J Glob Optim (2007) 38:405–419
DOI 10.1007/s10898-006-9090-4

Local saddle point and a class of convexification methods for nonconvex optimization problems

Tao Li · Yanjun Wang · Zhian Liang · Panos M. Pardalos

Received: 7 March 2006 / Accepted: 28 August 2006 / Published online: 7 October 2006

Abstract A class of general transformation methods are proposed to convert a nonconvex optimization problem to another equivalent problem. It is shown that under certain assumptions the existence of a local saddle point or local convexity of the Lagrangian function of the equivalent problem (EP) can be guaranteed. Numerical experiments are given to demonstrate the main results geometrically.

Keywords Nonconvex optimization · Local saddle point · Convexification · Local convexity

1 Introduction

We consider a constrained nonconvex optimization problem with the following form:

$$(P): \begin{cases} \min f_0(x) \\ \text{s.t. } f_i(x) \leq b_i,\ i = 1, \ldots, m, \\ x \in X, \end{cases}$$

where f_i: $R^n \to R$, $i = 0, \ldots, m$, are twice continuously differentiable functions and X is a nonempty subset of R^n. The Lagrangian function corresponding to Problem (P) is:

$$L(x, \lambda) = f_0(x) + \sum_{i=1}^{m} \lambda_i [f_i(x) - b_i], \quad \lambda = (\lambda_1, \ldots, \lambda_m) \geq 0.$$

T. Li (✉) · Y. Wang · Z. Liang
Department of Applied Mathematics, Shanghai University of Finance & Economics, Shanghai 200433, P.R. China
e-mail: taoli1981@yahoo.com

P. M. Pardalos
Department of Industrial and Systems Engineering, University of Gainesville, Gainesville, FL 32611, USA

It is well known that Lagrangian function plays an important role in many optimization problems such as the development of duality theory which is the foundation of primal-dual methods. So many researchers have concentrated on exploiting the properties of Lagrangian function. In our paper, the following contents about two aspects are studied. One is about the local saddle point of $L(x, \lambda)$; the other is about the local convexity of $L(x, \lambda)$. As for the first aspect, we know that the condition that the pair (x^*, λ^*) becomes a local saddle point of $L(x, \lambda)$ is the inequality

$$L(x^*, \lambda) \leq L(x^*, \lambda^*) \leq L(x, \lambda^*) \quad \text{for all } x \in N_{x^*} \bigcap X \quad \lambda \geq 0, \tag{1.1}$$

holds where N_{x^*} is a neighborhood of x^*. However, (1.1) does not always hold for many nonconvex problems, thus the existence of a local saddle point is not guaranteed (see Example 4.1) on many occasions. For the second aspect, local duality theory for problem(P) has been developed under a basic local convexity assumption on $L(x, \lambda)$. As shown in [1, 2], a crucial condition to apply the local duality theorem at x^* is that the Hessian matrix of Lagrangian function $\nabla^2 L(x^*, \lambda^*)$ should be positive definite. But lots of nonconvex problems do not satisfy the requirement. For example the Lagrangian function of linearly constrained concave programing and indefinite quadratic programing do not satisfy the local convexity requirement.

Recently, some transformation schemes have been adopted to transform Problem (P) into another equivalent problem whose Lagrangian function has the properties desired. For example, in [3], a special convexification method was proposed for Problem (P) by using p-power transformation to derive an equivalent form, and a local saddle point result under some reasonable assumptions is obtained. Subsequently, another convexification method was presented in [4] by using partial p-power transformation to Problem (P), i.e., applying the p-power only to the constraint functions, and under assumptions that are weaker than those in [3], the same result holds. In addition, the same p-power and partial p-power formulations were adopted to derive the equivalent problems in [5], and the paper proved that under certain assumptions the Hessian matrix of the Lagrangian function of the equivalent problem (EP) becomes positive definite in a neighborhood of a local optimal point of the primal problem. However, the restrictive conditions under which the mentioned transformations can be successfully done limit the range the problems they can tackle.

The main purpose of this paper is to present a class of general transformation methods including the methods in [3–5] as special cases, then for the EP derived, under certain assumptions, we can obtain a local saddle point result and prove that the Lagrangian function of the equivalent problem is locally convex. This paper not only generalizes the results obtained in [3–5], but also expands considerably the class of nonconvex problems for which some important optimization theories such as local dual search methods can be guaranteed.

This paper is organized as follows: in Sect. 2, we propose a class of convexification methods for Problem (P) to derive an EP. Then in Sects. 3 and 4 a local saddle point result is stated, and some sufficient conditions that guarantee the local convexity of the Lagrangian function are presented. Thereafter some relevant corollaries follow. Furthermore, numerical experiments are presented to interpret the main results of this paper in Sect. 5. Finally in Sect. 6 the conclusion is given.

2 A class of general transformation methods

Consider the following transformation of the objective function and constraint functions of Problem (P):

$$f_{ip}(x) = \psi_{ip}(f_i(x)), \quad i = 0, \ldots, m,$$

where ψ_{ip}: $Z_{ip} \to R$, in which $Z_{ip} \subseteq R$ and $f_i(X) \subseteq Z_{ip}$, $i = 0, 1, \ldots, m$, and $p > 0$ is a parameter.

We further assume that $\psi_{ip}(y)$, $i = 0, 1, \ldots, m$, have the following properties:

(1) $\psi_{ip}(y) \in C^2(R)$, $i = 0, 1, \ldots, m$.
(2) $\psi_{ip}(y)$, $i = 0, 1, \ldots, m$, are strictly monotonously increasing functions, which means $\psi'_{ip}(y) > 0$.
(3) There exists a N_{x^*} such that, for $N > 0$, there is a p_0 satisfying $\dfrac{\psi''_{0p}}{\psi'_{0p}} > -N$ when $p > p_0$, for $x \in N_{x^*} \bigcap X$, where ψ''_{0p} and ψ'_{0p} denote $\dfrac{\mathrm{d}^2\psi_{0p}(y)}{\mathrm{d}y^2}|_{y=f_0(x)}$ and $\dfrac{\mathrm{d}\psi_{0p}(y)}{\mathrm{d}y}|_{y=f_0(x)}$, respectively, for simplicity.
(4) For $x \in N_{x^*} \bigcap X$, $\lim_{p\to+\infty} \dfrac{\psi''_{ip}}{\psi'_{ip}} = +\infty$, $i = 1, \ldots, m$, where ψ''_{ip} and ψ'_{ip} represent $\dfrac{\mathrm{d}^2\psi_{ip}(y)}{\mathrm{d}y^2}|_{y=f_i(x)}$ and $\dfrac{\mathrm{d}\psi_{ip}(y)}{\mathrm{d}y}|_{y=f_i(x)}$, respectively, for simplicity.

