J Optim Theory Appl (2015) 167:644–652
DOI 10.1007/s10957-015-0715-y

Outer Approximation Algorithm for One Class of Convex Mixed-Integer Nonlinear Programming Problems with Partial Differentiability

Zhou Wei · M. Montaz Ali

Received: 26 November 2013 / Accepted: 12 February 2015 / Published online: 4 March 2015

Abstract In this paper, we mainly study one convex mixed-integer nonlinear programming problem with partial differentiability and establish one outer approximation algorithm for solving this problem. With the help of subgradients, we use the outer approximation method to reformulate this convex problem as one equivalent mixed-integer linear program and construct an algorithm for finding optimal solutions. The result on finite steps convergence of the algorithm is also presented.

Keywords MINLP · Outer approximation · Subgradient · Nonlinear program

Mathematics Subject Classification 90C11 · 90C25 · 90C30

1 Introduction

Many optimization problems involve both discrete and continuous variables and can be modelled as mixed-integer nonlinear programming problems (MINLPs) that arise from practical applications. Problems defined by convex functions are known as convex MINLPs, but they are still non-convex problems due to the presence of discrete variables. Over the past decades, convex MINLPs have become an active research area and several methods have been developed. Readers are invited to consult references [1–6] for more details.

Z. Wei
Department of Mathematics, Yunnan University, Kunming 650091, People's Republic of China
e-mail: wzhou@ynu.edu.cn

Z. Wei · M. M. Ali (✉)
School of Computational and Applied Mathematics, University of the Witwatersrand, Wits, Johannesburg 2050, South Africa
e-mail: Montaz.Ali@wits.ac.za

Convex MINLP with continuously differentiable functions has been extensively studied, while convex but not continuously differentiable functions appear in optimization problems. Therefore, it is natural to study convex MINLP by relaxing the differentiability assumption. Motivated by this, we consider convex MINLP with partial differentiability, one version of relaxed differentiability, and use outer approximation method to establish an appropriate algorithm for solving it.

Duran and Grossmann [1] introduced outer approximation (OA) method to deal with a class of MINLPs, whose functions are dependent linearly on discrete variables. An extension of this OA method was given by Fletcher and Leyffer [2]. They generalized OA method to deal with one wider class of MINLPs defined by convex and continuously differentiable functions. This OA method was mainly studied in [7,8] to equivalently reformulate convex MINLP as one MILP master program. Along the line given in [1,2,7,8], we study the OA method for solving convex MINLP with partial differentiability and use subgradients and KKT conditions to establish one outer approximation algorithm for finding optimal solutions.

2 Preliminaries

Let $\langle \cdot, \cdot \rangle$ denote the duality pairing between two elements of $\mathbb{R}^n$. Let Ω be a closed and convex set of $\mathbb{R}^n$ and $x \in \Omega$. We denote by $T(\Omega, x)$ the contingent cone of Ω at x; that is, $v \in T(\Omega, x)$ if and only if there exist a sequence $\{v_k\}$ in $\mathbb{R}^n$ converging to v and a sequence t_k in $]0, +\infty[$ decreasing to 0 such that $x + t_k v_k \in \Omega$ for all $k \in \mathbb{N}$. Let $N(\Omega, x)$ denote the normal cone of Ω at x, that is

$$N(\Omega, x) := \{\gamma \in \mathbb{R}^n : \langle \gamma, z - x \rangle \leq 0 \ \text{ for all } z \in \Omega\}.$$

It is known that normal cone and contingent cone are the polar of each other.

