Optimization: A Journal of Mathematical Programming and Operations Research
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gopt20

Zero duality and saddle points of a class of augmented Lagrangian functions in constrained non-convex optimization

Qian Liu [a] & Xinmin Yang [b]
[a] Department of Mathematics , Shandong Normal University , Jinan, P.R. China;
[b] Department of Mathematics , Chongqing Normal University , Chongqing, P.R. China
Published online: 28 Oct 2009.

Optimization
Vol. 57, No. 5, October 2008, 655–667

Taylor & Francis
Taylor & Francis Group

Zero duality and saddle points of a class of augmented Lagrangian functions in constrained non-convex optimization

Qian Liu[a] and Xinmin Yang[b]*

[a]*Department of Mathematics, Shandong Normal University, Jinan, P.R. China;*
[b]*Department of Mathematics, Chongqing Normal University, Chongqing, P.R. China*

(*Received 26 September 2006; final version received 21 November 2007*)

In this article, we introduce a unified class of augmented Lagrangian functions for constrained non-convex optimization problems which include many types of the augmented Lagrangians. We first get the zero duality gap property between the primal problem and the augmented Lagrangian dual problem. Then, under second-order sufficiency conditions, we prove that this class of augmented Lagrangian functions possesses local saddle points. Finally, we show the existence of global saddle points without requiring the compactness of X and the uniqueness of the global solution.

Keywords: non-convex optimization; augmented Lagrangian functions; duality gap; saddle point

AMS Subject Classifications: 90C29; 90C46

1. Introduction

We consider the constrained optimization problem:

$$(P) \quad \min\{f(x): x \in X_0\},$$

where $X_0 = \{x \in X \mid g_i(x) \leq 0, i = 1, \ldots, m\}$, f, $g_i: R^n \to R$, $i = 1, \ldots, m$ are continuous functions and $X \subseteq R^n$ is a non-empty closed set.

Among all the available methods for solving (P), the Lagrangian dual methods have received much attention. The classical Lagrangian function of (P) is defined as:

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x), \quad \lambda = (\lambda_1, \ldots, \lambda_m)^T \geq 0.$$

The corresponding Lagrangian dual problem of (P) is:

$$\sup_{\lambda \geq 0} d(\lambda),$$

*Corresponding author. Email: xmyang@cqnu.edu.cn

ISSN 0233–1934 print/ISSN 1029–4945 online

DOI: 10.1080/02331930802355416
http://www.informaworld.com

where the dual function of (P) is given by:

$$d(\lambda) = \inf_{x \in X} L(x, \lambda).$$

A key issue in the Lagrangian dual methods for solving (P) is the existence of the zero duality gap between the primal problem and the Lagrangian dual problem. It is well known that the existence of zero duality gap is equivalent to the existence of a saddle point of the Lagrangian function [9]. The existence of a saddle point thus plays a critical role in a successful adoption of Lagrangian dual methods for solving (P). A pair (x^*, λ^*) is said to be a global saddle point of $L(x, \lambda)$ if

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*) \tag{1.1}$$

for all $x \in X$ and $\lambda \geq 0$. If there exists a $\delta > 0$ such that (1.1) holds for all $\lambda \geq 0$ and $x \in X \cap N(x^*, \delta)$, where $N(x^*, \delta) = \{x \in R^n \mid \|x - x^*\| \leq \delta\}$, then (x^*, λ^*) is said to be a local saddle point of $L(x, \lambda)$. Here $\|\cdot\|$ denotes the two-norm of vectors.

Obviously, for any feasible point x and $\lambda \geq 0$, the weak duality $d(\lambda) \leq f(x)$ always holds. However, when the objective function or constraint functions, or both in problem (P) are non-convex, neither a zero duality gap nor the existence of a saddle point can be ensured using the classical Lagrangian function. In order to overcome these drawbacks, non-linear Lagrangian functions [3,18,19] and augmented Lagrangian functions [4,5,12–14,18] were introduced. The first augmented Lagrangian method was independently proposed by Hestenes [2] and Powell [11] for equality-constrained problems. Later, the approach for dealing with inequality constraints was introduced and thoroughly investigated by Rockafellar [12], where the gap was removed and the existence of global saddle points were obtained by passing to the essentially quadratic Lagrangian. Recently, Rockafellar and Wets [13] introduced a convex augmenting Lagrangian function, which includes the essentially quadratic Lagrangian functions as special cases. They established the zero duality gap property and gave a necessary and sufficient condition for the exact penalty representation. More recently, Huang and Yang [4] generalized these results by providing a generalized non-convex augmenting function.