In the remarks of Sect. 4, examples of admissible functions are given.

Then Problem (P) could be transformed into an EP which reads:

$$(\text{EP}): \begin{cases} \min \psi_{0p}(f_0(x)), \\ \text{s.t. } \psi_{ip}(f_i(x)) \le \psi_{ip}(b_i), \quad i = 1, 2, \ldots, m, \\ x \in X. \end{cases}$$

The Lagrangian function associated with Problem (EP) is:

$$L_p(x, \mu) = \psi_{0p}(f_0(x)) + \sum_{i=1}^{m} \mu_i[\psi_{ip}(f_i(x)) - \psi_{ip}(b_i)], \tag{2.1}$$

where $p > 0$ and $\mu_i \ge 0$, $i = 1, \ldots, m$. And the Hessian matrix of $L_p(x, \mu)$ could be written as:

$$\nabla^2 L_p(x, \mu) = \psi''_{0p} \nabla f_0(x) \nabla f_0(x)^T + \psi'_{0p} \left(\sum_{i \in J(x)} \frac{\psi''_{ip}}{\psi'_{ip}} \lambda_i \nabla f_i(x) \nabla f_i(x)^T + \nabla^2 L(x, \lambda) \right). \tag{2.2}$$

3 Local saddle point result

Definition 3.1 *Let S be a nonempty set in R^n and $x^* \in clS$. $T(x^*)$ is called the tangent cone of S at x^* when it is the set of all directions d such that*

$$d = \lim_{k\to+\infty} \lambda_k(x_k - x^*),$$

where $\lambda_k > 0$, $x_k \in S$, *for each* k *and* $\lim_{k\to+\infty} x_k = x^*$. *More properties of tangent cone are given in [6].*

Definition 3.2 *A nonempty set* X *is called locally convex near* x^*, *if there is a neighborhood* N_{x^*} *of* x^* *such that* $X \bigcap N_{x^*}$ *is convex.*

Suppose that there exist $x^* \in X$ *and* $\lambda^* \in R^m_+$ *such that*

$$\left[\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*)\right]^T d \geq 0, \quad \forall d \in T(x^*), \tag{3.1a}$$

$$\sum_{i=1}^{m} \lambda_i^* [f_i(x^*) - b_i] = 0, \tag{3.1b}$$

where $T(x^*)$ *is the tangent cone of* X *at* x^*.

Let

$$J(x^*) = \{i : \lambda_i^* > 0,\ i = 1, \ldots, m\},$$
$$N(x^*) = \{d \in R^n :\ \nabla f_0(x^*)^T d = 0 \text{ and } \nabla f_i(x^*)^T d = 0,\ i \in J(x^*)\},$$
$$\widehat{M}(x^*) = T(x^*) \bigcap N(x^*).$$

Then we have the following theorem:

Theorem 3.1 *Suppose that* x^* *is feasible to Problem* (*P*) *and* $J(x^*) \neq \emptyset$, X *is locally convex near* x^*, *there exists a* $\lambda^* \in R^m_+$ *satisfying* (3.1) *and the Hessian matrix of Lagrangian function of Problem (P) associated with the* λ^*

$$\nabla^2 L(x^*, \lambda^*) = \nabla^2 f_0(x^*) + \sum_{i \in J(x^*)} \lambda_i^* \nabla^2 f_i(x^*) \tag{3.2}$$

is positive definite on $\widehat{M}(x^*)$. *Then there is a* p_1 *such that the Lagrangian function of EP defined in* (2.1) *has a local saddle point* (x^*, μ_p^*) *when* $p \geq p_1$ *where* $\mu_p^* \in R^m_+$.

Proof Since Problem (P) and (EP) are equivalent and x^* is feasible to problem(P), then by (2), we have

$$\psi_{ip}(f_i(x^*)) \leq \psi_{ip}(b_i), \quad i = 1, \ldots, m. \tag{3.3}$$

For any $p > 0$, define $\mu_p^* \in R^m_+$ as

$$\mu_{ip}^* = \begin{cases} \dfrac{\psi_{0p}^{'} \lambda_i^*}{\psi_{ip}^{'}}, & i \in J(x^*), \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

From (3.1b) and (3.4), we get

$$\mu_{ip}^* (\psi_{ip}(f_i(x^*)) - \psi_{ip}(b_i)) = 0, \quad i = 1, \ldots, m. \tag{3.5}$$

By (2.1), (3.3), and (3.5), we obtain

$$\begin{aligned} L_p(x^*, \mu_p^*) &= \psi_{0p}(f_0(x)) \geq L_p(x^*, \mu) \\ &= \psi_{0p}(f_0(x)) + \sum_{i=1}^{m} \mu_i [\psi_{ip}(f_i(x)) - \psi_{ip}(b_i)], \end{aligned} \tag{3.6}$$

where $\mu_i \geq 0,\ i = 1, \ldots, m$. If $X \bigcap N_{x^*} = \{x^*\}$, by (3.6) the result already holds in this case.

In the following, we will show that the result still holds when $X \bigcap N_{x^*} \neq \{x^*\}$.

