Journal of Global Optimization (2005) 31:193–210

Further Study on Augmented Lagrangian Duality Theory*

X. X. HUANG[1,2] and X. Q. YANG[2]
[1]*Department of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047, P.R. China*
[2]*Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong*

(Received 30 March 2004; accepted in final form 1 April 2004)

Abstract. In this paper, we present a necessary and sufficient condition for a zero duality gap between a primal optimization problem and its generalized augmented Lagrangian dual problems. The condition is mainly expressed in the form of the lower semicontinuity of a perturbation function at the origin. For a constrained optimization problem, a general equivalence is established for zero duality gap properties defined by a general nonlinear Lagrangian dual problem and a generalized augmented Lagrangian dual problem, respectively. For a constrained optimization problem with both equality and inequality constraints, we prove that first-order and second-order necessary optimality conditions of the augmented Lagrangian problems with a convex quadratic augmenting function converge to that of the original constrained program. For a mathematical program with only equality constraints, we show that the second-order necessary conditions of general augmented Lagrangian problems with a convex augmenting function converge to that of the original constrained program.

Key words: augmented Lagrangian, constraint qualification, optimality condition, perturbation function, zero duality gap

1. Introduction

Augmented Lagrangian with a convex quadratic augmenting function was formally introduced by Rockafellar [11, 12] to eliminate the duality gap between the primal constrained optimization problem and its conventional (linear) Lagrangian dual problem. Augmented Lagrangian method has been widely and successfully used in the solution of constrained optimization problems (see, e.g. [1]). Moreover, as noted in Teo et al [16], any constrained optimal control problem can be reduced to a mathematical programming problem by using the control parametrization technique. As a result, augmented Lagrangian method can be employed to solve constrained optimal control problems. Recent interesting applications of augmented Lagrangian to the study of linear programming can be found

*This research is supported by the Research Grants Council of Hong Kong (PolyU B-Q359.)

in [17, 8, 10]. More recently, a general (convex) augmented Lagrangian was studied in [13]. Under suitable conditions, a zero duality gap theorem was established without any convexity assumption (see [13], Theorem 11.59). A necessary and sufficient condition for the exact penalty representation in the framework of the augmented Lagrangian was also obtained, see ([13], Theorem 11.61). Rubinov et al. [14] introduced nonlinear Lagrangians for a constrained optimization problem. Under mild conditions, a zero duality gap property was derived between the original constrained program and its nonlinear Lagrangian dual problem. In [18], for a constrained program, under some conditions, an equivalence in terms of zero duality gap properties was established between a class of nonlinear Lagrangian dual problems and a class of augmented Lagrangian dual problems. In [9], generalized augmented Lagrangian was introduced by relaxing the convexity on the augmenting function. Zero duality gap and exact penalization results were established under weaker conditions than those of [13]. For a constrained optimization problem, under weaker conditions than those of [18], an equivalence in terms of zero duality gap properties was obtained between a class of generalized augmented Lagrangian dual problems and a class of nonlinear Lagrangian dual problems.

Another direction that is worth noting in the study of augmented Lagrangian is the so-called exact augmented Lagrangian for inequality constrained nonlinear programming (see, e.g. [5–7]). Unlike the augmented Lagrangian we mentioned above in which the penalty term only considers the feasibility of 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 [5–7]).

In this paper, we shall not discuss exact augmented Lagrangian. Our attention will be restricted to the augmented Lagrangian discussed in [11–13, 9]. So, whenever we mention the term ‘‘augmented Lagrangian’’, it should refer to the augmented Lagrangian discussed in [11–13, 9].

In [11], in the context of a constrained optimization problem, a necessary and sufficient condition for a zero duality gap of the quadratic augmented Lagrangian dual problem was given in terms of the lower semicontinuity of a perturbation function (see ([11], Theorems 2 and 4)). Most recently, Rubinov et al. [15] introduced a very general augmented Lagrangian and nonlinear Lagrangian and established necessary and sufficient conditions for the zero duality gap properties based on these two types of Lagrangian functions in terms of the lower semicontinuity of certain perturbation functions, respectively. In this paper, a necessary and sufficient condition for the zero duality gap property via a class of generalized augmented Lagrangians will be given in the form of the lower semicontinuity of a perturbation function at the origin. This result combined with the necessary and sufficient condition for a zero duality gap between a constrained program and a nonlinear Lagrangian dual problem improve the equivalence result in terms of zero duality gap property between the nonlinear Lagrangian dual problem and a class of generalized augmented Lagrangian dual problems for a constrained program established in [9].

Despite the popularity of augmented Lagrangian in solving constrained optimization problems, it is worth noting that the convergence of second-order necessary optimality conditions of augmented Lagrangian problems to that of the original constrained optimization problem has never been investigated. Note that for a nonconvex program, conventional optimization methods only generate points that satisfy (first-order or second-order) necessary optimality conditions. Thus, it is both interesting and important to consider the convergence of optimality conditions of the augmented Lagrangian problems.

In this paper, in the context of a mathematical program with both equality and inequality constraints, we shall carry out convergence analysis of the first-order and second-order necessary conditions for the augmented Lagrangian with a convex quadratic augmenting function considered in [11, 12]. In the context of a mathematical program with only equality constraints, we show that the second-order necessary conditions of the augmented Lagrangian problems converge to that of the original constrained program for general augmented Lagrangians with convex augmenting functions. It should be mentioned that in both cases, explicit expressions of these augmented Lagrangians can be derived. However, there is generally no explicit expression for the augmented Lagrangian with a general convex augmenting function if the constrained program contains both equality and inequality constraints. So there exists some technical difficulty in carrying out the convergence analysis for a general augmented Lagrangian with a convex augmenting function if the constrained program contains both equality and inequality constraints. Hence, this paper concentrates on the convergence analysis for the above two cases where the augmented Lagrangian can be explicitly written down.

