A CUTTING PLANE METHOD FOR SOLVING CONVEX OPTIMIZATION PROBLEMS OVER THE CONE OF NONNEGATIVE POLYNOMIALS

Iurie Caraus
Moldova State University
Fac. of Mathematics and Informatics
Mateevici 60 str., Chisinau
Republic of Moldova, MD-2009
caraush@usm.md

Ion Necoara
University Politehnica Bucharest
Fac. of Automation and Computers
Automation and System Engineering Dept.
Spl. Independentei 313
060042 Bucharest, Romania
i.necoara@yahoo.com

Abstract: Many practical problems can be formulated as convex optimization problems over the cone of nonnegative univariate polynomials. We use a cutting plane method for solving this type of optimization problems in primal form. Therefore, we must be able to verify whether a polynomial is nonnegative, i.e. if it does not have real roots or all real roots are multiple of even order. In this paper an efficient method is derived to determine a scalar value for which the polynomial is negative and in the case that such a value exists a feasible cut is constructed. Our method is based on Sturm theorem, which allows to determine the number of distinct roots of a polynomial on a given interval, in combination with the bisection method. For numerical stability we construct the associated Sturm sequence using Chebyshev basis, and thus we can work with high degree polynomials, up to hundreds. Numerical results show the efficiency of our new approach.

Key–Words: convex optimization problems over the cone of nonnegative polynomials, cutting plane method, Chebyshev polynomials, Sturm sequence, feasible cut.

1 Introduction

Nonnegative polynomials on the real line play a fundamental role in systems and control theory: e.g., robust control [11], filter design problems [10], etc. The set of nonnegative univariate polynomials form a convex set, more precisely a convex cone and was recently studied in the area of convex optimization [8]. Convex optimization techniques allow us to efficiently treat the cone of nonnegative polynomials, which can be parameterized by semidefinite matrices [9, 15]. In [8] it was shown that for a convex optimization problem over this particular cone, the corresponding dual problem leads to a special semidefinite matrix structure, more specifically a dual Hankel-matrix structure. Therefore, various interior-point and cutting plane algorithms exploiting this dual semidefinite Hankel-structure were proposed in [8].
However, for large-scale problems the existing semidefinite interior-point based algorithms are inefficient since we must solve linear systems with large dimension matrices. Moreover, the methods mentioned above utilize the classical basis for polynomials which limit the polynomial degree. This originates from the fact that the polynomials are described using the natural powers $1, x, x^2, \cdots$, which lead to ill-conditioned Hankel matrices.

In this article we also consider convex problems over the cone of nonnegative polynomials on the real line, subject to linear constraints. We restrict ourselves to first-order methods, i.e. methods that use only first-order information (function evaluation and gradient evaluation). In particular we propose a cutting plane algorithm [5] for solving this type of problems in primal form without embedding it into a semidefinite format corresponding to the dual formulation. Tackling the problem directly in the primal form permits us to remove some of the disadvantages of the methods mentioned above, in particular it allows us to take into account the structure of the

given problem such as sparsity. Our contribution is twofold.
First, for the numerical stability of our method it is important that the polynomials are represented with respect to some orthogonal basis, e.g. the Chebyshev basis. This approach permits to work with high degree polynomials, up to hundreds. The explanation is derived from the observation that the Chebyshev polynomials have Sturm sequences with small coefficients and thus our method is numerically stable. Moreover, using Chebyshev basis for representing nonnegative polynomials, the corresponding dual cone leads to a semidefinite Toeplitz-plus-Hankel matrix structure, i.e. with well-behaved matrices.
Second, for using a cutting plane algorithm we need an efficient method to construct a feasible cut. In our case this reduces to checking nonnegativity of a polynomial. If the polynomial does not have real roots or all real roots are multiple of even order, then it is nonnegative. If the polynomial is negative then, we propose an efficient method for constructing a feasible cut using bisection in combination with Sturm theorem for determining some scalar for which the polynomial value is negative. We construct the Sturm sequence in the Chebyshev basis and with Sturm theorem we can determine the number of distinct real zeros of a polynomial in a given interval. We determine a scalar value for which the polynomial is negative by Sturm theorem with bisection and with this value a feasible hyperplane is constructed and used in the cutting plane method.

The layout of the paper is as follows. Section 2 gives the problem formulation, the main ingredients of a cutting plane method and the mathematical definitions which are used to derive Sturm theorem. In Section 3 we derive a method for counting the real zeros of a given polynomial within an interval. The method uses the notion of Cauchy index, replacing the problem of counting zeros lying in an interval, by the evaluation of a Cauchy index of a rational function associated with that polynomial in the Chebyshev form. In Section 4 we propose an efficient method for constructing a feasible cut based on bisection and Sturm theorem. We conclude with an application in Section 5.

2 Preliminaries

In this section we introduce the problem that we are going to solve and the mathematical tools that will be used in the paper.

