J Glob Optim (2010) 47:583–595
DOI 10.1007/s10898-009-9480-5

Global optimality conditions for cubic minimization problem with box or binary constraints

Yanjun Wang · Zhian Liang

Received: 3 August 2008 / Accepted: 19 October 2009 / Published online: 4 November 2009

Abstract In this article, we provide optimality conditions for global solutions to cubic minimization problems with box or binary constraints. Our main tool is an extension of the global subdifferential approach, developed by Jeyakumar et al. (J Glob Optim 36:471–481, 2007; Math Program A 110:521–541, 2007). We also derive optimality conditions that characterize global solutions completely in the case where the cubic objective function contains no cross terms. Examples are given to demonstrate that the optimality conditions can effectively be used for identifying global minimizers of certain cubic minimization problems with box or binary constraints.

Keywords Cubic minimization problem · Global optimality conditions · Box or binary constraints

1 Introduction

Consider the following cubic minimization problem with box constraints:

$$CP : \begin{cases} \min f(x) = \sum_{i=1}^{n} \frac{1}{3}\beta_i x_i^3 + \frac{1}{2}x^T Ax + b^T x \\ \text{s.t. } x \in X = \prod_{i=1}^{n}[u_i, v_i]. \end{cases}$$

where $x = (x_1, \ldots, x_n)^T$ is the vector of decision variables, $\beta_i, u_i, v_i \in R$, $u_i \leq v_i$, $b \in R^n$, $A \in S^n$ and S^n is the set of all symmetric $n \times n$ matrices.

The cubic optimization problem has spawned a variety of applications, especially in cubic polynomial approximation optimization [1], convex optimization [2], engineering design, and structural optimization [3]. Moreover, research results about cubic optimization problem can be applied to quadratic programming problems, which have been widely studied because of its broad applications, to enrich quadratic programming theory.

Y. Wang (✉) · Z. Liang
Department of Applied Mathematics, Shanghai University of Finance and Economics, Shanghai 200433, People's Republic of China
e-mail: mwangyj@hotmail.com

There are several general approaches used to establish optimality conditions for solutions to optimization problems. These approaches can be broadly classified into three groups: convex duality theory [4], local sub-differentials by linear functions [5–7] and global L-subdifferentials by quadratic functions [8–11]. The third approach, which we extend in this article, is often adopted to develop optimality conditions for special optimization forms: quadratic minimization with box or binary constraints, quadratic minimization with quadratic constraints, and bivalent quadratic minimization with inequality constraints, etc.

In this article, we focus our attention on more general problems, i.e., cubic minimization problems with box or binary constraints. Our main tool is an extension of global L-subdifferential approach from quadratic function to cubic function forms. By exploring some fundamental properties of the problems, we establish sufficient conditions under which a feasible point will be a global solution to CP. In particular, we present optimality conditions that characterize global solutions completely in the case where the cubic objective function contains no cross terms.

The layout of this paper is as follows. In Sect. 2, we extend global L-subdifferential approach and present sufficient optimality conditions for global solutions to CP. We also develop optimality conditions which characterize global solutions completely in the case where the cubic objective function contains no cross terms. In Sect. 3, we provide global optimality conditions for the cubic minimization problem subject to binary constraints. Finally, examples are given to show the effectiveness of the proposed global optimality conditions in Sect. 4.

2 L-subdifferentials and cubic minimization problem

We begin with basic definitions and notations that will be used throughout the article. The real line is denoted by R and the n-dimensional Euclidean space is denoted by R^n. The notation $A \succeq B$ means that the matrix $A - B$ is positive semi definite. A diagonal matrix with diagonal elements $c_1, c_2, \ldots, c_n$ is denoted by $diag(c_1, \ldots, c_n)$. Let L be a set of real-valued functions defined on R^n.

L-subdifferentials: Let $f : R^n \to R$ and $x_0 \in R^n$. An element $l \in L$ is called an L-subgradient of f at a point x_0 if

$$f(x) \geq f(x_0) + l(x) - l(x_0), \forall x \in R^n$$

The set $\partial_L f(x)$ of all L-subgradients of f at x_0 is referred as L-subdifferential of f at x_0.

Note that if L is the set of all linear functions defined on R^n, then for any real-valued convex function f defined on R^n, $\partial_L f(x) = \partial f(x)$, where $\partial f(x)$ is the subdifferential in the sense of convex analysis [12].

Throughout the rest of the article, we use the following specific choice of L defined by

$$L := \left\{ \sum_{i=1}^{n} \frac{1}{3}\beta_i x_i^3 + \frac{1}{2}x^T Qx + d^T x | Q = diag(c_1, \ldots, c_n), c_i \in R, d \in R^n \right\},$$

where $\beta_i \in R$ $(i = 1, \ldots, n)$ as before. The following result shows that, for the specific choice of L, the representation of $\partial_L f(x)$ ($f(x)$ is not necessarily convex) can be calculated explicitly.

Proposition 1 *Let* $f(x)=\sum_{i=1}^{n}\frac{1}{3}\beta_i x_i^3+\frac{1}{2}x^TAx+b^Tx$, $\bar{x}=(\bar{x}_1,\ldots,\bar{x}_n)^T\in R^n$. *Then,*

$$\partial_L f(\bar{x})=\left\{\sum_{i=1}^{n}\frac{1}{3}\beta_i x_i^3+\frac{1}{2}x^TQx+d^Tx\,\middle|\,\begin{array}{c}A\succeq Q,\ Q=diag(c_1,\ldots,c_n),\ c_i\in R,\\ d=(A-Q)\bar{x}+b.\end{array}\right\} \quad (1)$$

Proof Let $l_0(x)\in L$, i.e., $l_0(x)$ can be written as

$$l_0(x)=\sum_{i=1}^{n}\frac{1}{3}\beta_i x_i^3+\frac{1}{2}x^TQx+d^Tx,$$

where $Q=diag(c_1,\ldots,c_n)$, $c_i\in R$, $d\in R^n$.

