A note on the existence of saddle points of p-th power Lagrangian for constrained nonconvex optimization

H.X. Wu & H.Z. Luo

a Department of Mathematics, College of Science, Hangzhou Dianzi University, Hangzhou, Zhejiang 310018, China
b Department of Applied Mathematics, College of Science, Zhejiang University of Technology, Hangzhou, Zhejiang 310032, China

A note on the existence of saddle points of $p$-th power Lagrangian for constrained nonconvex optimization

H.X. Wu$^a$ and H.Z. Luo$^b$*

$^a$Department of Mathematics, College of Science, Hangzhou Dianzi University, Hangzhou, Zhejiang 310018, China; $^b$Department of Applied Mathematics, College of Science, Zhejiang University of Technology, Hangzhou, Zhejiang 310032, China

(Received 12 December 2009; final version received 14 February 2011)

Li and Sun [D. Li and X.L. Sun, Existence of a saddle point in nonconvex constrained optimization, J. Global Optim. 21 (2001), pp. 39–50; 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], present the existence of a global saddle point of the $p$-th power Lagrangian functions for constrained nonconvex optimization, under second-order sufficiency conditions and additional conditions that the feasible set is compact and the global solution of the primal problem is unique. In this article, it is shown that the same results can be obtained under additional assumptions that do not require the compactness of the feasible set and the uniqueness of global solution of the primal problem.

Keywords: constrained nonconvex optimization; $p$-th power Lagrangian functions; saddle points; compactness and uniqueness

AMS Subject Classifications: 90C26; 90C46

1. Introduction

Consider the following inequality constrained nonconvex optimization problem:

$$(P) \quad \min f(x)$$

$$\text{s.t. } g_i(x) \leq b_i, \quad i = 1, 2, \ldots, m,$$

$$x \in X,$$

where $f, g_i : \mathbb{R}^n \rightarrow \mathbb{R}, \quad i = 1, \ldots, m$ are twice continuously differentiable functions and $X$ is a nonempty and closed subset in $\mathbb{R}^n$. We can replace the constraint $g_i(x) \leq b_i$ by $\exp[g_i(x)] \leq \exp[b_i]$, if necessary. So, we assume throughout this article that $g_i(x) \geq 0$ for all $x \in X$ and $b_i > 0, \quad i = 1, \ldots, m$.

*Corresponding author. Email: hzluo@zjut.edu.cn

ISSN 0233–1934 print/ISSN 1029–4945 online
© 2012 Taylor & Francis
http://dx.doi.org/10.1080/02331934.2011.564620
http://www.tandfonline.com
Li [3] introduced the following $p$-th power Lagrangian to eliminate the duality gap in $(P)$:

$$L_1(x, \lambda, p) = [f(x)]^p + \sum_{i=1}^{m} \lambda_i([g_i(x)]^p - b_i^p), \quad x \in X, \quad \lambda \in \mathbb{R}_+^m,$$

(1)

where $p > 0$ is a penalty parameter and $f(x)$ is assumed to be positive on $X$. The function $L_1$ is the classical Lagrangian of the following reformulation of $(P)$:

$$\min [f(x)]^p \quad \text{s.t. } [g_i(x)]^p \leq b_i^p, \quad i = 1, \ldots, m, \quad x \in X.$$  

(2)

It was proved in [3] that under certain conditions, the perturbation function of (2) is locally convex in a neighbourhood of $b^p = (b_1^p, \ldots, b_m^p)^T$ when $p$ is large enough. A local saddle point is thus guaranteed to exist for the $p$-th power formulation (2). By applying the $p$-th power transformation only to the constraints, Xu [12] proposed the following partial $p$-th power reformulation of $(P)$:

$$\min f(x) \quad \text{s.t. } [g_i(x)]^p \leq b_i^p, \quad i = 1, \ldots, m, \quad x \in X.$$  

(3)

The associated Lagrangian of (3) is

$$L_2(x, \lambda, p) = f(x) + \sum_{i=1}^{m} \lambda_i([g_i(x)]^p - b_i^p), \quad x \in X, \quad \lambda \in \mathbb{R}_+^m.$$  

(4)

Note that $L_2(x, \lambda, p)$ reduces to the exponential-type augmented Lagrangian [1] when setting $g_i(x) = \exp(c_i(x))$ and $b_i = 1$ for $i = 1, \ldots, m$. It was shown in [12] that local saddle points and convexification of $L_1$ and $L_2$ can be obtained under certain weaker conditions. In [4], under the same weaker conditions, it was proved that the Hessian of $L_1$ or of $L_2$ becomes positive definite in a neighbourhood of a local optimal point of the original problem. The existence of global saddle points of $L_1$ and $L_2$ was investigated in [5,6] for nonconvex constrained optimization problems. Based on these results, problem $(P)$ can be solved by using the primal-dual method [1] for the equivalent problem. However, the existence theorems in [5,6] require two restrictive conditions: $X$ is compact and the global solution of $(P)$ is unique. These two conditions were also assumed in [1,8,9,11] for the existence of global saddle points for other types of augmented Lagrangian functions.

