Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 1, No. 1, pp. 81–99, January 2005

GENERALIZED AUGMENTED LAGRANGIAN METHODS FOR EQUALITY CONSTRAINED OPTIMIZATION PROBLEMS*

X.X. Huang and X.Q. Yang

Tribute to R.T. Rockafellar.†

Abstract: In this paper, we consider generalized augmented Lagrangian methods, including a classical augmented Lagrangian method and some "lower order" augmented Lagrangian methods as special cases, for a mathematical program with only equality constraints. Since generalized augmented Lagrangians are in general not differentiable or even not locally Lipschitz, we carry out convergence analysis of first-order and second-order stationary points of generalized augmented Lagrangian methods by applying the Borwein-Preiss approximate smooth variational principle.

Key words: *equality constrained optimization, augmented Lagrangian, constraint qualification, optimality condition*

Mathematics Subject Classification: *90C30, 65K05*

1 Introduction

In the literature, three types of augmented Lagrangians have attracted extensive attention: (i) classical augmented Lagrangian with a convex quadratic augmenting function; (ii) general augmented Lagrangian with a convex augmenting function; and (iii) generalized augmented Lagrangian with a (nonconvex) level-bounded augmenting function. Their implication relations are (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii).

Classical augmented Lagrangian method was first proposed by Hestenses [9] and Powell [14] to solve a mathematical program with only equality constraints. It was later extended by Rockafellar to solve optimization problems with both equality and inequality constraints, see, e.g., [1, 15, 16]. As noted in [1], in comparison with the traditional (quadratic) penalty method for constrained optimization problems, convergence of augmented Lagrangian method usually does not require that the penalty parameter tends to infinity. This important advantage results in elimination or at least moderation of the ill-conditioning problem in the traditional penalty method. Another important advantange of

*The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 5141/01E).

†This paper is dedicated to Terry Rockafellar in appreciation of his sustained contributions to optimization. Our recent work in this and other papers on the topic of augmented Lagrangian duality theory has been inspired by his early work.

augmented Lagrangian method over the traditional penalty method is that its convergence rate is considerably better than that of the traditional penalty method.

Recently, Rockafellar and Wets [17] introduced a general augmented Lagrangian with a convex augmenting function for an extended real-valued nonconvex optimization problem and established a strong duality result. Moreover, a necessary and sufficient condition for the exact penalty representation in the framework of the general augmented Lagrangian was also obtained.

More recently, a generalized augmented Lagrangian was introduced in [10] by relaxing the convexity on the augmenting function. This relaxation allows for the unification of some unconstrained methods for constrained and unconstrained optimization problems, e.g., the general augmented Lagrangian method discussed in [17], the "lower order" penalty function methods considered in [12] and [13] as well as a class of nonlinear penalty methods studied in [18].

It was shown in [19] that a "lower order" nonlinear penalty function usually admits a smaller least exact penalty parameter than the ordinary l_1 penalty function(see, e.g., [4] and [3]). As shown in [10], generalized augmented Lagrangian with a "lower order" augmenting function generally requires weaker conditions to guarantee its global exact penalty representation property than the general augmented Lagrangian in [17]. It will be shown in Section 2 of this paper that the generalized augmented Lagrangian with a "lower order" augmenting function also admits a smaller least local exact penalty representation paramter than the general augmented Lagrangian. These results motivate us to further study generalized augmented Lagrangian methods.

Another direction in the study of augmented Lagrangian is the so-called exact augmented Lagrangians for inequality constrained nonlinear programming problems (see, e.g., [6, 7, 8]). Unlike the (generalized) augmented Lagrangian we mentioned above in which the penalty term only considers the feasibility of the the original constrained program, exact augmented Lagrangian takes into account both the feasibility and the KKT conditions of the original constrained program. Under certain conditions, the relationship in terms of optimality conditions, local/global optimal solutions of the augmented Lagrangian function and that of the original constrained optimization problem has been established (see [6, 7, 8]).

It is clear that augmented Lagrangian methods for constrained optimization are a class of unconstrained methods. That is, they are used to solve a constrained optimization problem by converting it into one or a sequence of unconstrained optimization problems. However, it should be cautioned that for a nonconvex program, usual optimization methods only generate stationary points, i.e., points that satisfy first-order or second-order necessary optimality conditions. Thus, it is both interesting and significant to investigate whether the first-order or second-order stationary points of the unconstrained mathematical programs converge to that of the original constrained mathematical program. In [11], in the context of a mathematical program with both equality and inequality constraints, we proved that the first-order and second-order stationary points of the classical augmented Lagrangian problems considered in [15, 16] converge to that of the original constrained mathematical program, respectively. In the framework of a mathematical program with only equality constraints, we showed that the second-order stationary points of the augmented Lagrangian problems converge to that of the original constrained mathematical program for general augmented Lagrangians with convex augmenting functions studied in [17].

It is worth noting that there is generally no explicit expression for a general augmented Lagrangian for a mathematical program with inequality constraints and thus there exists some techinical difficulty in deriving first-order and second-order necessary optimality conditions for the corresponding augmented Lagrangian problems [11]. So we considered mathematical programs with only equality constraints and their general augmented Lagrangian problems. In this paper, for the same reason, we shall restrict our attention to a class of generalized augmented Lagrangian methods for mathematical programs with only equality constraints. We shall investigate their convergence properties. More specifically, we shall prove that first-order and second-order stationary points of the generalized augmented Lagrangian problems converge to that of the original constrained mathematical program, respectively. Since these generalized augmented Lagrangians are not differentiable or even not locally Lipschitz, we approximate these generalized augmented Lagrangians by certain smooth functions. We also show that the first-order and second-order stationary points of the smooth approximate problems converge to that of the original constrained optimization problem, respectively.

2 Generalized Augmented Lagrangian

We recall the definition of the generalized augmented Lagrangian. For details, see [10].

Consider the following primal optimization problem:

$$\text{(P)} \qquad \inf_{x \in R^n} \varphi(x),$$

where $\varphi : R^n \to \bar{R} = R \bigcup \{-\infty, +\infty\}$ is an extended real-valued function.

Suppose that $\bar{f}(x,u) : R^n \times R^m \to \bar{R}$ is a dualizing parameterization function of φ, i.e.,

$$\bar{f}(x,0) = \varphi(x), \quad x \in R^n.$$

Let $\sigma : R^m \to \bar{R}$ be a generalized augmenting function, i.e., it is proper, lower semi-continuous (lsc in short), level-bounded (the set $\{u \in R^m : \sigma(u) \leq \alpha\}$ is always bounded for any $\alpha \in R$), and attains its minimum 0 at the origin $0 \in R^m$. In [17], the augmenting function σ is required to be convex. Clearly, an augmenting function, which is lsc, convex and attains its minimum 0 at its unique minimizer $0 \in R^m$, is level-bounded.

The generalized augmented Lagrangian is defined as

$$l(x,y,r) = \inf_{u \in R^m} \{\bar{f}(x,u) - \langle y,u \rangle + r\sigma(u)\}, \quad x \in R^n, y \in R^m, r > 0.$$

Based on the generalized augmented Lagrangian for unconstrained programs, we introduce generalized Lagrangian for constrained programs. Since there exists some technical difficulty in deriving an explicit expression of a generalized augmented Lagrangian for a mathematical program with inequality constraints, we will only discuss in this paper mathematical programs with equality constraints.