If $l_0(x)\in\partial_L f(\bar{x})$, then

$$l_0(x)-l_0(\bar{x})\le f(x)-f(\bar{x}),\quad \forall x\in R^n,$$

i.e.,

$$\phi(x):=f(x)-l_0(x)=\frac{1}{2}x^TAx+b^Tx-\left(\frac{1}{2}x^TQx+d^Tx\right)\ge f(\bar{x})-l_0(\bar{x}).$$

Thus, $\phi(x)$ is bounded below and attains its minimum at $\bar{x}$. According to necessary optimality condition for unconstrained optimization problem, we have $\nabla\phi(\bar{x})=0$, $\nabla^2\phi(\bar{x})\succeq 0$, i.e., $A\succeq Q$, $d=(A-Q)\bar{x}+b$, which shows that

$$\partial_L f(\bar{x})\subset\left\{\sum_{i=1}^{n}\frac{1}{3}\beta_i x_i^3+\frac{1}{2}x^TQx+d^Tx\,\middle|\,\begin{array}{c}A\succeq Q,\ Q=diag(c_1,\ldots,c_n),\ c_i\in R,\\ d=(A-Q)\bar{x}+b.\end{array}\right\}.$$

Now, we show that converse inclusion is also valid. Suppose that $l_0(x)$ belongs to the set in right-hand side of (1), i.e., $l_0(x)$ has an expression as

$$l_0(x)=\sum_{i=1}^{n}\frac{1}{3}\beta_i x_i^3+\frac{1}{2}x^TQx+d^Tx,$$

where $A\succeq Q$, $d=(A-Q)\bar{x}+b$. Then, let

$$\phi(x):=f(x)-l_0(x)=\frac{1}{2}x^T(A-Q)x+(b-d)^Tx.$$

Since,

$$\nabla\phi(\bar{x})=(A-Q)\bar{x}+(b-d)=0,\quad \nabla^2\phi(\bar{x})=A-Q\succeq 0,$$

$\phi(x)$ is a convex function on R^n, and $\phi(x)$ attains its global minimum at $\bar{x}$. That means

$$\phi(x)=f(x)-l_0(x)\ge f(\bar{x})-l_0(\bar{x}),\quad \forall x\in R^n.$$

Hence, we have $l_0(x)\in\partial_L f(\bar{x})$. □

For $\bar{x}=(\bar{x}_1,\ldots,\bar{x}_n)^T\in X$, suppose that there exists a matrix $Q=diag(c_1,\ldots,c_n)$, $c_i\in R$, such that $A-Q\succeq 0$, and let $d=(A-Q)\bar{x}+b$. Then, by Proposition 1, we have

$$l(x)=\sum_{i=1}^{n}\frac{1}{3}\beta_i x_i^3+\frac{1}{2}x^TQx+d^Tx\in\partial_L f(\bar{x}).$$

For deriving optimality conditions of global solutions to CP, we present some definitions and notations. Define

$$p_i(x_i) = \frac{1}{3}\beta_i(x_i - \bar{x}_i)^2 + \frac{1}{2}c_i(x_i - \bar{x}_i) + (d_i + c_i\bar{x}_i + \beta_i x_i\bar{x}_i),$$
$$t_i = -\frac{1}{2}\bar{x}_i - \frac{3c_i}{4\beta_i}.$$

We observe that $p_i(x_i)$ is a quadratic function of x_i which attains the minimum (if $\beta_i > 0$) or maximum (if $\beta_i < 0$) at the point t_i.

Let $I = \{i|\bar{x}_i = u_i \text{ or } v_i\}$, $J = \{i|\bar{x}_i \in (u_i, v_i)\}$. For $i \in I$, we present the following definitions:

$$\breve{x}_i = \begin{cases} v_i & \text{if } \bar{x}_i = u_i, \\ u_i & \text{if } \bar{x}_i = v_i, \end{cases}$$

$$\tilde{x}_i = \begin{cases} \bar{x}_i & \text{if } t_i \le u_i \text{and } \beta_i > 0, \\ \breve{x}_i & \text{if } t_i \ge v_i \text{ and } \beta_i > 0, \\ t_i & \text{if } t_i \in (u_i, v_i) \text{and } \beta_i > 0, \\ \breve{x}_i & \text{if } t_i \le u_i \text{ and } \beta_i < 0, \\ \bar{x}_i & \text{if } t_i \ge v_i \text{ and } \beta_i < 0, \\ t_i & \text{if } t_i \in (u_i, v_i) \text{ and } \beta_i < 0, \\ \bar{x}_i & \text{if } \beta_i = 0, \end{cases}$$

$$\gamma_i = \begin{cases} 1 & \text{if } \bar{x}_i = u_i, \\ -1 & \text{if } \bar{x}_i = v_i, \end{cases}$$

$$\omega_i = \begin{cases} 0 & \text{if } t_i \in (u_i, v_i), \beta_i < 0, \bar{x}_i = u_i, \\ 0 & \text{if } t_i \in (u_i, v_i), \beta_i > 0, \bar{x}_i = v_i, \\ 0 & \text{if } \beta_i = 0, \\ 1 & \text{else}, \end{cases}$$

$$\tilde{\tau}_i = \gamma_i\omega_i,$$

and

$$\tau_i = \begin{cases} 1 & \text{if } t_i \in (u_i, v_i), \beta_i < 0, \bar{x}_i = u_i, \\ 1 & \text{if } t_i \in (u_i, v_i), \beta_i > 0, \bar{x}_i = v_i, \\ 1 & \text{if } \beta_i = 0, \\ 0 & \text{else.} \end{cases}$$

For $i \in J$, we define

$$\hat{x}_i = \begin{cases} \bar{x}_i & \text{if } t_i \le u_i \text{ and } \beta_i > 0, \\ u_i & \text{if } t_i \in (u_i, \bar{x}_i] \text{ and } \beta_i > 0, \\ \bar{x}_i & \text{if } t_i \ge v_i \text{ and } \beta_i < 0, \\ v_i & \text{if } t_i \in [\bar{x}_i, v_i) \text{ and } \beta_i < 0, \\ \bar{x}_i & \text{if } c_i \ge 0 \text{ and } \beta_i = 0, \end{cases}$$

and

$$\hat{\tau}_i = \begin{cases} 1 & \text{if } t_i \in (u_i, \bar{x}_i] \text{ and } \beta_i > 0, \\ -1 & \text{if } t_i \in [\bar{x}_i, v_i) \text{ and } \beta_i < 0, \\ 0 & else. \end{cases}$$

In this section, we make the following assumption at $\bar{x}$.