Recently, by means of the Image Space Analysis [2], local and global saddle point conditions for a general augmented Lagrangian function proposed by Rockafellar and Wets [10] were investigated in [7]. It was shown that the existence of a saddle point is equivalent to a non-linear separation of two suitable subsets of the Image Space associated with the given problem. Under second order sufficiency conditions in the Image Space [2], it was proved that the augmented Lagrangian admits a local saddle point. The existence of a global saddle point was then obtained under additional assumptions that do not require the compactness of the feasible set. We point out that $L_1(x, \lambda, p)$ or $L_2(x, \lambda, p)$ is not a subclass of the augmented Lagrangian of Rockafellar and Wets [10].
The purpose of this article is to study the existence of global saddle points of the $p$-th power Lagrangian functions $L_j(x, \lambda, p)$ $(j = 1, 2)$. We prove that under the second-order sufficiency conditions and additional mild conditions, the global saddle point of $L_j(x, \lambda, p)$ $(j = 1, 2)$ exists without requiring the compactness condition of $X$ and the uniqueness condition of global solution of $(P)$. Our results can be regarded as generalizations of the existence theorems of a saddle point for the $p$-th power Lagrangian functions [5,6] and for the exponential-type Lagrangian functions [11].

2. Main results

Let $x^*$ be a global solution to $(P)$. We consider the following assumption about $x^*$ from [12].

**Assumption 2.1** (Second-order sufficiency conditions)

(a) There exists $\lambda^* \in \mathbb{R}^m$ such that

$$
\begin{bmatrix}
\nabla f(x^*) + \sum_{i=1}^{m} \lambda^*_i \nabla g_i(x^*)
\end{bmatrix}^T d \geq 0 \quad \forall d \in T_X(x^*),
$$

$$
\lambda^*_i [g_i(x^*) - b_i] = 0, \quad i = 1, \ldots, m,
$$

where $T_X(x^*)$ denotes the tangent cone of $X$ at $x^*$.

(b) The Hessian matrix of Lagrangian of $(P)$:

$$
H(x^*) = \nabla^2 f(x^*) + \sum_{i \in J(x^*)} \lambda^*_i \nabla^2 g_i(x^*)
$$

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

$$
J(x^*) = \{i \mid \lambda^*_i > 0, \ i = 1, \ldots, m\},
$$

$$
N(x^*) = \{d \in \mathbb{R}^n \mid \nabla f(x^*)^T d = 0, \ \nabla g_i(x^*)^T d = 0, \ i \in J(x^*)\},
$$

$$
M(x^*) = N(x^*) \cap T(x^*).
$$

(c) The set $X$ is locally convex around $x^*$.

We remark that, if $x^*$ is an interior point of $X$, then (5) is equivalent to

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

and, moreover,

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

For any $p > 0$, define $\mu^*_p = (\mu^*_1, \ldots, \mu^*_n) \in \mathbb{R}^m_+$ by

$$
\mu^*_p = \begin{cases}
[f(x^*)]^{p-1} (\lambda^*_i / b_i^{p-1}), & i \in J(x^*), \\
0, & \text{otherwise},
\end{cases} \quad (\text{for } L_1),
$$

$$
\mu^*_j = \begin{cases}
\sum_{i \in J(x^*)} \mu^*_i / [f(x^*)]^{p-1} (\lambda^*_i / b_i^{p-1}), & j \in J(x^*), \\
0, & \text{otherwise},
\end{cases} \quad (\text{for } L_2).
$$
\[ \mu^*_p = \begin{cases} \frac{\lambda^*_i}{(p b^*_i)^{p-1}}, & i \in J(x^*), \\ 0, & \text{otherwise}, \end{cases} \quad (\text{for } L_2). \] (9)

The following theorem [12, Theorems 2.1 and 2.3] shows that \((x^*, \mu^*_p)\) is a local saddle point of \(L_j (j=1, 2)\) when \(p\) is sufficiently large.

**Theorem 2.2** Suppose that \(x^* \in X\) is a global solution to \((P)\) satisfying Assumption 2.1. For \(L_1\), suppose also that \(f(x)\) is positive over \(X\). Then there exist \(q_0 > 0\) and \(\delta > 0\) such that

\[ L_j (x^*, \mu, p) \leq L_j (x^*, \mu^*_p, p) \leq L_j (x, \mu^*_p, p), \quad (j = 1, 2) \]

for all \(x \in X \cap N(x^*, \delta)\) and \(\mu \in \mathbb{R}_+^m\) when \(p \geq q_0\), where \(N(x^*, \delta) = \{x \in \mathbb{R}^n | \|x - x^*\| \leq \delta\}\).

Let \(J^-(x^*) = \{1, \ldots, m\}\setminus J(x^*)\) and \(\epsilon \geq 0\). Define

\[
\begin{align*}
F(\epsilon) &= \{x \in X \mid g_i(x) \leq b_i + \epsilon, \ i = 1, \ldots, m\}, \\
F_1(\epsilon) &= \{x \in X \mid g_i(x) \leq b_i + \epsilon, \ i \in J(x^*)\}, \\
F_2 &= \{x \in X \mid g_i(x) \leq b_i, \ i \in J^-(x^*)\}, \\
U(\epsilon) &= \{x \in X \mid f(x) \leq f(x^*) + \epsilon\}.
\end{align*}
\]

The following is the existence theorem for global saddle points of \(L_j (j=1, 2)\) [6, Theorem 2.2] [5, Theorem 4].