2.1 Problem formulation

In this paper the following convex problem is considered:

$$\begin{aligned} &\min_f \; G(f) \\ &\text{s.t. } \; Af \le b \\ &\qquad f \in \mathcal{K}_{\mathcal{R}}, \end{aligned} \tag{1}$$

where $G : \mathcal{R}^{n+1} \to \mathcal{R}$ is a convex function, A is a $\mathcal{R}^{m\times(n+1)}$ matrix and $b \in \mathcal{R}^m$. We consider that n is even and in general $m \le n+1$. We assume that the set $\{f \in \mathcal{R}^{n+1} : Af \le b\}$ is bounded. Moreover, $\mathcal{K}_{\mathcal{R}}$ is the convex cone of coefficients of polynomials nonnegative on the interval (a, b), i.e. for $f = [f_0 \cdots f_n]^T$

$$\mathcal{K}_{\mathcal{R}} = \{f \in \mathcal{R}^{n+1} : f_0 + f_1 x + \cdots + f_n x^n \ge 0, \quad \forall x \in (a,\ b)\}.$$

There exist different methods for solving convex problems of the form (1), such as interior-point based methods in primal and dual form [8, 9], first-order methods [7], etc.

In this paper we use a cutting plane method for solving (1). The basic scheme of a cutting plane method for solving a general convex problem

$$\min_{x\in Q} g(x),$$

where g is a convex function and Q is a convex set in some vector space endowed with a scalar product $\langle\cdot,\cdot\rangle$, consists in the following steps:

Algorithm 1 (*cutting plane method* for solving convex problem $g^* = \min_{x\in Q} g(x)$)

1. choose an accuracy $\varepsilon > 0$ and a starting point x^0

2. k^{th} iteration $(k \ge 0)$:

2.1 if $x^k \notin Q$ (feasibility cut)

$$\exists a_k, b_k \text{ s.t. } \langle a_k, x - x^k\rangle + b_k \le 0 \quad \forall x \in Q$$

2.2 if $x^k \in Q$ (optimality cut)

$$g(x) \ge g(x^k) + \langle \nabla g(x^k), x - x^k\rangle$$
$$\forall x \in \mathrm{dom}(g)$$

3. compute x^{k+1}

4. test $g(x^{k+1}) - g^* \leq \varepsilon$.

Here, $\nabla g(x)$ denotes the (sub)gradient of the function g and ε is the required accuracy of the approximate solution. There are different strategies to update x^{k+1} in step 3 of a cutting plane method (Algorithm 1): e.g. center of gravity [9], ellipsoid [6], analytic center [7], etc.

2.2 Sturm theorem

Note that in order to solve the optimization problem (1) using a cutting plane method we should check if a polynomial belongs to the convex cone $\mathcal{K}_{\mathcal{R}}$. Sturm theorem allows us to decide if our polynomial has real roots in the interval $(a,\ b)$. In the sequel we introduce the basic ingredients for deriving Sturm theorem (see [3, 4] for more details). Let $f(x)$ be a real polynomial, i.e. in the classical basis it can be written as:

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

where f_k are real numbers. In the next definition we introduce the so-called *Cauchy index*:

Definition 1 *The Cauchy index of a real rational function $R(x)$ between the limits a and b (notation $I_a^b R(x)$; a and b are real numbers or $\pm\infty$) is the difference between the number of jumps of $R(x)$ from $-\infty$ to $+\infty$ and that of jumps from $+\infty$ to $-\infty$ as the argument changes from a to b. In counting the number of jumps, the extreme values a and b are not included.*

One of the methods of computing the Cauchy index $I_a^b R(x)$ is based on the classical Sturm sequence:

Definition 2 *We consider a sequence of polynomials:*

$$f_0(x),\ f_1(x), \ldots, f_m(x). \tag{2}$$

Such a sequence is called a Sturm sequence on an interval (a,b), where either a or b may be infinite, if

(i) $f_m(x)$ does not vanish in (a,b);

(ii) at any zero of $f_k(x)$, $k = 1, \cdots, m-1$, the two adjacent functions are nonzero in (a,b) and have opposite signs; that is,

$$f_{k-1}(x) f_{k+1}(x) < 0.$$

Definition 3 *Let $\{f_i(x)\}_{i=0,\ldots,m}$ be a Sturm sequence on (a,b), and let x_0 be a point of (a,b) at which $f_0(x_0) \neq 0$. We define $V(x_0)$ to be the number of changes of sign of $\{f_i(x_0)\}_i$, zero values being ignored. If*

(i) a is finite, then $V(a)$ is defined as $V(a+\varepsilon)$, where ε is such that no $f_i(x)$ vanishes in $(a, a+\varepsilon)$ and similarly for b when b is finite;

(ii) $a = -\infty$, then $V(a)$ is defined to be the number of changes of sign of $\{\lim_{x\to-\infty} f_i(x)\}_i$ and similarly for $V(b)$ when $b = \infty$.

