Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2010, Article ID 324812, 13 pages
doi:10.1155/2010/324812

Research Article

Existence of Local Saddle Points for a New Augmented Lagrangian Function

Wenling Zhao, Jing Zhang, and Jinchuan Zhou

Department of Mathematics, School of Science, Shandong University of Technology, Zibo 255049, China

Correspondence should be addressed to Jing Zhang, zhangjingsecond@163.com

Received 27 March 2010; Revised 14 July 2010; Accepted 13 September 2010

Academic Editor: Joaquim J. Júdice

We give a new class of augmented Lagrangian functions for nonlinear programming problem with both equality and inequality constraints. The close relationship between local saddle points of this new augmented Lagrangian and local optimal solutions is discussed. In particular, we show that a local saddle point is a local optimal solution and the converse is also true under rather mild conditions.

1. Introduction

Consider the nonlinear optimization problem

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad i = 1, 2, \ldots, m, \\ & h_j(x) = 0, \quad j = 1, 2, \ldots, l, \\ & x \in X, \end{aligned} \tag{P}$$

where $f, g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, \ldots, m$ and $h_j : \mathbb{R}^n \to \mathbb{R}$ for $j = 1, 2, \ldots, l$ are twice continuously differentiable functions and $X \subseteq \mathbb{R}^n$ is a nonempty closed subset.

The classical Lagrangian function associated with (P) is defined as

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{l} \mu_j h_j(x), \tag{1.1}$$

where $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)^T \in \mathbb{R}_+^m$ and $\mu = (\mu_1, \mu_2, \ldots, \mu_l)^T \in \mathbb{R}^l$.

The Lagrangian dual problem (D) is presented:

$$\begin{aligned} \max \quad & \theta(\lambda, \mu), \\ \text{s.t.} \quad & \lambda \geq 0, \end{aligned} \tag{D}$$

where

$$\theta(\lambda, \mu) = \inf_{x \in X} L(x, \lambda, \mu). \tag{1.2}$$

Lagrange multiplier theory not only plays a key role in many issues of mathematical programming such as sensitivity analysis, optimality conditions, and numerical algorithms, but also has important applications, for example, in scheduling, resource allocation, engineering design, and matching problems. According to both analysis and experiments, it performs substantially better than classical methods for solving some engineering projects, especially for medium-sized or large projects.

Roughly speaking, the augmented Lagrangian method uses a sequence of iterate point of unconstrained optimization problems, which are constructed by utilizing the Lagrangian multipliers, to approximate the optimal solution of the original problem. Toward this end, we must ensure that the zero dual gap property holds between primal and dual problems. Therefore, saddle point theory received much attention, due to its equivalence with zero dual gap property. It is well known that, for convex programming problems, the zero dual gap holds by using the above classical Lagrangian function. However, the nonzero duality gap may appear for nonconvex optimization problems. The main reason is that the classical Lagrangian function is linear with respect to the Lagrangian multiplier. To overcome this drawback, various types of nonlinear Lagrangian functions and augmented Lagrangian functions have been developed in recent years. For example, Hestenes [1] and Powell [2] independently proposed augmented Lagrangian methods for solving equality constrained problems by incorporating the quadratic penalty term in the classical Lagrangian function. This was extended by Rockafellar [3] to the constrained optimization problem with both equality and inequality constraints. A convex augmented function and the corresponding augmented Lagrangian with zero duality gap property were introduced by Rockafellar and Wets in [4]. This was further extended by Huang and Yang by removing the convexity assumption imposed on the augmented functions as in [4]; see [5, 6] for the details. Wang et al. [7] proposed two classes of augmented Lagrangian functions, which are simpler than those given in [4, 5], and discussed the existence of saddle points. For other kinds of augmented Lagrangian methods refer to [8–16]; for saddle points theory and multiplier methods, refer to [17–20]. It should be noted that the sufficient conditions given in the above papers for the existence of local saddle points of augmented Lagrangian functions all require the standard second-order sufficient conditions. So, a natural question arises: whether we can exploit local saddle points under rather mild assumptions, other than the standard second-order sufficient conditions. Motivated by this, in this paper, we propose a new augmented Lagrangian function and establish the close relationship between local saddle points and local optimal solutions of the original problem. In particular, we show that this property holds under weak second-order sufficient conditions.

The paper is organized as follows. After introducing some basic notation and definitions, we mainly present sufficient conditions for the existence of a local saddle point and discuss the close relationship between a local saddle point and a local optimal solution of the original problem. Finally, an example to illustrate our result is given.

