Global solutions to general polynomial benchmark optimization problems

Xiaojun Zhou · David Yang Gao · Chunhua Yang

Received: date / Accepted: date

Abstract  The goal of this paper is to solve a class of high-order polynomial benchmark optimization problems, including the Goldstein-Price problem and the Three Hump Camel Back problem. By using a generalized canonical duality theory, we are able to transform the nonconvex primal problems to concave dual problems over convex domain (without duality gap), which can be solved easily to obtain global solutions.

Keywords  Global optimization · Canonical duality theory · Polynomial benchmark problem

1 Introduction

Polynomial optimization problems have been widely studied in various fields such as nonlinear algebra, semidefinite programming, and operations research, with extensive applications in production planning, location and distribution, engineering design, risk management, water treatment and distribution, chemical process design, pooling and blending, structural design, signal processing, robust stability analysis, design of chips, and much more (see [1,2]).

Xiaojun Zhou
School of Science, Information Technology and Engineering, University of Ballarat, Victoria 3353, Australia.
School of Information Science and Engineering, Central South University, Changsha 410083, China.

David Yang Gao
School of Science, Information Technology and Engineering, University of Ballarat, Victoria 3353, Australia.

Chunhua Yang
School of Information Science and Engineering, Central South University, Changsha 410083, China.
Due to the nonconvexity, traditional direct methods for solving polynomial optimization problems are usually very difficult, or even impossible. For example, the algebraic method for the task is to find all of the critical points firstly and then to identify the global minimizer(s) among all these critical points. This approach becomes inefficient when there exist numerous local minima. Also, linearization and relaxation techniques were used to compute an approximate optimal solution of the primal problem. However, the approximate optimal solution was not guaranteed to be the actual global optimum \[3\] \[4\]. In \[5\], the so-called Z-eigenvalue methods were proposed to solve the best rank-one approximation problem, but they can be applied only for third-order polynomials. Besides these deterministic methods, stochastic techniques have also made significant contributions to the optimization applications of this kind \[6\] \[7\]. For example, the evolutionary computation method, could solve general problems in low dimension, but it failed to do well for large scale ones \[8\] \[9\] \[10\]. Generally speaking, due to the lack of a theory for identify the global minimizer(s), many polynomial optimization problems are considered to be \(NP\)-hard.

The canonical duality theory was originated in the late 1980s by Gao and Strang, and has developed significantly in recent years, both theoretically and practically \[11\] \[12\]. Actually, the canonical duality theory has been successfully applied to solve some special polynomial optimization problems. In \[13\], a special polynomial minimization problem called canonical polynomial was completely solved by the canonical duality theory. In \[14\], the theory was used to solve a special 8th order polynomial minimization problem. Recently, canonical dual solutions to sum of fourth-order polynomials minimization problems have also achieved \[15\]. This paper aims to solve some general polynomial benchmark problems by using a generalized canonical duality theory. Experimental results show that these polynomial benchmark problems can be solved completely by the canonical duality theory.

### 2 A brief review of the canonical duality theory

Let’s consider the following general polynomial optimization problem (primal problem)

\[
(P) : \min_{x \in \mathbb{R}^n} \left\{ P(x) = \frac{1}{2} x^T A x - x^T f + W(x) \right\},
\]

where, \(A \in \mathbb{R}^{n \times n}\) is a given symmetrical indefinite matrix, \(f \in \mathbb{R}^n\) is a given vector, \(W(x) : \mathbb{R}^n \to \mathbb{R}\) is a general nonconvex \(C^2\) function.

The main procedures of general methodology of the canonical duality theory can be summarized as the following three steps:

**Step 1: Canonical dual transformation**

Introducing a nonlinear operator (a Gâteaux differentiable geometrical measure)

\[
\xi = \Lambda(x) : \mathbb{R}^n \to \mathcal{E}_a \subset \mathbb{R}^m
\]
and a convex function $V : E_a \rightarrow \mathbb{R}$ such that $W(x)$ can be recast by $W(x) = V(\Lambda(x))$. Then the primal problem can be rewritten as the canonical form:

$$\min_{x \in \mathbb{R}^n} \left\{ P(x) = V(\Lambda(x)) - U(x) \right\},$$  \hspace{1cm} (3)

where $U(x) = -\frac{1}{2}x^T Ax + x^T f$.

