Journal of Global Optimization **26:** 419–433, 2003.

Multivariate Polynomial Minimization and Its Application in Signal Processing

LIQUN QI and KOK LAY TEO*
Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong.
(e-mail: maqilq@polyu.edu.hk; mateokl@polyu.edu.hk)

Abstract. We make a conjecture that the number of isolated local minimum points of a $2n$-degree or $(2n+1)$-degree r-variable polynomial is not greater than n^r when $n \leqslant 2$. We show that this conjecture is the minimal estimate, and is true in several cases. In particular, we show that a cubic polynomial of r variables may have at most one local minimum point though it may have 2^r critical points. We then study the global minimization problem of an even-degree multivariate polynomial whose leading order coefficient tensor is positive definite. We call such a multivariate polynomial a normal multivariate polynomial. By giving a one-variable polynomial majored below a normal multivariate polynomial, we show the existence of a global minimum of a normal multivariate polynomial, and give an upper bound of the norm of the global minimum and a lower bound of the global minimization value. We show that the quartic multivariate polynomial arising from broad-band antenna array signal processing, is a normal polynomial, and give a computable upper bound of the norm of the global minimum and a computable lower bound of the global minimization value of this normal quartic multivariate polynomial. We give some sufficient and necessary conditions for an even order tensor to be positive definite. Several challenging questions remain open.

Key words: Multivariate Polynomial, Minimization, Signal Processing, Tensor, The Bézout Theorem

1. Introduction

The global equality constrained quadratic minimization problem arises from broad-band antenna array signal processing. The problem has a positive definite quadratic objective function, a large number of variables, a large number of linear equality constraints, and a few quadratic equality constraints each having very low rank quadratic constraint matrices [1, 3, 10, 11]. In [11], Thng, Cantoni and Leung proposed a Quadratic Constraint Structure (QCS) algorithm to solve this problem. They converted this problem into a *quartic* minimization problem either unconstrained as

$$f^* := \min f(x) \tag{1}$$

* The work of this author was supported by the Research Grant Council of Hong Kong.

or constrained as

$$\begin{aligned} f^{**} := \min \quad & f(x) \\ \text{subject to} \quad & h(x) = 0, \end{aligned} \tag{2}$$

where $x \in \Re^r$, $f : \Re^r \to \Re$ is a quartic polynomial of x, $h : \Re^r \to \Re^p$ is a system of linear or quadratic polynomials. According to the assumptions and (P_3)-A, (P_3)-B of [11], the quartic polynomial f can be written as

$$f(x) = g(x)^T G g(x) + c^T g(x) \tag{3}$$

by dropping the constant term, where $g : \Re^r \to \Re^m$ is a set of linear or quadratic polynomials, G is an $m \times m$ symmetric positive definite matrix, and $m > r$. The dimension r is very low. In a 5-tuple example of [11], $r = 1$. In a 70-tuple example of [11], $r = 2$. Here, 5 and 70 are the dimensions of the original linear and quadratically constrained quadratic minimization problem considered in [11]. In [11], the quartic minimization problem is converted into a system of cubic polynomial equations, by writing out the optimality condition of the quartic minimization problem. For (1), this system of cubic polynomial equations is as follows.

$$\nabla f(x) \equiv 2g(x)^T G \nabla g(x) + c^T \nabla g(x) = 0. \tag{4}$$

For (2), this system of cubic polynomial equations is as follows.

$$\begin{aligned} &\nabla f(x) + \sum_{i=1}^{p} u_i \nabla h_i(x) = 0, \\ &h(x) = 0. \end{aligned} \tag{5}$$

Then, the U-Resultant algorithm [8] is used to solve the system (4) or (5). The system (5) has r equations of $r + p$ variables, and p equations of r variables. We may treat (4) as a special case of (5) with $p = 0$. According to the Bézout theorem [12], the system (5) have at most $2^p \times 3^r$ zeros. Comparing the values of f at all the real zeros of the cubic polynomial system, the global minimum of (1) or (2) is found.

Motivated by the above practical problem from signal processing, in this paper, we study two issues of multivariate polynomial optimization.

The first issue is the number of local minimum points of a multivariate polynomial. This issue was studied by Durfee, Kronenfeld, Munson, Roy and Westby [2] for $r = 2$ in 1993. They pointed out that this issue is related to Hilbert's 16th problem on the arrangements of ovals of real algebraic curves. Since then, to the best of our knowledge, no further results appear in the literature. This shows that this is a tough issue.