By (2.1) and (3.4), we obtain

$$\nabla L_p(x^*, \mu_p^*) = \psi'_{0p} \nabla f_0(x^*) + \sum_{i \in J(x^*)} \mu_{ip}^* \psi'_{ip} \nabla f_i(x^*). \tag{3.7}$$

Using (3.1a), (3.4), and (3.7), we get

$$\nabla L_p(x^*, \mu_p^*)^T d \geq 0, \quad d \in T(x^*). \tag{3.8}$$

Now we prove that there exists a $p_1 > 0$ such that when $p > p_1$ we can find some $\delta_p > 0$ satisfying

$$L_p(x^*, \mu_p^*) < L_p(x, \mu_p^*),\ x \in X \bigcap O(x^*, \delta_p) \subseteq X \bigcap N_{x^*} \quad \text{and } x \neq x^*, \tag{3.9}$$

where $O(x^*, \delta_p) \equiv \{x \in R^n, \|x - x^*\| < \delta_p\}$ and $\mu_p^* \in R_+^m$ defined in (3.4).

Suppose contrary to the result, i.e. there is a sequence $\{x_p\} \subset X$ with $x_p \to x^*$ such that

$$L_p(x_p, \mu_p^*) \leq L_p(x^*, \mu_p^*).$$

Then it follows that

$$\begin{aligned} 0 &\geq L_p(x_p, \mu_p^*) - L_p(x^*, \mu_p^*) \\ &= \nabla L_p(x^*, \mu_p^*)^T (x_p - x^*) + \frac{1}{2}(x_p - x^*)^T \nabla^2 L_p(\xi_p, \mu_p^*)(x_p - x^*), \end{aligned} \tag{3.10}$$

where $\xi_p \in O(x^*, \|x_p - x^*\|)$ and $\xi_p \to x^*$, as $p \to +\infty$.

Dividing both sides of (3.10) by $\|x_p - x^*\|^2$ and let

$$d_p = (x_p - x^*)/\|x_p - x^*\|, \tag{3.11}$$

we obtain

$$0 \geq \nabla L_p(x^*, \mu_p^*)^T d_p/\|x_p - x^*\| + \frac{1}{2} d_p^T \nabla^2 L_p(\xi_p, \mu_p^*) d_p\ . \tag{3.12}$$

Observe that $d_p \in S_n$, where S_n denotes the unit sphere in R^n. By the compactness of S_n , we let $d_p \to \overline{d}$ as $p \to +\infty$. Thus combining (3.11) with the definition of $T(x^*)$, we have

$$\overline{d} \in T(x^*) \quad \text{and} \quad \overline{d} \neq 0. \tag{3.13}$$

Since X is locally convex near x^*, thus we get

$$d_p \in T(x^*), \quad \text{for } p \text{ large enough.} \tag{3.14}$$

By (2.2) and (3.7), we rewrite (3.12) as

$$\begin{aligned} 0 \geq{}& \left[\nabla f_0(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) \right]^T d_p/\|x_p - x^*\| + \frac{1}{2} \frac{\psi''_{0p}}{\psi'_{0p}} (\nabla f_0(\xi_p)^T d_p)^2 \\ &+ \frac{1}{2} d_p^T\, \nabla^2\, L(\xi_p, \lambda^*) d_p + \frac{1}{2} \sum_{i \in J(x^*)} \lambda_i^* \frac{\psi''_{ip}}{\psi'_{ip}} (\nabla f_i(\xi_p)^T d_p)^2. \end{aligned}$$

By (3), for $N > 0$, there exists a p_0' such that when $p > p_0'$ we have

$$0 \geq \left[\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*)\right]^T d_p/\|x_p - x^*\| - \frac{1}{2}N(\nabla f_0(\xi_p)^T d_p)^2$$
$$+ \frac{1}{2} d_p^T \nabla^2 L(\xi_p, \lambda^*) d_p + \frac{1}{2} \sum_{i \in J(x^*)} \lambda_i^* \frac{\psi_{ip}''}{\psi_{ip}'} (\nabla f_i(\xi_p)^T d_p)^2. \tag{3.15}$$

Clearly $\nabla f_0(\xi_p)^T d_p$ and $d_p^T \nabla^2 L(\xi_p, \lambda^*) d_p$ have finite limit as $p \to +\infty$, and by (4), (3.1a) and (3.14) for p large enough, $[\nabla f_0(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*)]^T d_p/\|x_p - x^*\|$ and $\lambda_i^* \frac{\psi_{ip}''}{\psi_{ip}'} (\nabla f_i(\xi_p)^T d_p)^2$, $i \in J(x^*)$ are nonnegative.

In order to complete the proof, we consider the following two cases.

Case 1 $\quad \overline{d} \,\overline{\in}\, N(x_*)$.

In this case, if there exists some index $i_0 \in J(x^*)$ satisfying

$$(\nabla f_{i_0}(x^*)^T \overline{d})^2 > 0, \tag{3.16}$$

then we have

$$\lim_{p \to +\infty} (\nabla f_{i_0}(\xi_p)^T d_p)^2 = (\nabla f_{i_0}(x^*)^T \overline{d})^2 > 0. \tag{3.17}$$

Moreover, by (3.15) and (3.1a), we have

$$-N(\nabla f_0(\xi_p)^T d_p)^2 + d_p^T \nabla^2 L(\xi_p, \lambda^*) d_p \leq 0. \tag{3.18}$$

However, by (3.17) and (4), we get

$$\lim_{p \to +\infty} \lambda_{i_0}^* \frac{\psi_{i_0 p}''}{\psi_{i_0 p}'} (\nabla f_{i_0}(\xi_p)^T d_p)^2 = +\infty,$$

which means that by making p appropriately large, we can obtain a contradiction to (3.18).

If there is no $i_0 \in J(x^*)$ satisfying (3.16), then we must have

$$\nabla f_0(x^*)^T \overline{d} \neq 0, \ \nabla f_i(x^*)^T \overline{d} = 0, \quad i \in J(x^*). \tag{3.19}$$

By (3.1a), (3.13), and (3.19), we have

$$\left[\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*)\right]^T \overline{d} = \nabla f_0(x^*)^T \overline{d} > 0.$$

Thus we get

$$\lim_{p \to +\infty} \left[\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*)\right]^T d_p/\|x_p - x^*\| = +\infty. \tag{3.20}$$

Then, we can obtain a contradiction by making p appropriately large in (3.15). Therefore, (3.9) holds when $\overline{d} \,\overline{\in}\, N(x^*)$.