**Theorem 2.3** Let \(X\) in \((P)\) be compact and \(x^* \in X\) be a global optimal solution of \((P)\) at which Assumption 2.1 is satisfied. For \(L_1\), suppose that \(f(x)\) is positive over \(X\) and the following conditions hold:

(i) \(x^*\) is the unique global solution of \((P)\);

(ii) If \(J^-(x^*) \neq \emptyset\), then there exist positive \(\epsilon_1\) and \(\epsilon_2\) such that

\[ f(x) > f(x^*) + \epsilon_1 \quad \forall x \in F_1(\epsilon_2) \setminus F_2. \] (10)

For \(L_2\), suppose that \(F_1(0) \cap U(0) = \{x^*\}\). Then there exists \(q_1 > 0\) such that for all \(p \geq q_1\), \((x^*, \mu^*_p)\) is a global saddle point of \(L_j (x, \mu^*_p, p) (j = 1, 2)\).

We note that in Theorem 2.3, the conditions (i), (ii) can be implied by the condition \(F_1(0) \cap U(0) = \{x^*\}\) [6, Proposition 2.1]. It was pointed out in [6] that the condition (ii) is indispensable to ensure a saddle point to be generated by the \(p\)-th power Lagrangian formulation.

In the following, we will mainly generalize Theorem 2.3 to the case where \(X\) is not necessarily a compact set and the global solution of \((P)\) is not unique.

First, we consider the \(p\)-th power Lagrangian \(L_1\). We need the following assumptions:

**Assumption 2.4** \(f = \inf_{x \in X} f(x) > 0\).

**Assumption 2.5** There exists \(\epsilon_0 > 0\) such that \(F(\epsilon_0) \cap U(\epsilon_0)\) is bounded.

Obviously, Assumption 2.4 is satisfied if \(f(x)\) is positive over \(X\) and \(X\) is a compact set. Also, Assumption 2.5 is satisfied if \(X\) is a compact set.
This contradiction indicates that Claim 2.1 holds true.

Then there exists $q_1 > 0$ such that for all $p \geq q_1$, $(x^\star, \mu^\star_p)$ is a global saddle point of $L_1$

$$L_1(x^\star, \mu, p) \leq L_1(x^\star, \mu^\star_p, p) \leq L_1(x, \mu^\star_p, p)$$ \hspace{1cm} (11)

for all $x \in X$ and $\mu \in \mathbb{R}^m_+$, where $\mu^\star_p$ is defined by (8).

**Proof** It is easy to see that the first inequality in (11) always holds by noting the complementarity and feasibility of $x^\star$. We now prove that there exists $q_1$ such that the second inequality in (11) holds for all $x \in X$ when $p \geq q_1$. By contradiction, suppose that there exist sequences $\{p_k\}$ with $p_k \to \infty$ and $\{x^k\} \subseteq X$ with $x^k \neq x^\star$ for all $k$ such that

$$L_1(x^k, \mu^\star_{p_k}, p_k) < L_1(x^\star, \mu^\star_{p_k}, p_k) = [f(x^\star)]^{p_k},$$ \hspace{1cm} (12)

where the equality follows from $g_i(x^\star) = b_i$ for $i \in J(x^\star)$. Thus, from (8) and the definition of $L_1$, we can rewrite (12) as

$$\sum_{i \in J(x^\star)} \lambda^\star_i \left[ g_i(x^k)/b_i \right]^{p_k-1} g_i(x^k) - b_i < \lambda^\star_i b_i \leq \left(1 + \tilde{\epsilon}/b_i \right)^{p_k-1} (b_i + \tilde{\epsilon}) - \sum_{i \in J(x^\star)} \lambda^\star_i b_i$$

$$\to \infty, \quad (k \to \infty).$$

This contradiction indicates that Claim 2.1 holds true.

**Claim 2.1** \(\forall \epsilon > 0, \exists k_\epsilon > 0\) such that $x^k \in F_1(\epsilon)$ for all $k \geq k_\epsilon$.

Suppose, on the contrary, that there exist $\tilde{\epsilon} > 0$ and an infinite subsequence $\mathcal{K}_1 \subset \{1, 2, \ldots \}$ such that $x^k \in F_1(\tilde{\epsilon})$ for all $k \in \mathcal{K}_1$. Then there exist $i_0 \in J(x^\star)$ such that $g_{i_0}(x^k) > b_{i_0} + \tilde{\epsilon}$ for all $k \in \mathcal{K}_1$. Note that $g_i(x) \geq 0$ for all $x \in X$ and $b_i > 0$, $i = 1, \ldots, m$. Also, by Assumption 2.4, $f(x) > 0$ for all $x \in X$. Therefore, we have from (13) that

$$f(x^\star) > \lambda^\star_{i_0} \left[ g_{i_0}(x^k)/b_{i_0} \right]^{p_k-1} g_{i_0}(x^k) - \sum_{i \in J(x^\star)} \lambda^\star_i b_i$$

$$> \lambda^\star_{i_0} (1 + \tilde{\epsilon}/b_{i_0})^{p_k-1} (b_{i_0} + \tilde{\epsilon}) - \sum_{i \in J(x^\star)} \lambda^\star_i b_i$$

$$\to \infty, \quad (k \to \infty).$$

This contradiction indicates that Claim 2.1 holds true.

**Claim 2.2** \(\forall \epsilon > 0, \exists k_\epsilon > 0\) such that $x^k \in U(\epsilon)$ for all $k \geq k_\epsilon$.