2. Notation and Definition

We first introduce some basic notation and definitions, which will be used in the sequel. Let $\mathbb{R}^n_+$ be the nonnegative orthant. For notational simplification, let

$$\begin{gathered} G(x) = \left(g_1(x), g_2(x), \ldots, g_m(x)\right)^T, \\ \lambda = (\lambda_1, \ldots, \lambda_m)^T, \qquad t = (t_1, \ldots, t_m)^T, \\ \mu = \left(\mu_1, \ldots, \mu_l\right)^T, \qquad u = (u_1, \ldots, u_m)^T. \end{gathered} \tag{2.1}$$

Definition 2.1. A pair (x^*, λ^*, μ^*) is said to be a global saddle point of $L_r(x, \lambda, \mu)$ for some $r > 0$, if

$$L_r(x^*, \lambda, \mu) \leq L_r(x^*, \lambda^*, \mu^*) \leq L_r(x, \lambda^*, \mu^*) \tag{2.2}$$

whenever $(x, \lambda, \mu) \in X \times \mathbb{R}^m_+ \times \mathbb{R}^l$. If there exists some positive scalar $\delta > 0$ such that the above inequality holds for all $(x, \lambda, \mu) \in (X \cap N(x^*, \delta)) \times \mathbb{R}^m_+ \times \mathbb{R}^l$, where $N(x^*, \delta) = \{x \in \mathbb{R}^n \mid \|x - x^*\| \leq \delta\}$, then (x^*, λ^*, μ^*) is said to be a local saddle point of $L_r(x, \lambda, \mu)$ for $r > 0$.

Definition 2.2 (weak second-order sufficient conditions). Let x^* be a feasible solution.

(1) Suppose that the KKT conditions hold at x^*; that is, there exist scalars $\lambda^*_i \geq 0$ for $i = 1, \ldots, m$ and μ^*_j for $j = 1, \ldots, l$ such that

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda^*_i \nabla g_i(x^*) + \sum_{i=1}^{m} \mu^*_j \nabla h_j(x^*) = 0, \tag{2.3}$$

$$\lambda^*_i g_i(x^*) = 0, \quad i = 1, \ldots, m. \tag{2.4}$$

(2) The Hessian matrix

$$\nabla^2_{xx} L(x^*, \lambda^*, \mu^*) = \nabla^2 f(x^*) + \sum_{i=1}^{m} \lambda^*_i \nabla^2 g_i(x^*) + \sum_{j=1}^{l} \mu^*_j \nabla^2 h_j(x^*) \tag{2.5}$$

is positive definite on the cone

$$M(x^*) = \left\{ d \in \mathbb{R}^n,\ d \neq 0 \mid \begin{array}{c} \nabla h_j(x^*)^T d = 0,\ j = 1, \ldots, l, \\ \nabla g_i(x^*)^T d = 0,\ i \in J(x^*), \\ \nabla g_i(x^*)^T d \leq 0,\ i \in I(x^*) \setminus J(x^*) \end{array} \right\}, \tag{2.6}$$

where

$$I(x^*) = \{i \mid g_i(x^*) = 0,\ i = 1, \ldots, m\},$$
$$J(x^*) = \{i \in I(x^*) \mid \lambda_i^* > 0,\ i = 1, \ldots, m\}. \tag{2.7}$$

Clearly, the above cone is included in the cone $M'(x^*) = \{d \in \mathbb{R}^n,\ d \neq 0 \mid \nabla h_j(x^*)^T d = 0,\ j = 1, \ldots, l;\ \nabla g_i(x^*)^T d = 0,\ i \in J(x^*)\}$, which is involved in the second-order sufficient condition. Hence, we refer to above conditions as weak second-order sufficient conditions.

