SIAM J. OPTIM.
Vol. 11, No. 4, pp. 1119–1144
© 2001 Society for Industrial and Applied Mathematics

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

A NONLINEAR LAGRANGIAN APPROACH TO CONSTRAINED OPTIMIZATION PROBLEMS*

X. Q. YANG† AND X. X. HUANG‡

Abstract. In this paper we study nonlinear Lagrangian functions for constrained optimization problems which are, in general, nonlinear with respect to the objective function. We establish an equivalence between two types of zero duality gap properties, which are described using augmented Lagrangian dual functions and nonlinear Lagrangian dual functions, respectively. Furthermore, we show the existence of a path of optimal solutions generated by nonlinear Lagrangian problems and show its convergence toward the optimal set of the original problem. We analyze the convergence of several classes of nonlinear Lagrangian problems in terms of their first and second order necessary optimality conditions.

Key words. augmented Lagrangian, nonlinear Lagrangian, zero duality gap, optimal path, necessary optimality condition, smooth approximate variational principle

AMS subject classifications. 90C30, 49J52, 49M35

PII. S1052623400371806

1. Introduction.

It is well known that unconstrained optimization methods, such as the Lagrangian dual and penalty methods, have been extensively studied in order to solve constrained optimization problems. A zero duality gap can be guaranteed if conventional Lagrangian functions are used to define the dual problem under convexity or generalized convexity assumptions. Nevertheless, for a nonconvex constrained optimization problem, a nonzero duality gap may occur between the original problem and the conventional Lagrangian dual problem. To overcome this drawback, various approaches have been proposed in the literature. The convex conjugate framework in [16] was extended in [3, 13] for nonconvex optimization problems. In [17], a general augmented Lagrangian function was introduced, and it was shown that the general augmented dual problem constructed with an appropriately selected perturbation function yields a zero duality gap result. Recently, nonlinear Lagrangian functions were introduced using increasing functions for solving constrained optimization problems. A zero duality gap result is established between a nonconvex constrained optimization problem and the dual problem defined by using a nonlinear Lagrangian function in [10, 14, 18, 19]. In passing, we mention that exact penalization-type results were established for the augmented Lagrangian function in [17], for nonlinear Lagrangian functions under generalized calmness-type conditions for scalar optimization problems in [19], and for vector optimization problems in [12].

Noting the fact that, for nonconvex constrained optimization problems, both zero duality gap results in terms of augmented Lagrangian dual functions in [17] and nonlinear Lagrangian dual functions in [19] were established under very mild conditions, it is interesting to investigate whether there is a connection between these two

*Received by the editors May 5, 2000; accepted for publication (in revised form) November 11, 2000; published electronically May 16, 2001. This work was partially supported by the Research Grants Council of Hong Kong (grant PolyU B-Q359).

http://www.siam.org/journals/siopt/11-4/37180.html

†Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong (mayangxq@polyu.edu.hk).

‡Department of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047, China. Current address: Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong (mahuangx@polyu.edu.hk).

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

results. Therefore, the first goal of this paper is to establish an equivalence between zero duality gap properties, which are described using a class of augmented Lagrangian functions with specially structured perturbation functions, and nonlinear Lagrangian functions, respectively.

Recently, a wide class of penalty and barrier methods was studied in [2], including a number of specific functions in the literature (see [5, 9]). For convex programming problems, the existence of a path of optimal solutions generated by these penalty methods was established and its convergence toward the optimal set of the original problem was given. Hence, the second goal of this paper is to show, for nonconvex inequality constrained optimization problems, the existence of a path of optimal solutions generated by a general nonlinear Lagrangian function and to show its convergence toward the optimal set of the original problem. Moreover, we illustrate that this result can be specialized to convex programming problems, and thus a parallel result to that in [2] is obtained.

We then investigate the convergence analysis of nonlinear Lagrangian methods in terms of first and second order necessary optimality conditions, where the multipliers are independent of vectors in the tangential subspace of the active constraints. This follows the usual method, as in [1, 22]. Thus we need to derive, for example, corresponding second order necessary conditions for nonlinear Lagrangian problems. However, for cases where nonlinear Lagrangian functions are not twice differentiable, the derivation of this type of second order optimality condition of nonlinear Lagrangian problems is by no means an easy task. For example, one of the nonlinear Lagrangian functions to be considered is of the minimax type. Thus, the resulting problem is an unconstrained minimax optimization problem or, more generally, a convex composite optimization problem. Second order necessary conditions for convex composite optimization problems were established in [4, 7, 13, 23]. However, in these conditions the multipliers depend on the choice of the vector in the tangential subspace of the active constraints. These second order conditions are not applicable in our cases. Nevertheless, we are able to derive the required first and second order necessary conditions for these nonlinear Lagrangian problems by means of a higher order smooth approximation and the smooth approximate variational principle in [6, 8].

The outline of the paper is as follows. In section 2, we review the zero duality gap properties, which are obtained using augmented Lagrangian functions and nonlinear Lagrangian functions. In section 3, we show that if the dual problem which is constructed with an augmented Lagrangian and a specially structured perturbation function yields a zero duality gap, then the dual problem defined by nonlinear Lagrangian dual functions also yields a zero duality gap, and vice versa. In section 4, we show the existence of a path of optimal solutions generated by nonlinear Lagrangian problems and show its convergence to the optimal set of the original problem. In section 5, we carry out convergence analysis of this method for several classes of nonlinear Lagrangians in terms of first and second order necessary optimality conditions.

2. Zero duality gaps.

In this section, we introduce some definitions and recall the zero duality gap properties, which are described by augmented Lagrangian functions and nonlinear Lagrangian functions, respectively. Consider the following inequality constrained optimization problem (P):

$$\begin{array}{ll} \inf & f(x) \\ \text{s.t.} & x \in X, \quad g_j(x) \leq 0, \ j = 1, \dots, q, \end{array}$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where $X \subset R^p$ is a nonempty and closed set, and $f,\ g_j : X \to R^1$ $(j = 1, \ldots, q)$ are real-valued functions. Denote by M_P the infimum of (P) and by X_0 the feasible set of (P):

$$X_0 = \{x \in X \ : \ g_j(x) \le 0 \quad \forall j = 1, \ldots, q\}.$$

In this paper, we assume that $X_0 \neq \emptyset$.

Throughout this paper, we also assume that

$$f(x) \ge 0 \quad \forall x \in X.$$

Note that this assumption is not very restrictive. Otherwise, we may replace the objective function $f(x)$ with $1 + e^{f(x)}$, which satisfies the assumption; $\inf_{x\in X} f(x) > 0$ also holds; and the resulting constrained optimization problem has the same set of (local) solutions as that of (P).

Let $c : R^1_+ \times R^q \to R^1$ be a real-valued function. c is said to be *increasing* on $R^1_+ \times R^q$ if, for any $y^1, y^2 \in R^1_+ \times R^q$, $y^2 - y^1 \in R^{q+1}_+$ implies that $c(y^1) \le c(y^2)$. We will consider increasing and lower semicontinuous (l.s.c.) functions c defined on $R^1_+ \times R^q$, which enjoy the following properties:

(A) There exist positive real numbers $a_j, j = 1, \ldots, q$, such that, for any $y = (y_0, y_1, \ldots, y_q) \in R^1_+ \times R^q$, we have

$$c(y) \ge \max\{y_0, a_1 y_1, \ldots, a_q y_q\}.$$

(B) For any $y_0 \in R^1_+$,

$$c(y_0, 0, \ldots, 0) = y_0.$$

Let $y^+ = \max\{y, 0\}$ for $y \in R$. The following are some examples of function c (see [18]):

$$c(y) = \max\{y_0, y_1, \ldots, y_q\},$$
$$c(y) = \left(y_0^k + \sum_{j=1}^{q} {y_j^+}^k\right)^{1/k}, \quad k \in (0, +\infty).$$

The convergence analysis of optimality conditions for nonlinear Lagrangian dual problems defined by these functions (see below) will be given in section 5.

Let c be an increasing function defined as above, and

$$F(x, d) = (f(x), d_1 g_1(x), \ldots, d_q g_q(x)) \quad \forall x \in X,\ d = (d_1, \ldots, d_q) \in R^q_+.$$

The function defined by

$$L(x, d) = c(F(x, d))$$

is called a *nonlinear Lagrangian* corresponding to c.

The *nonlinear Lagrangian dual function* for (P) corresponding to c is defined by

$$\phi(d) = \inf_{x\in X} L(x, d), \quad d \in R^q_+.$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The *nonlinear Lagrangian dual problem* (D_N) for (P) corresponding to c is defined by

$$\sup_{d\in R^q_+} \phi(d).$$

Denote by M_N the supremum of problem (D_N). It can be easily verified [18, 19] that the following weak duality result holds:

$$M_N \leq M_P. \tag{1}$$

DEFINITION 2.1. *Let c be an increasing function satisfying properties* (A) *and* (B)*. The zero duality gap property with respect to c between* (P) *and* (D_N) *is said to hold if $M_N = M_P$.*

DEFINITION 2.2 (see [2]). *Let $X \subset R^p$ be unbounded. The function $h : X \to R^1$ is said to be* 0*-coercive on X if*

$$\lim_{x\in X, \|x\|\to+\infty} h(x) = +\infty.$$

Let

$$\begin{aligned} G(x) &= \max\{g_1(x), \ldots, g_q(x)\}, \quad x \in X, \\ h(x) &= \max\{f(x), G(x)\}, \quad x \in X. \end{aligned} \tag{2}$$

THEOREM 2.3. *Suppose that h, defined by* (2)*, is* 0*-coercive if X is unbounded. If the functions $f, g_1, \ldots, g_q$ are l.s.c., then the zero duality gap property with respect to c between* (P) *and (D_N) holds.*

Proof. It is clear that $L(x, d)$ is an increasing function of d. The result follows from Theorem 4.2 in section 4. $\square$

Let us recall the definition of the augmented Lagrangian function for (P) (for details, see Chapter 11, section K^* in [17]). Let $\varphi : R^p \to R^1 \bigcup\{+\infty\}$:

$$\varphi(x) = \begin{cases} f(x) & \text{if } x \in X_0; \\ +\infty & \text{otherwise.} \end{cases}$$

Let $\overline{f} : R^p \times R^q \to R^1 \bigcup\{+\infty\}$ be a perturbation function [17, p. 519] such that $\overline{f}(x, 0) = \varphi(x),\ x \in R^p$. Let σ be an *augmenting function*, namely, a proper, l.s.c., and convex function with the unique minimum at 0 and $\sigma(0) = 0$. The corresponding *augmented Lagrangian* $\bar{l} : R^p \times R^q \times (0, +\infty) \to R^1 \bigcup\{+\infty, -\infty\}$ with parameter $r > 0$ is defined by

$$\bar{l}(x, y, r) = \inf\{\overline{f}(x, u) + r\sigma(u) - \langle y, u\rangle : u \in R^q\},$$

where $\langle y, u\rangle$ denotes the inner product of y and u.

The corresponding *augmented Lagrangian dual function* is

$$\psi(y, r) = \inf\{\bar{l}(x, y, r) : x \in R^p\},$$

and the *augmented Lagrangian dual problem* (D_A) is

$$\sup_{(y,r)\in R^q\times(0,+\infty)} \psi(y, r).$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Let M_A denote the supremum of the dual problem (D_A). The following weak duality for (P) and (D_A) holds (see [17]):

$$M_A \leq M_P. \tag{3}$$

Definition 2.4. *Let $\overline{f} : R^p \times R^q \to R^1 \bigcup\{+\infty\}$ be a perturbation function and σ be an augmenting function. The zero duality gap property with respect to $\overline{f}$ and σ between* (P) *and* (D_A) *is said to hold if $M_A = M_P$.*

Definition 2.5 (see [17]). *A function $h : R^p \times R^q \to R^1 \bigcup\{+\infty, -\infty\}$ with values $h(x,u)$ is said to be level-bounded in x and locally uniform in u if, for each $\overline{u} \in R^q$ and $\alpha \in R^1$, there exists a neighborhood $V(\overline{u})$ of $\overline{u}$, along with a bounded set $D \subset R^p$, such that $\{x \in R^p : h(x,v) \leq \alpha\} \subset D \ \forall v \in V(\overline{u})$.*

Theorem 2.6 (see [17]). *Assume that the perturbation function $\overline{f} : R^p \times R^q \to R^1 \bigcup\{+\infty\}$ is proper and l.s.c., and that $\overline{f}(x,u)$ is level-bounded in x and locally uniform in u. Let σ be an augmenting function. Suppose further that there exist $\overline{y} \in R^q$ and $\overline{r} > 0$ such that*

$$\inf\{\overline{f}(x,u) + \overline{r}\sigma(u) - \langle \overline{y}, u\rangle : x \in R^p, u \in R^q\} > -\infty. \tag{4}$$

Then $M_A = M_P$.

