, \\ p_3(x) &= p_4(x) - p_2(x). \end{aligned}$$

Note that $p_c(x) \in \partial K_{n,4}$ and $p_2(x) \in \operatorname{int} K_{n,4}$ in view of Corollary 1. At the same time,

$$p_3(x) = \langle (nI_n - ee^T)[x]^2, [x]^2 \rangle,$$

where e is the vector of ones and $[x]^2 \in R^n$ is the vector with the components $(x^{(i)})^2$. Therefore $p_3(x) \in K_{n,4}$ and it is a boundary point of this cone since it has proper roots. As a consequence we also get $p_4(x) \in \operatorname{int} K_{n,4}$.

Thus, the geometric interpretation of the function $p(x)$ is as follows: it is a sum of two boundary points of the cone $K_{n,4}$. The sufficient condition for $p(x) > 0$, $x \neq 0$, is that $p(x) \in \operatorname{int} K_{n,4}$. It is interesting to study the sharpness of that condition. In any case, it seems reasonable to suppose that the existence of solutions of the boolean equation $\langle c, x \rangle = 0$ is coded in some way in the structure of the boundary of the cone $K_{n,4}$ at the point $p_3(x)$.

Another interesting application of the cone $K_{n,4}$ is related to the following NP-hard problem:

$$\text{Describe the cone } M_+^n = \{ A \in R^{n \times n} : \; \langle Ax, x \rangle \geq 0 \quad \forall x \geq 0 \in R^n \}.$$

It is clear that the cone M_+^n contains the following cones:

- The cone S_+^n of positive semidefinite $(n \times n)$-matrices.
- The cone $R_+^{n \times n}$ of $(n \times n)$-matrices with non-negative coefficients.

Let us consider the following matrix cone

$$P_+^n = \{P \in R^{n\times n} : \; p(x) \equiv \langle P[x]^2, [x]^2 \rangle \in K_{n,4}\}.$$

It can be checked easily that

$$S_+^n + R_+^{n\times n} \subseteq P_+^n \subseteq M_+^n.$$

Thus, an interesting question is how sharp is the last relation.

4.6 Minimization of multivariate polynomials.

In Section 4.5 we have seen that the problem of minimizing a multivariate polynomial is extremely difficult. However, the results of Section 2 suggest a kind of semidefinite relaxation for the problems of that type. Indeed, instead of finding a global minimum of such polynomial, we can find a smallest value of the constant term (with other coefficients being fixed) which keeps the possibility to represent the polynomial as a sum of squares. Such value can be computed in polynomial time and it gives some estimate for the global minimum of the polynomial. In one-dimensional case this estimate is exact. It seems very interesting to study its sharpness in multi-dimensional case.

On the other hand, it is clear that the problem of minimizing a multivariate polynomial can be easily solved if we are able to prove its non-negativity at any x. Indeed, in such a case we can emloy a trivial dichotomy scheme as applied to the value of the constant term of the initial polynomial. At the same time, in accordance with the seventeenth Hilbert hypotesis (which was proved for some dimensions) a multi-variate polynomial is non-negative if and only if it can be represented as a sum of squares of rational functions:

$$p(x) = \sum_{i=1}^{N} \left[\frac{r_i(x)}{q_i(x)}\right]^2, \tag{26}$$

where $r_i(x)$ and $q_i(x)$ are some polynomials. Unfortunately, no reasonable bounds are known for the number of terms N in the above sum and for the degree of polynomials $p_i(x)$ and $q_i(x)$. However, since we are interested even in an approximate solution of this problem, we can impose some artificial bounds on these values. If the bounds are small absoulte constants, then we get a feasibility problem which can be solved in polynomial time.

4.7 Duality relations

Let K be a proper cone. In Sections 2, 3 we have seen that for any $p \in \operatorname{int} K$ we can define a dual object $p^* \in \operatorname{int} K^*$, which is a unique solution of the following nonlinear equation:

$$p = -F_*'(p^*),$$

where $F_*(\cdot)$ is a self-concordant barrier for the dual cone K^*. It seems that such type of dual relations does not appear in the traditional theory of polynomials. At the same time, we may expect that some properties of the polynomial $p(x)$ can have a natural explanation through its dual counterpart. In this paper we restrict ourselves by a simple example of such an interaction.