**Step 2: Generalized complementary function**

The dual variable $\varsigma$ to $\xi$ is defined by the duality mapping

$$\varsigma = \nabla V(\xi) : E_a \rightarrow E^*_a \subset \mathbb{R}^m,$$  \hspace{1cm} (4)

which should be invertible, due to the convexity of $V(\xi)$. Then the Legendre conjugate $V^*(\varsigma)$ of $V(\xi)$ can be uniquely defined by the Legendre transformation

$$V^*(\varsigma) = \text{sta}\{ \xi^T \varsigma - V(\xi) | \xi \in E_a \}$$  \hspace{1cm} (5)

and the following canonical duality relations hold on $E_a \times E^*_a$:

$$\varsigma = \nabla V(\xi) \Leftrightarrow \xi = \nabla V^*(\varsigma) \Leftrightarrow V(\xi) + V^*(\varsigma) = \xi^T \varsigma.$$  \hspace{1cm} (6)

Replacing $W(x) = V(\Lambda(x))$ by $\Lambda(x)^T \varsigma - V^*(\varsigma)$, we obtain the following generalized complementary function:

$$\Xi(x, \varsigma) = \Lambda(x)^T \varsigma - V^*(\varsigma) - U(x) : \mathbb{R}^n \times E^*_a \rightarrow \mathbb{R}.$$  \hspace{1cm} (7)

**Step 3: Canonical dual function**

By using the generalized complementary function, the canonical dual function $P^d(\varsigma)$ can be formulated as

$$P^d(\varsigma) = \text{sta}\{ \Xi(x, \varsigma) | x \in \mathbb{R}^n \} = U^\Lambda(\varsigma) - V^*(\varsigma),$$  \hspace{1cm} (8)

where $U^\Lambda(\varsigma)$ is defined by

$$U^\Lambda(\varsigma) = \text{sta}\{ \Lambda(x)^T \varsigma - U(x) | x \in \mathbb{R}^n \}.$$  \hspace{1cm} (9)

Let $S_a \subset E^*_a$ be a dual feasible space such that $U^\Lambda(\varsigma)$ is well-defined, and the canonical dual problem can be obtained as

$$(P^d) : \text{sta}\{ P^d(\varsigma) | \varsigma \in S_a \}.$$  \hspace{1cm} (10)

**Theorem 1 (Complementary-Dual Principle)**\cite{12}. The problem $(P^d)$ is canonically dual to the primal problem $(P)$ in the sense that if $(\bar{x}, \bar{\varsigma})$ is a critical point of $\Xi(x, \varsigma)$, then $\bar{x}$ is a feasible solution of $(P)$, $\bar{\varsigma}$ is a feasible solution of $(P^d)$, and

$$P(\bar{x}) = \Xi(\bar{x}, \bar{\varsigma}) = P^d(\bar{\varsigma}).$$  \hspace{1cm} (11)

In many applications, the geometrical operator $\Lambda(x)$ is usually quadratic

$$\Lambda(x) = \left\{ \frac{1}{2}x^TC_kx + x^T b_k \right\} : \mathbb{R}^n \rightarrow E_a \subset \mathbb{R}^m,$$  \hspace{1cm} (12)
where \( C_k \in \mathbb{R}^{n \times n} \) and \( b_k \in \mathbb{R}^n \) are given. In this case, the canonical dual function can be formulated in the form of

\[
P_d(\varsigma) = -\frac{1}{2} F^T(\varsigma) G^{-1}(\varsigma) F(\varsigma) - V^*(\varsigma), \tag{13}
\]

which is well defined on

\[
S_a = \{ \varsigma \in \mathbb{R}^m | F(\varsigma) \in \text{Col}(G(\varsigma)) \}, \tag{14}
\]

where \( G(\varsigma) = A + \sum_{k=1}^{m} \varsigma_k C_k \), \( F(\varsigma) = f - \sum_{k=1}^{m} \varsigma_k b_k \), and \( \text{Col}(G(\varsigma)) \) denotes the column space of \( G(\varsigma) \).

