A Global Homogeneous Polynomial Optimization Problem over the Unit Sphere

by

Liqun Qi

Department of Applied Mathematics
The Hong Kong Polytechnic University

Outline

- *The Problem*
- *Applications of This Problem*
- *Exact Z-Eigenvalue Methods*
- *Biquadrate Tensors*
- *Pseudo-Canonical Form Methods*
- *Numerical Results*

1. The Problem

In this talk, we consider the following global homogeneous polynomial minimization problem

$$\begin{aligned} &\min f(x) = \sum_{i_1,i_2,\cdots,i_m=1}^{n} a_{i_1 i_2 \cdots i_m} x_{i_1} x_{i_2} \cdots x_{i_m} \\ &\text{subject to } \; x^T x = 1, \end{aligned} \tag{1}$$

where $x \in \Re^n$, $m, n \geq 2$, f is a homogeneous polynomial of degree m with n variables.

This problem has wide applications in engineering and sciences. However, these applications are scattered in journals of very different disciplines and used not to be recognized as a global polynomial optimization problem in the form (1). In the next section, we will describe four such applications and analyze their relationships with problem (1). The first one is the multivariate form positive definiteness problem in automatic control, where m is even. The second one is the best rank-one approximation problem in statistical data analysis, where m is small but n can be very large. The third one is the strong ellipticity problem in solid mechanics, where $m = 4$ and $n = 2$ (in the plane) or 3 (in the space). The fourth one is the diffusion kurtosis imaging problem in biomedical engineering, where $m = 4$ and $n = 3$.

2. Applications of This Problem

- The Multivariate Form Definiteness Problem
- The Best Rank-One Approximation Problem
- The Strong Ellipticity Problem
- The Diffusion Kurtosis Imaging Problem

2.1. The Multivariate Form Definiteness Problem

Suppose that $f(x) = \sum_{i_1, i_2, \cdots, i_m = 1}^{n} a_{i_1 i_2 \cdots i_m} x_{i_1} x_{i_2} \cdots x_{i_m}$. In automatic control, such a homogeneous polynomial is called a multivariate form. If $f(x) > 0$ as long as $x \neq 0$, then we say that $f(x)$ is positive definite. Clearly, this definition is only meaningful when m, the order of f, is even. The problem to identify if an even order multivariate form is positive definite or not plays an important role in the stability study of nonlinear autonomous systems via Liapunov's direct method in automatic control.

The multivariate form $f(x)$ is positive definite if and only if the global optimal objective function value of (1) is positive. Hence, if we solve (1), we solve the multivariate form positive definiteness problem. On the other hand, to solve the multivariate form positive definiteness problem, we do not need to find a global minimizer of (1) or the exact global optimal objective function value of (1). Hence, the multivariate form positive definiteness problem is a little easier than the global minimization problem (1).

2.2. Study on the Positive Definiteness

[1]. B.D. Anderson, N.K. Bose and E.I. Jury, "Output feedback stabilization and related problems-solutions via decision methods", *IEEE Trans. Automat. Contr.* **AC20** (1975) 55-66.

[2]. N.K. Bose and P.S. Kamt, "Algorithm for stability test of multidimensional filters", *IEEE Trans. Acoust., Speech, Signal Processing,* **ASSP-22** (1974) 307-314.

[3]. N.K. Bose and A.R. Modaress, "General procedure for multivariable polynomial positivity with control applications", *IEEE Trans. Automat. Contr.* **AC21** (1976) 596-601.

[4]. N.K. Bose and R.W. Newcomb, "Tellegon's theorem and multivariate realizability theory", *Int. J. Electron.* **36** (1974) 417-425.

[5]. M. Fu, "Comments on 'A procedure for the positive definiteness of forms of even-order' ", *IEEE Trans. Autom. Contr.* **43** (1998) 1430.

[6]. M.A. Hasan and A.A. Hasan, “A procedure for the positive definiteness of forms of even-order”, *IEEE Trans. Autom. Contr.* **41** (1996) 615-617.

[7]. J.C. Hsu and A.U. Meyer, Modern Control Principles and Applications, McGraw-Hill, New York, 1968.

[8]. E.I. Jury and M. Mansour, “Positivity and nonnegativity conditions of a quartic equation and related problems” *IEEE Trans. Automat. Contr.* **AC26** (1981) 444-451.

[9]. W.H. Ku, “Explicit criterion for the positive definiteness of a general quartic form”, *IEEE Trans. Autom. Contr.* **10** (1965) 372-373.

[10]. Q. Ni, L. Qi and F. Wang, “An eigenvalue method for the positive definiteness identification problem”, to appear in: *IEEE Transactions on Automatic Control.*

[11]. F. Wang and L. Qi, “Comments on ‘Explicit criterion for the positive definiteness of a general quartic form’ ”, *IEEE Trans. Autom. Contr.* **50** (2005) 416- 418.

2.3. The Best Rank-One Approximation Problem

The best rank-one approximation to a supersymmetric tensor has applications in signal processing, wireless communication systems, signal and image processing, data analysis, higher-order statistics, as well as independent component analysis. An mth order n-dimensional real supersymmetric tensor $\mathcal{A}$ is an m-way array whose entries are addressed via m indices, and it is said to be supersymmetric if its entries $a_{i_1 \cdots i_m}$ are invariant under any permutation of their indices $\{i_1, \cdots, i_m\}$. Given a higher order supersymmetric tensor $\mathcal{A}$, if there exist a scalar λ and a unit-norm vector u such that the rank-one tensor $\bar{\mathcal{A}} \overset{\triangle}{=} \lambda u^m$ minimizes the least-squares cost function

$$\tau(\bar{\mathcal{A}}) = \|\mathcal{A} - \bar{\mathcal{A}}\|_F^2$$

over the manifold of rank-one tensors, where $\|\cdot\|_F$ is the Frobenius norm, then λu^m is called the best rank-one approximation to tensor $\mathcal{A}$.