Let $K \subset R^n$ be a cone of polynomials representable as a sum of squares. We have seen that the natural barrier for the dual cone K^* is as follows:

$$F_*(c) = -\ln \det \Lambda(c).$$

Therefore, the duality relation between $p \in \operatorname{int} K$ and $p^* \in \operatorname{int} K^*$ is given by the following equation:

$$p = \Lambda^*([\Lambda(p^*)]^{-1}).$$

As a simple consequence of that equation we get

$$\langle p^*, p \rangle = \langle p^*, \Lambda^*([\Lambda(p^*)]^{-1}) \rangle = \langle \Lambda(p^*), [\Lambda(p^*)]^{-1} \rangle = n.$$

Lemma 2 *Let $s_0 \in \operatorname{int} K^*$ and $\bar{s} \in R^n$. Assume that the hyperplane*

$$\{p \in R^n : \langle \bar{s}, p \rangle = 1\}$$

intersects the interior of the cone K. Then

$$\begin{array}{lll} f_* \equiv & \min\limits_p \langle s_0, p \rangle & = [\lambda_{\max}([\Lambda(s_0)]^{-1/2}\Lambda(\bar{s})[\Lambda(s_0)]^{-1/2})]^{-1}, \\ s.t. & \langle \bar{s}, p \rangle = 1, & \\ & p \in K & \end{array} \tag{27}$$

where $\lambda_{\max}(\cdot)$ is the maximal eigenvalue of the corresponding matrix.

If in addition $\Lambda(\bar{s})$ is a rank-one positive semidefinite matrix: $\Lambda(\bar{s}) = \bar{y}\bar{y}^T$ for some $\bar{y} \in R^n$, then $f_* = \langle [\Lambda(s_0)]^{-1}\bar{y}, \bar{y} \rangle^{-1/2}$ and the unique solution of the above minimization problem is given by the polynomial $p(x) = \langle q, u(x) \rangle^2$ where

$$q = \frac{[\Lambda(s_0)]^{-1}\bar{y}}{\langle [\Lambda(s_0)]^{-1}\bar{y}, \bar{y} \rangle^{1/2}}.$$

Proof:
In view of the conic duality (see [6]) the problem dual to (27) has the following form:

$$\begin{array}{ll} & \max\limits_\tau \quad \tau \\ \text{s.t.} & s_0 - \tau \bar{s} \in K^*. \end{array}$$

Since the problem (27) is strictly feasible, the optimal values of the primal and dual problems coincide. Note that the constraint of the dual problem can be written as

$$\Lambda(s_0) \succeq \tau \Lambda(\bar{s}).$$

Since $\Lambda(s_0) \succ 0$, we conclude that

$$f^* = \frac{1}{\lambda_{\max}([\Lambda(s_0)]^{-1/2}\Lambda(\bar{s})[\Lambda(s_0)]^{-1/2})}.$$

Finally, if $\Lambda(\bar{s}) = \bar{y}\bar{y}^T$ then

$$\lambda_{\max}([\Lambda(s_0)]^{-1/2}\Lambda(\bar{s})[\Lambda(s_0)]^{-1/2}) = \| [\Lambda(s_0)]^{-1/2}\bar{y} \|^2 = \langle [\Lambda(s_0)]^{-1}\bar{y}, \bar{y} \rangle.$$

Let $q = [\Lambda(s_0)]^{-1}\bar{y}/\langle[\Lambda(s_0)]^{-1}\bar{y}, \bar{y}\rangle$ and $p(x) = \langle q, u(x)\rangle^2$. Then, in view of (2) we have:

$$\begin{aligned} \langle p, v(x)\rangle &\equiv p(x) = \langle q, u(x)\rangle^2 = \langle qq^T, u(x)u(x)^T\rangle \\ &= \langle qq^T, \Lambda(v(x))\rangle = \langle \Lambda^*(qq^T), v(x)\rangle. \end{aligned}$$

Therefore $p = \Lambda^*(qq^T)$ and

$$\begin{aligned} \langle \bar{s}, p\rangle &= \langle \bar{s}, \Lambda^*(qq^T)\rangle = \langle \Lambda(\bar{s}), qq^T\rangle = \langle \bar{y}, q\rangle^2 = 1, \\ \langle s_0, p\rangle &= \langle \Lambda(s_0)q, q\rangle = \langle[\Lambda(s_0)]^{-1}\bar{y}, \bar{y}\rangle^{-1} = f_*. \end{aligned}$$

The solution is unique since the matrix $\bar{y}\bar{y}^T$ has only one eigenvector for its maximal eigenvalue. □

Let us fix some $x_0 \in \Delta$. Then, in view of (2) the matrix $\Lambda(\bar{s})$ with $\bar{s} = v(x_0)$ has the following form:

$$\Lambda(\bar{s}) = \Lambda(v(x_0)) = u(x_0)u(x_0)^T.$$

On the other hand, in this case we have $\langle \bar{s}, p\rangle = p(x_0)$. Let $p \in \operatorname{int} K$. Then the statement of Lemma 2 can be rewritten as follows:

$$\frac{\langle s_0, p\rangle}{p(x_0)} \geq \frac{1}{\langle[\Lambda(s_0)]^{-1}u(x_0), u(x_0)\rangle} = \frac{1}{p_0(x_0)},$$

where p_0 is a dual polynomial for $s_0 \in \operatorname{int} K^*$. Therefore we get the following result.