We make a conjecture that the number of isolated local minimum points of a $2n$-degree or $(2n + 1)$-degree r-variable polynomial is not greater than n^r when $n \leqslant 2$. We show that this conjecture is the minimal estimate, and is true in several cases. In particular, we show that a cubic polynomial of r variables may have at most one local minimum point though it may have 2^r critical points.

We then study the global minimization problem of an even-degree multivariate polynomial whose leading order coefficient tensor is positive definite. We call such a multivariate polynomial a **normal multivariate polynomial**. We show that f defined by (3) is a normal multivariate quartic polynomial. By tensor analysis, we give a one-variable polynomial majored below a normal multivariate polynomial. With the help of such a majored below one-variable polynomial, we show the existence of a global minimum of the normal multivariate polynomial, and give an upper bound of the norm of the global minimum and a lower bound of the global minimization value. In the case of f defined by (3), we give a computable upper bound of the norm of the global minimum and a computable lower bound of the global minimization value. We give some sufficient and necessary conditions for an even order tensor to be positive definite.

Several challenging questions are presented in the final concluding section.

2. The Number of Local Minima of a Multivariate Polynomial

Our conjecture is below.

CONJECTURE 1. *A $2n$-degree or $2n+1$-degree polynomial of r variables has at most n^r isolated local minima when $n \leqslant 2$.*

Clearly, this conjecture is true for $r = 1$. This conjecture is also true when f is separable, i.e.,

$$f(x) = \sum_{i=1}^{r} f_i(x_i),$$

where f_i is a polynomial of x_i, and the degree of f_i is not higher than $2n+1$. Then f_i has at most n local minima. Since the x_i component of a local minimum of f must be a local minimum of f_i, the conjecture is true in this case. It also shows that this conjecture is the minimal estimate of this number, as we can construct a $2n$-degree $2n+1$-degree separable polynomial of r variables, which has n^r isolated local minima. This conjecture is also true for a quadratic polynomial, since it has at most one isolated critical point. We now show that this conjecture is true for a cubic multivariate polynomial, i.e., a cubic multivariate polynomial has at most one isolated local minimum. This is nontrivial as a cubic multivariate polynomial may have 2^r isolated critical points.

Let the degree of f be N, where $N = 2n$ or $2n + 1$.

THEOREM 1.
(i) *A $2n$-degree or $2n+1$-degree polynomial of r variables has at most n isolated local minima in a line.*

(ii) *A cubic polynomial of r variables may have* 2^r *critical points. However, it has at most and may have one local minimum and one local maximum.*

Proof.

(i) Let f be a $2n$-degree or $2n+1$-degree polynomial of r variables. In a line $a + th$, where $a, h \in \Re^r$, $t \in \Re$, define

$$\phi(t) = f(a + th).$$

Then an isolated local minimum of f in this line is an isolated local minimum of ϕ. Since ϕ has at most n isolated local minima, the claim (i) holds.

(ii) By (i), A cubic polynomial f of r variables has at most one local minimum in a line. This shows that it has at most one local minimum in the whole space. Consider $-f$. Then it has at most one local maximum. By the Bézout Theorem, it has at most 2^r critical points.

The following example shows that 2^r critical points are achievable. Let f be defined by

$$f(x) = \sum_{i=1}^{r} x_i^3 - 3x_i.$$

Clearly, f has 2^r critical points

$$x = (x_1, \cdots, x_r)^T,$$

where $x_i = 1$ or -1. However, f has only one local minimum

$$x = (1, \cdots, 1)^T,$$

and one local maximum

$$x = (-1, \cdots, -1)^T. \qquad \square$$

We have proved the conjecture is true for the following three cases:

(a). $N \leqslant 3$;

(b). $r = 1$;

(c). f is separable.

In 1993, Durfee, Kronenfeld, Munson, Roy and Westby [2] studied the case when $r = 2$. They prove that

$$N_{\max} + N_{\min} \leqslant \frac{1}{2}N^2 - N + 1,$$

where $N_{\max}$ is the number of local maxima, $N_{\min}$ is the number of local minima, N is the degree of the polynomial. When $N = 4$, their bound is 5. This is actually the true upper bound, i.e., a two variable quartic polynomial has at most 5 local maxima and local minima. An example is f defined by

$$f(x) = x_1^4 + x_2^4 - 2x_1^2 - 2x_2^2.$$

They also made a conjecture that if $N = 2n$ is even,

$$N_{\max} \leqslant \frac{3}{2}n(n-1) + 1.$$

If we take $N = 4$, i.e., $n = 2$, we have

$$N_{\max} \leqslant 4.$$

If we multiply the polynomial by -1, this implies that

$$N_{\min} \leqslant 4.$$

This is the same as our conjecture.