2.4. Its Relation with Problem (1)

Denote

$$\mathcal{A}x^m = \sum_{i_1,i_2,\cdots,i_m=1}^{n} a_{i_1 i_2 \cdots i_m} x_{i_1} x_{i_2} \cdots x_{i_m}.$$

The best rank-one approximation to tensor $\mathcal{A}$ can be obtained by solving the global polynomial minimization problem (1). When m is odd, a global minimizer x of (1) and its corresponding objective function value $\lambda = \mathcal{A}x^m$ form the best rank-one approximation λx^m to $\mathcal{A}$. When m is even, let y and z be a global minimizer and a global maximizer of (1), respectively. Let $\lambda_1 = \mathcal{A}y^m$ and $\lambda_2 = \mathcal{A}z^m$. If $|\lambda_1| \geq |\lambda_2|$, let $x = y$ and $\lambda = \lambda_1$; otherwise let $x = z$ and $\lambda = \lambda_2$. Then λx^m is the best rank-one approximation to $\mathcal{A}$. Note that we may change the sign of $\mathcal{A}$ in (1) and solve the problem to find z. Hence, if we solve (1), then we may solve the best rank-one approximation problem. On the other hand, it is not difficult to show that if we solve the best rank-one approximation problem, we may also solve problem (1). Hence, we may say that these two problems are mathematically equivalent.

2.5. Study on the Best Rank-One Approximation Problem

[12]. J.F. Cardoso, "High-order contrasts for independent component analysis", *Neural Computation* 11 (1999) 157-192.

[13]. P. Comon, "Independent component analysis, a new concept?" *Signal Processing* 36 (1994) 287-314.

[14]. P. Comon, G. Golub, L-H. Lim and B. Mourrain, "Symmetric tensors and symmetric tensor rank", to appear in: *SIAM J. Matrix Anal. Appl.*

[15]. L. De Lathauwer, B. De Moor and J. Vandewalle, "On the best rank-1 and rank-$(R_1, R_2, \cdots, R_N)$ approximation of higher-order tensor", *SIAM J. Matrix Anal. Appl.* 21 (2000) 1324-1342.

[16]. L. De Lathauwer, P. Comon, B. De Moor and J. Vandewalle, "Higher-order power method—application in indepedent component analysis", in Procedings of the International Symposium on Nonlinear Theory and its Applications (NOLTA'95), Las Vegas, NV, 1995, pp. 91-96.

[17]. V.S. Grigorascu and P.A. Regalia, “Tensor displacement structures and polyspectral matching”, Chapter 9 of *Fast Reliable Algorithms for Structured Matrices*, T. Kailath and A.H. Sayed, eds., SIAM Publications, Philadeliphia, 1999.

[18]. E. Kofidis and P.A. Regalia, “On the best rank-1 approximation of higher-order supersymmetric tensors”, *SIAM J. Matrix Anal. Appl.* 23 (2002) 863-884.

[19]. C.L. Nikias and A.P. Petropulu, Higher-Order Spectra Analysis, A Nonlinear Signal Processing Framework, Prentice-Hall, Englewood Cliffs, NJ, 1993.

[20]. Y. Wang and L. Qi, “On the Successive Supersymmetric Rank-1 Decomposition of Higher Order Supersymmetric Tensors”, *Numerical Linear Algebra with Applications* 14 (2007) 503-519.

[21]. T. Zhang and G.H. Golub, “Rank-1 approximation of higher-order tensors”, *SIAM J. Matrix Anal. Appl.* 23 (2001) 534-550.

2.6. The Strong Ellipticity Problem

The elasticity tensor $\mathcal{E}$ is a fourth order tensor of dimension two (in the plane) or three (in the space). It is not supersymmetric. Its entries e_{ijkl} satisfy the following symmetry: for any i, j, k, l, we have $e_{ijkl} = e_{kjli} = e_{iklj}$. The strong ellipticity is a very important property in solid mechanics. Recently, Qi, Dai and Han [34] identified that this property holds if and only if the global optimal objective function value of the following minimization problem

$$\begin{aligned} \min g(x, y) \equiv \mathcal{E}xyxy \equiv \sum_{i,j,k,l=1}^{n} e_{ijkl} x_i y_j x_k y_l \\ \text{subject to } x^T x = 1,\ y^T y = 1, \end{aligned} \tag{2}$$

where $x, y \in \Re^n$, $n = 2$ (in the plane) or 3 (in the space). Comparing with problem (1), the dimension of problem (2) is low ($n = 2$ or 3), but the additional variable y makes the problem a little complicated. If we let $x = y$ in (2), then we have (1) with $m = 4$. Roughly speaking, the difficulty of problem (2) when $n = 2$ is equivalent to the difficulty of problem (1) when $n = 3$. When $n = 3$, in the case of anisotropic elastic materials, Han, Dai and Qi [27] show that problem (2) can be solved by solving three instances of problem (1) with $m = 2$ and $n = 3$ (these are 3×3 matrices), three instances of problem (1) with $m = 4$ and $n = 3$, and one instance of problem (1) with $m = 6$ and $n = 3$. This links the strong ellipticity problem with problem (1).

2.7. Study on the Strong Ellipticity Problem

[22]. R.C. Abeyaratne, "Discontinuous deformation gradients in plane finite elastostatic of imcompressible materials", *Journal of Elasticity* 10 (1980) 255-293.

[23]. S. Chiriţă and M. Ciarletta, "Spatial estimates for the constrained anisotropic elastic cylinder", *Journal of Elasticity* 85 (2006) 189-213.

[24]. S. Chiriţă, A. Danescu and M. Ciarletta, "On the strong ellipticity of the anisotropic linearly elastic materials", *Journal of Elasticity* 87 (2007) 1-27.

[25]. B. Dacorogna, "Necessary and sufficient conditions for strong ellipticity for isotropic functions in any dimension", *Dynamical Systems* 1B (2001) 257-263.

[26]. M.E. Gurtin, "The linear theory of elasticity", In Truesdell, C. (ed.) *Handbuch der Physik*, vol. VIa/2. Springer, Berlin, 1972.