Case 2 $\overline{d} \in N(x^*)$.

In this case, by (3.13), we get

$$\bar{d} \in \widehat{M}(x^*). \tag{3.21}$$

Moreover, from (3.15), it is easy to get

$$d_p^T \bigtriangledown^2 L(\xi_p, \lambda^*) d_p - N(\bigtriangledown f_0(\xi_p)^T d_p)^2 + \sum_{i \in J(x^*)} \lambda_i^* \frac{\psi''_{ip}}{\psi'_{ip}} (\bigtriangledown f_i(\xi_p)^T d_p)^2 \le 0, \tag{3.22}$$

Then, let $p \to +\infty$ in (3.22), and we get

$$\bar{d}^T \nabla^2 L(x^*, \lambda^*) \bar{d} \le 0. \tag{3.23}$$

The combination of (3.21) and (3.23) contradicts the positive definiteness assumption of the Hessian matrix in (3.2). Thus (3.9) is true in this case.

Therefore, we prove (3.9) holds. By (3.6) and (3.9), we get the conclusion and complete the proof.

4 Convexification for nonconvex optimization problem

In this section, we assume that $x^* \in X$ and $\lambda^* \in R_+^m$ satisfy

$$\bigtriangledown f_0(x^*) + \sum_{i=1}^m \lambda_i^* \bigtriangledown f_i(x^*) = 0, \tag{4.1a}$$

$$\sum_{i=1}^m \lambda_i^* [f_i(x^*) - b_i] = 0 \tag{4.1b}$$

and let

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

Note that if x^* is an interior point of X, then (3.1) and (4.1) are equivalent and $\widehat{M}(x^*) = M(x^*)$.

From (4.1a) and (2.1), we obtain

$$\nabla L_p(x^*, \mu_p^*) = 0,$$

where μ_p^* is defined as (3.4).

Now we present some results about the local convexity of the Lagrangian function given in (2.1).

Theorem 4.1 *Suppose that x^* and λ^* satisfy* (4.1) *where $x^* \in X$ and $\lambda^* \in R_+^m$. Further we assume that $J(x^*) \neq \emptyset$ and the Hessian matrix in* (3.2) *is positive definite on $M(x^*)$, then there is a p_1 such that the Hessian matrix is positive definite when $p > p_1$, where μ_p^* is given by* (3.4).

Proof Suppose contrary to the result, i.e. there exists a sequence $\{d_p\}_{p=1}^{+\infty} \subseteq S_n$ such that

$$d_p^T \bigtriangledown^2 L_p(x^*, \mu^*) d_p \le 0, \tag{4.2}$$

where S_n denotes the unit sphere in R^n. By the compactness of S_n, there exists a convergent subsequence of $\{d_p\}_{p=1}^{+\infty} \subseteq S_n$. Without loss of generality, we assume that it is $\{d_p\}_{p=1}^{+\infty}$ itself and $\lim_{p\to+\infty} d_p = \overline{d}$.

Thus, from (2.2) and (4.2), we have

$$\psi'_{0p}\left\{\frac{\psi''_{0p}}{\psi'_{0p}}(\nabla f_0(x^*)^T d_p)^2 + d_p^T \nabla^2 L(x^*,\lambda^*)d_p + \Sigma_{i\in J(x^*)}\lambda_i^* \frac{\psi''_{ip}}{\psi'_{ip}}(\nabla f_i(x^*)^T d_p)^2\right\} \leq 0.$$

Then we can prove Theorem 4.1 by the similar way used in the proof of Theorem 3.1. Therefore, Theorem 4.1 holds.

Corollary 4.1 *Suppose that x^* and λ^* satisfy the conditions in Theorem* 4.1*, and x^* is feasible to Problem (P), then there is a p_1 such that x^* is a strictly local minimum solution of $L_p(x,\mu_p^*)$ when $p > p_1$, where μ_p^* is given by* (3.4).

Proof The result could be obtained directly from Theorem 4.1. □

Corollary 4.2 *Let x^* and λ^* satisfy the conditions in Corollary* 4.1*, then there exists a p_1 such that (x^*,μ_p^*) is a local saddle point of $L_p(x,\mu)$ when $p > p_1$, where μ_p^* is given by* (3.4).

Proof By Theorem 4.1 and Corollary 4.1, we get the conclusion.

Define the Lagrangian dual function for problem EP as

$$\max \theta_p(\mu) \\ \text{s.t. } \mu \geq 0,$$

where

$$\theta_p(\mu) = \min_{x\in N_{x^*}} L_p(x,\mu).$$

Then we get the following corollary.

Corollary 4.3 *Let x^* and λ^* satisfy the conditions in Corollary* 4.1*, then there exists a $p_1 > 0$ such that the Lagrangian dual problem has a local solution μ_p^* given by* (3.4) *when $p > p_1$ and further we have $L_p(x^*,\mu_p^*) = \theta_p(\mu_p^*)$.*

Proof By Corollary 4.1 and the definition of $\theta_p(\mu)$, in some neighborhood of x^*, we have

$$\theta_p(\mu_p^*) = \min_{x\in N_{x^*}} L_p(x,\mu) = L_p(x^*,\mu_p^*). \tag{4.3}$$

From (3.4), μ_p^* is a feasible solution of the Lagrangian dual function of Problem EP. Further, by Corollary 4.2, we get that (x^*,μ_p^*) is a local saddle point of $L_p(x,\mu)$. Thus we have

$$\begin{aligned} L_p(x^*,\mu_p^*) &= \psi_{0p}(f_0(x^*)) + \sum_{i=1}^m \mu_{ip}^*[\psi_{ip}(f_i(x^*)) - \psi_{ip}(b_i)] \\ &\geq L_p(x^*,\mu) = \psi_{0p}(f_0(x^*)) + \sum_{i=1}^m \mu_i[\psi_{ip}(f_i(x^*)) - \psi_{ip}(b_i)], \end{aligned} \tag{4.4}$$

holds locally, where $\mu_i \geq 0,\ i = 1,\ldots,m$. Then we have

$$\psi_{ip}(f_i(x^*)) \leq \psi_{ip}(b_i),\ i = 1, 2, \ldots, m \tag{4.5}$$

otherwise (4.4) would be violated by making the corresponding μ_i appropriately large. Thus, x^* is a feasible solution of Problem EP. By (3.4), (4.3), and (4.5) we get the conclusion and complete the proof. □