Let $\psi : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ be a continuous and convex function and $(\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^p$. Recall that a vector $(C, D) \in \mathbb{R}^{n \times p}$ is said to be a subgradient of ψ at $(\bar{x}, \bar{y})$ iff

$$\psi(x, y) \geq \psi(\bar{x}, \bar{y}) + (C, D)^T \begin{pmatrix} x - \bar{x} \\ y - \bar{y} \end{pmatrix} \ \text{ for all } (x, y) \in \mathbb{R}^n \times \mathbb{R}^p.$$

where $(C, D)^T$ is the transpose of matrix (C, D). The set of all such subgradients, denoted by $\partial\psi(\bar{x}, \bar{y})$, is said to be the subdifferential of ψ at $(\bar{x}, \bar{y})$. When $\bar{y}$ is fixed, the subdifferential of $\psi(\cdot, \bar{y})$ at $\bar{x}$ is defined by

$$\partial\psi(\cdot, \bar{y})(\bar{x}) := \left\{C \in \mathbb{R}^n : \psi(x, \bar{y}) \geq \psi(\bar{x}, \bar{y}) + \langle C, x - \bar{x} \rangle \ \text{ for all } x \in \mathbb{R}^n\right\}.$$

It is easy to verify that, for any $(C, D) \in \partial\psi(\bar{x}, \bar{y})$, one has $C \in \partial\psi(\cdot, \bar{y})(\bar{x})$ and $D \in \partial\psi(\bar{x}, \cdot)(\bar{y})$.

3 Main Results

In this section, we mainly study convex MINLP with partial differentiability and reformulate this MINLP as one equivalent MILP master program. Then by solving a finite sequence of relaxed MILP master programs, we establish one outer approximation algorithm for this MINLP.

Let us make some assumptions:

(A1) *X is a nonempty, compact and convex set in $\mathbb{R}^n$ and Y is a discrete set of $\mathbb{R}^p$, and functions $f, g_i : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ $(i = 1, \ldots, m)$ are convex and continuous.*

(A2) *Moreover $f(\cdot, y)$ and $g(\cdot, y)$ are differentiable functions on $\mathbb{R}^n$ for any fixed $y \in Y$.*

The class of convex MINLPs considered in the whole paper is defined as follows:

$$\min_{x,y} \; f(x, y), \quad \text{s.t.} \;\; g(x, y) \leq 0, \quad x \in X, \quad y \in Y. \tag{1}$$

Let us set:

$$V := \left\{ y \in Y : g(x, y) \leq 0 \;\text{ for some } x \in X \right\}$$

is the set of all discrete assignments that give rise to feasible subproblems. For any fixed $y \in Y$, we consider the following subproblem $P(y)$:

$$P(y) \qquad \min_{x} \; f(x, y) \;\text{ s.t. }\; g(x, y) \leq 0, \quad x \in X.$$

For the equivalent reformulation of problem (P), we assume that problem (P) satisfies the following Slater constraint qualification (A3):

(A3) *for any $y \in Y$ producing feasible subproblem $P(y)$, the following Slater constraint qualification holds:*

$$g(\hat{x}, y) < 0 \;\textit{ for some } \hat{x} \in X.$$

For the convex MINLP, whose functions are continuously differentiable, Fletcher and Leyffer [2] and Bonami et al. [7] used gradients and first-order Taylor expansion to linearly approximate functions f, g and established one equivalent MILP master program with the help of KKT conditions. Under the partial differentiability assumption, we substitute gradients with subgradients and equivalently reformulate problem (P) of (1) along the line given in [2,7].

Let $y_j \in V$ be fixed and let x_j be an optimal solution to $P(y_j)$. Take any subgradients $(A_j, B_j) \in \partial f(x_j, y_j)$, $(C_j^i, D_j^i) \in \partial g_i(x_j, y_j)$ for all $i = 1, \cdots, m$ and set $C_j := (C_j^1, \ldots, C_j^m)$, $D_j := (D_j^1, \ldots, D_j^m)$. We first study the following problem:

$$LP(x_j, y_j) \begin{cases} \min\limits_{x} & f(x_j, y_j) + \langle (A_j, B_j), (x - x_j, 0)^T \rangle \\ \text{s.t.} & g(x_j, y_j) + \langle (C_j, D_j), (x - x_j, 0)^T \rangle \leq 0, \\ & x \in X. \end{cases} \tag{2}$$

The following theorem establishes the equivalence between $LP(x_j, y_j)$ and $P(y_j)$. This equivalence plays a key role in the reformulation of problem (P).