2. Zero Duality Gap Via Generalized Augmented Lagrangians

Consider the following primal optimization problem:

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

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

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

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

Let $\sigma{:}R^m \to \bar{R}$ be a generalized augmenting function, i.e. it is proper, lsc, level-bounded (the set $\{u \in R^m : \sigma(u) \leqslant \alpha\}$ is always bounded for any $\alpha \in R$), and attains its minimum 0 only at the origin $0 \in R^m$.

The generalized augmented Lagrangian is defined as

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

The generalized augmented Lagrangian dual function is

$$\psi(y,r) = \inf_{x \in R^n} l(x,y,r), \quad y \in R^m, \ \ r > 0.$$

The generalized augmented Lagrangian dual problem is

$$\text{(D)} \qquad \sup_{(y,r) \in R^m \times (0,+\infty)} \psi(y,r).$$

Denote by M_{P} and M_{D} the optimal values of (P) and (D), respectively. It is clear from [9] that weak duality holds

$$M_{\mathrm{D}} \leqslant M_{\mathrm{P}}.$$

Consequently, if $M_{\mathrm{P}} = -\infty$, then $M_{\mathrm{D}} = M_{\mathrm{P}} = -\infty$. So we shall assume that $M_{\mathrm{P}} > -\infty$ in the study of necessary and sufficient conditions for $M_{\mathrm{P}} = M_{\mathrm{D}}$ to hold.

Define the perturbation function

$$p(u) = \inf_{x \in R^n} f(x,u), \quad u \in R^m.$$

Obviously, $p(0) = M_{\mathrm{P}}$.

THEOREM 2.1. *Suppose that $M_{\mathrm{P}} > -\infty$. Then the zero duality gap property $M_{\mathrm{P}} = M_{\mathrm{D}}$ holds iff*

(a) *the perturbation function $p(u)$ is lsc at $0 \in R^m$,*

(b) *there exist $\bar{y}, \bar{r}$ such that*

$$\psi(\bar{y},\bar{r}) = \inf_{x \in R^n} l(x,\bar{y},\bar{r}) > -\infty.$$

((b) $\Leftrightarrow MD > -\infty$).

Proof. The conclusion follows immediately from Theorems 2.2–2.4 and Proposition 2.2 in [15]. $\square$

Remark. 2.1. A sufficient condition that guarantees (b) is that $f(\cdot,\cdot)$ is bounded below on $R^n \times R^m$. In this case, we can take $\bar{y}=0$ and any $r>0$ as $\bar{r}$.

Consider the following constrained program:

$$\begin{aligned}
\text{(CP)}\quad & \inf f_0(x), \\
& \text{s.t.} \in, \\
& \quad g_j(x) \leqslant 0, \quad j=1,\ldots,m_1, \\
& \quad g_j(x) = 0, \quad j=m_1+1,\ldots,m,
\end{aligned}$$

where X is a subset of R^n, $f_0, g_j\colon X \to R$, $j=1,\ldots,m$ are real-valued functions. Denote by X_0 the feasible set of (CP), i.e.

$$X_0 = \{x \in X \,:\, g_j(x) \leqslant 0, j=1,\ldots,m_1, g_j(x)=0, j=m_1+1,\ldots,m\}$$

and by M_{CP} the optimal value of problem (CP).

Let

$$\varphi(x) = \begin{cases} f_0(x) & \text{if } x \in X_0, \\ +\infty & \text{if } x \in R^n \backslash X_0. \end{cases} \tag{1}$$

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

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

Define the dualizing parameterization function

$$f_{\mathrm{CP}}(x,u) = f_0(x) + \delta_{R_-^{m_1} \times \{0_{m-m_1}\}}(G(x)+u) + \delta_X(x), \quad x \in R^n, \quad u \in R^m, \tag{2}$$

where 0_{m-m_1} is the origin of R^{m-m_1}, $G(x) = (g_1(x),\ldots,g_m(x))$, and δ_{D} is the indicator function of the set D, i.e.

$$\delta_{\mathrm{D}}(z) = \begin{cases} 0 & \text{if } z \in \mathrm{D}, \\ +\infty & \text{else}. \end{cases}$$

Thus, a class of generalized augmented Lagrangians for (CP) with the dualizing parameterization function f_{CP} defined by (2) can be expressed as

$$l_{\mathrm{CP}}(x,y,r) = \inf\{f_{\mathrm{CP}}(x,u) - \langle y,u\rangle + r\sigma(u) \,:\, u \in R^m\}, \tag{3}$$

where σ is a generalized augmenting function. The corresponding generalized augmented Lagrangian dual function is

$$\psi_{\mathrm{CP}}(y,r) = \inf\{l_{\mathrm{CP}}(x,y,r) \,:\, x \in R^n\}, \quad y \in R^m, \quad r \in (0,+\infty). \tag{4}$$

The corresponding generalized augmented Lagrangian dual problem is

$$(\mathrm{D}_1) \qquad \sup_{(y,r)\in R^m\times(0,+\infty)} \psi_{\mathrm{CP}}(y,r). \tag{5}$$

The optimal value of (D_1) is denoted by M_{D_1}.

Define the perturbation function

$$\begin{aligned} p_1(u) &= \inf\{f_{\mathrm{CP}}(x,u) \,:\, x\in R^n\} \\ &= \inf\{f_0(x) \,:\, x\in X, g_j(x)+u_j\leqslant 0, \\ &\quad j=1,\ldots,m_1, g_j(x)+u_j=0, j=m_1+1,\ldots,m\}, \end{aligned}$$

where $u=(u_1,\ldots,u_m)$.