Remark 4.1 If we set $\psi_{ip}(y) = y^p$, $i = 0, \ldots, m$, it is easy to verify that $\psi_{ip}(y)$ satisfy properties (1)–(4). And the equivalent problem can be written as

$$\begin{aligned} &\min f_0^p(x),\\ &\text{s.t. } f_i^p(x) \leq b_i^p, \quad i = 1, \ldots, m,\\ &\quad x \in X. \end{aligned} \tag{4.6}$$

Note that (4.6) is exactly the transformation proposed in [3]. Moreover, different from the assumptions in [3], the regularity of x^* is not required in Theorems 3.1 and 4.1. Thus our results can be extended to the problems in which equality and inequality constraints are involved.

Remark 4.2 Let $\psi_{0p}(y) = y$ and $\psi_{ip}(y) = y^p$, $i = 1, \ldots, m$. Observe that properties (1)–(4) are also satisfied and Problem (P) could be transformed into

$$\begin{aligned} &\min f_0(x),\\ &\text{s.t. } f_i^p(x) \leq b_i^p, \quad i = 1, \ldots, m,\\ &\quad x \in X. \end{aligned} \tag{4.7}$$

Note that (4.7) is exactly the transformation proposed in [4]. Thus the main results obtained in [4] could be viewed as special cases of this paper.

Remark 4.3 Let $\psi_{0p}(y) = y$ and $\psi_{ip}(y) = e^{py}$, $i = 1, \ldots, m$. Clearly $\psi_{ip}(y)$, $i = 0, \ldots, m$, have properties (1)–(4), then Problem (P) could be transformed into the following equivalent form

$$\begin{aligned} &\min f_0(x),\\ &\text{s.t. } e^{pf_i(x)} \leq e^{pb_i}, \quad i = 1, \ldots, m,\\ &\quad x \in X \end{aligned} \tag{4.8}$$

(4.8) is the transformation proposed in [7]. And the result in [7] is a special case of Corollary 4.2.

Remark 4.4 We can derive other types of transformations which are different from those proposed in previous papers by constructing many specific function forms possessing properties (1)–(4). For example, each of functions $-e^{-\frac{y}{p}}$, e^{y^p} $(y > 0)$, $\ln^p(y+k)$ (k should be chosen to ensure $(y + k) > 1$), and $(1 + \frac{x}{p})^{p^k}$ $(k \geq 2)$ could be used as $\psi_{0p}(y)$ and the latter three could also be used as ψ_{ip}, $i = 1, \ldots, m$, where k is an integral number.

5 Numerical experiments

In this section, we give three examples to demonstrate Theorems 3.1 and 4.1 geometrically. The transformation we use in each example is different from those proposed before. More specifically, in Example 1, we use a specific transformation to guarantee

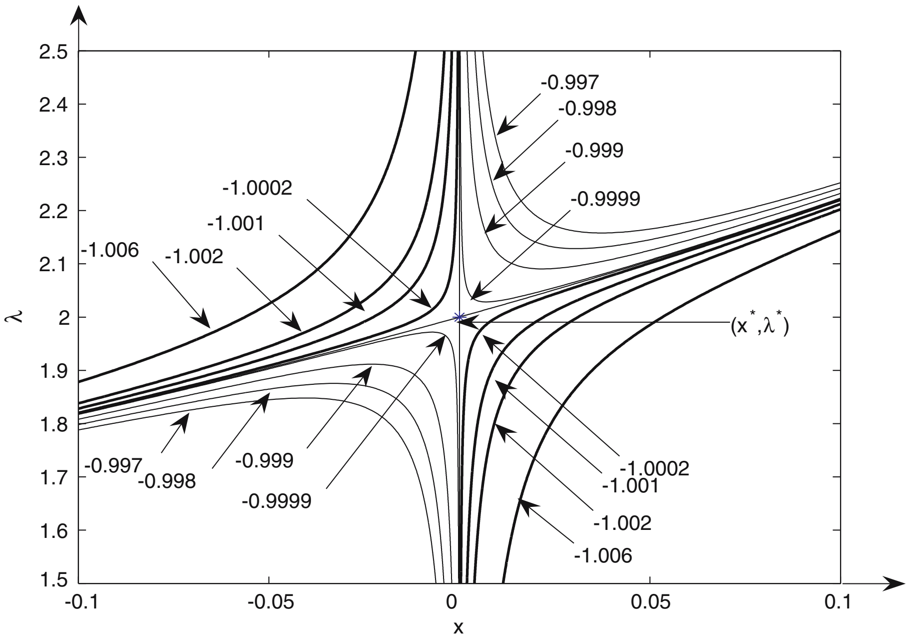

Fig. 1 Contours of the Lagrangian function $L(x, \lambda)$ for Problem (1P) near (x^*, λ^*)

the existence of a local saddle point. In Examples 2 and 3, convexification transformations are applied to ensure the local convexity of the Lagrangian function.

Example 5.1 Consider the following problem

$$(1P) : \begin{cases} \min f_0(x) = \dfrac{x+1}{x-1} \\ \text{s.t. } f_1(x) = x \leq 0, \\ x \in \ X = \{x \in R : x < 1\}. \end{cases}$$

The Lagrangian function of Problem (1P) is

$$L(x, \lambda) = \frac{x+1}{x-1} + \lambda x. \tag{5.1}$$

Figure 1 shows contours of $L(x, \lambda)$. It is easy to verify that $x^* = 0$ and $\lambda^* = 2$ satisfy the conditions in Theorem 3.1.