Let the positive domain

\[
S_a^+ = \{ \varsigma \in S_a | G(\varsigma) \succeq 0 \} \tag{15}
\]

where \( G(\varsigma) \succeq 0 \) indicates that \( G(\varsigma) \) is a positive semi-definite matrix.

**Theorem 2** (Global Optimality condition) \[12\]. Suppose \( \bar{\varsigma} \) is a critical point of \( P_d \) and \( \bar{x} = G^{-1}(\bar{\varsigma}) F(\bar{\varsigma}) \). If \( \bar{\varsigma} \in S_a^+ \), then \( \bar{\varsigma} \) is a global maximizer of \( (P_d) \) on \( S_a^+ \) if and only if \( \bar{x} \) is a global minimizer of \( (P) \) on \( \mathbb{R}^n \), i.e.,

\[
P(\bar{x}) = \min_{x \in \mathbb{R}^n} P(x) \iff \max_{\varsigma \in S_a^+} P_d(\varsigma) = P_d(\bar{\varsigma}). \tag{16}
\]

Some polynomial problems have already been given to testify the effectiveness of canonical duality theory, but most of them are no more than fourth degree (see \[15\]). Although some larger degree polynomial problems are completely solved by the same theory, they belong to a special case (see \[13,14\]). In the next two sections, we are to solve some general polynomial benchmark optimization problems (they are also not obvious to find the global optimum by observation), aiming to expand the use of canonical duality methodology.

### 3 Application for Goldstein-Price problem

The Goldstein-Price problem is given in the form of \[16\]:

\[
\min_{x,y} f_1(x, y) = [1 + (x + y + 1)^2(19 - 14x + 3x^2 - 14y + 6xy + 3y^2)] \times \\
[30 + (2x - 3y)^2(18 - 32x + 12x^2 + 48y - 36xy + 27y^2)].
\]

The landscape and contour of Goldstein-Price function are given in Fig.1 and we can find that there exist a few extrema. Due to the nonconvexity of the problem, it is not easy to find the global minimum.

By using the following linear transformation

\[
\begin{pmatrix} s \\ t \end{pmatrix} = T(x, y) = \begin{pmatrix} x + y \\ 2x - 3y \end{pmatrix}, \tag{17}
\]

the Goldstein-Price function can be rewritten to

\[
f_1(x, y) = h(s)g(t) \tag{18}
\]
Global solutions to general polynomial benchmark optimization problems

Fig. 1 Graph and contour of Goldstein-Price function

where

\[ h(s) = 1 + (s + 1)^2(19 - 14s + 3s^2) \]  

(19)

and

\[ g(t) = 30 + t^2(18 - 16t + 3t^2) \]  

(20)

**Proposition 1** Under the linear transformation \( T(x, y) \), the Goldstein-Price problem is equivalent to the decoupled problems as follows

\[ \min_{x,y} f_1(x, y) = \min_s h(s) \min_t g(t) \]  

(21)

**Proof.** Since the linear transformation in (11) is independent, and \( h(s), g(t) \) are bounded below, it is easy to see the proposition follows. \( \square \)

Next, we will solve \( \min_s h(s) \) and \( \min_t g(t) \) separately.

For \( h(s) \), we can find that \( \forall s, h(s) = (1 + (s + 1)^2((\sqrt{3}s - \frac{7}{\sqrt{3}})^2 + \frac{8}{3})) > 0 \), and there is only one critical point \( s = -1 \) for \( h(s) \), that is to say, \( s^* = -1 \) is the global minimum of \( h(s) \).