Clearly, $p_1(0)=M_{\mathrm{CP}}$.

Now we apply Theorem 2.1 to (P′) and the generalized augmented Lagrangian dual problem (D_1) and obtain the following result concerning a necessary and sufficient condition for the zero duality gap between the constrained program (CP) and its generalized augmented Lagrangian dual problem (D_1).

THEOREM 2.2. *Consider the problem* (CP) *and its generalized augmented Lagrangian dual problem* (D_1). *Suppose that* $M_{\mathrm{CP}}>-\infty$. *Then the zero duality gap property* $M_{\mathrm{CP}}=M_{\mathrm{D}_1}$ *holds iff*

(a) $p_1(u)$ *is lsc at* $u=0\in R^m$;

(b) *there exist* $\bar{y},\bar{r}$ *such that*

$$\psi_{\mathrm{CP}}(\bar{y},\bar{r})>-\infty.$$

((b) $\Leftrightarrow M_{\mathrm{D}_1}>-\infty$).

Remark 2.2. If f_0 is bounded below on X, then (b) holds. In this case, $M_{\mathrm{CP}}=M_{\mathrm{D}_1}$ iff $p_1(u)$ is lsc at $0\in R^m$, which has nothing to do with the choice of the generalized augmenting function σ. In other words, if f_0 is bounded below on X, and any one of the generalized augmented Lagrangian dual problem yields a zero duality gap, then all the generalized augmented Lagrangian dual problems yield a zero duality gap regardless of the choice of the generalized augmenting function σ.

To compare the zero duality gap property of the generalized augmented Lagrangian dual problem with that of the nonlinear Lagrangian dual problem, we recall the nonlinear Lagrangian for (CP) and its nonlinear Lagrangian dual problem (for details, see [14, 9]).

A basic assumption in the definition of a nonlinear Lagrangian is

$$f_0(x)\geqslant 0, \quad \forall x\in X.$$

Let $c: R_+ \times R^{m_1} \times R_+^{m-m_1} \to R$ be a real-valued function. c is said to be *increasing* if, for any $y^1, y^2 \in R_+ \times R^{m_1} \times R_+^{m-m_1}, y^2 - y^1 \in R_+^{m+1}$ implies that $c(y^1) \leqslant c(y^2)$.

Consider increasing and lsc functions c defined on $R_+ \times R^{m_1} \times R_+^{m-m_1}$ having the following properties:

(A) $\exists a_j > 0, j = 1, \ldots, m$ such that, for any $y = (y_0, y_1, \ldots, y_{m_1}, y_{m_1+1}, \ldots, y_m) \in R_+ \times R^{m_1} \times R_+^{m-m_1}$, we have

$$c(y) \geqslant \max\{y_0, a_1 y_1, \ldots, a_{m_1} y_{m_1}, a_{m_1+1} y_{m_1+1}, \ldots, a_m y_m\}.$$

(B) For any $y_0 \in R_+$, we have $c(y_0, 0, \ldots, 0) = y_0$.

Let c be an increasing function with properties (A) and (B), and

$$F(x, d) = (f(x), d_1 g_1(x), \ldots, d_{m_1} g_{m_1}(x), d_{m_1+1}|g_{m_1+1}(x)|, \ldots, d_m|g_m(x)|),$$

where $d = (d_1, \ldots, d_m) \in R_+^m$ and $x \in X$.

The function defined by

$$L(x, d) = c(F(x, d)), \quad x \in X, \ d \in R_+^m$$

is called a *nonlinear Lagrangian* corresponding to c.

The nonlinear Lagrangian dual function corresponding to c is defined as

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

The nonlinear Lagrangian dual problem is defined by

$$(\mathrm{D}_2) \qquad \sup_{d \in R_+^m} \phi(d). \tag{6}$$

Denote by M_{D_2} the optimal value of problem (D_2). Let $u = (u_1, \ldots, u_m) \in R^m$. Denote

$$X_u = \{x \in X : g_j(x) + u_j \leqslant 0, j = 1, \ldots, m_1, g_j(x) + u_j = 0, \\ j = m_1 + 1, \ldots, m\}.$$

The following result is concerned with a necessary and sufficient condition for a zero duality gap between (CP) and its nonlinear Lagrangian dual problem (D_2). For the proof, we refer readers to [15].

THEOREM 2.3. *Consider the constrained program* (CP) *and its nonlinear Lagrangian dual problem* (D_2). *Assume that* $f_0(x) \geqslant 0, \forall x \in X$. *Suppose that the increasing function c defining the nonlinear Lagrangian is continuous. Then a necessary and sufficient condition for the zero duality gap* $M_{\mathrm{CP}} = M_{\mathrm{D}_2}$ *to hold is that* $p_1(u)$ *is lsc at* $u = 0 \in R^n$.

The following equivalence result follows immediately from Theorems 2.2 and 2.3.

THEOREM 2.4. *Assume that* $f_0(x) \geqslant 0, \forall x \in X$. *Consider the constrained program* (CP) *and its generalized augmented Lagrangian dual problem* (D_1) *and its nonlinear Lagrangian dual problem* (D_2). *Suppose that the increasing function c defining the nonlinear Lagrangian is continuous. Then the following two statements are equivalent*:

(a) *the generalized augmented Lagrangian dual problem* (D_2) *yields a zero duality gap*;
(b) *the nonlinear Lagrangian dual problem* (D_2) *yields a zero duality gap.*

Remark 2.3. In [9], the equivalence between zero duality gaps in terms of generalized augmented Lagrangian and the nonlinear Lagrangian was established under the continuity of the generalized augmenting function. Thus, Theorem 2.4 improves this equivalence result.