Assumption 1 For each $i \in J$, the following five cases do not happen

$$\begin{cases} (1) & t_i \geq v_i, \beta_i > 0, \\ (2) & t_i \in (\bar{x}_i, v_i), \beta_i > 0, \\ (3) & t_i \leq u_i, \beta_i < 0, \\ (4) & t_i \in (u_i, \bar{x}_i), \beta_i < 0, \\ (5) & c_i < 0, \beta_i = 0. \end{cases}$$

Note that, for $\hat{x}_i$, there is no definition in the five cases which appears in Assumption 1. Then, only under Assumption 1 we can establish sufficient optimality conditions for global solutions to CP.

Theorem 2 (Sufficient global optimality conditions) *Let* $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)^T \in X$ *satisfy Assumption 1. Suppose that there exists a diagonal matrix*

$$Q = diag(c_1, \ldots, c_n), c_i \in R, \text{ such that } A - Q \succeq 0.$$

If it holds that

$$\tilde{\tau}_i p_i(\tilde{x}_i) + \tau_i \min\{\gamma_i p_i(u_i), \gamma_i p_i(v_i)\} \geq 0, \text{ for each } i \in I, \tag{2}$$

$$\hat{\tau}_i p_i(\hat{x}_i) \leq 0 \text{ and } p_i(\bar{x}_i) = 0, \text{ for each } i \in J, \tag{3}$$

then $\bar{x}$ *is a global minimizer of the problem CP.*

Proof Let $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)^T \in X$ satisfy Assumption 1. Suppose that there exists a diagonal matrix

$$Q = diag(c_1, \ldots, c_n), c_i \in R, \text{ such that } A - Q \succeq 0.$$

Let $d = b + (A - Q)\bar{x}$. By Proposition 1, we immediately have that

$$l(x) = \sum_{i=1}^{n} \frac{1}{3}\beta_i x_i^3 + \frac{1}{2}x^T Q x + d^T x \in \partial_L f(\bar{x}),$$

i.e.,

$$f(x) - f(\bar{x}) \geq l(x) - l(\bar{x}) \ \ \forall x \in R^n.$$

Obviously, it suffices to prove that $\bar{x}$ is a global solution of $l(x)$ over X.

Note that

$$l(x) - l(\bar{x}) = \sum_{i=1}^{n}\{\frac{1}{3}\beta_i(x_i - \bar{x}_i)^3 + \frac{1}{2}c_i(x_i - \bar{x}_i)^2 + (d_i + c_i\bar{x}_i + \beta_i x_i \bar{x}_i)(x_i - \bar{x}_i)\}.$$

Therefore, $\bar{x}$ is a global minimizer of $l(x)$ over X if and only if for each $i = 1, \ldots, n$,

$$\frac{1}{3}\beta_i(x_i - \bar{x}_i)^3 + \frac{1}{2}c_i(x_i - \bar{x}_i)^2 + (d_i + c_i\bar{x}_i + \beta_i x_i \bar{x}_i)(x_i - \bar{x}_i) \geq 0, \text{ for all } x_i \in [u_i, v_i]. \tag{4}$$

Indeed, if (4) holds for each $i = 1, \ldots, n$, it is evident that $l(x) - l(\bar{x}) \geq 0$ for all $x \in X$. Conversely, suppose that $\bar{x}$ is a global minimizer. If there exist i_0 and x_{i_0} such that (4) is not fulfilled, then by taking $x^* = (x_1^*, \ldots, x_n^*)^T$ such that $x_i^* = \bar{x}_i, i \neq i_0$ and $x_{i_0}^* = x_{i_0}$, then

$$l(x^*) - l(\bar{x}) = \frac{1}{3}\beta_{i_0}(x_{i_0} - \bar{x}_{i_0})^3 + \frac{1}{2}c_{i_0}(x_{i_0} - \bar{x}_{i_0})^2 + (d_{i_0} + c_{i_0}\bar{x}_{i_0} + \beta_{i_0}x_{i_0}\bar{x}_{i_0})(x_{i_0} - \bar{x}_{i_0}) < 0,$$

which contradicts the fact that $\bar{x}$ is a global minimizer.

From the above analysis, it is sufficient to prove for each $i = 1, \ldots, n$, under conditions (2) or (3), (4) holds. We now consider the following three cases.

Case 1 If $\bar{x}_i = u_i$, then $i \in I$. Notice that in this case $x_i - \bar{x}_i \geq 0$ for all $x_i \in [u_i, v_i]$. Thus, for proving (4), it suffices to prove, under condition (2),

$$p_i(x_i) = \frac{1}{3}\beta_i(x_i - \bar{x}_i)^2 + \frac{1}{2}c_i(x_i - \bar{x}_i) + (d_i + c_i\bar{x}_i + \beta_i x_i \bar{x}_i) \geq 0, \quad \text{for all } x_i \in [u_i, v_i]. \tag{5}$$

Notice that in (2) the value of $\tilde{x}_i$ is defined according to the values of β_i and t_i. And there are seven cases in the definition. Here, we only consider the first case, i.e., $\beta_i > 0$ and $t_i \leq u_i$. Then, it follows that $\tilde{x}_i = u_i$, $\gamma_i = 1$, $\tilde{\tau}_i = 1$, and $\tau_i = 0$. Substituting these values into condition (2), we have $p_i(u_i) \geq 0$.