Moreover, the results of saddle points of different kinds of augmented Lagrangian were also considered. In [1–8], a so-called p-th power reformulation of (P) and the corresponding pth augmented Lagrangian function were introduced for constrained nonconvex optimization. When p is large enough, the existence of saddle point of p-th augmented Lagrangian function can be obtained. After that, Sun et al. [15] considered four classes of augmented Lagrangian: essentially quadratic augmented Lagrangian, exponential-type augmented Lagrangian, modified barrier augmented Lagrangian and the penalized exponential-type augmented Lagrangian. They showed the existence of local and global saddle points for the four classes of augmented Lagrangians, respectively. Recently, Wang and Li [18] presented a class of approximate augmented Lagrangian method. Using this approach, the existence theorem of global saddle points was extended without requiring the compactness of X and the uniqueness of the global solution. In this article, we consider a unified class of augmented Lagrangian functions which include many types of augmented Lagrangian functions, such as Modified Courant-type augmented Lagrangian, p-th power augmented Lagrangian and augmented Lagrangians mentioned in [15], as special cases. Without using the approximate augmented Lagrangian method, we get a zero duality gap property between the primal problem and the augmented Lagrangian dual problem. Under second-order sufficiency conditions, we show that this class of augmented Lagrangian functions possesses local saddle points. Finally, neither with the compactness of X nor with the uniqueness of the global optimal solution, we prove the existence of a global saddle point for a class of augmented Lagrangians which include three types considered in [15]. The existence of a global saddle point of the exponential-type augmented Lagrangian can be ensured under an additional condition [15].

The rest of this article is organized as follows. In the next section, we present a class of augmented Lagrangian functions and derive a zero duality gap property for the primal problem and its augmented Lagrangian dual problem. In Section 3, some sufficient conditions for the existence of local and global saddle points are obtained, respectively.

2. Augmented Lagrangian and zero duality gap

At the beginning, we need to introduce some notations. Let $R_+ = \{x | x \geq 0\}$, $R_{++} = \{x | x > 0\}$ and the set $G(\alpha) = \{x \in X | g_i(x) \leq \alpha, \quad i = 1, 2, \dots, m\}$. When $\alpha = 0$, $G(0) = X_0$. The perturbation function is denoted by $v(\alpha) = \min\{f(x) | x \in G(\alpha)\}$. Define $F(\alpha) = \{x \in X | f(x) \leq v(0) + \alpha\}$ and G^* as the optimal solution set of problem (P). We assume that the set X_0 is not empty.

The augmented Lagrangian dual problem of (P) is defined by:

$$(D) \quad \max_{(\lambda, c) \geq 0} d(\lambda, c),$$

where $d(\lambda, c) = \min_{x \in X} L(x, \lambda, c)$. Here the augmented Lagrangian function $L: R^n \times R^m_+ \times R_{++} \to R$ is defined as:

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

where $x \in X$, $\lambda = (\lambda_1, \dots, \lambda_m)^T \geq 0$ and $P(s, t, c): R \times R_+ \times R_{++} \to R$ is a continuous function. We denote the objective value of problems (P) and (D) with val(P) and val(D), respectively. In this article, we will assume that the function $P(s, t, c)$ satisfies the following conditions:

(A1) $\forall\ s > 0$, $\lim_{c \to +\infty} P(s, t, c) = +\infty$; $\forall\ s \leq 0$, $\lim_{c \to +\infty} P(s, t, c) = 0$.
(A1′) $\forall\ s > 0$ and $\forall\ t > 0$, $\lim_{c \to +\infty} P(s, t, c) = +\infty$; $\forall\ s \leq 0$, $\lim_{c \to +\infty} P(s, t, c) = 0$.
(A2) $P(s, t, c)$ is non-decreasing on s and c satisfying $P(s, 0, c) \geq 0$; $P(0, t, c) = 0$.
(A3) $\forall\ s > 0$, $P(s, t, c) \to +\infty (t \to +\infty)$.
(A4) $P'_s(0, t, c) = t$.
(A5) $P(s, t, c)$ is twice continuously differentiable at $0 \in R^m$ for s and $\forall t > 0$ and $c > 0$, $P''_{ss}(0, t, c) \to +\infty (c \to +\infty)$.

We find that the following examples are special cases of $L(x, \lambda, c)$, which satisfy the assumptions (A1)–(A5).

Example 2.1 Modified Courant-type augmented Lagrangian function $L_1(x, \lambda, c)$, where

$$P(s, t, c) = \frac{1}{2c}\left(\max\{0, \sigma(cs) + t\}^2 - t^2\right).$$

The function $\sigma: R \to R$ is twice continuously differentiable and strictly convex satisfying $\sigma(0) = 0$, $\sigma'(0) = 1$, $\lim_{s \to -\infty} \sigma'(s) > 0$ and $\sigma''(0) \geq 0$. Especially, choosing $\sigma(s) = s$, we get the Modified Courant-type augmented Lagrangian function [1,2,11].

Example 2.2 Essentially quadratic augmented Lagrangian function $L_2(x,\lambda,c)$, where

$$P(s,t,c)=\begin{cases} ts+\dfrac{1}{c}\phi(cs), & t+\phi'(cs)\geq 0,\\ \min\limits_{\tau\in R}\left\{t\tau+\dfrac{1}{c}\phi(c\tau)\right\}, & t+\phi'(cs)<0.\end{cases}$$

The function $\phi:R\to R$ is twice continuously differentiable and strictly convex satisfying $\phi(0)=0$, $\phi'(0)=0$, $\phi''(0)>0$ and $\phi(s)/s\to+\infty$ $(s\to+\infty)$. Especially, choosing $\phi(s)=(1/2)s^2$, we get the augmented Lagrangian function introduced by Rockafellar [12].