3. Convergence of Optimality Conditions

In this section, we consider the constrained optimization problem (CP). We assume that $X = R^n$ and all the functions involved in (CP) are twice continuously differentiable. We shall discuss the convergence of second-order necessary conditions of the augmented Lagrangian problems in two cases:

(a) The CP contains both equality and inequality constraints and the augmenting function $\sigma(u) = 1/2\Sigma_{j=1}^{m} u_j^2$ is used. The corresponding augmented Lagrangian is called a proximal Lagrangian (see [13]).
(b) The CP contains only equality constraints and the augmenting function $\sigma(\mathrm{u})$ can be any convex function.

3.1. CONVERGENCE ANALYSIS FOR PROXIMAL LAGRANGIAN PROBLEMS

Suppose that $\{y^k\} \subset R^m$ is bounded and $0 < r_k$. Consider the following proximal Lagrangian problems:

$$(\mathrm{P}_\mathrm{k}^1) \qquad \inf_{\mathrm{x} \in R^\mathrm{n}} l_2(x, y^k, r_k),$$

where

$$l_2(x,y,r) = f_0(x) + \sum_{j=1}^{m_1} \begin{cases} y_j g_j(x) + \frac{r}{2} g_j^2(x) & \text{if } \; g_j(x) \geqslant -y_j/r \\ \frac{-y_j^2}{2r} & \text{if } \; g_j(x) < -y_j/r \end{cases} + \sum_{j=m_1+1}^{m} \left[y_j g_j(x) + \frac{r}{2} g_j^2(x) \right], \quad r > 0. \tag{7}$$

Let $\bar{x} \in X_0$. We denote

$$J_1(\bar{x}) = \{j : g_j(\bar{x}) = 0, j = 1, \ldots, m_1\}.$$

We say that the linear independence constraint qualification (LICQ) for (CP) holds at $\bar{x}$ if $\{\nabla g_j(\bar{x}) : j \in J_1(\bar{x})\} \cup \{\nabla g_j(\bar{x}) : j = m_1 + 1, \ldots, m\}$ are linearly independent.

Suppose that $\bar{x} \in R^n$ is a local optimal solution to (CP) and the LICQ holds for (CP) at $\bar{x}$. Then, the first-order necessary optimality condition is that $\exists \mu_j \geqslant 0, j \in J_1(\bar{x})$ and $\mu_j, j = m_1 + 1, \ldots, m$ such that

$$\nabla f_0(\bar{x}) + \sum_{j \in J_1(\bar{x})} \mu_j \nabla g_j(\bar{x}) + \sum_{j=m_1+1}^{m} \mu_j \nabla g_j(\bar{x}) = 0 \tag{8}$$

and the second-order necessary optimality condition is that the first-order necessary condition holds, and for any $\bar{d} \in R^n$ satisfying

$$\begin{aligned} &\nabla g_j(\bar{x})\bar{d} = 0, \quad j \in J_1(\bar{x}), \\ &\nabla g_j(\bar{x})\bar{d} = 0, \quad j = m_1 + 1, \ldots, m, \end{aligned} \tag{9}$$

we have

$$\bar{d}^T \nabla^2 f_0(\bar{x})\bar{d} + \sum_{j \in J_1(\bar{x})} \mu_j \bar{d}^T \nabla^2 g_j(\bar{x})\bar{d} + \sum_{j=m_{1+1}}^{m} \mu_j \bar{d}^T \nabla^2 g_j(\bar{x})\bar{d} \geqslant 0. \tag{10}$$

It is known from [12] that l_2 may be only $C^{1,1}$ (the first-order derivative is only locally Lipschitz) in x no matter how smooth the functions involved in (CP) are. In what follows, we first derive the first-order and second-order necessary optimality conditions for (P_k^1).

Suppose that x^k is a local optimal solution of (P_k^1). Denote

$$\begin{aligned} J_1^{+k} &= \{j : g_j(x^k) > -y_j^k / r_k, j = 1, \ldots, m_1\}, \\ J_1^k &= \{j : g_j(x^k) = -y_j^k / r_k, j = 1, \ldots, m_1\}. \end{aligned}$$

The following first-order condition for (P_k^1) can be straightforwardly derived.

LEMMA 3.1 (First-order condition). *Let $x^k \in R^n$ be a local optimal solution of (P_k^1). Then*

$$\begin{aligned} \nabla_x l_2(x^k, y^k, r_k) = \nabla f_0(x^k) + &\sum_{j \in J_1^{+k} \cup J_1^k} (y_j^k + r_k g_j(x^k)) \nabla g_j(x^k) \\ &\sum_{j=m_1+1}^{m} (y_j^k + r_k g_j(x^k)) \nabla g_j(x^k) = 0. \end{aligned} \tag{11}$$

The following definition of second-order directional derivative was introduced by Ben-Tal and Zowe in [2, 3] (see also [19]).

DEFINITION 3.1. Let $h : R^n \to R$ be a continuously differentiable function. The second-order directional derivative of h at x in the direction d is defined as

$$h''(x; d, d) = \lim_{s \to 0^+} s^{-2}(h(x + sd + s^2 d) - h(x) - s\nabla h(x)d)$$

if the right hand side limit exists.