3. Equivalence of zero duality gaps.

In this section, we establish an equivalence of zero duality gap properties between a class of augmented Lagrangian dual problems and the nonlinear Lagrangian dual problem.

Denote the indicator function of a set $D \subset R^q$ by

$$\delta_D(y) = \begin{cases} 0 & \text{if } y \in D; \\ +\infty & \text{otherwise.} \end{cases}$$

It is easy to check that (P) is equivalent to the following problem:

$$\inf_{x \in X} f(x) + \delta_{R^q_-}(g_1(x), \ldots, g_q(x))$$

in the sense that the two problems have the same sets of (locally) optimal solutions and optimal values. Let

$$H(x) = (g_1(x), \ldots, g_q(x)),$$

$$\overline{f}(x,u) = f(x) + \delta_{R^q_-}(H(x) + u) + \delta_X(x). \tag{5}$$

Then, for $x \in R^p$, $\overline{f}(x,0) = \varphi(x)$. Thus, $\overline{f}(x,u)$ is a perturbation function.

Lemma 3.1. *Let the perturbation function be defined by* (5), *σ an augmenting function, and $v = (v_1, \ldots, v_q)$. Then*

$$\bar{l}(x,y,r) = \begin{cases} f(x) + \sum_{j=1}^q y_j g_j(x) + \inf_{v \geq 0} \left\{ \sum_{j=1}^q y_j v_j + r\sigma(-g_1(x) - v_1, \ldots, -g_q(x) - v_q) \right\} & \textit{if } x \in X, \\ +\infty & \textit{otherwise.} \end{cases}$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. Let $x \in X$.

$$\begin{aligned}\bar{l}(x,y,r) &= \inf\{\overline{f}(x,u) + r\sigma(u) - \langle y,u\rangle : u \in R^q\}\\ &= \inf_{v\geq 0}\left\{f(x) + \sum_{j=1}^{q} y_j(g_j(x)+v_j) + r\sigma(-g_1(x)-v_1,\ldots,-g_q(x)-v_q)\right\}\\ &= f(x) + \sum_{j=1}^{q} y_j g_j(x) + \inf_{v\geq 0}\left\{\sum_{j=1}^{q} y_j v_j + r\sigma(-g_1(x)-v_1,\ldots,-g_q(x)-v_q)\right\}.\end{aligned}$$

Let $x \notin X$. It is clear that $\overline{f}(x,u) = +\infty$. Thus $\bar{l}(x,y,r) = +\infty$. ▯

The following proposition summarizes some properties of augmented Lagrangian $\bar{l}$, where $\overline{f}$ is defined by (5), and the nonlinear Lagrangian L.

LEMMA 3.2. *Let the perturbation function $\overline{f}(x,u)$ be defined by* (5). *Then, the following properties of augmented Lagrangian function $\bar{l}$ hold:*

(I) $\bar{l}(x,y,r) \leq f(x)\ \forall x \in X_0,\ y \in R^q,\ r > 0,$ *and* $\bar{l}(x,0,r) = f(x)\ \forall x \in X_0,$ $r > 0.$

(II) $\bar{l}(x,0,r) \geq f(x)\ \forall x \in X.$

(III) *For any* $x \in X\backslash X_0,\ y \in R^q,\quad \bar{l}(x,y,r) \to +\infty$ *as* $r \to +\infty,$

and the following properties of nonlinear Lagrangian function L hold:

(I′) $L(x,d) = f(x)\ \forall x \in X_0.$

(II′) $L(x,d) \geq f(x)\ \forall x \in X.$

(III′) *For any* $x \in X\backslash X_0,\quad L(x,d) \to +\infty$ *as* $d \to +\infty.$

Here the notation $d = (d_1,\ldots,d_q) \to +\infty$ *means that* $d_j \to +\infty$ *for each* $j \in \{1,\ldots,q\}$.

It follows from Lemma 3.2 that $\bar{l}(x,0,r)$ behaves very similarly to $L(x,re)$, where $e = (1,\ldots,1) \in R^q_+$. For any $x \in R^p$, let

$$J^+(x) = \{j \in \{1,\ldots,q\} : g_j(x) > 0\},\ J(x) = \{j \in \{1,\ldots,q\} : g_j(x) = 0\}.$$

PROPOSITION 3.3. *Let augmenting function σ be a finite and l.s.c. function which attains its minimum* 0 *at* $0 \in R^q$. *Let the perturbation function $\overline{f}(x,u)$ defining the augmented Lagrangian be selected as* (5). *If $M_A = M_P$, then $M_N = M_P$.*

Proof. If $M_N = M_P$ fails to hold by weak duality (1) of the nonlinear Lagrangian, then there exists $\epsilon_0 > 0$ such that $M_N \leq M_P - \epsilon_0$.

By the assumption, we get

$$M_A = \sup_{(y,r)\in R^q\times(0,+\infty)} \inf_{x\in X} \bar{l}(x,y,r) = M_P.$$

Then, for $\frac{\epsilon_0}{4} > 0$, there exist $\bar{y} \in R^q$ and $\bar{r} > 0$ such that $\bar{l}(x,\bar{y},\bar{r}) \geq M_P - \frac{\epsilon_0}{4}\ \forall x \in X$. That is, for any $x \in X$,

(6)
$$f(x) + \sum_{j=1}^{q} \bar{y}_j g_j(x) + \inf_{v\geq 0}\left\{\sum_{j=1}^{q} \bar{y}_j v_j + \bar{r}\sigma(-g_1(x)-v_1,\ldots,-g_q(x)-v_q)\right\} \geq M_P - \frac{\epsilon_0}{4}.$$

Let $d_n = (d_{1,n},\ldots,d_{q,n}) \to +\infty$. Thus,

$$\inf_{x\in X} L(x,d_n) = q(d_n) \leq M_P - \epsilon_0.$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

There then exists $x_n \in X$, such that

$$0 \le f(x_n) \le L(x_n, d_n) \le M_P - \frac{\epsilon_0}{2} \tag{7}$$

and

$$0 < \max\{a_1 d_{1,n} g_1(x_n), \ldots, a_q d_{q,n} g_q(x_n)\} \le L(x_n, d_n) \le M_P - \frac{\epsilon_0}{2}. \tag{8}$$

Equation (7) implies

$$\begin{aligned} f(x) + \sum_{j=1}^{q} \bar{y}_j g_j(x) + \sum_{j=1}^{q} \bar{y}_j v_j + \bar{r}\sigma(-g_1(x) - v_1, \ldots, -g_q(x) - v_q) \\ \ge M_P - \frac{\epsilon_0}{4} \quad \forall v \ge 0. \end{aligned} \tag{9}$$

Let $x = x_n$ in (9), $v_{j,n} = -g_j(x_n)$ if $g_j(x_n) \le 0$, and $v_{j,n} = 0$ if $g_j(x_n) > 0$, $j = 1, \ldots, q$. We get

$$f(x_n) + \sum_{j \in J^+(x_n)} \bar{y}_j g_j(x_n) + \bar{r}\sigma(-v^*_{1,n}, \ldots, -v^*_{q,n}) \ge M_P - \frac{\epsilon_0}{4}, \tag{10}$$

where $v^*_{j,n} = g_j(x_n), j \in J^+(x_n)$, and $v^*_{j,n} = 0$ otherwise.

By the assumption on σ, we know that σ is locally Lipschitz around $0 \in R^q$.

Equation (8) and $d_n \to +\infty$ yield that $0 < \max_{j \in J^+(x_n)}\{g_j(x_n)\} \to 0$ as $n \to +\infty$.

Therefore, there exist $\beta > 0$ and $n_0 > 0$ such that for $n \ge n_0$,

$$\sigma(-v^*_{1,n}, \ldots, -v^*_{q,n}) \le \beta \sum_{j=1}^{q} |v^*_j|.$$

Consequently, the facts above and (10) jointly yield

$$\begin{aligned} f(x_n) &+ \left(\sum_{j \in J^+(x_n)} (|\bar{y}_j| + \bar{r}\beta) \right) \max_{j \in J^+(x_n)} g_j(x_n) \\ &\ge f(x_n) + \sum_{j \in J^+(x_n)} (\bar{y}_j + \bar{r}\beta) g_j(x_n) \\ &= f(x_n) + \sum_{j \in J^+(x_n)} \bar{y}_j g_j(x_n) + \bar{r}\beta \sum_{j=1}^{m} |v^*_j| \\ &\ge f(x_n) + \sum_{j \in J^+(x_n)} \bar{y}_j g_j(x_n) + \bar{r}\sigma(-v^*_{1,n}, \ldots, -v^*_{q,n}) \\ &\ge M_P - \frac{\epsilon_0}{4}. \end{aligned}$$

Let $\gamma = \sum_{j=1}^{q} |\bar{y}_j| + q\bar{r}\beta$. Then

$$f(x_n) + \gamma \max_{j \in J^+(x_n)}\{g_j(x_n)\} \ge M_P - \frac{\epsilon_0}{4}. \tag{11}$$

On the other hand, let $\lambda_n = \min_{1 \le j \le q}\{a_j d_{j,n}\}$. It follows from (8) that

$$\lambda_n \max\{g_1(x_n), \ldots, g_q(x_n)\} \le L(x_n, d_n) \le M_P - \epsilon_0/2.$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Thus,

$$\max_{j\in J^+(x_n)}\{g_j(x_n)\} \le \frac{M_P-\epsilon_0/2}{\lambda_n}.$$

By (11), we have

$$\begin{aligned}M_P-\frac{\epsilon_0}{4} &\le f(x_n)+\frac{\gamma}{\lambda_n}\left(M_P-\frac{\epsilon_0}{2}\right)\\ &\le M_P-\frac{\epsilon_0}{2}+\frac{\gamma}{\lambda_n}\left(M_P-\frac{\epsilon_0}{2}\right),\end{aligned}$$

where the last inequality follows from (7).