[27]. D. Han, H.H. Dai and L. Qi, "Conditions for strong ellipticity of anisotropic elastic materials", Preprint, Department of Applied Mathematics, The Hong Kong Polytechnic University, August 2007.

[28]. J.K. Knowles and E. Sternberg, "On the ellipticity of the equations of non-linear elastostatics for a special material", *J. Elasticity* 5 (1975) 341-361.

[29]. J.K. Knowles and E. Sternberg, "On the failure of ellipticity of the equations for finite elastostatic plane strain", *Arch. Ration. Mech. Anal.* 63 (1977) 321-336.

[30]. J. Merodio and R.W. Ogden, "Instabilities and loss of ellipticity in fiber-reinforced compressible nonlinearly elastic solids under plane deformation", *International Journal of Solids Structure* 40 (2003) 4707-4727.

[31]. R.W. Ogden, "Elements of the theory of finite elasticity", In: *Nonlinear Elasticity: Theory and Applications* (eds. Y. Fu and R.W. Ogden), Cambridge University Press, Cambridge, 2001, pp. 1-57.

[32]. C. Padovani, "Strong ellipticity of transversely isotropic elasticity tensors", *Meccanica* 37 (2002) 515-525.

[33]. R.G. Payton, *Elastic wave propagation in transversely isotropic media*, Martinus Nijhoff Publishers, Boston, 1983.

[34]. L. Qi, H.H. Dai and D. Han, "Conditions for Strong Ellipticity", Preprint, Department of Applied Mathematics, The Hong Kong Polytechnic University, July 2007.

[35]. P. Rosakis, "Ellipticity and deformations with discontinuous deformation gradients in finite elastostatics", *Arch. Ration. Mech. Anal.* 109 (1990) 1-37.

[36]. H.C. Simpson and S.J. Spector, "On copositive matrices and strong ellipticity for isotrropic elstic materials", *Arch. Rational Mech. Anal.*, 84 (1983) 55-68.

[37]. J.R. Walton and J.P. Wilber, "Sufficient conditions for strong ellipticity for a class of anisotropic materials", *International Journal of Non-Linear Mechanics* 38 (2003) 441-455.

[38]. Y. Wang and M. Aron, "A reformulation of the strong ellipticity conditions for unconstrained hyperelastic media", *Journal of Elasticity,* 44 (1996) 89-96.

[39]. L. Zee and E. Sternberg, "Ordinary and strong ellipticity in the equilibrium theory of impressible hyperelastic solids", *Archive for Rational Mechanics and Analysis* 83 (1983) 53-90.

2.8. The Diffusion Tensor Imaging

A popular magnetic resonance imaging (MRI) model in medical engineering is called diffusion tensor imaging (DTI. The MR measurement of an effective diffusion tensor of water in tissues can provide unique biologically and clinically relevant information that is not available from other imaging modalities. A diffusion tensor D is a second order three dimensional fully symmetric tensor. It has six independent elements. After obtaining the values of these six independent elements by MRI techniques, the medical engineering researchers will further calculate some characteristic quantities of this diffusion tensor. These characteristic quantities are rotationally invariant, independent from the choice of the laboratory coordinate system. They include the three eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ of D, the mean diffusivity (M_D), the fractional anisotropy (FA), etc. The largest eigenvalue λ_1 describes the diffusion coefficient in the direction parallel to the fibres in the human tissue. The other two eigenvalues describe the diffusion coefficient in the direction perpendicular to the fibres in the human tissue.

2.9. The Diffusion Kurtosis Imaging Problem

However, DTI is known to have a limited capability in resolving multiple fibre orientations within one voxel. This is mainly because the probability density function for random spin displacement is non-Gaussian in the confining environment of biological tissues and, thus, the modeling of self-diffusion by a second order tensor breaks down. Recently, a new MRI model is presented by medical engineering researchers. They propose to use a fourth order three dimensional fully symmetric tensor, called the diffusion kurtosis (DK) tensor, to describe the non-Gaussian behavior. The values of the fifteen independent elements of the DK tensor W can be obtained by the MRI technique. The diffusion kurtosis imaging (DKI) has important biological and clinical significance.

What are the coordinate system independent characteristic quantities of the DK tensor W? Are there some type of eigenvalues of W, which can play a role here?

2.10. The D-Eigenvalues

Qi, Wang and Wu [45] answered these two questions. They defined D-eigenvalues for the DK tensor $W = (W_{ijkl})$. Here, "D" stands for the word diffusion. D-eigenvalues are invariant under co-ordinate system rotations. In particular, the smallest and the largest D-eigenvalues and their D-eigenvectors correspond to the smallest and the largest diffusion kurtosis coefficients and their directions. The smallest and the largest D-eigenvalues are the global minimizer and the global maximizer of the following problem:

$$\begin{aligned} \min f(x) &= \sum_{i,j,k,l=1}^{3} w_{ijkl} y_i y_j y_k y_l \\ \text{subject to } & y^T D y = 1. \end{aligned} \tag{3}$$

If we let $x = D^{\frac{1}{2}} y$, we may convert (3) to (1) with $m = 4$ and $n = 3$. Here, D is positive definite.

2.11. Study on Diffusion Tensor Imaging and Diffusion Kurtosis Imaging

[40]. P.J. Basser and D.K. Jones, "Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review", *NMR in Biomedicine,* 15 (2002) 456-467.

[41]. J.H. Jensen, J.A. Helpern, A. Ramani, H. Lu and K. Kaczynski, "Diffusional kurtosis imaging: The quantification of non-Gaussian water diffusion by means of maganetic resonance imaging", *Magnetic Resonance in Medicine,* 53 (2005) 1432-1440.

[42]. D. Li, S. Bao, C. Zhu and L. Ma, "Computing the measures of DTI based on PC and Matlab", *Chinese Journal of Medical Imaging Technology*, 20 (2004) 90-94. (in Chinese)

[43]. C. Liu, R. Bammer, B. Acar and M.E. Mosely, "Characterizing non-Gaussian diffusion by generalized diffusion tensors", *Magnetic Resonance in Medicine,* 51 (2004) 924-937.