The next theorem provides a way to compute the Cauchy index:

Theorem 4 *(Sturm theorem) [3] If $\{f_i(x)\}_i$ is a Sturm sequence in (a,b) and $V(x)$ is the number of variations of sign in the sequence, then*

$$I_a^b \frac{f_1(x)}{f_0(x)} = V(a) - V(b). \tag{3}$$

It is easy to see that by means of Sturm theorem the number of distinct real roots of the polynomial $f(x)$ in the interval (a,b) is $I_a^b \frac{f'(x)}{f(x)}$, where $f'(x)$ denotes the derivative of $f(x)$.

3 Sturm sequence in Chebyshev basis

Any algorithm for devising the Sturm sequence (e.g. Euclidian algorithm) is numerically unstable when the polynomials are represented in the classical basis since in general we have to perform polynomial division. Therefore, in this section we derive the Sturm sequence associated to a given polynomial using the shifted Chebyshev basis. This type of polynomials are orthogonal on the given interval and satisfies a three term recurrence relation.

3.1 Constructing Sturm sequence using shifted Chebyshev polynomials

We define shifted Chebyshev polynomials on a given interval $(a,\ b)$ using the following linear transformation [2]

$$s(x) = \frac{2x-(a+b)}{b-a},\ \ x \in (a,\ b).$$

The shifted Chebyshev polynomials of the first kind of degree i on (a,b) denoted $T_i(x)$ satisfy the following three-recurrence relation:

$$T_0(x) = 1$$
$$T_1(x) = s(x)$$
$$T_{i+1}(x) = 2s(x)T_i(x) - T_{i-1}(x),\ \ \forall i \geq 1.$$

Shifted Chebyshev polynomials of the second kind $U_i(x)$ are defined as follows:

$$U_0(x) = 1$$
$$U_1(x) = 2s(x)$$
$$U_{i+1}(x) = 2s(x)U_i(x) - U_{i-1}(x),\ \ \forall i \geq 1.$$

Given a polynomial $f(x)$ using the shifted Chebyshev basis, a Sturm sequence may be formed using Euclid algorithm for computing the greatest common divisor between $f_0(x) = f(x)$ and $f_1(x) = f'(x)$. Divide $f_0(x)$ by $f_1(x)$ and take $f_2(x)$ to be the negative of the resulting remainder. Proceeding in this fashion, the degree of $f_i(x)$ is always less than $f_{i-1}(x)$.

Let us assume that our polynomial can be represented in the shifted Chebyshev basis of the first kind as:

$$f(x) = \sum_{k=0}^{n} a_k T_k(x),\ \ \ x \in (a,\ b).$$

It is easy to show the following relation between the first and second kind Chebyshev polynomials:

$$[T_k(x)]' = \frac{2k}{b-a}U_{k-1}(x).$$

Based on this relation we find that

$$f_1(x) = \sum_{k=0}^{n-1} b_k U_k(x),$$

where

$$b_k = \frac{2(k+1)}{b-a}a_{k+1},\ \ k = 0,\dots,n-1. \tag{4}$$

The remainder $f_2(x)$ can be computed as $f_2(x) = (\alpha_1^1 T_1(x) + \alpha_2^1)f_1(x) - f_0(x)$, i.e. takes the following form:

$$\begin{aligned} f_2(x) =& [\frac{1}{2}a_2 - a_0 + \frac{1}{2}\alpha_1^1 b_1 + \alpha_2^1 b_0]U_0(x) + \\ & \sum_{k=1}^{n-2}[\frac{1}{2}a_{k+2} + \frac{1}{2}\alpha_1^1 b_{k+1} + \\ & \qquad \frac{1}{2}\alpha_1^1 b_{k-1} + \alpha_2^1 b_k - \frac{1}{2}a_k]U_k(x), \end{aligned}$$

where $\alpha_1^1 = \frac{a_n}{b_{n-1}}$ and $\alpha_2^1 = \frac{a_{n-1} - b_{n-2}\alpha_1^1}{2b_{n-1}}$.

For the general case (i.e. $j \geq 2$) we assume that $f_{j-1}(x) = \sum_{k=0}^{n-j+1} c_k U_k(x)$ and $f_j(x) = \sum_{k=0}^{n-j} d_k U_k(x)$. Writing the remainder $f_{j+1}(x) = (\alpha_1^j U_1(x) + \alpha_2^j)f_j(x) - f_{j-1}(x)$ we find that

$$f_{j+1}(x) = [\alpha_1^j d_1 + \alpha_2^j d_0 - c_0]U_0(x)$$
$$+ \sum_{k=1}^{n-j-1}[\alpha_1^j d_{k-1} + \alpha_1^j d_{k+1} + \alpha_2^j d_k - c_k]U_k(x),$$

where

$$\alpha_1^j = \frac{c_{n-j+1}}{d_{n-j}},\ \alpha_2^j = \frac{c_{n-j} - d_{n-j-1}\alpha_1^j}{d_{n-j}}. \tag{5}$$

It is important to see that for any scalar $\beta \in (a,\ b)$, the evaluation of the Sturm sequence at this value can be done recursively, without explicitly evaluating $f_j(x)$ at β. Indeed, given $f_0(\beta)$ and $f_1(\beta)$ (these values can be computed since we know explicitly the polynomial $f(x)$), we have the following recurrent relation:

$$f_{j+1}(\beta) = (\alpha_1^j U_1(\beta) + \alpha_2^j)f_j(\beta) - f_{j-1}(\beta)$$

for all $j \geq 2$. Therefore, for finding the number of changes of sign of the sequence $\{f_i(\beta)\}_{i\geq 0}$ we need to know the following two sequences $\{\alpha_1^i\}_{i\geq 0}$ and $\{\alpha_2^i\}_{i\geq 0}$. From the above discussion we can conclude that the computational complexity for finding the Sturm sequence is $\mathcal{O}(n^2)$.