For \( g(t) \), we rewrite it to the following canonical form

\[ g(t) = V(\Lambda(t)) - U(t) \]  

(22)

where, \( V(\Lambda(t)) = 3(t^2 - \frac{8}{3}t - 2) - 9(t^2 - \frac{8}{3}t - 2) \), and \( U(t) = -\frac{53}{3}t^2 + 56t \).

Introducing a nonlinear operator

\[ \xi = \Lambda(t) = t^2 - \frac{8}{3}t - 2 = (t - \frac{4}{3})^2 - \frac{34}{9} \geq -\frac{34}{9} \]  

(23)

then

\[ V(\xi) = 3\xi^2 - 9\xi, \varsigma = 6\xi - 9 \geq \frac{95}{3}, V^*(\varsigma) = \frac{(9 + \varsigma)^2}{12}, \]  

(24)
therefore, we get the generalized complementary function

$$g(\varsigma, t) = A(t)\varsigma - V^*(\varsigma) - U(t)$$

$$= (t^2 - \frac{8}{3}t - 2)\varsigma - \frac{\varsigma^2 + 18\varsigma + 81}{12} + \frac{53}{3}t^2 - 56t$$

$$= (\varsigma + \frac{53}{3})t^2 - (\frac{8}{3}\varsigma + 56)t - \frac{\varsigma^2 + 18\varsigma + 81}{12} - 2\varsigma, \quad (25)$$

For a given $\varsigma$, the criticality condition $g_t(\varsigma, t) = 0$ leads to the canonical equilibrium equation

$$2(\varsigma + \frac{53}{3})t = \frac{8}{3}\varsigma + 56. \quad (26)$$

Substituting $t = (\frac{8}{3}\varsigma + 56)/(2(\varsigma + \frac{53}{3}))$ into $g(\varsigma, t)$, we obtain the dual function

$$P^d(\varsigma) = \frac{1}{12}(-\varsigma^2 - 18\varsigma - 81) - \frac{(\frac{8}{3}\varsigma + 56)^2}{4(\varsigma + \frac{53}{3})} - 2\varsigma, \quad (27)$$

which is concave in the positive domain

$$S^+_u = \{\varsigma | \varsigma + \frac{53}{3} > 0\}. \quad (28)$$

The plots of $g(t)$ and $P^d(\varsigma)$ are illustrated in Fig. 2.

![Fig. 2 The primal and dual of g(t)](image)

Using the sequential quadratic programming method from the Optimization Toolbox within the MATLAB environment for $P^d(\varsigma)$ over $S^+_u$, we can get $\varsigma = -15$. Then we get the corresponding $t^* = (\frac{8}{3}\varsigma + 56)/(2(\varsigma + \frac{53}{3})) = 3$ by using the canonical equilibrium equation (26).
By the inverse linear transformation of (17), we can obtain the global minimum to \( f_1(x, y) \)

\[
\begin{pmatrix} x^* \\ y^* \end{pmatrix} = T^{-1}(s^*, t^*) = \begin{pmatrix} (3s^* + t)/5 \\ (2s^* - t)/5 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix},
\]

which is indeed the global minimum with the result given in [16].

4 Application for Three Hump Camel Back problem

The Three Hump Camel Back problem is given in the form of [17]:

\[
\min_{x, y} f_2(x, y) = 2x^2 - 1.05x^4 + \frac{x^6}{6} + xy + y^2.
\]

The landscape and contour of Three hump camel back function are given in Fig. 3, and we can find that there also exist a few extrema. The nonconvexity also makes it difficult to find the global minimum.

Firstly, we rewrite \( f_2(x, y) \) to the following canonical form

\[
f_2(x, y) = \frac{1}{6}(V_1(\Lambda_1(x, y)) - U_1(x, y)),
\]

where, \( V_1(\Lambda_1(x, y)) = (x^3 - 3.2x)^2 \), \( U_1(x, y) = -0.1x^4 - 1.76x^2 - 6xy - 6y^2 \).

