J Glob Optim (2009) 44:433–458
DOI 10.1007/s10898-008-9347-1

Unified theory of augmented Lagrangian methods for constrained global optimization

Chang-Yu Wang · Duan Li

Received: 6 May 2007 / Accepted: 23 August 2008 / Published online: 17 September 2008

Abstract We classify in this paper different augmented Lagrangian functions into three unified classes. Based on two unified formulations, we construct, respectively, two convergent augmented Lagrangian methods that do not require the global solvability of the Lagrangian relaxation and whose global convergence properties do not require the boundedness of the multiplier sequence and any constraint qualification. In particular, when the sequence of iteration points does not converge, we give a sufficient and necessary condition for the convergence of the objective value of the iteration points. We further derive two multiplier algorithms which require the same convergence condition and possess the same properties as the proposed convergent augmented Lagrangian methods. The existence of a global saddle point is crucial to guarantee the success of a dual search. We generalize in the second half of this paper the existence theorems for a global saddle point in the literature under the framework of the unified classes of augmented Lagrangian functions.

Keywords Augmented Lagrangian function · Duality theory · Global convergence · Global optimization · Nonconvex optimization · Nonlinear programming · Saddle point

1 Introduction

We consider in this paper the following inequality constrained global optimization problem:

$$(P) \qquad \min f(x) \quad \text{s. t. } g_i(x) \le 0,\ i = 1, \ldots, m,$$

C.-Y. Wang
Institution of Operations Research, Qufu Normal University, Rizhao, Shandong 276826, China

D. Li (✉)
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, NT, Hong Kong
e-mail: dli@se.cuhk.edu.hk

where $x \in \mathbb{R}^n$ and f and $g_i, i = 1, \ldots, m$, are all continuously differentiable functions on $\mathbb{R}^n$. Denote X^* as the set of the optimal solutions to problem (P).

The Lagrangian function of (P) is defined as follows for $\lambda = (\lambda_1, \ldots, \lambda_m)^T \geq 0$,

$$L(x, \lambda) = f(x) + \sum_{i=1}^{n} \lambda_i g_i(x),$$

while the dual function of (P) is given by

$$d(\lambda) = \min_{x \in \mathbb{R}^n} L(x, \lambda).$$

The Lagrangian dual problem of (P) is then to search for a multiplier vector $\lambda^* \in \mathbb{R}^m_+$ that maximizes $d(\lambda)$ for all $\lambda \in \mathbb{R}^m_+$:

$$(D) \qquad \max_{\lambda \in \mathbb{R}^m_+} d(\lambda).$$

The classical Lagrangian method based on the above dual formulation has been successfully applied to convex optimization as the zero duality gap can be achieved between the primal problem (P) and the dual problem (D). However, in nonconvex situations, nonzero duality gap often exists. Thus, the classical Lagrangian method may fail to identify the global optimal solution of problem (P) via the dual search. Augmented Lagrangian methods have been developed to remedy the notorious problem of nonzero duality gap. The first augmented Lagrangian method was independently proposed by Hestenes [10] and Powell [24] for equality-constrained problems by incorporating a quadratic penalty term in the conventional Lagrangian function. This method avoids the ill-conditional behavior and slow convergence of the early penalty methods. However, no convergence properties were presented for this method. The method by Hestenes [10] and Powell [24] was then extended to inequality constrained convex optimization problems by Bertsekas [2] with a proof of its global convergence. Rockafellar [25,26] proposed essentially-quadratic augmented Lagrangian method and proved its global convergence for nonconvex optimization problems. However, these augmented Lagrangian functions in [25,26] are not twice continuously differentiable with respect to x, thus preventing the use of Newton-type methods for solving their corresponding unconstrained Lagrangian relaxation problems. Exponential-type augmented Lagrangian functions overcome this weakness, as they are twice continuously differentiable with respect to x (see [2,13,33]). To avoid the difficulties arising in the convergence analysis and the ill-conditional numerical behavior of the exponential-type Lagrangian methods, exponential-type penalty function (see [2,18,33]) and modified barrier function (see [19]) were developed. Promising numerical computation results of the modified barrier function were reported in [1,21].

The convergence properties of augmented Lagrangian methods have been studied for decades. Local convergence properties have been analyzed in [7,8,12,17–19], while global convergence of convex optimization has been analyzed in [2,9,22,23,27,33]. Global convergence properties for nonconvex optimization problems have been analyzed in [3,20,25,26]. However, the convergence analysis of the above methods requires an essential assumption, namely, the boundedness of the multiplier sequence or certain constraint qualifications for the primal problem, such as the linear independence constraint qualification or Mangasarian-Fromoritz constraint qualification, which are sufficient conditions for the boundedness of the multiplier sequence. This assumption confines applications of augmented Lagrangian methods. The first research objective of this paper is to remove this restrictive condition.

On the other hand, the duality and exact penalty properties of many augmented Lagrangian and nonlinear Lagrangian have been studied in [4–6,11,12,14,16,30,31,35,36]. The existence of a saddle point of augmented Lagrangian functions plays a crucial role in ensuring a success of the Lagrangian dual methods. The existence of the global saddle point of p-power Lagrangian functions for inequality constrained nonconvex problems has been analyzed in [13–15]. The existence of a local saddle point of modified barrier Lagrangian functions and essentially-quadratic augmented Lagrangians has been studied in [19,28,29]. Rockafellar [26] has presented an existence theorem of a global saddle point for essentially quadratic augmented Lagrangian under second order sufficient condition and the assumption of the uniqueness of the optimal solution. Recently, Sun et al [32] have considered four types of augmented Lagrangian functions including two classes of augmented Lagrangian functions in [19,28,29], and have systematically investigated the properties of local and global saddle points. Under the second order sufficient condition and the assumption of the uniqueness of the global optimal solution, they have proved the existence of a global saddle point for three classes of augmented Lagrangian functions over a compact set. Using an extra separation condition, they have further proved the existence of a global saddle point for the class of exponential augmented Lagrangian functions. The second objective of this paper is to extend the existing results on the existence of saddle points without requiring the uniqueness of the global optimal solution and that the feasible region of the Lagrangian relaxation be compact.

To achieve the above two objectives, we first present in this paper three unified classes of augmented Lagrangian functions which include six important augmented Lagrangian functions in the literature as their special cases. We then develop two convergent augmented Lagrangian methods, respectively, based on the first two unified augmented Lagrangian functions. These two methods do not require the global solvability of the corresponding relaxation problems. Furthermore, the global convergence of the two methods does not require the boundedness of the multiplier sequence and other constraint qualifications of the primal problem. In particular, when there is no limit point of the iteration points, we establish the necessary and sufficient condition for the convergence of the sequence of the objective value. On the other hand, we propose two multiplier algorithms based on the proposed convergent augmented Lagrangian methods and prove the convergence of the proposed multiplier algorithms using the same conditions as for the convergent augmented Lagrangian methods. We next derive a multiplier algorithm based on the third unified class of augmented Lagrangian functions and prove its convergence without requiring the boundedness of multiplier sequence and additional constraint qualification. In the second half of the paper, we extend the existence results of the saddle points in the literature by relaxing the requirement of the compactness of X and the uniqueness of the global optimal solution.

The paper is organized as follows. In Sect. 2, we present three unified classes of augmented Lagrangian functions and their special instances. In Sect. 3, we propose two convergent augmented Lagrangian methods based on the first two unified classes of augmented Lagrangian functions, prove their global convergence properties and present the necessary and sufficient condition under which the sequence of the objective value converges to the optimal value of the primal problem. In Sect. 4, we derive Lagrangian multiplier algorithms, prove their global convergence properties, and derive the necessary and sufficient condition under which the sequence of the objective values is convergent to the optimal value. We develop next in Sect. 5 a multiplier algorithm based on the third class of augmented Lagrangian functions and prove its global convergence. Finally in Sect. 6, we generalize the existence results of global saddle points in the literature.

2 Unified formulations of augmented Lagrangian functions

We propose in this section three unified formulations of augmented Lagrangian functions and demonstrate each unified formulation by some illustrative examples.

Denote $\mathbb{R}_+ = [0, +\infty), \mathbb{R}_a = (-\infty, a)$ for $a \in (0, +\infty]$, and $\Omega_c(a) = \{x \in \mathbb{R}^n | cg_i(x) < a, 1 \le i \le m\}$ for $c > 0$. Note that $\mathbb{R}_\infty = \mathbb{R}$.

2.1 P-type augmented Lagrangian function

The first type of unified augmented Lagrangian functions is given as follows:

$$L_P(x, \lambda, c) = \begin{cases} f(x) + \frac{1}{c}\sum_{i=1}^{m} P(cg_i(x), \lambda_i), & x \in \Omega_c(a) \\ +\infty, & x \notin \Omega_c(a) \end{cases}$$

where $P(s, t)$ is continuous on $\mathbb{R}_a \times \mathbb{R}_+$ and continuously differentiable with respect to the first variable $s \in \mathbb{R}_a$.

We assume that function $P(s, t)$ possesses the following properties.

(H_1) P$(\cdot, t)$ is monotonically increasing with respect to s and satisfies:

$$P(0, t) = 0, \ \forall\, t \in \mathbb{R}_+;$$
$$P(s, 0) \ge 0, \ \forall\, s \in \mathbb{R}_a;$$
$$P(s, t) \to +\infty \ (t \to +\infty), \ \text{for } s > 0.$$

(H_2) There exists a continuous function $r(t)$ such that

$$P(s, t) \ge r(t), \ \forall\, (s, t) \in \mathbb{R}_a \times \mathbb{R}_+.$$

(H_3) If $a = +\infty$, then $\frac{P(s,t)}{s} \to +\infty (s \to +\infty)$ holds uniformly for any $t \in \mathbb{R}_+$.
(H_4) $P_s'(s, t) \le t, \forall s < 0$ and $P_s'(s, t) \to 0 \ (s \to -\infty)$ holds uniformly for any $t \in S \subset \mathbb{R}_+$,

where S is any bounded set.

Definition 1 The augmented Lagrangian function $L_P(x, \lambda, c)$ with $P(s, t)$ satisfying (H_1)–(H_4) is called P-type augmented Lagrangian.

The following examples illustrate some special cases of L_P.

Example 1 Modified Courant-type augmented Lagrangian function:

$$L_{P_1}(x, \lambda, c) = f(x) + \frac{1}{c}\sum_{i=1}^{m} P_1(cg_i(x), \lambda_i),$$

$(a = +\infty)$, where $P_1(s, t) = (\max\{0, \phi_1(s) + t\})^2 - t^2$, $(s, t) \in \mathbb{R} \times \mathbb{R}_+$, and the function $\phi_1(\cdot)$ satisfies the following conditions:

(A_1) $\phi_1(\cdot)$ is a twice continuously differentiable and convex function on $\mathbb{R}$;
(A_2) $\phi_1(0) = 0, \phi_1'(0) = 1$;
(A_3) $\lim_{s \to -\infty} \phi_1'(s) > 0$.

