Positivity Of Trigonometric Polynomials

Alexandre MEGRETSKI
35-418, LIDS, EECS, MIT
Cambridge MA 02139
ameg@mit.edu

Abstract

The paper introduces a modification of the well-known sum-of-squares relaxation scheme for semi-algebraic programming by Shor [1], based on replacing the ordinary polynomials by their trigonometric counterparts. It is shown that the new scheme has certain theoretical advantages over the classical one: in particular, a trigonometric polynomial is positive if and only if it can be represented as a sum of squares of a finite number of trigonometric polynomials. A dual version of the SOS relaxation is also introduced and discussed. An example of a quantized finite horizon optimal control application with state constraints, typical for model predictive control, is discussed.

Key Words: non-convex optimization, semidefinite relaxations, model predictive control.

1 Introduction

A number of problems in system analysis and design are closely related to the task of checking positivity of polynomials. A real-valued polynomial of several real variables is positive if it can be represented as a sum of squares of polynomials. Shor [1] proposed using this simple idea to generate a set of conservative but relatively cheap tests of polynomial positivity, with computations relying on semidefinite programming. Parrilo [2] argued for extensive use of the "sum of squares" (SOS) idea in solving problems of system analysis and design.

Theoretically, the SOS analysis suffers from a serious flaw: a positive polynomial of many variables is *not* necessarily a finite sum of squares of polynomials. This, at least in principle, forces the user to consider representations of a given polynomial as a sum of *ratios* of squares of polynomials, which is known [3] to be necessary and sufficient for positivity. However, using the rational SOS has is own disadvantages, especially when solving the problem of *finding* a positive element in a given subset of polynomials.

The objective of this paper is to point out that a minor modification of the polynomial SOS argument removes this particular difficulty (and, potentially, provides some other benefits). The alteration involves replacing ordinary polynomials with *trigonometric* polynomials, i.e. with polynomials with respect to cosins and sins of a given set of real arguments. One can easily show that a positive trigonometric polynomial is *always* a finite sum of squares of trigonometric polynomials. A similar statement for positivity on the set of common zeros of given trigonometric polynomials also follows from the basic theorems of functional analysis.

2 Trigonometric Sums of Squares

Let $\mathbf{T}$ be unit circle $\mathbf{T} = \{z \in \mathbf{C} : \ |z| = 1\}$ in the complex plane. Let $\mathbf{Z}$ denote the subset of integers in $\mathbf{R}$. A *trigonometric polynomial* on $\mathbf{T}^n$ is a function $F : \ \mathbf{T}^n \to \mathbf{C}$ defined by

$$F(z) = \sum_{k \in S} f(k) z^k, \tag{2.1}$$

where S is a given *finite* subset of $\mathbf{Z}^n$, $f : \ \mathbf{Z}^n \to \mathbf{C}$ is a given function with a *finite* set of non-zero values, and z^k is defined by

$$z^k = z_1^{k_1} z_2^{k_2} \cdot \dots \cdot z_n^{k_n} \text{ for } k = (k_1, \dots, k_n) \in \mathbf{Z}^n, \ z = (z_1, \dots, z_n) \in \mathbf{T}^n.$$

When $z_i \in \mathbf{T}$ are parameterized by $z_i = \exp(jt_i)$, $F(z)$ becomes an ordinary polynomial with respect to $\cos(t_i)$ and $\sin(t_i)$, hence the name "trigonometric". In this paper, $\mathcal{P}_n$ refers to the set of all trigonometric polynomials on $\mathbf{T}^n$, while $\mathcal{P}_n^m$ denotes the set of m-vectors with elements from $\mathcal{P}_n = \mathcal{P}_n^1$.

It is easy to see that a trigonometric polynomial $F = F(z)$ defined by (2.1) is real-valued if and only if $f(-k) = \bar{f}(k)$ for all $k \in \mathbf{Z}^n$. The sets of real-valued trigonometric polynomials on $\mathbf{T}^n$ will be denoted by $\mathcal{RP}_n$, with a similar definition for its vector version $\mathcal{RP}_n^m$.