Introducing a nonlinear operator

\[
\xi_1 = \Lambda_1(x) = x^3 - 3.2x,
\]

then

\[
V_1(\xi_1) = \xi_1^2, \quad \varsigma_1 = 2\xi_1, \quad V_1^*(\varsigma_1) = \frac{1}{4}\varsigma_1^2,
\]
thus, we can get the first generalized complementary function
\[
    f_2(c_1, x, y) = \frac{1}{6}(A_1(x)c_1 - V_1^*(c_1) - U_1(x, y))
\]
\[
    = \frac{1}{6}\left((x^3 - 3.2x)c_1 - \frac{1}{4}c_1^2 + 0.1x^4 + 1.76x^2 + 6xy + 6y^2\right). \quad (32)
\]

Again, the \( f_2(c_1, x, y) \) can be rewritten to
\[
    f_2(c_1, x, y) = \frac{1}{60} (V_2(A_2(x, y, c_1)) - U_2(x, y, c_1)), \quad (33)
\]
where, \( V_2(A_2(x, y, c_1)) = (x^2 + 5c_1x)^2, U_2(x, y, c_1) = (25c_1^2 - 17.6)x^2 + 32c_1x - 60xy - 60y^2 + 2.5c_1^2 \).

Then, we introduce another nonlinear operator
\[
    \xi_2 = \Lambda_2(x, c_1) = x^2 + 5c_1x,
\]
thus
\[
    V_2(\xi_2) = \xi_2^2, \quad \xi_2 = 2\xi_2, \quad V_2^*(\xi_2) = \frac{1}{4}\xi_2^2, \quad (35)
\]

consequently, we obtain the final generalized complementary function
\[
    f_2(c_1, c_2, x, y) = \frac{1}{60}(\Lambda_2(x, c_1)c_2 - V_2^*(c_2) - U_2(x, y, c_1))
\]
\[
    = \left(\frac{22}{75} - \frac{5}{12}c_1^2 + \frac{c_2}{60}\right)x^2 + y^2 + xy + \left(\frac{1}{12}c_1c_2 - \frac{8}{15}c_1\right)x - \frac{1}{24}c_1^2 - \frac{1}{240}c_2^2 \quad (36)
\]

For given \( c_1 \) and \( c_2 \), the criticality condition \( \nabla \Xi_1(c_1, c_2, x, y) = 0 \) leads to the following canonical equilibrium equations
\[
\begin{align*}
    \left\{ \begin{array}{l}
    2\left(\frac{22}{75} - \frac{5}{12}c_1^2 + \frac{c_2}{60}\right)x + y + \left(\frac{1}{12}c_1c_2 - \frac{8}{15}c_1\right) = 0, \\
    2y + x = 0
    \end{array} \right. \quad (37)
\end{align*}
\]

and finally we get the canonical dual function
\[
    P^d(\varsigma) = \frac{-1250c_1^4 - 50c_1^2(31c_2 - 105) + c_2^2(5c_2 + 13)}{240(125c_1^2 - 5c_2 - 13)}. \quad (38)
\]

which is concave in the positive domain
\[
    S^+_a = \left\{ \varsigma \left\lvert \begin{array}{c}
    \left(\frac{22}{75} - \frac{5}{12}c_1^2 + \frac{c_2}{60}\right) 0.5 \\
    \frac{1}{2}
    \end{array} \right. \geq 0 \right\}. \quad (39)
\]