By contradiction, suppose that $\exists \tilde{\epsilon} > 0$ and $\exists \mathcal{K}_2 \subset \{1, 2, \ldots \}$ such that $x^k \notin U(\tilde{\epsilon})$ for all $k \in \mathcal{K}_2$. Then

$$f(x^k) > f(x^\star) + \tilde{\epsilon}, \quad \text{for all } k \in \mathcal{K}_2.$$

By Assumption 2.4, $f = \inf_{x \in X} f(x) > 0$. Thus, from (13) and by $p_k \to \infty$, we have

$$- \sum_{i \in J(x^\star)} \lambda^\star_i b_i < \lambda^\star_i b_i \leq \left(1 + \tilde{\epsilon}/f(x^\star) \right)^{p_k-1} f$$

$$\to -\infty, \quad (k \to \infty),$$

which gives a contradiction.
Claim 2.3 $x^k \in \bar{F}_2$ for sufficiently large $k$, where $\bar{F}_2$ is defined by

$$
\bar{F}_2 = \begin{cases} 
F_2, & \text{if } J^- (x^*) \neq \emptyset, \\
X, & \text{if } J^- (x^*) = \emptyset.
\end{cases}
$$

Clearly, this claim holds true if $J^- (x^*) = \emptyset$. Now suppose that $J^- (x^*) \neq \emptyset$. By contradiction, assume that $\exists k_3 \in \{1, 2, \ldots\}$ such that $x^k \notin \bar{F}_2$ for all $k \in k_3$. By Claim 2.1 and (14), $\exists k_{e_2} > 0$ such that $x^k \in F(\epsilon_2) \setminus F_2$ for all $k \in k_3$ and $k \geq k_{e_2}$. Hence, from (10), we have

$$
f(x^k) > f(x^*) + \epsilon_1 \quad \text{for all } k \in k_3 \text{ and } k \geq k_{e_2}.
$$

Using similar arguments as in the proof of Claim 2.2, we can derive a contradiction.

Next, from Claims 2.1–2.3, we have that

$$
x^k \in F_1(\epsilon) \cap U(\epsilon) \cap \bar{F}_2 \quad \forall \epsilon \in (0, \epsilon_0]
$$

(15)

for $k$ large enough. Note that $F_1(\epsilon) \cap \bar{F}_2 \subseteq F(\epsilon)$ for $\epsilon > 0$. Hence, $x^k \in F(\epsilon_0) \cap U(\epsilon_0)$ when $k$ is large enough. By Assumption 2.5, the sequence $\{x^k\}$ is bounded and hence it admits at least a limit point. Without loss of generality, we assume that

$$
\lim_{k \to \infty} x^k = \hat{x}.
$$

Since $f$ and each $g_i$ are continuous and the set $X$ is closed, we obtain from (15) that

$$
\hat{x} \in F_1(\epsilon) \cap U(\epsilon) \cap \bar{F}_2 \quad \forall \epsilon \in (0, \epsilon_0].
$$

Since $\epsilon > 0$ is arbitrary, we have that $\hat{x} \in F_1(0) \cap U(0) \cap \bar{F}_2$. Note that $F(0) = F_1(0) \cap \bar{F}_2$. Thus, $\hat{x} \in F(0) \cap U(0)$ and so $\hat{x}$ is a global optimal solution to $(P)$.

Therefore, by the condition of this theorem, Assumption 2.1 is satisfied at $(\hat{x}, \lambda^*)$. Hence, by Theorem 2.2, there exist $p_{\hat{x}} > 0$ and $\delta_{\hat{x}} > 0$ such that

$$
L_1(x, \mu_{\hat{x}}^*, p) \geq L_1(\hat{x}, \mu_{p_{\hat{x}}}^*, p) = [f(\hat{x})]^p = [f(x^*)]^p
$$

(17)

for all $x \in N(\hat{x}, \delta_{\hat{x}}) \cap X$ when $p \geq p_{\hat{x}}$, where $\mu_{p_{\hat{x}}}^*$ is defined by (8). Since $p_k \to \infty$, we have $p_k \geq p_{\hat{x}}$ for $k$ sufficiently large. Therefore, inequality (17) leads to

$$
[f(x^*)]^p \leq L_1(x, \mu_{p_k}^*, p_k) \quad \forall x \in N(\hat{x}, \delta_{\hat{x}}) \cap X,
$$

when $k$ is sufficiently large. This together with (12) gives rise to

$$
x^k \in X \setminus N(\hat{x}, \delta_{\hat{x}})
$$

for $k$ sufficiently large, which contradicts (16).

Next, we discuss the existence of a global saddle point of $L_2$.

Assumption 2.7 $\ell = \inf_{x \in X} f(x) > -\infty$.

Similar to the proof of Theorem 2.6, we have the following result for $L_2$.

Theorem 2.8 Suppose that the conditions of Theorem 2.6 are satisfied, where Assumption 2.4 is replaced by Assumption 2.7. Then there exists $q_1 > 0$ such that for all
\( p \geq q_1, (x^*, \mu^*_p) \) is a global saddle point of \( L_2 \)

\[
L_2(x^*, \mu, p) \leq L_2(x^*, \mu^*_p, p) \leq L_2(x, \mu^*_p, p)
\]  

(18)

for all \( x \in X \) and \( \mu \in \mathbb{R}^m_+ \), where \( \mu^*_p \) is defined by (9).