[44]. H. Lu, J.H. Jensen, A. Ramani and J.A. Helpern, "Three-dimensional characterization of non-Gaussian water diffusion in humans using diffusion kurtosis imaging", *NMR in Biomedicine,* 19 (2006) 236-247.

[45]. L. Qi, Y. Wang and E.X. Wu, "D-eigenvalues of diffusion kurtosis tensors", to appear in: *Journal of Computational and Applied Mathematics.*

3. Exact Z-Eigenvalue Methods

We may solve problem (1) by a general global polynomial optimization method, for example, the sum of squares (SOS) method.
When $n = 2$ or 3, some other methods can also be considered. As stated before, these two cases are especially useful for the strong ellipticity problem in solid mechanics. In the case that $n = 2$, if m is odd, the SOS method needs to solve an SDP (semi-definite programming) problem of size $m + 1$, and if $m = 2d$ is even, the SOS method needs to solve an SDP problem of size $d + 1$. While the direct Z-eigenvalue method given by Qi, Wang and Wang in [51] for this case needs to solve a one-dimensional polynomial of degree $m+1$, whose coefficients are explicitly given. This work is comparable with that of the SOS method for this case. We will use this method as a subroutine for the method in the higher dimensional case.
In the case that $n = 3$, a direct Z-eigenvalue method to solve problem (1) was proposed by Qi, Wang and Wang in [51] for $m = 3$ and extended to any m in [51]. In this method, we need to calculate a determinant of size $(2m - 1)$ to find a one-dimensional polynomial of degree $(m^2 - m + 1)$, and solve it. This work is in the same order as that of the SOS method. This method is an exact method to find a global minimizer of problem (1), while the SOS method is not an exact method in general in this case. We will also use this method as a subroutine for the method in the higher dimensional case.

3.1. Eigenvalues of Tensors

The theory of eigenvalues of tensors was developed in the following papers:

[46]. L. Qi, "Eigenvalues of a real supersymmetric tensor", *Journal of Symbolic Computation* 40 (2005) 1302-1324.

[47]. L. Qi, "Rank and eigenvalues of a supersymmetric tensor, a multivariate homogeneous polynomial and an algebraic surface defined by them", *Journal of Symbolic Computation* 41 (2006) 1309-1327.

[48]. L. Qi, "Eigenvalues and invariants of tensors", *Journal of Mathematical Analysis and Applications* 325 (2007) 1363-1377.

[49]. G. Ni, L. Qi, F. Wang and Y. Wang, "The degree of the E-characteristic polynomial of an even order tensor", *J. Math. Anal. Appl.* 329 (2007) 1218-1229.

[50]. L-H. Lim, "Singular values and eigenvalues of tensors: A variational approach", Proceedings of the First IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), December 13-15, 2005, pp. 129-132.

3.2. Z-Eigenvalue Methods

Z-eigenvalue methods were developed in the following two papers:

[51]. L. Qi, F. Wang and Y. Wang, "Z-Eigenvalue methods for a global polynomial optimization problem", to appear in: *Mathematical Programming*.

[52]. L. Qi, Y. Wang and F. Wang, "A global homogeneous polynomial problem over the unit sphere", Department of Applied Mathematics, The Hong Kong Polytechnic University, August 2007.

This talk is based on [52].

3.3. Z-Eigenvalues

Let $\mathcal{A}$ be an mth order n-dimensional real supersymmetric tensor. Let $\mathcal{A}x^{m-1}$ be a vector in $\Re^n$ with its ith component as

$$(\mathcal{A}x^{m-1})_i = \sum_{i_2,\cdots,i_m=1}^{n} a_{ii_2\cdots i_m} x_{i_2} \cdots x_{i_m}.$$

Obviously, the critical points of (1) satisfy the following equations for some $\lambda \in \Re$:

$$\begin{cases} \mathcal{A}x^{m-1} = \lambda x, \\ x^T x = 1. \end{cases} \tag{4}$$

A real number λ satisfying (4) with a real vector x is called a **Z-eigenvalue** of $\mathcal{A}$, and the real vector x is called a **Z-eigenvector** of $\mathcal{A}$ associated with the Z-eigenvalue λ. In this sense, problem (1) is equivalent to finding the smallest Z-eigenvalue $\lambda_{\min}$ and the corresponding Z-eigenvector.

3.4. A Direct Z-Eigenvalue Method for $n = 2$

Denote

$$\alpha_j = a_{i_1 \cdots i_m},$$

where $i_1 = \cdots = i_{m-j} = 1$, $i_{m-j+1} = \cdots = i_m = 2$ and $0 \le j \le m$. The following theorem was given in [44].

Theorem 3.1 *Suppose that $n = 2$.*
If $\alpha_1 = a_{11\ldots12} = 0$, then $\lambda = \alpha_0 = a_{11\ldots1}$ is a Z-eigenvalue of $\mathcal{A}$, with a Z-eigenvector $x = (1,0)^T$. If furthermore m is odd, then $\lambda = -a_{11\ldots1}$ is also a Z-eigenvalue of $\mathcal{A}$, with a Z-eigenvector $x = (-1,0)^T$.
The other Z-eigenvalues and corresponding Z-eigenvectors of $\mathcal{A}$ can be found by finding real roots of the following one dimensional polynomial equation of t:

$$\sum_{j=0}^{m-1} \binom{m-1}{j} \left[\alpha_j t^{m-j} - \alpha_{j+1} t^{m-j+1}\right] = 0, \tag{5}$$

and substituting such real values of t to

$$x_1 = \pm\frac{t}{\sqrt{1+t^2}}, \quad x_2 = \pm\frac{1}{\sqrt{1+t^2}},$$

and

$$\lambda = \sum_{j=0}^{m} \binom{m}{j} \alpha_j x_1^{m-j} x_2^j.$$

Equation (5) has at most $m + 1$ real roots. After finding all the Z-eigenvalues of $\mathcal{A}$, and the Z-eigenvectors associated with them, we may easily solve (1).