Theorem 3.1 *Let $LP(x_j, y_j)$ be defined as (2). Then x_j is one optimal solution of $LP(x_j, y_j)$, and $f(x_j, y_j)$ is the optimal value of $LP(x_j, y_j)$.*

Proof In order to prove Theorem 3.1, it suffices to show that

$$(A_j, B_j)^T \begin{pmatrix} x - x_j \\ 0 \end{pmatrix} \geq 0, \ \ \forall x \in X \text{ with } g(x_j, y_j) + (C_j, D_j)^T \begin{pmatrix} x - x_j \\ 0 \end{pmatrix} \leq 0. \quad (3)$$

Let $x \in X$ be such that

$$g(x_j, y_j) + (C_j, D_j)^T \begin{pmatrix} x - x_j \\ 0 \end{pmatrix} \leq 0. \quad (4)$$

Using assumption (A2), one has that $A_j = \nabla_x f(x_j, y_j)$ and $C_j^i = \nabla_x g_i(x_j, y_j)$ for all $i = 1, \ldots, m$. This and (4) imply that

$$(A_j, B_j)^T \begin{pmatrix} x - x_j \\ 0 \end{pmatrix} = \langle \nabla_x f(x_j, y_j), x - x_j \rangle$$

and

$$g_i(x_j, y_j) + \langle \nabla_x g_i(x_j, y_j), x - x_j \rangle \leq 0, \ \ \forall i = 1, \ldots, m. \quad (5)$$

Noting that x_j solves $P(y_j)$ and the Slater constraint qualification for $g(\cdot, y_j)$ holds, it follows from KKT conditions that there exist $(\lambda_1, \ldots, \lambda_m) \in \mathbb{R}_+^m$ and $\gamma \in N(X, x_j)$ such that $\lambda_i g_i(x_j, y_j) = 0$ ($\forall i = 1, \ldots, m$) and

$$\nabla_x f(x_j, y_j) + \sum_{i \in I(x_j)} \lambda_i \nabla_x g_i(x_j, y_j) + \gamma = 0 \quad (6)$$

where $I(x_j) := \big\{i \in \{1, \ldots, m\} : g_i(x_j, y_j) = 0\big\}$. Using (4) and (5), one has

$$\langle \nabla_x g_i(x_j, y_j), x - x_j \rangle \leq 0, \ \ \forall i \in I(x_j). \quad (7)$$

Since $x - x_j \in T(X, x_j)$ by the convexity of X, it follows from (6), (7) and $\gamma \in N(X, x_j)$ that

$$\nabla_x f(x_j, y_j)^T (x - x_j) = - \sum_{i \in I(x_j)} \lambda_i \nabla_x g_i(x_j, y_j)^T (x - x_j) - \gamma^T (x - x_j) \geq 0.$$

Hence (3) holds. The proof is complete. □

Let us set

$$T := \big\{ j : P(y_j) \text{ is feasible and } x_j \text{ is an optimal solution to } P(y_j) \big\}.$$

For any $j \in T$, take any subgradients $(A_j, B_j) \in \partial f(x_j, y_j)$, $(C_j^i, D_j^i) \in \partial g_i(x_j, y_j)$ for all $i = 1, \ldots, m$ and set $C_j := (C_j^1, \ldots, C_j^m)$, $D_j := (D_j^1, \ldots, D_j^m)$. We consider the following MILP master program (M_V):

$$(M_V) \begin{cases} \min\limits_{x,y,\eta} \quad \eta \\ \text{s.t.} \quad f(x_j, y_j) + (A_j, B_j)^T \begin{pmatrix} x - x_j \\ y - y_j \end{pmatrix} \leq \eta \ \ \forall j \in T, \\ \qquad g(x_j, y_j) + (C_j, D_j)^T \begin{pmatrix} x - x_j \\ y - y_j \end{pmatrix} \leq 0 \ \ \forall j \in T, \\ \qquad x \in X, y \in V \text{ discrete variable.} \end{cases} \tag{8}$$