One objective of this paper is to analyze an SOS relaxation scheme for optimization problem "X" defined as follows: given an "explicitly defined" finite dimensional convex subset $\Omega \in \mathcal{RP}^n$ (e.g. a linear subspace, a polytope, a disc, etc.), and a vector trigonometric polynomial $Q \in \mathcal{RP}_n^m$, find an element $F \in \Omega$ which has the largest minimum over the zero set $Z_Q = \{z : Q(z) = 0\}$ of Q. The *SOS relaxed problem*, following the general philosophy of [1, 2], is defined as that of maximizing y subject to the equality constraint

$$F(z) = y + B(z)'HQ(z) + V(z)'WV(z) \quad \forall\ z \in \mathbf{T}^n, \tag{2.2}$$

where the decision parameters, in addition to y and the coefficients of F, are the elemens of matrices H and W, with W being constrained by $W = W' \geq 0$. In (2.2), B is a fixed vector of real-valued trigonometric polynomials (a natural suggestion would be to use the "monomials" $\mathrm{Re}(z^k)$ and $\mathrm{Im}(z^k)$ as elements of B). Similarly, V is a fixed vector with trigonometric polynomial elements which are allowed to be complex-valued (and hence the "true" monomials z^k can be used in constructing V.

The following mathematical result gives some justification to the trigonometric SOS relaxation scheme described above.

Theorem 2.1 *Let $F, Q_1, \dots, Q_m \in \mathcal{RP}_n$ be such that $F(z) > 0$ for all z satisfying $Q_1(z) = Q_2(z) = \dots = Q_m(z) = 0$. Then there exist $H_1, \dots, H_m \in \mathcal{RP}_n$, and $V_1, \dots, V_r \in \mathcal{P}_n$ such that*

$$F(z) = \sum_{i=1}^{m} H_i(z)Q_i(z) + \sum_{k=1}^{r} |V_k(z)|^2 \ \forall\ z \in \mathbf{T}^n. \tag{2.3}$$

Theorem 2.1 with $m = 0$ (no equality constraints) can be found in the functional analysis textbooks (see, for example, [4, 5]). The more general case can be proven similarly. For completeness, and since the proof is a useful demonstration of the use of fundamental theorems of functional analysis, it is given in the Appendix.

The proof of Theorem 2.1 suggests a different (in some sense, "dual") relaxation scheme for the case $m = 0$, described by the following statement.

Theorem 2.2 *Let $F \in \mathcal{RP}_n$ be given by (2.1). Let $h_1, h_2, \dots$ be some enumeration of the elements of $\mathbf{Z}^n$, i.e. for every $k \in \mathbf{Z}^n$ there exists i such that $h_i = k$, and $h_i \neq h_s$ when $i \neq s$. For $r = 1, 2, \dots$ let W^r be the r-by-r Hermitian matrix with entries $W^r_{i,s} = f(k_i - k_s)$. Then the (non-increasing)* *sequence of the smallest eigenvalues of* W^r *converges to the minimum of* F *on* $\mathbf{T}^n$ *as* $r \to \infty$

Theorem 2.2 can be used directly to obtain another relaxation scheme for problem "X" with $m = 0$ as the problem of maximizing (over F) the minimal eigenvalue of $W^r = W^r(F)$ for some fixed $r, (h_i)_{i=1}^r$. When equality constraints are present, the theorem can be applied to the modified problem, with F replaced by $F_*(z) = F(z) + R|H(z)|^2$ where $R > 0$ is sufficiently large.