Corollary 2 *For any $p\ q \in \operatorname{int} K$ and $x \in \Delta$ we have*

$$\frac{p(x)}{\langle q^*, p\rangle} \leq q(x) \leq \langle p^*, q\rangle p(x).$$

The right-hand side of this inequality is valid also for $q \in \partial K$.

References

[1] R.Hettich and K.O.Kortanek, "Semi-infinite programming: theory, methods and applications," *SIAM Review*, 35(1993) No 3, 380–429.

[2] S.Karlin and W.J.Studden, *Tchebyshev Systems: with Applications in Analysis and Statistics*, Interscience Publishers, John Willey & Sons, New York, 1966.

[3] Lukacs. Verscharfung der ersten Mittelwertsatzes der Integralrechnung fur rationale Polynome, Math. Zeitschrift, 2 (1918), 229-305."

[4] J.C.Maithuber. "On Haar's theorem concerning Chebyshev approximation problems having unique solutions, Proc. amer. Math. Soc. 7(1956), 609–615.

[5] Markov, A.A. Lecture notes on functions with the least deviation from zero, (1906). Reprinted in "Markov A.A. Selected papers.", (ed. N.Achiezer), GosTechIzdat, 1948, Moscow (in Russian), 244-291.

[6] Yu. Nesterov and A. Nemirovsky, *Interior Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.

[7] Yu. Nesterov, M. J. Todd. Self-scaled Barriers and Interior-Point Methods for Convex Programming. *Mathematics of Operation Research*, 22(1997) 1, 1-42.

[8] Yu. Nesterov, M. J. Todd. Primal-Dual Interior-Point Methods for Self-scaled Cones. To be published in *SIAM Journal of Optimization*, 1995.

[9] S.-P.Wu, S. Boyd and L.Vandenberghe, "FIR Filter Design via Special Factorization and Convex Optimization", to appear in Chapter 1, *Applied Computational Control, Signal and Communications*, Beswa Datta ed., Birkhauser, 1997.

Appendix

Let Q be a closed convex set with nonempty interior endowed with a ν-self-concordant barrier $F(x)$. For $x \in \operatorname{int} Q$ denote

$$\| p \|_x = \langle F''(x)p, p \rangle^{1/2}, \quad p \in R^n.$$

Consider $\bar{x} \in \operatorname{int} Q$. Assume that there exists a set of *recession* directions $\{p_1, \ldots, p_k\}$:

$$\bar{x} + \alpha p_i \in Q \quad \forall \alpha \geq 0, \quad i = 1, \ldots, k.$$

The following statement sometimes is useful for getting a lower bound for the parameter of a self-concordant barrier for the set Q.

Lemma 3 *Let the positive coefficients $\{\beta_i\}_{i=1}^k$ satisfy the condition*

$$\bar{x} - \beta_i \, p_i \notin \operatorname{int} Q, \quad i = 1, \ldots, k.$$

If for some positive $\alpha_1, \ldots, \alpha_k$ we have $\bar{y} = \bar{x} - \sum\limits_{i=1}^{k} \alpha_i p_i \in Q$, then the parameter ν of any self-concordant barrier for Q satisfies the inequality:

$$\nu \geq \sum_{i=1}^{k} \frac{\alpha_i}{\beta_i}.$$

Proof:
Let $F(x)$ be a ν-self-concordant barrier for the set Q. Since p_i is a recession direction, we have:

$$\langle F'(\bar{x}), -p_i \rangle \geq \langle F''(\bar{x})p_i, p_i \rangle^{1/2} \equiv \| p_i \|_{\bar{x}},$$

(since otherwise the function $f(t) = F(\bar{x} + tp)$ attains its minimum; see Theorem 2.2.2(i) in [6]).

Note that $\bar{x} - \beta_i \, p_i \notin Q$. Therefore, in view of Theorem 2.1.1(ii) [6] the norm of the direction p_i is large enough: $\beta_i \| p_i \|_{\bar{x}} \geq 1$. Hence, in view of Proposition 2.3.2(i.2) [6] we obtain:

$$\nu \geq \langle F'(\bar{x}), \bar{y} - \bar{x} \rangle = \langle F'(\bar{x}), -\sum_{i=1}^{k} \alpha_i p_i \rangle \geq \sum_{i=1}^{k} \alpha_i \| p_i \|_{\bar{x}} \geq \sum_{i=1}^{k} \frac{\alpha_i}{\beta_i}.$$

□