Theorem 3.2 *Master program (M_V) in (8) and problem (1) are equivalent in the sense that they have the same optimal value and that the optimal solution $(\bar{x}, \bar{y})$ to problem (1) corresponds to the optimal solution $(\bar{x}, \bar{y}, \bar{\eta})$ to (M_V) with $\bar{\eta} = f(\bar{x}, \bar{y})$.*

For reformulating the problem (1) completely, it only suffices to represent suitably the constraint $y \in V$. As pointed out by Fletcher and Leyffer [2] and Bonami et al. [7], it is necessary to include information from infeasible subproblems and then exclude discrete assignments producing infeasible subproblems.

Let $y_l \in Y$ be such that $P(y_l)$ is infeasible and let J_l be one subset of $\{1, \ldots, m\}$ such that there is some $\bar{x} \in X$ satisfying

$$g_i(\bar{x}, y_l) < 0, \quad \forall i \in J_l. \tag{9}$$

Denote $J_l^{\perp} := \{1, \ldots, m\} \backslash J_l$. To detect the infeasibility, we study the following nonlinear program

$$F(y_l) \begin{cases} \min\limits_{x} \quad \sum\limits_{i \in J_l^{\perp}} [g_i(x, y_l)]_+ \\ \text{s.t.} \quad g_i(x, y_l) \leq 0 \ \ \forall i \in J_l, \\ \qquad x \in X, \end{cases} \tag{10}$$

where $[g(x, y_l)]_+ := \max\{g(x, y_l), 0\}$.

Theorem 3.3 *Let $y_l \in Y$ be such that $P(y_l)$ is infeasible and x_l solve nonlinear program $F(y_l)$. Then for any subgradients $(C_l^i, D_l^i) \in \partial g_i(x_l, y_l) (\forall i = 1, \ldots, m)$, y_l is infeasible to the following constraint*

$$g_i(x_l, y_l) + (C_l^i, D_l^i)^T \begin{pmatrix} x - x_l \\ y - y_l \end{pmatrix} \leq 0, \quad \forall i \in J_l^{\perp} \cup J_l,$$
$$x \in X, y \in Y.$$

Proof Since X is compact and g is continuous, one has $\sum_{i \in J_l^{\perp}} [g_i(x_l, y_l)]_+ > 0$. Noting that $(C_l^i, D_l^i) \in \partial g_i(x_l, y_l) (\forall i = 1, \ldots, m)$, it follows that

$$C_l^i = \nabla_x g_i(x_l, y_l), \quad \forall i \in \{1, \ldots, m\}. \tag{11}$$

Suppose to the contrary that there exists $\hat{x} \in X$ such that

$$g_i(x_l, y_l) + \langle \nabla_x g_i(x_l, y_l), \hat{x} - x_l \rangle \leq 0, \quad \forall i \in J_l^{\perp} \cup J_l. \tag{12}$$

Since x_l solves $F(y_l)$ and (9) holds, by KKT conditions, there exist $\lambda_i \geq 0$ such that $\lambda_i g_i(x_l, y_l) = 0$ $(\forall i \in J_l)$ and

$$0 \in \sum_{i \in J_l^{\perp}} \partial[g_i(\cdot, y_l)] + (x_l) + \sum_{i \in J_l} \lambda_i \nabla_x g_i(x_l, y_l) + N(X, x_l). \tag{13}$$

Denote $J^1 := \{i \in J_l^{\perp} : g_i(x_l, y_l) = 0\}$ and $J^2 := \{i \in J_l^{\perp} : g_i(x_l, y_l) > 0\}$. By virtue of [9, Theorem 2.4.18] and (13), there exist $t_i \in [0, 1](i \in J^1)$ and $\gamma \in N(X, x_l)$ such that

$$\sum_{i \in J^1} t_i \nabla_x g_i(x_l, y_l) + \sum_{i \in J^2} \nabla_x g_i(x_l, y_l) + \sum_{i \in J_l} \lambda_i \nabla_x g_i(x_l, y_l) + \gamma = 0. \tag{14}$$