It is necessary to point out that (x^*, λ^*) is not a local saddle point of the Lagrangian function of Problem (1P). Suppose the contrary, then we have

$$L(0, 2) \leq L(x, 2), \ \ x \in N_{x^*} \bigcap X,$$

where N_{x^*} is a neighborhood of x^*. In this example it is clearly that $N_{x^*} \bigcap X \supset \{0\}$. From (5.1), we get $-1 \leq \dfrac{x+1}{x-1} + 2x$. Since $x \in X$, then we have

$$x^2 \leq 0, \ x \in N_{x^*} \bigcap X,$$

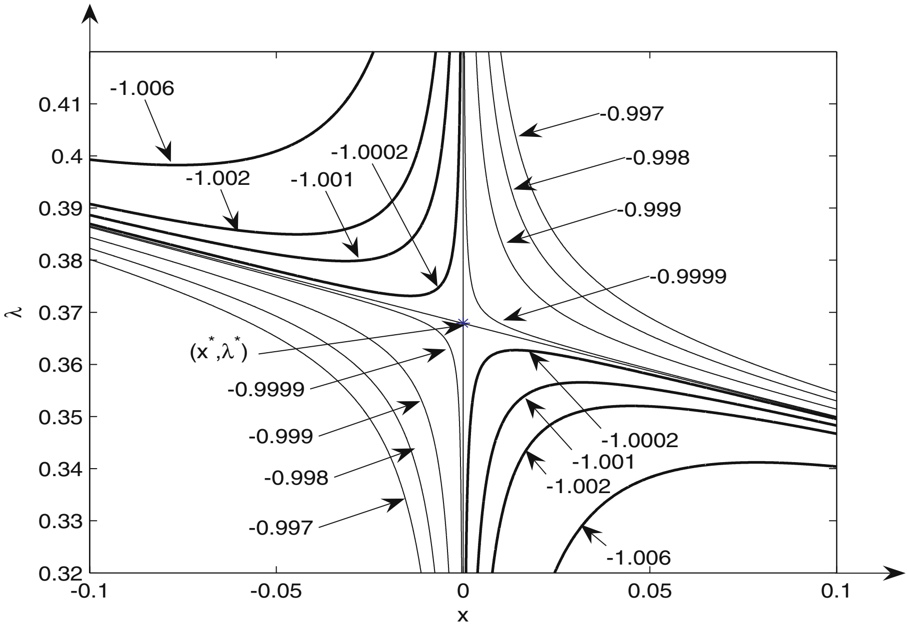

Fig. 2 Contours of the Lagrangian function $L_2(x,\mu)$ for Problem (1EP) near (x^*,μ_2^*)

which is a contradiction. And we set $\psi_{0p}(y) = y$ and $\psi_{1p} = e^{(y+1)^p}$. Then Problem (1P) is equivalent to the following problem:

$$(1EP): \begin{cases} \min \ \dfrac{x+1}{x-1} \\ \text{s.t. } e^{(x+1)^p} \le e, \\ x \in \ X = \{x \in R : x < 1\}. \end{cases}$$

From (3.4), we have $\mu_p^* = \dfrac{2}{pe}$ and $L_P(x,\mu) = \dfrac{x+1}{x-1} + \mu(e^{(x+1)^p} - e)$. By calculation, we get that when $p = 2$, (x^*,μ_2^*) is a local saddle point of $L_2(x,\mu)$. Figure 2 depicts contours of $L_2(x,\mu)$ in the neighborhood of (x^*,μ_2^*).

Example 5.2 Consider the following problem

$$(2P): \begin{cases} \min f_0(x) = (4-x_1)(4-x_2) \\ \text{s.t. } f_1(x) = 2x_1 - x_2 \le 2, \\ f_2(x) = x_2 \le 2, \\ x \in \ X = [1,4]^2. \end{cases}$$

Note that $x^* = (2,2)$ and $\lambda^* = (0.5, 0.125)$ is one solution to (4.1). It can be verified that $M(x^*) = \{0\}$ and the Hessian of the Lagrangian function is

$$\nabla^2 L(x^*,\lambda^*) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

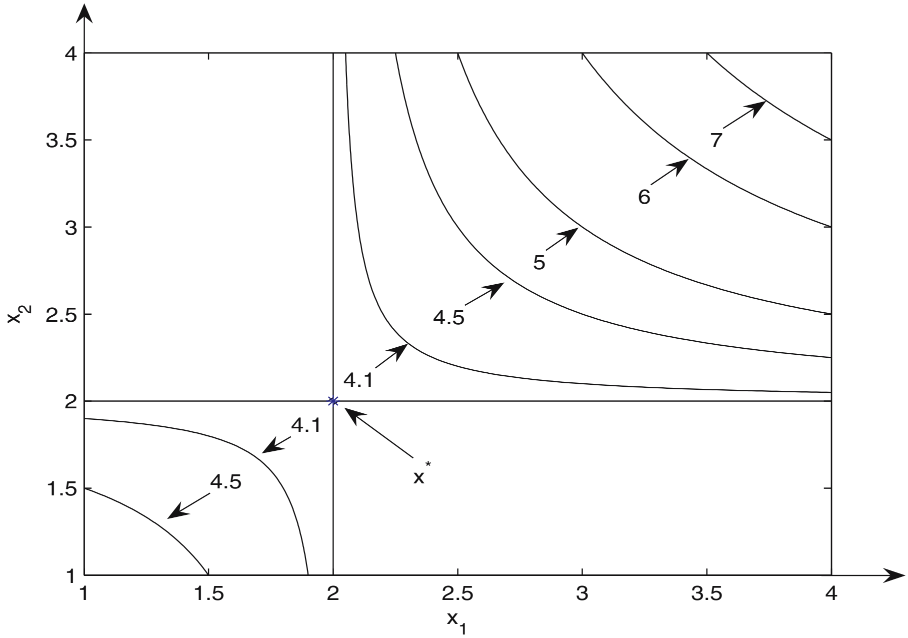

Fig. 3 Contours of the Lagrangian function $L(x, \lambda^*)$ for Problem (2P)

which is indefinite. Let $\psi_{0p}(y) = -e^{-\frac{y}{p}}$ and $\psi_{ip} = e^{py},\ i = 1, 2$, then Problem (2P) is transformed into the following EP:

$$(2\text{EP}) : \begin{cases} \min -e^{(-\frac{(4-x_1)(4-x_2)}{p})} \\ \text{s.t. } e^{p(2x_1-x_2)} \leq e^{2p}, \\ e^{px_2} \leq e^{2p}, \\ x \in X = [1, 4]^2. \end{cases}$$