Notice that the quadratic function $p_i(x_i)$ will attain the minimum at the point t_i in the case of $\beta_i > 0$. By combining the fact that $t_i \leq u_i$, we can see that $p_i(x_i)$ is monotonically increasing over $[u_i, v_i]$. Since $p_i(u_i) \geq 0$, then (5) holds.

The other cases in the definition of $\tilde{x}_i$ can be considered analogously.

Case 2 If $\bar{x}_i = v_i$, then $i \in I$. Notice that in this case $x_i - \bar{x}_i \leq 0$ for all $x_i \in [u_i, v_i]$. Thus, for proving (4), it suffices to prove $p_i(x_i) \leq 0$ for all $x_i \in [u_i, v_i]$ under condition (2). Similarly, we consider only the case in the definition of $\tilde{x}_i$: $\beta_i > 0$ and $t_i \leq u_i$. It follows that $\tilde{x}_i = v_i$, $\gamma_i = -1$, $\tilde{\tau}_i = -1$, and $\tau_i = 0$. Substituting these values into condition (2), we have $p_i(v_i) \leq 0$. From the analysis in case 1, when $\beta_i > 0$ and $t_i \leq u_i$, $p_i(x_i)$ is monotonically increasing over $[u_i, v_i]$. Since $p_i(v_i) \leq 0$, then $p_i(x_i) \leq 0$ for all $x_i \in [u_i, v_i]$.

The other cases in the definition of $\tilde{x}_i$ can be considered analogously.

Case 3 If $\bar{x}_i \in (u_i, v_i)$, then $i \in J$. For proving (4), it suffices to prove, under the condition (3),

$$\begin{cases} p_i(x_i) \leq 0, & \text{for } x_i \in [u_i, \bar{x}_i), \\ p_i(x_i) \geq 0, & \text{for } x_i \in (\bar{x}_i, v_i]. \end{cases} \tag{6}$$

Similarly, we consider only the case in the definition of $\hat{x}_i$: $\beta_i > 0$ and $t_i \leq u_i$. It follows that $\hat{x}_i = \bar{x}_i$ and $\hat{\tau}_i = 0$. Substituting these values into condition (3), we have $p_i(\bar{x}_i) = 0$. From the analysis in case 1, when $\beta_i > 0$ and $t_i \leq u_i$, $p_i(x_i)$ is monotonically increasing over $[u_i, v_i]$. Since $p_i(\bar{x}_i) = 0$, it is obvious that (6) holds.

The other cases under which $\hat{x}_i$ has a definition can be considered analogously.

Combining the above three cases yields the desired result. □

We now consider a special class of cubic minimization problems where the cubic objective function in CP contains no cross terms. The problem has the forms as

$$(CP_0): \begin{cases} \min f_0(x) = \sum_{i=1}^{n}(\frac{1}{3}\beta_i x_i^3 + \frac{a_i}{2}x_i^2 + b_i x_i) \\ s.t. \ \ x \in X = \prod_{i=1}^{n}[u_i, v_i]. \end{cases}$$

We will present complete characterization of global optimality for the above problem. Define

$$q_i(x_i) = \frac{1}{3}\beta_i(x_i - \bar{x}_i)^2 + \frac{1}{2}a_i(x_i - \bar{x}_i) + (b_i + a_i\bar{x}_i + \beta_i x_i \bar{x}_i).$$

Notice that $q_i(x_i)$ has similar form with $p_i(x_i)$. In $p_i(x_i)$, replacing c_i by a_i yields $q_i(x_i)$. In addition, we need make some definitions of $\tilde{\tau}_i$, τ_i, $\tilde{x}_i$, γ_i, $\hat{\tau}_i$ and $\hat{x}_i$, which are similar to the previous ones. However, only note the following two points:

1. in the definition of $\hat{x}_i$ and the case (5) in Assumption 1, c_i should be replaced by a_i.
2. t_i , which appears in all the definitions and Assumption 1, should be rewritten as

$$t_i = -\frac{1}{2}\bar{x}_i - \frac{3a_i}{4\beta_i}.$$

Corollary 3 *For* (CP_0)*, let* $\bar{x} \in X$ *satisfy Assumption 1. Then* $\bar{x}$*, is a global minimizer of* (CP_0) *if and only if,*

$$\tilde{\tau}_i q_i(\tilde{x}_i) + \tau_i \min\{\gamma_i q_i(u_i), \gamma_i q_i(v_i)\} \geq 0,\ \textit{for}\ i \in I, \tag{7}$$

$$\hat{\tau}_i q_i(\hat{x}_i) \leq 0\ \textit{ and }\ q_i(\bar{x}_i) = 0,\ \ \textit{for}\ i \in J. \tag{8}$$

Proof Let $\bar{x}$ satisfy Assumption 1. By definition, $\bar{x}$ is a global minimizer of (CP_0) if and only if

$$f_0(x) - f_0(\bar{x}) = \sum_{i=1}^{n} \left[\frac{1}{3}\beta_i(x_i - \bar{x}_i)^3 + \frac{1}{2}a_i(x_i - \bar{x}_i)^2 + (b_i + a_i\bar{x}_i + \beta_i x_i \bar{x}_i)(x_i - \bar{x}_i)\right] \geq 0,\ \text{for } x \in X. \tag{9}$$

From the proof of Theorem 2, we can derive (9) is equivalent to for each $i = 1, \ldots, n$,

$$\frac{1}{3}\beta_i(x_i - \bar{x}_i)^3 + \frac{1}{2}a_i(x_i - \bar{x}_i)^2 + (b_i + a_i\bar{x}_i + \beta_i x_i\bar{x}_i)(x_i - \bar{x}_i) \geq 0,\ \text{for all } x_i \in [u_i, v_i]. \tag{10}$$

Therefore, it suffices to prove, for each $i = 1, \ldots, n$, (7) or (8) holds if and only if (10) holds.

We now consider the following three cases.