**Proof** We only prove that there exists \( q_1 \) such that the second inequality in (18) holds for all \( x \in X \) when \( p \geq q_1 \). By contradiction, suppose that there exist sequences \( \{p_k\} \) with \( p_k \to \infty \) and \( \{x^k\} \subseteq X \) with \( x^k \neq x^* \) for all \( k \), such that

\[
L_2(x^k, \mu^*_p, p_k) < L_2(x^*, \mu^*_p, p_k) = f(x^*).
\]  

(19)

From (9) and the definition of \( L_2 \), (19) can be rewritten as

\[
\sum_{i \in J(x^*)} \lambda^*_i b_i \frac{\left( \left( g_i(x^k) \right)^{p_k} \right)}{b_i} - 1 < f(x^*) - f(x^k).
\]  

(20)

**Claim 2.4** \( \forall \epsilon > 0, \exists k^\epsilon > 0 \) such that \( x^k \in F_1(\epsilon) \) for all \( k \geq k^\epsilon \).

Suppose, by contradiction, that there exist \( \tilde{\epsilon} > 0 \) and an infinite subsequence \( K_1 \subset \{1, 2, \ldots\} \) such that \( x^k \notin F_1(\tilde{\epsilon}) \) for all \( k \in K_1 \). Then there exist \( i_0 \in J(x^*) \) such that \( g_i(x^k) > b_{i_0} + \tilde{\epsilon} \) for all \( k \in K_1 \). Note that each \( g_i \) is nonnegative over \( X \) and \( b_j > 0 \) for all \( i \). Also, by Assumption 2.7, \( \underline{f} = \inf_{x \in X} f(x) > -\infty \). Therefore, we have from (20) that

\[
f(x^*) - f > \sum_{i \in J(x^*)} \lambda^*_i b_i \frac{\left( \left( g_i(x^k) \right)^{p_k} \right)}{b_i} - 1
\]

\[
\geq \frac{\lambda^*_i b_i}{p_k} \left[ \left( 1 + \frac{\tilde{\epsilon}}{b_{i_0}} \right)^{p_k} - 1 \right] - \sum_{i \in J(x^*)} \lambda^*_i b_i \frac{\left( \left( g_i(x^k) \right)^{p_k} \right)}{b_i}
\]

\[
\geq \frac{\lambda^*_i b_i}{p_k} \left[ 1 + p_k \ln \left( 1 + \frac{\tilde{\epsilon}}{b_{i_0}} \right) + \frac{1}{2} p_k^2 \ln^2 \left( 1 + \frac{\tilde{\epsilon}}{b_{i_0}} \right) - 1 \right] - \sum_{i \in J(x^*)} \lambda^*_i b_i \frac{\left( \left( g_i(x^k) \right)^{p_k} \right)}{b_i}
\]

\[
= \frac{\lambda^*_i b_i}{p_k} \ln \left( 1 + \frac{\tilde{\epsilon}}{b_{i_0}} \right) \left[ 1 + \frac{\tilde{\epsilon} p_k \ln \left( 1 + \frac{\tilde{\epsilon}}{b_{i_0}} \right)}{b_{i_0}} \right] - \sum_{i \in J(x^*)} \lambda^*_i b_i \frac{\left( \left( g_i(x^k) \right)^{p_k} \right)}{b_i}
\]

\[
\to \infty \quad (k \to \infty),
\]

which is a contradiction.

**Claim 2.5** \( \forall \epsilon > 0, \exists k^\epsilon > 0 \) such that \( x^k \in U(\epsilon) \) for all \( k \geq k^\epsilon \).

Suppose that \( \exists \tilde{\epsilon} > 0 \) and \( \exists K_2 \subset \{1, 2, \ldots\} \) such that \( x^k \notin U(\tilde{\epsilon}) \) for all \( k \in K_2 \). Then

\[
f(x^k) > f(x^*) + \tilde{\epsilon} \quad \text{for all } k \in K.
\]

By (20), we have that

\[
-\tilde{\epsilon} > - \sum_{i \in J(x^*)} \lambda^*_i b_i \frac{\left( \left( g_i(x^k) \right)^{p_k} \right)}{b_i}.
\]  

(21)

Taking the limit in (21) and using \( p_k \to \infty \) gives rise to \( \tilde{\epsilon} \leq 0 \), contradicting \( \tilde{\epsilon} > 0 \).
Using a proof similar to that of Claim 2.3 in the proof of Theorem 2.6, we have the following claim.

**Claim 2.6** \( x^k \in \tilde{F}_2 \) for sufficiently large \( k \), where \( \tilde{F}_2 \) is defined by (14).

The rest of the proof is similar to that of Theorem 2.6, and we can deduce a contradiction.

We end this section with applying the existence of saddle points for the partial \( p \)-th power Lagrangian to the box and linear inequality constrained indefinite quadratic programming problems of the form

\[
\text{(CQP)} \quad \min f(x) = \frac{1}{2} x^T Q x + q^T x \quad \text{s.t. } g_i(x) = \alpha_i^T x \leq b_i, \quad i = 1, 2, \ldots, m, \quad x \in X = [l, u]^n,
\]

where \( Q \) is a symmetric \( n \times n \) indefinite matrix, \( q \in \mathbb{R}^m \), \( \alpha_i \in \mathbb{R}^n \) with \( \alpha_i \neq 0 \), \( b_i > 0 \), \( i = 1, \ldots, m \), \( l < u \in \mathbb{R}^n \), and it is assumed that \( \beta_i = \min_{x \in \mathbb{R}^n} \alpha_i^T x \geq 0 \) for \( i = 1, \ldots, m \).