3 Example

Consider the task of finite horizon quantized optimal control of a (discrete time) double integrator, defined by the linear equations

$$x(t+1) = Ax(t) + Bu(t),\ y(t) = Cx(t),\ v(t) = Lx(t),\ x(0) = x_0, \quad (3.4)$$

the constraints

$$u(t) \in \{-1, 1\},\ |v(t)| \leq M,\ (t = 0, 1, \dots, T), \quad (3.5)$$

and the objective

$$J = \sum_{t=1}^{T} y(t)^2 \to \min, \quad (3.6)$$

where

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},\ B = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\ C = \begin{bmatrix} 1 & 0 \end{bmatrix},\ L = \begin{bmatrix} 0 & 1 \end{bmatrix},$$

x_0, M, T are known parameters, and $u(t)$ is to be optimized on the time interval $0 \leq t < T$. This simplified academic example is styled after a common model predictive control setup.

The SOS relaxation introduced above can be used to compute lower bounds of the cost functional, while the smallest eigenvalue relaxations of Theorem 2.2 can provide upper bounds for its minimum.

To reduce the problem to equality-constrained minimization of trigonometric polynomials, re-parameterize $u(t)$ by

$$u(t) = \frac{z_{t+1} + z_{t+1}^{-1}}{2}, \quad t = 0, 1, \dots, T-1,$$

where the complex variables $z_i \in \mathbf{T}$ are constrained by $z_i^2 = 1$ for $i = 1, 2, \ldots, T$. This parameterization fixes the required range of $u(t)$. Similarly, the other constrained variable $v(t) = Lx(t)$ should be re-parameterized by

$$v(t) = \frac{M}{2}(z_{t+T} + z_{t+T}^{-1}), \quad t = 1, 2, \ldots, T,$$

where the complex variables $z_i \in \mathbf{T}$ are otherwise unconstrained for $i = T+1, T+2, \ldots, 2T$. Finally, equalities $v(t) = v(t-1) + u(t-1)$, where $t = 1, 2, \ldots, T$, provide the additional constraints relating z_t, z_{t+T-1}, and z_{t+T} (except for the case $t = 0$, when $v_0 = Lx_0$ is not a decision variable but a known constant). Equalities $y(t) = y(t-1) + v(t-1)$ allow one to define the cost as a quadratic functional with respect to $v(t)$, which makes J a trigonometric polynomial of degree 2 with respect to $z = (z_i)$.

Consider the particular relaxation (2.2), where H_k, V_k are limited to first degree trigonometric polynomials, i.e.

$$V(z) = B(z) = \begin{bmatrix} 1 \\ \frac{\xi+\bar{\xi}}{2} \\ \frac{\xi-\bar{\xi}}{2j} \\ \frac{\eta+\bar{\eta}}{2} \\ \frac{\eta-\bar{\eta}}{2j} \end{bmatrix}, \text{ where } \eta = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_T \end{bmatrix}, \quad \xi = \begin{bmatrix} z_{T+1} \\ z_{T+2} \\ \vdots \\ z_{2T} \end{bmatrix}.$$

There are certain symmetries which allow one to reduce dramatically the number of decision variables in the resulting semidefinite program. First, since $\eta = \bar{\eta}$ for all z, the components related to $(\eta - \bar{\eta})/2j$ can be ignored. Second, $F(z)$ and $Q_i(z)$, as polynomials with respect to $(\xi + \bar{\xi})/2$ and $(\eta + \bar{\eta})/2$, do not depend on $(\xi - \bar{\xi})/2j$. This suggests that the vector variable $(\xi - \bar{\xi})/2j$ can be removed as well. However, since its components enter into the a-priori assumed equality constraint

$$|(\xi_i + \bar{\xi}_i)/2|^2 + |(\xi_i - \bar{\xi}_i)/2j|^2 = 1,$$

the term

$$\frac{1}{4}(\xi + \bar{\xi})' D_\xi (\xi + \bar{\xi}) - \text{trace}(D)$$

, where D is an arbitrary diagonal T-by-T matrix with non-negative coefficients, should be added to $F(z)$ instead.