LEMMA 3.2. *Let $\lambda_i \in R^1, i = 1, \dots, q$. Suppose that*

(i) $f_i\colon R^n \to R^1 (i = 1, \dots, q)$ *are* $C^{1,1}$ *and, for some positive integer* $q_1 \leqslant q, \dots f_i, i = 1, \dots, q_1$ *are* $\dots C_2$;

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

$$\nabla f_i(x)d = 0,\ i = q_1 + 1, \dots, q, \tag{12}$$

$f_i''(x; d, d)$ *exists,* $i = q_1 + 1, \dots, q$;

(iii) x *locally minimizes* $f = \sum_{i=1}^q \lambda_i f_i$ *over* R^n.

Then

(a)

$$\sum_{i=1}^q \lambda_i \nabla f_i(x) = 0;$$

(b) $\forall d \in R^n$,

$$\sum_{i=1}^{q_1} \lambda_i d^T \nabla^2 f_i(x) d + 2 \sum_{i=q_1+1}^{q} \lambda_i f_i''(x; d, d) \geqslant 0.$$

Proof. (a) is obvious.

(b) can be straightforwardly proved by applying Theorem 5.2 (i) of [19] and formula (1.1) in [19]. □

LEMMA 3.3. *Let $h : R^n \to R$ be twice continuously differentiable. Then, $(h^+)^2$ is $C^{1,1}$ and*

$$[(h^+)^2]''(x; d, d) = \begin{cases} 2h(x)\nabla h(x)d + 2h(x)(d^T\nabla^2 h(x)d) & \text{if } h(x) > 0, \\ 2(\nabla h(x)d)^2 & \text{if } h(x) = 0, \\ 0 & \text{otherwise .} \end{cases}$$

Proof. The conclusion follows from Proposition 3.3 in [3]. □

It is known from ([9], Example 2.1) that

$$
\begin{aligned}
l_2(x,y,r) = & f_0(x) + \frac{r}{2}\left[\sum_{j=1}^{m_1}\left(\frac{y_j}{r}+g_j(x)\right)^{+2} - \sum_{j=1}^{m_1}\left(\frac{y_j}{r}\right)^2\right] \\
& + \sum_{j=m_1+1}^{m}\left[y_j g_j(x) + \frac{r}{2}g_j^2(x)\right].
\end{aligned} \tag{13}
$$

LEMMA 3.4. (Second-order condition). *Suppose that x^k is a local optimal of* (P_k^1). *Then, the first-order necessary optimality condition* (11) *holds, and for any* $d \in R^n$

$$
\begin{aligned}
& d^T\nabla^2 f_0(x^k)d + \sum_{j\in J_1^{+k}}[r_k(\nabla g_j(x^k)d)^2 + (y_j^k + r_k g_j(x^k))d^T\nabla^2 g_j(x^k)d] \\
& + \sum_{j\in J_1^k}[(\nabla g_j(x^k)d)^+]^2 + \sum_{j=m_1+1}^{m}[r_k(\nabla g_j(x^k)d)^2 \\
& + (y_j^k + r_k g_j(x^k))d^T\nabla^2 g_j(x^k)d] \geqslant 0.
\end{aligned} \tag{14}
$$

Proof. Using Lemmas 3.2 and 3.3 and formula (13), the conclusion can be directly proved. □

The following lemma is needed in the derivation of the convergence of the second-order conditions.

LEMMA 3.5. *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 $k \geqslant \bar k$, there exists $d^k \in R^n$ satisfying $(c_i^k)^T d^k = 0$, $i = 1,\ldots,q$ and $d^k \to \bar d$.

Proof. It follows directly from Corollary II.3.4 of [4]. □

THEOREM 3.1 (Convergence of first-order conditions). *Suppose that* $\{y^k\} \subset R^m$ *is bounded,* $0 < r_k \to +\infty$, x^k *satisfies the first-order necessary optimality condition of* (P_k^1) *stated in Lemma* 3.1, *and* $x^k \to \bar x \in X_0$. *Furthermore, suppose that the LICQ for (CP) holds at* $\bar x$. *Then the first-order necessary condition for (CP) holds at* $\bar x$.

Proof. Since $x^k \to \bar x \in X_0$, we deduce that

$$J_1^{+k} \cup J_1^k \subset J_1(\bar x),$$

when k is sufficiently large. In the following, we assume that k is sufficiently large. Set

$$\begin{aligned} &\mu_j^k = y_j^k + r_k g_j(x^k), \quad j \in J_1(\bar{x}) \cap J_1^{+k}, \\ &\mu_j^k = 0, \quad j \in J_1(\bar{x}) \backslash J_1^{+k}, \\ &\mu_j^k = y_j^k + r_k g_j(x^k), \quad j = m_1 + 1, \ldots, m. \end{aligned} \tag{15}$$

Then

$$\mu_j^k \geqslant 0, \quad j \in J_1(\bar{x}), \tag{16}$$

and (11) becomes

$$\nabla f_0(x^k) + \sum_{j \in J_1(\bar{x})} \mu_j^k \nabla g_j(x^k) + \sum_{j=m_1+1}^{m} \mu_j^k \nabla g_j(x^k) = 0. \tag{17}$$

Now we prove by contradiction that the sequence $\{\sum_{j \in J(\bar{x})} \mu_j^k + \sum_{j=m_1+1}^{m} |\mu_j^k|\}$ is bounded. Otherwise, assume without loss of generality that

$$\sum_{j \in J(\bar{x})} \mu_j^k + \sum_{j=m_1+1}^{m} |\mu_j^k| \to +\infty$$

and

$$\frac{\mu_j^k}{\sum_{j \in J(\bar{x})} \mu_j^k + \sum_{j=m_1+1}^{m} |\mu_j^k|} \to \mu_j', \quad j \in J_1(\bar{x}) \cup \{m_1 + 1, \ldots, m\}.$$

From (16), it is clear that

$$\mu_j' \geqslant 0, \quad j \in J_1(\bar{x}).$$

Dividing (17) by $\sum_{j \in J(\bar{x})} \mu_j^k + \sum_{j=m_1+1}^{m} |\mu_j^k|$ and passing to the limit, we get

$$\sum_{j \in J_1(\bar{x})} \mu_j' \nabla g_j(\bar{x}) + \sum_{j=m_1+1}^{m} \mu_j' \nabla g_j(\bar{x}) = 0.$$

This contradicts the LICQ of (CP) at $\bar{x}$ since $\sum_{j \in J_1(\bar{x})} \mu_j' + \sum_{j=m_1+1}^{m} |\mu_j'| = 1$.