An interesting open question is: Can we prove that a quartic two variable polynomial has at most 4 local minima, or construct a quartic two variable polynomial having 5 local minima?

Durfee, Kronenfeld, Munson, Roy and Westby [2] also pointed out that this issue is related to Hilbert's 16th problem on the arrangements of ovals of real algebraic curves. By [$\geqslant$], conjecture 1 is not valid when $n \geqslant 3$.

3. Tensor Analysis

We use $A^{(n)}$ to denote an nth order tensor and use $A^{(n)}_{i_1 \cdots i_n}$ to denote its elements. We assume $i_k = 1, \cdots, r$ for $k = 1, \cdots, n$. We assume that $A^{(n)}$ is **totally symmetric**, i.e.,

$$A^{(n)}_{i_1 \cdots i_n} = A^{(n)}_{j_1 \cdots j_n}$$

if $\{i_1, \cdots, i_n\}$ is any reordering of $\{j_1, \cdots, j_n\}$. Let $x \in \Re^r$. Define

$$A^{(n)}x^n := \sum_{i_1, \cdots, i_n}^{r} A^{(n)}_{i_1 \cdots i_n} x_{i_1} \cdots x_{i_n}.$$

Let $\| \cdot \|$ be a norm in $\Re^r$. Define the induced norm of $A^{(n)}$ as

$$\|A^{(n)}\| := \max\{|A^{(n)}x^n| : \|x\| = 1\}. \tag{6}$$

Then the norm of an nth order totally symmetric tensor is well-defined. When $n = 1$, it is the dual vector norm of the norm $\| \cdot \|$. When $n = 2$ and $\| \cdot \|$ is the 2-norm, it reduces to the usual symmetric matrix norm and $\|A^{(2)}\|$ is the largest absolute value of the eigenvalue of the symmetric matrix $A^{(2)}$. For any $\alpha \in \Re$, we have

$$\|A^{(n)}(\alpha x)^n\| = |\alpha|^n \|A^{(n)}x^n\|. \tag{7}$$

We say that an even-order tensor $A^{(2n)}$ is **positive semidefinite** if

$$A^{(2n)}x^{2n} \geqslant 0$$

for all $x \in \Re^r$. We say that an even-order tensor $A^{(2n)}$ is **positive definite** if

$$A^{(2n)}x^{2n} > 0$$

for all $x \neq 0, x \in \Re^r$. For $n = 1$, these two definitions are the same as the definitions of positive semidefinite and positive definite symmetric matrices. For an even-order tensor $A^{(2n)}$, we define

$$[A^{(2n)}] := \min\{A^{(2n)}x^{2n} : \|x\| = 1\}. \tag{8}$$

Clearly, $A^{(2n)}$ is positive semidefinite if and only if $[A^{(2n)}] \geqslant 0$, is positive definite if and only if $[A^{(2n)}] > 0$. When $\|\cdot\|$ is the 2-norm, $[A^{(2)}]$ is the smallest eigenvalue of the symmetric matrix $A^{(2)}$.

Since $\|A^{(2)}\|$ and $[A^{(2)}]$ are the largest absolute value of the eigenvalue, and the smallest eigenvalue of the symmetric matrix $A^{(2)}$ respectively when $\|\cdot\|$ is the 2-norm, there are various computational methods to calculate them. A natural question is: whether similar facts exist for $\|A^{(n)}\|, n \geqslant 3$ and $[A^{(2n)}], n \geqslant 2$?

If there are m nth order tensors $B_1^{(n)}, \cdots, B_m^{(n)}$ such that for all $x \in \Re^r$,

$$A^{(2n)}x^{2n} = \sum_{i=1}^{m} \left(B_i^{(n)}x^n\right)^2, \tag{9}$$

then $A^{(2n)}$ is positive semidefinite. If furthermore the family $\{B_1^{(n)}, \cdots, B_m^{(n)}\}$ is **regular** in the sense that if $x \neq 0$ then at least one of $B_1^{(n)}x^n, \cdots, B_m^{(n)}x^n$ is non-zero, then $A^{(2n)}$ is positive definite. On the other hand, if $n = 1$ and $A^{(2n)}$ is positive definite, we know that there are $m = r$ vectors $B_1^{(1)}, \cdots, B_m^{(1)}$ such that they form a regular family and (9) holds. Later, we will see that this is true when $n = 2$, $A^{(4)}$ is the leading coefficient tensor of the quartic polynomial f defined by (3), and g in (3) satisfies a regularity condition. For a general even-order tensor $A^{(2n)}$, are the following two statements true? (i). $A^{(2n)}$ is positive semidefinite if and only if there are m nth order tensors $B_1^{(n)}, \cdots, B_m^{(n)}$ such that (9) holds; and (ii). $A^{(2n)}$ is positive definite if and only if there are m nth order tensors $B_1^{(n)}, \cdots, B_m^{(n)}$ such that they form a regular family and (9) holds. Hilbert has investigated the first question [9]. He showed that the answer is fully positive only for the following cases [9]:

(a). $n = 1$;

(b). $r = 2$;

(c). $n = 2, r = 3$.