It is clear that function L_{P_1} includes the modified Courant-type augmented Lagrangian function, i.e. the Rockefeller's *essentially quadratic* augmented Lagrangian function, as a special case by setting $\phi_1(s) = s$ (see [2,10,24,27]).

Example 2 Essentially quadratic augmented Lagrangian function (see [32]):

$$L_{P_2}(x, \lambda, c) = f(x) + \frac{1}{c}\sum_{i=1}^{m} P_2(cg_i(x), \lambda_i),$$

$(a = +\infty)$, where $P_2(s,t) = \min_{\tau \geq s}\{t\tau + \phi_2(\tau)\}$, $(s,t) \in \mathbb{R} \times \mathbb{R}_+$, and the function $\phi_2(\cdot)$ satisfies the following conditions:

(B_1) $\phi_2(\cdot)$ is a twice continuously differentiable and convex function on $\mathbb{R}$;
(B_2) $\phi_2(0) = 0, \phi_2'(0) = 0, \phi_2''(0) > 0$;
(B_3) $\frac{\phi_2(s)}{|s|} \to +\infty, (|s| \to +\infty)$.

When setting $\phi_2(s) = \frac{1}{2}s^2$, this special case of L_{P_2} becomes the same as the special case of L_{P_1} by setting $\phi_1(s) = s$. Note that function L_{P_2} is, in general, different in nature from function L_{P_1}.

Example 3 Exponential-type augmented Lagrangian [32]:

$$L_{P_3}(x, \lambda, c) = f(x) + \frac{1}{c}\sum_{i=1}^{m} P_3(cg_i(x), \lambda_i),$$

$(a = +\infty)$, where $P_3(s,t) = t\phi_3(s) + \xi(s)$, $(s,t) \in \mathbb{R} \times \mathbb{R}_+$, the function $\phi_3(\cdot)$ satisfies the following conditions:

(C_1) $\phi_3(\cdot)$ is a twice continuously differentiable and convex on $\mathbb{R}$;
(C_2) $\phi_3(0) = 0, \phi_3'(0) = 0, \phi_3''(0) > 0$;
(C_3) $\lim_{s \to -\infty}\phi_3(s) > -\infty, \lim_{s \to -\infty}\phi_3'(s) = 0$,

and the function $\xi(\cdot)$ satisfies the following conditions:

(C_1') $\xi(\cdot)$ is a twice continuously differentiable convex function on $\mathbb{R}$;
(C_2') $\xi(s) = 0$ for $s \leq 0$ and $\xi(s) > 0$ for $s > 0$;
(C_3') $\frac{\xi(s)}{s} \to +\infty, \ (s \to +\infty)$.

Note that function L_{P_3} is a generalization of *exponential-type* augmented Lagrangian function (see [2,18,33]).

Example 4 Modified Carroll barrier-type augmented Lagrangian function (see [32]):

$$L_{P_4}(x, \lambda, c) = \begin{cases} f(x) + \frac{1}{c}\sum_{i=1}^{m} P_4(cg_i(x), \lambda_i), & x \in \Omega_c(1) \\ +\infty, & x \notin \Omega_c(1) \end{cases}$$

$(a = 1)$, where $P_4(s,t) = t\phi_4(s)$, $(s,t) \in \mathbb{R} \times \mathbb{R}_+$, and the function $\phi_4(\cdot)$ satisfies the following conditions:

(D_1) $\phi_4(\cdot)$ is a twice continuously differentiable and convex function on $\mathbb{R}_1$.
(D_2) $\phi_4(0) = 0, \phi_4'(0) = 1, \phi_4''(0) > 0$.
(D_3) $\lim_{s \to -\infty}\phi_4(s) > -\infty, \lim_{s \to -\infty}\phi_4'(s) = 0$.

Taking $\phi_4(s) = \frac{1}{1-s} - 1$ in L_{P_4} gives rise to the *modified Carroll* function (see [19]).

2.2 R-type augmented lagrangian function

The second type of unified augmented Lagrangian functions is

$$L_R(x, \lambda, c) = \begin{cases} f(x) + \frac{1}{c} \sum_{i=1}^{m} R(cg_i(x), \lambda_i), & x \in \Omega_c(a) \\ +\infty, & x \notin \Omega_c(a) \end{cases}$$

where $R(s, t)$ is continuous on $\mathbb{R}_a \times \mathbb{R}_+$ and at least first order continuously differentiable with respect to $s \in \mathbb{R}_a$.

We assume that function $R(s, t)$ possesses the following properties:

- (H_1') is the same as (H_1);
- (H_2') For any given $t \in \mathbb{R}_+$, $\frac{R(s,t)}{s} \to 0$ $(s \to -\infty)$;
- (H_3') is the same as (H_3);
- (H_4') is the same as (H_4).

Definition 2 The augmented Lagrangian function $L_R(x, \lambda, c)$ with $R(s, t)$ satisfying (H_1'), (H_2'), (H_3') and (H_4') is called R-type augmented Lagrangian.

It is easy to check that the following example is a special case of L_R.

Example 5 Modified Frish barrier augmented Lagrangian function [32]:

$$L_{R_1}(x, \lambda, c) = \begin{cases} f(x) + \frac{1}{c} \sum_{i=1}^{m} R_1(cg_i(x), \lambda_i), & x \in \Omega_c(1) \\ +\infty, & x \notin \Omega_c(1) \end{cases}$$

$(a = 1)$, where $R_1(s, t) = t\varphi(s)$, $(s, t) \in \mathbb{R} \times \mathbb{R}_+$, and the function $\psi(\cdot)$ satisfies the following conditions:

- (E_1) $\varphi(\cdot)$ is a twice continuously differentiable and convex function on $\mathbb{R}_1$;
- (E_2) $\varphi(0) = 0, \varphi'(0) = 1, \varphi''(0) > 0$;
- (E_3) $\lim_{s\to-\infty} \frac{\varphi(s)}{s} = 0, \quad \lim_{s\to-\infty} \varphi'(s) = 0$.

Let $\varphi(s) = -\ln(1 - s)$ in L_{R_1}, we get *modified Frish* augmented Lagrangian function [19].

2.3 V-type augmented lagrangian function

The third type of unified augmented Lagrangian functions is

$$L_V(x, \lambda, c) = \begin{cases} f(x) + \frac{1}{c} \sum_{i=1}^{m} V(cg_i(x), \lambda_i), & x \in \Omega_c(a) \\ +\infty, & x \notin \Omega_c(a) \end{cases}$$

where $V(s, t)$ is continuous on $\mathbb{R}_a \times \mathbb{R}_+$ and at least first order continuously differentiable with respect to $s \in \mathbb{R}_a$.

We assume that function $V(s, t)$ possesses the following properties:

- (H_1'') $V(\cdot, t)$ is monotonically increasing and convex in s and satisfies,

$$V(0, t) = 0, \forall t \in \mathbb{R}_+; V(s, t) \geq st, \forall\ (s, t) \in \mathbb{R}_a \times \mathbb{R}_+;$$

- (H_2'') The same as (H_2);

(H_3'') If $a = +\infty$, then $\frac{V(s,t)}{s} \to +\infty (s \to +\infty)$ holds uniformly for any $t \in S \subset \mathbb{R}_+$, where S is any closed unbounded set of $\mathbb{R}_+$;

(H_4'') $V'(s,t) > 0, \forall t > 0$ and $V_s'(s,t) \to 0$ $(s \to -\infty)$ holds uniformly for any $t \in S \subset \mathbb{R}_+$, where S is any unbounded set of $\mathbb{R}_+$.

Definition 3 The augmented Lagrangian function $L_v(x,\lambda,c)$ with $R(s,t)$ satisfying (H_1''), (H_2''), (H_3''), (H_4'') is called V-type augmented Lagrangian.

We can verify that the following example is a special case of L_V.

Example 6 Modified exponential augmented Lagrangian function [32]:

$$L_{V_1}(x,\lambda,c) = f(x) + \frac{1}{c}\sum_{i=1}^{m} V_1(cg_i(x),\lambda_i)$$

with $(a = +\infty)$, where $V_1(s,t) = t\psi(s)$, $(s,t) \in \mathbb{R} \times \mathbb{R}_+$, and the function $\psi(\cdot)$ satisfies the following conditions:

(F_1) $\psi(\cdot)$ is a twice continuously differentiable and strictly convex function on $\mathbb{R}$;

(F_2) $\psi(0) = 0$, $\psi'(0) = 1$, $\psi''(0) > 0$;

(F_3) $\lim_{s\to+\infty} \frac{\psi(s)}{s} = +\infty$, $\lim_{s\to-\infty} \psi(s) > -\infty$, $\lim_{s\to-\infty} \psi'(s) = 0$.

Taking $\psi(s) = e^s - 1$ in L_{V_1} gives rise to the *e*xponential augmented Lagrangian function (see [2,13,33]).

3 Convergent augmented Lagrangian methods

Augmented Lagrangian methods are devised to eliminate the duality gap in nonconvex optimization when applying the traditional Lagrangian method and, eventually, to find a global optimal solution of (P) via globally solving a series of augmented Lagrangian relaxation problems. Although the augmented Lagrangian relaxation problem is an unconstrained minimization problem, the existence of its global optimal solution, in general, cannot be guaranteed, since $\Omega_c(a)$ is an open set. To cope with this circumstance, we develop in this section two *convergent augmented Lagrangian methods* that only require an approximate solution of the augmented Lagrangian relaxation in the iteration. These two convergent augmented Lagrangian methods are, respectively, based on L_P and L_R, the two types of the unified augmented Lagrangian functions proposed in the previous section. We will then analyze their respective convergence properties.

The convergent augmented Lagrangian methods based on L_P and L_R are denoted as $PALM$ and $RALM$, respectively.

3.1 $PALM$ and its convergence

Consider the Lagrangian relaxation problem associated with the augmented Lagrangian L_P,

$$(L_{P,\lambda,c}) \quad \min_{x\in\mathbb{R}^n} L_P(x,\lambda,c).$$

Definition 4 For any given $\lambda \geq 0$, $c > 0$ and $\varepsilon > 0$, if there exists an $x(\lambda,c,\varepsilon) \in \mathbb{R}^n$ such that

$$L_P(x(\lambda,c,\varepsilon),\lambda,c) \leq \inf_{x\in\mathbb{R}^n} L_P(x,\lambda,c) + \varepsilon,$$

then $x(\lambda,c,\varepsilon)$ is said to be an ε-optimum of $(L_{P,\lambda,c})$.