Noticing that $\lambda_n \to +\infty$ as $n\to\infty$ and letting $n\to\infty$, we obtain

$$M_P-\frac{\epsilon_0}{4}\le M_P-\frac{\epsilon_0}{2},$$

which is a contradiction. □

PROPOSITION 3.4. *Let function c defining the nonlinear Lagrangian L be continuous. If $M_P = M_N$, then $M_P = M_A$.*

Proof. By the weak duality (3) of the augmented Lagrangian, $M_A \le M_P$. Suppose to the contrary that there exists $\epsilon_0 > 0$ such that

$$M_A = \sup_{(y,r)\in R^q\times(0,+\infty)}\inf_{x\in X}\bar{l}(x,y,r)\le M_P-\epsilon_0.$$

Thus,

$$\inf_{x\in X}\bar{l}(x,y,r)\le M_P-\epsilon_0 \quad \forall (y,r)\in R^q\times(0,+\infty).$$

In particular,

$$\inf_{x\in X}\bar{l}(x,0,r)\le M_P-\epsilon_0 \quad \forall r\in(0,+\infty).$$

Let $r_n\to+\infty$. There then exists $n_0>0$ such that, for $n\ge n_0$ and some $x_n\in X$, $\bar{l}(x_n,0,r_n)\le M_P-\frac{\epsilon_0}{2}$. Thus,

$$f(x_n)+\inf_{v\in R^q_+}\{r_n\sigma(-g_1(x_n)-v_1,\ldots,-g_q(x_n)-v_q)\}\le M_P-\frac{\epsilon_0}{2}.$$

Furthermore, there exists $v_n=(v_{1,n},\ldots,v_{q,n})\in R^q_+$ such that

$$f(x_n)+r_n\sigma(-g_1(x_n)-v_{1,n},\ldots,-g_q(x_n)-v_{q,n})\le M_P-\frac{\epsilon_0}{4},\quad n\ge n_0. \tag{12}$$

Noticing that $f(x_n)\ge 0\ \forall n$, we deduce from (12) that

$$\sigma(-g_1(x_n)-v_{1,n},\ldots,-g_q(x_n)-v_{q,n})\le\frac{M_P-\epsilon_0/4}{r_n}.$$

Thus

$$\limsup_{n\to+\infty}\sigma(-g_1(x_n)-v_{1,n},\ldots,-g_q(x_n)-v_{q,n})=0.$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Since σ is a convex function with a unique minimum at 0 with $\sigma(0) = 0$, it follows that

$$g_j(x_n) + v_{j,n} \to 0 \quad \text{ as } n \to +\infty, (j = 1, \dots, q).$$

Let $\epsilon_n = \max_{1 \le j \le q} g_j(x_n)$. Then $\epsilon_n > 0$ and $\epsilon_n \to 0$ as $n \to +\infty$. It follows from (12) and $f(x_n) \ge 0$ that

$$0 \le f(x_n) \le M_P - \frac{\epsilon_0}{4}, \quad n \ge n_0. \tag{13}$$

Without loss of generality, we assume that

$$f(x_n) \to t_0 \ge 0 \quad \text{ as } n \to +\infty. \tag{14}$$

The combination of (13) and (14) yields $0 \le t_0 \le M_P - \frac{\epsilon_0}{4}$. Let $d = (d_1, \dots, d_q) \in R^q_+$. Then, by the monotonicity of c,

$$c(f(x_n), d_1 g_1(x_n), \dots, d_q g_q(x_n)) \le c(f(x_n), d\epsilon_n, \dots, d\epsilon_n).$$

Taking the upper limit as $n \to +\infty$ and applying the continuity of c, we obtain

$$\limsup_{n \to +\infty} c(f(x_n), d_1 g_1(x_n), \dots, d_q g_q(x_n)) \le c(t_0, 0, \dots, 0) = t_0 \le M_P - \frac{\epsilon_0}{4}.$$

Hence, for each $d \in R^q_+$, $\exists n(d) > 0$ such that

$$c(f(x_{n(d)}), d_1 g_1(x_{n(d)}), \dots, d_q g_q(x_{n(d)})) \le M_P - \frac{\epsilon_0}{8}.$$

It follows that

$$\inf_{x \in X} c(f(x), d_1 g_1(x), \dots, d_q g_q(x)) \le M_P - \frac{\epsilon_0}{8}.$$

As $d \in R^q_+$ is arbitrary, we conclude that $M_N \le M_P - \frac{\epsilon_0}{8}$, which contradicts the assumption $M_N = M_P$. The proof is complete. ∎

The relationships are summarized below between the zero duality properties of the augmented Lagrangian dual problem (D_A), with the perturbation function $\overline{f}(x,u)$ selected as (5), and the nonlinear Lagrangian dual problem (D_N).

THEOREM 3.5. *Consider the problem* (P)*, the nonlinear Lagrangian dual problem* (D_N)*, and the augmented Lagrangian dual problem* (D_A)*. If the function c defining the nonlinear Lagrangian L is continuous, the perturbation function $\overline{f}(x,u)$ defining the augmented Lagrangian is selected as* (5)*, and the augmenting function σ is finite, l.s.c., and convex, attaining its minimum* 0 *at* $0 \in R^q$*, then the following two statements are equivalent:*

(i) $M_A = M_P$*;*

(ii) $M_N = M_P$*.*

The following example verifies Theorem 3.5.

Example 3.1. Consider the problem

$$\begin{array}{ll} \inf & f(x) \\ \text{s.t.} & x \in X, \ g(x) \le 0, \end{array}$$

where $X = [0, +\infty)$, $f(x) = 1/(x+1)$ $\forall x \in X$; $g(x) = x - 1$ if $0 \le x \le 1$; $g(x) = 1/\sqrt{x} - 1/x$ if $1 < x < +\infty$. Then $M_P = 1/2$.

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Let $c(y_1, y_2) = \max\{y_1, y_2\}$ $\forall y_1 \geq 0, y_2 \in R^1$. It is easy to check that $M_N = 0$. Hence $M_N < M_P$.

Let

$$\overline{f}(x,u) = f(x) + \delta_{R^1_-}(g(x)+u) + \delta_X(x)$$

be defined as in (5). Let $\sigma(u) = 1/2u^2, u \in R^1$. Then $M_A = 0$. Indeed, by Lemma 3.1,

$$\bar{l}(x,y,r) = f(x) + yg(x) + \inf_{v\geq 0}\{yv + r/2(g(x)+v)^2\} \quad \forall x \in X, y \in R^1, r > 0. \tag{15}$$

By the definition of M_A, for any $\epsilon > 0$, there exist $\bar{y} \in R^1$ and $\bar{r} > 0$ such that

$$M_A < \bar{l}(x,\bar{y},\bar{r}) + \epsilon \quad \forall x \in X. \tag{16}$$

The combination of (15) and (16) yields

$$M_A < f(x) + \bar{y}(g(x)+v) + \bar{r}/2(g(x)+v)^2 + \epsilon \quad \forall x \in X,\ v \geq 0. \tag{17}$$

Setting $v = 0$ in (17) gives us

$$M_A < f(x) + \bar{y}g(x) + \bar{r}/2g^2(x) + \epsilon \quad \forall x \in X. \tag{18}$$

Note that, for any $x \in (1,+\infty)$, (18) becomes

$$M_A < \frac{1}{x+1} + \left(\frac{1}{\sqrt{x}} - \frac{1}{x}\right)\bar{y} + \bar{r}/2\left(\frac{1}{\sqrt{x}} - \frac{1}{x}\right)^2 + \epsilon. \tag{19}$$

Taking the limit in (19) as $x \to +\infty$, we obtain $M_A \leq \epsilon$. By the arbitrariness of $\epsilon > 0$, we deduce that $M_A \leq 0$. However, it is obvious that $M_A \geq 0$. Hence $M_A = 0$. Consequently, $M_A < M_P$. Thus, Theorem 3.5 is verified.

It is worth noting that the following conditions in Theorems 2.3 and 2.6 are not satisfied:

(i) The condition $\lim_{x\in X,\|x\|\to+\infty} \max\{f(x), g(x)\} \to +\infty$ in Theorem 2.3 does not hold.

(ii) $\overline{f}(x,u)$ is not level-bounded in x and locally uniform in u. In fact, for any sufficiently small $\epsilon > 0$, we cannot find a bounded set $D_0 \subset R^1$ such that $\{x \in X : \overline{f}(x,u) \leq 1\} \subset D_0$ holds for all u satisfying $|u| < \epsilon$.

The following examples show that, if the perturbation function is not defined by (5), then Theorem 3.5 may not hold.

Example 3.2. Consider the same problem as in Example 3.1. Then $M_N < M_P$. But if we let $\varphi(x) = f(x)$, if $x \in X_0$, and $\varphi(x) = +\infty$ otherwise. Define $\overline{f}(x,u) = \varphi(x)$; if $x \in X_0$ and $u = 0$, $\overline{f}(x,u) = +\infty$ otherwise. It is then easy to check that $\overline{f}(x,u)$ is a perturbation function, but is different from (5). On the other hand, the augmented Lagrangian $\bar{l}(x,y,r) = f(x)$ $\forall x \in X_0,\ y \in R^1,\ r > 0$, and $\bar{l}(x,y,r) = +\infty,\ x \notin X_0$. Thus $M_A = M_P$.

Example 3.3. Let $p = q = 1$. Let $X = [0,+\infty),\ f(x) = x,\ x \in X$, and $g(x) = x - 1, x \in X$. Then we have

$$\sigma(u) = |u| \quad \forall u \in R^1,$$

$$\overline{f}(x,u) = \begin{cases} f(x) - u^2 & \text{if } g(x) \leq u, x \in X; \\ +\infty & \text{otherwise.} \end{cases}$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

It is easy to verify that

$$\overline{f}(x,0) = \begin{cases} f(x) & \text{if } x \in X_0 = [0,1]; \\ +\infty & \text{otherwise.} \end{cases}$$

Let us look at the augmented Lagrangian function

$$\bar{l}(x,y,r) = \inf\{f(x) - (v+g_1(x))^2 + r|v+g_1(x)| - y(g_1(x)+v) : v \geq 0\} \equiv -\infty.$$

Thus, (4) does not hold and $M_A = -\infty$. However, $M_P = 0$. It follows that $M_A < M_P$. On the other hand, $M_N = 0$. Hence $M_N = M_P$.

4. A nonlinear Lagrangian method. Let $d \in R^q_+$. Consider the following unconstrained optimization problem (Q_d):

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

where $L(x,d)$ is a nonlinear Lagrangian function. Under certain conditions, we show the existence of a path of optimal solutions generated by unconstrained optimization problems (Q_{d^k}) (where $\{d^k\} \subset R^q_+$ and $d^k \to +\infty$ as $k \to +\infty$) and show its convergence to the optimal set of (P).

Let S denote the optimal solution set of (P), S_d the optimal solution set of (Q_d), and v_d the optimal value of (Q_d).

LEMMA 4.1 (see [12]). *Let $d \in R^q_+$. If the functions defining* (P) *are l.s.c., then $L(\cdot,d)$ is l.s.c. on X.*

THEOREM 4.2. *Consider the problem* (P)*. Let $h(x)$ defined by* (2) *be 0-coercive on X if X is unbounded. Then S is nonempty and compact. For each $d \in R^q_+ + e$, S_d is nonempty and compact. Furthermore, for each selection $x_d \in S_d$ as $d \to +\infty$, $\{x_d\}$ is bounded, its limit points belong to S, and $\lim_{d\to+\infty} v_d = M_P$.*

Proof. Let $\overline{x} \in X_0$. By the 0-coercivity and l.s.c. of h,

$$X_1 = \{x \in X_0 : f(x) \leq f(\overline{x})\} = \{x \in X : h(x) \leq f(\overline{x})\} \cap X_0$$

is nonempty and compact. It follows that S is nonempty. In addition, $S \subset X_1$; therefore, S is bounded. As $S = \bigcap_{x \in X_0}[\{x^* \in X : f(x^*) \leq f(x)\} \bigcap X_0]$ is closed by the lower semicontinuity of f, S is nonempty and compact.