Example 2.3 The penalized exponential-type augmented Lagrangian $L_3(x,\lambda,c)$, where

$$P(s,t,c)=\frac{t\psi(cs)+\xi(cs)}{c}.$$

The function $\psi:R\to R$ is a twice continuously differentiable and strictly convex function satisfying $\psi(0)=0$, $\psi'(0)=1$, $\psi''(0)>0$, $\lim_{s\to-\infty}\psi(s)>-\infty$ and $\lim_{s\to-\infty}\psi'(s)=0$. $\xi:R\to R$ is a twice continuously differentiable and strictly convex function satisfying $\xi(s)=0$ $(s\leq 0)$, $\xi(s)>0$ $(s>0)$, $\xi''(0)=0$ and $\xi(s)/s\to+\infty$ $(s\to+\infty)$.

Example 2.4 The modified barrier augmented Lagrangian $L_4(x,\lambda,c)$, where

$$P(s,t,c)=\begin{cases}\dfrac{t}{c}\varphi(cs), & s<1,\\ +\infty, & s\geq 1.\end{cases}$$

The function $\varphi:R\to R$ is a twice continuously differentiable and strictly convex function satisfying $\varphi(0)=0$, $\varphi'(0)=1$, $\varphi''(0)>0$, $\lim_{s\to-\infty}\varphi'(s)=0$ and $\lim_{s\to-\infty}\varphi(s)/s=0$. Taking $\varphi(s)=-\ln(1-t)$ or $\varphi(s)=1-1/(1-t)$, $L_3(x,\lambda)$ gives rise to the modified Frish function or the modified Carroll function [10].

Since condition (A1) implies (A1′), the above examples satisfy (A1′) too. The converse implication is not true as indicated in the following example.

Example 2.5 The exponential-type augmented Lagrangian $L_5(x,\lambda,c)$, where

$$P(s,t,c)=\frac{t}{c}\psi(cs).$$

The function ψ satisfy the same conditions as Example 2.2. When setting $\psi(s)=e^s-1$, $L_3(x,\lambda,c)$ is the exponential penalty function [1,16].

Example 2.6 The p-th power augmented Lagrangian $L_6(x,\lambda,c)$, where

$$P(s,t,c)=\begin{cases}\dfrac{t}{c}\left[\dfrac{(s+b)^c}{b^{c-1}}-b\right], & s\geq -b,\\ 0, & s\leq -b.\end{cases}$$

Here $b>0$ is a constant. If $\inf_{x\in R^n}\min_{1\leq i\leq m}g_i(x)>-\infty$ and b is sufficiently large, we get p-th power augmented Lagrangian [8].

When the function $P(s, t, c)$ satisfies condition (A1), the weak duality between problem (P) and augmented Lagrangian dual problem (D) holds, that is, $\text{val}(D) \leq \text{val}(P)$. If $v(\alpha)$ is lower semicontinuous at $\alpha = 0$ and the following assumptions are satisfied:

Assumption 1 $f_* = \inf_{x\in X} f(x) > -\infty$,

Assumption 2 $g_* = \inf_{x\in X} \min_{1\leq i\leq m} g_i(x) > -\infty$,

then the zero duality gap property can be obtained. Since X_0 is non-empty, we get $g_* \leq 0$.

THEOREM 2.1 *Suppose Assumptions* 1 *and* 2 *hold, and suppose that the function* $P(s, t, c)$ *satisfies* (A1)(A2) *or* (A1′)(A2). *If* $v(\alpha)$ *is lower semicontinuous at* $\alpha = 0$, *then the zero duality gap property between problem* (P) *and augmented Lagrangian dual problem* (D) *holds, that is,* $\text{val}(D) = \text{val}(P)$.

Proof We first choose $\bar{\lambda} \geq 0$ or $\bar{\lambda} > 0$. Since (A1) or (A1′) holds, $\forall \alpha > 0$ and $i \in \{1, 2, \ldots, m\}$, we have $\lim_{c\to+\infty} P(\alpha, \bar{\lambda}_i, c) = +\infty$ and $\lim_{c\to+\infty} P(g_*, \bar{\lambda}_i, c) = 0$, which derive that

$$\lim_{c\to+\infty} \min_{i\in\{1,2,\ldots,m\}} P(\alpha, \bar{\lambda}_i, c) = +\infty \tag{2.1}$$

and

$$\lim_{c\to+\infty} \min_{i\in\{1,2,\ldots,m\}} P(g_*, \bar{\lambda}_i, c) = 0. \tag{2.2}$$

Thus from (2.1) and (2.2), there exists $c_0 > 0$ such that

$$\begin{aligned}
\min_{x\in X\backslash G(\alpha)} L(x, \bar{\lambda}, c) &= \min_{x\in X\backslash G(\alpha)} \left\{ f(x) + \sum_{i=1}^{m} P(g_i(x), \bar{\lambda}_i, c) \right\} \\
&\geq f_* + \min_{i\in\{1,2,\ldots,m\}} P(\alpha, \bar{\lambda}_i, c) + (m-1) \min_{i\in\{1,2,\ldots,m\}} P(g_*, \bar{\lambda}_i, c) \\
&> \min_{x\in X} L(x, \bar{\lambda}, c)
\end{aligned}$$

for all $c > c_0$. Hence,

$$\begin{aligned}
\min_{x\in X} L(x, \bar{\lambda}, c) &= \min_{x\in G(\alpha)} L(x, \bar{\lambda}, c) \\
&= \min_{x\in G(\alpha)} \left\{ f(x) + \sum_{i=1}^{m} P(g_i(x), \bar{\lambda}_i, c) \right\} \\
&\geq \min_{x\in G(\alpha)} \left\{ f(x) + mP(g_*, \bar{\lambda}_i, c) \right\}.
\end{aligned} \tag{2.3}$$