Denote the set of all ε-optima for $(L_{P,\lambda,c})$ by

$$S_P^*(\lambda, c, \varepsilon) = \{x \in \mathbb{R}^n | L_P(x, \lambda, c) \le \inf_{x\in\mathbb{R}^n} L_P(x, \lambda, c) + \varepsilon\}.$$

when $\varepsilon = 0$, $S_P^*(\lambda, c, 0)$ reduces to $S_P^*(\lambda, c)$, the set of the optimal solutions of the Lagrangian relaxation problem $(L_{P,\lambda,c})$.

Assumption 1 $f_* = \inf_{x\in\mathbb{R}^n} f(x) > -\infty$.

Note that Assumption 1 is not restrictive at all. In fact, if $f_* = -\infty$, we can replace $f(x)$ with $e^{f(x)}$ in problem (P).

Under Assumption 1, for any given $\lambda \ge 0$ and $c > 0$, property (H_2) implies

$$\inf_{x\in\mathbb{R}^n} L_P(x, \lambda, c) \ge f_* + \frac{1}{c}\sum_{i=1}^{m} r(\lambda_i).$$

Therefore, no matter whether a global optimal solution exists or not, under Assumption 1, for any $\lambda \ge 0$, $c > 0$ and $\varepsilon > 0$, $S_p^*(\lambda, c, \varepsilon) \ne \emptyset$ always holds. Based on this property, we propose the following convergent augmented Lagrangian method associated with augmented Lagrangian function L_P that does not require an exact solution of $(L_{P,\lambda,c})$.

PALM: Given sequence $\{\varepsilon_k\}$ with $\varepsilon_k \to 0^+$ $(k \to \infty)$ and sequence $\{\lambda^k\} \subset R_+^m$. Choose the penalty parameter $c_k \to 0$ $(k \to \infty)$ to satisfy the following condition:

$$\lim_{k\to\infty} \frac{1}{c_k}\sum_{i=1}^{m} |r(\lambda_i^k)| = 0. \tag{1}$$

Find an $x(\lambda^k, c_k, \varepsilon_k) \in S_P^*(\lambda^k, c_k, \varepsilon_k)$, and let $x^k = x(\lambda^k, c_k, \varepsilon_k)$.

If the relaxation problem $(L_{P,\lambda,c})$ is globally solvable, i.e., $S_P^*(\lambda^k, c_k) \ne \emptyset$, then take any $x(\lambda^k, c_k) \in S_P^*(\lambda_k, c_k)$, and let $x^k = x(\lambda^k, c_k)$.

We prove next the convergence of the sequence $\{x^k\}$ produced by *PALM*.

Define the following for $\alpha > 0$,

$$G(\alpha) = \{x \in \mathbb{R}^n | g_i(x) \le \alpha, 1 \le i \le m\}.$$

when $\alpha = 0$, $G(0)$ is the feasible set of the primal problem (P). We assume, in this paper, that $G(0) \ne \emptyset$.

Define the perturbation function of (P) as:

$$\beta_f(\alpha) = \inf\{f(x) | x \in G(\alpha)\}.$$

when $\alpha = 0, \beta_f(0)$ is the optimal value of (P). It is easy to verify that $\beta_f(\cdot)$ is upper semi-continous at point 0.

Let

$$F(\alpha) = \{x \in \mathbb{R}^n | f(x) \le \beta_f(0) + \alpha\}.$$

Lemma 1 *For any* $\lambda \ge 0$, $c > 0$ *and* $\varepsilon > 0$, *we have*

$$S_p^*(\lambda, c, \varepsilon) \subseteq \{x \in \mathbb{R}^n | L_P(x, \lambda, c) \le \beta_f(0) + \varepsilon\}.$$

Proof when $x \in G(0)$, we have the following from (H_1),

$$P(cg_i(x), \lambda_i) \le 0, \ (1 \le i \le m). \tag{2}$$

Therefore, for any $\bar{x} \in S_P^*(\lambda, c, \varepsilon)$, we have the following by (2),

$$\begin{aligned} L_P(\bar{x}, \lambda, c) &\leq \inf\{L_P(x, \lambda, c) | x \in \mathbb{R}^n\} + \varepsilon \\ &\leq \inf\{L_P(x, \lambda, c) | x \in G(0)\} + \varepsilon \\ &= \inf\{f(x) + \frac{1}{c}\sum_{i=1}^{m} P(cg_i(x), \lambda_i) | x \in G(0)\} + \varepsilon \\ &\leq \beta_f(0) + \varepsilon. \end{aligned}$$

□

Lemma 2 *Suppose that Assumption 1 is satisfied. Then, for any $\varepsilon > 0$, there exists k_ε such that*

$$\{x \in \mathbb{R}^n | L_P(x, \lambda^k, c_k) \leq \beta_f(0) + \varepsilon\} \subseteq G(\varepsilon),$$

when $k \geq k_\varepsilon$, where the selection of c_k satisfies (1), the condition specified in $PALM$.

Proof We prove it by contradiction. Suppose that there exist an $\varepsilon_0 > 0$ and an infinite subsequence $N \subseteq \{1, 2, \ldots\}$ such that, for $k \in N$, we have

$$z^k \in \{x \in \mathbb{R}^n | L_P(x, \lambda^k, c_k) \leq \beta_f(0) + \varepsilon_0\}, \tag{3}$$

$$z^k \notin G(\varepsilon_0). \tag{4}$$

From (4), there exist an $i_0 \in \{1, 2, \ldots, m\}$ and an infinite subsequence $N_0 \subseteq N$ such that, for $k \in N_0$, we have

$$g_{i_0}(z^k) > \varepsilon. \tag{5}$$

If $a < +\infty$, from (4), we get $z^k \in \Omega_{c_k}(a)$, i.e. $g_{i_0}(z^k) < \frac{\varepsilon}{c_k}$. Therefore, $\limsup_{k\to\infty} g_{i_0}(z^k) \leq 0$, which is a contradiction to (5). Now, let $a = +\infty$ and therefore, for any $k \in N_0$, we get the following from (4), (5), (H_1) and (H_2),

$$\begin{aligned} \beta_f(0) + \varepsilon &\geq L_P(z^k, \lambda^k, c_k) \\ &= f(z^k) + \frac{1}{c_k}\sum_{i=1}^{m} P(c_k g_i(z^k), \lambda_i^k) \\ &\geq f_* + \frac{1}{c_k} P(c_k g_{i_0}(z^k), \lambda_{i_0}^k) + \frac{1}{c_k}\sum_{i\neq i_0} P(c_k g_i(z^k), \lambda_i^k) \\ &\geq f_* + \frac{1}{c_k} P(c_k \varepsilon_0, \lambda_{i_0}^k) + \frac{1}{c_k}\sum_{i\neq i_0} r(\lambda_i^k) \\ &\geq f_* + \frac{1}{c_k} P(c_k \varepsilon_0, \lambda_{i_0}^k) - \frac{1}{c_k}\sum_{i=1}^{m} |r(\lambda_i^k)|. \end{aligned}$$

Taking the limit with respect to $k \in N_0$ in the last inequality and using (1) and (H_3) give rise to a contradiction. □

Lemma 3 *For any $\varepsilon > 0$, there exists a k_ε such that the following holds when $k \geq k_\varepsilon$,*

$$\{x \in \mathbb{R}^n | L_P(x, \lambda^k, c_k) \leq \beta_f(0) + \frac{\varepsilon}{2}\} \subseteq F(\varepsilon),$$

where the selection of c_k satisfies (1), the condition specified in $PALM$.

Proof From (1), there exists a k_ε such that the following holds when $k \geq k_\varepsilon$,

$$\frac{1}{c_k}\sum_{i=1}^{m}|r(\lambda^k)| \leq \frac{\varepsilon}{2}. \tag{6}$$

Therefore, for all $k \geq k_\varepsilon$ and all $\bar{x} \in \{x \in \mathbb{R}^n | L_P(x, \lambda, c) \leq \beta_f(0) + \frac{\varepsilon}{2}\}$, we get the following from (6),

$$\begin{aligned} f(\bar{x}) &= L_P(\bar{x}, \lambda^k, c_k) - \frac{1}{c_k}\sum_{i=1}^{m} P(c_k g_i(\bar{x}), \lambda_i^k) \\ &\leq \beta_f(0) + \frac{\varepsilon}{2} - \frac{1}{c_k}\sum_{i=1}^{m} r(\lambda_i^k) \\ &\leq \beta_f(0) + \frac{\varepsilon}{2} + \frac{1}{c_k}\sum_{i=1}^{m} |r(\lambda_i^k) \\ &\leq \beta_f(0) + \varepsilon. \end{aligned}$$

□

Next, we prove the global convergence theorem for $PALM$.

Theorem 1 *Suppose that Assumption 1 is satisfied and* $\{x^k\}$ *generated by* $PALM$ *admits limit points.*

(a) *For any* $\varepsilon > 0$, $S_P^*(\lambda^k, c_k, \varepsilon_k) \subseteq G(\varepsilon) \cap F(\varepsilon)$ *when* k *is large enough;*
(b) *For any limit point,* x^*, *of* $\{x^k\}$, $x^* \in X^*$.

Proof (a) Since $\varepsilon_k \to 0^+$ $(k \to \infty)$, for any $\varepsilon > 0$, we have $\varepsilon_k \leq \frac{\varepsilon}{2}$ when k is large enough. Therefore, from Lemmas 1 – 3, we have the following when k is large enough,

$$\begin{aligned} S_P^*(\lambda^k, c_k, \varepsilon_k) &\subseteq S_P^*(\lambda^k, c_k, \frac{\varepsilon}{2}) \\ &\subseteq \{x \in \mathbb{R}^n | L_P(x, \lambda^k, c_k) \leq \beta_f(0) + \frac{\varepsilon}{2}\} \\ &\subseteq G(\varepsilon) \cap F(\varepsilon). \end{aligned}$$

(b) For any $\varepsilon > 0$, when k is large enough, we have the following from (a),

$$x^k \in G(\varepsilon) \cap F(\varepsilon). \tag{7}$$

Since f and g_i $(1 \leq i \leq m)$ are continuous, $G(\varepsilon)$ and $F(\varepsilon)$ are closed sets. Therefore, we have $x^* \in G(\varepsilon) \cap F(\varepsilon)$ from (7). Since $\varepsilon > 0$ is arbitrary, we obtain $x^* \in G(0)$ and $f(x^*) \leq \beta_f(0)$, i.e., $x^* \in X^*$. □

If $\{x^k\}$ has no limit point, i.e., $\{x^k\}$ is divergent (see [26]), we consider the convergence of sequence $\{f(x^k)\}$. In the following, we present the necessary and sufficient condition under which $\{f(x^k)\}$ is convergent to $\beta_f(0)$, the optimal value of problem (P).