Consider the following constrained program:

$$\begin{array}{lll} \text{(CP)} \qquad & \inf & f(x) \\ & \text{s.t.} & x \in R^n \\ & & g_j(x) = 0, \quad j = 1, \ldots, m, \end{array}$$

where $f, g_j : R^n \to R, j = 1, \ldots, m$ are twice continuously differentiable functions.

Denote by X_0 the feasible set of (CP), i.e.,

$$X_0 = \{x \in R^n : g_j(x) = 0, j = 1, \ldots, m\}.$$

Let M_{CP} denote the optimal value of the problem (CP).

Set

$$\varphi(x) = \begin{cases} f(x), & \text{if } x \in X_0, \\ +\infty, & \text{if } x \in R^n \setminus X_0. \end{cases} \tag{1}$$

It is clear that (CP) is equivalent to the following unconstrained problem (P') in the sense that the two problems have the same set of (locally) optimal solutions and the same optimal value:

$$\text{(P')} \qquad \inf_{x \in R^n} \varphi(x).$$

Define the dualizing parameterization function:

$$\bar{f}_{CP}(x,u) = f(x) + \delta_{\{0_m\}}(G(x)+u), \quad x \in R^n, u \in R^m, \tag{2}$$

where 0_m is the origin of R^m and $G(x) = (g_1(x), \ldots, g_m(x))$. Thus, a class of generalized augmented Lagrangians for (CP) with the dualizing parameterization function f_{CP} defined by (2) can be expressed as

$$\bar{l}_{CP}(x,y,r) = \inf\{\bar{f}_{CP}(x,u) - \langle y,u \rangle + r\sigma(u) : u \in R^m\}, \tag{3}$$

where σ is a generalized augmenting function.

It can be easily computed from (3) that

$$\bar{l}_{CP}(x,y,r) = f(x) + \sum_{j=1}^{m} y_j g_j(x) + r\sigma(-g_1(x), \ldots, -g_m(x)). \tag{4}$$

Definition 2.1[10]. Consider the constrained program (CP) and the associated generalized augmented Lagrangian $\bar{l}_{CP}(x,y,r)$. A vector $\bar{y} \in R^m$ is said to support a global exact penalty representation if there exists $\bar{r} > 0$ such that

$$M_{CP} = \inf\{\bar{l}_{CP}(x,\bar{y},r) : x \in R^n\}, \quad \forall r \geq \bar{r}$$

and

$$\text{argmin (CP)} = \text{argmin}_x \bar{l}_{CP}(x,\bar{y},r), \quad \forall r \geq \bar{r},$$

where argmin (CP) and $\text{argmin}_x \bar{l}_{CP}(x,\bar{y},r)$ denote the set of optimal solutions of (CP) and the set of optimal solutions to the problem of minimizing $\bar{l}_{CP}(x,\bar{y},r)$ over $x \in R^n$, respectively.

Definition 2.2. Consider the constrained program (CP) and the associated generalized augmented Lagrangian $\bar{l}_{CP}(x,y,r)$. Let $\bar{x} \in X_0$ be a local solution to (CP). $\bar{y} \in R^m$ is said to support a local exact penalty representation for (CP) at $\bar{x}$ in the framework of the generalized augmented Lagrangian $\bar{l}_{CP}(x,y,r)$ if there exist $\bar{r} > 0$ and $\delta > 0$ such that

$$f(\bar{x}) \leq \bar{l}_{CP}(x,\bar{y},\bar{r}), \quad \forall x \in V_\delta = \{x \in R^n : \|x - \bar{x}\| \leq \delta\}. \tag{5}$$

In the sequel, we restrict our attention to two classes of generalized augmenting functions σ for (CP). More specifically, they are stated as follows:

(a) the generalized augmenting function is $\sigma(u) = \left[\sum_{j=1}^m |u_j|\right]^\alpha$, where $\alpha > 0$;

(b) the generalized augmenting function is $\sigma(u) = [\max\{|u_j| : j = 1, \ldots, m\}]^\alpha$, where $\alpha > 0$.

It is obvious from (4) that the generalized augmented Lagrangian in case (a) is

$$L^\alpha(x,y,r) = f(x) + \sum_{j=1}^m y_j g_j(x) + r\left[\sum_{j=1}^m |g_j(x)|\right]^\alpha,$$

and the generalized augmented Lagrangian in case (b) is

$$\bar{L}^\alpha(x,y,r) = f(x) + \sum_{j=1}^m y_j g_j(x) + r\left[\max\{|g_j(x)| : j = 1,\ldots,m\}\right]^\alpha.$$

The generalized augmented Lagrangian problems in case (a) are

$$(P^\alpha_{y,r}) \qquad \min_{x\in R^n} L^\alpha(x,y,r), \tag{6}$$

and the generalized augmented Lagrangian problems in case (b) are

$$(\bar{P}^\alpha_{y,r}) \qquad \min_{x\in R^n} \bar{L}^\alpha(x,y,r). \tag{7}$$

It is clear from ([10], Theorems 5.3 and 5.4) that if $0 < \alpha' \leq \alpha$, weaker conditions are needed to guarantee that $\bar{y} \in R^m$ supports a global exact penalty representation for (CP) in the framework of the generalized augemented Lagrangian $L^{\alpha'}(x,y,r)$ than in the case when the generalized augmented Lagrangian $L^\alpha(x,y,r)$ is used. Next we show that if $0 < \alpha' \leq \alpha$ and $\bar{y}$ supports a local exact penalty representation for (CP) in the framework of $L^\alpha(x,y,r)$, then $\bar{y}$ also supports a local exact penalty representation for (CP) in the framework of $L^{\alpha'}(x,y,r)$, and in some sense, $L^{\alpha'}(x,y,r)$ admits a smaller least local exact penalty representation parameter than $L^\alpha(x,y,r)$.

Suppose that $\bar{x} \in X_0$ is a local solution to (CP) and $\bar{y} \in R^m$ supports a local exact penalty representation for (CP) at $\bar{x}$ in the framework of the generalized augmented Lagrangian $L^\alpha(x,y,r)$. Let $\delta > 0$ in Definition 2.2 (with $\bar{l}_{CP}$ replaced by L^α) be fixed. Define the least local exact penalization representation parameter by

$$r^*_\alpha(\bar{y},\delta) = \inf\{\bar{r} > 0 : f(\bar{x}) \leq L^\alpha(x,\bar{y},\bar{r}), \quad \forall x \in V_\delta\}.$$

Define the following locally perturbed problem of (CP):

$$\begin{array}{lll} (CP_s(\delta)) & \inf & f(x) \\ & \text{s.t.} & x \in V_\delta \\ & & g_j(x) = s_j, \quad j = 1,\ldots,m, \end{array}$$

where $s = (s_1,\ldots,s_m) \in R^m$. Let $\beta(s,\delta)$ denote the optimal value of $(CP_s(\delta))$. If $(CP_s(\delta))$ does not have a feasible solution, $\beta(s,\delta) = +\infty$. Set

$$S(\delta) = \{s : s \in R^m, s \neq 0, s_j = g_j(x), j = 1,\ldots,m, \text{ for some } x \in V_\delta\}.$$

We have the following proposition concerning the calculation of $r^*(\bar{y},\delta)$.

Proposition 2.1.

$$r^*_\alpha(\bar{y},\delta) = \max\{0, \sup_{s\in S(\delta)} \frac{f(\bar{x}) - \beta(s,\delta) - \sum_{j=1}^m \bar{y}_j s_j}{\|s\|_1^\alpha}\},$$

where $\|s\|_1 = \sum_{j=1}^m |s_j|, \forall s = (s_1,\ldots,s_m) \in R^m$.