Remark 5 *Note that we can also have the case when the difference in the degree of two consecutive polynomials in the Sturm sequence is larger than one. In this case it is obvious that the quotient is not linear. However, the derivation of the coefficients of the quotient are computed similarly as above for the linear case.*

From the previous discussion we can derive the following theorem:

Theorem 6 *The number N of distinct real roots inside the interval $(a,\ b)$ of a given polynomial $f(x)$ written in the Chebyshev basis is given by*

$$N = V(a) - V(b).$$

3.2 Destroying multiplicity

In this section we derive the Sturm sequence in the case when the polynomial has multiple roots. If we have multiple roots, we add a small positive constant ε to our polynomial in order to destroy its multiplicity. We construct the Sturm sequence as in the previous section: let $f_0^\varepsilon(x)$ be the perturbed polynomial, i.e.

$$f_0^\varepsilon(x) = \sum_{k=1}^{n} a_k T_k(x) + [\varepsilon + a_0]T_0(x).$$

Note that $f_1^\varepsilon(x) = \sum_{k=0}^{n-1} b_k U_k(x)$, where b_k is defined in (4). As in the previous section we obtain

$$\begin{aligned} f_2^\varepsilon(x) =& [\frac{1}{2}a_2 - (a_0+\varepsilon) + \alpha_1 b_1 \frac{1}{2} + \alpha_2 b_0]U_0(x) \\ &+ \sum_{k=1}^{n-2}[\frac{1}{2}a_{k+2} + \alpha_1 b_{k+1}\frac{1}{2} + \alpha_1 b_{k-1}\frac{1}{2} + \\ &\alpha_2 b_k - \frac{1}{2}a_k]U_k(x), \end{aligned}$$

where $\alpha_1 = \dfrac{a_n}{b_{n-1}}$ and $\alpha_2 = \dfrac{a_{n-1} - b_{n-2}\alpha_1}{2b_{n-1}}$.

By induction we can show the following:

(i) if $j = 1 \cdots n/2$

$$f_{j+1}^\varepsilon(x) = \sum_{k=0}^{j-1} r_k(\varepsilon)U_k(x) + \sum_{k=j}^{n-j-1} r_k U_k(x),$$

where for the coefficients $r_k(\varepsilon)$ the dependence from ε is linear.

(ii) if $j = n/2+1 \cdots n$, taking into account (5) we obtain

$$f_{j+1}^\varepsilon(x) = \sum_{k=j+1}^{n} r_{n-k}(\varepsilon)U_{n-k}(x),$$

where for the coefficients $r_k(\varepsilon)$ the dependence from ε is rational, i.e. $r_k(\varepsilon) = P(\varepsilon)/Q(\varepsilon)$, with $P(\varepsilon)$ and $Q(\varepsilon)$ are real polynomials, with $\deg(P) = \deg(Q)+1$.

Theorem 6 can also be applied in this case, which allow us to determine the number of distinct real roots of our original polynomial.

3.3 Comrade Matrix Approach

Using the Sturm sequence method developed in Section 3 we can check if a given polynomial has real roots or not on some interval. If the associated Cauchy index $I_a^b \dfrac{f'}{f} = 0$ then the polynomial is nonnegative on the interval (a,b). Another method to decide if the polynomials is nonnegative or not on the interval (a,b) is given by the comrade matrix approach.

For an orthogonal basis of polynomials satisfying a 3-term recurrence relation

$$\begin{gathered} f_0(x) = \alpha_0, \\ f_1(x) = \alpha_1 x + \beta_1, \\ f_i(x) = (\alpha_i x + \beta_i)f_{i-1}(x) - \gamma_i f_{i-2}(x), \\ i \geq 2 \quad (\text{with } \gamma_i > 0), \end{gathered}$$

and a polynomial f represented in this basis as

$$f(x) = \delta(f_n(x) + a_1 f_{n-1}(x) + \dots a_n f_0(x)),$$

we define the comrade matrix

$$C = \begin{bmatrix} -\frac{\beta_1}{\alpha_1} & \frac{1}{\alpha_1} & 0 & \dots & \dots & -\frac{a_n}{\alpha_n} \\ \frac{\gamma_2}{\alpha_2} & -\frac{\beta_2}{\alpha_2} & \frac{1}{\alpha_2} & \dots & \dots & -\frac{a_{n-1}}{\alpha_n} \\ 0 & \frac{\gamma_3}{\alpha_3} & -\frac{\beta_3}{\alpha_3} & \frac{1}{\alpha_3} & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & -\frac{a_3}{\alpha_n} \\ \dots & \dots & \dots & \dots & \dots & -\frac{a_2+\gamma_n}{\alpha_n} \\ \dots & \dots & \dots & \dots & \frac{\gamma_{n-1}}{\alpha_{n-1}} & -\frac{a_1+\beta_n}{\alpha_n} \end{bmatrix}$$