Taking limit on the third inequality of (2.3) as $c \to +\infty$, we get

$$\begin{aligned}
\lim_{c\to+\infty} \min_{x\in X} L(x, \bar{\lambda}, c) &\geq \lim_{c\to+\infty} \min_{x\in G(\alpha)} \left\{ f(x) + mP(g_*, \bar{\lambda}_i, c) \right\} \\
&= \min_{x\in G(\alpha)} f(x) = v(\alpha).
\end{aligned} \tag{2.4}$$

Since α is arbitrary and $v(\alpha)$ is lower semicontinuous at $\alpha = 0$,

$$\lim_{c\to+\infty} \min_{x\in X} L(x, \bar{\lambda}, c) \geq \liminf_{\alpha\to 0} v(\alpha) \geq v(0) = \text{val}(P).$$

Therefore,

$$\begin{aligned}\mathrm{val}(D) &= \max_{(\lambda,c)\geq 0} d(\lambda, c)\\ &\geq \lim_{c\to+\infty} d(\overline{\lambda}, c)\\ &= \lim_{c\to+\infty}\min_{x\in X} L(x, \overline{\lambda}, c)\\ &\geq \mathrm{val}(P).\end{aligned}$$

The weak duality implies $\mathrm{val}(D)\leq\mathrm{val}(P)$. So we have $\mathrm{val}(D)=\mathrm{val}(P)$, that is, the zero duality gap property holds.

3. Existence of saddle points

As the same as in the case of Lagrangian, the saddle point condition of augmented Lagrangian is a sufficient condition for the optimality of (P).

THEOREM 3.1 *Suppose $P(s,t,c)$ satisfies* (A2) *and* (A3). *If (x^*,λ^*) is a global (local) saddle point of $L(x,\lambda,c)$ for some $c>0$, then x^* is a global (local) optimal solution to* (P).

Proof Let (x^*,λ^*) be a global (local) saddle point of $L(x,\lambda,c)$, we have

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

That is,

$$\begin{aligned} f(x^*)+\sum_{i=1}^m P(g_i(x^*),\lambda_i,c) &\leq f(x^*)+\sum_{i=1}^m P(g_i(x^*),\lambda_i^*,c)\\ &\leq f(x)+\sum_{i=1}^m P(g_i(x),\lambda_i^*,c), \end{aligned}\tag{3.1}$$

for all $x\in X(X\cap N(x^*,\delta))$ and $\lambda\geq 0$. First, we claim that x^* is a feasible solution to (P). Suppose on the contrary $g_{i_0}(x^*)>0$ for some i_0. Choose $\lambda_{i_0}>0$, $\lambda_i=0$ $(\forall i\in\{1,\dots,m, i\neq i_0\})$. Since $P(s,0,c)\geq 0$, we have

$$\sum_{i=1}^m P(g_i(x^*),\lambda_i,c)\geq P(g_{i_0}(x^*),\lambda_{i_0},c).$$

By assumption (A3), when $\lambda_{i_0}\to+\infty$, $P(g_{i_0}(x^*),\lambda_{i_0},c)\to+\infty$. Thus,

$$\sum_{i=1}^m P(g_i(x^*),\lambda_i,c)\to+\infty,$$

which contradicts the first inequality of (3.1). So x^* is a feasible solution to (P).

For any feasible x, since $g_i(x)\leq 0$, $\forall i=1,\dots,m$ and assumption (A2), we get

$$P(g_i(x),\lambda_i^*,c)\leq 0,\quad \forall i=1,\dots,m. \tag{3.2}$$

The second inequality of (3.1) and (3.2) imply that, when $x\in X(X\cap N(x^*,\delta))$ is feasible,

$$f(x^*)+\sum_{i=1}^m P(g_i(x^*),\lambda_i^*,c)\leq f(x). \tag{3.3}$$

Using (A2) and the first inequality of (3.1),

$$\sum_{i=1}^{m} P(g_i(x^*), \lambda_i^*, c) \geq \sum_{i=1}^{m} P(g_i(x^*), 0, c) \geq 0. \tag{3.4}$$

Thus, (3.2) and (3.4) yield

$$\sum_{i=1}^{m} P(g_i(x^*), \lambda_i^*, c) = 0. \tag{3.5}$$

By (3.3) and (3.5), we obtain

$$f(x^*) \leq f(x),$$

whenever $x \in X(X \cap N(x^*, \delta))$ is feasible. Therefore, x^* is a global (local) optimal solution of (P).

Considering the existence of local and global saddle point of augmented Lagrangian functions, we assume that f and $g_i, i = 1, \ldots, m$ are twice continuously differentiable and the following conditions at x^* hold.