By (3.4), we have $\mu_p^* = (\frac{1}{p^2}e^{(-2p-4/p)}, \frac{3}{p^2}e^{(-2p-4/p)})$, and

$$L_p(x, \mu_p^*) = -e^{-\frac{(4-x_1)(4-x_2)}{p}} + \frac{1}{p^2}e^{(2px_1-px_2-2p-4/p)} + \frac{3}{p^2}e^{(px_2-2p-4/p)} - e^{-4/p}.$$

By direct calculation, we get

$$\nabla^2 L_p(x^*, \mu_p^*) = e^{-4/p}\begin{pmatrix} 4 - 4/p^2 & -4/p^2 + 1/p - 2 \\ -4/p^2 + 1/p - 2 & 4 - 4/p^2 \end{pmatrix}.$$

Observe that when $p = 2$, $\mu_2^* = \{6.25 \times 10^{-4}, 1.9 \times 10^{-3}\}$, and $\nabla^2 L_2(x^*, \mu_2^*) = \begin{pmatrix} 3e^{-2} & -2.5e^{-2} \\ 2.5e^{-2} & 3e^{-2} \end{pmatrix}$ is positive definite. Figures 3 and 4 depict contours of $L(x, \lambda^*)$ for Problem (2P) and $L_2(x, \mu_2^*)$ for 2EP, respectively.

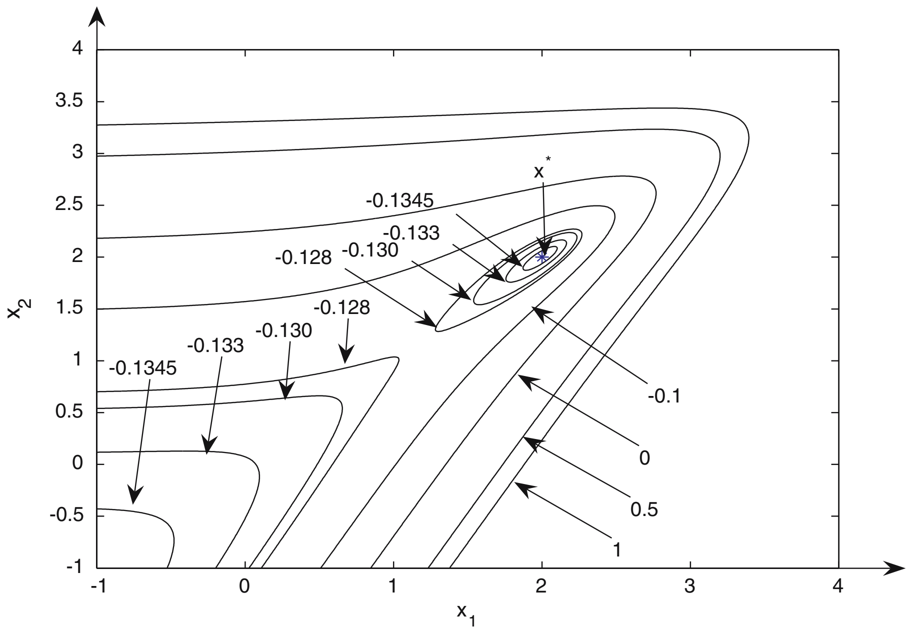

Fig. 4 Contours of the Lagrangian function $L_2(x, \mu_2^*)$ for Problem (2EP)

Example 5.3 Consider the following problem:

$$(3P): \begin{cases} \min f_0(x) = 4 - x_1 x_2 \\ \text{s.t. } f_1(x) = x_1 + 4x_2 \le 1, \\ x \in X = [0,1]^2, \end{cases}$$

$x^* = (0.5, 0.125)^T$ and $\lambda^* = 0.125$ satisfy (4.1). It can be verified that $M(x^*) = \{d \in R^2 | d_1 + 4d_2 = 0\}$ and the Hessian matrix of the Lagrangian function at (x^*, λ^*) is

$$\nabla^2 L(x^*, \lambda^*) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix},$$

which is an indefinite matrix and positive definite on $M(x^*)$.

Set $\psi_{0p} = y$ and $\psi_{1P} = (1 + \frac{y}{p})^{p^2}$, then Problem (3P) is equivalent to the following problem:

$$(3\text{EP}): \begin{cases} \min\ 4 - x_1 x_2 \\ \text{s.t. } \left(\frac{1 + (x_1 + 4x_2)}{p}\right)^{p^2} \le (1 + \frac{1}{p})^{p^2}. \\ x \in X = [0,1]^2. \end{cases}$$

By (3.4) we have $\mu_p^* = \dfrac{1}{8p(1+1/p)^{p^2-1}}$, then we get

$$L_p(x, \mu_p^*) = 4 - x_1 x_2 + \frac{1}{8p(1+1/p)^{p^2-1}} \left[\left(1 + \frac{x_1 + 4x_2}{p}\right)^{p^2} - (1 + 1/p)^{p^2} \right].$$