3.5. A Direct Z-Eigenvalue Method for $n = 3$

For $n = 3$, a direct Z-eigenvalue method was proposed to solve (1) for $m = 3$ in [51] and extended to any m in [52]. In this method, we calculate a determinant of size $2m - 1$ to find a one-dimensional polynomial of degree $m^2 - m + 1$, and solve it. This work is in the same order as that of the SOS method. Since this method is an exact method, it is usable if we wish to assure finding a global minimizer of (1) in this case. In [52], we use this method as a subroutine for an algorithm solving the problem with a larger dimension.

Let α_j be the same as defined before. For $0 \le i, j \le m - 1$, denote

$$\binom{m-1}{i,j} = \frac{(m-1)!}{i!j!(m-1-i-j)!},$$

$\beta_j = a_{3i_1\cdots i_{m-1}}$, for $i_1 = \cdots = i_{m-1-j} = 1,\ i_{m-j} = \cdots = i_{m-1} = 2$, and denote $a_{k\underbrace{1\cdots 1}_{i}\underbrace{2\cdots 2}_{j}\underbrace{3\cdots 3}_{(m-1-i-j)}}$ by $\gamma_{k,i,j}$ for $k = 1, 2, 3$.

Theorem 3.2 *Suppose that* $n = 3$. *Then the following statements hold.*
(a). If $a_{11\cdots 12} = a_{11\cdots 13} = 0$, *then* $\lambda = a_{11\cdots 1}$ *is a Z-eigenvalue of* $\mathcal{A}$ *with a Z-eigenvector* $x = (1,0,0)^T$. *If furthermore* m *is odd, then* $\lambda = -a_{11\cdots 1}$ *is also a Z-eigenvalue of* $\mathcal{A}$, *with a Z-eigenvector* $x = (-1,0,0)^T$.
(b). For any real root t *of the following equations:*

$$\begin{cases} \sum_{j=0}^{m-1} \binom{m-1}{j} \left[\alpha_j t^{m-j-1} - \alpha_{j+1} t^{m-j}\right] = 0, \\ \sum_{j=0}^{m-1} \binom{m-1}{j} \beta_j t^{m-j-1} = 0, \end{cases} \tag{6}$$

$$x = \pm \frac{1}{\sqrt{t^2+1}} (t, 1, 0)^T \tag{7}$$

is a Z-eigenvector of $\mathcal{A}$ *with the Z-eigenvalue* $\lambda = \mathcal{A}x^m$.
(c). The other Z-eigenvalues and corresponding Z-eigenvectors of $\mathcal{A}$ *can be found by finding real solutions of the following polynomial equations in* u *and* v*:*

$$\begin{cases} b^3_{m-1}(v) u^m + \sum_{i=1}^{m-1} \left[b^3_{i-1}(v) - b^1_i(v)\right] u^i - b^1_0(v) = 0, \\ \sum_{i=0}^{m-1} \left[b^3_i(v) v - b^2_i(v)\right] u^i = 0, \end{cases} \tag{8}$$

where

$$b_i^k(v) = \sum_{j=0}^{m-1-i} \binom{m-1}{i,j} \gamma_{k,i,j} v^j, \qquad k = 1,2,3, \;\; i = 0,1,\cdots,m-1,$$

and substituting such real values of $(u,v)^T$ *to*

$$x = \pm \frac{1}{\sqrt{u^2+v^2+1}}(u,v,1)^T \tag{9}$$

and $\lambda = \mathcal{A}x^m$.

We regard the polynomial equation system (8) as equations of u. It has complex solutions if and only if its resultant attains zero. Note that its resultant is a one-dimensional polynomial equation of v which can be obtained by computing the determinant of a $(2m-1)$ square matrix defined by coefficients in system (8). Hence, we may find all the real roots of this one-dimensional polynomial, and substitute them to (8) to find all the real solutions of u. Since these solutions correspond to E-eigenvalues of $\mathcal{A}$ (E-eigenpairs are complex solutions of (4)), by [49], the degree of this one-dimensional polynomial is not greater than (m^2-m+1) when m is even. We believe that this conclusion is also true when m is odd.

4. Biquadrate Tensors

In [52], we also propose a direct Z-eigenvalue method to solve (1) in the case of biquadrate tensors. A biquadrate tensor is a special fourth order n-dimensional supersymmetric tensor. Its dimension n can be arbitrary such that it can be used as a testing example for the method proposed in [52] for higher dimensions.

Suppose that $\mathcal{A}$ is a fourth order n-dimensional supersymmetric tensor. We call $\mathcal{A}$ a biquadrate tensor if its elements satisfy the following conditions: for $i_1 \le i_2 \le i_3 \le i_4$,

$$a_{i_1 i_2 i_3 i_4} = 0, \quad \text{if } i_1 \ne i_2 \text{ or } i_3 \ne i_4.$$

For the sake of simplicity, we denote

$$c_{ij} = \begin{cases} a_{iiii}, & \text{for} \quad i = 1, 2, \cdots, n, \\ 3a_{iijj}, & \text{for} \quad i \ne j,\ i, j = 1, 2, \cdots, n. \end{cases}$$

Certainly, they are the only possible nonzero elements of $\mathcal{A}$.

Suppose that $\mathcal{A}$ is a real biquadrate tensor. Then problem (1) reduces to the following quadratic problem:

$$\min \sum_{i,j=1}^{n} c_{ij} y_i y_j$$
$$\text{s.t.} y_1 + \cdots + y_n = 1, y_i \geq 0, \quad i = 1, 2, \cdots, n.$$

In the nonconvex case, this problem is not trivial.
The following theorem presents a method for computing all the Z-eigenvalues of a real biquadratic tensor $\mathcal{A}$.