Second-order sufficient conditions: Let $x^* \in X$ be a local solution to (P). There exists $\lambda^* \geq 0$ such that

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) = 0, \tag{3.6}$$

$$\lambda_i^* g_i(x^*) = 0, \quad \forall i = 1, 2, \ldots, m \tag{3.7}$$

$$g_i(x^*) = 0 \Rightarrow \lambda_i^* > 0, \quad \forall i = 1, 2, \ldots, m \tag{3.8}$$

and the Hessian matrix

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

is positive definite on the cone $M(x^*)$, where

$$M(x^*) = \left\{ d \in R^n, \ \mathrm{d} \neq 0 | d^T \nabla g_i(x^*) = 0, i \in J(x^*) \right\},$$

$$J(x^*) = \{ i | \lambda_i^* > 0, i = 1, \ldots, m \}.$$

From Lemma 1.25 in [2], it is easy to prove that the following Lemma holds.

Lemma 3.1 *Let A be a symmetric $n \times n$ matrix and B be a positive semidefinite symmetric $n \times n$ matrix. Suppose the function $\eta(c): R \to R$ satisfies $\eta(c) \to +\infty$ $(c \to +\infty)$. Assume that $y^T A y > 0$ for any $y \neq 0$ satisfying $By = 0$. Then there exists $c_0 > 0$ such that $A + \eta(c) B$ is positive definite when $c \geq c_0$.*

Theorem 3.2 *Let x^* be a local optimal solution to (P). Suppose that second-order sufficient conditions are satisfied at x^* and $P(s, t, c)$ satisfies* (A2)–(A5). *Then there exist $c_0 > 0$ and $\delta > 0$ such that for all $c \geq c_0$, $\lambda \geq 0$ and $x \in N(x^*, \delta)$*

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

Proof From assumption (A2) and feasibility of x^*, we have

$$P(g_i(x^*), \lambda_i, c) \leq 0, \quad \forall i \in 1, \ldots, m, \tag{3.9}$$

for any $c > 0$ and $\lambda_i \geq 0$. If $i \notin J(x^*)$, then $\lambda_i^* = 0$. By noting that $P(s, 0, c) \geq 0$,

$$P(g_i(x^*), \lambda_i^*, c) = 0, \quad \forall i \notin J(x^*). \tag{3.10}$$

For $i \in J(x^*)$, second-order sufficient conditions imply that $g_i(x^*) = 0$. Hence,

$$P(g_i(x^*), \lambda_i^*, c) = 0, \quad \forall i \in J(x^*). \tag{3.11}$$

From (3.10) and (3.11), we obtain

$$L(x^*, \lambda^*, c) = f(x^*) + \sum_{i=1}^{m} P(g_i(x^*), \lambda_i^*, c) = f(x^*). \tag{3.12}$$

Thus, (3.9) and (3.12) yield

$$L(x^*, \lambda, c) = f(x^*) + \sum_{i=1}^{m} P(g_i(x^*), \lambda_i, c) \leq f(x^*) = L(x^*, \lambda^*, c).$$

Now we show that the second inequality holds. If $J(x^*) = \emptyset$, then $\lambda_i^* = 0$, which implies $g_i(x^*) < 0$, for any $i \in \{1, \ldots, m\}$. Combing with x^* be a local optimal solution to (P), there exists $\delta > 0$ such that $g_i(x) \leq 0$ and $f(x^*) \leq f(x)$, for any $x \in X \cap N(x^*, \delta)$. From assumption (A2), we have $P(g_i(x), \lambda_i^*, c) \geq 0$. So the following inequality holds:

$$f(x^*) \leq f(x) + \sum_{i=1}^{m} P(g_i(x), \lambda_i^*, c), \quad \forall x \in X \cap N(x^*, \delta),$$

that is, $L(x^*, \lambda^*, c) \leq L(x, \lambda^*, c)$. We assume in the following that $J(x^*) \neq \emptyset$. Since $P(g_i(x), 0, c) \geq 0$ for any $x \in X$, it is suffices to prove that there exist $c_0 > 0$ and $\delta > 0$ such that

$$f(x^*) \leq f(x) + \sum_{i \in J(x^*)} P(g_i(x), \lambda_i^*, c),$$

for any $c \geq c_0$ and $x \in X \cap N(x^*, \delta)$. Set

$$\widetilde{L}(x, \lambda^*, c) = f(x) + \sum_{i \in J(x^*)} P(g_i(x), \lambda_i^*, c).$$

Since $P_s'(0, t, c) = t$ by (A4), we have

$$\begin{aligned}
\nabla_x \widetilde{L}(x^*, \lambda^*, c) &= \nabla f(x^*) + \sum_{i \in J(x^*)} \nabla g_i(x^*) P'(g_i(x^*), \lambda_i^*, c) \\
&= \nabla f(x^*) + \sum_{i \in J(x^*)} \nabla g_i(x^*) P'(0, \lambda_i^*, c) \\
&= \nabla f(x^*) + \sum_{i \in J(x^*)} \lambda_i^* \nabla g_i(x^*) \\
&= \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) \\
&= 0.
\end{aligned}$$