Proof. Let

$$A=\sup_{s\in S(\delta)}\frac{f(\bar{x})-\beta(s,\delta)-\sum_{j=1}^{m}\bar{y}_j s_j}{\|s\|_1^{\alpha}}.$$

First we prove that $r^*_\alpha(\bar{y},\delta)\geq A$.

Let $\bar{r}>0$ satisfy

$$f(\bar{x})\leq L^{\alpha}(x,\bar{y},\bar{r}),\quad \forall x\in V_\delta. \tag{8}$$

Then for any $s\in S(\delta)$, there exists an x feasible to $(CP_s(\delta))$ such that

$$f(\bar{x})\leq f(x)+\sum_{j=1}^{m}\bar{y}_j s_j+\bar{r}\|s\|_1^{\alpha}.$$

Consequently,

$$f(\bar{x})\leq \beta(s,\delta)+\sum_{j=1}^{m}\bar{y}_j s_j+\bar{r}\|s\|_1^{\alpha}.$$

Thus,

$$\bar{r}\geq A.$$

By the definition of $r^*_\alpha(\bar{y},\delta)$, we see that $r^*_\alpha(\bar{y},\delta)\geq A$.

Now we show that

$$f(\bar{x})\leq L^{\alpha}(x,\bar{y},A),\quad \forall x\in V_\delta.$$

Suppose to the contrary that there exist $x_0\in V_\delta$ and $t_0>0$ such that

$$f(x_0)+\sum_{j=1}^{m}\bar{y}_j g_j(x_0)+A\left[\sum_{j=1}^{m}|g_j(x_0)|\right]^{\alpha}\leq f(\bar{x})-t_0. \tag{9}$$

We assert that $x_0\notin X_0$. Otherwise, from (8), we have

$$f(\bar{x})\leq f(x_0). \tag{10}$$

On the other hand, from (9) we obtain

$$f(x_0)\leq f(\bar{x})-t_0,$$

contradicting (10). Let

$$s_j=g_j(x_0),\quad j=1,\ldots,m. \tag{11}$$

Then from $x_0\notin X_0$, we conclude that $s\neq 0$. Moreover, it is apparent that $s\in S(\delta)$. Substituting (11) into (9), we get

$$f(x_0)+\sum_{j=1}^{m}\bar{y}_j s_j+A\|s\|_1^{\alpha}\leq f(\bar{x})-t_0.$$

This further implies that

$$\beta(s,\delta)+\sum_{j=1}^{m}\bar{y}_j s_j+A\|s\|_1^{\alpha}\leq f(\bar{x})-t_0.$$

As a result,

$$A\leq\frac{f(\bar{x})-t_0-\beta(s,\delta)-\sum_{j=1}^{m}\bar{y}_j s_j}{\|s\|_1^{\alpha}},$$

contradicting the definition of A. It follows that $r^*_\alpha(\bar{y},\delta)\leq A$. So the conclusion holds. □

Now assume that $\bar{y}$ supports a local exact penalty representation for (CP) at $\bar{x} \in X_0$ in the framework of the generalized augmented Lagrangian $L^\alpha(x,y,r)$. Then there exists $\delta > 0$ such that (5) holds. Suppose that $0 < \theta < 1$. By the continuity of g_j $(j = 1,\ldots,m)$, there exists $\delta' > 0$ such that

$$\sum_{j=1}^{m} |g_j(x)| \le \theta, \quad \forall x \in V_{\delta'} = \{x \in R^n : \|x - \bar{x}\| \le \delta'\}. \tag{12}$$

Set $\delta'' = \min\{\delta, \delta'\}$. Then we have

$$f(\bar{x}) \le L^\alpha(x, \bar{y}, \bar{r}), \quad \forall x \in V_{\delta''} = \{x \in R^n : \|x - \bar{x}\| \le \delta''\}. \tag{13}$$

By Proposition 2.1, we obtain

$$r^*_\alpha(\bar{y}, \delta'') = \max\{0, \sup_{s \in S(\delta'')} \frac{f(\bar{x}) - \beta(s,\delta'') - \sum_{j=1}^m \bar{y}_j s_j}{\|s\|_1^\alpha}\}. \tag{14}$$

Suppose that $0 < \alpha' < \alpha$. Then (13) also holds when α is replaced by α'. Again, by Proposition 2.1, we have

$$r^*_{\alpha'}(\bar{y}, \delta'') = \max\{0, \sup_{s \in S(\delta'')} \frac{f(\bar{x}) - \beta(s,\delta'') - \sum_{j=1}^m \bar{y}_j s_j}{\|s\|_1^{\alpha'}}\}. \tag{15}$$

By our assumption (12) and the definition of δ'', we see that

$$\|s\|_1 < \theta < 1, \quad \forall s \in S(\delta'').$$

Consequently, it is easily deduced from (14) and (15) that

$$r^*_{\alpha'}(\bar{y}, \delta'') \le \theta^{\alpha - \alpha'} r^*_\alpha(\bar{y}, \delta'').$$

Note that $\theta^{\alpha-\alpha'} < 1$. So we conclude that the generalized augmented Lagrangian $L^{\alpha'}(x,y,r)$ admits a smaller least local exact penalty representation parameter than $L^\alpha(x,y,r)$ if $0 < \alpha' < \alpha$.

Similar arguments motivate our interest in the generalized augmented Lagrangian $\bar{L}^\alpha(x, y, r)$, particularly when $\alpha > 0$ is small.

3 Convergence Analysis

In this section, we consider the constrained optimization problem (CP). We shall discuss the convergence of first-order and second-order necessary optimality conditions of the generalized augmented Lagrangian problems $(P^\alpha_{y,r})$ and $(\bar{P}^\alpha_{y,r})$, respectively.

3.1 Optimality Conditions for Generalized Lagrangian Problems

It is clear that both $L^\alpha(x,y,r)$ and $\bar{L}^\alpha(x,y,r)$ are generally nonsmooth functions. Furthermore, when $0 < \alpha < 1$, they may not even be locally Lipschitz or convex composite functions. It is true that there have already been some results concerning first-order and second-order necessary conditions for nonsmooth or even non-Lipschitz optimization problems in the literature (see, e.g., [17]). However, we have not found one that is suitable for our purpose of convergence analysis of the generalized augmented Lagrangian methods. Fortunately, we are able to derive appropriate first-order and second-order necessary conditions for the optimization problems (6) and (7) by combining the technique of smooth approximation and the approximate smooth variational principle due to Borwein and Preiss [2].

The next lemma follows immediately from the approximate smooth variational principle ([2], Theorem 2.6).

Lemma 3.1 Let $X \subset R^n$ be nonempty and closed. Let $h : X \to R$ be lsc and bounded below on X. Suppose that $\bar{x}$ is a point such that $h(\bar{x}) < \inf_{x\in X} h(x) + \epsilon$, where ϵ is a positive number. Then, for any $\lambda > 0$, there exists $x_\epsilon \in X$ such that

$$\|x_\epsilon - \bar{x}\| < \lambda,$$

$$h(x_\epsilon) < \inf_{x\in X} h(x) + \epsilon$$

and

$$h(x_\epsilon) \leq h(x) + \epsilon/\lambda^2 \|x - x_\epsilon\|^2, \quad \forall x \in X.$$

We need also the following lemma to derive second-order necessary conditions for $(P^\alpha_{y,r})$ and $(\bar{P}^\alpha_{y,r})$.