Let \( x^* \in \text{int} X \) be a feasible solution of (CQP). Suppose that there exists \( \lambda^* \in \mathbb{R}_+^m \) such that

\[
Q x^* + q + \sum_{i \in J(x^*)} \lambda_i^* \alpha_i = 0, \quad (22)
\]

\[
d^T Q d > 0 \quad \forall d \in M(x^*), \quad d \neq 0, \quad (23)
\]

where \( J(x^*) = \{ i : \lambda_i^* > 0, i = 1, \ldots, m \} \) and \( M(x^*) = \{ d \in \mathbb{R}^n : \alpha_i^T d = 0, \ i \in J(x^*) \} \).

Now, consider the partial \( p \)-th power Lagrangian defined in (4) for problem (QCP), i.e.

\[
\hat{L}(x, \mu, p) = \frac{1}{2} x^T Q x + q^T x + \sum_{i=1}^m \mu_i [(\alpha_i^T x)^p - b_i^p]. \quad (24)
\]

By Theorem 2.2 above or Corollary 2.1 in [12] or Theorem 2.1 in [4], we can obtain the following corollary, but the proof here is different from that of Theorem 2.2 in [12] and of Theorem 2.1 in [4].

**Corollary 2.9** Let \( x^* \in \text{int} X \) be a feasible solution of (CQP) and suppose that there exists \( \lambda^* \in \mathbb{R}_+^m \) such that (22) and (23) holds. Then there exists \( \hat{q}_0 > 0 \) such that for all \( p \geq \hat{q}_0 \), \( (x^*, \mu^*_p) \) is a local saddle point of \( \hat{L}(x, \mu, p) \), where \( \mu^*_p \) is defined by (9), and \( \hat{q}_0 \) is given by (29) below.

**Proof** It suffices to show that there exists \( \hat{q}_0 > 0 \) such that the Hessian matrix \( \nabla_{xx}^2 \hat{L}(x^*, \mu^*_p, p) \) of \( \hat{L}(x, \mu, p) \) at \( x^* \) is positive definite on \( \mathbb{R}^n \) when \( p \geq \hat{q}_0 \).

From (24) and (9), the direct computation gives rise to

\[
\nabla_{xx}^2 \hat{L}(x^*, \mu^*_p, p) = Q + (p - 1) \sum_{j \in J(x^*)} \mu_j^* \alpha_j \alpha_j^T, \quad (25)
\]

where \( \mu_j^* = \lambda_j^* / b_j, \ j \in J(x^*) \). Let

\[
B_n = \{ d \in \mathbb{R}^n : \| d \| = 1 \}, \quad K = \{ d \in B_n : d^T Q d \leq 0 \}. \quad (26)
\]
We consider two cases.

**Case 1** \( K = \emptyset \). Obviously, \( d^T Q d > 0 \) for all \( d \in B_n \). It then follows from (25) that
\[
d^T \nabla^2_{xx} \hat{L}(x^*, \mu^*_p, p) d = d^T Q d + (p - 1) \sum_{j \in J(x^*)} \mu^*_j (\alpha^T d)^2 > 0 \quad \forall d \in B_n,
\]
which implies that \( \nabla^2_{xx} \hat{L}(x, \mu^*_p, p) \) is positive definite on \( \mathbb{R}^n \) when \( p \geq 1 \).

**Case 2** \( K \neq \emptyset \). Denote
\[
\eta = \min_{d \in K} d^T Q d,
\]
\[
\tau = \min_{d \in K} \sum_{j \in J(x^*)} \mu^*_j (\alpha^T d)^2.
\]
Clearly, \( \eta \leq 0 \) and \( \tau \geq 0 \). Now we claim that \( \tau > 0 \). Otherwise, if \( \tau = 0 \), then there exists \( d \in K \) such that \( \sum_{j \in J(x^*)} \mu^*_j (\alpha^T d)^2 = 0 \), which, by \( \mu^*_j > 0 \) for \( j \in J(x^*) \), implies that \( \alpha^T d = 0 \), which, by \( d \in M(x^*) \), thus, \( d \in M(x^*) \). Also, \( \|d\| = 1 \) due to \( d \in K \). Therefore, from condition (23), we have that \( \hat{d}^T Q \hat{d} > 0 \), contradicting to \( d \in K \).

Let
\[
\hat{q}_0 = -\frac{\eta}{\tau} + 1.
\]
It is clear that \( \hat{q}_0 \geq 1 \) since \( \eta \leq 0 \) and \( \tau > 0 \). If \( d \in K \), then it follows from (25) and (27)–(29) that when \( p > \hat{q}_0 \),
\[
d^T \nabla^2_{xx} \hat{L}(x^*, \mu^*_p, p) d = d^T Q d + (p - 1) \sum_{j \in J(x^*)} \mu^*_j (\alpha^T d)^2
\]
\[
\geq \eta + (p - 1) \tau > \eta + \left(-\frac{\eta}{\tau}\right) \cdot \tau = 0.
\]
If \( d \in B_n \setminus K \), then \( d^T Q d > 0 \), this means that when \( p \geq 1 \),
\[
d^T \nabla^2_{xx} \hat{L}(x^*, \mu^*_p, p) d > 0 \quad \forall d \in B_n \setminus K.
\]
Combined with (30), it leads to the positive definiteness of \( \nabla^2_{xx} \hat{L}(x^*, \mu^*_p, p) \) on \( \mathbb{R}^n \).