Theorem 2 *Suppose that Assumption 1 is satisfied. Then*

$$\lim_{k\to\infty} f(x^k) = \beta_f(0) \tag{8}$$

if and only if $\beta_f(\alpha)$ *is lower semi-continuous at* $\alpha = 0$.

Proof Sufficiency: Suppose that $\beta_f(\alpha)$ is lower semi-continuous at $\alpha = 0$. From Theorem 1, for any $\varepsilon > 0$, relation (7) holds when k is large enough. Therefore,

$$\beta_f(\varepsilon) \le f(x^k) \le \beta_f(0) + \varepsilon. \tag{9}$$

Since $\beta_f(\alpha)$ is lower semi-continuous at point $\alpha = 0$, we get the following from (9),

$$\begin{aligned}\beta_f(0) &\le \lim_{\varepsilon\to 0^+} \inf \beta_f(\varepsilon) \le \lim_{k\to\infty} \inf f(x^k)\\ &\le \lim_{k\to\infty} \sup f(x^k) \le \beta_f(0).\end{aligned}$$

Thus,

$$\lim_{k\to\infty} f(x^k) = \beta_f(0).$$

Necessity: We prove it by contradiction. Suppose that (8) is satisfied and there exist $\delta_0 > 0$ and $\varepsilon_j \to 0^+$ $(j \to \infty)$ such that, for any j, we have

$$\beta_f(\varepsilon_j) \le \beta_f(0) - \delta_0. \tag{10}$$

By the continuity of $P(\cdot,\cdot)$ and (H_1), for each k, we can get a sufficiently large j_k such that

$$c_k\varepsilon_{j_k} < a \quad \text{and} \quad P(c_k\varepsilon_{j_k}, \lambda_i^k) \le \eta, (1 \le i \le m). \tag{11}$$

Here, $\eta > 0$ is a constant. Then, for each k, there exists $z^k \in G(\varepsilon_{j_k})$ such that

$$f(z^k) \le \beta_f(\varepsilon_{j_k}) + \frac{\delta_0}{2}. \tag{12}$$

From (10) and (12), we obtain

$$f(z^k) \le \beta_f(0) - \frac{\delta_0}{2}. \tag{13}$$

Furthermore, we get from (11) the following for $i = 1, 2, \ldots, m$,

$$c_k g_i(z^k) \le c_k\varepsilon_{j_k} < a.$$

We can now conclude $z^k \in \Omega_c(a)$. We also have the following from (11), (13), (H_1) and (H_2),

$$\begin{aligned}f(x^k) &= L_P(x^k, \lambda^k, c_k) - \frac{1}{c_k}\sum_{i=1}^m P(c_k g_i(x^k), \lambda_i^k)\\ &\le \inf_{x\in\mathbb{R}^n} L_P(x, \lambda^k, c_k) + \varepsilon_k - \frac{1}{c_k}\sum_{i=1}^m r(\lambda_i^k)\\ &\le f(z^k) + \frac{1}{c_k}\sum_{i=1}^m P(c_k g_i(z^k), \lambda_i^k) + \varepsilon_k + \frac{1}{c_k}\sum_{i=1}^m |r(\lambda_i^k)\\ &\le \beta_f(0) - \frac{1}{2}\delta_0 + \frac{1}{c_k}\sum_{i=1}^m P(c_k\varepsilon_k, \lambda_i^k) + \varepsilon_k + \frac{1}{c_k}\sum_{i=1}^m |r(\lambda_i^k)\\ &\le \beta_f(0) - \frac{1}{2}\delta_0 + \frac{1}{c_k}\eta + \varepsilon_k + \frac{1}{c_k}\sum_{i=1}^m |r(\lambda_i^k)|.\end{aligned}$$

Therefore, we get

$$f(x^k) \leq \beta_f(0) - \frac{1}{2}\delta_0 + \frac{1}{c_k}\eta + \varepsilon_k + \frac{1}{c_k}\sum_{i=1}^{m}|r(\lambda_i^k)|. \tag{14}$$

Taking the limit in both sides of (14), we get the following from (1) and (8),

$$\beta_f(0) = \lim_{k\to\infty} f(x^k) \leq \beta_f(0) - \frac{\delta_0}{2},$$

which is a contradiction. The proof is completed. □

3.2 *RALM* and its convergence

Consider the following relaxation problem of L_R,

$$(L_{R,\lambda,c}) \quad \min_{x\in\mathbb{R}^n} L_R(x, \lambda, c).$$

Define the set of ε-optimal solutions of $(L_{R,\lambda,c})$ as:

$$S_R^*(\lambda, c, \varepsilon) = \{x \in \mathbb{R}^n | L_R(x, \lambda, c) \leq \inf_{x\in\mathbb{R}^n} L_R(x, \lambda, c) + \varepsilon\},$$

and the set of optima of $(L_{R,\lambda,c})$ as $S_R^*(\lambda, c)$.

Assumption 2 $g_* = \inf_{x\in\mathbb{R}^n} \min_{1\leq i\leq m} g_i(x) > -\infty$.

If there exists i_0 such that $\inf_{x\in R^n} g_{i_0}(x) = -\infty$, we can replace $g_{i_0}(x) \leq 0$ with $e^{g_{i_0}(x)} - 1 \leq 0$ in problem (P).

Note that under Assumptions 1 and 2, we have the following using the property (H_1'),

$$\inf_{x\in\mathbb{R}^n} L_R(x, \lambda, c) \geq f_* + \frac{1}{c}\sum_{i=1}^{m} R(cg_*, \lambda_i) > -\infty, \ \forall\, \lambda \geq 0, \ \forall\, c > 0.$$

Therefore, $S_R^*(\lambda, c, \varepsilon) \neq \emptyset$. Based on this property, we present the following convergent augmented Lagrangian method associated with L_R that does not require an exact solution of $(L_{R,\lambda,c})$.

RALM: Given any parameters $r \in (-\infty, g_*)$, sequence $\{\varepsilon_k\}$ with $\varepsilon_k \to 0^+$ $(k \to \infty)$ and sequence $\{\lambda^k\} \subset \mathbb{R}_+^m$. Choose a penalty parameter $c_k \to +\infty$ $(k \to \infty)$ that satisfies the following conditions:

$$\lim_{k\to\infty} \frac{1}{c_k}\sum_{i=1}^{m} R(rc_k, \lambda_i^k) = 0. \tag{15}$$

Find an $x(\lambda^k, c_k, \varepsilon_k) \in S_R^*(\lambda^k, c_k, \varepsilon_k)$, and let $x^k = x(\lambda^k, c_k, \varepsilon_k)$.

Note that we have $g_* \leq 0$ and $r < 0$ from $G(0) \neq \emptyset$. Therefore, from the property (H_2') of R, *RALM* is well-posed under Assumptions 1 and 2.

Compared with *PALM*, an additional assumption, i.e., Assumption 2, is needed for *RALM*, since condition (H_2') is weaker than (H_2). As the proof of its convergence theorem is similar to that of *PALM*, we give the following theorems of convergence results for *RALM* with their proofs omitted.

Theorem 3 *Suppose that Assumptions 1 and 2 are satisfied and $\{x^k\}$ admits limit points. Then*

(a) $S_R^*(\lambda^k, c_k, \varepsilon_k) \subseteq G(\varepsilon) \cap F(\varepsilon)$ *for any* $\varepsilon > 0$ *and sufficiently large* k;
(b) *For any limit point*, x^*, *of* $\{x^k\}$, $x^* \in X^*$.

Theorem 4 *Suppose that Assumptions 1 and 2 are satisfied. Then*

$$\lim_{k\to\infty} f(x^*) = \beta_f(0)$$

if and only if $\beta_f(\alpha)$ *is lower semi-continous at* $\alpha = 0$.

4 Multiplier algorithms based on $PALM$ and $RALM$

In this section, we derive two augmented Lagrangian multiplier algorithms by applying, respectively, the convergent augmented Lagrangian methods $PALM$ and $RALM$ developed in the previous section.

Algorithm 1 (Multiplier algorithm based on $PALM$): Take a parameter $\eta \in (0, a)$.

Step 0. Select initial points $x^0 \in \mathbb{R}^n$, $\lambda^0 \geq 0$ and $c_0 > 0$. Set $k := 0$
Step 1. Compute

$$\lambda_i^{k+1} = P_s'(c_k g_i(x^k), \lambda_i^k),\ 1 \leq i \leq m,$$
$$c_{k+1} \geq (k+1) \max\left\{1, \sum_{i=1}^{m} |r(\lambda_i^{k+1})|\right\}, \tag{16}$$

where P_s' is the derivative of $P(s,t)$ with respect to its first variable s.
Step 2. Find $x^{k+1} \in S_P^*(\lambda^{k+1}, c_{k+1})$.
Step 3. Set $k := k+1$ and go to Step 1.

From (16) in Step 1, we can verify that the condition given in (1) is satisfied. Therefore, from Theorems 1 and 2, we obtain the following global convergence of Algorithm 1.

Theorem 5 *Suppose that Assumption 1 is satisfied. Then the sequence of iteration points,* $\{x^k\}$, *generated by Algorithm 1 has the following properties:*

(a) *If* $\{x^k\}$ *admits a limit point* x^*, *then* $x^* \in X^*$;
(b) $\lim_{k\to\infty} f(x^k) = \beta_f(0)$ *if and only if* $\beta_f(\alpha)$ *is lower semi-continuous at* $\alpha = 0$.

Next, we analyze some other convergence properties of Algorithm 1. Let $x^* \in G(0)$ be a K-T point of primal problem (P), i.e., there exists $\lambda^* \in \mathbb{R}^m$ with $\lambda^* \geq 0$, such that

$$f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) = 0, \tag{17}$$

$$\lambda_i^* g_i(x^*) = 0,\ 1 \leq i \leq m, \tag{18}$$

where $\nabla g_i(\cdot)$ is the gradient of $g_i(\cdot)$. Let

$$I(x^*) = \{1 \leq i \leq m | g_i(x^*) = 0\},\quad J(x^*) = \{i \in I(x^*) | \lambda_i^* > 0\}.$$

The linear independence constraint qualification holds at x^* if $\{\nabla g_i(x^*)\}_{i\in I(x^*)}$ are linearly independent, while the M-F constraint qualification holds at x^* if there exists $d \in \mathbb{R}^n$ such that

$$\nabla g_i(x^*)^T d < 0,\quad i \in I(x^*). \tag{19}$$

Note that the linear independence constraint qualification implies the M-F constraint qualification.