Using the sequential quadratic programming method from the Optimization Toolbox within the MATLAB environment for \( P^d(\varsigma) \) over \( S^+_a \), we can get \( c_1 = 0, c_2 = 0 \). According to the canonical equilibrium equations [37], we can obtain the corresponding
\[
    \begin{pmatrix} x^* \\ y^* \end{pmatrix} = \left(\frac{2}{3}(\frac{22}{75} - \frac{5}{12}c_1^2 + \frac{c_2}{60}) 1\right)^{-1} \begin{pmatrix} 8c_1^2 - \frac{1}{12}c_1c_2 \\ 0 \end{pmatrix} = \left(\begin{array}{c}
    0 \\
    0
\end{array} \right),
\]
which is indeed the global minimum with the result given in [17].
5 Conclusion

When the canonical duality methodology is applied to a special class of polynomial optimization problem, it has the ability to solve the class of problem completely. On the other hand, for a general polynomial problem, we can also design appropriate canonical dual transformation to achieve the goal. As for Goldstein-Price problem, we transform it to decoupled minimization problems, and then solve them separately. While for Three hump camel back problem, we can utilize two-level canonical dual transformations. The completely solutions of the general polynomial benchmark functions have witnessed the powerfulness of the canonical duality methodology again.

References

1. Mevissen, M.: Introduction to concepts and advances in polynomial optimization. Tutorial at the ETH Zürich summer school New Algorithm Paradigms in Optimization (2008)
2. Tuy, H.: Polynomial optimization: a robust approach. Pacific Journal of Optimization. 1, 357–374 (2005)
3. Parrilo, P.A. and Sturmfels, B.: Minimizing polynomial functions, In Algorithmic and Quantitative Real Algebraic geometry. DIMACS Series in Discrete Mathematics and Theoretical Computer Science. 60, 83-99 (2003)
4. Kojima, M., Kim, S.Y. and Waki, H.: A general framework for convex relaxation of polynomial optimization problems over cones. Journal of the Operations Research. 46(2), 125-144 (2003)
5. Qi, L.Q., Wang, F. and Wang, Y.J.: Z-eigenvalue methods for a global polynomial optimization problem. Math. Program., Ser. A. 118, 301-316 (2009)
6. Fleming, P.J., Purshouse, R.C.: Evolutionary algorithms in control systems engineering: a survey. Control Engineering Practice. 10, 1223-1241 (2002)
7. Ashlock, D.: Evolutionary computation for modeling and optimization. Springer-Verlag, New York (2006)
8. Kelly, C.T.: Iterative methods for optimization. SIAM publications, Philadelphia (1999)
9. Hendrix, E.M.T., Toth, B.G.: Introduction to Nonlinear and Global Optimization. Springer-Verlag, New York, NY USA (2010)
10. Ho, S.H., Shu, L.S. and Chen, J.H.: Intelligent evolutionary algorithms for large parameter optimization problems. IEEE Transaction on evolutionary computation. 8(6), 522-541 (2004)
11. Gao, D.Y. and Strang G.: Geometric nonlinearity: potential energy, complementary energy, and the gap function. Quarterly journal of applied mathematics. XLVII(3), 487-504 (1989)
12. Gao, D.Y.: Canonical duality theory: Unified understanding and generalized solution for global optimization problems. Computers and Chemical Engineering. 33, 1964-1972 (2009)
13. Gao, D.Y.: Complete solutions and extremality criteria to polynomial optimization problems. Journal of Global Optimization. 35, 131-143 (2006)
14. Gao, K.T.: Solutions to 8th order polynomial minimization problem. The Electronic Journal of Mathematics and Technology. 1(3), 271-276 (2007)
15. Gao, D.Y., Ruan, N. and Pardalos, P.M.: Canonical dual solutions to sum of fourth-order polynomials minimization problems with applications to sensor network localization. in Sensors: Theory, Algorithms and Applications. 61(1), 37-54 (2012)
16. Molga, M., and Smutnicki, C.: Test functions for optimization needs. [http://www.zsl.ict.pwr.wroc.pl/files/docs/functions.pdf] (2005)
17. Mishra, S.K.: Some new test functions for global optimization and performance of repulsive particle swarm method. Social Science Research Network (SSRN) Working Papers Series, [http://ssrn.com/abstract=927134] (2006)