Lemma 3.2. Let $\{c_i^k\}_{k=1}^\infty \subset R^n, i = 1, \ldots, q$ be sequences such that

$$\lim_{k\to+\infty} c_i^k = c_i, \quad i = 1, \ldots, q.$$

Suppose that $\{c_i : i = 1, \ldots, q\}$ are linearly independent. Then $\forall \bar{d} \in \{d \in R^n : c_i^T d = 0, i = 1, \ldots, q\}$, there exists $\bar{k} > 0$ such that, when $k \geq \bar{k}$, there exists $d^k \in R^n$ satisfying $(c_i^k)^T d^k = 0, \quad i = 1, \ldots, q$ and $d^k \to \bar{d}$.

Proof. It follows directly from Corollary II.3.4 of [5] (see also Lemma 5.1 of [20]). $\square$

The next proposition gives first-order and second-order necessary condition for a local minimum of $(P^\alpha_{y,r})$.

Proposition 3.1. Consider the generalized augmented Lagrangian problem $(P^\alpha_{y,r})$. Suppose that $\bar{x}$ is a local minimum of $(P^\alpha_{y,r})$ and $\{\bigtriangledown g_j(\bar{x}) : j = 1, \ldots, m\}$ are linearly independent. Then there exist $\nu_j, j = 1, \ldots, m$ such that

(i)

$$\bigtriangledown f(\bar{x}) + \sum_{j=1}^m \nu_j \bigtriangledown g_j(\bar{x}) = 0; \tag{16}$$

and

(ii) for any $d \in R^n$ satisfying

$$\bigtriangledown g_j(\bar{x}) d = 0, \quad j = 1, \ldots, m, \tag{17}$$

there holds

$$d^T \bigtriangledown^2 f(\bar{x})d + \sum_{j=1}^{m} \nu_j d^T \bigtriangledown^2 g_j(\bar{x})d \geq 0. \tag{18}$$

Proof. Since $\bar{x}$ is a local minimum of $(P^{\alpha}_{y,r})$, there exists a neighbourhood $U_\delta = \{x \in R^n : \|x - \bar{x}\| \leq \delta\}$ such that

$$L^{\alpha}(\bar{x}, y, r) \leq L^{\alpha}(x, y, r), \quad \forall x \in U_\delta.$$

Let k be an integer and

$$s_k(x) = f(x) + \sum_{j=1}^{m} y_j g_j(x) + r \left[\sum_{j=1}^{m} \sqrt{g_j^2(x) + 1/k^2} \right]^{\alpha}, \quad x \in R^n.$$

It is not hard to check that for $x \in U_\delta$, there holds

$$s_k(\bar{x}) \leq s_k(x) + \begin{cases} rm^{\alpha}/k^{\alpha}, & \text{if } \sum_{j=1}^{m} |g_j(\bar{x})| = 0, \\ r\alpha/k^{\alpha} \left[\sum_{j=1}^{m} |g_j(\bar{x})| \right]^{\alpha-1}, & \text{if } 0 < \alpha < 1 \text{ and } \sum_{j=1}^{m} |g_j(\bar{x})| > 0, \\ r\alpha/k^{\alpha} \left[\sum_{j=1}^{m} \sqrt{g_j^2(\bar{x}) + 1} \right]^{\alpha-1}, & \text{if } \alpha > 1 \text{ and } \sum_{j=1}^{m} |g_j(\bar{x})| > 0. \end{cases}$$

Set

$$\epsilon_k = 2\max\{m/k, \alpha/k \left[\sum_{j=1}^{m} |g_j(\bar{x})| \right]^{\alpha-1}, \ \alpha/k \left[\sum_{j=1}^{m} \sqrt{g_j^2(\bar{x}) + 1} \right]^{\alpha-1} \}.$$

Then we have

$$s_k(\bar{x}) < s_k(x) + \epsilon_k, \quad \forall x \in U_\delta. \tag{19}$$

It is clear that $\epsilon_k \downarrow 0$. So we assume without loss of generality that $\epsilon_k^{1/4} < \delta$.

From (19) and Lemma 3.1 (taking $\lambda = \epsilon_k^{1/4}$), we obtain $\bar{x}'_k \in U_\delta$ such that

$$\|\bar{x}'_k - \bar{x}\| < \epsilon_k^{1/4}$$

and $\bar{x}'_k$ is a minimum of the problem:

$$\min_{x \in U_\delta} s_k(x) + \epsilon_k^{1/2} \|x - \bar{x}'_k\|^2. \tag{20}$$

Since $\|\bar{x}'_k - \bar{x}\| < \epsilon_k^{1/4} < \delta$, it follows that $\bar{x}'_k \in \mathrm{int} U_\delta$. By the first-order necessary optimality condition for problem (20), we get

$$\bigtriangledown s_k(\bar{x}'_k) = 0.$$

That is,

$$\bigtriangledown f(\bar{x}'_k) + \sum_{j=1}^{m} y_j \bigtriangledown g_j(\bar{x}'_k) + r\alpha \left[\sum_{j=1}^{m} \sqrt{g_j^2(\bar{x}'_k) + 1/k^2} \right]^{\alpha-1}$$
$$\sum_{j=1}^{m} \left[g_j^2(\bar{x}'_k) + 1/k^2 \right]^{-1/2} g_j(\bar{x}'_k) \bigtriangledown g_j(\bar{x}'_k) = 0$$

or

$$\bigtriangledown f(\bar{x}'_k) + \sum_{j=1}^{m} \{y_j + r\alpha \left[\sum_{j=1}^{m} \sqrt{g_j^2(\bar{x}'_k) + 1/k^2} \right]^{\alpha-1}$$
$$\sum_{j=1}^{m} \left[g_j^2(\bar{x}'_k) + 1/k^2 \right]^{-1/2} g_j(\bar{x}'_k) \} \bigtriangledown g_j(\bar{x}'_k) = 0. \tag{21}$$

Let

$$\nu_j^k = y_j + r\alpha \left[\sum_{j=1}^{m} \sqrt{g_j^2(\bar{x}'_k) + 1/k^2} \right]^{\alpha-1}$$
$$\sum_{j=1}^{m} \left[g_j^2(\bar{x}'_k) + 1/k^2 \right]^{-1/2} g_j(\bar{x}'_k), \quad j = 1, \ldots, m. \tag{22}$$

Then (21) becomes

$$\bigtriangledown f(\bar{x}'_k) + \sum_{j=1}^{m} \nu_j^k \bigtriangledown g_j(\bar{x}'_k) = 0. \tag{23}$$

We assert that $\tau_k = \sum_{j=1}^{m} |\nu_j^k|$ is bounded. Indeed, suppose to the contrary that $\{\tau_k\}$ is unbounded. Assume without loss of generality that $\tau_k \to +\infty$ and

$$\lim_{k \to +\infty} \nu_j^k / \tau_k = \nu'_j, \quad j = 1, \ldots, m. \tag{24}$$

Dividing (23) by τ_k and passing to the limit as $k \to +\infty$, we obtain

$$\sum_{j=1}^{m} \nu'_j \bigtriangledown g_j(\bar{x}) = 0. \tag{25}$$

The combination of (24) and (25) contradicts the linear independence of $\{\bigtriangledown g_j(\bar{x}) : j = 1, \ldots, m\}$. Hence, each sequence $\{\nu_j^k\}$, $j = 1, \ldots, m$ is bounded. Assume without loss of generality that

$$\nu_j^k \to \nu_j, \quad j = 1, \ldots, m. \tag{26}$$

Taking the limit in (23) as $k \to +\infty$, we obtain (16). This proves (i).