Case 1 If $\bar{x}_i = u_i$, then $i \in I$. It suffices to prove (7) holds if and only if (10) holds. We only consider the case in the definition of $\tilde{x}_i$: $\beta_i > 0$ and $t_i \leq u_i$. Since $\beta_i > 0$ and $t_i \leq u_i$, we have $\tilde{x}_i = u_i$, $\gamma_i = 1$, $\tilde{\tau}_i = 1$ and $\tau_i = 0$. Substituting these values into (7) yields

$$q_i(u_i) \geq 0. \tag{11}$$

Therefore, it suffices to prove (10) and (11) are equivalent.

Note that $q_i(x_i)$ is monotonically increasing over $[u_i, v_i]$ in the case of $\beta_i > 0$ and $t_i \leq u_i$. Therefore, if (11) holds, it follows that $q_i(x_i) \geq 0$ for all $x_i \in [u_i, v_i]$. Noting that $x_i - \bar{x}_i \geq 0$ for all $x_i \in [u_i, v_i]$, we have (10) holds.

Conversely, if (10) holds, then it follows that

$$q_i(x_i) = \frac{1}{3}\beta_i(x_i - \bar{x}_i)^2 + \frac{1}{2}a_i(x_i - \bar{x}_i) + (b_i + a_i\bar{x}_i + \beta_i x_i\bar{x}_i) \geq 0,\ \text{for all } x_i \in (u_i, v_i]$$

by dividing (10) by $x_i - \bar{x}_i$ $(x_i \in (u_i, v_i])$. By combining the fact that $q_i(x_i)$ is continuous on R, we can derive (11).

The other cases in the definition of $\tilde{x}_i$ can be considered analogously to show the equivalence of (7) and (10).

Case 2 If $\bar{x}_i = v_i$, then $i \in I$. It suffices to prove (7) holds if and only if (10) holds. Similarly, we consider only the case in the definition of $\tilde{x}_i$: $\beta_i > 0$ and $t_i \leq u_i$. Since $\beta_i > 0$ and $t_i \leq u_i$, we have $\tilde{x}_i = v_i$, $\gamma_i = -1$, $\tilde{\tau}_i = -1$ and $\tau_i = 0$. Substituting these values into (7) yields

$$q_i(v_i) \leq 0. \tag{12}$$

Therefore, it suffices to prove (10) and (12) are equivalent.

Note that $q_i(x_i)$ is monotonically increasing over $[u_i, v_i]$ in the case of $\beta_i > 0$ and $t_i \leq u_i$. Therefore, if (12) holds, it follows that $q_i(x_i) \leq 0$ for all $x_i \in [u_i, v_i]$. Noting that $x_i - \bar{x}_i \leq 0$ for all $x_i \in [u_i, v_i]$, we have (10) holds.

Conversely, if (10) holds, then it follows that

$$q_i(x_i) = \frac{1}{3}\beta_i(x_i - \bar{x}_i)^2 + \frac{1}{2}a_i(x_i - \bar{x}_i) + (b_i + a_i\bar{x}_i + \beta_i x_i \bar{x}_i) \leq 0, \text{ for all } \ x_i \in (u_i, v_i]$$

by dividing (10) by $x_i - \bar{x}_i$ ($x_i \in [u_i, v_i)$). By combining the fact that $q_i(x_i)$ is continuous on R, we can derive (12).

The other cases in the definition of $\tilde{x}_i$ can be considered analogously to show the equivalence of (7) and (10).

Case 3 If $\bar{x}_i \in (u_i, v_i)$, then $i \in J$. It suffices to prove (8) holds if and only if (10) holds. Similarly, we consider only the case in the definition of $\hat{x}_i$: $\beta_i > 0$ and $t_i \leq u_i$. Since $\beta_i > 0$ and $t_i \leq u_i$, we have $\hat{x}_i = \bar{x}_i$, and $\hat{\tau}_i = 0$. Substituting these values into (8) yields

$$q_i(\bar{x}_i) = 0. \tag{13}$$

Therefore, it suffices to prove (10) and (13) are equivalent.

Note that $q_i(x_i)$ is monotonically increasing over $[u_i, v_i]$ in the case of $\beta_i > 0$ and $t_i \leq u_i$. Therefore, if (13) holds, it follows that

$$\begin{cases} q_i(x_i) \leq 0, & \text{for } \ x_i \in [u_i, \bar{x}_i], \\ q_i(x_i) \geq 0, & \text{for } \ x_i \in [\bar{x}_i, v_i]. \end{cases}$$

Noticing also that $x_i - \bar{x}_i \leq 0$ for all $x_i \in [u_i, \bar{x}_i]$ and $x_i - \bar{x}_i \geq 0$ for all $x_i \in [\bar{x}_i, v_i]$, we have (10) holds.

Conversely, if (10) holds, then it follows that

$$\begin{cases} q_i(x_i) \leq 0 & \text{for } \ x_i \in [u_i, \bar{x}_i), \\ q_i(x_i) \geq 0 & \text{for } \ x_i \in (\bar{x}_i, v_i]. \end{cases} \tag{14}$$

Combining the fact that $q_i(x_i)$ is continuous on R, we can derive (13).

The other cases under which $\hat{x}_i$ has a definition can be considered analogously to show the equivalence of (8) and (10).

Combining the above three cases leads to the desired result. □

In the following corollary, we will see that Assumption 1 plays an important role in necessary global optimality conditions for (CP_0).