3. Existence of Local Saddle Points

For inequality constrained optimization, Sun et al. [21] introduced a class of the generalized augmented Lagrangian function $L_r : \mathbb{R}^n \times \mathbb{R}_+^m \to \mathbb{R}$,

$$L_r(x, \lambda) = f(x) + P_r(-G(x), \lambda), \tag{3.1}$$

where $P_r(s, t) : \mathbb{R}^m \times \mathbb{R}_+^m \to \mathbb{R}$ is defined as

$$P_r(s, t) = \inf_{u - s \leq 0} \left\{ -u^T t + r\sigma(u) \right\}. \tag{3.2}$$

The function $\sigma(u)$ satisfies the following assumptions:

(A1) $\sigma(0) = 0$, and $\sigma(u) > 0$ for all $u \neq 0$;

(A2) for each $i \in \{1, 2, \ldots, m\}$, $\sigma(u)$ is nondecreasing on $u_i \geq 0$ and nonincreasing on $u_i < 0$;

(A3) $\sigma(u)$ is continuously differentiable and $\nabla\sigma(0) = 0$;

(A4) $\sigma(u)$ is twice continuously differentiable in a neighborhood of $0 \in \mathbb{R}^m$ and $h^T \nabla^2 \sigma(0) h > 0$ for all nonzero $h \in \mathbb{R}^m$.

We extend this function to treat the optimization problems with equality and inequality constraints. Consider a new augmented Lagrangian function $L_r(x, \lambda, \mu) : \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}^l \to \mathbb{R}$

$$L_r(x, \lambda, \mu) = f(x) + \sum_{j=1}^{l} \mu_j h_j(x) + \frac{r}{2} \sum_{j=1}^{l} h_j^2(x) + P_r(-G(x), \lambda), \tag{3.3}$$

where $P_r(-G(x), \lambda)$ is defined as above and the function $\sigma(u)$ satisfies (A1)–(A4). Several important augmented functions satisfy the above assumptions, as for example:

Example 3.1. $\sigma(u) = \sum_{i=1}^{m} u_i^2$.

Example 3.2. $\sigma(u) = e^{\|u\|^2} - 1$.

Under the weak second-order sufficient conditions (instead of the standard second-order sufficient conditions), we show that a local optimal solution is also a local saddle point of the augmented Lagrangian function.

Theorem 3.3. *Let x^* be a local optimal solution to problem (P). If the weak second-order sufficient conditions are satisfied at x^*, then there exist $r_0 > 0$ and $\delta > 0$ such that for any $r \geq r_0$,*

$$L_r(x^*,\lambda,\mu) \leq L_r(x^*,\lambda^*,\mu^*) \leq L_r(x,\lambda^*,\mu^*), \quad \forall(x,\lambda,\mu) \in (X \cap N(x^*,\delta)) \times \mathbb{R}^m_+ \times \mathbb{R}^l. \tag{3.4}$$

Proof. Since x^* is a feasible solution to problem (P), then $h_j(x^*) = 0$ and

$$P_r(-G(x^*),\lambda) = \inf_{u+G(x^*)\leq 0}\left\{-u^T\lambda + r\sigma(u)\right\} \leq 0^T\lambda + r\sigma(0) = 0, \quad \forall\lambda \geq 0. \tag{3.5}$$

It follows from (A1) and (2.4) that

$$\begin{aligned}
P_r(-G(x^*),\lambda^*) &= \inf_{u+G(x^*)\leq 0}\left\{-u^T\lambda^* + r\sigma(u)\right\} \\
&\geq \inf_{u+G(x^*)\leq 0}\left\{-u^T\lambda^*\right\} \\
&= \sum_{i=1}^m g_i(x^*)\lambda_i^* \\
&= 0.
\end{aligned} \tag{3.6}$$

Combining the last two inequalities yields $P_r(-G(x^*),\lambda^*) = 0$. Hence,

$$\begin{aligned}
L_r(x^*,\lambda,\mu) &= f(x^*) + \sum_{j=1}^l \mu_j h_j(x^*) + \frac{r}{2}\sum_{j=1}^l h_j^2(x^*) + P_r(-G(x^*),\lambda) \\
&\leq f(x^*) \\
&= L_r(x^*,\lambda^*,\mu^*).
\end{aligned} \tag{3.7}$$

We obtain the left inequality of (2.2) as desired.

To show the right inequality of (2.2), it suffices to prove that it holds for some r_0 and δ, since $L_r(x,\lambda,\mu)$ is nondecreasing in $r > 0$. Suppose the contrary that such r_0 and δ do not exist. Then for each positive integer k, there must exist $x^k \in X$ such that $x^k \to x^*$ and

$$L_k\left(x^k,\lambda^*,\mu^*\right) < L_k(x^*,\lambda^*,\mu^*) = f(x^*). \tag{3.8}$$