Theorem 4.1 *Suppose that $\mathcal{A}$ is a real biquadratic tensor. Then all the Z-eigenvectors $x = (x_1, \cdots, x_n)^\top$ of $\mathcal{A}$ can be found by solving the following system of linear systems*

$$\begin{cases} \sum_{j \in S} c_{ij} y_j = \lambda, \quad i \in S, \ \ y_j = 0, \quad j \notin S, \quad y_j \geq 0, \quad j \in S \\ \sum_{i \in S} y_i = 1, \end{cases} \tag{10}$$

where $S \subset \{1, 2, \cdots, n\}$ and $|S| \geq 1$. Using $\lambda = \mathcal{A}x^4$, we find the corresponding Z-eigenvalues. Solving (10) for each subset S of $\{1, 2, \cdots, n\}$ with $|S| \geq 2$, we find all the other Z-eigenvalues of $\mathcal{A}$.

5. Pseudo-Canonical Form Methods

For the case that $n \geq 16$ and $m \geq 3$, it is beyond the practical limit of the SOS method. Hence, for such a case, in [52], we propose an r-th order pseudo-canonical form method which uses lower-dimensional methods as subroutines.

5.1. Orthogonal Similarity

Let $\mathcal{A}$ be an mth order n-dimensional supersymmetric tensor, $P = (p_{ij})$ be an $n \times n$ real matrix. Define $\mathcal{B} = P^m\mathcal{A}$ as another mth order n-dimensional tensor with entries

$$b_{i_1 i_2 \cdots i_m} = \sum_{j_1, j_2, \cdots, j_m = 1}^{n} p_{i_1 j_1} p_{i_2 j_2} \cdots p_{i_m j_m} a_{j_1 j_2 \cdots j_m}.$$

If P is an orthogonal matrix, then we say that A and B are **orthogonally similar**. This is a reminiscence of the orthogonal transformation for symmetric matrices. By [46], we have the following theorem.

Theorem 5.1 *Suppose that $\mathcal{A}$ is an mth order n-dimensional supersymmetric tensor, $\mathcal{B} = P^m\mathcal{A}$, P is an $n \times n$ orthogonal matrix. Then $\mathcal{A}$ and $\mathcal{B}$ have the same Z-eigenvalues. If λ is a Z-eigenvalue of $\mathcal{A}$ with a Z-eigenvector x, then λ is a Z-eigenvalue of $\mathcal{B}$ with a Z-eigenvector $y = Px$.*

5.2. Pseudo-Canonical Forms

Suppose that λ is a Z-eigenvalue of $\mathcal{A}$ with a Z-eigenvector x. Let P be an orthogonal matrix with x^T as its first row. Let $\mathcal{B} = P^m\mathcal{A}$. Then we see that $y = Px = e^{(1)}$. By (4), we see that

$$b_{11\cdots 1} = \lambda, \quad b_{11\cdots 1i} = 0, \quad \text{for } i = 2, \cdots, n.$$

An mth order n-dimensional supersymmetric tensor $\mathcal{B}$ is said to be a pseudo-canonical form of another mth order n-dimensional supersymmetric tensor $\mathcal{A}$ if $\mathcal{A}$ and $\mathcal{B}$ are orthogonally similar and

$$b_{ii\cdots ij} = 0$$

for all $1 \le i < j \le n$. In this case, we say that $\mathcal{B}$ is a pseudo-canonical form.

5.3. rth Order Pseudo-Canonical Forms

Suppose that r is an integer satisfying $2 \leq r \leq 15$ and $r < n$. Let $1 \leq j_1 < j_2 < \cdots j_r \leq n$. We use $\mathcal{B}(j_1, j_2, \cdots, j_r)$ to denote the mth order r-dimensional supersymmetric tensor whose entries are $b_{i_1 i_2, \cdots i_m}$ for $i_1, i_2, \cdots, i_m = j_1, j_2, \cdots, j_r$. We also use $[\mathcal{B}(j_1, j_2, \cdots, j_r)]_{\min}$ to denote the smallest Z-eigenvalue of $\mathcal{B}(j_1, j_2, \cdots, j_r)$.

An mth order n-dimensional supersymmetric tensor $\mathcal{B}$ is called an r-th order pseudo-canonical form of another mth order n-dimensional supersymmetric tensor $\mathcal{A}$ if it is a pseudo-canonical form of $\mathcal{A}$ and

$$b_{111\cdots 1} = \min_{1 \leq j_1 < j_2 < \cdots < j_r \leq n} [\mathcal{B}(j_1, j_2, \cdots, j_r)]_{\min}.$$

If we find an rth order pseudo-canonical form $\mathcal{B} = P^m \mathcal{A}$, then $b_{111\cdots 1}$ and the first row vector of P are approximations to the smallest Z-eigenvalue of $\mathcal{A}$ and its corresponding Z-eigenvector. In our designed algorithm below, we will try to find such an rth order pseudo-canonical form of tensor $\mathcal{A}$ by using the orthogonal transformation technique combined with lower-dimensional methods and some optimization method.

5.4. An rth Order Pseudo-Canonical Form Method

Throughout the algorithm, we need to compute global minimizers of lower-dimensional minimization problems, and use the obtained solutions as initial points to find local minimizers of problem (1). To find global minimizers of lower-dimensional minimization problems, we use the exact Z-eigenvalue methods for $n = 2, 3$ and the SOS method for $4 \leq n \leq 15$. Then with the obtained solutions as initial points we use the projected gradient method to find local minimizers of problem (1) as the projection from $\Re^n$ to the unit ball can easily be obtained. For simplicity, we denote the first method by Algorithm M_1 and the second method, i.e., the projected gradient method, by Algorithm M_2.
The basic ideas of our algorithm are as follows.
In Step 1, we first fix the values of $n - r$ variables as zeros and use Algorithm M_1 to solve a reduced version of problem (1) with dimension r, and then take the obtained solution as the initial point and use Algorithm M_2 to find a local minimizer of problem (1), say $x^{(0)}$. Construct an orthogonal matrix Q based on vector $x^{(0)}$ and let $P = Q$, where the orthogonal matrix P denotes the orthogonal transformation made to tensor $\mathcal{A}$ during the iterations. In the iterative step, it contains two procedures which mainly concern the following transformed problem