Corollary 4 (Necessary global optimality condition) *For (CP_0), if $\bar{x}$ is a global solution, then Assumption 1 holds.*

Proof Let $\bar{x}$ is global solution of (CP_0). Suppose that Assumption 1 does not hold. Then there exists some $i \in J = \{i|\bar{x}_i \in (u_i, v_i)\}$ such that at least one case in Assumption 1 happens. For instance, case (1) happens, i.e., $t_i \geq v_i$, $\beta_i > 0$. Notice that in this case $q_i(x_i)$ is decreasing over $[u_i, v_i]$. Therefore, it is evident that condition (14) is not satisfied. This implies condition (10) does not hold. Since $\bar{x}$ is global solution of (CP_0) if and only if (10) holds, then $\bar{x}$ is not global solution of (CP_0), which is a contradiction. For other four cases in Assumption 1, we can also get a contradiction analogously. Thus, the desired result follows. □

3 Bivalent cubic minimization problem

In this section, we will apply the method, developed in Sect. 2, to bivalent cubic minimization problem of the following form

$$(BCP): \begin{cases} \min \sum_{i=1}^{n} \frac{1}{3}\beta_i x_i^3 + \frac{1}{2}x^T Ax + b^T x \\ \text{s.t. } \ x \in \prod_{i=1}^{n}\{u_i, v_i\}. \end{cases}$$

where $u_i, v_i, \beta_i \in R$, $u_i < v_i$, $i = 1, 2, \ldots, n$, $b \in R^n$ and $A \in S^n$. Let $S_B^i := \{u_i, v_i\}$, $i = 1, \ldots, n$, and let $S_B := \prod_{i=1}^{n}\{u_i, v_i\}$. Using the same line of arguments as in the proof of Theorem 2, we derive sufficient global optimality condition for BCP.

Theorem 5 *Let $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)^T \in S_B$. Suppose that there exists a diagonal matrix $Q := diag(c_1, \ldots, c_n)$, $c_i \in R$, $i = 1, \ldots, n$, such that $A - Q \succeq 0$, and for each $i = 1, \ldots, n$,*

$$\frac{1}{3}\gamma_i\beta_i(v_i - u_i)^2 + \frac{c_i}{2}(v_i - u_i) + \gamma_i(d_i + c_i\bar{x}_i + \beta_i v_i u_i) \geq 0 \tag{15}$$

Then $\bar{x}$ is a global minimizer of BCP.

Proof Let $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n)^T \in S_B$, $Q = diag(c_1, \ldots, c_n)$ such that $A - Q \succeq 0$, and $d = b + (A - Q)\bar{x}$. By Proposition 1, we immediately have

$$l(x) = \sum_{i=1}^{n} \frac{1}{3}\beta_i x_i^3 + \frac{1}{2}x^T Qx + d^T x \in \partial_L f(\bar{x}),$$

i.e.

$$f(x) - f(\bar{x}) \geq l(x) - l(\bar{x}) \ \ \forall x \in R^n.$$

Thus, it suffices to prove $l(x) - l(\bar{x}) \geq 0$ for all $x \in S_B$, i.e., $\bar{x}$ is a global minimizer of $l(x)$ over S_B. Note that

$$l(x) - l(\bar{x}) = \sum_{i=1}^{n}\{\frac{1}{3}\beta_i(x_i - \bar{x}_i)^3 + \frac{1}{2}c_i(x_i - \bar{x}_i)^2 + (d_i + c_i\bar{x}_i + \beta_i x_i\bar{x}_i)(x_i - \bar{x}_i)\}.$$

Thus, from the proof of Theory 2, we have $\bar{x}$ is a global minimizer of $l(x)$ over S_B if and only if for each $i = 1, \ldots, n$, for each $x_i \in \{u_i, v_i\}$,

$$\frac{1}{3}\beta_i(x_i - \bar{x}_i)^3 + \frac{1}{2}c_i(x_i - \bar{x}_i)^2 + (d_i + c_i\bar{x}_i + \beta_i x_i\bar{x}_i)(x_i - \bar{x}_i) \geq 0. \tag{16}$$

Therefore, it suffices to prove, for each $i = 1, \ldots, n$, if condition (15) holds, then for each $x_i \in \{u_i, v_i\}$ (16) holds.

We now consider the following two cases:

Case 1 If $\bar{x}_i = u_i$, then $\gamma_i = 1$. Substituting this value into (15) leads to

$$\frac{1}{3}\beta_i(v_i - u_i)^2 + \frac{1}{2}c_i(v_i - u_i) + (d_i + c_i u_i + \beta_i v_i u_i) \geq 0. \tag{17}$$

Multiplying (17) by $(v_i - u_i)$ and replacing u_i by $\bar{x}_i$, we have (16) holds for $x_i = v_i$. Note that for $x_i = u_i$, (16) holds obviously. Therefore, it follows that for each $x_i \in \{u_i, v_i\}$, (16) holds.

Case 2 If $\bar{x}_i = v_i$, then $\gamma_i = -1$. Substituting this value into (15) leads to

$$-\frac{1}{3}\beta_i(v_i - u_i)^2 + \frac{1}{2}c_i(v_i - u_i) - (d_i + c_i v_i + \beta_i v_i u_i) \geq 0 \tag{18}$$

Multiplying (18) by $(v_i - u_i)$ and replacing v_i by $\bar{x}_i$, we have (16) holds for $x_i = u_i$. Note that, for $x_i = v_i$, (16) holds obviously. Therefore, it follows that, for each $x_i \in \{u_i, v_i\}$, (16) holds.

Combining the above two cases yields the desired result. □

Next, we will consider a special case of BCP which has the form as

$$(BCP_0): \begin{cases} \min f_0(x) = \sum_{i=1}^n \frac{1}{3}\beta_i x_i^3 + \sum_{i=1}^n \frac{a_i}{2} x_i^2 + \sum_{i=1}^n b_i x_i \\ s.t. \;\; x \in S_B = \prod_{i=1}^n \{u_i, v_i\}. \end{cases}$$

By exploring the properties of (BCP_0), we present optimality conditions which completely characterize global solutions to (BCP_0) in the following corollary.