In the next section, we discuss more on positive definite even-order tensors. Here we give a proposition on a property of a positive definite even-order tensor.

Let $E^{(n)}$ denote the nth order tensor whose elements satisfy

$$E^{(n)}_{i_1 \cdots i_n} = \begin{cases} 1, & \text{if } i_1 = \cdots = i_n, \\ 0, & \text{otherwise.} \end{cases}$$

Then $E^{(1)}$ is the vector whose elements are all 1's, while $E^{(2)}$ is the unit matrix.

PROPOSITION 1. *If a $2n$-order tensor $A^{(2n)}$ is positive definite, then there exists a positive constant α such that $A^{(2n)} - \alpha E^{(2n)}$ is still positive definite. In particular, this is true for any α satisfying*

$$0 < \alpha < [A^{(2n)}] \tag{10}$$

if $[A^{(2n)}]$ is defined by $2n$-norm in (8).

Proof. By (8), for any $x \in \Re^r$ with $\|x\| = 1$,

$$A^{(2n)}x^{2n} \geqslant [A^{(2n)}]\|x\|^{2n} > \alpha\|x\|_{2n}^{2n} = \alpha E^{(2n)}x^{2n}, \tag{11}$$

where α is an appropriate positive constant, and it can be any constant satisfying (10) if $[A^{(2n)}]$ is defined by $2n$-norm in (8). Now, for any $x \in \Re^r$, $x \neq 0$, let

$$\bar{x} = \frac{x}{\|x\|}.$$

Then $\|\bar{x}\| = 1$. By (11), we have

$$\left(A^{(2n)} - \alpha E^{(2n)}\right)x^{2n} = \|x\|^{2n}\left(A^{(2n)}\bar{x}^{2n} - \alpha E^{(2n)}\bar{x}^{2n}\right) > 0.$$

This shows that $A^{(2n)} - \alpha E^{(2n)}$ is positive definite. □

4. Normal Multivariate Polynomial

Let f be a $2n$-degree multivariate polynomial defined by

$$f(x) = \sum_{k=0}^{2n} A^{(2n-k)}x^{2n-k}, \tag{12}$$

where $A^{(2n-k)}$ is a $2n-k$ totally symmetric tensor. If $A^{(2n)}$ is positive definite, then we call f a **normal multivariate polynomial**.

Let f be a $2n$-degree normal multivariate polynomial as defined by (12). Define a one-variable polynomial ϕ by

$$\phi(\alpha) := [A^{(2n)}]\alpha^{2n} - \sum_{k=1}^{2n-1} \|A^{(2n-k)}\|\alpha^{2n-k} + A^{(0)}, \tag{13}$$

and call ϕ the **characterized polynomial** of f.

THEOREM 2. *Suppose that f is a $2n$-degree normal multivariate polynomial and ϕ is its characterized polynomial. Then f is bounded below by ϕ in the sense that for all $x \in \Re^r$, we have*

$$f(x) \geqslant \phi(\|x\|). \tag{14}$$

When $\|x\| \to \infty$, we have

$$f(x) \to \infty. \tag{15}$$

These imply that f has a global minimum. Moreover, if x^* is a global minimum of f, then it satisfies

$$\|x^*\| \leqslant L := \max\left\{1, \frac{\sum_{k=1}^{2n-1} \|A^{(2n-k)}\|}{[A^{(2n)}]}\right\}. \tag{16}$$

Proof. By (6)-(13), we have (14). Since

$$\phi(\alpha) \to \infty$$

as $\alpha \to \infty$, by (14), we have (15) as $\|x\| \to \infty$. Since ϕ is bounded below, f is also bounded below. These imply that f has a global minimum. When $\|x\| > L$,

$$f(x) \geqslant \phi(\|x\|) > A^{(0)} = f(0).$$

Hence, x cannot be a global minimum of f. This proves (16). □

The number L defined in (16) gives an upper bound of the norm of the global minimum of f. We may find a lower bound $\phi_{\min}$ for ϕ such that for all α,

$$\phi(\alpha) \geqslant \phi_{\min}.$$

Then, by (14), $\phi_{\min}$ is a lower bound of f^*.