Theorem 6 *Suppose that Assumption 1 is satisfied. Then, the sequence of iteration points, $\{x^k\}$, generated by Algorithm 1 satisfies the following properties:*

(a) *If $\lim_{k\to\infty} x^k = x^*$ and the M-F constraint qualification is satisfied at x^*, then $\{\lambda^k\}$ is bounded and any of its limit points, λ^*, is such that (x^*, λ^*) is a K-T point.*
(b) *If $\lim_{k\to\infty} x^k = x^*$ and the linearly independent constraint qualification holds at x^*, then*

$$\lim_{k\to\infty} \lambda^k = \lambda^*$$

and (x^*, λ^*) is a K-T point.

Proof Let $\lim_{k\to\infty} x^k = x^*$. Theorem 5 implies $x^* \in X^*$. (a): When $i \notin I(x^*)$, we have $g_i(x^*) < 0$. Since $x^k \to x^*$ $(k \to \infty)$, there exist $\varepsilon_0 > 0$ and k_0 such that $g_i(x^k) \leq -\varepsilon_0$ when $k \geq k_0$. Therefore, when $i \in I(x^*)$, we have

$$\lim_{k\to\infty} c_k g_i(x^k) = -\infty. \tag{20}$$

From (H_4) and Step 1, we have the following when $k \geq k_0$,

$$\lambda_k^{k+1} = P_s'(c_k g_i(x^k), \lambda_i^k) \leq \lambda_i^k.$$

We can conclude that $\{\lambda_i^k\}$ is bounded when $i \notin I(x^*)$. From (20) and (H_4),

$$\lim_{k\to\infty} \lambda_i^k = 0, \tag{21}$$

when $i \notin I(x^*)$. Next, we prove that $\{\lambda_i^k\}$ is bounded when $i \in I(x^*)$. If this is not true, there exists an infinite subsequence $N \subseteq \{1, 2, \ldots\}$ such that

$$T_k = \sum_{i\in I(x^*)} \lambda_i^k \to +\infty,\ k \in N, k \to \infty. \tag{22}$$

From $x^{k-1} \in S_P^*(\lambda^{k-1}, c_{k-1})$ and Step 1, we get

$$\nabla f(x^{k-1}) + \sum_{i=1}^{m} \lambda_i^k \nabla g_i(x^{k-1}) = 0. \tag{23}$$

Without loss of generality, we may assume that $\frac{\lambda_i^k}{T_k} \to \tilde{\lambda}_i^*$, $(k \in N, k \to \infty)$. Dividing both sides of (23) by T_k, taking limit with respect to $k \in N$, and using (21) and (22), we get

$$\sum_{i\in I(x^*)} \tilde{\lambda}_i^* \nabla g_i(x^*) = 0.$$

As not all $\tilde{\lambda}_i^*, i \in I(x^*)$, are zero, this leads to a contradiction to the M-F constraint qualification. Therefore, $\{\lambda^k\}$ is bounded. Suppose that λ^* is a limit point of $\{\lambda^k\}$. From (21), $\lambda_i^* = 0$ for $i \notin I(x^*)$, and from (23) we get

$$\nabla f(x^*) + \sum_{i\in I(x^*)} \lambda_i^* \nabla g_i(x^*) = 0, \tag{24}$$

which proves (a).

For (b), since the M-F constraint qualification is implied by linearly independent constraint qualification, from (a), (x^*, λ^*) is a K-T point of primal problem (P), where λ^* is any limit point of $\{\lambda^k\}$. From (24) and the linear independence of $\{\nabla g_i(x^*)\}_{i\in I(x^*)}$, we further conclude that the bounded infinite sequence $\{\lambda^k\}$ has a unique limit point, i.e, $\lambda^k \to \lambda^*, k \to \infty$. □

Next, we give the multiplier algorithm based on $RALM$.

Algorithm 2 (Multiplier algorithm based on $RALM$): Take parameters $r \in (-\infty, g_*)$.

Step 0. Select initial points $x^0 \in \mathbb{R}^n$, $\lambda^0 \geq 0$, and $c_0 > 0$. Let $k := 0$.
Step 1. Compute

$$\lambda_i^{k+1} = R_s'(c_k g_i(x^k), \lambda_i^k), 1 \leq i \leq m,$$
$$c_{k+1} \geq (k+1) \max \left\{ 1, \sum_{i=1}^{m} R(rc_{k+1}, \lambda_i^{k+1}) \right\}, \tag{25}$$

where R_s' is the derivative of $R(s,t)$ with respect to its first variable s.
Step 2. Find an $x^{k+1} \in S_R^*(\lambda^{k+1}, c_{k+1})$.
Step 3. Set $k := k+1$ and go to Step 1.

Note from (25) in Step 1 that the condition given in (15) is satisfied. Therefore, a satisfaction of Theorems 3 and 4 guarantees the global convergence of Algorithm 2, as stated in the following theorem.

Theorem 7 *Suppose that Assumptions 1 and 2 are satisfied, then the sequence of the iteration points, $\{x^k\}$, generated by Algorithm 2 has the following properties:*

(a) *If $\{x^k\}$ admits a limit point x^*, then $x^* \in X^*$;*
(b) $\lim_{k\to\infty} f(x^k) = \beta_f(0)$ *if and only if $\beta_f(\alpha)$ is lower semi-continous at $\alpha = 0$.*

Similar to the proof of Theorem 6, we can get the following theorem.

Theorem 8 *Suppose that Assumptions 1 and 2 are satisfied. Then, the sequence of iteration points, $\{x^k\}$, generated by Algorithm 2 satisfies the following properties.*

(a) *If $\lim_{k\to\infty} x^k = x^*$, and the M-F constraint qualification is satisfied at x^*, then $\{\lambda^k\}$ is bounded and any limit point of $\{\lambda^k\}$, λ^*, is such that (x^*, λ^*) is a K-T point.*
(b) *If $\lim_{k\to\infty} x^k = x^*$, and the linear independence qualification constraint holds at x^*, then*

$$\lim_{k\to\infty} \lambda^k = \lambda^*$$

and (x^*, λ^*) is a K-T point.

Remark 1 The global convergence properties of the augmented Lagrangian multiplier algorithms in [2,9,18,26] all require the assumption that the multiplier sequence is bounded, while our Algorithms 1–2 remove this restrictive condition. Theorems 5 and 7 show that any limit point of $\{x^k\}$ is the optimal solution of the primal problem without a necessary satisfaction of the boundedness condition of the multiplier sequence.

Remark 2 Rockafellar [26] showed that if the boundedness of the multiplier sequence is not assumed, then the *essential quadratic* augmented Lagrangian method may be divergent, that is, the sequence $\{x^k\}$ may not have a limit point. Theorems 5 and 7 show that if $\{x^k\}$ has no limit point, then the sequence of the objective value is convergent to the optimal value if and only if the perturbation function is lower semi-continuous at zero.

Table 1 Solution process by PALM with essentially quadratic augmented Lagrangian function

k	c_k	λ_k	x_k	$f(x_k)$
1	10	(1, 1, 1)	(0.9765, −1.0720, 1.1080)	−7.9545
2	10	(1, 1, 2)	(0.9729, −1.0445, 1.0931)	−7.6604
3	1000	(10, 10, 10)	(1.0028, −0.9978, 0.9930)	−6.9615
4	1000	(20, 20, 20)	(1.0067, −0.9946, 0.9836)	−6.9088

Table 2 Solution process by the multiplier algorithm based on PALM with essentially quadratic augmented Lagrangian function

k	c_k	λ_k	x_k	$f(x_k)$
1	10	(1, 1, 1)	(0.9765, −1.0720, 1.1080)	−7.9545
2	100	(2.2181, 0, 3.1461)	(1.0004, −1.0006, 0.9993)	−7.0018
3	100	(1.9993, 0, 3.3294)	(1.0003, −1.0000, 0.9997)	−7.0000
4	100	(2.0001, 0, 3.3298)	(1.0005, −1.0000, 0.9995)	−7.0000

Example 7 We consider the following example problem in [34],

$$\begin{aligned}
&\min 5x_1x_2x_3 - \frac{1}{2}x_1^2 + 10(x_1-1)^2 - 2x_2x_3 - x_3 - \frac{3}{2}x_2^2 - x_3^2\\
&\text{s.t.} - x_1^2 - x_3^2 - x_1 - 2x_2 - x_3 + 2 \geq 0\\
&\quad x_1 + \frac{3}{4} \geq 0\\
&\quad (x_1 - x_3)^2 + x_2^3 - 0.1x_1 + 0.05x_1^2 + 1.05 \geq 0.
\end{aligned}$$

It can be verified that the optimal solution of this example is $x^* = (1, -1, 1)$. Adopting the essentially quadratic augmented Lagrangian function $L_{P_2}(x, \lambda, c)$ with $P_2(s, t) = \min_{\tau \geq s}\{t\tau + \phi_2(\tau)\}$ in algorithm $PALM$ generates the solution sequence in Table 1, where initial point $x_0 = (0, 0, 0)$ is chosen. Adopting the multiplier algorithm based on PALM, with essentially quadratic augmented Lagrangian function discussed in Sect. 4, yields the solution process given in Table 2, where initial point $x_0 = (0, 0, 0)$ is chosen. The optimal multiplier of this example is $\lambda^* = (2, 0, \frac{10}{3})$.

5 Augmented Lagrangian algorithm based on L_V

The augmented Lagrangian algorithm associated with L_V is similar to Algorithm 1 except for the ways of choosing initial points and the penalty parameter c_k.

Assumption 3 The relaxation problem of augmented Lagrangian L_V is globally solvable.

Algorithm 3

Step 0. Select initial points $x^0 \in \mathbb{R}^n$, $\lambda^0 > 0$ and $c_0 > 0$. Set $k := 0$.

Step 1. Compute

$$\lambda_i^{k+1} = V_s'(c_k g_i(x^k), \lambda_i^k),\ 1 \le i \le m,$$
$$c_{k+1} \ge (k+1) \max\left\{1, \sum_{i=1}^{m} |r(\lambda_i^{k+1})|\right\}, \tag{26}$$

where V_s' is the derivative of $V(s,t)$ with respect to its first variable s.
Step 2. Find an $x^{k+1} \in S_V^*(\lambda^{k+1}, c_{k+1})$.
Step 3. Set $k := k+1$ and go to Step 1.