Let $h_1(x) = \max\{f(x), [\min_{1\leq j \leq q} a_j]g(x)\}$. Then

$$L(x,d) \geq \max\{f(x), a_1 d_1 g_1(x), \ldots, a_q d_q g_q(x)\} \geq h_1(x) \quad \forall x \in X,\ d \in R^q_+ + e.$$

It is easy to see that $h_1(x)$ is l.s.c. and 0-coercive. Let $X_2 = \{x \in X : h_1(x) \leq f(\overline{x})\}$. Then X_2 is nonempty and compact. For each $d \in R^q_+ + e$, let $X^d = \{x \in X : L(x,d) \leq L(\overline{x},d)\}$. By Lemma 3.2(I′), we have $X^d = \{x \in X : L(x,d) \leq f(\overline{x})\}$. Moreover, since $L(x,d) \geq h_1(x)$ $\forall x \in X$, it follows that $X^d \subseteq X_2$ is nonempty and compact. Hence, S_d is nonempty and bounded. It follows from Lemma 4.1 that $L(\cdot,d)$ is l.s.c. on X. Thus, S_d is closed. So S_d is nonempty and compact for any $d \in R^q_+ + e$. Moreover,

$$S_d \subseteq X^d \subseteq X_2 \quad \forall d \in R^q_+ + e.$$

It follows that, for each selection $x_d \in S_d$, $\{x_d\}$ is bounded. Suppose that x^* is a limit point of $\{x_d\}$, namely, $\exists d^k = (d^k_1, \ldots, d^k_m) \to +\infty$ and $x_{d^k} \to x^*$ as $k \to +\infty$. Arbitrarily fix an $x \in X_0$. Then we have

$$\max\{f(x_{d^k}), a_1 d^k_1 g_1(x_{d^k}), \ldots, a_q d^k_q g_q(x_{d^k})\} \leq L(x_{d^k}, d^k) \leq L(x, d^k) = f(x). \tag{20}$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Thus,

$$f(x_{d^k}) \le f(x) \tag{21}$$

and

$$\left[\min_{1\le j\le q} a_j\right] \cdot \left[\min_{1\le j\le q} d_j^k\right] \cdot g(x_{d^k}) \le f(x). \tag{22}$$

Equation (22) implies

$$g(x_{d^k}) \le \frac{f(x)}{\left[\min_{1\le j\le q} a_j\right] \cdot \left[\min_{1\le j\le q} d_j^k\right]}.$$

Taking the lower limit and using the lower semicontinuity of g, we have $g(x) \le 0$, i.e., $x \in X_0$. Taking the lower limit in (21) and applying the lower semicontinuity of f, we obtain $f(x^*) \le f(x)$. By the arbitrariness of $x \in X_0$, we conclude that $x^* \in S$.

Furthermore, arbitrarily taking $\{d^k\} \subset R^q_+ + e$ with $d^k \to +\infty$ as $k \to +\infty$, suppose that $x_{d^k} \to x^* \in S$. It follows from (20) (setting $x = x^*$) that $f(x_{d^k}) \le v_{d^k} \le f(x^*)$. Therefore,

$$v = f(x^*) \le \liminf_{k\to+\infty} f(x_{d^k}) \le \liminf_{k\to+\infty} v_{d^k}$$

and $\limsup_{k\to+\infty} v_{d^k} \le f(x^*) = M_P$. Consequently, $\lim_{k\to+\infty} v_{d^k} = M_P$. Thus $\lim_{d\to+\infty} v_d = M_P$. ∎

Remark 4.1. It is clear that if f is 0-coercive on X, then h is also 0-coercive. Theorem 4.2 holds if the 0-coercivity of h is replaced with the 0-coercivity of f.

As a byproduct, we apply Theorem 4.2 to obtain a corollary for the case that (P) is a convex programming problem, which is parallel to [2, Theorem 2.2]. In the following, we assume that f, g_j are finite, l.s.c., and convex functions defined on a nonempty, closed, and convex set $X \subseteq R^p$. Let $F : R^p \to R^1 \bigcup \{+\infty\}$ be an extended real-valued convex function. The *recession function* F^∞ of F is defined by

$$epi(F^\infty) = [epi(F)]^\infty,$$

where $epi(F) = \{(x, r) \in R^p \times R^1 : F(x) \le r\}$ is the epigraph of F. It is known [2] that

$$F^\infty(y) = \inf\left\{\liminf_{k\to+\infty} \frac{F(t_k x_k)}{t_k} : t_k \to +\infty, x_k \to y\right\},$$

where $\{t_k\}$ and $\{x_k\}$ are sequences in R^1 and R^p, respectively.

LEMMA 4.3. *Let f, g_j be finite, l.s.c., and convex functions defined on a nonempty, closed, and convex set X. If the optimal solution set S of* (P) *is nonempty and compact, then $h(x)$ is a finite, l.s.c., convex, and* 0*-coercive function on X.*

Proof. Let us set

$$\hat f(x) = \begin{cases} f(x) & \text{if } x \in X; \\ +\infty & \text{otherwise,} \end{cases}$$

$$\hat g_j(x) = \begin{cases} g_j(x) & \text{if } x \in X; \\ +\infty & \text{otherwise.} \end{cases}$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Then (P) is equivalent to the following convex programming problem (P'):

$$\min\{\hat{f}(x) : x \in C\},$$

where $C = \{x \in R^p : \hat{g}_j(x) \leq 0, j = 1, \ldots, q\}$.

It follows from the assumptions and [2] that S is nonempty and compact if and only if

$$\hat{f}_\infty(w) \leq 0,\ (\hat{g}_j)_\infty(w) \leq 0,\ j = 1, \ldots, q, \quad w \in R^p \Rightarrow w = 0. \tag{23}$$

Since S is nonempty and compact, (23) holds.

Now we show by contradiction that h is 0-coercive. Suppose that there exists $\{x_k\} \subset X$ such that $\|x_k\| \to +\infty$ and $h(x_k) \leq M$ for some $M > 0$. Then $f(x_k) \leq M\ \forall k$ and $g_j(x_k) \leq M\ \forall j, k$. Since $\{\frac{x_k}{\|x_k\|}\}$ is bounded, without loss of generality we assume that $w_k = \frac{x_k}{\|x_k\|} \to w$ as $k \to +\infty$. Clearly, $w \neq 0$ since $\|w\| = 1$. It follows from the definition of a recession function that

$$\hat{f}_\infty(w) \leq \liminf_{k\to+\infty} \frac{f(\|x_k\|w_k)}{\|x_k\|} \leq \lim_{k\to+\infty} \frac{M}{\|x_k\|} = 0, \tag{24}$$

$$(\hat{g}_j)_\infty(w) \leq \liminf_{k\to+\infty} \frac{g_j(\|x_k\|w_k)}{\|x_k\|} \leq \lim_{k\to+\infty} \frac{M}{\|x_k\|} = 0. \tag{25}$$

Thus, $w \neq 0$, and (24) and (25) contradict (23). ∎

Remark 4.2. Let f, g_j, X be as in Lemma 4.3. If X is unbounded, then S is nonempty and compact if and only if h is 0-coercive. This can be regarded as a characterization of the nonemptiness and compactness of the optimal solution set S of the constrained convex programming problem (P).

COROLLARY 4.4. *Let X be a nonempty, closed, and convex subset of R^p. Let f, g_j be finite, l.s.c., and convex functions on X. If S is nonempty and compact, then for each $d \in R^q_+ + e$, S_d is nonempty and compact. Furthermore, for each selection $x_d \in S_d$, $\{x_d\}$ is bounded and its limit points belong to S and* $\lim_{d\to+\infty} v_d = M_P$.

Proof. The proof follows from Theorem 4.2 and Lemma 4.3. ∎

Next we apply Theorem 4.2 to develop a method to seek a so-called ϵ-quasi-solution of (P) when (P) may not have an optimal solution.

Let $\epsilon > 0$. The following various definitions of approximate solutions are cited from [15].

DEFINITION 4.5. $x^* \in X_0$ *is called an ϵ-solution of* (P) *if*

$$f(x^*) \leq f(x) + \epsilon \quad \forall x \in X_0.$$

DEFINITION 4.6. $x^* \in X_0$ *is called an ϵ-quasi-solution of* (P) *if*

$$f(x^*) \leq f(x) + \epsilon\|x - x^*\| \quad \forall x \in X_0.$$

Remark 4.3. An ϵ-quasi-solution is also a local ϵ-solution. In fact, x^* is an ϵ-solution of f on $\{x \in X_0 : \|x - x^*\| \leq 1\}$.

DEFINITION 4.7. *Let $\epsilon > 0$. If $x^* \in X_0$ is both an ϵ-solution and an ϵ-quasi-solution of* (P)*, we say that x^* is a regular ϵ-solution of* (P).

Vavasis [20] gave an algorithm for seeking a local approximate solution via the Ekeland variational principle to a problem that contains only box constraints. Specifically, the following optimization problem (P'') is considered:

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & \alpha_i \leq x_i \leq \beta_i, \quad i = 1, \ldots, p, \end{array}$$

where α_i, β_i, $i = 1, \dots, p$, are real numbers and $x = (x_1, \dots, x_p)$. The algorithm in [20] attempted to find a feasible solution x^*, such that $\| \bigtriangledown f(x^*)\| \le \epsilon$, which is a necessary condition for x^* to be an ϵ-quasi-solution of (P''), where $\epsilon > 0$ is a given precision value.

In the following, we give a model algorithm to find an ϵ-quasi-solution by using a nonlinear Lagrangian. Let $\epsilon > 0$ and $x_0 \in X$. Define

$$f^1(x) = f(x) + \epsilon \|x - x_0\|, \quad x \in X.$$

Consider the following optimization problem (P_ϵ):

$$\begin{array}{ll} \min & f^1(x) \\ \text{s.t.} & x \in X,\ g_j(x) \le 0,\ j = 1, \dots, q, \end{array}$$

and the following unconstrained optimization problem (Q_d^ϵ):

$$\min \overline{L}(x, d) \text{ s.t. } x \in X,$$

where $\overline{L}(x, d) = c(f^1(x), d_1 g_1(x), \dots, d_q g_q(x))\ \forall x \in X,\ d = (d_1, \dots, d_q) \in R_+^q$, and c is defined as in section 2.

Let $\overline{S}_\epsilon$ and $\overline{S}_d^\epsilon$ denote the optimal solution sets of (P_ϵ) and (Q_d^ϵ), respectively. Let $\overline{v}_\epsilon$ and $\overline{v}_d^\epsilon$ denote the optimal values of (P_ϵ) and (Q_d^ϵ), respectively.

THEOREM 4.8. *Let $f(x)$ be 0-coercive on X if X is unbounded. We have the following:*

(i) *$\overline{S}_\epsilon$ is a nonempty and compact set and, for each $d \in R_+^q + e$, $\overline{S}_d^\epsilon$ is a nonempty and compact set.*

(ii) *Let $\overline{x}_d \in \overline{S}_d^\epsilon$, $d \in R_+^q$. Then $\{\overline{x}_d\}$ is bounded, every limit point belongs to $\overline{S}_\epsilon$, and $\lim_{d \to +\infty} \overline{v}_d^\epsilon = \overline{v}_\epsilon$.*

(iii) *Furthermore, any $x^* \in \overline{S}_\epsilon$ is an ϵ-quasi-solution of* (P).

(iv) *If $x_0 \in X_0$, then*

$$f(x^*) \le f(x_0) - \epsilon \|x_0 - x^*\|. \tag{26}$$

Proof. It is clear that f^1 is 0-coercive on X if X is unbounded. Applying Theorem 4.2 by replacing f with f^1, (P) with (P_ϵ), and (Q_d) with (Q_d^ϵ), we conclude that $\overline{S}_\epsilon$ is nonempty and compact; that for each $d \in R_+^q + e$, $\overline{S}_d^\epsilon$ is nonempty and compact; that for each selection $\overline{x}_d \in \overline{S}_d^\epsilon$, $\{\overline{x}_d\}$ is bounded; and that each limit point of $\{\overline{x}_d\}$ belongs to $\overline{S}_\epsilon$ and $\lim_{d \to +\infty} \overline{v}_d^\epsilon = \overline{v}_\epsilon$. Thus (i) and (ii) hold.

Furthermore, for $x^* \in \overline{S}_\epsilon$, we have

$$f(x^*) + \epsilon \|x^* - x_0\| \le f(x) + \epsilon \|x - x_0\| \ \forall x \in X_0. \tag{27}$$

It follows that

$$f(x^*) \le f(x) + \epsilon(\|x - x_0\| - \|x^* - x_0\|) \le f(x) + \epsilon \|x - x^*\| \ \forall x \in X_0.$$

That is, x^* is an ϵ-quasi-solution of (P). Thus, (iii) holds. Moreover, if $x_0 \in X_0$, then by (27) (taking $x = x_0$), we get (26). The proof is complete. □

Remark 4.4. The last assertion (26) tells us that even if we already obtained an ϵ-quasi-solution x_0 of (P), it is still possible to apply Theorem 4.8 to seek a "better" ϵ-quasi-solution x^* of (P) (if the resulting $x^* \neq x_0$).