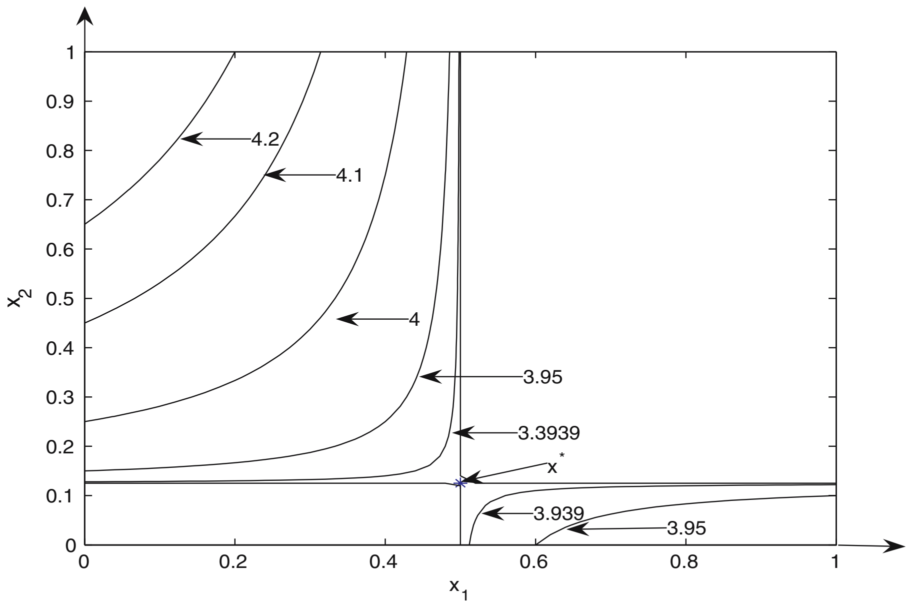

Fig. 5 Contours of the Lagrangian function $L(x, \lambda^*)$ for Problem (3P)

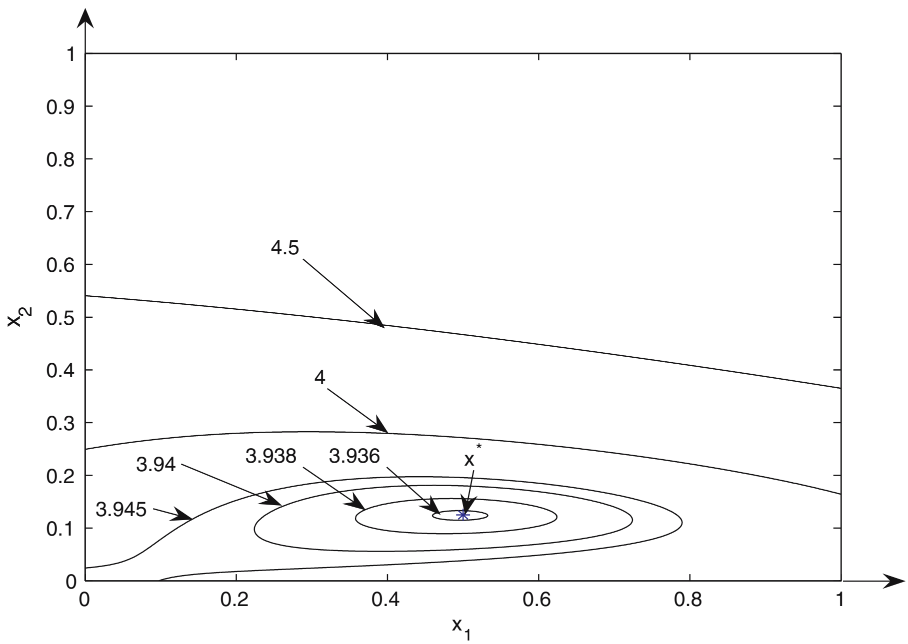

Fig. 6 Contours of the Lagrangian function $L_3(x, \mu_3^*)$ for Problem (3EP)

The Hessian matrix of $L_p(x,\mu)$ at (x^*,μ_p^*) is

$$\nabla^2 L(x^*,\mu_p^*) = \begin{pmatrix} (p^2-1)(1+1/p)^{p^2-2} & 4(p^2-1)(1+1/p)^{p^2-2} \\ 4(p^2-1)(1+1/p)^{p^2-2} & 16(p^2-1)(1+1/p)^{p^2-2} \end{pmatrix}.$$

Note that when p=3, $\mu_3^* = 0.004$, and $\nabla^2 L_3(x^*,\mu_3^*) = \begin{pmatrix} 0.25 & 0 \\ 0 & 4 \end{pmatrix}$, which is positive definite.

Figures 5 and 6 depict contours of $L(x,\lambda^*)$ for Problem (3P) and $L_3(x,\mu_3^*)$ for 3EP, respectively.

6 Conclusions

In this paper, we propose a general class of transformation methods including the transformations presented in [3–5,7] as special cases. Then under certain assumptions, we prove the local saddle point result and the local convexification of the the Lagrangian function. Hence, the main results obtained in this paper could be viewed as an expansion of the results in [3–5,7]. The important aspects of this paper lie not only in theory but also in practice. Since it provides us with more specific transformations to obtain an EP with better properties, we can tackle some practical problems that can not be done with by using the transformations presented in [3–5,7], or can do with them more efficiently.

However, about the transformation method, there are still problems that need to be pursued. The most important and challenging one among them may be how to select the most suitable transformation function for a given function of interest. Besides, how to identify the lower bound of p that has to be used in the transformation and how the value of p affects the equivalent function obtained are also important topics. It is evident that the solving of these problems would be very helpful in practice. But until now few papers have addressed these problems systematically. So we suggest that more efforts should be devoted to them in the future.

Acknowledgment Project 10261005 supported by National Natural Science Foundation of China.

References

1. Wolfe P.: A duality theorem for nonlinear programming. Q. Appl. Math. **19**, 239–244 (1961)
2. Luenberger D.G.: Linear and Nonlinear Programming, 2nd edn. Addison-Wesley, Reading, MA (1984)
3. Li D.: Zero Duality gap for a class of nonconvex optimization problems. J. Optim. Theory Appl. **85**, 309–324 (1995)
4. Xu Z.K.: Local saddle points and convexification for nonconvex optimization problems. J. Optim. Theory Appl. **94**, 739–746 (1997)
5. Li D., Sun, X.L.: Local convexification of Lagrangian function in nonconvex optimization. J. Optim. Theory Appl. **104**, 109–120 (2000)
6. Mokhtar S.B., Hanif D.S., Shetty C.M.: Nonlinear Programming, Theory and Algorithms, 2nd edn. Wiley Inc, New York, (1993)
7. Xu Z.K.: Introduction to Mathematical Programming(in Chinese). Science Publisinghouse, Beijing, China, (1999)