The resulting relaxed semidefinite program will be that of maximizing y subject to the single matrix inequality

$$\left[\begin{array}{ccc} a^2 - \mathrm{trace}(D_\xi + D_\eta) - y & Mae'E & 0 \\ MaE'e & M^2E'E + D_\xi & 0 \\ 0 & 0 & D_\eta \end{array}\right] + \sum_{t=1}^{T}(G_t'H_t + H_t'G_t) \geq 0,$$

where D_ξ, D_η, H_t and y are the decision parameters (D_ξ diagonal and non-negative, D_η arbitrary diagonal, H_t arbitrary real row vectors), $a = v_0 + y_0$, and

$$\left[\begin{array}{c} G_1 \\ G_2 \\ \vdots \\ G_T \end{array}\right] = \left[\begin{array}{cccccccccccc} v_0 & -M & 0 & 0 & \dots & & 0 & 1 & 0 & \dots & 0 \\ 0 & M & -M & 0 & \dots & & 0 & 0 & 1 & \dots & 0 \\ & & & & & \dots & & & & & \\ 0 & 0 & & & \dots & M & -M & 0 & \dots & & 1 \end{array}\right].$$

4 Appendix

In this Appendix, proofs of the main theorems of this paper are given. We will need two fundamental results of functional analysis: the Hahn-Banach theorem and the Bochner's theorem.

Let X be a real vector space (possibly infinite dimensional, no topological properties assumed). A vector $z \in X$ is said to be an *interior point* of a subset $Y \subset X$ if for every $\delta \in X$ there exist $\epsilon > 0$ such that $z + t\delta \in Y$ for all $t \in [0, \epsilon]$. The Hahn-Banach theorem can be formulated as follows.

Theorem 4.3 *Let Y be a convex subset of a vector space X. Assume that Y has an interior point, but does not contain 0. Then there exists a non-zero linear functional $L: X \to \mathbf{R}$ such that $L(y) \geq 0$ for all $y \in Y$.*

In other words, a "fat" convex set not containing zero lies on one side of some hyperplane containing zero. Note that, in the infinite dimensional case, the condition of existence of an interior point cannot be dropped.

A complex-valued function $g: \mathbf{Z}^n \to \mathbf{C}$ is called *positive definite* if the matrix $(g(h_i - h_s))_{i,j=1}^{r}$ is Hermitian positive semidefinite for all finite subsets $h_1, \dots, h_r$ of Z^n. The Bochner's Theorem can be formulated as follows.

Theorem 4.4 *A function $g:\ \mathbf{Z}^n \to \mathbf{C}$ is positive semidefinite if and only if there exists a Borel measure $\mu \geq 0$ on T^n such that*

$$g(k) = \int_{\mathbf{T}^n} z^k d\mu(z) \quad \forall\ k \in \mathbf{Z}^n. \tag{4.7}$$

To prove Theorem 2.1 assume that F cannot be represented in the form (2.3). Then $0 \in \mathcal{RP}_n$ is not an element of the set $\mathcal{D} \subset \mathcal{RP}_n$ of all real trigonometric polynomials Δ representable in the form

$$\Delta = -F + \sum_{i=1}^{m} H_i(z)Q_i(z) + \sum_{k=1}^{r} |V_k(z)|^2,$$

where r is not fixed.

The next step is to apply the Hahn-Banach Theorem in the case when $X = \mathcal{RP}_n$ and $Y = \mathcal{D}$. By construction, $\mathcal{RP}_n$ is a real vector space, and $\mathcal{D}$ is a convex subset of $\mathcal{RP}_n$. By assumption $0 \notin \mathcal{D}$. Let us show that $1 - F(z)$ is an interior point of $\mathcal{D}$. Indeed, for every

$$C(z) = \sum_k c(k)z^k \in \mathcal{RP}_n \tag{4.8}$$

we have

$$1 - F(z) + tC(z) = -F(z) + 1 - \frac{t}{2}\sum_k (1 + |c(k)|^2) + \frac{t}{2}\sum_k |1 + c(k)z^k|^2,$$

which is in $\mathcal{D}$ when t is close enough to zero. Therefore, by Theorem 4.3 there exists a non-zero linear functional $L:\ \mathcal{RP}_n \to \mathbf{R}$ such that $L(\Delta) \geq 0$ for all $\Delta \in \mathcal{D}$.