Now we apply second-order necessary condition to problem (20) and see that for any $d \in R^n$, there holds

$$d^T \bigtriangledown^2 s_k(\bar{x}'_k) d + 2\epsilon_k^{1/2} d^T d \geq 0.$$

That is,

$$d^T \bigtriangledown^2 f(\bar{x}'_k) d + \sum_{j=1}^{m} \nu_j^k d^T \bigtriangledown^2 g_j(\bar{x}'_k) d$$
$$+ r\alpha(\alpha - 1) \sum_{j=1}^{m} \left[\sum_{j=1}^{m} \sqrt{g_j^2(\bar{x}'_k) + 1/k^2} \right]^{\alpha-2} \left[g_j^2(\bar{x}'_k) + 1/k^2 \right]^{-1} g_j^2(\bar{x}'_k) (\bigtriangledown g_j(\bar{x}'_k) d)^2$$

$$
\begin{aligned}
&+r\alpha\sum_{j=1}^{m}\left[\sum_{j=1}^{m}\sqrt{g_j^2(\bar{x}'_k)+1/k^2}\right]^{\alpha-1}\left[g_j^2(\bar{x}'_k)+1/k^2\right]^{-3/2}g_j^2(\bar{x}'_k)(\bigtriangledown g_j(\bar{x}'_k)d)^2\\
&+r\alpha\sum_{j=1}^{m}\left[\sum_{j=1}^{m}\sqrt{g_j^2(\bar{x}'_k)+1/k^2}\right]^{\alpha-1}\left[g_j^2(\bar{x}'_k)+1/k^2\right]^{-1/2}(\bigtriangledown g_j(\bar{x}'_k)d)^2\\
&+2\epsilon_k^{1/2}d^Td\geq 0, \qquad (27)
\end{aligned}
$$

where ν_j^k, $j=1,\ldots,m$ are as in (22).

Note that $\bar{x}'_k\to\bar{x}$ and $\{\bigtriangledown g_j(\bar{x}):j=1,\ldots,m\}$ are linearly independent. By Lemma 3.2, for any $d\in R^n$ satisfying (17), there exist $d_k\in R^n$ such that $d_k\to d$ as $k\to+\infty$ and

$$\bigtriangledown g_j(\bar{x}'_k)d_k=0,\quad j=1,\ldots,m. \qquad (28)$$

Substituting (28) into (27) (with d repalced by d_k), we obtain

$$d_k^T\bigtriangledown^2 f(\bar{x}'_k)d_k+\sum_{j=1}^{m}\nu_j^k d_k^T\bigtriangledown^2 g_j(\bar{x}'_k)d_k+2\epsilon_k^{1/2}d_k^Td_k\geq 0.$$

Passing to the limit as $k\to+\infty$ and applying (26), we obtain (18). The proof is complete. □

The first-order and second-order necessary conditions for a local minimum of $(\bar{P}^{\alpha}_{y,r})$ are given in the following proposition.

Proposition 3.2. Consider the generalized augmented Lagrangian problem $(\bar{P}^{\alpha}_{y,r})$. Suppose that $\bar{x}$ is a local minimum of $(\bar{P}^{\alpha}_{y,r})$ and $\{\bigtriangledown g_j(\bar{x}):j=1,\ldots,m\}$ are linearly independent. Then there exist $\nu_j, j=1,\ldots,m$ such that (16) holds and for any $d\in R^n$ satisfying (17), (18) holds.

Proof. Since $\bar{x}$ is a local minimum of $(\bar{P}^{\alpha}_{y,r})$, there exists a neighbourhood $U_\delta=\{x\in R^n:\|x-\bar{x}\|\leq\delta\}$ such that

$$\bar{L}^{\alpha}(\bar{x},y,r)\leq\bar{L}^{\alpha}(x,y,r),\quad\forall x\in U_\delta.$$

Let k be an integer and

$$s_k(x)=f(x)+\sum_{j=1}^{m}y_jg_j(x)+r\left[\sum_{j=1}^{m}(g_j^2(x)+1/k^2)^k\right]^{\frac{\alpha}{2k}},\quad x\in R^n.$$

Clearly, s_k is twice continuously differentiable. Moreover, it is routine to verify that for $x\in U_\delta$, there holds

$$s_k(\bar{x})\leq s_k(x)$$

$$+\begin{cases} \frac{m}{k^\alpha}, & \text{if } \max_{1\le j\le m}|g_j(\bar{x})| = 0, \\ r\alpha[\max_{1\le j\le m}|g_j(\bar{x})|]^{\alpha-1}[\frac{1}{k^2} + (\sum_{j=1}^m g_j^{2k}(\bar{x}))^{\frac{1}{2k}} - \max_{1\le j\le m}|g_j(\bar{x})|], & \\ & \text{if } \max_{1\le j\le m}|g_j(\bar{x})| > 0 \\ & \text{and } 0 < \alpha < 1, \\ r\alpha[\sum_{j=1}^m g_j^2(\bar{x}) + 1]^{\frac{\alpha}{2}}[\frac{1}{k^2} + (\sum_{j=1}^m g_j^{2k}(\bar{x}))^{\frac{1}{2k}} & \\ \quad - \max_{1\le j\le m}|g_j(\bar{x})|], & \text{if } \max_{1\le j\le m}|g_j(\bar{x})| > 0 \\ & \text{and } \alpha > 1. \end{cases}$$

Let

$$\epsilon_k = 2\max\{\frac{m}{k^\alpha}, r\alpha[\max_{1\le j\le m}|g_j(\bar{x})|]^{\alpha-1}[\frac{1}{k^2} + (\sum_{j=1}^m g_j^{2k}(\bar{x}))^{\frac{1}{2k}} - \max_{1\le j\le m}|g_j(\bar{x})|],$$
$$r\alpha[\sum_{j=1}^m g_j^2(\bar{x}) + 1]^{\frac{\alpha}{2}}[\frac{1}{k^2} + (\sum_{j=1}^m g_j^{2k}(\bar{x}))^{\frac{1}{2k}} - \max_{1\le j\le m}|g_j(\bar{x})|]\}.$$

Obviously, $\epsilon_k \downarrow 0$. Assume without loss of generality that $\epsilon_k^{1/4} < \delta, \quad \forall k$.