In Section 5, we will give a practical way to compute an upper bound of the norm of the global minimum of f, and a lower bound of f^*, when f is defined by (3).

Let

$$f_0(x) = A^{(2n)}x^{2n},$$

$$y(i) = (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)^T$$

and

$$f_i(y(i)) = f_0(x_1, \cdots, x_{i-1}, 1, x_{i+1}, \cdots, x_r).$$

THEOREM 3. *Let $A^{(2n)}$ be a $2n$th order totally symmetric tensor, y and f_i be defined as above. Then the following statements are true.*

(i) *$A^{(2n)}$ is positive definite if and only if f_i are strictly positive polynomials for $i = 1, \cdots, r$. $A^{(2n)}$ is positive semidefinite if and only if f_i are nonnegative polynomials for $i = 1, \cdots, r$.*

(ii) *$A^{(2n)}x^{2n}$ can be written in the form of (9), if for $i = 1, \cdots, r$, f_i can be represented as a sum of squares of polynomials. $A^{(2n)}x^{2n}$ can be written in the form of (9) such that $B_1^{(1)}, \cdots, B_m^{(1)}$ form a regular family, if for $i = 1, \cdots, r$, f_i can be represented as a sum of squares of polynomials and a positive constant, i.e.,*

$$f_i(y(i)) = \sum_{j=1}^{m_i} \left(g_{ij}(y(i))\right)^2 + \alpha_i^2, \tag{17}$$

where g_{ij} are polynomials of $y(i)$, $\alpha_i \neq 0$.

(iii) *If $r = 2$, then $A^{(2n)}$ is positive semidefinite if and only if $A^{(2n)}x^{2n}$ can be represented by (9), and $A^{(2n)}$ is positive definite if and only if $A^{(2n)}x^{2n}$ can be represented by (9) such that $B_1^{(1)}, \cdots, B_m^{(1)}$ form a regular family.*

(iv) *If $r \geqslant 3$, then even if $A^{(2n)}$ is positive semidefinite, $A^{(2n)}x^{2n}$ may not be able to be represented by (9).*

Proof.

(i) Suppose that $i = 1, \cdots, r$. Let

$$x = (x_1, \cdots, x_{i-1}, 1, x_{i+1}, \cdots, x_r)^T.$$

If $A^{(2n)}$ is positive definite, then for any $y(i) \in \Re^{r-1}$,

$$f_i(y(i)) = f_0(x) = A^{(2n)}x^{2n} > 0,$$

i.e., f_i is a strictly positive polynomial.

On the other hand, suppose that f_i are strictly positive polynomials for $i = 1, \cdots, r$. Let $x \in \Re^r$, $x \neq 0$. Then there is an i such that $x_i \neq 0$. Let

$$y(i) = \left(\frac{x_1}{x_i}, \cdots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \cdots, \frac{x_n}{x_i}\right)^T.$$

Then,

$$A^{(2n)}x^{2n} = f_0(x) = x_i^{2n} f_i(y(i)) > 0,$$

i.e., $A^{(2n)}$ is positive definite.

The statement on positive semidefinite totally symmetric tensors and nonnegative polynomials can be proved similarly.

(ii) In the second case, we have

$$\begin{aligned} A^{(2n)}x^{2n} &= f_0(x) = \frac{1}{r}\sum_{i=1}^{r} x_i^{2n} f_i(y(i)) \\ &= \frac{1}{r}\sum_{i=1}^{r} x_i^{2n}\left[\sum_{j=1}^{m_i}\left(g_{ij}(y(i))\right)^2 + \alpha_i^2\right] \\ &= \frac{1}{r}\sum_{i=1}^{r}\left[\sum_{j=1}^{m_i}\left(x_i^n g_{ij}(y(i))\right)^2 + \left(\alpha_i x_i^n\right)^2\right]. \end{aligned}$$

The first case can be proved similarly. This proves (ii).

(iii) The "if" part is true for any positive integer r, as stated at the end of the last section. Therefore, we only need to prove the "only if" part. Suppose $r = 2$ and $A^{(2n)}$ is positive definite. By (i), f_i for $i = 1, \cdots, r$ are strictly positive polynomials. Since $r = 2$, f_i are polynomials of one variable. It is well-known [6, 9] that a one-variable strictly positive polynomial can be represented in the form (17). By (ii), we get the conclusion. The conclusion for positive semidefinite totally symmetric tensors can be proved similarly.