Corollary 6 *Let $\bar{x} \in S_B$. Then $\bar{x}$, is a global minimizer of (BCP_0) if and only if for each $i = 1, \ldots, n$,*

$$\frac{1}{3}\gamma_i\beta_i(v_i - u_i)^2 + \frac{a_i}{2}(v_i - u_i) + \gamma_i(b_i + a_i\bar{x}_i + \beta_i v_i u_i) \geq 0. \tag{19}$$

Proof Let $\bar{x} \in S_B$. By definition, $\bar{x}$ is a global minimizer of (BCP_0) if and only if, for each $x \in S_B$,

$$f_0(x) - f_0(\bar{x}) = \sum_{i=1}^n [\frac{1}{3}\beta_i(x_i - \bar{x}_i)^3 + \frac{1}{2}a_i(x_i - \bar{x}_i)^2 + (b_i + a_i\bar{x}_i + \beta_i x_i \bar{x}_i)(x_i - \bar{x}_i)] \geq 0. \tag{20}$$

Note that (20) holds for all $x \in S_B$ if and only if for each $i = 1, \ldots, n$, for each $x_i \in \{u_i, v_i\}$,

$$\frac{1}{3}\beta_i(x_i - \bar{x}_i)^3 + \frac{1}{2}a_i(x_i - \bar{x}_i)^2 + (b_i + a_i\bar{x}_i + \beta_i x_i\bar{x}_i)(x_i - \bar{x}_i) \geq 0. \tag{21}$$

Therefore, it suffices to prove, for each $i = 1, \ldots, n$, (19) holds if and only if, for each $x_i \in \{u_i, v_i\}$, (21) holds. We now consider two cases.

Case 1 If $\bar{x}_i = u_i$, then $\gamma_i = 1$. Substituting the value into (19) yields

$$\frac{1}{3}\beta_i(v_i - u_i)^2 + \frac{1}{2}a_i(v_i - u_i) + (b_i + a_i u_i + \beta_i v_i u_i) \geq 0. \tag{22}$$

It suffices to prove (22) holds if and only if (21) holds for each $x_i \in \{u_i, v_i\}$.

Indeed, if (21) holds for each $x_i \in \{u_i, v_i\}$, we can take $x_i = v_i$ and replace $\bar{x}_i$ by u_i. Thus, (22) holds by dividing (21) by $(v_i - u_i)$. Conversely, suppose that (22) holds. Multiplying (22) by $(v_i - u_i)$ and replacing u_i by $\bar{x}_i$, we have, for $x_i = v_i$, (21) holds. Note that, for $x_i = u_i$, (21) holds obviously. Therefore, it follows that, for each $x_i \in \{u_i, v_i\}$, (21) holds.

Case 2 If $\bar{x}_i = v_i$, then $\gamma_i = -1$. Substituting the value into (19) yields

$$-\frac{1}{3}\beta_i(v_i - u_i)^2 + \frac{1}{2}a_i(v_i - u_i) - (b_i + a_i v_i + \beta_i v_i u_i) \geq 0. \tag{23}$$

It suffices to prove (23) holds if and only if (21) holds for each $x_i \in \{u_i, v_i\}$.

Indeed, if (21) holds for each $x_i \in \{u_i, v_i\}$, we can take $x_i = u_i$ and replace $\bar{x}_i$ by v_i. Thus (23), holds by dividing (21) by $(v_i - u_i)$. Conversely, suppose that (23) holds. Multiplying (23) by $(v_i - u_i)$ and replacing v_i by $\bar{x}_i$, we have, for $x_i = u_i$, (21) holds. Note that, for $x_i = v_i$, (21) holds obviously. Therefore, it follows that, for each $x_i \in \{u_i, v_i\}$, (21) holds.

Combining the above two cases yields the desired result. □

4 Numerical experiment

Example 1 Consider the following problem

$$\begin{cases} \min f(x) = \frac{2}{3}x_1^3 + x_2^3 - 2x_4^3 - \frac{1}{2}x_1^2 + 2x_1x_2 + x_1x_4 - \frac{1}{2}x_2^2 + x_2x_3 + 3x_3^2 - x_3x_4 - x_4^2 \\ \qquad\qquad +4x_1 + \frac{9}{2}x_2 - x_3 - x_4 \\ \text{s.t.} \qquad x \in \prod_{i=1}^4 [-1, 1]. \end{cases}$$

In [8], Jeyakumar, Rubinov and Wu proved that $\bar{x} = (-1, -1, \frac{1}{2}, 1)^T$ is a global minimizer of the following quadratic program with box constraints:

$$\begin{cases} \min f(x) = -\frac{1}{2}x_1^2 + 2x_1x_2 + x_1x_4 - \frac{1}{2}x_2^2 + x_2x_3 + 3x_3^2 - x_3x_4 - x_4^2 \\ \qquad\qquad +4x_1 + \frac{9}{2}x_2 - x_3 - x_4 \\ \text{s.t.} \qquad x \in \prod_{i=1}^4 [-1, 1]. \end{cases}$$

By adding the cubic terms to the objective function of above problem, we can obtain Example 1. Note that $\bar{x}$ is also a global minimizer of the following problem:

$$\min \left\{ \frac{2}{3}x_1^3 + x_2^3 - 2x_4^3 \mid x \in \prod_{i=1}^4 [-1, 1] \right\}.$$

Hence, we can conclude that $\bar{x}$ is a global minimizer of Example 1.

In our method, $\beta = (2, 3, 0, -6)^T$, $b = (4, \frac{9}{2}, -1, -1)^T$, $I = \{1, 2, 4\}$, $J = \{3\}$, and

$$A = \begin{pmatrix} -1 & 2 & 0 & 1 \\ 2 & -1 & 1 & 0 \\ 0 & 1 & 6 & -1 \\ 1 & 0 & -1 & -2 \end{pmatrix}.$$

Obviously Assumption 1 holds. Let $Q = diag(c_i) = diag(-4, -4, 0, -4)$. Then it is clear that $A - Q \succeq 0$, $d = (A - Q)\bar{x} + b = (0, 0, 0, -1/2)^T$. For $i = 1 \in I$, it holds that $t_1 = -\frac{1}{2}\bar{x}_1 - \frac{3c_1}{4\beta_1} = 2$. Since $t_1 \geq v_1$, $\beta_1 > 0$, then $\tilde{\tau}_1 = 1$, $\tilde{x}_1 = v_1$, $\tau_1 = 0$, $\gamma_1 = 1$. So we have

$$\tilde{\tau}_1 p_1(\tilde{x}_1) + \tau_1 \min\{\gamma_1 p_1(u_1), \gamma_1 p_1(v_1)\} = \frac{2}{3} > 0$$

Similarly, we can prove that the conditions in Theorem 2 hold for $i = 2, 3, 4$. From Theorem 2, we can conclude that $\bar{x}$ is a global minimizer.