Hence, $\{\sum_{j \in J_1(\bar{x})} \mu_j^k + \sum_{j=m_1+1}^{m} |\mu_j^k|\}$ is bounded. Thus, without loss of generality, we assume that

$$\mu_j^k \to \mu_j, \quad j \in J_1(\bar{x}) \cup \{m_1 + 1, \ldots, m\}. \tag{18}$$

Clearly, from (16) we have

$$\mu_j' \geqslant 0, \quad j \in J_1(\bar{x}).$$

Taking the limit in (17) as $k \to +\infty$ and applying (18), we obtain the first-order necessary condition of (CP). □

THEOREM 3.2 (Convergence of second-order conditions). *Assume the same conditions as in Theorem* 3.1. *Furthermore, assume the second-order conditions stated in Lemma* 3.4 *hold. Then $\bar{x}$ satisfies the second-order necessary condition of* (CP).

Proof. First we note from Theorem 3.1 that $\bar{x}$ satisfies the first-order condition of (CP). Since the LICQ holds for (CP) at $\bar{x}$, it follows from Lemma 3.5 that for any $\bar{d} \in R^n$ such that (9) holds, there exist $d^k \in R^n$ such that

$$\begin{aligned} &\nabla g_j(x^k)d^k = 0, \quad j \in J_1(\bar{x}), \\ &\nabla g_j(x^k)d^k = 0, \quad j = m_1 + 1, \ldots, m \end{aligned} \tag{19}$$

and

$$d^k \to \bar{d}. \tag{20}$$

Note that

$$J_1^{+k} \cup J_1^k \subset J_1(\bar{x}),$$

when k is sufficiently large. As a result, (14) can be written as

$$\begin{aligned} d^T \nabla^2 f_0(x^k)d + &\sum_{j \in J_1(\bar{x}) \cap J_1^{+k}} [r_k(\nabla g_j(x^k)d)^2 + (y_j^k + r_k g_j(x^k))d^T \nabla^2 g_j(x^k)d] \\ &+ \sum_{j \in J_1(\bar{x}) \cap J_1^k} [(\nabla g_j(x^k)d)^+]^2 + \sum_{j=m_1+1}^{m} [r_k(\nabla g_j(x^k)d)^2 \\ &+ (y_j^k + r_k g_j(x^k))d^T \nabla^2 g_j(x^k)d] \geqslant 0. \end{aligned} \tag{21}$$

Substituting (19) into (21) (with d replaced by d^k), we obtain

$$\begin{aligned} [d^k]^T \nabla^2 f_0(x^k)d^k + &\sum_{j \in J_1(\bar{x}) \cap J_1^{+k}} (y_j^k + r_k g_j(x^k))[d^k]^T \nabla^2 g_j(x^k)d^k \\ + \sum_{j=m_1+1}^{m} &(y_j^k + r_k g_j(x^k))[d^k]^T \nabla^2 g_j(x^k)d^k \geqslant 0. \end{aligned}$$

Using (15), we get

$$[d^k]^T\nabla^2 f_0(x^k)d+\sum_{j\in J_1(\bar{x})}\mu_j^k[d^k]^T\nabla^2 g_j(x_k)d^k$$
$$+\sum_{j=m_1+1}^{m}\mu_j^k[d^k]^T\nabla^2 g_j(x^k)d^k \geqslant 0. \tag{22}$$

Taking the limit in (22) as $k\to+\infty$, and applying (18) and (20), we obtain (10). The proof is complete. □

3.2. CONVERGENCE ANALYSIS FOR GENERAL AUGMENTED LAGRANGIAN PROBLEMS

In this subsection, we assume that $m_1=0$. That is, the CP has only equality constraints. In this case, its augmented Lagrangian with convex augmenting function (in the sense of [13]) σ can be written as

$$l_{CP}(x,y,r)=f_0(x)+\sum_{j=1}^{m}y_j g_j(x)+r\sigma(-g_1(x),\dots,-g_m(x)),\quad x\in R^n, \tag{23}$$

where $y\in R^m, r>0$, and $\sigma : R^m\to\bar{R}$ is a proper convex augmenting function.

Suppose that $\{y^k\}$ is a sequence in R^m and $0<r_k$. The general augmented Lagrangian problems are

$$(\mathrm{P}_k^2)\qquad \inf_{x\in R^n} l_{CP}(x,y^k,r_k).$$

The following lemma establishes a second-order necessary condition for a local optimal solution to (P_k^2).