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

5. Convergence analysis of the nonlinear Lagrangian method in terms of necessary optimality conditions. In this section, we investigate the convergence of first and second order necessary optimality conditions that are obtained from nonlinear Lagrangian problems. Specifically, we shall consider the following classes of nonlinear Lagrangians:

(i) $L^\infty(x,d) = \max\{f(x), d_1g_1(x), \ldots, d_qg_q(x)\}$, $x \in X$;

(ii) $L^k(x,d) = (f(x)^k + \sum_{j=1}^q d_j^k {g_j^+(x)}^k)^{1/k}$, $x \in X$, where $2 \le k < \infty$;

(iii) $L^k(x,d)$ is as in (ii) with $0 < k < 2$,

where properties (A) and (B) are satisfied with $a_j = 1, j = 1,\ldots,q$.

Throughout this section, we further assume

(A1) $X = R^p$;

(A2) $\beta = \inf_{x\in R^p} f(x) > 0$;

(A3) $f,\ g_j,\ j = 1,\ldots,q$, are $C^{1,1}$, namely, they are differentiable and their gradients are locally Lipschitz; and

(A4) $\max\{f(x), g_1(x), \ldots, g_q(x)\} \to +\infty$ as $\|x\| \to +\infty$.

Let f be a $C^{1,1}$ function. We denote by $\partial^2 f(x)$ the generalized Hessian of f at x; see [11, 23]. It is noted that the set-valued mapping $x \to \partial^2 f(x)$ is upper semicontinuous.

We consider the following type of optimality conditions which were derived in [11, 21]. It is worth noting that in these conditions the multipliers do not depend on the choice of vectors in the tangential subspace of the active constraints.

DEFINITION 5.1. *Let $x^* \in X_0$. The first order necessary condition of* (P) *is said to hold at x^* if there exist $\lambda,\ \mu_j \ge 0,\ j \in J(x^*)$, such that*

$$\lambda \bigtriangledown f(x^*) + \sum_{j\in J(x^*)} \mu_j \bigtriangledown g_j(x^*) = 0. \tag{28}$$

The second order necessary condition of (P) *is said to hold at x^* if* (28) *holds and, for any $u^* \in R^p$ satisfying*

$$\bigtriangledown g_j(x^*)^\top u^* = 0, \quad j \in J(x^*), \tag{29}$$

there exist $F \in \partial^2 f(x^*),\ G_j \in \partial^2 g_j(x^*),\ j \in J(x^*)$, such that

$$u^{*T}\left(\lambda F + \sum_{j\in J(x^*)} \mu_j G_j\right) u^* \ge 0. \tag{30}$$

We need the following lemma.

LEMMA 5.2. *Let $k \in (0, +\infty],\ z \in X_0$, and $d_n = (d_{1,n}, \ldots, d_{q,n}) (\in R^q_+) \to +\infty$ as $n \to +\infty$. If the sequence $\{x_n\} \subset X$ satisfies $L^k(x_n, d_n) \le f(z)\ \forall n$, then $\{x_n\}$ is bounded and its limit points belong to X_0.*

Proof. It is known that $\max\{f(x_n), d_{1,n}g_1(x_n), \ldots, d_{q,n}g_q(x_n)\} \le L^k(x_n, d_n)$. Thus,

$$\max\{f(x_n), d_{1,n}g_1(x_n), \ldots, d_{q,n}g_q(x_n)\} \le f(z). \tag{31}$$

Suppose that $\{x_n\}$ is unbounded. Without loss of generality, assume that $\|x_n\| \to +\infty$. By assumption (A4), we get

$$\max\{f(x_n), g_1(x_n), \ldots, g_q(x_n)\} \to +\infty \text{ as } n \to +\infty. \tag{32}$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Since $d_{j,n} \to +\infty$ as $n \to +\infty$ $(j = 1, \ldots, q)$, we see that $d_{j,n} > 1$ $(j = 1, \ldots, q)$ when n is sufficiently large. Hence, for sufficiently large n,

$$\max\{f(x_n), g_1(x_n), \ldots, g_q(x_n)\} \leq \max\{f(x_n), d_{1,n}g_1(x_n), \ldots, d_{q,n}g_q(x_n)\}.$$

This fact, combined with (32), contradicts (31). So the sequence $\{x_n\}$ is bounded.

Now we show that any limit point of $\{x_n\}$ belongs to X_0. Without loss of generality, we assume that $x_n \to x^*$. Suppose that $x^* \notin X_0$. There exists $\gamma_0 > 0$ such that $\max\{g_1(x^*), \ldots, g_q(x^*)\} \geq \gamma_0 > 0$. It follows that $\max\{g_1(x_n), \ldots, g_q(x_n)\} \geq \gamma_0/2$ for sufficiently large n. Moreover, it follows from (31) that

$$\begin{aligned} f(z) &\geq L^k(x_n, d_n) \geq \max\{d_{1,n}g_1(x_n), \ldots, d_{q,n}g_q(x_n)\} \\ &\geq \min_{1\leq j\leq q}\{d_{j,n}\}\max\{g_1(x_n), \ldots, g_q(x_n)\} \geq \frac{\gamma_0}{2}\min_{1\leq j\leq q}\{d_{j,n}\}, \end{aligned}$$

which is impossible, as $n \to +\infty$. □

Define

$$J^*(\overline{x}) = \begin{cases} J^+(\overline{x}) \cup J(\overline{x}) & \text{if } k \in (0,2), \\ J(\overline{x}) & \text{if } k \in [2,\infty), \\ J^+(\overline{x}) & \text{if } k = \infty. \end{cases}$$

Lemma 5.3 (see [22]). *Suppose that $\{\nabla g_j(x)\}_{j\in J^*(x)}$ is linearly independent for any $x \in X_0$ and that $\overline{x}_n \to x^*$ as $n \to +\infty$ and $x^* \in X_0$. Then, for $u^* \in R^p$ satisfying* (29), *there exists a sequence $\{u_n\} \subset R^p$ such that $\nabla g_j(\overline{x}_n)^\top u_n = 0,\ j \in J^*(x^*)$, and $u_n \to u^*$.*

As shown in [1, 22], if $x \in X_0$ and $x_n \to x$, then, for sufficiently large n,

$$J(x_n) \subseteq J(x), \quad J^+(x_n) \subseteq J(x). \tag{33}$$

We shall carry out the convergence analysis by considering the following two cases.

Case 1. $2 \leq k < +\infty$.

Case 2. $k = +\infty$ or $k \in (0,2)$.

5.1. Case 1. $2 \leq k < +\infty$. When $2 \leq k < +\infty$, the nonlinear Lagrangian function $L^k(x,d)$ is $C^{1,1}$. Thus, the first and second order necessary optimality conditions of (Q_{d_n}) can be easily derived.

Let $d_n = (d_{1,n}, \ldots, d_{q,n})(\in R^q_+) \to +\infty$ as $n \to +\infty$.

Let $\overline{x}_n$ be a local minimum of (Q_{d_n}). Thus, the first order necessary condition for $\overline{x}_n$ to be a local minimum of (Q_{d_n}) can be written as $\nabla L^k(\overline{x}_n, d_n) = 0$, or

$$a_n^{\frac{1}{k}-1}\left[f^{k-1}(\overline{x}_n) \bigtriangledown f(\overline{x}_n) + \sum_{j\in J^+(\overline{x}_n)} d^k_{j,n}(g^+_j(\overline{x}_n))^{k-1} \bigtriangledown g_j(\overline{x}_n)\right] = 0, \tag{34}$$

where $a_n = [L^k(\overline{x}_n, d_n)]^k$.

The second order necessary condition is that, for every $u \in R^p$, $u^\top Mu \geq 0$ for some $M \in \partial^2 L^k(\overline{x}_n, d_n)$; thus there exist $F_n \in \partial^2 f(\overline{x}_n), G_{j,n} \in \partial^2 g_j(\overline{x}_n), j \in J^+(\overline{x}_n)$, such that

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

$$
\begin{aligned}
&\left(\frac{1}{k}-1\right) a_n^{\frac{1}{k}-2}\Bigg[\alpha(n)(\bigtriangledown f(\overline{x}_n)^\top u)^2 + \sum_{j\in J^+(\overline{x}_n)} \beta_{j,1}(n)(\bigtriangledown g_j(\overline{x}_n)^\top u)^2 \\
&\qquad + \sum_{j\in J^+(\overline{x}_n)} \beta_{j,2}(n)(\bigtriangledown f(\overline{x}_n)^\top u)(\bigtriangledown g_j(\overline{x}_n)^\top u) \\
&\qquad + \sum_{i\in J^+(\overline{x}_n)} \sum_{j\in J^+(\overline{x}_n)} \beta_{j,3}(n)(\bigtriangledown g_i(\overline{x}_n)^\top u)(\bigtriangledown g_j(\overline{x}_n)^\top u)\Bigg] \\
&+ a_n^{\frac{1}{k}-1}(k-1)\Bigg[\xi(n)(\bigtriangledown f(\overline{x}_n)^\top u)^2 + \sum_{j\in J(\overline{x}_n)} \eta_{j,1}(n)[(\bigtriangledown g_j(\overline{x}_n)^\top u)^+]^2 \\
&\qquad + \sum_{j\in J^+(\overline{x}_n)} \eta_{j,2}(n)(\bigtriangledown g_j(\overline{x}_n)^\top u)^2\Bigg] \\
&+ a_n^{\frac{1}{k}-1} u^T \left[f^{k-1}(\overline{x}_n)F_n + \sum_{j\in J^+(\overline{x}_n)} d_{j,n}^k (g_j^+(\overline{x}_n))^{k-1} G_{j,n}\right] u \geq 0, \tag{35}
\end{aligned}
$$

where $\alpha(n),\ \beta_{i,1}(n),\ \beta_{i,2}(n),\ \beta_{i,3}(n),\ \xi(n),\ \eta_{i,1}(n)$, and $\eta_{i,2}(n)$ are real numbers.

We have the following convergence result.

THEOREM 5.4. *Suppose that* $\{\nabla g_j(x)\}_{j\in J(x)}$ *is linearly independent for any* $x \in X_0$. *Let* $2 \leq k < +\infty$ *and* $d_n \in R_+^q$ *be such that* $d_n \to +\infty$. *Let* $\overline{x}_n$ *be generated by some descent method for* (Q_{d_n}) *starting from a point* $z \in X_0$ *and* $\overline{x}_n$ *satisfy first order necessary condition* (34) *and second order necessary condition* (35). *Then* $\{\overline{x}_n\}$ *is bounded and every limit point of* $\{\overline{x}_n\}$ *is a point of* X_0 *satisfying first order necessary optimality condition* (28) *and second order necessary optimality condition* (30) *of* (P).

Proof. It follows from Lemma 5.2 that $\{\overline{x}_n\}$ is bounded and every limit point of $\{\overline{x}_n\}$ belongs to X_0. Without loss of generality, we assume that $\overline{x}_n \to x^*$. Let

$$a_n = [L^k(\overline{x}_n, d_n)]^k > 0; \quad b_n = a_n^{\frac{1}{k}-1}\left(f^{k-1}(\overline{x}_n) + \sum_{j\in J^+(\overline{x}_n)} d_{j,n}^k g_j^+(\overline{x}_n)^{k-1}\right) > 0.$$

Thus,

$$\frac{a_n^{\frac{1}{k}-1} f^{k-1}(\overline{x}_n)}{b_n} + \sum_{j\in J^+(\overline{x}_n)} \frac{a_n^{\frac{1}{k}-1} d_{j,n}^k (g_j^+(\overline{x}_n))^{k-1}}{b_n} = 1.$$

Without loss of generality, we assume that

$$\frac{a_n^{\frac{1}{k}-1} f^{k-1}(\overline{x}_n)}{b_n} \to \lambda, \tag{36}$$

$$\frac{a_n^{\frac{1}{k}-1} d_{j,n}^k (g_j^+(\overline{x}_n))^{k-1}}{b_n} \to \mu_j, \quad j \in J(x^*). \tag{37}$$

Then by (33),

$$\lambda \geq 0,\ \mu_j \geq 0,\ j \in J(x^*),\ \text{and}\ \lambda + \sum_{j\in J(x^*)} \mu_j = 1. \tag{38}$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Dividing (34) by b_n and taking the limit, we obtain

$$\lambda \bigtriangledown f(x^*) + \sum_{j\in J(x^*)} \mu_j \bigtriangledown g_j(x^*) = 0.$$

Since $\{\bigtriangledown g_j(x^*)\}_{j\in J(x^*)}$ is linearly independent, it follows that $\lambda > 0$.