It is known that the zeros of the polynomial f are the eigenvalues of the above matrix C. Recently methods with $O(n^2)$ complexity were developed to solve the comrade eigenvalue problem [17], [18]. The comrade matrix approach consists in the following steps:

- compute the eigenvalues of the matrix C associated to our polynomial
- take the middle points and evaluate f in these points
- if all these values are nonnegative then f belongs to $\mathcal{K}_{\mathcal{R}}$.

3.4 Example

We consider the polynomial

$$f = \sum_{k=0}^{n} a_k T_k(x)$$

on the interval $(a,\ b)$. We take a polynomial of degree $n = 24$ on the interval $(1,\ 10)$. The corresponding Chebyshev coefficients for this polynomial are given in the following table:

$a_0 = 2.63$	$a_1 = 1.70$	$a_2 = 0.28$
$a_3 = 0.27$	$a_4 = 0.29$	$a_5 = 0.37$
$a_6 = 0.25$	$a_7 = 0.47$	$a_8 = 0.18$
$a_9 = 0.52$	$a_{10} = 0.15$	$a_{11} = 0.53$
$a_{12} = 0.24$	$a_{13} = 0.51$	$a_{14} = 0.47$
$a_{15} = 0.47$	$a_{16} = 0.69$	$a_{17} = 0.39$
$a_{18} = 0.72$	$a_{19} = 0.26$	$a_{20} = 0.50$
$a_{21} = 0.12$	$a_{22} = 0.22$	$a_{23} = 0.03$
$a_{24} = 0.04$		

For $a = 1$ the values of the Sturm sequence are displayed in the following table and the number of sign changes is 13:

$f_0 = 1$	$f_1 = -65.54$	$f_2 = 11.12$
$f_3 = -84.61$	$f_4 = 9.105$	$f_5 = -59.44$
$f_6 = 6.60$	$f_7 = -36.87$	$f_8 = 4.38$
$f_9 = -19.62$	$f_{10} = 1.95$	$f_{11} = -4.74$
$f_{12} = -0.44$	$f_{13} = -44.73$	$f_{14} = 0.062$
$f_{15} = 71.57$	$f_{16} = -0.02$	$f_{17} = -182.48$
$f_{18} = -0.14$	$f_{19} = -819.89$	$f_{20} = -1.14$
$f_{21} = -4987.3$	$f_{22} = -0.051$	$f_{23} = -1445.83$
$f_{24} = -0.026$		

For $b = 10$ the values of the Sturm sequence are given below and the number of sign changes are 11:

$f_0 = 12.42$	$f_1 = 341.76$	$f_2 = 52.53$
$f_3 = 344.69$	$f_4 = 36.54$	$f_5 = 209.95$
$f_6 = 22.01$	$f_7 = 104.07$	$f_8 = 8.68$
$f_9 = 7.87$	$f_{10} = -6.25$	$f_{11} = -88.40$
$f_{12} = -21.44$	$f_{13} = -2882.2$	$f_{14} = 1.18$
$f_{15} = -1962.3$	$f_{16} = 0.70$	$f_{17} = -1718.1$
$f_{18} = 0.66$	$f_{19} = -1464.6$	$f_{20} = 0.25$
$f_{21} = 4587.4$	$f_{22} = -0.09$	$f_{23} = 2935.0$
$f_{24} = -0.02$		

In conclusion, $I_1^{10}\frac{f'}{f} = V(1) - V(10) = 13 - 11 = 2$ and so our polynomial has 2 real roots on this interval.

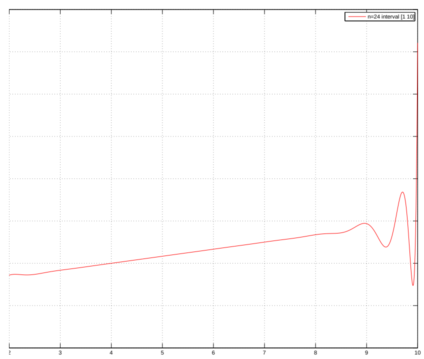

Figure 1: A polynomial of degree $n = 24$

From Figure 1 we can see that our polynomial has indeed two real roots in the interval $(1,\ 10)$, in the neighborhood of $x = 1$.

For the same polynomial, represented using shifted Chebyshev basis on the interval $(a,\ b)$, we calculated the zeros of this polynomial using comrade matrix approach and we obtain explicitly the following two roots: 1.02445 and 1.06061.