Since Assumptions 2.5 and 2.7 can be implied by the compactness of \( X \), the following corollary can be directly obtained from Theorem 2.8.

**Corollary 2.10** Suppose that there exists \( \lambda^* \in \mathbb{R}^m_+ \) such that for any global solution \( \hat{x} \) of \( (CQP) \), (22) and (23) is satisfied at \( (\hat{x}, \lambda^*) \). Let \( x^* \in X \) be a global solution of \( (CQP) \) and condition (ii) in Theorem 2.3 holds. Then there exists \( \hat{q}_1 > 0 \) such that for all \( p \geq \hat{q}_1 \), \( (x^*, \mu^*_p) \) is a global saddle point of \( \hat{L}(x, \mu, p) \), where \( \mu^*_p \) is defined by (9).

It is pointed out that in the literature most of the papers on the existence of global saddle points for constrained nonconvex optimization need a sufficiently large penalty parameter and do not give a lower bound of it \([1,5–9,11]\). So, we take the risk of saying that it would be difficult to determine a lower bound of \( \hat{q}_1 \) in Corollary 2.10.
3. Illustrated examples

In this section, we give two examples to illustrate the existence results of global saddle points derived in Section 2, in which the global solution of problem is not unique and \( X = \mathbb{R}^n \).

**Example 3.1** Consider the following constrained nonconvex optimization:

\[
\begin{align*}
\min & \quad f(x) = (x_2 - 1)^2 + 8e^{-\frac{1}{4}(x_1^4 - 1)} + 1 \\
\text{s.t.} & \quad g_1(x) = x_1^2 + (x_2 - 1)^2 + 1 \leq 2, \\
& \quad x \in X = \mathbb{R}^2.
\end{align*}
\]  

(31)

It is easy to check that \( f(x) \) is a nonconvex function on \( \mathbb{R}^2 \). This problem has two global optimal solutions: \( x^{*,1} = (-1, 1)^T \) and \( x^{*,2} = (1, 1)^T \), with the optimal value \( f^* = 9 \). The optimal Lagrange multipliers corresponding to these two global solutions are both \( \lambda^* = 2 \). The feasible region of this problem is shown in Figure 1. Since

\[
\nabla f(x^{*,j}) = ((-1)^{j+1}4, 0)^T, \quad \nabla g_1(x^{*,j}) = ((-1)^j2, 0)^T,
\]

we have \( \nabla_x L_1(x^{*,j}, \lambda^*, 1) = 0 \), and

\[
\nabla^2_{xx} L_1(x^{*,j}, \lambda^*, 1) = \text{diag}(-6, 6), \quad j = 1, 2.
\]

It is thus clear that there does not exist a saddle point for the conventional Lagrangian function \( L_1(x, \lambda, 1) \). Figure 2 illustrates the picture of the classical Lagrangian \( L_1(x, \lambda^*, 1) \). On the other hand, since problem (31) has two global solutions and \( X = \mathbb{R}^2 \), the condition in Theorem 2.3 is not satisfied.

Note that \( M(x^{*,j}) = \{ d \in \mathbb{R}^2 : d_1 = 0 \} \) (j = 1, 2). So, for any \( d \in M(x^{*,j}) \) with \( d \neq 0 \), it holds that \( d^T \nabla^2_{xx} L_1(x^{*,j}, \lambda^*) d = 6d_2^2 > 0 \) for \( j = 1, 2 \). Hence the second-order sufficient conditions hold at \( (x^{*,1}, \lambda^*) \) and \( (x^{*,2}, \lambda^*) \), respectively. Moreover, we see that \( F(\epsilon_0) \) is bounded and hence \( F(\epsilon_0) \cap U(\epsilon_0) \) is bounded for some \( \epsilon_0 > 0 \). Note that

![Figure 1](image-url)
$\ell = \inf_{x \in X} f(x) > 0$ and $g_1(x)$ is positive on $X$. Also, $J^{-}(x^*) = \emptyset$. Thus the conditions in Theorems 2.6 and 2.8 are fulfilled.

Consider the $p$-th power Lagrangian $L_1$ defined in (1) for this example. Let $\mu_p^*$ be defined by (8), then $\mu_p^* = 2\left(\frac{2}{q}\right)^{p-1}$. Direct calculation gives the Hessian of the $p$-th power Lagrangian,

$$\nabla_{xx}^2 L_1(x^{*j}, \mu_p^*, p) = p^{9p-1} \begin{pmatrix} \frac{52}{9}p - \frac{106}{9} & 0 \\ 0 & 6 \end{pmatrix}, \quad j = 1, 2,$$

which is a positive definite matrix when $p > \frac{53}{26}$.

Consider again the partial $p$-th power Lagrangian $L_2$ defined in (4) for this problem. From (9), we have $\mu_p^* = \frac{2}{p^{2p-1}}$ and a direct calculation gives the Hessian of the partial $p$-th power Lagrangian,

$$\nabla_{xx}^2 L_2(x^{*j}, \mu_p^*, p) = \begin{pmatrix} 4p - 10 & 0 \\ 0 & 6 \end{pmatrix}, \quad j = 1, 2,$$

which is a positive definite matrix when $p > 2.5$.

Now, we take $p = 3$ for $L_1$ and $p = 4$ for $L_2$, then $\mu_p^* = 40.5$ for $L_1$ and $\mu_p^* = \frac{1}{10}$ for $L_2$. It can be also verified that $x^{*j}$ solves problem $\min_{x \in X} L_i(x, \mu_p^*, p)$ globally and therefore $(x^{*j}, \mu_p^*)$ is a global saddle point of $L_i, (i, j = 1, 2)$. Figures 3 and 4 illustrate the pictures of $L_i(x, \mu_p^*, p), i = 1, 2$, respectively.