By Lemma 5.3, we deduce that, for any $u^* \in R^p$ satisfying (29), we can find $u_n \in R^p$ such that

$$\bigtriangledown g_j(\overline{x}_n)^\top u_n = 0,\ j \in J(x^*) \tag{39}$$

and

$$u_n \to u^*. \tag{40}$$

Furthermore, for every u_n satisfying (39) and (40), we can find $F_n \in \partial^2 f(\overline{x}_n), G_{j,n} \in \partial^2 g_j(\overline{x}_n), j \in J^+(\overline{x}_n)$, such that (35) holds with u replaced by u_n.

Substituting (39) into (34), we get

$$\bigtriangledown f(\overline{x}_n)^\top u_n = 0. \tag{41}$$

Substituting (39)–(41) into (35), we have

$$a_n^{\frac{1}{k}-1} u_n^\top \left(f^{k-1}(\overline{x}_n)F_n + \sum_{j\in J^+(\overline{x}_n)} d_{j,n}^k (g_j^+(\overline{x}_n))^{k-1} G_{j,n} \right) u_n \geq 0. \tag{42}$$

Since $\overline{x}_n \to x^*$ as $n \to \infty$, $\partial^2 f(\cdot), \partial^2 g_j(\cdot)$ are upper semicontinuous at x^* and $\partial^2 f(x^*), \partial^2 g_j(x^*)$ are compact, without loss of generality we can assume that

$$F_n \to F \in \partial^2 f(x^*),\ G_{j,n} \to G_j \in \partial^2 g_j(x^*),\ j \in J(x^*). \tag{43}$$

Dividing (42) by b_n and taking the limit, applying (36), (37), (40), and (43), we obtain

$$u^{*T} \left(\lambda F + \sum_{j\in J(x^*)} \mu_j G_j \right) u^* \geq 0 \ \text{ and } \lambda > 0.$$

5.2. Case 2. $k = +\infty$ or $k \in (0, 2)$. When $k = +\infty$, problem (Q_{d_n}) is a minimax optimization problem and thus a convex composite optimization problem. However, the second order necessary conditions for a convex composite optimization problem given in [4, 23] are not applicable, as the multipliers depend on the choice of the vector in the tangential subspace of the active constraints. When $k \in (0, 2)$, function $g_j^+(x)^k$ and thus $L^k(x, d)$ is not $C^{1,1}$. Thus, the existing optimality conditions in the literature are not applicable. However, we are able to derive optimality conditions for (Q_{d_n}) by applying the smooth approximate variational principle, which is due to Borwein and Preiss [6] (see also [8, Theorem 5.2]).

LEMMA 5.5 (approximate smooth variational principle [8, Theorem 5.2]). *Let X be a Hilbert space. Let $g : X \to (-\infty, +\infty]$ be l.s.c. and bounded below with $dom(g) \neq \emptyset$. Let $\overline{x}$ be a point such that $g(\overline{x}) < \inf_{x\in X} g(x) + \epsilon$, where $\epsilon > 0$. Then, for any $\lambda > 0$, there exist y_ϵ, z_ϵ with $\|y_\epsilon - z_\epsilon\| < \lambda, \|z_\epsilon - \overline{x}\| < \lambda$, $g(y_\epsilon) < \inf_{x\in X} g(x) + \epsilon$, and having the property that the function $y \to g(y) + (\epsilon/\lambda^2)\|y - z_\epsilon\|^2$ has a unique minimum over X at $y = y_\epsilon$.*

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Remark 5.1. If the Hilbert space X in Lemma 5.5 is replaced with a nonempty and closed subset X_1, then the conclusion also holds. As a matter of fact, if $g : X_1 \to (-\infty, +\infty]$ is l.s.c. and bounded below on X_1, we can define a function $\overline{g} : X \to (-\infty, +\infty]$ as follows: $\overline{g}(x) = g(x)$ if $x \in X_1$ and $\overline{g}(x) = +\infty$ otherwise. It is easy to verify that $\overline{g}$ is l.s.c. and bounded below on X. Applying Lemma 5.3 to $\overline{g}$, the conclusion for g follows.

Next we present first and second order necessary conditions for $\overline{x}$ to be a local minimum of $L^k(x, d)$ under the linear independence assumption. The proof is given in the appendix.

PROPOSITION 5.6. *Let $k \in (0, 2)$ or $k = +\infty$. Let $\overline{x}$ be a local minimum of $L^k(x, d)$ and $\{\bigtriangledown g_j(\overline{x})\}_{j \in J^*(\overline{x})}$ be linearly independent. Then there exist $\lambda > 0,\ \mu_j \geq 0,\ j \in J^*(\overline{x})$, with $\lambda + \sum_{j \in J^*(\overline{x})} \mu_j = 1$ such that*

$$\lambda \bigtriangledown f(\overline{x}) + \sum_{j \in J^*(\overline{x})} \mu_j \bigtriangledown g_j(\overline{x}) = 0.$$

Furthermore, for each $u \in R^p$ satisfying

$$\bigtriangledown g_j(\overline{x})^\top u = 0, \quad j \in J^*(\overline{x}), \tag{44}$$

there exist $F \in \partial^2 f(\overline{x}),\ G_j \in \partial^2 g_j(\overline{x}),\ j \in J^*(\overline{x})$, such that

$$u^T \left(\lambda F + \sum_{j \in J^*(\overline{x})} \mu_j G_j \right) u \geq 0.$$

THEOREM 5.7. *Suppose that $\{\nabla g_j(x)\}_{j \in J^*(x)}$ is linearly independent for any $x \in X_0$. Let $k \in (0, 2)$ or $k = +\infty$. Let $d_n (\in R^q_+) \to +\infty$ as $n \to +\infty$. Let $\overline{x}_n$ be generated by some descent method for (Q_{d_n}) starting from a point $z \in X_0$. Then $\{\overline{x}_n\}$ is bounded and every limit point of $\{\overline{x}_n\}$ is a point of X_0 satisfying first order necessary condition* (28) *and second order necessary condition* (30) *of* (P)*, respectively.*

Proof. It follows from Lemma 5.2 that $\{\overline{x}_n\}$ is bounded and every limit point of $\{\overline{x}_n\}$ belongs to X_0. Without loss of generality, suppose that $\overline{x}_n \to x^* \in X_0$ and that $J^+(\overline{x}_n) \cup J(\overline{x}_n) \subset J(x^*)$ for sufficiently large n. That $\{\bigtriangledown g_j(x^*)\}_{j \in J(x^*)}$ is linearly independent implies that $\{\bigtriangledown g_j(\overline{x}_n)\}_{j \in J^+(\overline{x}_n) \cup J(\overline{x}_n)}$ is linearly independent when n is sufficiently large. In other words, the assumptions in Proposition 5.6 hold (with $\overline{x}$ replaced by $\overline{x}_n$) when n is sufficiently large. Thus, we assume that $\{\bigtriangledown g_j(\overline{x}_n)\}_{j \in J^+(\overline{x}_n) \cup J(\overline{x}_n)}$ is linearly independent for all n.

The first order necessary optimality conditions in Proposition 5.6 can be written as

$$\lambda_n \bigtriangledown f(\overline{x}_n) + \sum_{j \in J(x^*)} \mu_{j,n} \bigtriangledown g_j(\overline{x}_n) = 0, \tag{45}$$

where $\lambda_n > 0, \mu_{j,n} \geq 0, j \in J(x^*)$, with $\mu_{j,n} = 0\ \forall j \in J(x^*) \backslash J(\overline{x}_n)$ and $\lambda_n + \sum_{j \in J(x^*)} \mu_j = 1$. Without loss of generality, we assume that $\lambda_n \to \lambda, \mu_{j,n} \to \mu_j, j \in J(x^*)$, as $n \to +\infty$. Taking the limit in (45) gives us

$$\lambda \bigtriangledown f(x^*) + \sum_{j \in J(x^*)} \mu_j \bigtriangledown g_j(x^*) = 0.$$

By the linear independence of $\{\bigtriangledown g_j(x^*)\}_{j \in J(x^*)}$, we see that $\lambda > 0$. That is, (28) holds.

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Let $u^* \in R^p$ satisfy (29). Since $\{\bigtriangledown g_j(x^*)\}_{j\in J(x^*)}$ is linearly independent and $\overline{x}_n \to x^*$, by Lemma 5.3, we obtain $\overline{u}_n \in R^p$ such that

$$\bigtriangledown g_j(\overline{x}_n)^T \overline{u}_n = 0, \qquad j \in J(x^*), \tag{46}$$

and $\overline{u}_n \to u^*$.

Thus, if $\overline{x}_n$ satisfies any one of the second order necessary conditions in Proposition 5.6, then, for every $\overline{u}_n$ satisfying (46), there exist $F_n \in \partial^2 f(\overline{x}_n), G_{j,n} \in \partial^2 g_j(\overline{x}_n), j \in J(x^*)$,

$$\overline{u}_n{}^T \left(\lambda_n F_n + \sum_{j\in J(x^*)} \mu_{j,n} G_{j,n} \right) \overline{u}_n \geq 0, \tag{47}$$

where $\lambda_n, \mu_{j,n}$ are as in (45).

By the upper semicontinuity of $\partial^2 f(\cdot), \partial^2 g_j(\cdot)$ and the nonemptiness and compactness of $\partial^2 f(x^*), \partial^2 g_j(x^*)(j = 1, \ldots, q)$, without loss of generality we assume that

$$F_n \to F \in \partial^2 f(x^*),\ G_{j,n} \to G_j \in \partial^2 g_j(x^*),\ j \in J(x^*),$$

as $n \to +\infty$. Taking the limit in (47), we get

$$u^{*T} \left(\lambda F + \sum_{j\in J(x^*)} \mu_j G_j \right) u^* \geq 0,$$

where $\lambda > 0$. Thus, (30) follows. The proof is complete. $\square$

Appendix. Proof of Proposition 5.6. We consider the following two cases.

Case 1. $k = \infty$. In this case, $J^*(\overline{x}) = J^+(\overline{x})$. Since $\overline{x} \in X, f(\overline{x}) > 0$. Thus, it follows that $L^\infty(\overline{x}, d) = \max\{f(\overline{x}), d_j g_j(\overline{x})\}_{j\in J^+(\overline{x})}$. Since $\overline{x}$ is a local minimum of $L^\infty(x, d)$, there exists $\delta > 0$ such that

$$L^\infty(\overline{x}, d) \leq L^\infty(x, d) = \max\{f(x), d_j g_j(x)\}_{j\in J^+(\overline{x})} \quad \forall x \in U_\delta,$$

where $U_\delta = \{x \in R^p : \|x - \overline{x}\| \leq \delta\}$ $(X = R^p)$.