(iv) Consider the sixth order totally symmetric tensor $A^{(6)}$ for $r = 3$, defined by

$$A^{(6)}x^6 = x_1^2x_2^2(x_1^2 + x_2^2 - x_3^2) + x_3^6.$$

If $A^{(6)}x^6$ can be represented by (9), then

$$A^{(6)}x^6 = x_1^2x_2^2(x_1^2 + x_2^2 - x_3^2) + x_3^6 = \sum_{i=1}^{m}\left(B_i^{(3)}x^3\right)^2.$$

Let $x_3 = 1$. Then

$$h(x) = x_1^2x_2^2(x_1^2 + x_2^2 - 1) + 1$$

can be represented by squares of polynomials. But this is impossible [6]. Hence, (iv) holds. □

It is noted that this issue is related with Hilbert's 17th problem on the representation of nonnegative polynomials [6, 9].

5. Quartic Polynomial from Signal Processing

We now turn to the quartic polynomial f defined by (3). Firstly, we may assume that each component of $g(x)$ does not include the constant term, i.e., it has only linear and quadratic terms. If not, we may write

$$g(x) = \begin{pmatrix} g_{00} \\ g_{10} + g_1(x) \end{pmatrix},$$

where g_{00} and g_{10} are constant vectors and each component of g_1 has only linear and quadratic terms, the dimension of g_1 may be lower than m. We may partition c and G accordingly as

$$c = \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}$$

and

$$G = \begin{pmatrix} G_{00} & G_{01} \\ G_{10} & G_{11} \end{pmatrix},$$

such that the dimensions of c_1 and G_{11} are consistent with the dimension of g_1. Then we have

$$f(x) = g_1(x)^T G_{11} g_1(x) + (2G_{10}g_{00} + 2G_{11}g_{10} + c_1)^T g_1(x)$$

by dropping the constant term of f. Here G_{11} is symmetric and positive definite. Again, f has the form (3) but the dimension m may be reduced. Hence, we may assume that each component of $g(x)$ has only linear and quadratic terms.

In this way, we have

$$f(0) = 0. \tag{18}$$

Denote $g_2(x)$ as the quadratic term part of $g(x)$. Here we allow some components of g_2 to be zero such that g_2 has the same dimension as g. We say that g is **regular** if $g_2(x) \neq 0$ as long as $x \neq 0$.

THEOREM 4. *If g is regular, then f defined by (3) is a normal quartic polynomial. Then f has a global minimum point. Let the quartic term of f be denoted as $A^{(4)}x^4$. Then there are m $r \times r$ symmetric square matrices $B_1, \cdots, B_m$ such that for all $x \in \Re^r$,*

$$A^{(4)}x^4 = \sum_{i=1}^{m} \left(x^T B_i x\right)^2, \tag{19}$$

and the family $\{B_1, \cdots, B_m\}$ is **regular** *in the sense that if $x \neq 0$ then at least one of $x^T B_1 x, \cdots, x^T B_m x$ is non-zero.*

Proof. We may write f in the tensor form (12). Then we have

$$f(x) = A^{(4)}x^4 + A^{(3)}x^3 + A^{(2)}x^2 + A^{(1)}x + A^{(0)}. \tag{20}$$

By (3), we see that

$$A^{(4)}x^4 = g_2(x)^T G g_2(x).$$

If $x \neq 0$ and g is regular, we have $g_2(x) \neq 0$ and

$$A^{(4)}x^4 = g_2(x)^T G g_2(x) > 0$$

as G is positive definite. This shows that f is a normal quartic polynomial. The last conclusion of the theorem now follows from Theorem 1. Furthermore, by the properties of positive definite symmetric matrix G, $A^{(4)}x^4$ may be written in the form of (19), with

$$x^T B_i x = \sqrt{\lambda_i}(u^i)^T g(x),$$

for $i = 1, \cdots, m$, where λ_i is the ith eigenvalue of G, $u^i \in \Re^m$ is the corresponding eigenvector, $u^1, \cdots, u^m$ are orthogonal to each other. Since g is regular, the family $\{B_1, \cdots, B_m\}$ is also regular. □

Let us consider the 70-tuple example of [11], where $r = 2$ and the last term of $g(x)$ is $2x_1^2 + x_2^2$, (See (C11) of [11]). Clearly, g is regular. By Theorem 2, f is also normal. Hence, f has a global minimum point.

We now give a computable upper bound of the norm of global minima of f and a computable lower bound of the value of f. Since g is regular, we see that $\|g(x)\| \to \infty$ if $\|x\| \to \infty$.