Let $t_i \equiv 1$ for all $i \in J^2$. Using (12) and (14), one has

$$\begin{aligned} 0 &\geq \sum_{i \in J^1 \cup J^2} t_i g_i(x_l, y_l) + \Big(\sum_{i \in J^1 \cup J^2} t_i \nabla_x g_i(x_l, y_l) + \sum_{i \in J_l} \lambda_i \nabla_x g_i(x_l, y_l) + \gamma \Big)^T (\hat{x} - x_l) \\ &\geq \sum_{i \in J^2} g_i(x_l, y_l) = \sum_{i \in J_l^{\perp}} [g_i(x_l, y_l)]_+, \end{aligned}$$

which contradicts $\sum_{i \in J_l^{\perp}} [g_i(x_l, y_l)]_+ > 0$. The proof is complete. □

Let us set

$$S := \big\{ l : P(y_l) \text{ is infeasible and } x_l \text{ solves } F(y_l) \big\}.$$

Using Theorem 3.3, we obtain the following theorem which enables us to eliminate discrete assignments producing infeasible subproblems.

Theorem 3.4 *Let $l \in S$ and take any subgradients $(C_l^i, D_l^i) \in \partial g_i(x_l, y_l)$ for all $i = 1, \ldots, m$. Then the following constraints*

$$g(x_l, y_l) + (C_l, D_l)^T \begin{pmatrix} x - x_l \\ y - y_l \end{pmatrix} \leq 0, \quad \forall l \in S \tag{15}$$

exclude all discrete assignments $y_l \in Y$ satisfying $P(y_l)$ is infeasible.

From Theorem 3.4, we add linearization from $F(y_l)$ where $P(y_l)$ is infeasible to correctly represent the constraints $y \in V$ in (8) and give rise to the MILP master program (MP) that is equivalent to problem (1).

For any $j \in T$, take any subgradients $(A_j, B_j) \in \partial f(x_j, y_j)$, $(C_j^i, D_j^i) \in \partial g_i(x_j, y_j)$ for all $i = 1, \ldots, m$ and set $C_j := (C_j^1, \ldots, C_j^m)$, $D_j := (D_j^1, \ldots, D_j^m)$, while for any $l \in S$, take any subgradients $(C_l^i, D_l^i) \in \partial g_i(x_l, y_l)$ for all $i = 1, \ldots, m$ and set $C_l := (C_l^1, \ldots, C_l^m)$, $D_l := (D_l^1, \ldots, D_l^m)$. We consider the following MILP master problem:

$$\text{(MP)} \begin{cases} \min\limits_{x,y,\eta} & \eta \\ \text{s.t.} & f(x_j, y_j) + (A_j, B_j)^T \begin{pmatrix} x - x_j \\ y - y_j \end{pmatrix} \le \eta \ \ \forall j \in T, \\ & g(x_j, y_j) + (C_j, D_j)^T \begin{pmatrix} x - x_j \\ y - y_j \end{pmatrix} \le 0 \ \ \forall j \in T, \\ & g(x_l, y_l) + (C_l, D_l)^T \begin{pmatrix} x - x_l \\ y - y_l \end{pmatrix} \le 0 \ \ \forall l \in S, \\ & x \in X, y \in Y \text{ discrete variable.} \end{cases} \tag{16}$$

Theorem 3.5 *Master program (MP) in (16) is equivalent to problem (1) in the sense that they have the same optimal value and that the optimal solution $(\bar{x}, \bar{y})$ to problem (1) corresponds to the optimal solution $(\bar{x}, \bar{y}, \bar{\eta})$ to (MP) with $\bar{\eta} = f(\bar{x}, \bar{y})$.*

Remark 3.1 As one extension of [2, Theorem1] and [7, Theorem1], Theorem 3.5 shows that the OA method can be used to equivalently reformulate problem (1) and any optimal solution of problem (1) is that of (MP) in (16). However, the converse may not hold necessarily. Consider the following MINLP:

$$\begin{aligned} \min_{x,y} \quad & f(x, y) := x_2, \\ \text{s.t.} \quad & g(x, y) := x_1^2 + x_2^2 + |y| - 2 \le 0 \\ & x = (x_1, x_2) \in [-1, 1] \times [-2, 2], \\ & y \in \{-1, 0, 1, 3\}. \end{aligned} \tag{17}$$