4 An efficient method for constructing a feasible cut

In the introduction we have described the main steps of a cutting plane method (Algorithm 1) for solving convex problems. As we have seen, the main difficulty in application of the cutting plane method is to derive an efficient method for constructing the feasible cut. In this section we provide such a method in the case when we solve convex problems over the cone of nonnegative polynomials of the form (1). Our method combines bisection with the Sturm theorem (see Theorem 6) in order to find a scalar value for which the polynomial takes a negative value. Once this scalar is determined, we can derive a feasible hyperplane in an straightforward manner. We also discuss the complexity of our method.

4.1 Checking nonnegativity of a polynomial

In the sequel we present an algorithm for verifying the nonnegativity of a given polynomial $f(x)$[1]. The output of this algorithm is either a scalar value α for which the polynomial value is negative or our polynomial is nonnegative.

Algorithm 2

1. Given a polynomial $f(x)$ written in the shifted Chebyshev basis on the interval (a, b) as $f(x) = \sum_{k=0}^{n} c_k T_k(x)$, where $c_0 \cdots c_n$ are real coefficients.

2. Find $N = V(a) - V(b)$ the number of distinct real roots of $f(x)$ in (a, b), using Theorem 6.

- **2.1** if $N = 0$ then our polynomial is nonnegative and stop.
- **2.2** if $N = n$ then go to step 4.
- **2.3** else use Theorem 6 for the perturbed polynomial $f(x) + \varepsilon$ and find $N_\varepsilon = V_\varepsilon(a) - V_\varepsilon(b)$ (where $\varepsilon > 0$ is sufficiently small).
  - **2.3.1** if $N_\varepsilon = 0$ then $f(x)$ is nonnegative but has multiple roots and stop.
  - **2.3.2** if $N_\varepsilon = N$ go to step 4.
  - **2.3.3** else (i.e. $0 < N_\varepsilon < N$) go to step 4 for $f(x) = f_\varepsilon(x)$.

3. if this interval contains more than one root, we divide it into two subintervals $\left(a, \dfrac{(a+b)}{2}\right]$ and $\left(\dfrac{(a+b)}{2}, b\right]$. We evaluate the given polynomial in $\dfrac{(a+b)}{2}$. If $f\left(\dfrac{(a+b)}{2}\right) < 0$, then $\alpha = \dfrac{(a+b)}{2}$. We interrupt this process and we construct the feasibility cut using $\alpha = \dfrac{(a+b)}{2}$.

4. Obtain an interval $(a_i, b_i]$.

- **4.1** if the interval has more than one root we evaluate $f(x)$ in $\dfrac{a_i + b_i}{2}$ and proceed as in step 3.
- **4.2** if the interval has one root we will verify the function values in a_i and b_i. If the function values in a_i and b_i are positive, then we will discard this interval. Otherwise α is either a_i or b_i.
- **4.3** else continue this process discarding intervals that contain no roots.

If we did not receive a negative function value during this algorithm, then it means that we have a nonnegative polynomial. Note that our Algorithm 2 is different from the bisection method for isolation of the roots of a polynomial since we interrupt the iteration at the first time when an α, for which our polynomial has a negative value, is found.

4.2 Feasibility/Optimality cuts for the cone of nonnegative polynomials

We assume that the convex optimization problem (1) has the following reformulation when using shifted Chebyshev basis:

$$\begin{aligned} \min_{f} \ & G_s(p) \\ \text{s.t. } & A_s p \le b_s \\ & p \in \mathcal{K}_s, \end{aligned} \tag{6}$$

where $p = [c_0 \cdots c_n]^T$ and

$$\mathcal{K}_s = \{p \in \mathcal{R}^{n+1} : c_0 T_0(x) + c_1 T_1(x) + \cdots + c_n T_n(x) \ge 0, \ \forall x \in (a, b)\}.$$

[1]Note that any polynomial written in the classical basis can be also represented using shifted Chebyshev basis for a given interval.

Note that optimization problem (6) is also convex since the relation between the coefficients of a polynomial in the classical basis and shifted Chebyshev basis is linear, i.e. if $f(x)$ can be represented as $f(x) = f_0 + f_1 x + \cdots + f_n x^n$ in the classical basis and as $f(x) = c_0 T_0(x) + c_1 T_1(x) + \cdots + c_n T_n(x)$ in the shifted Chebyshev basis, then there exists an invertible matrix B such that $Bp = f$.