Let $m > 0$ be an integer and

$$s_m(x) = \left[f^m(x) + \sum_{j\in J^+(\overline{x})} d_j^m g_j^m(x) \right]^{\frac{1}{m}},\ x \in U_\delta,$$
$$\epsilon_m = \left[(q+1)^{\frac{1}{m}} - 1 \right] L^\infty(\overline{x}, d).$$

Then $0 \leq s_m(x) - L^\infty(x, d)$ $\forall x \in U_\delta$ and $s_m(\overline{x}) \leq [(q+1)^{\frac{1}{m}}]L^\infty(\overline{x}, d)$. Thus,

$$\begin{aligned} s_m(\overline{x}) &\leq L^\infty(\overline{x}, d) + [(q+1)^{\frac{1}{m}} - 1]L^\infty(\overline{x}, d) \\ &\leq L^\infty(x, d) + [(q+1)^{\frac{1}{m}} - 1]L^\infty(\overline{x}, d) \\ &\leq s_m(x) + [(q+1)^{\frac{1}{m}} - 1]L^\infty(\overline{x}, d) \\ &= s_m(x) + \epsilon_m \quad \forall x \in U_\delta. \end{aligned}$$

Note that $\epsilon_m \downarrow 0$ as $m \to +\infty$. Without loss of generality, we assume that $2\epsilon_m^{1/4} < \delta\ \forall m$. Applying Lemma 5.5 by setting $\lambda = \epsilon_m^{1/4}$, we obtain $\overline{x}'_m, \overline{x}''_m \in U_\delta$ such that

$$\|\overline{x}'_m - \overline{x}''_m\| < \epsilon_m^{1/4} \quad \text{and} \quad \|\overline{x}''_m - \overline{x}\| < \epsilon_m^{1/4}$$

and $\overline{x}'_m$ is a unique minimum of the problem

$$\min\ v_m(x) = s_m(x) + \epsilon_m^{1/2}\|x - \overline{x}''_m\|^2 \quad \text{s.t. } x \in U_\delta. \tag{48}$$

Note that $\|\overline{x}'_m - \overline{x}\| \le \|\overline{x}'_m - \overline{x}''_m\| + \|\overline{x}''_m - \overline{x}\| \le 2\epsilon_m^{1/4} < \delta$. It follows that $\overline{x}'_m \in \mathrm{int}U_\delta$. Applying the first order necessary optimality condition to problem (48), we get $\triangledown v_m(\overline{x}'_m) = 0$. That is,

(49)

$$a_m^{\frac{1}{m}-1}\left[f^{m-1}(\overline{x}'_m) \triangledown f(\overline{x}'_m) + \sum_{j\in J^+(\overline{x})} d_j^m g_j^{m-1}(\overline{x}'_m) \triangledown g_j(\overline{x}'_m)\right] + 2\epsilon_m^{1/2}(\overline{x}'_m - \overline{x}''_m) = 0,$$

where $a_m = [s_m(\overline{x}'_m)]^m$.

Let

$$b_m = a_m^{\frac{1}{m}-1}\left[f^{m-1}(\overline{x}'_m) + \sum_{j\in J^+(\overline{x})} d_j^m g_j^{m-1}(\overline{x}'_m)\right].$$

It is clear that there exists $\alpha > 0$ such that $b_m \ge \alpha > 0\ \forall m$. Without loss of generality, we can assume that

$$\frac{a_m^{\frac{1}{m}-1} f^{m-1}(\overline{x}'_m)}{b_m} \to \lambda,\ \frac{a_m^{\frac{1}{m}-1} d_j^m g_j^{m-1}(\overline{x}'_m)}{b_m} \to \mu_j,\ j \in J^+(\overline{x}). \tag{50}$$

Thus

$$\lambda \ge 0,\ \mu_j \ge 0,\ j \in J^+(\overline{x}), \quad \text{and } \lambda + \sum_{j\in J^+(\overline{x})} \mu_j = 1.$$

Dividing (50) by b_m and taking the limit as $m \to +\infty$, it follows from (50) that

$$\lambda \triangledown f(\overline{x}) + \sum_{j\in J^+(\overline{x})} \mu_j \triangledown g_j(\overline{x}) = 0.$$

Since $\{\triangledown g_j(\overline{x})\}_{j\in J^+(\overline{x})}$ is linearly independent, it follows that $\lambda > 0$.

Now we apply the second order necessary optimality condition to (48). For any $u \in R^p$, there exists $V_m \in \partial^2 v_m(\overline{x}'_m)$ such that $u^\top V_m u \ge 0$. That is, there exist $F_m \in \partial^2 f(\overline{x}'_m)$ and $G_{j,m} \in \partial^2 g_j(\overline{x}'_m),\ j \in J^+(\overline{x})$, such that

$$\begin{aligned}&\left(\frac{1}{m} - 1\right) a_m^{\frac{1}{m}-2}\left(f^{m-1}(\overline{x}'_m) \triangledown f(\overline{x}'_m)^\top u + \sum_{j\in J^+(\overline{x})} d_j^m g_j^{m-1}(\overline{x}'_m) \triangledown g_j(\overline{x}'_m)^\top u\right)^2 \\ &+(m-1)a_m^{\frac{1}{m}-1}\left(f^{m-2}(\overline{x}'_m)(\triangledown f(\overline{x}'_m)^\top u)^2 + \sum_{j\in J^+(\overline{x})} d_j^m g_j^{m-2}(\overline{x}'_m)(\triangledown g_j(\overline{x}'_m)^\top u)^2\right) \\ &+a_m^{\frac{1}{m}-1} u^\top\left(f^{m-1}(\overline{x}'_m)F_m + \sum_{j\in J^+(\overline{x})} d_j^m (g_j^+(\overline{x}'_m))^{m-1} G_{j,m}\right) u + 2\epsilon_m^{1/2} u^T u \ge 0.\end{aligned}$$

(51)

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Since $\{\nabla g_j(\overline{x})\}_{j\in J^+(\overline{x})}$ is linearly independent and $\overline{x}'_m \to \overline{x}$, from Lemma 5.3, for any $\overline{u} \in R^p$ satisfying (44), there exists a sequence $\{u_m\}$, such that

$$\nabla g_j(\overline{x}'_m)^\top u_m = 0, \quad j \in J^+(\overline{x}), \tag{52}$$

and $u_m \to \overline{u}$.

The combination of (51) (setting $u = u_m$) and (52) yields

$$\left(\frac{1}{m}-1\right) a_m^{\frac{1}{m}-2} \left(f^{m-1}(\overline{x}'_m) \nabla f(\overline{x}'_m)^\top u_m\right)^2 + (m-1)a_m^{\frac{1}{m}-1} f^{m-2}(\overline{x}'_m)(\nabla f(\overline{x}'_m)^\top u_m)^2$$
$$+a_m^{\frac{1}{m}-1} u_m^T \left[f^{m-1}(\overline{x}'_m)F_m + \sum_{j\in J^+(\overline{x})} d_j^m g_j^{m-1}(\overline{x}'_m) G_{j,m}\right] u_m + 2\epsilon_m^{1/2} u_m^T u_m \geq 0. \tag{53}$$

From (50) (setting $u = u_m$) and (52), we have

$$\begin{aligned}&\left|\left(\frac{1}{m}-1\right) a_m^{\frac{1}{m}-2} \left(f^{m-1}(\overline{x}'_m) \nabla f(\overline{x}'_m)^\top u_m\right)^2 / b_m\right| \\ &\quad = 4\epsilon_m [(\overline{x}'_m - \overline{x}''_m)^\top u_m]^2 \left(1-\frac{1}{m}\right) / (a_m^{1/m} b_m) \leq \frac{4\epsilon_m^{\frac{3}{2}}}{(\alpha\beta)} \|u_m\|^2.\end{aligned}$$

Therefore,

$$\left(\frac{1}{m}-1\right) a_m^{1/m-2} \left(f^{m-1}(\overline{x}'_m) \nabla f(\overline{x}'_m)^\top u_m\right)^2 / b_m \to 0 \quad \text{ as } m \to \infty.$$

The first formula in (50) guarantees that, when m is sufficiently large,

$$a_m^{\frac{1}{m}-1} f^{m-1}(\overline{x}'_m)/b_m > \lambda/2 > 0.$$

Thus, the combination of (50) (letting $u = u_m$) and (52) also yields

$$\begin{aligned}&(m-1)a_m^{\frac{1}{m}-1} f^{m-2}(\overline{x}'_m) \left(\nabla f(\overline{x}'_m)^\top u_m\right)^2 / b_m \\ &\quad = \frac{1}{f(\overline{x}'_m)}(m-1)4\epsilon_m[(\overline{x}'_m - \overline{x}''_m)^\top u_m]^2 / [(a_m^{\frac{1}{m}-1} f^{m-1}(\overline{x}'_m)/b_m) b_m^2] \\ &\quad \leq \frac{1}{\beta\alpha^2}\|u_m\|^2 4(m-1)\epsilon_m^{3/2}/(\lambda/2).\end{aligned}$$

Noting that

$$4(m-1)\epsilon_m^{3/2} \leq 4(m-1)\left((q+1)^{1/m}-1\right)^{3/2} [L^\infty(\overline{x}, d)]^{3/2},$$

we deduce that

$$(m-1)a_m^{\frac{1}{m}-1} f^{m-2}(\overline{x}'_m) \left(\nabla f(\overline{x}'_m)^\top u_m\right)^2 / b_m \to 0 \quad \text{ as } m \to \infty.$$

Since $\partial^2 f(\cdot)$, $\partial^2 g_j(\cdot)$ are upper semicontinuous at $\overline{x}$ and $\partial^2 f(\overline{x}), \partial^2 g_j(\overline{x})$ are nonempty and compact, we obtain $F \in \partial^2 f(\overline{x})$, $G_j \in \partial^2 g_j(\overline{x})$, $j \in J^+(\overline{x})$, such that

$$F_m \to F,\ G_m \to G,\ j \in J^+(\overline{x}) \ \text{ as } m \to \infty.$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Thus, dividing (53) by b_m and taking the limit, we have

$$\overline{u}^T\left(\lambda F+\sum_{j\in J^+(\overline{x})}\mu_j G_j\right)\overline{u}\geq 0 \quad \text{ and } \quad \lambda>0.$$

Case 2. $k\in(0,2)$. In this case, $J^*(\overline{x})=J^+(\overline{x})\bigcup J(\overline{x})$. Since $\overline{x}$ is a local minimum of $L^k(x,d)$, there exists $\delta>0$ such that $L^k(\overline{x},d)\leq L^k(x,d)$ $\forall x\in U_\delta$. Then

$$\left(f^k(\overline{x})+\sum_{j\in J^+(\overline{x})\cup J(\overline{x})}d_j^k {g_j^+}^k(\overline{x})\right)^{1/k}\leq\left(f^k(x)+\sum_{j\in J^+(\overline{x})\cup J(\overline{x})}d_j^k {g_j^+}^k(x)\right)^{1/k}.$$

Let

$$t_m(x)=\left(f^k(x)+\frac{1}{2^k}\sum_{j\in J^+(\overline{x})\cup J(\overline{x})}\left(d_jg_j(x)+\sqrt{d_j^2g_j^2(x)+1/m}\right)^k\right)^{1/k}.$$

It is not hard to prove that $0\leq t_m(\overline{x})-L^k(\overline{x},d)\leq\epsilon_m$ and $L^k(x,d)\leq t_m(x)$ $\forall x\in U_\delta$, where

$$\epsilon_m=\begin{cases}\frac{q}{k}L^k(\overline{x},d)^{\frac{1}{k}-1}\frac{1}{m^{k/2}} & \text{if } k\in(0,1];\\ \frac{1}{2\sqrt{m}}q^{1/k} & \text{if } k\in(1,2).\end{cases}$$

Thus,

$$t_m(\overline{x})\leq L^k(\overline{x},d)+\epsilon_m\leq L^k(x,d)+\epsilon_m\leq t_m(x)+\epsilon_m \quad \forall x\in U_\delta.$$

Since $\epsilon_m\downarrow 0$ as $m\to+\infty$, without loss of generality we assume that $2\epsilon_m^{1/4}<\delta$ $\forall m$. Applying Lemma 5.5 by setting $\lambda=\epsilon_m^{1/4}$, there exist $\overline{x}'_m,\overline{x}''_m\in U_m$ with $\|\overline{x}'_m-\overline{x}''_m\|<\epsilon_m^{1/4}$, and $\|\overline{x}''_m-\overline{x}\|<\epsilon_m^{1/4}$, such that $\overline{x}'_m$ is the unique minimum of the optimization problem

$$\min\ w_m(x)=t_m(x)+\epsilon_m^{1/2}\|x-\overline{x}''_m\|^2 \quad \text{s.t. } x\in U_\delta. \tag{54}$$