Clearly, we have

$$s_k(\bar{x}) < s_k(x) + \epsilon_k, \quad x \in U_\delta.$$

Applying Lemma 3.1, we obtain $\bar{x}'_k \in U_\delta$ such that

$$\|\bar{x}'_k - \bar{x}\| < \epsilon_k^{1/4}$$

and $\bar{x}'_k$ is a minimum of the problem:

$$\min_{x\in U_\delta} s_k(x) + \epsilon_k^{1/2}\|x - \bar{x}''_k\|^2. \tag{29}$$

Note that $\|\bar{x}'_k - \bar{x}\| < \epsilon_k^{1/4} < \delta$. It follows that $\bar{x}'_k \in \text{int}U_\delta$. By the first-order necessary optimality condition for problem (20), we get

$$\bigtriangledown s_k(\bar{x}'_k) = 0.$$

That is,

$$\bigtriangledown f(\bar{x}'_k) + \sum_{j=1}^m y_j \bigtriangledown g_j(\bar{x}'_k) + r\alpha\left[\sum_{j=1}^m \left(g_j^2(\bar{x}'_k) + 1/k^2\right)^k\right]^{\frac{\alpha}{2k}-1}$$
$$\sum_{j=1}^m \left(g_j^2(\bar{x}'_k) + 1/k^2\right)^{k-1} g_j(\bar{x}'_k) \bigtriangledown g_j(\bar{x}'_k) = 0,$$

or

$$\bigtriangledown f(\bar{x}'_k) + \sum_{j=1}^m \{y_j + r\alpha\left[\sum_{j=1}^m \left(g_j^2(\bar{x}'_k) + 1/k^2\right)^k\right]^{\frac{\alpha}{2k}-1}$$
$$\sum_{j=1}^m \left(g_j^2(\bar{x}'_k) + 1/k^2\right)^{k-1} g_j(\bar{x}'_k)\} \bigtriangledown g_j(\bar{x}'_k) = 0. \tag{30}$$

Let

$$\begin{aligned}\nu_j^k &= y_j + r\alpha \left[\sum_{j=1}^m \left(g_j^2(\bar{x}_k') + 1/k^2\right)^k\right]^{\frac{\alpha}{2k}-1} \\ &\quad \sum_{j=1}^m \left(g_j^2(\bar{x}_k') + 1/k^2\right)^{k-1} g_j(\bar{x}_k'), \quad j = 1, \ldots, m. \end{aligned} \tag{31}$$

Then (30) becomes

$$\bigtriangledown f(\bar{x}_k') + \sum_{j=1}^m \nu_j^k \bigtriangledown g_j(\bar{x}_k') = 0. \tag{32}$$

As argued in the proof of Proposition 3.1, we can show that each sequence $\{\nu_j^k\}$, $j = 1, \ldots, m$ is bounded. As a result, we can assume without loss of generality that

$$\lim_{k \to +\infty} \nu_j^k = \nu_j, \quad j = 1, \ldots, m. \tag{33}$$

Taking the limit in (32) and applying (33), we obtain (16).

Applying second-order necessary optimality condition to the problem (29), we have that for each $d \in R^n$, there holds

$$d^T \bigtriangledown^2 s_k(\bar{x}_k')d + 2\epsilon_k^{1/2} d^T d \geq 0.$$

That is,

$$d^T \bigtriangledown^2 f(\bar{x}_k')d + \sum_{j=1}^m \nu_j^k d^T \bigtriangledown^2 g_j(\bar{x}_k')d + \sum_{j=1}^m \beta_j^k (\bigtriangledown g_j(\bar{x}_k')d)^2 + 2\epsilon_k^{1/2} d^T d \geq 0, \tag{34}$$

where $nu_j^k, j = 1, \ldots, m$ are as in (31) and $\beta_j^k, j = 1, \ldots, m$ are some real numbers.

Since $\{\bigtriangledown g_j(\bar{x}) : j = 1, \ldots, m\}$ are linearly independent and $\bar{x}_k' \to \bar{x}$ as $k \to +\infty$, by Lemma 3.2, for any $d \in R^n$ satisfying (17), there exist $d_k \in R^n$ such that $d_k \to d$ as $k \to +\infty$ and

$$\bigtriangledown g_j(\bar{x}_k')d_k = 0, \quad j = 1, \ldots, m. \tag{35}$$

Substituting (35) into (34) (with d repalced by d_k), we obtain

$$d_k^T \bigtriangledown^2 f(\bar{x}_k')d_k + \sum_{j=1}^m \nu_j^k d_k^T \bigtriangledown^2 g_j(\bar{x}_k')d_k + 2\epsilon_k^{1/2} d_k^T d_k \geq 0.$$

Passing to the limit as $k \to +\infty$ and applying (33), we obtain (18). The proof is complete. □

Remark 3.1. In [3], the exact penalization technique was employed in order to derive first-order and second-order optimality conditions for a constrained optimization problem. It was assumed there that the orginal constrained program admits a local exact penalization at a local solution and the penalty function is a convex composite function. That is, the local minimizer of the constrained program is also a local minimizer of the convex composite penalty function. Then, from the previously developed first-order and second-order optimality conditions for a local minimizer of a convex composite function, one can obtain first-order and second-order optimality conditions for a local minimizer of a constrained optimization problem. The goal of Propositions 3.1 and 3.2 is purely to derive first-order and second-order conditions for local minimizers of problems $(P^{\alpha}_{y,r})$ and $(\bar{P}^{\alpha}_{y,r})$, respectively. It is worth noting that we did not assume that $\bar{x}$ is feasible to the original constrained program (CP). Moreover, when $0 < \alpha < 1$, neither of the generalized augmented Lagrangian problems $(P^{\alpha}_{y,r})$ and $(\bar{P}^{\alpha}_{y,r})$ is a convex composite optimization problem.

$\boxed{3.2}$ Convergence Results

Definition 3.1. Consider the constrained program (CP). Let $x^* \in X_0$. Suppose that $\bigtriangledown\{g_j(x^*) : j = 1, \ldots, m\}$ are linearly independent. The *first-order necessary optimality condition* is that $\exists \nu_j, j = 1, \ldots, m$ such that

$$\bigtriangledown f(x^*) + \sum_{j=1}^{m} \nu_j \bigtriangledown g_j(x^*) = 0; \tag{36}$$

and the *second-order necessary optimality condition* is that the first-order necessary condition (36) holds, and for any $d \in R^n$ satisfying

$$\bigtriangledown g_j(x^*)d = 0, \quad j = 1, \ldots, m, \tag{37}$$

we have

$$d^T \bigtriangledown^2 f(x^*)d + \sum_{j=1}^{m} \nu_j d^T \bigtriangledown^2 g_j(x^*)d \geq 0. \tag{38}$$

Theorem 3.1. Let the sequence $\{y_k\} \subset R^m$ be bounded and $0 < r_k \uparrow +\infty$. Let each $x_k^* \in R^n$ be generated by some method for solving $(P^{\alpha}_{y_k,r_k})$ (or $(\bar{P}^{\alpha}_{y_k,r_k})$). Assume that there exist $m_0, M \in R$ such that $f(x_k^*) \geq m_0, \quad \forall k$ and $L^{\alpha}(x_k^*, y_k, r_k) \leq M, \quad \forall k$ (or $\bar{L}^{\alpha}(x_k^*, y_k, r_k) \leq M, \quad \forall k$). Then every limit point x^* of $\{x_k^*\}$ is feasible to the original constrained program (CP). Let x^* be a limit point of $\{x_k^*\}$ and suppose that $\{\bigtriangledown g_j(x^*) : j = 1, \ldots, m\}$ are linearly independent. Further assume that each x_k^* satisfies the necesary optimality conditions presented in Proposition 3.1 (or 3.2). Then x^* satisfies the first-order and second-order necessary optimality conditions of (CP).

Proof. We only prove the case of $(P^{\alpha}_{y,r})$. The case of $(\bar{P}^{\alpha}_{y,r})$ can be analogously proved.