By assumptions (A4) and (A5), we obtain

$$\begin{aligned}\nabla^2_{xx}\widetilde{L}(x^*,\lambda^*,c) &= \nabla^2 f(x^*) + \sum_{i\in J(x^*)}\left\{P''(0,\lambda_i^*,c)\nabla g_i(x^*)\nabla^T g_i(x^*) + P'(0,\lambda_i^*,c)\nabla^2 g_i(x^*)\right\}\\ &= \nabla^2 f(x^*) + \sum_{i\in J(x^*)}\left\{P''(g_i(x^*),\lambda_i^*,c)\nabla g_i(x^*)\nabla^T g_i(x^*) + \lambda_i^*\nabla^2 g_i(x^*)\right\}\\ &\geq \nabla^2_{xx}L(x^*,\lambda^*) + \eta(c)\sum_{i\in J(x^*)}\nabla g_i(x^*)\nabla^T g_i(x^*),\end{aligned}$$

where $\eta(c)=\min\{P''(0,\lambda_i^*,c)|i\in J(x^*)\}>0$. It is easy to see that $\eta(c)\to+\infty(c\to+\infty)$. From Lemma 3.1, there exists $c_0>0$ such that $\nabla^2_{xx}\widetilde{L}(x^*,\lambda^*,c)$ is positive definite for any $c\geq c_0$. Thus, x^* is a local optimal solution of $\widetilde{L}(x^*,\lambda^*,c)$. So, there exists $\delta>0$ such that

$$f(x^*)\leq f(x)+\sum_{i\in J(x^*)}P(g_i(x),\lambda_i^*,c)$$

for any $c\geq c_0$ and $x\in X\cap N(x^*,\delta)$.

Now we will further discuss the global saddle point property of $L(x,\lambda,c)$.

Theorem 3.3 *Suppose Assumptions* 1 *and* 2 *hold and the function P satisfies* (A1) *and* (A2). *If* (i) $\exists\lambda^*\geq 0$, $\forall x^*\in G^*$, *there exists* $\bar{c}>0$ *such that* (x^*,λ^*) *is a local saddle point for all* $c\geq\bar{c}$; (ii) *there exists* $\alpha_0>0$ *such that* $G(\alpha_0)\cap F(\alpha_0)$ *is bounded. Then for any* $x^*\in G^*$ *there exists* $c^*>0$ *such that* (x^*,λ^*) *is a global saddle point for all* $c\geq c^*$.

Proof From condition (i), for any $x^*\in G^*$, there exists $\bar{c}>0$ and $\bar{\delta}>0$ such that $x\in N(x^*,\bar{\delta})$, $c\geq\bar{c}$ and $\lambda\geq 0$ satisfying

$$L(x^*,\lambda,c)\leq L(x^*,\lambda^*,c)\leq L(x,\lambda^*,c). \tag{3.13}$$

Since $x^*\in G(0)$ and assumption (A2) holds, for any $i\in 1,2,\ldots,m$,

$$P(g_i(x^*),\lambda_i,c)\leq 0. \tag{3.14}$$

The first inequality of (3.13) combined with (3.14) and assumption (A2) implies that

$$P(g_i(x^*),\lambda_i^*,c)=0.$$

So $L(x^*,\lambda^*,c)=f(x^*)=\mathrm{val}(P)$. We consider the following three cases:

Case 1 $x\notin G(\alpha_0)$. Then there exists an i_0, such that $g_{i_0}>\alpha_0$. By assumption (A1), $\lim_{c\to+\infty}\min_{i\in\{1,2,\ldots,m\}}P(\alpha_0,\lambda_i^*,c)=+\infty$ and $\lim_{c\to+\infty}\min_{i\in\{1,2,\ldots,m\}}P(g_*,\lambda_i^*,c)=0$. So there exists $c_1>0$ such that for all $c\geq c_1$,

$$\begin{aligned}L(x,\lambda^*,c) &= f(x)+\sum_{i=1}^m P(g_i(x),\lambda_i^*,c)\\ &\geq f_*+P(g_{i_0}(x),\lambda_{i_0}^*,c)+\sum_{i\neq i_0}P(g_i(x),\lambda_i^*,c)\\ &\geq f_*+\min_{i\in\{1,2,\ldots,m\}}P(\alpha_0,\lambda_i^*,c)+(m-1)P(g_*,\lambda_i^*,c)\\ &\geq \mathrm{val}(P).\end{aligned}$$

Case 2 $x \notin F(\alpha_0)$. Then $f(x) > v(0) + \alpha_0$. By assumption (A1) there exists $c_2 > 0$ such that for any $c > c_2$,

$$\begin{aligned} L(x, \lambda^*, c) &= f(x) + \sum_{i=1}^{m} P(g_i(x), \lambda_i^*, c) \\ &> v(0) + \alpha_0 + mP(g_*, \lambda_i^*, c) \\ &\geq \text{val}(P). \end{aligned}$$

Case 3 $x \in G(\alpha_0) \cap F(\alpha_0)$. We will prove there exists $c_3 > 0$ such that $L(x, \lambda^*, c) \geq \text{val}(P)$ for any $c \geq c_3$. Conversely, there exists a sequence $\{c_k\} \to +\infty$ and $\{z^k\} \in G(\alpha_0) \cap F(\alpha_0)$ such that

$$L(z^k, \lambda^*, c_k) < \text{val}(P). \tag{3.15}$$

Since $G(\alpha_0) \cap F(\alpha_0)$ is bounded, the sequence $\{z^k\}$ has at least a cluster point $\bar{z}$. Without loss of generality, we can choose $z^k \to \bar{z}$ as $k \to +\infty$. In the following, we will prove that $\bar{z} \in G^*$. First, the point $\bar{z}$ is feasible. In fact, if $\bar{z}$ is not feasible, there exists i_0 such that $g_{i_0}(\bar{z}) = \beta > 0$. Since $z^k \to \bar{z}$ and g_{i_0} is continuous, there is a $K_1 > 0$ such that $g_{i_0}(z^k) > (\beta/2)$ for any $k \geq K_1$. Assumption (A1) implies that $\lim_{c\to+\infty} P((\beta/2), \lambda_{i_0}^*, c) = +\infty$ and $\lim_{c\to+\infty} P(g_*, \lambda_i, c) = 0$ $(i = 1, \ldots, m)$. So there exists $K_2 > 0$ such that