Example 2 Consider the following problem:

$$\begin{cases} \min f(x) = \frac{1}{3}x_1^3 - x_2^3 + 2x_3^3 - \frac{1}{2}x_1^2 - x_2^2 - \frac{3}{2}x_3^2 + x_1 - x_2 + 2x_3 \\ s.t. \ \ x \in \prod_{i=1}^3 [-1, 1]. \end{cases}$$

In [8], Jeyakumar, Rubinov and Wu proved that $\bar{x} = (-1, 1, -1)^T$ is a global minimizer of the following quadratic program:

$$\begin{cases} \min f(x) = -\frac{1}{2}x_1^2 - x_2^2 - \frac{3}{2}x_3^2 + x_1 - x_2 + 2x_3 \\ \text{s.t.} \qquad x \in \prod_{i=1}^3 [-1, 1]. \end{cases}$$

By adding the cubic terms to the objective function of above problem, we can obtain Example 2. Note that $\bar{x}$ is also a global minimizer of the following problem:

$$\min \left\{ \frac{1}{3}x_1^3 - x_2^3 + 2x_3^3 \mid x \in \prod_{i=1}^3 [-1, 1] \right\}.$$

Hence, we can conclude that $\bar{x}$ is a global minimizer of Example 2. Obviously $\bar{y} = (-1, -1, -1)^T$ is not the global minimizer of Example 2.

In our method, for $\bar{x} = (-1, 1, -1)^T$, $I = \{1, 2, 3\}$. For $i = 1$, it holds that $\beta_1 = 1, a_1 = -1, b_1 = 1$, so $t_1 = -\frac{1}{2}\bar{x}_1 - \frac{3a_1}{4\beta_1} = \frac{5}{4}$. Since $t_1 \geq v_1, \beta_1 > 0$, then $\tilde{\tau}_1 = 1, \tilde{x}_1 = v_1, \tau_1 = 0, \gamma_1 = 1$. Therefore, we have

$$\tilde{\tau}_1 q_1(\tilde{x}_1) + \tau_1 \min\{\gamma_1 q_1(u_1), \gamma_1 q_1(v_1)\} = \frac{4}{3} > 0.$$

Similarly, we can prove that the conditions in Corollary 3 hold for $i = 2, 3$. From Corollary 3, we can conclude that $\bar{x}$ is a global minimizer.

For $\bar{y} = (-1, -1, -1)^T$, the index set $I = \{1, 2, 3\}$. Since $\bar{y}_1 = \bar{x}_1$ and $\bar{y}_3 = \bar{x}_3$, then it suffices to check on $\bar{y}_2$. Since

$$\tilde{\tau}_2 q_2(\tilde{y}_2) + \tau_2 \min\{\gamma_2 q_2(u_2), \gamma_2 q_2(v_2)\} = -2 < 0,$$

from Corollary 3, it is known that the point $\bar{y}$ is not the global minimizer.

Acknowledgment The authors are grateful to the referees for constructive comments and suggestions which have contributed to the final preparation of the paper. This research was supported by National Nature Science Foundation of China under Project 10601030, the Cultivation Fund of the Key Scientific and Technical Innovation Project of Ministry of Education of China (NO708040) and Leading Academic Discipline Program, the 10th five year plan of 211 Project for Shanghai University of Finance and Economics.

References

1. Canfield, R.A.: Multipoint cubic surrogate function for sequential approximate optimization. Struct. Multidiscip. Optim. **27**, 326–336 (2004)
2. Nesterov, Yu.: Accelerating the cubic regularization of Newtons method on convex problems. Math. Program. **112**(1), 159–181 (2008)
3. Lin, C.-S., Chang, P.-R., Luh, J.Y.S.: Formulation and optimization of cubic polynomial joint trajectories for industrial robots. IEEE Transact. Autom. Control **AC-28**(12), 1066–1074 (1983)
4. Beck, A., Teboulle, M.: Global optimality conditions for quadratic optimization problems with binary constraints. SIAM J. Optim. **11**, 179–188 (2000)
5. Hiriart-Urruty, J.B.: Global optimality conditions in maximizing a convex quadratic function under convex quadratic constraints. J. Glob. Optim. **21**, 445–455 (2001)
6. Hiriart-Urruty, J.B.: Conditions for global optimality 2. J. Glob. Optim. **13**, 349–367 (1998)
7. Strekalovsky, A.: Global optimality conditions for nonconvex optimization. J. Glob. Optim. **12**, 415–434 (1998)
8. Jeyakumar, V., Rubinov, A.M., Wu, Z.Y.: Sufficient global optimality conditions for non-convex quadratic minimization problems with box constraints. J. Glob. Optim. **36**, 471–481 (2006)
9. Wu, Z.Y., Jeyakumar, V., Rubinov, A.M.: Sufficient conditions for global optimality of bivalent nonconvex quadratic programs with inequality constraints. J. Optim. Theory Appl. **133**, 123–130 (2007)

10. Jeyakumar, V., Rubinov, A.M., Wu, Z.Y.: Nonconvex quadratic minimization with quadratic constraints: global optimality conditions. Math. Program. A **110**(3), 521–541 (2007)
11. Wu, Z.Y.: Sufficient global optimality conditions for weakly convex minimization problems. J. Glob. Optim. **39**, 427–440 (2007)
12. Bertsekas, D.P., Nedic, A., Ozdaglar, A.E.: Convex Analysis and Optimization. Athena Scientific, Belmont, MA, USA (2003)