Theorem 9 *Suppose that $\{x^k\}$ is the iteration sequence generated by Algorithm* 3. *If*

$$\lim_{k\to\infty} x^k = x^*, \tag{27}$$

then $x^* \in X^*$.

Proof Let $\bar{x} \in G(0)$. From $x^k \in S_V^*(\lambda^k, c_k)$ and (H_1''), we get

$$\begin{aligned} f(x^k) &= L_V(x^k, \lambda^k, c_k) - \frac{1}{c_k}\sum_{i=1}^{m} V(c_k g_i(x^k), \lambda_i^k) \\ &\le L_V(\bar{x}, \lambda^k, c_k) - \frac{1}{c_k}\sum_{i=1}^{m} V(c_k g_i(x^k), \lambda_i^k) \\ &\le f(\bar{x}) - \frac{1}{c_k}\sum_{i=1}^{m} V(c_k g_i(x^k), \lambda_i^k). \end{aligned} \tag{28}$$

We claim that $x^* \in G(0)$. If $a < +\infty$, it is easy to get $x^* \in G(0)$. Now let $a = +\infty$. If $x^* \notin G(0)$, there exists index i_0 such that $g_{i_0}(x^*) > 0$. From (27), there exist $\varepsilon_0 > 0$ and k_0 such that

$$g_{i_0}(x^k) \ge \varepsilon_0, \tag{29}$$

when $k \ge k_0$. From the convexity of $V(\cdot, t)$ and (H_1''), we get

$$\begin{aligned} \lambda_{i_0}^k c_k g_{i_0}(x^k) &\le V(c_k g_{i_0}(x^k), \lambda_{i_0}^k) \\ &= V(c_k g_{i_0}(x^k), \lambda_{i_0}^k) - V(0, \lambda_{i_0}^k) \\ &\le V_s'(c_k g_{i_0}(x^k), \lambda_{i_0}^k) \cdot c_k g_{i_0}(x^k). \end{aligned}$$

From the above inequalities, (29) and (26), we get

$$\lambda_{i_0}^k \le V_s'(c_k g_{i_0}(x^k), \lambda_{i_0}^k) = \lambda_{i_0}^{k+1}, \tag{30}$$

when $k \ge k_0$. Due to the positive initial multiplier λ^0 and (H_4''), from Step 1, we get that $\lambda^k > 0$ for all k. From (30), we also have $\lambda_{i_0}^k \ge \lambda_{i_0}^{k_0} > 0$ for all $k \ge k_0$. Therefore, $\{\lambda_{i_0}^k\}_{k\ge k_0} \subset [\lambda_{i_0}^{k_0}, +\infty)$, i.e. $\{x^k\}_{k\ge k_0}$ is in a closed set. From (H_3''), we get

$$\frac{V(c_k\varepsilon_0, \lambda_{i_0}^k)}{c_k} \to +\infty,\ (k \to \infty). \tag{31}$$

From the last inequality of (28), combining (H_1''), (H_2'') and (29) gives the following for $k \geq k_0$,

$$\begin{aligned} f(x^k) &= f(\bar{x}) - \frac{1}{c_k} V(c_k g_{i_0}(x^k), \lambda_{i_0}^k) - \frac{1}{c_k} \sum_{i \neq i_0} V(c_k g_i(x^k), \lambda_i^k) \\ &\leq f(\bar{x}) - \frac{1}{c_k} V(c_k \varepsilon_0, \lambda_{i_0}^k) - \frac{1}{c_k} \sum_{i \neq i_0} r(\lambda_i^k) \\ &\leq f(\bar{x}) - \frac{1}{c_k} V(c_k \varepsilon_0, \lambda_{i_0}^k) + \frac{1}{c_k} \sum_{i=1}^{m} |r(\lambda_i^k). \end{aligned}$$

Taking $k \to \infty$ in the above inequality, we get a contradiction from (27), (31) and Step 1. Thus, $x^* \in G(0)$. From (H_2'') and the last inequality of (28), we get

$$f(x^k) \leq f(\bar{x}) + \frac{1}{c_k} \sum_{i=1}^{m} |r(\lambda_i^k)|.$$

Taking $k \to \infty$ in the above inequality, we get the following from Step 1,

$$f(x^*) \leq f(\bar{x}).$$

From the arbitrariness of $\bar{x} \in G(0)$, we get $x^* \in X^*$. □

As the same as in Algorithms 1 and 2, the convergence result for Algorithm 3 does not need the condition of the boundedness of $\{\lambda^k\}$.

Similar to the proof of Theorem 6, we have the following theorem.

Theorem 10 *Suppose that* $\{x^k\}$ *is the iteration sequence generated by Algorithm* 3.

(a) *If* $\lim_{k\to\infty} x^k = x^*$ *and the M-F constraint qualification is satisfied at* x^**, then* $\{\lambda^k\}$ *is bounded and any limit point of* $\{\lambda^k\}$, λ^*, *is such that* (x^*, λ^*) *is a K-T point.*
(b) *If* $\lim_{k\to\infty} x^k = x^*$ *and the linear independence constraint qualification holds at* x^**, then*

$$\lim_{k\to\infty} \lambda^k = \lambda^*$$

and (x^*, λ^*) *is a K-T point.*

6 Existence theorems of global saddle point of L_P, L_R and L_V

In this section, we generalize the existence theorems for global saddle point in [32] and [26]. More specifically, we prove the existence of a global saddle point for L_P and L_R under the framework of $PALM$ or $RALM$. Using a similar method, we prove the existence of a global saddle point for L_V with an additional condition. Finally, for six important augmented Lagrangian functions classified in L_P, L_R and L_V, respectively, we discuss the second-order sufficient condition which ensures the existence of their local saddle points.

We first give the definition of a saddle point of L_P.

Definition 5 (x^*, λ^*) is said to be a global saddle point of L_P for $c_0 > 0$, if for all $x \in \mathbb{R}^n$ and all $\lambda \geq 0$,

$$L_P(x^*, \lambda, c_0) \leq L_P(x^*, \lambda^*, c_0) \leq L_P(x, \lambda^*, c_0). \tag{32}$$

If $x^* \in \Omega_{c_0}(a)$ and there exists a neighborhood $N(x^*, \delta)$ of x^* such that for all $x \in N(x^*, \delta)$ and all $\lambda \geq 0$, (32) is satisfied, then (x^*, λ^*) is said to be a local saddle point of L_P for $c_0 > 0$. (x^*, λ^*) is said to be a local saddle point of L_P for $c \geq c_0$, if there exists a neighborhood $N(x^*, \delta)$ of x^*, which is independent of c, such that for all $x \in N(x^*, \delta)$, all $\lambda \geq 0$, and all $c \geq c_0$, (32) is satisfied.

We can define a global (local) saddle point of L_R and L_V in similar ways.

Applying the property (H_1) of $P(s, t)$ that $P(s, t) \to +\infty$ $(t \to +\infty)$ for $s > 0$, we can get the following theorem without a detailed proof, as it is similar to a one in [32].

Theorem 11 *Suppose that (x^*, λ^*) is a global (local) saddle point of L_P for c_0, then x^* is a global (local) optimal solution of the primal problem (P).*

Results similar to Theorem 11 can be also obtained for L_R and L_V. Theorem 11 shows that the saddle point condition of L_P is sufficient for the optimal solution of the primal problem, which is the same as in the classical Lagrangian. It is well recognized that saddle point often does not exist for the classical Lagrangian.

Next, we will present an existence theorem for the global saddle point of L_P.

Theorem 12 *Suppose that Assumption 1 is satisfied. Assume that the following assumptions also hold.*

(a) *There exists $\lambda^* \geq 0$ such that for any $x^* \in X^*$, there exists $c_* > 0$ such that for any $c \geq c_*$, (x^*, λ^*) is a local saddle point of L_P.*
(b) *There exists $\alpha_0 > 0$ such that $G(\alpha_0) \cap F(\alpha_0)$ is bounded.*

Then, for any $x^* \in X^*$, there exists $c^* > 0$ such that for all $c \geq c^*$, (x^*, λ^*) is the global saddle point of L_P.

Proof Let $x^* \in X^*$ be arbitrary. From assumption (a) of the theorem, there exist a neighborhood $N(x^*, \delta)$ of x^* and $c_* > 0$ such that for all $x \in N(x^*, \delta)$, all $\lambda \geq 0$ and all $c \geq c_*$,

$$L_P(x^*, \lambda, c) \leq L_P(x^*, \lambda^*, c) \leq L_P(x, \lambda^*, c).$$

From the first inequality given above and (H_1), for $c \geq c_*$, we have

$$P(cg_i(x^*), \lambda^*) = 0, \qquad (1 \leq i \leq m).$$

Then,

$$L_P(x^*, \lambda^*, c) = f(x^*), \ \forall\, c \geq c_*.$$

Due to the above equality, we now only need to prove that there exists $c^* \geq c_*$ such that for any $c \geq c^*$,

$$f(x^*) \leq L_P(x, \lambda^*, c), \ \forall\, x \in \mathbb{R}^n.$$

Suppose on the contrary, there exist $c_k \to \infty (k \to \infty)$ and $\varepsilon_k \to 0^+$ $(k \to \infty)$ such that

$$\inf_{x \in \mathbb{R}^n} L_P(x, \lambda^*, c_k) + \varepsilon_k < f(x^*). \tag{33}$$

Choose an ε_k- optimal solution $z^k \in S_p^*(\lambda^*, c_k, \varepsilon_k)$. From (33), for any k, we have

$$L_P(z^k, \lambda^*, c_k) < f(x^*). \tag{34}$$

On the other hand, for $\alpha_0 > 0$, the following holds true from Theorem 1,

$$z^k \in S_p^*(\lambda^*, c_k, \varepsilon_k) \subseteq G(\alpha_0) \cap F(\alpha_0),$$

when k is sufficiently large. From assumption (b) of the theorem, we know that $\{z^k\}$ is bounded and it has an accumulation point x^*. Without loss of generality, we set

$$\lim_{k\to\infty} z^k = z^*. \tag{35}$$

From Theorem 1, $z^* \in X^*$. Then from assumption (a) of the theorem, there exist $c_{**} > 0$ and a neighborhood $N(z^*, \delta^*)$ of z^* such that for all $x \in N(z^*, \delta^*)$, all $\lambda \geq 0$, and all $c \geq c_{**}$, we have

$$L_P(z^*, \lambda, c) \leq L_P(z^*, \lambda^*, c) \leq L_P(x, \lambda^*, c). \tag{36}$$

Adopting the same approach as above, the first inequality of (36) and $f(x^*) = f(z^*)$ imply that for $c \geq c_{**}$,

$$L_P(z^*, \lambda^*, c) = f(x^*). \tag{37}$$

From (35), when k is large enough, $z^k \in N(z^*, \delta^*)$. Taking $x = z^k$ in (36), (37) yields

$$f(x^*) \leq L_P(z^k, \lambda^*, c), \;\; c \geq c_{**},$$

when k is sufficiently large, which is a contradiction of (34). □

Note that $P(s,t)$ and $R(s,t)$ are convex with respect to s in all examples given in Sect. 2. The following theorem shows that under the assumption that $P(s,t)$ is convex in s, Theorem 12 can be simplified.