$$\begin{aligned} L(z^k, \lambda^*, c_k) &\geq f_* + P\left(\frac{\beta}{2}, \lambda_{i_0}^*, c_k\right) \sum_{i \neq i_0} P(g_*, \lambda_i^*, c_k) \\ &\geq \text{val}(P), \end{aligned}$$

for any $k \geq K_2$, which is a contradiction with condition (3.15). Hence, the point $\bar{z}$ is feasible.

Let us prove that $\bar{z}$ is a optimal solution of problem (P). Conversely, $f(\bar{z}) > \text{val}(P)$. We set $\delta = f(\bar{z}) - \text{val}(P) > 0$. Since f is continuous, $z^k \to \bar{z}$ as $k \to +\infty$ and $\lim_{c\to+\infty} P(g_*, \lambda_i^*, c) = 0$ for all $i = 1, 2, \ldots, m$, there exists $K_3 > 0$ such that $f(z^k) > \text{val}(P) + (\delta/2)$ and $\sum_{i=1}^{m} P(g_*, \lambda_i^*, c_k) > -(\delta/4)$ for all $k \geq K_3$. So

$$\begin{aligned} L(z^k, \lambda^*, c_k) &= f(z^k) + \sum_{i=1}^{m} P(g_*, \lambda_i^*, c_k) \\ &\geq \text{val}(P) + \frac{\delta}{2} - \frac{\delta}{4} \\ &\geq \text{val}(P) + \frac{\delta}{4}, \end{aligned}$$

which contradicts with condition (3.15). So we get $\bar{z} \in G^*$. By condition (i) for $\bar{z} \in G^*$ there exist $\bar{c} > 0$ and $\delta > 0$ such that for any $c \geq \bar{c}$, $\lambda \geq 0$ and $z \in N(\bar{z}, \delta)$ the following inequalities holds:

$$L(\bar{z}, \lambda, c) \leq L(\bar{z}, \lambda^*, c) = \text{val}(P) \leq L(z, \lambda^*, c). \tag{3.16}$$

Since $z^k \to \bar{z}$ as $k \to +\infty$ and condition (3.16), when k is large enough, $c_k \geq \bar{c}$ and the point $z^k \in N(\bar{z}, \delta)$. So we have

$$\text{val}(P) \leq L(z^k, \lambda^*, c_k).$$

This is a contradiction with condition (3.15). So if $x \in G(\alpha_0) \cap F(\alpha_0)$, there exists $c_3 > 0$ such that $L(x, \lambda^*, c) \geq \mathrm{val}(P)$ for any $c \geq c_3$. From the above three cases, we get that for any $x^* \in G^*$, the following inequalities hold:

$$L(x^*, \lambda, c) \leq L(\bar{x}, \lambda^*, c) \leq L(x, \lambda^*, c),$$

for all $c \geq c^*$, $\lambda \geq 0$ and $x \in X$, where $c^* = \max\{\bar{c}, c_1, c_2, c_3\}$. The proof is completed.

In the the above theorem of global saddle point, we do not require the compactness of X and the uniqueness of the global solution. Let us consider an example to illustrate that these two conditions are indeed not necessary to guarantee the existence of a global saddle point.

Example 3.1

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

The problem is non-convex and has only two global solutions: $x^{*,1} = (1,0)^T$ and $x^{*,2} = (-1,0)^T$. Obviously, Assumptions 1 and 2 are satisfied by $f(x)$ and $g_i(x)$ $(i = 1,2)$, respectively. For any $\alpha > 0$, the sets

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

and

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

are both unbounded. But the set $G(\alpha) \cap F(\alpha)$ is bounded. Set $\lambda^* = (-1,0)^T$, then we can obtain that, for $(x^{*,i}, \lambda^*)$ $(i = 1,2)$,