This convex MINLP satisfies assumptions (A1)–(A3) and has the optimal value $-\sqrt{2}$. Note that

$$\partial g\left(\left(0, -\sqrt{2}\right), 0\right) = \left\{\left(0, -2\sqrt{2}\right)\right\} \times [-1, 1].$$

By taking any $t \in [-1, 1]$ and using the reformulation as (16), problem (17) can be equivalently rewritten as

$$\begin{aligned} \min_{x,y,\eta} \quad & \eta \\ \text{s.t} \quad & x_2 \le \eta, \\ & -2\sqrt{2}x_2 - 4 + ty \le 0, \\ & -2x_2 + y - 3 \le 0, \\ & -2x_2 - y - 3 \le 0, \\ & 1 + y - 3 \le 0, \\ & x = (x_1, x_2) \in [-1, 1] \times [-2, 2], \\ & y \in \{-1, 0, 1, 3\}. \end{aligned} \tag{18}$$

Clearly for any $x_1 \neq 0$, $(x, y, \eta) = \left(\left(x_1, -\sqrt{2}\right), 0, -\sqrt{2}\right)$ is an optimal solution to problem (18), but $(x, y) = \left(\left(x_1, -\sqrt{2}\right), 0\right)$ is infeasible to problem (17).

Now, by solving relaxations of master program (MP) in (16), we formally state the outer approximation algorithm for problem (1) as follow:

Algorithm 1 (Outer approximation algorithm for problem (1))

Step 1. Given an initial discrete variable $y_1 \in Y$, set $\mathrm{UBD}^0 := +\infty$, $T_0 = S_0 := \emptyset$ and let $k := 1$.

Step 2. Solve subproblem $P(y_k)$ and let the solution be x_k; or solve nonlinear program $F(y_k)$ if $P(y_k)$ is infeasible and let the solution be x_k. If $P(y_k)$ is feasible, set $T_k := T_{k-1} \cup \{k\}$, $S_k := S_{k-1}$ and $UBD^k := \min\{f(x_k, y_k), UBD^{k-1}\}$; otherwise set $T_k := T_{k-1}$, $S_k := S_{k-1} \cup \{k\}$ and $UBD^k := UBD^{k-1}$.

Step 3. If $k \in T_k$, take any subgradients $(A_k, B_k) \in \partial f(x_k, y_k)$, $(C_k^i, D_k^i) \in \partial g_i(x_j, y_j)(\forall i = 1, \ldots, m)$ and set $C_k := (C_k^1, \ldots, C_k^m)$, $D_k := (D_k^1, \ldots, D_k^m)$; otherwise $(k \in S_k)$, take any subgradients $(C_k^i, D_k^i) \in \partial g_i(x_j, y_j)(\forall i = 1, \ldots, m)$ and set $C_k := (C_k^1, \ldots, C_k^m)$, $D_k := (D_k^1, \ldots, D_k^m)$. Solve the following relaxed MILP master program MP^k:

$$MP^k \left\{ \begin{array}{l} \min\limits_{x,y,\eta} \quad \eta \\ \text{s.t.} \quad \eta < UBD^k \\ \qquad f(x_j, y_j) + (A_j, B_j)^T \begin{pmatrix} x - x_j \\ y - y_j \end{pmatrix} \leq \eta \;\; \forall j \in T_k, \\ \qquad g(x_j, y_j) + (C_j, D_j)^T \begin{pmatrix} x - x_j \\ y - y_j \end{pmatrix} \leq 0 \;\; \forall j \in T_k, \\ \qquad g(x_l, y_l) + (C_l, D_l)^T \begin{pmatrix} x - x_l \\ y - y_l \end{pmatrix} \leq 0 \;\; \forall l \in S_k, \\ \qquad x \in X, \, y \in Y \text{ discrete variable.} \end{array} \right. \tag{19}$$

Denote $(\hat{x}, \hat{y}, \hat{\eta})$ the optimal solution to MP^k. Let $y_{k+1} := \hat{y}$, obtaining a new discrete assignment, and set $k := k + 1$. Return to **Step 2** until MP^{k+1} is infeasible.