Applying the first order necessary optimality condition to $w_m(x)$ and noticing that $\overline{x}'_m\in\text{int}U_\delta$, we have $\bigtriangledown w_m(\overline{x}'_m)=0$. That is,

$$\begin{aligned}&a_m^{\frac{1}{k}-1}\Bigg(f^{k-1}(\overline{x}'_m)\bigtriangledown f(\overline{x}'_m)\\ &\qquad+\frac{1}{2^k}\sum_{j\in J^+(\overline{x})\cup J(\overline{x})}d_jc_m^{k-1}(1+d_jg_j(\overline{x}'_m)(d_j^2g_j^2(\overline{x}'_m)+1/m)^{-1/2})\bigtriangledown g_j(\overline{x}'_m)\Bigg)\\ &\qquad+\epsilon_m^{1/2}(\overline{x}'_m-\overline{x}''_m)=0,\end{aligned}\tag{55}$$

where

$$a_m=(t_m(\overline{x}'_m))^k;\ c_m=d_jg_j(\overline{x}'_m)+\sqrt{d_j^2g_j^2(\overline{x}'_m)+1/m}.$$

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Let

$$b_m = a_m^{\frac{1}{k}-1}\left[f^{k-1}(\overline{x}'_m) + \frac{1}{2^k}\sum_{j\in J^+(\overline{x})\cup J(\overline{x})} d_j c_m^{k-1}\left(1 + d_j g_j(\overline{x}'_m)\left(d_j^2 g_j^2(\overline{x}'_m) + \frac{1}{m}\right)^{-1/2}\right)\right].$$

Without loss of generality, we assume that

$$\frac{a_m^{\frac{1}{k}-1} f^{k-1}(\overline{x}'_m)}{b_m} \to \lambda,$$

$$c_{j,m}/b_m \to \mu_j, \quad j \in J^+(\overline{x}) \cup J(\overline{x}), \tag{56}$$

where, for $j \in J^+(\overline{x}) \cup J(\overline{x})$,

$$c_{j,m} = \frac{a_m^{\frac{1}{k}-1}}{2^k} d_j c_m^{k-1}\left(1 + d_j g_j(\overline{x}'_m)\left(d_j^2 g_j^2(\overline{x}'_m) + \frac{1}{m}\right)^{-1/2}\right).$$

It is easy to see that $\mu_j = 0, j \in J(\overline{x})$, if $k > 1$. Thus we obtain $\lambda \geq 0, \mu_j \geq 0$ with $\lambda + \sum_{j\in J^*(\overline{x})} \mu_j = 1$.

Dividing (55) by b_m and taking the limit, we get

$$\lambda \bigtriangledown f(\overline{x}) + \sum_{j\in J^+(\overline{x})\cup J(\overline{x})} \mu_j \bigtriangledown g_j(\overline{x}) = 0.$$

Applying the second order necessary optimality condition to (54), we know that, for every $u \in R^p$, there exist $F_m \in \partial^2 f(\overline{x}'_m), G_{j,m} \in \partial^2 g_j(\overline{x}'_m), j \in J^+(\overline{x}) \cup J(\overline{x})$ such that

$$\begin{aligned}
&\left(\frac{1}{k}-1\right) a_m^{\frac{1}{k}-2}\left(f^{k-1}(\overline{x}'_m) \bigtriangledown f(\overline{x}'_m)^\top u + \sum_{j\in J^+(\overline{x})\cup J(\overline{x})} \alpha_j(m) \bigtriangledown g_j(\overline{x}'_m)^\top u\right)^2\\
&+a_m^{\frac{1}{k}-1}\left((k-1)f^{k-2}(\overline{x}'_m)(\bigtriangledown f(\overline{x}'_m)^\top u)^2 + \sum_{j\in J^+(\overline{x})\cup J(\overline{x})} \theta_j(m)(\bigtriangledown g_j(\overline{x}'_m)^\top u)^2\right)\\
&+a_m^{\frac{1}{k}-1} u^\top\left(f^{k-1}(\overline{x}'_m)F_m + \frac{1}{2^k}\sum_{j\in J^+(\overline{x})\cup J(\overline{x})} d_j\left[d_j g_j(\overline{x}'_m) + \sqrt{d_j^2 g_j^2(\overline{x}'_m) + \frac{1}{m}}\right]^{k-1}\right.\\
&\left.\left(1 + d_j g_j(\overline{x}'_m)\sqrt{d_i^2 g_j^2(\overline{x}'_m) + \frac{1}{m}}\right) G_{j,m}\right) u \geq 0,
\end{aligned} \tag{57}$$

where $\alpha_j(n), \theta_j(n)$ are real numbers. Since $\{\bigtriangledown g_j(\overline{x})\}_{j\in J^*(\overline{x})}$ is linearly independent, i.e., $\{\bigtriangledown g_j(\overline{x})\}_{j\in J^+(\overline{x})\cup J(\overline{x})}$ is linearly independent, and $\overline{x}'_m \to \overline{x}$, by Lemma 5.3, we conclude that, for every $\overline{u} \in R^p$ satisfying (44), there exists $u_m \in R^p$, such that

$$\bigtriangledown g_j(\overline{x}'_m)^\top u_m = 0, \quad j \in J^*(\overline{x}), \tag{58}$$

and $u_m \to \overline{u}$.

Furthermore, for every u_m satisfying (58), we obtain $F_m \in \partial^2 f(\overline{x}'_m)$, $G_{j,m} \in \partial^2 g_j(\overline{x}'_m)$, $j \in J^+(\overline{x}) \bigcup J(\overline{x})$, such that (57) holds (with u replaced by u_m).

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The combination of (58) and (55) gives us

$$a_m^{\frac{1}{k}-1} f^{k-1}(\overline{x}'_m) \bigtriangledown f(\overline{x}'_m)^\top u_m = -\epsilon_m^{\frac{1}{2}}(\overline{x}'_m - \overline{x}''_m)^\top u_m.$$

Thus

$$\left| \left(\frac{1}{k}-1\right) a_m^{\frac{1}{k}-2} \left(f^{k-1}(\overline{x}'_m) \bigtriangledown f(\overline{x}'_m)^\top u_m\right)^2 \right| \leq \frac{1}{q^*}\epsilon_m^{\frac{3}{4}}\|u_m\|^2$$

and

$$\left|(k-1)a_m^{\frac{1}{k}-1} f^{k-2}(\overline{x}'_m)(\bigtriangledown f(\overline{x}'_m)^\top u_m)^2\right| \leq \frac{1-k}{q^*}\epsilon_m^{\frac{3}{4}}\|u_m\|^2.$$

Noting that $b_m \geq 1$, we obtain, as $m \to +\infty$,

$$\frac{1}{b_m}\left(\frac{1}{k}-1\right) a_m^{\frac{1}{k}-2} \left(f^{k-1}(\overline{x}'_m) \bigtriangledown f(\overline{x}'_m)^\top u_m\right)^2 \to 0, \tag{59}$$

$$\frac{1}{b_m}(k-1)a_m^{\frac{1}{k}-1} f^{k-2}(\overline{x}'_m)(\bigtriangledown f(\overline{x}'_m)^\top u_m)^2 \to 0. \tag{60}$$

By the upper semicontinuity of $x \to \partial^2 f(x),\ x \to \partial^2 g_j(x)(j = 1,\ldots,q)$ and the nonemptiness and compactness of $\partial^2 f(\overline{x})$ and $\partial^2 g_j(\overline{x})$, without loss of generality we can assume that $F_m \to F \in \partial^2 f(\overline{x}),\ G_{j,m} \to G_j \in \partial^2 g_j(\overline{x}),\ j \in J^+(\overline{x}) \cup J(\overline{x})$.

Letting $u = u_m$ in (57) and substituting (58) into it, dividing (57) by b_m and taking the limit, and applying (56), (59), and (60), we obtain

$$\overline{u}^T \left(\lambda F + \sum_{j\in J^+(\overline{x})\cup J(\overline{x})} \mu_j G_j\right)\overline{u} \geq 0,$$

where $\lambda > 0$. ◻

Acknowledgments. The authors are grateful to the two referees for their detailed comments and suggestions which have improved the presentation of this paper.

REFERENCES

[1] A. Auslender, *Penalty methods for computing points that satisfy second order necessary conditions*, Math. Programming, 17 (1979), pp. 229–238.

[2] A. Auslender, R. Cominetti, and M. Haddou, *Asymptotic analysis for penalty and barrier methods in convex and linear programming*, Math. Oper. Res., 22 (1997), pp. 43–62.

[3] E. J. Balder, *An extension of duality-stability relations to nonconvex optimization problems*, SIAM J. Control Optim., 15 (1977), pp. 329–343.

[4] A. Ben-Tal and J. Zowe, *Necessary and sufficient optimality conditions for a class of nonsmooth minimization problems*, Math. Programming, 24 (1982), pp. 70–91.

[5] D. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.

[6] J. M. Borwein and D. Preiss, *A smooth variational principle with applications to subdifferentiability and differentiability*, Trans. Amer. Math. Soc., 303 (1987), pp. 517–527.

[7] C. Charalambous, *On conditions for optimality of the nonlinear l_1 problem*, Math. Programming, 17 (1979), pp. 123–135.

[8] F. H. Clarke, Y. S. Ledyaev, and P. R. Wolenski, *Proximal analysis and minimization principles*, J. Math. Anal. Appl., 196 (1995), pp. 722–735.

[9] A. Fiacco and G. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York, 1968.

Downloaded 08/01/13 to 158.132.161.52. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

[10] C. J. Goh and X. Q. Yang, *A nonlinear Lagrangian theory for nonconvex optimization*, J. Optim. Theory Appl., 109 (2001), pp. 99–121.

[11] J. B. Hiriart-Urruty, J. J. Strodiot, and V. Hien Nguyen, *Generalized Hessian matrix and second-order optimality conditions for problems with $C^{1,1}$ data*, Appl. Math. Optim., 11 (1984), pp. 43–56.

[12] X. X. Huang and X. Q. Yang, *Nonlinear Lagrangian for Multiobjective Optimization and Application to Duality and Exact Penalization*, preprint, Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, 2001.

[13] A. D. Ioffe, *Necessary and sufficient conditions for a local minimum.* III: *Second-order conditions and augmented duality*, SIAM J. Control Optim., 17 (1979), pp. 266–288.

[14] D. Li, *Zero duality gap for a class of nonconvex optimization problems*, J. Optim. Theory Appl., 85 (1995), pp. 309–324.

[15] P. Loridan, *Necessary conditions for ϵ-optimality*, Math. Programming Stud., 19 (1982), pp. 140–152.

[16] R. T. Rockafellar, *Conjugate Duality and Optimization*, SIAM, Philadelphia, PA, 1974.

[17] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998.

[18] A. M. Rubinov, B. M. Glover, and X. Q. Yang, *Modified Lagrangian and penalty functions in continuous optimization*, Optimization, 46 (1999), pp. 327–351.

[19] A. M. Rubinov, B. M. Glover, and X. Q. Yang, *Decreasing functions with applications to penalization*, SIAM J. Optim., 10 (1999), pp. 289–313.

[20] S. A. Vavasis, *Black-box complexity of local minimization*, SIAM J. Optim., 3 (1993), pp. 60–79.

[21] X. Q. Yang, *Second-order conditions of $C^{1,1}$ optimization with applications*, Numer. Funct. Anal. Optim., 14 (1993), pp. 621–632.

[22] X. Q. Yang, *An exterior point method for computing points that satisfy second order necessary conditions for a $C^{1,1}$ optimization problem*, J. Math. Anal. Appl., 87 (1994), pp. 118–133.

[23] X. Q. Yang, *Second-order global optimality conditions for convex composite optimization*, Math. Programming, 81 (1998), pp. 327–347.