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

and the Hessian matrix

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

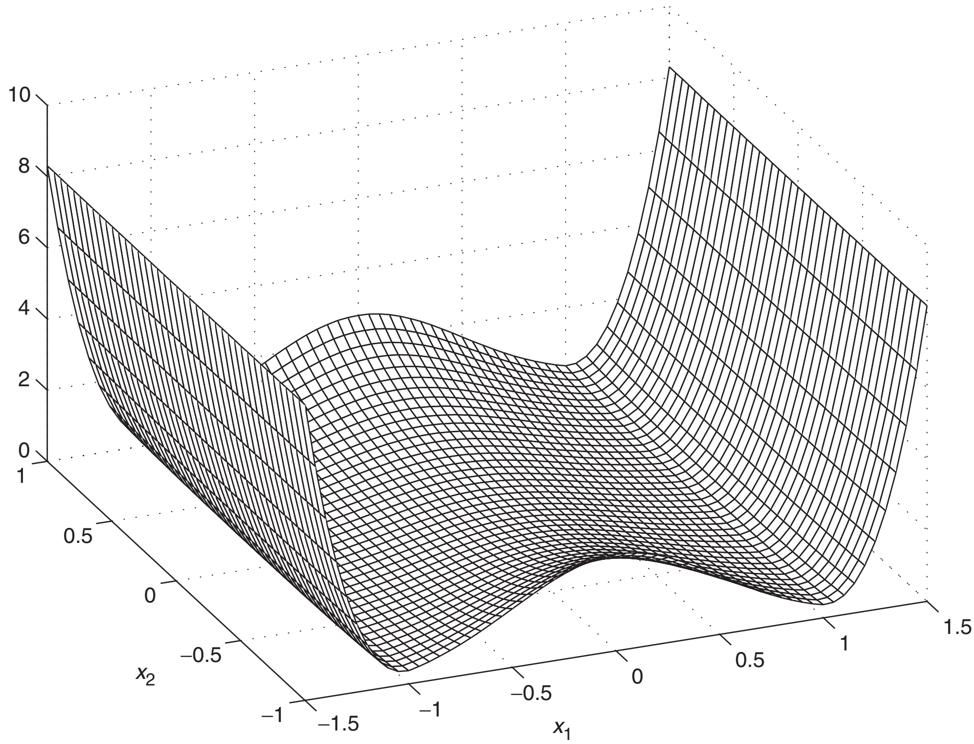

Figure 1. $L(x, \lambda^*, c)$ of example 3.1.

is definite. Thus, the second sufficient conditions are satisfied at $(x^{*,i}, \lambda^*)$ $(i = 1, 2)$. By Theorem 3.3, there exists $\bar{c} > 0$ such that for all $c \geq \bar{c}$, $(x^{*,1}, \lambda^*)$ and $(x^{*,2}, \lambda^*)$ are global saddle points for augmented Lagrangian function $L(x, \lambda, c)$. Now, we choose

$$L(x, \lambda, c) = f(x) + \sum_{i=1}^{2} P(g_i(x), \lambda_i, c),$$

where $P(s, t, c) = (1/2c)(\max\{0, cs + t\}^2 - t^2)$ and $c = 10$. Figure 1 illustrate the pictures of $L(x, \lambda^*, c)$.

Acknowledgement

This work is supported by National Natural Science Foundation under Grants 10771228 and 10571106.

References

[1] D.P. Bertsekas, *Constrained Optimization and Lagrangian Multiplier Methods*, Academic Press, New York, 1982.
[2] M.R. Hestenes, *Multiplier and gradient methods*, J. Optim. Theory App. 4 (1969), pp. 303–320.

[3] X.X. Huang and X.Q. Yang, *Approximate optimal solutions and nonlinear Lagrangian functions*, J. Global Optim. 21 (2001), pp. 51–65.
[4] ———, *A unified augmented Lagrangian approach to duality and exact penalization*, Math. Oper. Res. 28 (2003), pp. 524–532.
[5] ———, *Further study on augmented Lagrangian duality theory*, J. Global Optim. 31 (2005), pp. 193–210.
[6] D. Li, *Zero duality gap for a class of nonconvex optimization problems*, J. Optim. Theory Appl. 85 (1995), pp. 309–323.
[7] ———, *Saddle-point generation in nonlinear nonconvex optimization*, Nonlinear Anal. 30 (1997), pp. 4339–4344.
[8] D. Li and X.L. Sun, *Convexification and existence of saddle point in a p-th-power reformulation for nonconvex constrained optimization*, Nonlinear Anal. 47 (2001), pp. 5611–5622.
[9] M. Minoux, *Mathematical Programming: Theory and Algorithms*, John Wiley & Sons, New York, 1986.
[10] R. Polyak, *Modified barrier functions: Theory and methods*, Math. Program. 54 (1992), pp. 177–222.
[11] M.J.D. Powell, *A method for nonlinear constraints inminimization problems*, in *Optimization*, R. Fletcher, ed., Academic Press, New York, 1969, pp. 283–298.
[12] R.T. Rockafellar, *Augmented Lagrange multiplier functions and duality in nonconvex programming*, SIAM J. Control Optim. 12 (1974), pp. 268–285.
[13] R.T. Rockafellar and J.B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998.
[14] A.M. Rubinov, X.X. Huang, and X.Q. Yang, *The zero duality gapproperty and lower semicontinuity of the perturbation function*, Math. Oper. Res. 27 (2002), pp. 775–791.
[15] X.L. Sun, D. Li, and K. Mckinnon, *On saddle points of augmented Lagrangians for constrained nonconvex optimization*, SIAM J. Optim. 15 (2005), pp. 1128–1146.
[16] P. Tseng, *On the convergence of the exponential multiplier method for convex programming*, Math. Program. 60 (1993), pp. 1–9.
[17] C.Y. Wang, X.Q. Yang, and X.M. Yang, *Nonlinear Lagrange duality theorems and penalty function methods in continuous optimization*, J. Global Optim. 27 (2003), pp. 473–484.
[18] C.Y. Wang and D. Li, *Approximate augmented Lagrangian method in constrained global optimization*, J. Global Optim. (to appear), The Hongkong Polytechnic University, 2, 2006.
[19] C.Y. Wang, X.Q. Yang, and X.M. Yang, *Unified nonlinear Lagrangian approach to duality and optimal paths*, J. Optimiz. Theory Appl. 135 (2007), pp. 85–100.