THEOREM 5. *For all x,*

$$f(x) \geqslant -\frac{1}{2}c^T Gc. \tag{21}$$

Let $\lambda_{\min}$ be the smallest eigenvalue of G. Let L be a positive number such that if $\|x\| > L$, then

$$\|g(x)\| > \frac{\|c\|}{\lambda_{\min}}. \tag{22}$$

If x^* is a global minimum point of f, then it satisfies

$$\|x^*\| \leqslant L. \tag{23}$$

Proof. Define a function of y, say ψ, as

$$\psi(y) = y^T Gy + c^T y.$$

Then

$$f(x) = \psi(g(x))$$

and ψ has a global minimum point $y^0 = -\frac{1}{2}G^{-1}c$. For any $y \in \Re^m$, we have

$$\psi(y) \geqslant \psi(y^0) = -\frac{1}{2}c^T Gc.$$

Let $y = g(x)$, we have (21). If $\|x\| > L$, by (22), we have

$$f(x) = \psi(g(x)) \geqslant \lambda_{\min}\|g(x)\|^2 - \|c\| \cdot \|g(x)\| > 0 = f(0).$$

Hence, x cannot be a global minimum of f. This proves (23). □

The formula (21) gives a computable lower bound of the value of f. In practice, it is also not difficult to find L satisfying (22). For example, in the 70-tuple example of [11], since $r = 2$ and the last term of $g(x)$ is $2x_1^2 + x_2^2$, it is not difficult to compute L. Thus, (23) also gives a computable upper bound of the norm of global minima of f.

We now turn to the constrained problem (2). Clearly, the formula (21) still gives a computable lower bound of the global minimum value of (2). But (23) may not be true for (2) since 0 may not be feasible to (2). Assume that x_0 is a feasible point of (2). Shift $x - x_0$ to x. Clear up the constant terms of the components of the reformed function g. Then (18) is still true and 0 is feasible. Then we can have (22) as a computable upper bound of the norm of global minima of (2).

The formula (21) gives a computable lower bound of the global minimum value of (1) as well as (2), while the formula (23) gives a computable upper bound of the norm of global minima of (1) as well as (2). This makes the task of finding a global minimum of (1) or (2) much easier [4].

6. Concluding Remarks

In this paper, we discussed the global minimization problem of a multivariate polynomial, in particular, a normal multivariate polynomial. The global minimization problem of a normal quartic polynomial from signal processing is the motivation of this paper.

In 1998, Shor in his book [9] studied multivariate polynomial minimization by relating it to the 17th Hilbert problem. He reduced the multivariate polynomial minimization problem to a special quadratic equality constrained quadratic minimization problem. By considering the Lagrange function and the dual problem, Shor [9] established a dual lower bound for f^*. This dual lower bound can be obtained by solving the convex dual problem. Shor showed that the dual lower bound is equal to f^* if and only if $f(x) - f^*$ can be represented as a sum of squares of polynomials.

It is interesting to note that Shor's approach is in the opposite direction of the approach used in [11], the latter converts a high dimensional quadratic equality constrained quadratic minimization problem into a low dimensional multivariate polynomial minimization problem. It may be worth exploring the possibility to apply Shor's approach to the problem in [11] directly.

Nesterov [7] discussed the multivariate polynomial minimization problem in 2000.

In 2001, Lasserre [6] further studied this problem. By using the theory of moments, which is in duality with the theory of nonnegative polynomials and the 17th Hilbert problem, Lasserre reduced this problem to solving an (often finite) sequence of positive semidefinite programs. He showed that (1) is equivalent to a positive semidefinite program if $f(x) - f^*$ can be represented as a sum of squares of polynomials.

Recently, Kojima [5] also studied the multivariate polynomial optimization problem.

The topic of solving a system of multivariate polynomial equations is a popular topic in computational mathematics. It has applications in chemical engineering, power engineering, robot engineering and economics. The topic of multivariate polynomial interpolation is also a popular topic in approximation theory. It is natural to expect that multivariate polynomial minimization may take a similar place in the development of optimization. In fact, multivariate quartic polynomial minimization may be the simplest multivariate nonconvex global optimization problem. With its abundant theoretic interests revealed in [2, 5, 6, 7, 9] and this paper, it deserves more attention in optimization. It will be interesting to know if it has some engineering applications other than the signal processing problem discussed in this paper.

Several challenging questions remain open.