$$
\begin{aligned}
&\min \mathcal{B}x^m \\
&\text{s.t.} x^T x = 1,
\end{aligned} \tag{11}
$$

where $\mathcal{B} = P^m \mathcal{A}$. First, we fix the values of variables x_1 and last $(n - r - 1)$ variables as zeros in problem (11) and use Algorithm M_1 to solve it. Second, based on the obtained point, we use Algorithm M_2 to find a local minimizer of the original problem (1) and a new problem obtained by adding constraint $x_1 = 0$ to problem (1), respectively. The two local minimizers are respectively denoted by $x^{(1)}$ and $y^{(1)}$. If $f(x^{(1)}) < f(x^{(0)})$, then replace $x^{(0)}$ by $x^{(1)}$ and go to Step 1. Otherwise, use $(e^{(1)})^T$ and $(y^{(1)})^T$ as the first two rows to construct another orthogonal matrix Q and let $P = QP$. Repeat this process, until it cannot be executed.

6. Numerical Results

The computation was done on a personal computer (Pentium IV, 2.8GHz) by running Matlab 7.0. To test the performance of the methods, we use three classes of examples where the objective functions assume the following forms:

TP1 $$f(x) = \sum_{i,j,k=1}^{n} a_{ijk} x_i x_j x_k,$$

TP2 $$f(x) = \sum_{i,j,k,l=1}^{n} a_{ijkl} x_i x_j x_k x_l,$$

TP3 $$f(x) = \sum_{i,j=1}^{n} c_{ij} x_i^2 x_j^2.$$

In the following, we use the 3rd, the 6th and the 9th order pseudo-canonical form methods to find global minimums and minimizers of (1). In our computation, we use the direct Z-eigenvalue method ($r = 3$) and the SOS method ($r = 6, 9$), as Algorithm M_1, to find global minimizers of lower-dimensional minimization subproblems, and we adopt the projected gradient method, as algorithm M_2, to find a local minimizer of (1).

6.1. Numerical Results for $m = 3$

TP1 $$f(x) = \sum_{i,j,k=1}^{n} a_{ijk} x_i x_j x_k.$$

We take $a_{ijk} = -i + \frac{j^3}{3} - \frac{1}{k}$ for $1 \leq i \leq j \leq k \leq n$. The other a_{ijk} are generated by the supersymmetry. By using the 3rd, the 6th and the 9th order pseudo-canonical form methods respectively, we have the following numerical results.

For $n = 10$, we obtain the global minimum of (1), $f^* = -3.3597 \times 10^3$, and a global minimizer of (1),

$$x^* = -(0.1936,\ 0.1921,\ 0.1939,\ 0.2022,\ 0.2213,$$

$$0.2552,\ 0.3076,\ 0.3791,\ 0.4619,\ 0.5305)^T.$$

This solution coincides with the solution obtained by the SOS method.

For $n = 20$, we get an approximate optimal value of (1), $\bar{f} = -7.0374 \times 10^4$, and an approximate global minimizer,

$$\bar{x} = -(0.1345,\ 0.1343,\ 0.1343,\ 0.1346,\ 0.1356,\ 0.1374,$$

$$0.1405,\ 0.1452,\ 0.1520,\ 0.1611,\ 0.1731,\ 0.1883,\ 0.2069,$$

$$0.2291,\ 0.2547,\ 0.2830,\ 0.3127,\ 0.3417,\ 0.3665,\ 0.3820)^T.$$

For $n = 30$, we get the following approximate minimizer,

$$\bar{x} = -(0.1093,\ 0.1092,\ 0.1092,\ 0.1092,\ 0.1094,\ 0.1097,$$
$$0.1102,\ 0.1111,\ 0.1123,\ 0.1140,\ 0.1162,\ 0.1191,$$
$$0.1228,\ 0.1273,\ 0.1328,\ 0.1393,\ 0.1469,\ 0.1558,$$
$$0.1660,\ 0.1775,\ 0.1902,\ 0.2042,\ 0.2193,\ 0.2352,$$
$$0.2515,\ 0.2678,\ 0.2832,\ 0.2968,\ 0.3074,\ 0.3135)^T$$

with the function value $\bar{f} = -4.2383 \times 10^5$.

6.2. Numerical Results for $m = 4$

TP2 $$f(x) = \sum_{i,j,k,l=1}^{n} a_{ijkl} x_i x_j x_k x_l.$$

For this class of examples, we take $a_{ijkl} = i^3 - j^2 + 3ijk - l^4$ for $1 \le i \le j \le k \le l \le n$. The other a_{ijkl} are generated by the supersymmetry.
For $n = 10$, our computed global minimum of (1) is $f^* = -6.2595 \times 10^5$ and a global minimizer is

$$x^* = (0.2947,\ 0.2913,\ 0.2878,\ 0.2843,\ 0.2815,$$

$$0.2812,\ 0.2869,\ 0.3058,\ 0.3521,\ 0.4545)^T.$$

This solution also coincides with the solution obtained by the SOS method. For $n = 20$, our computed optimal value of (1) is $\bar{f} = -3.7833 \times 10^7$ and a minimizer of (1) is

$$\bar{x} = (0.2031, 0.2026, 0.2020, 0.2014, 0.2008, 0.2002,$$

$$0.1997, 0.1991, 0.1988, 0.1987, 0.1990, 0.2002, 0.2024,$$

$$0.2065, 0.2132, 0.2236, 0.2395, 0.2633, 0.2985, 0.3505)^T.$$

For $n = 30$, our computed optimal value of (1) is $\bar{f} = -4.2116 \times 10^8$ and a minimizer of (1) is

$$\bar{x} = (0.1644,\ 0.1642,\ 0.1640,\ 0.1638,\ 0.1636,\ 0.1634,$$
$$0.1632,\ 0.1630,\ 0.1628,\ 0.1626,\ 0.1624,\ 0.1622,$$
$$0.1621,\ 0.1621,\ 0.1623,\ 0.1625,\ 0.1631,\ 0.1639,$$
$$0.1653,\ 0.1671,\ 0.1698,\ 0.1735,\ 0.1784,\ 0.1849,$$
$$0.1935,\ 0.2048,\ 0.2195,\ 0.2385,\ 0.2632,\ 0.2954)^T.$$