**Example 3.2** Consider the following constrained nonconvex problem

$$\min f(x) = -x_1^2 - x_2^2$$

s.t. $g_1(x) = 2x_1^2 + 3x_2^2 + 2x_1x_2 \leq 1,$

$g_2(x) = 3x_1^2 + 2x_2^2 - 4x_1x_2 \leq 1,$

$g_3(x) = x_1^2 + 6x_2^2 - 4x_1x_2 \leq 1,$

$x \in X = \mathbb{R}^2.$

---

Figure 2. The classical Lagrangian $L_1(x, \lambda^*, 1)$ of Example 3.1.
This problem has two global solutions: \( x^{*1} = (0.645, 0.1047)^T \) and \( x^{*2} = (-0.6450, -0.1047)^T \), with the optimal value \( f^* = -0.4270 \). The associated vector of Lagrange multipliers is both \( \lambda^* = (0.2776, 0.1495, 0)^T \). The feasible region of this example is shown in Figure 5. Since

\[
\nabla f(x^{*j}) = ((-1)^j 1.29, (-1)^j 0.2094)^T, \\
\nabla g_1(x^{*j}) = ((-1)^j 1.27894, (-1)^j 1.9182)^T, \\
\nabla g_2(x^{*j}) = ((-1)^j 1.34512, (-1)^j 2.1612)^T, \\
\nabla g_3(x^{*j}) = ((-1)^j 1.08712, (-1)^j 1.3236)^T, \\
\]

(33)
we have \( \nabla^2 L_2(x^{*,j}, \lambda^*, 1) = 0 \), and
\[
\nabla^2 L_2(x^{*,j}, \lambda^*, 1) = \begin{pmatrix}
0.0069 & -0.0427 \\
-0.0427 & 0.2631
\end{pmatrix} \triangleq A \ (j = 1, 2),
\]
which is not a positive definite matrix as \( \det(A) = -7.9 \times 10^{-6} < 0 \). Thus there does not exist a saddle point for the conventional Lagrangian function \( L_2(x, \lambda, 1) \). On the other hand, since problem (32) has two global solutions and \( X = \mathbb{R}^2 \), the condition in Theorem 2.3 is not satisfied.

From (9), we have that \( \inf_{x \in \mathbb{R}^2} g_i(x) \geq 0, i = 1, 2, 3, \) and we can see from Figure 5 that there exist positive \( \epsilon_1 \) and \( \epsilon_2 \) such that inequality (10) holds at \( x^{*,1} \) and \( x^{*,2} \), respectively. Thus conditions in Theorem 2.8 are fulfilled except for \( f^* = \inf_{x \in X} f(x) = -\infty \).

Consider the partial \( p \)-th power Lagrangian \( L_2 \) defined in (4) for this problem. From (9), we have that \( \mu_p^* = \frac{x^{*,j}}{p} \) and a direct calculation gives the Hessian of the partial \( p \)-th power Lagrangian,
\[
\nabla^2_{xx} L_2(x^{*,j}, \mu_p^*, p) = \begin{pmatrix}
0.0069 + 3.9397(p - 1) & -0.042 + 0.3703(p - 1) \\
-0.042 + 0.3703(p - 1) & 0.2631 + 1.7193(p - 1)
\end{pmatrix} \triangleq A(p),
\]
where \( j = 1, 2 \). Note that \( \det(A(p)) = 6.6364(p - 1)^2 + 1.08(p - 1) - 7.9 \times 10^{-6} > 0 \) when \( p > 1 + 7.314 \times 10^{-6} \). So, the matrix \( A(p) \) is positive definite when \( p > 1 + 7.314 \times 10^{-6} \). Now, we take \( p = 3 \). Then \( \mu_p^* = \frac{x^{*,j}}{p} = (0.0925, 0.0498, 0)^T \). It can be also verified that with \( p = 3 \), \( x^{*,j} \) solves problem \( \min_{x \in X} L_2(x, \mu_p^*, p) \) globally and therefore \( (x^{*,j}, \mu_p^*) \) is a global saddle point of \( L_2 (j = 1, 2) \). Figure 6 illustrates the picture of \( L_2(x, \mu_p^*, p) \) with \( p = 3 \).

### 4. Conclusions

In this article, we have presented the properties of the global saddle point of \( p \)-th power Lagrangian functions for constrained nonconvex optimization. We have
showed the existence of global saddle points of $p$-th power Lagrangian functions under the second-order sufficiency conditions and additional mild conditions that do not require the compactness condition of $X$ and the uniqueness of global solution of $(P)$, but with the additional assumption that any global solution of $(P)$ is associated with the same Lagrange multiplier.

One of the future research topics is to further investigate the existence of global saddle points under the assumption that any global solution of $(P)$ is associated with the different Lagrange multiplier.

Acknowledgements
The authors are grateful to the anonymous referees for their valuable comments and constructive suggestions for improving this article. This work was partially supported by the National Natural Science Foundation of China under grant No. 11071219, the Zhejiang Provincial Natural Science Foundation of China under grant Nos. Y6090080 and Y7080184 and the Postdoctoral Key Research Foundation of China under grant No. 201003242.

References