Let $P(s,t)$ be convex with respect to s on $\mathbb{R}_a \times \mathbb{R}_+$. Using the property of convex function, we note that for some $c_0 > 0$, $x \notin \Omega_{c_0}(a)$ implies $x \notin \Omega_c(a)$ for $c \geq c_0$. Then

$$\frac{1}{c}\sum_{i=1}^{m} P(cg_i(x), \lambda_i^*)$$

is increasing with respect to c for any $x \in \Omega_c(a)$, which implies $L_P(x, \lambda^*, c)$ is increasing with respect to c. Thus, from the inequality of the saddle point and (H_1), we get the following Lemma.

Lemma 4 *Let $P(s,t)$ be convex with respect to s on $\mathbb{R}_a \times \mathbb{R}_+$. If (x^*, λ^*) is a global (local) saddle point of L_P for some $c_0 > 0$, then (x^*, λ^*) is a global (local) saddle point for any $c \geq c_0$.*

Thus, under the assumption that $P(s,t)$ is convex with respect to s, Lemma 4 and Theorem 12 yield the following Theorem.

Theorem 13 *Suppose that Assumption 1 is satisfied. Assume also that the following hold.*

(a) *There exists $\lambda^* \geq 0$ such that for any $x^* \in X^*$, (x^*, λ^*) is a local saddle point of L_P for some $c_* > 0$;*
(b) *There exists $\alpha_0 > 0$ such that $G(\alpha_0) \cap F(\alpha_0)$ is bounded.*

Then, for any $x^* \in X^*$, there exists $c^* > 0$ such that for all $c \geq c^*$, (x^*, λ^*) is the global saddle point of L_P.

We can prove the existence of a global saddle point for L_R via $RALM$ in a manner similar to the way in proving Theorems 12–13.

Theorem 14 *Supposed that Assumptions 1 and 2 are satisfied. Assume also that the following hold.*

(a) *There exists $\lambda^* \geq 0$ such that for any $x^* \in X^*$, there exists $c_* > 0$ such that for all $c \geq c_*$, (x^*, λ^*) is a local saddle point of L_R.*
(b) *There exists $\alpha_0 > 0$ such that $G(\alpha_0) \cap F(\alpha_0)$ is bounded.*

Then for any $x^* \in X^*$, there exists $c^* > 0$ such that for all $c \geq c^*$, (x^*, λ^*) is a global saddle point of L_R.

It is obvious that when $R(s, t)$ is convex with respect to s, Lemma 4 is satisfied. Thus, we get the following simplified result under an assumption that $R(s, t)$ is convex with respect to s.

Theorem 15 *Suppose that Assumptions 1 and 2 are satisfied. Assume also that the following hold.*

(a) *There exists $\lambda^* \geq 0$ such that for any $x^* \in X^*$, (x^*, λ^*) is a local saddle point of L_R for some $c_* > 0$.*
(b) *There exists $\alpha_0 > 0$ such that $G(\alpha_0) \cap F(\alpha_0)$ is bounded.*

Then for any $x^* \in X^*$ there exists $c^* > 0$ such that for all $c \geq c^*$, (x^*, λ^*) is a global saddle point of L_R.

In the following, we consider the global saddle point of L_V. We generalize the separation condition in [32]. Under the generalized separation condition, we prove the existence of a global saddle point of L_V using an approach similar to the proof of Theorem 12.

Define the following for a given $\lambda^* \geq 0$,

$$J^* = \{1 \leq i \leq m | \lambda_i^* > 0\},$$

and

$$G_{J^*}(\alpha) = \{x \in \mathbb{R}^n | g_i(x) \leq \alpha, i \in J^*\}, (\alpha > 0).$$

Similar to the proofs of Lemma 2 and Lemma 3, we can obtain the following Lemma 5 by utilizing the property of $V(s, t)$.

Lemma 5 *Suppose that Assumption 1 is satisfied and $c_k \to +\infty$ $(k \to \infty)$. Then for any $\varepsilon > 0$, there exists k_ε such that*

$$\{x \in \mathbb{R}^n | L_V(x, \lambda^*, c_k) \leq \beta_f(0)\} \subseteq G_{J^*}(\varepsilon) \cap F(\varepsilon)$$

when $k \geq k_\varepsilon$.

Theorem 16 *Suppose that Assumption 1 is satisfied. Assume also that the following hold.*

(a) *There exists $\lambda^* \geq 0$ such that for any $x^* \in X^*$, (x^*, λ^*) is a local saddle point of L_V for some $c_* > 0$.*
(b) *There exists $\alpha_0 > 0$ such that $G_{J^*}(\alpha_0) \cap F(\alpha_0)$ is bounded.*
(c) *$G_{J^*}(0) \cap F(0) = X^*$.*

Then for any $x^* \in X^*$ there exist $c^* > 0$ such that for all $c \geq c^*$, (x^*, λ^*) is a global saddle point of L_V.

Proof Let $x^* \in X^*$ be arbitrary. (H_1'') implies that $V(s,t)$ is convex with respect s. Then Lemma 4 is also satisfied for L_V. Assumption (a) then implies that there exist $\lambda^* \geq 0$, a neighborhood $N(x^*, \delta)$ of x^* and $c_* > 0$ such that

$$L_V(x^*, \lambda, c) \leq L_V(x^*, \lambda^*, c) \leq L_V(x, \lambda^*, c),$$

for all $x \in N(x^*, \delta)$, all $\lambda \geq 0$ and all $c \geq c_*$. From the first inequality above and (H_1''), we have

$$L_V(x^*, \lambda^*, c) = f(x^*),$$

for $c \geq c_*$. By the above equality , we only need to prove that there exists $c^* > 0$ such that for any $x \in \mathbb{R}^n$,

$$f(x^*) \leq L_V(x, \lambda^*, c^*). \tag{38}$$

we consider the following two cases.

Case 1. $J^* = \emptyset$. Then $\lambda_i^* = 0$ $(1 \leq i \leq m)$ and $G_{J^*}(0) = \mathbb{R}^n$. From assumption (c) of the theorem, $F(0) = X^*$, that is , $x^* \in X^*$ is the optimal solution of problem $\min\limits_{x \in \mathbb{R}^n} f(x)$. On the other hand, (H_1'') implies that

$$V(cg_i(x), 0) \geq 0,$$

for $c > 0$ and $x \in \Omega_c(a)$. Therefore

$$f(x^*) \leq f(x) \leq L_V(x, 0, c)$$

for $x \in \mathbb{R}^n$ and $c > 0$. We can conclude that (38) is satisfied.

Case 2. $J^* \neq \emptyset$. Suppose on the contrary that (38) is unsatisfied. Then, there exist $c_k \to +\infty$ $(k \to \infty)$ and $z^k \in \Omega_{c_k}(a)$ such that

$$L_V(z^k, \lambda^*, c_k) < f(x^*). \tag{39}$$

Let $\varepsilon \in (0, \alpha_0]$ be arbitrary. From (39) and Lemma 5, when k is large enough, we have

$$z^k \in \{x \in \mathbb{R}^n | L_V(x, \lambda^*, c_k) \leq f(x^*)\} \subseteq G_{J^*}(\varepsilon) \cap F(\varepsilon). \tag{40}$$

From (40) and assumption (b), $\{z^k\}$ is bounded and then must have an accumulation point z^*. Without loss of generality, we set

$$\lim_{k \to \infty} z^k = z^*. \tag{41}$$

Since f and $g_i (1 \leq i \leq m)$ are continuous, $G_{J^*}(\varepsilon)$ and $F(\varepsilon)$ are bounded sets. Therefore (41) implies that

$$z^* \in G_{J^*}(\varepsilon) \cap F(\varepsilon).$$

By assumption (c) of the theorem and the arbitrariness of $\varepsilon > 0$, we obtain

$$z^* \in G_{J^*}(0) \cap F(0) = X^*. \tag{42}$$

Thus, similar to the proof of the second half of Theorem 12, (42) and Assumption (a) will lead to a contradiction of (39). □

The existence theorems of a global saddle point derived above are achieved under the existence condition of local saddle points. Thus, it is indispensable to discuss some sufficiency conditions for the existence of a local saddle point. Without doubt, such sufficiency conditions depend on specific structures of different situations under investigation. At this stage, we are unable to provide a unified sufficiency condition applied for all the three unified formulations as a whole. Instead, we will derive the second-order sufficiency conditions for the six specific examples discussed in Sect. 2, respectively.

Second-order sufficiency condition (A): Let $x^* \in G(0)$ and $\lambda^* \geq 0$. Assume that (x^*, λ^*) is a K-T point of problem (P) and the Hessian matrix of the conventional Lagrangian

$$\nabla^2_{xx} L(x^*, \lambda^*) = \nabla^2 f(x^*) + \sum_{i \in J(x^*)} \lambda^* \nabla^2 g_i(x^*)$$

is positive definite on the cone

$$\widetilde{K}(x^*)=\{d \in \mathbb{R}^n, d \neq 0 | \nabla g_i(x^*)^T d = 0,\ i \in J(x^*);\ \nabla g_i(x^*)^T d \leq 0,\ i \in I(x^*) \backslash J(x^*)\}.$$

Second-order sufficiency condition (B): Let $x^* \in G(0)$ and $\lambda^* \geq 0$. Assume that (x^*, λ^*) is a K-T point of problem (P) and the Hessian matrix of the conventional Lagrangian, $\nabla^2_{xx} L(x^*, \lambda^*)$ is positive definite on the cone

$$K(x^*) = \{d \in \mathbb{R}^n, d \neq 0 | \nabla g_i(x^*)^T d = 0,\ i \in J(x^*)\}.$$

Condition (A) is weaker than condition (B). For L_{P_i} $(i = 3, 4)$, L_{R_1} and L_{V_1}, condition (B) is proved to be the sufficiency condition for the existence of a local saddle point in [32]. We prove in this paper that condition (A) is a sufficiency condition for the existence of a local saddle point of Example L_{P_i} $(i = 1, 2)$.