Define $u_G^k = (u_{G_1}^k, u_{G_2}^k, \ldots, u_{G_m}^k)$ as follows:

$$u_{G_i}^k = \begin{cases} 0, & i \in \{1,2,\ldots,m\} \setminus I(x^*), \\ 0, & i \in J_1^k(x^*), \\ -g_i\left(x^k\right), & i \in J_2^k(x^*), \\ -g_i\left(x^k\right), & i \in J(x^*), \end{cases} \tag{3.9}$$

where

$$\begin{aligned} J_1^k(x^*) &= \{I(x^*) \setminus J(x^*)\} \cap \left\{i \mid g_i\left(x^k\right) \le 0\right\}, \\ J_2^k(x^*) &= \{I(x^*) \setminus J(x^*)\} \cap \left\{i \mid g_i\left(x^k\right) \ge 0\right\}. \end{aligned} \tag{3.10}$$

For x^k, we have

$$\begin{aligned} P_{\mathrm{r}}\left(-G\left(x^k\right), \lambda^*\right) &= \inf_{u_i + g_i(x^k) \le 0,\ 1 \le i \le m} \left\{ \sum_{i \in J(x^*)} - u_i^T \lambda_i^* + k\sigma(u) \right\} \\ &= \sum_{i \in J(x^*)} g_i\left(x^k\right) \lambda_i^* + k\sigma\left(u_G^k\right). \end{aligned} \tag{3.11}$$

Three cases may be considered.

Case 1. When $i \in \{\{1,2,\ldots,m\} \setminus I(x^*)\} \cup J_1^k(x^*)$, take u_i $(i \in J_2^k(x^*) \cup J(x^*))$. Since $\lambda_i^* = 0$, then the original point is a minimizer of

$$\inf_{u_i + g_i(x^k) \le 0,\ i \in \{\{1,2,\ldots,m\} \setminus I(x^*)\} \cup J_1^k(x^*)} \left\{ \sum_{i \in J(x^*)} - u_i^T \lambda_i^* + k\sigma(u) \right\}. \tag{3.12}$$

Case 2. When $i \in J_2^k(x^*)$, taking into account to the fact that the function $\sum_{i \in J(x^*)} -u_i^T \lambda_i^* + k\sigma(u)$ is decreasing on u_i in $(-\infty, -g_i(x^k)]$, then $u_i = -g_i(x^k)$ is a minimizer of

$$\inf_{u_i + g_i(x^k) \le 0, i \in J_2^k(x^*)} \left\{ \sum_{i \in J(x^*)} - u_i^T \lambda_i^* + k\sigma(u) \right\}. \tag{3.13}$$

Case 3. When $i \in J(x^*)$, let $\nabla\sigma(u)_i$ be the ith component of vector $\nabla\sigma(u)$, for any $0 \le u_i \le -g_i(x^k)$. We get from (A3) that

$$\nabla\sigma(u)_i \le \frac{\lambda_i^*}{k}, \quad \forall k \in N, \tag{3.14}$$

that is, $-\lambda_i^* + k\nabla\sigma(u)_i \le 0$, and this implies that $\sum_{i\in J(x^*)} -u_i^T\lambda_i^* + k\sigma(u)$ is decreasing on u_i in $(0, -g_i(x^k)]$. So $u_i = -g_i(x^k)$ is a minimizer of

$$\inf_{u_i+g_i(x^k)\le 0,\ i\in J(x^*)} \left\{ \sum_{i\in J(x^*)} - u_i^T\lambda_i^* + k\sigma(u) \right\}. \tag{3.15}$$

Therefore,

$$\begin{aligned}
0 &> f\left(x^k\right) - f(x^*) + \sum_{j=1}^{l}\mu_j^* h_j\left(x^k\right) + \frac{k}{2}\sum_{j=1}^{l} h_j^2\left(x^k\right) + \sum_{i\in J(x^*)} g_i\left(x^k\right)\lambda_i^* + k\sigma\left(u_G^k\right) \\
&= L\left(x^k, \lambda^*, \mu^*\right) - L(x^*, \lambda^*, \mu^*) + \frac{k}{2}\sum_{j=1}^{l}\left\{h_j(x^*) + \nabla h_j(x^*)\left(x^k - x^*\right) + o\left(\left\|x^k - x^*\right\|\right)\right\}^2 \\
&\quad + k\sigma(0) + k\nabla\sigma(0)^T u_G^k + \frac{k}{2} u_G^k \nabla^2\sigma(0) u_G^k + ko\left(\left\|u_G^k\right\|^2\right).
\end{aligned} \tag{3.16}$$