1. Is the conjecture in Sections 2 true? Prove it or give a counter example. In particular, can we prove that a quartic two variable polynomial has at most 4 local minima, or can we construct a quartic two variable polynomial having 5 local minima?

2. When 2-norm is used, $\|A^{(2)}\|$ is the largest eigenvalue of $A^{(2)}$. If 2-norm is used and $A^{(2)}$ is positive definite, $[A^{(2)}]$ is the smallest eigenvalue of $A^{(2)}$. Are there similar results for $\|A^{(3)}\|$, $\|A^{(4)}\|$ and $[A^{(4)}]$?

3. Are there practical ways to compute $\|A^{(3)}\|$, $\|A^{(4)}\|$ and $[A^{(4)}]$?

4. When $r \geqslant 3$, for a general even-order tensor $A^{(2n)}$, is the following statement true? That is, $A^{(2n)}$ is positive definite if and only if there are m nth order tensors $B_1^{(n)}, \cdots, B_m^{(n)}$ such that they form a regular family and (9) holds.

5. We showed that the quartic polynomial f defined by (3) is a normal polynomial. Can we represent any normal quartic polynomial in the form (3)? If the answer is true, then the formulas (21) and (23) can be applied to a general normal quartic polynomial. If the answer is negative, under what conditions for which a normal quartic polynomial is representable in the form (3)?

6. Let f be defined by (3). Under what conditions to be imposed on $f(x) - f^*$ such that f is representable as a sum of squares of quadratic polynomials? If so, by Shor [9] and Lasserre [6], it can be converted into a convex semidefinite program problem.

7. If we apply Shor's method and Lasserre's method to the problem (1) when f is defined by (3), can we obtain some better results?

8. Let $g_2(x)$ be an m-dimensional vector-valued function such that each component of g_2 is the sum of some quadratic product terms of $x_1, \cdots, x_r$. Can we give an algorithm to judge if g_2 is regular, i.e., $g_2(x) \neq 0$ as long as $x \neq 0$?

9. If g is regular, can we give an algorithm to calculate L satisfying (22)?

10. What is the relationship among multivariate polynomial optimization, multivariate polynomial equations, and multivariate polynomial interpolation?

Acknowledgments

The authors would like to thank Masao Fukushima, Masakazhu Kojima, Dayong Cai, Fengshan Bai, Jiu Ding, Xiaoshen Wang, Leung-fu Cheung and Zengxin Wei for discussion and comments. The authors are also thankful to the referees for their comments.

References

1. A. Cantoni and M.C. Blasikiewiez, Derivative constrained broad band antenna array processors revisited, IASTED 2nd Internat. Symp. on Signal Processing and Its Applications, Gold Coast, Queensland, Aug. 1990, pp. 319–326.
2. A. Durfee, N. Kronenfeld, H. Munsen, J. Roy and I. Westby, Counting critical points of real polynomials in two variables, *Amer. Math. Monthly*, 100 (1993), 255–271.
3. M.H. Er and A. Cantoni, *Techniques in broad band beamforming*, Control and Dynamic Systems, (C.T. Leondes ed.), Academic Press, San Diego, 1992.
4. C.A. Floudas, Deterministic Global Optimization: Theory, Methods and Applications, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.
5. M. Kojima, *A general framework for convex relaxation of polynomial optimization problems over cones*, Lecture in the Second International Conference on Optimization and Control with Applications, Huangshan, Anhui, China, Aug. 2002.
6. J.B. Lasserre, Global optimization with polynomials and the problem of moments, *SIAM J. Opt.*, 11 (2001), 796–817.
7. Y. Nesterov, Squared functional systems and optimization problems, High Performance Optimization, (H. Frenk, K. Roos, T. Terlaky and S. Zhang eds.), Dordrecht, The Netherlands, 2000. 1992.
8. J. Renegar, On the computational complexity and geometry of the first order theory of the reals. Part I: Introduction: preliminaries; the geometry of semi-algebraic sets; the decision problems for the existential theory of reals, J. Symbolic Comput., 13 (1992), 255–299.
9. N.Z. Shor, Nondifferentiable Optimization and Polynomial Problems, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.
10. I. Thng, A. Cantoni and Y.H. Leung, Derivative constrained optimum broad band antenna arrays, *IEEE Transaction on Signal Process*, 41 (1993), 2376–2388.
11. I. Thng, A. Cantoni and Y.H. Leung, Analytical solutions to the optimization of a quadratic cost function subject to linear and quadratic equality constraints, *Applied Math. Optim.*, 34 (1996) 161–382.
12. B.L. van der Waerden, Modern Algebra, Vol. 2, Ungar, New York, 1950.