Without loss of generality, suppose that $x_k^* \to x^*$. Note that $L^{\alpha}(x_k^*, y_k, r_k) \leq M, \quad \forall k$, namely,

$$f(x_k^*) + \sum_{j=1}^{m} y_j^k g_j(x_k^*) + r_k \left[\sum_{j=1}^{m} |g_j(x_k^*)| \right]^{\alpha} \leq M.$$

Moreover, $\{y_k\}$ is bounded and $f(x_k^*) \geq m_0$. Consequently, there exists $M_1 > 0$ such that

$$r_k \left[\sum_{j=1}^{m} |g_j(x_k^*)| \right]^{\alpha} \leq M_1.$$

Thus,

$$\left[\sum_{j=1}^{m} |g_j(x_k^*)| \right]^{\alpha} \leq M_1/r_k.$$

Passing to the limit as $k \to +\infty$, we see that

$$g_j(x^*) = 0, \quad j = 1, \ldots, m.$$

Therefore, x^* is a feasible solution to (CP). As each x_k^* satisfies optimality conditions in Proposition 3.1, so there exist $\nu_j^k, j = 1, \ldots, m$ such that

$$\bigtriangledown f(x^k) + \sum_{j=1}^{m} \nu_j^k \bigtriangledown g_j(x^k) = 0; \tag{39}$$

and for any $d \in R^n$ satisfying

$$\bigtriangledown g_j(x^k)d = 0, \quad j = 1, \ldots, m, \tag{40}$$

there holds

$$d^T \bigtriangledown^2 f(x^k)d + \sum_{j=1}^{m} \nu_j^k d^T \bigtriangledown^2 g_j(x^k)d \geq 0. \tag{41}$$

Arguing as in the proof of Proposition 3.1, we can prove that each sequence $\{\nu_j^k\}, j = 1, \ldots, m$ is bounded. Assume without loss of generality that

$$\nu_j^k \to \nu_j, \quad j = 1, \ldots, m. \tag{42}$$

Taking the limit in (39) while observing (42), we obtain (36). Now for any $d \in R^n$ satisfying (37), since $x_k^* \to x^*$ and $\{\bigtriangledown g_j(x^*) : j = 1, \ldots, m\}$ are linearly independent, by Lemma 3.2, we have $\{d_k\} \subset R^n$ such that $d_k \to d$ and (40)holds with d replaced by d_k. Hence, (41) holds with d replaced by d_k. That is,

$$d_k^T \bigtriangledown^2 f(x^k)d_k + \sum_{j=1}^{m} \nu_j^k d_k^T \bigtriangledown^2 g_j(x^k)d_k \geq 0. \tag{43}$$

Taking the limit in (43) as $k \to +\infty$, we obtain (38). The proof is complete. □

Remark 3.2. The linear independence assumption of $\{g_j(x^*), j = 1, \ldots, m\}$ is important to guarantee the correctness of Theorem 3.1. Otherwise, Theorem 3.1 may fail.

Example 3.1. Consider the following constrained program

$$\begin{array}{lll} \text{(CP)} \qquad & \min & f(x) = x, \\ & \text{s.t.} & x \in R^1, \\ & & g(x) = x^2 = 0. \end{array}$$

Let $y_k \equiv 0, \quad \forall k$ and $0 < r_k \uparrow +\infty$. Consider problem $(P^1_{y_k,r_k})$:

$$\min\ x + r_k x^2 \quad \text{s.t.} \quad x \in R^1.$$

Clearly, the optimal solution of $(P^1_{y_k,r_k})$ is $x_k = -\frac{1}{2r_k}$, which converges to $x^* = 0$ as $k \to +\infty$. Since $\bigtriangledown g(x_k) = 2x_k = -1/r_k \neq 0$, the linear independence assumption of Proposition 3.1 holds. Moreover, Let $\nu^k = r_k, \quad \forall k$. Then, $\bigtriangledown f(x_k) + \nu^k \bigtriangledown g(x_k) = 0$ and for any $d \in R^1$, $d^T \bigtriangledown^2 f(x_k)d + \nu^k d^T \bigtriangledown^2 g(x_k)d = 2r_k d^2 \geq 0$. It is easily checked that all the other conditions (except the linear independence condition of $\{g(x^*)\}$) of Theorem 3.1 also hold. However, it is routine to check that x^* is not a KKT point of (CP), i.e., x^* does not satisfy the first-order necessary optimality condition given in Definition 3.1.

4 Convergence of a Class of Approximate Augmented Lagrangian Methods

As shown in Section 3, the nondifferentiable generalized augmented Lagrangian methods (when $0 < \alpha < 1$) are convergent. However, the functions L^α and $\bar{L}^\alpha$ are not even locally Lipschitz when $0 < \alpha < 1$. Current algorithms for unconstrained optimization problems are not applicable to minimize L^α and $\bar{L}^\alpha$. In this section, we shall smooth L^α and $\bar{L}^\alpha$ by approximation. We prove that the smooth approximate problems corresponding to L^α and $\bar{L}^\alpha$ preserve the convergence properties obtained in Section 3.

Let $0 < \epsilon_k \downarrow 0$. Set

$$L^\alpha(x,y,r,\epsilon_k) = f(x) + \sum_{j=1}^m y_j g_j(x) + r\left[\sum_{j=1}^m \sqrt{g_j^2(x)+\epsilon_k^2}\right]^\alpha, \quad x \in R^n.$$

$$\bar{L}^\alpha(x,y,r,\epsilon_k) = f(x) + \sum_{j=1}^m y_j g_j(x) + r\left[\sum_{j=1}^m (g_j^2(x)+\epsilon_k^2)^k\right]^{\frac{\alpha}{2k}}, \quad x \in R^n.$$

Clearly, both $L^\alpha(x,y,r,\epsilon_k)$ and $\bar{L}^\alpha(x,y,r,\epsilon_k)$ are twice continuously differentiable.

Consider the following smooth problems:

$$(P^\alpha_{y,r,\epsilon_k}) \quad \min_{x\in R^n} L^\alpha(x,y,r,\epsilon_k), \tag{44}$$

and

$$(\bar{P}^\alpha_{y,r,\epsilon_k}) \quad \min_{x\in R^n} \bar{L}^\alpha(x,y,r,\epsilon_k). \tag{45}$$

It is routine to derive the next two propositions.

Proposition 4.1. Suppose that $\bar{x}$ is a local minimum of $(P^\alpha_{y,r,\epsilon_k})$. Then

$$\begin{aligned}\bigtriangledown_x L^\alpha(\bar{x},y,r,\epsilon_k) \quad &= \quad \bigtriangledown f(\bar{x}) + \sum_{j=1}^m \{y_j + r\alpha\left[\sum_{j=1}^m \sqrt{g_j^2(\bar{x})+\epsilon_k^2}\right]^{\alpha-1} \\ &\qquad \sum_{j=1}^m \left[g_j^2(\bar{x})+\epsilon_k^2\right]^{-1/2} g_j(\bar{x})\} \bigtriangledown g_j(\bar{x}) = 0,\end{aligned} \tag{46}$$

and

$$\begin{aligned}&d^T \bigtriangledown_x^2 L^\alpha(\bar{x},y,r,\epsilon_k)d = d^T \bigtriangledown^2 f(\bar{x})d + \sum_{j=1}^m \nu_j^k d^T \bigtriangledown^2 g_j(\bar{x})d \\ &+r\alpha(\alpha-1)\sum_{j=1}^m \left[\sum_{j=1}^m \sqrt{g_j^2(\bar{x})+\epsilon_k^2}\right]^{\alpha-2} \left[g_j^2(\bar{x})+\epsilon_k^2\right]^{-1} g_j^2(\bar{x})(\bigtriangledown g_j(\bar{x})d)^2 \\ &+r\alpha\sum_{j=1}^m \left[\sum_{j=1}^m \sqrt{g_j^2(\bar{x})+\epsilon_k^2}\right]^{\alpha-1} \left[g_j^2(\bar{x})+\epsilon_k^2\right]^{-3/2} g_j^2(\bar{x})(\bigtriangledown g_j(\bar{x})d)^2 \\ &+r\alpha\sum_{j=1}^m \left[\sum_{j=1}^m \sqrt{g_j^2(\bar{x})+\epsilon_k^2}\right]^{\alpha-1} \left[g_j^2(\bar{x})+\epsilon_k^2\right]^{-1/2} (\bigtriangledown g_j(\bar{x})d)^2 \geq 0, \quad d \in R^n\end{aligned} \tag{47}$$