Since $\nabla_x L(x^*, \lambda^*, \mu^*) = 0$, then

$$L\left(x^k, \lambda^*, \mu^*\right) - L(x^*, \lambda^*, \mu^*) = \frac{1}{2}\left(x^k - x^*\right)^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*)\left(x^k - x^*\right) + o\left(\left\|x^k - x^*\right\|\right)^2. \tag{3.17}$$

Hence

$$\begin{aligned}
0 &> \frac{1}{2}\left(x^k - x^*\right)^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*)\left(x^k - x^*\right) + o\left(\left\|x^k - x^*\right\|\right)^2 \\
&\quad + \frac{k}{2}\sum_{j=1}^{l}\left\{\nabla h_j(x^*)\left(x^k - x^*\right) + o\left(\left\|x^k - x^*\right\|\right)\right\}^2 \\
&\quad + \frac{k}{2} u_G^k \nabla^2\sigma(0) u_G^k + ko\left(\left\|u_G^k\right\|^2\right).
\end{aligned} \tag{3.18}$$

Set $d_k = (x^k - x^*)/\|x^k - x^*\|$, which is bounded, we can assume without loss of generality that d_k converges to d with $\|d\| = 1$. It follows from (3.18) that

$$
\begin{aligned}
0 > {} & \frac{1}{2} d_k^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) d_k + \frac{o(\|x^k - x^*\|)^2}{\|x^k - x^*\|^2} \\
& + \frac{k}{2} \sum_{j=1}^{l} \left\{ \nabla h_j(x^*) d_k + \frac{o(\|x^k - x^*\|)}{\|x^k - x^*\|} \right\}^2 \\
& + \frac{k}{2} \left(\frac{u_G^k}{\|x^k - x^*\|} \right)^T \nabla^2 \sigma(0) \left(\frac{u_G^k}{\|x^k - x^*\|} \right) + k \frac{o\left(\|u_G^k\|^2\right)}{\|x^k - x^*\|^2}.
\end{aligned}
\tag{3.19}
$$

Let l be the small eigenvalue of $\nabla^2 \sigma(0)$. Then $l > 0$ by assumption. We claim that $u_G^k/\|x^k - x^*\|$ converges to zero. Suppose the contrary, that for $x^k \to x^*$, we have

$$
\begin{aligned}
& \frac{k}{2} \left(\frac{u_G^k}{\|x^k - x^*\|} \right)^T \nabla^2 \sigma(0) \left(\frac{u_G^k}{\|x^k - x^*\|} \right) + ko\left(\left\| \frac{u_G^k}{x^k - x^*} \right\|^2 \right) \\
& \quad \geq \frac{kl}{2} \left\| \frac{u_G^k}{x^k - x^*} \right\|^2 + ko\left(\left\| \frac{u_G^k}{x^k - x^*} \right\|^2 \right).
\end{aligned}
\tag{3.20}
$$

Taking limits in the above inequality as $k \to +\infty$, the right hand converges to $+\infty$, which contradicts (3.19). So $u_G^k/\|x^k - x^*\| \to 0$ as claimed.

Noting that (3.9) amounts to

$$
u_{G_i}^k = \begin{cases}
0, & i \in \{1, 2, \ldots, m\} \setminus I(x^*), \\
0, & i \in J_1^k(x^*), \\
-\nabla g_i(x^*)^T (x^k - x^*) + o(\|x^k - x^*\|), & i \in J_2^k(x^*), \\
-\nabla g_i(x^*)^T (x^k - x^*) + o(\|x^k - x^*\|), & i \in J(x^*),
\end{cases}
\tag{3.21}
$$

then

$$
\frac{u_{G_i}^k}{\|x^k - x^*\|} = \begin{cases}
0, & i \in \{1, 2, \ldots, m\} \setminus I(x^*), \\
0, & i \in J_1^k(x^*), \\
-\nabla g_i(x^*)^T d_k + \dfrac{o(x^k - x^*)}{\|x^k - x^*\|}, & i \in J_2^k(x^*), \\
-\nabla g_i(x^*)^T d_k + \dfrac{o(x^k - x^*)}{\|x^k - x^*\|}, & i \in J(x^*).
\end{cases}
\tag{3.22}
$$