LEMMA 3.6. *Suppose that* $(-g_1(x^k),\dots,-g_m(x^k))\in\ \mathrm{dom}(\sigma)=\{u\in R^m : -\infty<\sigma(u)<+\infty\}$ *and that* x^k *is a local optimal solution of* (P_k^2) *at which the following condition holds:*

$$(C)\quad \sum_{j=1}^{m}w_j\nabla g_j(x^k)=0,\quad w\in N(-g(x^k)|\mathrm{dom}(\sigma))\Longrightarrow w=0,$$

where $w=(w_1,\dots,w_m)\in R^m, g(x)=(g_1(x),\dots,g_m(x))$, *and* $N(-g(x^k)|\mathrm{dom}(\sigma))$ *is the normal cone to* $\mathrm{dom}(\sigma)$ *at* $-g(x^k)$. *Furthermore, suppose that* $\{\nabla g_j(x^k) : j=1,\dots,m\}$ *are linearly independent. Then, the second-order necessary conditions of* (P_k^2) *hold: there exist* $\mu_j^k, j=1,\dots,m$ *such that*

$$\nabla f_0(x^k)+\sum_{j=1}^{m}\mu_j^k\,\nabla g_j(x^k)=0 \tag{24}$$

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

$$\nabla g_j(x^k)d = 0, \quad j = 1, \ldots, m, \tag{25}$$

there holds

$$d^T \nabla^2 f_0(x^k)d + \sum_{j=1}^{m} \mu_j^k d^T \nabla^2 g_j(x^k)d \geqslant 0. \tag{26}$$

Proof. It is easily seen that $l_{\mathrm{CP}}(x, y^k, r_k)$ is a convex composite function of the following two functions:

$$h(t,u) = t - \sum_{j=1}^{m} y_j^k u_j + r_k \sigma(u), \quad t \in R^1, \quad u \in R^m,$$
$$F(x) = (f_0(x), -g_1(x), \ldots, -g_m(x)), \quad x \in R^n,$$

where h is convex. That is,

$$l_{\mathrm{CP}}(x, y^k, r_k) = h(F(x)).$$

Obviously,

$$\mathrm{dom}(h) = R^1 \times \mathrm{dom}(\sigma).$$

In addition, it can be verified that $N(F(x^k)|\mathrm{dom}(h)) = \{0\} \times N(-g(x^k)|\mathrm{dom}(\sigma))$. Indeed, $\mathrm{dom}(h) = R^1 \times \mathrm{dom}(\sigma)$ is a covex subset of $R^1 \times R^m.v = (s, w) \in N(F(x^k)|\mathrm{dom}(h))$ if and only if for any $(s', w') \in R^1 \times \mathrm{dom}(\sigma)$, there holds

$$(s - f_0(x^k))(s' - f_0(x^k)) + (w + g(x^k))^T (w' + g(x^k)) \leqslant 0.$$

That is, $(s - f_0(x^k))(s' - f_0(x^k)) \leqslant 0$, $\forall s' \in R^1$ (by setting $w' = f_0(x^k)$) and $(w + g(x^k))^T (w' + g(x^k)) \leqslant 0$ (by setting $s' = f(x^k)$). The former means $s \in N(f_0(x^k)|R^1) = \{0\}$ while the latter means $w \in N(-g(x^k)|\mathrm{dom}(\sigma))$. Therefore, $s\nabla f_0(x^k) - \sum_{j=1}^{m} w_j \nabla g_j(x^k) = 0$ implies that $\sum_{j=1}^{m} w_j \nabla g_j(x^k) = 0$. Applying the condition (C), we have $w = 0$. Consequently, $v = 0$. Now we apply ([20], Theorem 3.1) to (P_k^2). There exists $z^k = (z_1^k, \ldots, z_m^k) \in \partial\sigma(-g_1(x^k), \ldots, -g_m(x^k))$ such that

$$\nabla f_0(x^k) + \sum_{j=1}^{m} y_j^k \nabla g_j(x^k) + r_k \sum_{j=1}^{m} z_j^k (-\nabla g_j(x^k)) = 0,$$

namely,

$$\nabla f_0(x^k) + \sum_{j=1}^{m} (y_j^k - r_k z_j^k) \nabla g_j(x^k) = 0 \tag{27}$$

and

$$\max\{d^T[\nabla^2 f_0(x^k) + \sum_{j=1}^{m}(y_j^k - r_k z_j^k)\nabla^2 g_j(x^k)]d : z^k \in \partial\sigma(-g_1(x^k), \ldots,$$
$$-g_m(x^k))\} \geqslant 0, \quad \forall d \in K(x^k), \tag{28}$$

where

$$K(x^k) = \{d \in R^n : \tau \nabla f_0(x^k)d + \tau \sum_{j=1}^{m} \nabla g_j(x^k)d$$
$$+ r_k\sigma(-g_1(x^k) - \tau \nabla g_1(x^k)d, \ldots, -g_m(x^k) - \tau \nabla g_m(x^k)d)$$
$$\leqslant r_k(-g_1(x^k), \ldots, -g_m(x^k)) \text{ for some } \tau > 0\}.$$

Let

$$\mu_j^k = y_j^k - r_k z_j^k, \quad j = 1, \ldots, m.$$

Then from (27), we have

$$\nabla f_0(x^k) + \sum_{j=1}^{m} \mu_j^k \nabla g_j(x^k) = 0. \tag{29}$$

By the linear independence of $\{\nabla g_j(x^k) : j = 1, \ldots, m\}$, we see that the μ_j^k $(j = 1, \ldots, m)$ are unique, having nothing to do with the choice of $z^k \in \partial\sigma(-g_1(x^k), \ldots, -g_m(x^k))$. As a result, (28) can be written as

$$d^T[\nabla^2 f_0(x^k) + \sum_{j=1}^{m} \mu_j^k \nabla^2 g_j(x^k)]d \geqslant 0, \quad \forall d \in K(x^k). \tag{30}$$

Note that any d satisfying (25) belongs to $K(x^k)$. Hence, for any d such that (25) holds, (26) holds. □

Remark 3.1. It can be checked that if σ is finite on R^m, then $\text{dom}(\sigma) = R^m$. It follows that $N(-g(x^k)|\text{dom}(\sigma)) = \{0\}$. Hence, condition (C) holds automatically.