Since the intersection of a polytope described by $A_s p \leq b_s$ with the convex cone $\mathcal{K}_s$ is a convex set, an oracle can be written for the convex set

$$Q_s = \left\{p \in \mathcal{R}^{n+1} : p \in \mathcal{K}_s, \;\; A_s p \leq b_s\right\}.$$

The steps 2.1 and 2.2 in the cutting plane method (Algorithm 1) for our problem (1) can be derived using the following reasoning: given $p^k = [c_0 \cdots c_n]^T$, if during the bisection process explained in Algorithm 2 we are able to find an interval $(a_i,\ b_i) \subseteq (a,\ b)$ that contains a scalar α for which

$$p^k(\alpha) = \langle p^k, \pi(\alpha)\rangle < 0,$$

where

$$\pi(\alpha) = [T_0(\alpha)\, T_1(\alpha) \cdots T_n(\alpha)]^T,$$

then for any $p \in \mathcal{K}_s$ define

$$\begin{aligned} 0 \leq & \min_{x\in(a,b)} p(x) = \\ & \min_{x\in(a,b)} [p(x) - p^k(\alpha) + p^k(\alpha)] \leq \\ & \min_{x\in(a,b)} [p(x) - p^k(\alpha)] + p^k(\alpha) \leq p(\alpha) - p^k(\alpha). \end{aligned}$$

In conclusion the following cuts can be constructed:

2.1 **Feasibility cut**: if $p^k \notin (Q_s \cap \mathcal{K}_s)$, then

$$\exists\, \alpha \text{ s.t. } \langle p^k, \pi(\alpha)\rangle < 0,$$

In this case we construct the following feasible hyperplane: for some $\delta \geq 0$ sufficiently small

$$\delta + \langle p^k, \pi(\alpha)\rangle \leq \langle p, \pi(\alpha)\rangle.$$

2.2 **Optimality cut**: if $p^k \in Q_s$, then the following hyperplane is constructed

$$G_s(p) \geq G_s(p^k) + \langle \nabla G_s(p^k), p - p^k\rangle.$$

There are several ways of enforcing the linear constrains $A_s p \leq b_s$. The simplest one is to explicitly take it into account in the query point generator.

4.3 Complexity

In this section we derive the complexity of Algorithm 2. First, the following lemma is an immediate consequence of our discussion from Section 3.1:

Lemma 7 *There is an $\mathcal{O}(n^2)$ time algorithm for computing the Sturm sequence in the Chebyshev basis.*

From the previous lemma we can see that for a given polynomial $f(x)$, the computational complexity of the number of changes of sign in the corresponding Sturm sequence is of order $\mathcal{O}(n^2)$. It remains to derive the complexity of the bisection method. However, given an irreducible polynomial with integer coefficients $f(x) = f_0 + f_1 x + \cdots + f_n x^n$, with the roots $x_1, \cdots, x_n$, the *minimum root separation* is defined as $\text{sep}(f) = \min_{i\neq j} |x_i - x_j|$. There is a well-known bound for the minimum root separation of such a polynomial [16]:

$$\text{sep}(f) > 2n^{-n/2-1/2}\|f\|^{-n+1},$$

where $\|f\| = \max_i |f_i|$ is the max norm of the polynomial. Note that such a bound can be extended easily to polynomials with rational coefficients as well. Moreover, the data in computer are represented in rational form. Since for a given interval $(a,\ b)$ and a minimum root separation $\text{sep}(f)$ for a polynomial $f(x)$, the complexity of the bisection method is $\log(\frac{b-a}{\text{sep}(f)})$, we obtain the following complexity for bisection in terms of the degree n of the polynomial:

$$\mathcal{O}(n \log n + n \log \|f\|).$$

Using a bisection search of $\mathcal{O}(n \log n)$ stages in Algorithm 2 (assuming $\|f\| \leq n$) we arrive at the following complexity per iteration for our method:

$$\mathcal{O}(n^3 \log n).$$

However, our numerical results show only a complexity of order $\mathcal{O}(n^2)$.

5 Application

In this section we consider the interval $(a,\ b) = (1,\ 10)$ and $T_k(x)$ are shifted Chebyshev polynomials of order k on this interval. Given scalars $\alpha_0, \cdots, \alpha_m \in (1,\ 10)$, let us define the vectors $a_i = [T_0(\alpha_i) \cdots T_n(\alpha_i)]^T$ for $i = 0, \cdots, m$.

We consider the following optimization problem

$$\min_{f} c_0^T p$$
$$\text{s.t. } a_i^T p = b_i \quad \forall i = 1, \cdots, m$$
$$p \in \mathcal{K}_{\mathcal{R}},$$

n	m	nr. of bisection iter.	outer iter.
22	10	2	46
52	30	6	141
72	60	5	90
102	40	12	128
222	80	25	251

The table displays the average number of iterations for the bisection and the number of iterations of the cutting plane method for different values of n (dimension of the polynomial) and m (number of interpolation constraints).

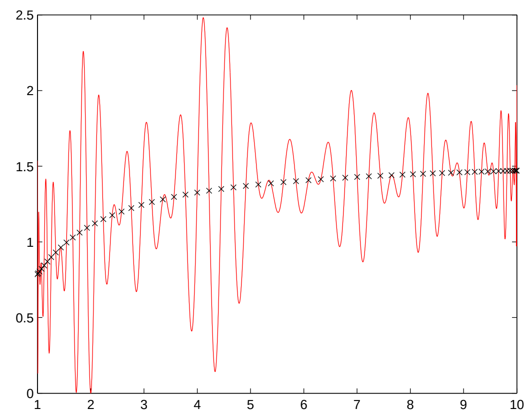

Figure 2: The optimal polynomial satisfying the given interpolation constraints.