So for any $i \in J(x^*) \cup J_2^k(x^*)$, taking limits in (3.19) with k approaches to $+\infty$, we must have

$$d^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) d \leq 0, \tag{3.23}$$

$$\nabla h_j(x^*)^T d = 0, \quad j = 1, \ldots, l, \tag{3.24}$$

$$\nabla g_i(x^*)^T d = 0, \quad i \in J(x^*) \cup J_2^k(x^*). \tag{3.25}$$

For $i \in I(x^*) \setminus \{J(x^*) \cup J_2^k(x^*)\}$, there is an infinite index set $N_0 \subseteq N$ such that $g_i(x^k) \leq 0$ for all $k \in N_0$. So

$$\nabla g_i(x^*)^T d = \lim_{k \to \infty} \nabla g_i\left(\xi_i^k\right) d_k = \lim_{k \to \infty} \frac{g_i(x^k) - g_i(x^*)}{\|x^k - x^*\|} \leq 0, \tag{3.26}$$

where ξ_i^k lies in the line segment between x^* and x^k. Putting (3.24)–(3.26) together implies that $d \in M(x^*)$. We get $d^T \nabla_{xx}^2 L(x^*, \lambda^*, \mu^*) d > 0$ by Definition 2.2, which is a contradiction with (3.23). So the right inequality of (2.2) holds. The proof is complete. □

The converse of Theorem 3.3 is given below.

Theorem 3.4. *If (x^*, λ^*, μ^*) is a local saddle point of $L_r(x, \lambda, \mu)$ for some $r > 0$, then x^* is a local optimal solution to the problem (P).*

Proof. Let (x^*, λ^*, μ^*) be a local saddle point of $L_r(x, \lambda, \mu)$ for some $r > 0$. Then

$$L_r(x^*, \lambda, \mu) \leq L_r(x^*, \lambda^*, \mu^*) \leq L_r(x, \lambda^*, \mu^*) \tag{3.27}$$

whenever $(x, \lambda, \mu) \in (X \cap N(x^*, \delta)) \times \mathbb{R}_+^m \times \mathbb{R}^l$. We first show that x^* is a feasible solution. If not, there must exist $g_{i_0}(x^*) > 0$ for some i_0 or $h_{j_0}(x^*) \neq 0$ for some j_0.

Case 1. There exists $g_{i_0}(x^*) > 0$ for some i_0. Note that

$$L_r(x^*, \lambda, \mu) = f(x^*) + \sum_{j=1}^{l} \mu_j h_j(x^*) + \frac{r}{2} \sum_{j=1}^{l} h_j^2(x^*) + P_r(-G(x^*), \lambda). \tag{3.28}$$

Choose $\lambda_{i_0} > 0$ and $\lambda_i = 0$ for all $i \neq i_0$. Then we get from (A1) that

$$\begin{aligned} P_r(-G(x^*), \lambda) &= \inf_{u + G(x^*) \leq 0} \left\{ -u^T \lambda + r\sigma(u) \right\} \\ &\geq \inf_{u + G(x^*) \leq 0} \left\{ -u^T \lambda \right\} \\ &\geq g_{i_0}(x^*) \lambda_{i_0}. \end{aligned} \tag{3.29}$$

Taking the limit as $\lambda_{i_0} \to +\infty$ yields $P_r(-G(x^*), \lambda) \to +\infty$, which is a contradiction with (2.2). So we have $g_i(x^*) \leq 0$ for all $i = 1, \ldots, m$.

Case 2. There exists $h_{j_0}(x^*) \neq 0$ for some j_0. Choose μ_{j_0} and $h_{j_0}(x^*)$ with the same signal and let μ_{j_0} approach to $+\infty$, which is a contradiction with (2.2). So we have $h_j(x^*) = 0$ for all $j \in \{1, \ldots, l\}$. Then x^* is a feasible solution as claimed.

Since x^* is feasible, we have $h_j(x^*) = 0$ and for $u \in \{u \in \mathbb{R}^m \mid u + G(x^*) \leq 0\}$. In particular for $u = 0$, we have

$$P_r(-G(x^*), 0) = \inf_{u+G(x^*)\leq 0} \left\{-u^T 0 + r\sigma(u)\right\} = r\sigma(0) = 0. \tag{3.30}$$