Define $g:\ \mathbf{Z}^n \to \mathbf{C}$ by

$$g(k) = \frac{1}{2}(L(z^k + z^{-k}) - jL(jz^k - jz^{-k})).$$

Then

$$L(C) = \sum_k g(k)c(k)$$

for every C defined by (4.8). The next step is to show that g is positive definite. Indeed, for a given finite subset $\{h_1, \ldots, h_r\} \in \mathbf{Z}^n$, for every $t \in \mathbf{R}$

and for every set of complex numbers $u_1, \ldots, u_r$ the trigonometric polynomial $-F + t^2 C$, where

$$C(z) = \left| \sum_{i=1}^{r} u_i z^{h_i} \right|^2 = \sum_{i,s=1}^{r} u_s' u_i z^{h_i - h_s},$$

belongs to $\mathcal{D}$. Hence $-L(F) + t^2 L(C) \geq 0$ for all $t \in \mathbf{R}$, which implies

$$L(C) = \sum_{i,s=1}^{r} u_s' u_i g(h_i - h_s) \geq 0.$$

Hence g is positive definite, and by Theorem 4.4 representation (4.7) holds for some measure $\mu \geq 0$. Note that $\mu \neq 0$ because otherwise $L = 0$.

To finish the proof of Theorem 2.1 note that (4.7) implies that

$$L(C) = \int_{\mathbf{T}^n} C(z) d\mu(z).$$

Since for every $H \in \mathcal{RP}_n$, $t \in \mathbf{R}$ we have $tHQ_i - F \in \mathcal{D}$ we conclude that $L(HQ_i) = 0$ for all $H \in \mathcal{RP}_n$. Hence

$$\int_{\mathbf{T}^n} H(z) Q_i(z)(z) d\mu(z) = 0 \quad \forall \; H \in \mathcal{RP}_n,$$

which means that $Q_i(z) d\mu(z) = 0$ for every i, i.e. that the measure μ is concentrated on the set Z_Q of common zeros of Q_i. Finally, $-F \in \mathcal{D}$, which should imply $L(F) \leq 0$. However, since F is assumed to be strictly positive on Z_Q,

$$L(F) = \int F(z) d\mu(z)$$

must be positive. The contradiction proves representation (2.3).

In fact, the same proof can be applied to the case when $\mathbf{T}^n$ is replaced by an arbitrary commutative compact group G, such as $\{-1,1\}^n$, $\mathbf{T}^a \times \{-1,1\}^b \times \{1,-1,j,-j\}^c$, etc., the polynomials on G being defined as finite linear combinations of group homomorphisms $G \to \mathbf{T}$. However, since every commutative compact group is isomorphic to a subgroup of $\mathbf{T}^n$ defined by simple polynomial equations of the form $z_i^{q_i} = 1$, there appears to be little practical benefit in doing that.

References

[1] N.Z. Shor. Class of global minimum bounds of polynomial functions. *Cybernetics*, 23(6): 731-734, 1987. (Russian orig.: Kibernetika, No. 6(1987), 9-11).

[2] Pablo A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization California Institute of Technology* (Ph.D. Thesis), Pasadena, CA, May 2000.

[3] J. Bochnak, M. Coste, and M-F. Roy. *Real Algebraic Geometry.* Springer, 1998.

[4] W. Rudin, *Fourier analysis on groups*, Interscience Publishers, 1962, 285 p.

[5] W. Rudin, *Function theory in polydiscs*, W. A. Benjamin, 1969, 188 p.