Proposition 4.2. Suppose that $\bar{x}$ is a local minimum of $(\bar{P}^{\alpha}_{y,r,\epsilon_k})$. Then

$$\begin{aligned}\nabla_x \bar{L}^{\alpha}(\bar{x},y,r,\epsilon_k) &= \nabla f(\bar{x}) + \sum_{j=1}^{m}\{y_j + r\alpha\left[\sum_{j=1}^{m}\left(g_j^2(\bar{x})+\epsilon_k^2\right)^k\right]^{\frac{\alpha}{2k}-1} \\ &\qquad \sum_{j=1}^{m}\left(g_j^2(\bar{x})+\epsilon_k^2\right)^{k-1} g_j(\bar{x})\} \nabla g_j(\bar{x}) = 0, \end{aligned} \tag{48}$$

and

$$\begin{aligned} d^T \nabla_x^2 \bar{L}^{\alpha}(\bar{x},y,r,\epsilon_k)d &= d^T \nabla^2 f(\bar{x})d + \sum_{j=1}^{m}\{y_j + r\alpha\left[\sum_{j=1}^{m}\left(g_j^2(\bar{x})+\epsilon_k^2\right)^k\right]^{\frac{\alpha}{2k}-1} \\ &\quad \sum_{j=1}^{m}\left(g_j^2(\bar{x})+\epsilon_k^2\right)^{k-1} g_j(\bar{x})\} d^T \nabla^2 g_j(\bar{x})d \\ &\quad +\{r(1-2k)\left[\sum_{j=1}^{m}\left(g_j^2(\bar{x})+\epsilon_k^2\right)^k\right]^{\frac{\alpha}{2k}-2} \left[\sum_{j=1}^{m}\left(g_j^2(\bar{x})+\epsilon_k^2\right)^{k-1}\right]^2 g_j^2(\bar{x}) \\ &\quad +r\left[\sum_{j=1}^{m}\left(g_j^2(\bar{x})+\epsilon_k^2\right)^k\right]^{\frac{\alpha}{2k}-1} \sum_{j=1}^{m}\left[2(k-1)\left(g_j^2((\bar{x})+\epsilon_k^2\right)^{k-2} g_j^2(\bar{x})\right. \\ &\quad \left. + \left(g_j^2((\bar{x})+\epsilon_k^2\right)^{k-1}\right]\}(\nabla g_j(\bar{x})d)^2 \geq 0, \quad d \in R^n. \end{aligned} \tag{49}$$

Theorem 4.1. Let the sequence $\{y_k\} \subset R^m$ be bounded, $0 < r_k \uparrow +\infty$ and $0 < \epsilon_k \downarrow 0$. Let each $x_k^* \in R^n$ be generated by some method for solving $(P^{\alpha}_{y_k,r_k,\epsilon_k})$ (or $(\bar{P}^{\alpha}_{y_k,r_k,\epsilon_k})$). Assume that there exist $m_0, M \in R$ such that $f(x_k^*) \geq m_0, \quad \forall k$ and $L^{\alpha}(x_k^*, y_k, r_k, \epsilon_k) \leq M, \quad \forall k$ (or $\bar{L}^{\alpha}(x_k^*, y_k, r_k, \epsilon_k) \leq M, \quad \forall k$). Then every limit point x^* of $\{x_k^*\}$ is feasible to the original constrained program (CP). Let x^* be a limit point of $\{x_k^*\}$ and suppose that $\{\nabla g_j(x^*) : j = 1, \ldots, m\}$ are linearly independent. Further assume that each x_k^* satisfies the necesary optimality conditions presented in Proposition 4.1 (or 4.2). Then x^* satisfies the first-order and second-order necessary optimality conditions of (CP).

Proof. The proof is almost the same as that of Theorem 3.1 and thus omited. □

References

[1] D.P. Bertsekas, *Constrained Optimization and Lagrangian Multiplier Methods*, Academic Press, New York, 1982.

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

[3] J.V. Burke, An exact penalization viewpoint of constrained optimization, *SIAM Journal on Control and Optimization* 29 (1991) 968–998.

[4] F.H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley, New York, 1983.

[5] G.B. Dantzig, J. Folkman and N. Shapiro, On continuity of the minimum set of a continuous function, *Journal of Mathematical Analysis and Applications* 17 (1967) 519–548.

[6] G. Di Pillo and L. Grippo, An augmented Lagrangian for inequality constraints in nonlinear programming problems, *Journal of Optimization Theory and Applications* 36 (1982) 495–519.

[7] G. Di Pillo and S. Lucidi, On exact augmented Lagrangian functions in nonlinear programming, in *Nonlinear Optimization and Applications*, G. Di Pillo and F. Giannessi (eds.), Plenum Press, New York, 1996, pp. 85–100.

[8] G. Di Pillo and S. Lucidi, An augmented Lagrangian function with improved exactness properties, *SIAM Journal on Optimization* 12 (2001) 376–406.

[9] M.R. Hestenes, Multiplier and gradient methods, *Journal of Optimization Theory and Applications* 4 (1969) 303–320.

[10] X.X. Huang and X.Q. Yang, A unified augmented Lagrangian approach to duality and exact penalization, *Mathematics of Operations Research* 28 (2003) 533–522.

[11] X.X. Huang and X.Q. Yang, Further study on augmented Lagrangian duality theory, Submitted for publication.

[12] Z.Q. Luo, J.S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, 1996.

[13] J.S. Pang, Error bounds in mathematical programming, *Math. Program.* 79 (1997) 299–332.

[14] M.J.D. Powell, A method for nonlinear constraints in minimization problems, in *Optimization*, R. Fletcher (ed.), Academic Press, New York, 1969, pp. 283–298.

[15] R.T. Rockafellar, Augmented Lagrangian multiplier functions and duality in nonconvex programming, *SIAM Journal on Control and Optimization* 12 (1974) 268–285.

[16] R.T. Rockafellar, Lagrangian multipliers and optimality, *SIAM Review* 35 (1993) 183–238.

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

[18] A.M. Rubinov, B.M. Glover and X.Q. Yang, Decreasing functions with applications to penalization, *SIAM Journal on Optimization* 10 (1999) 289–313.

[19] A.M. Rubinov, X.Q. Yang and A. Bagirov, Penalty functions with a small penalty parameter, *Optimization Methods and Software* 17 (2002) 931–964.

[20] X.Q. Yang, An exterior method for computing points that satisfy second-order necessary conditions for a $C^{1,1}$ optimization problem, *Journal of Mathematical Analysis and Applications* 187 (1994) 118–133.

Manuscript received 4 June 2004
revised 26 October 2004
accepted for publication 28 October 2004

X.X. Huang
Department of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047, China
E-mail address: `xuexianghuang@yahoo.com.cn`

X.Q. Yang
Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong
E-mail address: `mayangxq@polyu.edu.hk`