Figure 2 corresponds to an optimization problem with the following data: dimension of the polynomial is $n = 72$, the number of interpolation constraints is $m = 60$ and the accuracy of solution is $\varepsilon = 10^{-4}$. Moreover $\alpha_0 = 2$ and $\alpha_i = \dfrac{(a+b)}{2} + \dfrac{(b-a)}{2} \cos\left(\dfrac{(2i-1)\pi}{2m}\right)$, $b_i = \arctan(\alpha_i)$ for all $i = 1, \ldots, m$. The software[2] OBOE was used to solve the above problem.

[2]https://projects.coin-or.org/OBOE

6 Conclusions

We developed a cutting plane method for solving convex optimization problems over the cone of nonnegative polynomials with linear inequalities. The main difficulty in using a cutting plane method for solving this type of problems consists in providing an efficient method for checking nonnegativity of a polynomial. In this paper an efficient method was devised to determine a scalar value for which the polynomial is negative and in the case that such a value exists a feasible cut can be constructed. Our method is based on Sturm theorem, which allows to determine the number of distinct real roots of a polynomial on a given interval, in combination with a bisection method. For numerical stability we constructed the associated Sturm sequence using Chebyshev polynomials although other orthogonal bases could be also used such as Hermite polynomials, Laguerre polynomials, Legendre polynomials, etc. It turns out that the complexity of our method is of the order $O\left(n^3 \ln n\right)$. However, in practice we observed an $O\left(n^2\right)$ complexity. Numerical results show also the efficiency of our new approach.

Note that there are other options besides Sturm theorem to count the number of distinct real roots of a polynomial in a given interval such as methods based on Descartes rule of signs [1], Sylvester theorem [13], etc. In a future work we intend to investigate in more detail these methods. Furthermore, it will be interested to study the effect of the accuracy on the minimum root separation for the iterates.

Acknowledgment. The authors thank Prof. Y. Nesterov and Prof. M. van Barel for the many inspiring discussions on this topic.

References:

[1] G. Alkiviadis and G. Collins, *Polynomial real root isolation using Descartes rule of signs*, In the Proceedings of the 1976 ACM Symposium on Symbolic and Algebraic Computation, pp. 275-285 (1976).

[2] J. Maso and D. Handscomb, *Chebyshev polynomials*, Chapman and Hall, 2003.

[3] F. Gantmacher, *The theory of matrices*, Chelsea Publishing Company, 1974.

[4] A. Ralston and P. Rabinowotz, *A first course in numerical analysis*, McGraw, 1982.

[5] Y. Nesterov, *Introductory lectures on convex optimization. A Basic Course*, Kluwer Academic Publishers, 2003.

[6] A. Nemirovskii and D. Yudin, *Problem complexity and method efficiency in optimization*, Wiley-Interscience Series in Discrete Mathematics, John Wiley and Sons Inc., New Yourk, 1983.

[7] J. Goffin, A. Huarie and J. Vial, *Decomposition and nondifferentiable optimization with projective algorithm*, Management Science, 38, pp. 284–302, 1992.

[8] Y. Hachez, *Convex optimization over non-negative polynomials: structured algorithms and applications,* PhD thesis, Universite Catholique de Louvain, Belgium, 2003.

[9] L. Vandenberghe and S. Boyd, *Semidefinite programming,* SIAM Review, 38, pp. 49-95, 1996.

[10] Y. Genin, Y. Hachez, Y. Nesterov and P. Van Dooren, *Convex optimization over positive polynomials and filter design,* In the Proceedings of the International Conference on Control, Cambridge, 2000.

[11] K. Zhou, J. Doyle and K. Glover, *Robust and optimal control*, Prince Hall, 1996.

[12] K. Toh, *Some new search directions for primal-dual methods in semidefinite programming*, SIAM Journal of Optimization, 11, pp. 223-242, 2000.

[13] V. Prasolov, *Polynomials*, Algorithms and Computation in Mathematics, Springer, 2001.

[14] J. Reif, *An $\mathcal{O}(nlog^3 n)$ algorithm for the real root problem*, In the Proceedings of the 34th IEEE Conference on Foundations of Computer Science, Palo Alto, 1993.

[15] P. Parrilo, *Structured Semidefinite Programs and Semialgebraic Geometry: Methods in Robustness and Optimization*, PhD thesis, California Institute of Technology, USA, 2000.

[16] G. Collins, *Polynomial minimum root separation*, Journal of Symbolic Computation, 32, pp. 467–473, 2001.

[17] Bini, D. A. and Eidelman, Y. and Gemignani, L. and Gohberg, I. C. *Fast QR eigenvalue algorithms for Hessenberg matrices which are rank-one perturbations of unitary matrices*, SIAM Journal on Matrix Analysis and its Applications, 29(2), pp. 566-585, 2007

[18] Bini, D. A. and Gemignani, L. and Pan, V. Y. *Fast and stable QR eigenvalue algorithms for generalized companion matrices and secular equations.* Numerische Mathematik, 100(3), pp. 373-408, 2005.