Substituting (3.30) into (2.2) yields

$$P_r(-G(x^*), \lambda^*) \geq 0. \tag{3.31}$$

On the other hand, for any feasible x and $\lambda \geq 0$,

$$P_r(-G(x^*), \lambda) = \inf_{u+G(x^*)\leq 0} \left\{-u^T \lambda + r\sigma(u)\right\} \leq 0^T \lambda + r\sigma(0) = 0, \tag{3.32}$$

which, together with (3.31) and (3.32), implies that

$$P_r(-G(x^*), \lambda^*) = 0. \tag{3.33}$$

So

$$L_r(x^*, \lambda^*, \mu^*) = f(x^*). \tag{3.34}$$

Finally, for any feasible $x \in X \cap N(x^*, \delta)$, we have

$$\begin{aligned} f(x^*) &= L_r(x^*, \lambda^*, \mu^*) \leq L_r(x, \lambda^*, \mu^*) \\ &= f(x) + \sum_{j=1}^{l} \mu_j^* h_j(x) + \frac{r}{2} \sum_{j=1}^{l} h_j^2(x) + P_r(-G(x), \lambda^*) \\ &\leq f(x), \end{aligned} \tag{3.35}$$

which means that x^* is a local optimal solution of (P). □

Example 3.5. Consider the nonconvex programming problem

$$\begin{aligned}
\min \quad & f(x) = e^{x_1^2 - x_2^2}, \\
\text{s.t.} \quad & g_1(x) = 1 - x_2^2 \le 0, \\
& g_2(x) = e^{-x_1} - 1 \le 0, \\
& h(x) = x_1^{-2} - x_2^2 + 1 = 0, \\
& x \in X = R^2.
\end{aligned} \tag{3.36}$$

The optimal solutions of the above problem are $x^{*,1} = (0,1)^T$ and $x^{*,2} = (0,-1)^T$ with objective value e^{-1}. Setting $\lambda^* = (e^{-1}, 0)$ and $\mu^* = 0$, then we get from KKT conditions that

$$\begin{aligned}
\nabla_x L\left(x^{*,i}, \lambda^*, \mu^*\right) &= \nabla f\left(x^{*,i}\right) + \sum_{i=1}^{2} \lambda_i^* \nabla g_i\left(x^{*,i}\right) + \mu^* \nabla h\left(x^{*,i}\right) \\
&= \begin{pmatrix} 0 \\ -2x_2^{*,i} e^{-1} \end{pmatrix} + e^{-1} \begin{pmatrix} 0 \\ 2x_2^{*,i} \end{pmatrix} \\
&= 0.
\end{aligned} \tag{3.37}$$

The Hessian matrix

$$\begin{aligned}
\nabla_{xx}^2 L\left(x^{*,i}, \lambda^*, \mu^*\right) &= \nabla^2 f\left(x^{*,i}\right) + \sum_{i=1}^{2} \lambda_i^* \nabla^2 g_i\left(x^{*,i}\right) + \mu^* \nabla^2 h\left(x^{*,i}\right) \\
&= \begin{pmatrix} 2e^{-1} & 0 \\ 0 & 2e^{-1} \end{pmatrix} + e^{-1} \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \\
&= \begin{pmatrix} 2e^{-1} & 0 \\ 0 & 4e^{-1} \end{pmatrix}
\end{aligned} \tag{3.38}$$

is positive definite. The weak second-order sufficient conditions are satisfied at $(x^{*,i}, \lambda^*, \mu^*)$, $i = 1, 2$. By Theorem 3.4, $(x^{*,i}, \lambda^*, \mu^*)$ $(i = 1, 2)$ is a local saddle point for $L_r(x, \lambda, \mu)$, and hence $x^{*,1}$ and $x^{*,2}$ are the optimal solutions to (P).

Based on the above discussion, we know that, if (x^*, λ^*, μ^*) is a saddle point of L_r, then

$$\operatorname{val}(P) = L_r(x^*, \lambda^*, \mu^*) = \sup_{\lambda \ge 0, \mu \in \mathbb{R}^l} \inf_{x \in \mathbb{R}^n} L_r(x, \lambda, \mu), \tag{3.39}$$

where we denote by $\operatorname{val}(P)$ the optimal value of problem (P) and by (P') the problem given in the right-hand side. Note that the problem (P') has just the nonnegative constraints $\lambda \ge 0$ (rather simple constraints). Hence, we successful convert the original nonconvex problem to a simple constrained optimization problems by using the augmented Lagrangian function. Furthermore, many efficient algorithms for solving unconstrained optimization problems can be used to solve (P'), such as gradient-type algorithms. Therefore, our results, obtained with weaker conditions, provide a new insight and theoretical foundation for the use of augmented Lagrangian functions in constrained optimization problems.