Theorem 17 *Let x^* be a local optimal solution to problem (P). Assume that second-order sufficiency condition (A) is satisfied at x^*. Then there exists $c^* > 0$ such that for all $c \geq c^*$, (x^*, λ^*) is a local saddle point of L_{P_i} $(i = 1, 2)$.*

Proof Except for some minor differences, the proof is similar to a one in [32]. □

In the following, we consider an example problem in which the global optimal solution is not unique.

Example 8

$$\begin{aligned} \min\ & f(x) = e^{x_2^2 - x_1^2}, \\ s.t.\ & g_1(x) = x_1^2 - 1 \leq 0, \\ & g_2(x) = e^{-x_2} - 1 \leq 0. \end{aligned}$$

This example clearly is a nonconvex optimization problem. Note that Assumptions 1 and 2 are satisfied by $f(x)$ and $g_i(x)$ $(i = 1, 2)$, respectively, and there exist two global optimal solutions:

$$x^{*,1} = (1, 0), \quad x^{*,2} = (-1, 0).$$

It is easy to find out that $\beta_f(0) = e^{-1}$. For any $\alpha > 0$, both the sets

$$G(\alpha) = \{x \in \mathbb{R}^2 | x_1^2 - 1 \leq \alpha, e^{-x_2} - 1 \leq \alpha\}$$

and

$$F(\alpha) = \{x \in \mathbb{R}^2 | e^{x_2^2 - x_1^2} \leq e^{-1} + \alpha\}$$

are unbounded. However, the set $G(\alpha) \cap F(\alpha)$ is bounded. Set

$$\lambda_1^* = e^{-1}, \quad \lambda_2^* = 0.$$

We can verify the following for $(x^{*,i}, \lambda^*)$ $(i = 1, 2)$,

$$\begin{aligned}
\nabla_x L(x^{*,i}, \lambda^*) &= \nabla f(x^{*,i}) + \sum_{i=1}^{2} \lambda_i^* \nabla g_i(x^{*,i}) \\
&= \begin{pmatrix} -2x_1^{*,i} e^{-1} \\ 0 \end{pmatrix} + e^{-1} \begin{pmatrix} 2x_1^{*,i} \\ 0. \end{pmatrix} \\
&= 0
\end{aligned}$$

Note that the Hessian matrix

$$\begin{aligned}
\nabla_{xx}^2 L(x^{*,i}, \lambda^*) &= \nabla^2 f(x^{*,i}) + \sum_{i=1}^{2} \lambda_i^* \nabla^2 g_i(x^{*,i}) \\
&= \begin{pmatrix} 2e^{-1} & 0 \\ 0 & 2e^{-1} \end{pmatrix} + e^{-1} \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 4e^{-1} & 0 \\ 0 & 2e^{-1} \end{pmatrix}
\end{aligned}$$

is positive definite. Thus, the second sufficient condition (B) is satisfied at $(x^{*,i}, \lambda^*)$ $(i = 1, 2)$. By Theorem 17 and the corresponding theorems in [32], $(x^{*,i}, \lambda^*)$ $(i = 1, 2)$ are local saddle points for augmented Lagrangian L_{P_j} $(j = 1, 2, 3, 4)$, L_{R_1} and L_{V_1}. Notice that the separation condition is also satisfied, that is,

$$G_{J^*}(0) \cap F(0) = \{x \in \mathbb{R}^2 | x_1^2 - 1 \le 0\} \cap \{x \in \mathbb{R}^2 | e^{x_2^2 - x_1^2} \le e^{-1}\} = X^*$$

and for any $\alpha > 0$, $G_{J^*}(\alpha) \cap F(\alpha)$ is bounded. Therefore, from Theorem 13, 15 and 16, $(x^{*,i}, \lambda^*)$, $i = 1, 2$, are global saddle points for all the six classes of augmented Lagrangian functions.

Remark 3 In this section, we not only generalize the existence theorem of a global saddle point in [32], but also extend in Theorem 13 the corresponding result in [26] (see Theorem 6 and Corollary 6.1 in [26]). In fact, the assumption in [26] that x^* is the unique optimal solution to (P) in the strong sense implies a satisfaction of the Assumption (b) in Theorem 13, i.e., there exists $\alpha_0 > 0$ such that $G(\alpha_0) \cap F(\alpha_0)$ is bounded.

7 Conclusions

The convergence results for augmented Lagrangian presented in this paper offer new insights for duality theory, expand the domain for the existence of a global saddle point and facilitate the development of novel schemes for multiplier methods in global optimization. The implementation issues will be the subject of our future study, aiming for certain special classes of global optimization problems. The key to success is on identification of efficient global optimization algorithms for unconstrained Lagrangian relaxation problems with some special structures.

Acknowledgements This work was supported by the National Natural Science Foundation of China under grants 10571106 and 10701047 and by the Research Grants Council of Hong Kong under CUHK 4245/04E and 2150534. The authors are indebted to the reviewers for their valuable comments and suggestions which helped to improve the manuscript.

References

1. Ben-Tal, A., Zibuleysky, M.: Penalty/barrier multiplier methods for convex programming problems. SIAM J. Optim. **7**, 347–366 (1997)
2. Bertsekas, D.P.: Constrained Optimization and Lagrangian Multiplier Methods. Academic Press, New York (1982)
3. Conn, A.R., Gould, I.M., Toint, Ph.L.: A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bound. SIAM J. Numer. Anal. **28**, 545–572 (1991)
4. Di Pillo, G., Grippo, L.: An exact penalty function method with global convergence properties for nonlinear programming problems. Math. Prog. **36**, 1–18 (1986)
5. Di Pillio, G., Grippo, L.: Exact penalty functions in constrained optimization. SIAM J. Control Optim. **27**, 1333–1360 (1989)
6. Di Pillio, G., Lvcicd, S.: An augmented Lagrangian function with improved exactness properties. SIAM J. Optim. **12**, 376–406 (2001)
7. Goldfarb, D., Polyak, R., Scheinberg, K., Yuzefovich, I.: A modified barrier-augmented Lagrangian method for constrained minimization. Comp. Optim. Appl. **14**, 55–74 (1999)
8. Griva, I., Polyak, R.: Primal-dual nonliner rescaling method with dynamic scaling parameter update. Math. Prog. Ser. A **106**, 237–259 (2006)
9. Hartman, J.K.: Iterative determination of parameters for an exact penalty function. J. Optim. Theory Appl. **16**, 49–66 (1975)
10. Hestenes, M.R.: Multiplier and gradient methods. J. Optim. Theory. Appl. **4**, 303–320 (1969)
11. Huang, X.X., Yang, X.Q.: A unified augmented Lagrangian approach to duality and exact penalization. Math. Oper. Res. **28**, 533–552 (2003)
12. Huang, X.X., Yang, X.Q.: Further study on augmented Lagrangian duality theory. J. Glob. Optim. **31**, 193–210 (2005)
13. Kort, B.W., Bertsekas, D.P.: A new penalty method for constrained minimization. In: Proceddings of the 1972 IEEE Conference on Decision and Control, New Orleans, pp. 162–166 (1972)
14. Li, D.: Zero duality gap for a class of nonconvex optimization problems. J. Optim. Theory Appl. **85**, 309–324 (1995)
15. Li, D.: Saddle-point generation on nonlinear nonconvex optimization. Nonlinear Anal. **30**, 4339–4344 (1997)
16. Li, D., Sun, X.L.: Local convexification of the Lagrangian function in nonconvex optimization. J. Optim. Theory Appl. **104**, 109–120 (2000)
17. Mangasarian, O.L.: Unconstrained Lagrangians in nonlinear programming. SIAM J. Control **13**, 772–791 (1975)
18. Nguyen, V.H., Strodiot, J.J.: On the convergence rate of a penalty function method of exponential type. J. Optim. Theory Appl. **27**, 495–508 (1979)
19. Polyak, R.: Modified barrier functions: Theory and methods. Math. Prog. **54**, 177–222 (1992)
20. Polyak, R.: Log-sigmoid multipliers method in constrained optimization. Ann. Oper. Res. **101**, 427–460 (2001)
21. Polyak, R., Griva, I.: A Primal-dual nonlinear rescaling method with dynamic scaling parameter update, Technical report, SEOR Department, George Mason University, Fairfax, VA22030, Technical Report SEOR-11-02-2002
22. Polyak, R., Griva, I.: Nonlinear rescaling vs. smoothing technique in convex optimization. Math. Prog. **92**, 197–235 (2002)
23. Polyak, R., Griva, I.: Primal-dual nonlinear rescaling method for convex optimization. J. Optim. Theory Appl. **122**, 111–156 (2004)
24. Powell, M.J.D.: A method for nonlinear constraints in minimization problems. In: Fletcher, R. (ed.) Optimization, pp. 283–298. Academic Press, New York (1969)
25. Rockfellar, R.T.: A dual approach to solving nonlinear programming problems by unconstrained optimization. Math. Prog. **5**, 354–373 (1973)
26. Rockfellar, R.T.: Augmented Lagrange multiplier functions and duality in nonconvex programming. SIAM J. Control **12**, 268–285 (1974)

27. Rockfellar, R.T.: Augmented Lagrangians and applications of the proximal point algorithm in convex programming. Math. Oper. Res. **1**, 97–116 (1976)
28. Rockfellar, R.T.: Lagrange multipliers and optimality. SIAM Rev. **35**, 183–238 (1993)
29. Rockfellar, R.T., Wets, R.J.B.: Variational Analysis. Spriger-Verlag, Berlin (1998)
30. Rubinov, A.M., Huang, X.X., Yang, X.Q.: The zero duality gap property and lower semicontinuity of the perturbation function. Math. Oper. Res. **27**, 775–791 (2002)
31. Rubinov, A.M., Yang, X.Q.: Lagrangian-type Functions in Constrained Non-convex Optimization. Kluwer, Massachusettes (2003)
32. Sun, X.L., Li, D., Mckinnon, K.: On saddle points of augments Lagrangians for constrained nonconvex optimization. SIAM J. Optim. **15**, 1128–1146 (2005)
33. Tseng, P., Bertsekas, D.P.: On the convergence of the exponential multiplier method for convex programming. Math. Prog. **60**, 1–19 (1993)
34. Wright, M.H.: III-conditioning and computational error in interior methods for nonlinear programming. SIAM J. Optim. **9**, 84–111 (1998)
35. Xu, Z.K.: Local saddle points and convexification for nonconvex optimization problems. J. Optim. Thoeory Appl. **94**, 739–746 (1997)
36. Yang, X.Q., Huang, X.X.: A nonlinear Lagrangian approach to constrained optimization problems. SIAM J. Optim. **11**, 1119–1149 (2001)