6.3. Numerical Results for for Biquadrate Tensors

$$\textbf{TP3} \qquad f(x) = \sum_{i,j=1}^{n} c_{ij} x_i^2 x_j^2.$$

For this class of examples, we take $c_{ij} = c_{ji} = \frac{1}{2}(i + 1/j)$ for $1 \le i < j \le n$ and $c_{ii} = i + 1/i$ for $i = 1, 2, \cdots, n$. For $n = 10$, we used the SOS method to find a global minimizer of (1), but failed. For this instance, the SOS method can only provide an optimal value $f^* = 1.2805$.

When we use the direct method described in Section 4, we obtain the global minimum of (1), $f^* = 1.2805$, and a global minimizer of (1),

$$x^* = (0.67485,\ 0.48743,\ 0.34748,\ 0.25851,\ 0.20030,$$
$$0.16108,\ 0.13448,\ 0.11695,\ 0.10639,\ 0.10139)^T.$$

For this case, by using the 9th order pseudo-canonical form method, we obtain an approximate optimal value of (1) with relative error 1.06×10^{-11}, and an approximate global minimizer of (1),

$$\bar{x} = (0.67485,\ -0.48743,\ -0.34748,\ -0.25851,\ -0.20030,$$
$$-0.16108,\ 0.13448,\ -0.11695,\ 0.10638,\ 0.10140)^T.$$

For this case, by using the 3rd and the 6th order pseudo-canonical form methods, we obtain an approximate global minimum of (1), $\bar{f} = 1.2812$ and an approximately global minimizer of (1),

$$\bar{x} = (-0.67504,\ -0.48780,\ -0.34818,\ -0.25980,\ 0.20254,$$
$$-0.16478,\ 0.14030,\ -0.12558,\ 0.11843,\ 0)^T.$$

For $n = 20$, by using the direct method, we obtain the global minimizer of (1), $f^* = 1.2792$, and a global minimizer of (1),

$$x^* = (0.67458,\ 0.48691,\ 0.34650,\ 0.25672,\ 0.19714,\ 0.15577,$$

$$0.12593,\ 0.10375,\ 0.086875,\ 0.073798,\ 0.063538,\ 0.055429,\ 0.049017,$$

$$0.043986,\ 0.040117,\ 0.037258,\ 0.035296,\ 0.034145,\ 0.033730,\ 0.033979)^T.$$

When we use the 6th and the 9th order pseudo-canonical form methods, we obtain an approximate global minimum of (1), with relative error 6.53×10^{-11}, and an approximate minimizer of (1),

$$\bar{x} = (-0.67458,\ -0.48691,\ 0.34650,\ -0.25672,\ 0.19714,\ 0.15577,$$

$$0.12593,\ -0.10375,\ 0.086875,\ 0.073797,\ 0.063538,\ -0.055430,\ -0.049018,$$

$$-0.043991,\ -0.040107,\ 0.037271,\ -0.035309,\ 0.034157,\ -0.033695,\ -0.033980)^T.$$

When we use the 3rd order pseudo-canonical form method, we get an approximate global minimum of (1), $\bar{f} = 1.2800$, and an approximate global minimizer

$$\bar{x} = (0.67474,\ 0.48722,\ 0.34708,\ 0.25779,\ -0.19903,\ 0.15896,\ -0.13111,$$

$$-0.11182,\ 0.099001,\ 0.091296,\ 0.087792,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0)^T.$$

For $n = 30$, when we use the 3rd order pseudo-canonical form method, our computed optimal value of (1) is $\bar{f} = 1.2800$ and a minimizer of (1) is

$$\bar{x} = (0.67474,\ 0.48722,\ 0.34708,\ 0.25779,\ -0.19903,\ 0.15896,$$

$$-0.13111,\ -0.11182,\ 0.099001,\ 0.091296,\ 0.087792,\ 0,\ \cdots,\ 0)^T.$$

For this case, when we use the 6th order pseudo-canonical form method, our computed optimal value of (1) is $\bar{f} = 1.2792$ and a minimizer of (1) is

$$\bar{x} = (0.67458,\ 0.48691,\ -0.34650,\ 0.25672,\ -0.19714,$$

$$-0.15577,\ -0.12593,\ 0.10375,\ -0.086875,\ 0.073797,\ -0.063538,$$

$$0.055430,\ 0.049018,\ 0.043991,\ 0.040107,\ -0.037271,\ 0.035309,$$

$$0.034157,\ 0.033695,\ 0.033980,\ 0,\ 0,\ \cdots,\ 0)^T.$$

For this case, when we use the 9th order pseudo-canonical form method, our computed optimal value of (1) is $\bar{f} = 1.2792$ and a minimizer of (1) is

$$\bar{x} = (0.67458, \ -0.48690, \ -0.34648, \ -0.25669, \ -0.19710,$$

$$0.15569, \ -0.12580, \ 0.10355, \ 0.086555, \ 0.073315, \ -0.062823,$$
$$-0.054393, \ 0.047521, \ -0.041927, \ 0.037286, \ 0.033448, \ -0.030272,$$
$$0.027570, \ 0.025514, \ -0.023794, \ -0.022460, \ 0.021464, \ 0.020837,$$
$$-0.020471, \ -0.020426, \ 0.020659, \ 0.021060, \ 0, \ 0, \ 0)^T.$$

For this case, the direct method described could not give a global optimal minimizer of (1) because of its expensive computations.

The numerical results show that the rth order pseudo-canonical form method is a practical method to solve problem (1) in the case that $n \geq 16$.