Acknowledgments

The authors would like to give their sincere thanks to the anonymous referees for their helpful suggestions and valuable comments which improved the presentation of this paper. This research was supported by the National Natural Science Foundation of China (10971118, 10701047, 10901096).

References

[1] M. R. Hestenes, "Multiplier and gradient methods," *Journal of Optimization Theory and Applications*, vol. 4, pp. 303–320, 1969.
[2] M. J. D. Powell, "A method for nonlinear constraints in minimization problems," in *Optimization*, R. Fletcher, Ed., pp. 283–298, Academic Press, London, UK, 1969.
[3] R. T. Rockafellar, "Augmented Lagrange multiplier functions and duality in nonconvex programming," *SIAM Journal on Control and Optimization*, vol. 12, pp. 268–285, 1974.
[4] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, vol. 317 of *Grundlehren der Mathematischen Wissenschaften*, Springer, Berlin, Germany, 1998.
[5] X. X. Huang and X. Q. Yang, "A unified augmented Lagrangian approach to duality and exact penalization," *Mathematics of Operations Research*, vol. 28, no. 3, pp. 533–552, 2003.
[6] Q. Liu, W. M. Tang, and X. M. Yang, "Properties of saddle points for generalized augmented Lagrangian," *Mathematical Methods of Operations Research*, vol. 69, no. 1, pp. 111–124, 2009.
[7] C. Wang, J. Zhou, and X. Xu, "Saddle points theory of two classes of augmented Lagrangians and its applications to generalized semi-infinite programming," *Applied Mathematics and Optimization*, vol. 59, no. 3, pp. 413–434, 2009.
[8] A. Ben-tal and M. Zibulevsky, "Penalty/barrier multiplier methods for convex programming problems," *SIAM Journal on Optimization*, vol. 7, no. 2, pp. 347–366, 1997.
[9] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Computer Science and Applied Mathematics, Academic Press, New York, NY, USA, 1982.
[10] X. X. Huang and X. Q. Yang, "Approximate optimal solutions and nonlinear Lagrangian functions," *Journal of Global Optimization*, vol. 21, no. 1, pp. 51–65, 2001.
[11] X. X. Huang and X. Q. Yang, "Further study on augmented Lagrangian duality theory," *Journal of Global Optimization*, vol. 31, no. 2, pp. 193–210, 2005.
[12] E. Polak and J. O. Royset, "On the use of augmented Lagrangians in the solution of generalized semi-infinite min-max problems," *Computational Optimization and Applications*, vol. 31, no. 2, pp. 173–192, 2005.
[13] R. Polyak, "Modified barrier functions: theory and methods," *Mathematical Programming*, vol. 54, no. 2, pp. 177–222, 1992.
[14] A. Rubinov and X. Yang, *Lagrange-Type Functions in Constrained Non-Convex Optimization*, vol. 85 of *Applied Optimization*, Kluwer Academic Publishers, Boston, Mass, USA, 2003.
[15] P. Tseng and D. P. Bertsekas, "On the convergence of the exponential multiplier method for convex programming," *Mathematical Programming*, vol. 60, no. 1, pp. 1–19, 1993.
[16] Y. Y. Zhou and X. Q. Yang, "Augmented Lagrangian function, non-quadratic growth condition and exact penalization," *Operations Research Letters*, vol. 34, no. 2, pp. 127–134, 2006.
[17] D. Li, "Zero duality gap for a class of nonconvex optimization problems," *Journal of Optimization Theory and Applications*, vol. 85, no. 2, pp. 309–324, 1995.
[18] D. Li, "Saddle point generation in nonlinear nonconvex optimization," *Nonlinear Analysis: Theory, Methods and Applications*, vol. 30, no. 7, pp. 4339–4344, 1997.
[19] D. Li and X. L. Sun, "Existence of a saddle point in nonconvex constrained optimization," *Journal of Global Optimization*, vol. 21, no. 1, pp. 39–50, 2001.

[20] D. Li and X. L. Sun, "Convexification and existence of a saddle point in a pth-power reformulation for nonconvex constrained optimization," *Nonlinear Analysis: Theory, Methods and Applications*, vol. 47, no. 8, pp. 5611–5622, 2001.

[21] X. L. Sun, D. Li, and K. I. M. McKinnon, "On saddle points of augmented Lagrangians for constrained nonconvex optimization," *SIAM Journal on Optimization*, vol. 15, no. 4, pp. 1128–1146, 2005.