Remark 3.2 Constraint $\eta < UBD^k$ in MP^k (19) is used to prevent any y_j $(j \in T_k)$ from being the optimal solution to the relaxed master program MP^k. Further, as pointed out in [2], constraint $\eta < UBD^k$ would be substituted with $\eta \leq UBD^k - \varepsilon$ in practice for some tolerance parameter $\varepsilon > 0$.

The following theorem shows that Algorithm 1 stated above is able to detect feasibility or infeasibility of problem (1) and the procedure terminates after a finite number of steps under the assumption of finite discrete assignments. The proof, similar to that of [2, Theorem1], can be obtained by using Theorems 3.1 and 3.3.

Theorem 3.6 *Suppose assumptions* (A1)–(A3) *hold and the cardinality of discrete set Y is finite. Then either problem* (1) *is infeasible or Algorithm 1 terminates at k_0-th step for some $k_0 \in \mathbb{N}$ and there exists $j_0 \in T_{k_0-1} \cup \{k_0\}$ such that $f(x_{j_0}, y_{j_0})$ equals the optimal value of problem* (1).

4 Conclusions

The work in this paper is devoted to the study of the convex MINLP by relaxing the differentiability assumption and the construction of outer approximation algorithm for solving such MINLP. With the assumption of partial differentiability, the OA method and subgradients obtained from KKT conditions are used to linearize convex functions and reformulate this MINLP as an equivalent MILP master program. An example showing that some optimal solution of the reformulated master program is infeasible to the MINLP problem is given. By solving a finite sequence of subproblems and relaxed master programs, the outer approximation algorithm has been established to find optimal solutions of this MINLP. The convergence after a finite number of steps is also proved. This work is an extension of the OA method for solving convex MINLP under relaxed differentiability assumptions.

Acknowledgments The authors are grateful to the referee for careful reading this paper and valuable comments which help us to improve the original version. This research was supported by the National Natural Science Foundations of P. R. China (Grant Nos. 11401518 and 11261067) and the IRTSTYN, and by the Claude Leon Foundation of South Africa.

References

1. Duran, M., Grossmann, I.E.: An outer-approximation algorithm for a class of mixed-integer nonlinear programs. Math. Program. **36**, 307–339 (1986)
2. Fletcher, R., Leyffer, S.: Solving mixed-integer nonlinear programs by outer approximation. Math. Program. **66**, 327–349 (1994)
3. Geoffrion, A.M.: Generalized benders decomposition. J. Optim. Theory Appl. **10**(4), 237–260 (1972)
4. Grossmann, I.E.: Review of nonlinear mixed-integer and disjunctive programming techniques. Optim. Eng. **3**, 227–252 (2002)
5. Leyffer, S.: Integrating SQP and branch-and-bound for mixed integer nonlinear programming. Comput. Optim. Appl. **18**, 295–309 (2001)
6. Quesada, I., Grossmann, I.E.: An LP/NLP based branch and bound algorithm for convex MINLP optimization problems. Comput. Chem. Eng. **16**, 937–947 (1992)
7. Bonami, P., Biegler, L., Conn, A.R., Cornuéjols, G., Grossmann, I.E., Laird, C., Lee, J., Lodi, A., Margot, F., Sawaya, N., Wächter, A.: An algorithmic framework for convex mixed integer nonlinear programs. Discret. Optim. **5**(2), 186–204 (2008)
8. Eronen, V.-P., Makela, M.M., Westerlund, T.: On the generalization of ECP and OA methods to nonsmooth convex minlp problems. Optimization **63**, 1057–1073 (2014)
9. Zălinescu, C.: Convex Analysis in General Vector Spaces. World Scientific, Singapore (2002)