THEOREM 3.3. *Suppose that* $0 < r_k, x^k \to \bar{x} \in X_0$. $\{\nabla g_j(\bar{x}) : j = 1, \ldots, m\}$ *are linearly independent. Then,* $\{\nabla g_j(x^k) : j = 1, \ldots, m\}$ *are linearly independent when k is sufficiently large. Further suppose that* x^k *satisfies the second-order necessary optimality conditions of* (P_k^2) *stated in Lemma* 3.6. *Then* $\bar{x}$ *satisfies the second-order conditions of* (CP).

Proof. It is obvious that $\{\nabla g_j(x^k) : j = 1, \ldots, m\}$ are linearly independent when k is sufficiently large since $x^k \to \bar{x}$.

By Lemma 3.6, (29) holds. By similar arguments as in the proof of Theorem 3.1, we can prove that $\{\sum_{j=1}^{m}|\mu_j^k|\}$ is bounded. As a result, we can assume without loss of generality that

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

Taking the limit in (29) as $k\to+\infty$, we get

$$\nabla f_0(\bar{x})+\sum_{j=1}^{m}\mu_j\nabla g_j(\bar{x})=0.$$

Now let d satisfy

$$\nabla g_j(\bar{x})d=0,\quad j=1,\ldots,m.$$

Since $\{\nabla g_j(\bar{x}):j=1,\ldots,m\}$ are linearly independent and $x^k\to\bar{x}$, by Lemma 3.5, there exists $d_k\to d$ such that

$$d_k^T\nabla^2 f_0(x^k)d_k+\sum_{j=1}^{m}\mu_j^k d_k^T\nabla^2 g_j(x^k)d_k\geqslant 0.$$

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

$$d^T\nabla^2 f_0(\bar{x})d+\sum_{j=1}^{m}\mu_j d^T\nabla^2 g_j(\bar{x})d\geqslant 0.\qquad\square$$

4. Conclusions

We presented a necessary and sufficient condition for the zero duality gap property via generalized augmented Lagrangian. For a constrained program, an equivalence in terms of the zero duality gap property was established between a general nonlinear Lagrangian dual problem and a class of generalized augmented Lagrangian dual problems. In the context of a mathematical program with both equality and inequality constraints, we proved that the second-order conditions of the Lagrangian problems with a convex quadratic augmenting function converge to that of the original constrained problem. In the context of a mathematical program with only equality constraints, for general augmented Lagrangian problems with a convex augmenting function, we showed that the second-order necessary conditions of the augmented Lagrangian converge to that of the original constrained program.

However, for a mathematical program with both equality and inequality constraints, we still do not know whether the second-order conditions of general augmented Lagrangian problems with a convex augmenting function converge to that of the original constrained program or not. There may exist some technical difficulty in answering this question. That is, there is generally no such a simple explicit expression as (23) for an augmented Lagrangian if the original constrained program has both equality and inequality constraints.

References

1. Bertsekas, D.P. (1982), Constrained Optimization and Lagrange Multiplier Methods, Academic Press, New York.
2. Ben-Tal, A. and Zowe, J. (1982), A unified theory of first-order and second-order conditions for extremum problems in topological vector spaces, *Mathematical Programming Study*, 19,39–76.
3. Ben-Tal, A. and Zowe, J. (1982), Necessary and sufficient conditions for a class of nonsmooth optimization problems, *Mathematical Programmming*, 24, 70–91.
4. Dantzig, G.B., Folkman, J. and Shapiro, N. (1967), On continuity of the minimum set of a continuous function, *Journal of Mathematical Analysis and Applications*, 17, 519–548.
5. Di Pillo, G. and Grippo, L. (1982), An augmented Lagrangian for inequality constraints in nonlinear programming problems, *Journal of Optimization Theory and Applications*, 495–519.
6. Di Pillo, G. and Lucidi, S. (1996), On exact augmented Lagrangian functions in nonlinear programming, In: G. Di Pillo and F. Giannessi (eds), *Nonlinear Optimization and Applications*, Plenum Press, New York, pp. 85–100.
7. Di Pillo, G. and Lucidi, S. (2001), An augmented Lagrangian function with improved exactness properties, *SIAM Journal on Optimization*, 12, 376–406.
8. Guker, O. (1992), Augmented Lagrangian algorithms for linear programming, *Journal of Optimization Theory and Applications*, 75, 445–470.
9. Huang, X.X. and Yang, X.Q. (2003), A unified augmented Lagrangian approach to duality and exact penalization, Mathematics of Operations Research, 28, 533–552.
10. Pinar, Mustafa C. (1999), A characterization of optimal set of linear programs based on the augmented Lagrangian, *Journal of Information and Optimization Sciences*, 20, 299–308.
11. Rockafellar, R.T. (1954), Augmented Lagrangian multiplier functions and duality in nonconvex programming, *SIAM Journal on Control and Optimization*, 12, 268–285.
12. Rockafellar, R.T. (1993), Lagrangian multipliers and optimality, *SIAM Review*, 35, 183–238.
13. Rockafellar, R.T. and Wets, R.J.-B. (1998), Variational Analysis, Springer-Verlag, Berlin.
14. Rubinov, A.M., Glover, B.M. and Yang, X.Q. (1999), Modified Lagrangian and penalty functions in continuous optimization, *Optimization*, 46, 327–351.
15. Rubinov, A.M., Huang, X.X. and Yang, X.Q. (2002), The zero duality gap property and lower semicontinuity of the perturbation function, Mathematics of Operations Research 27, 775–791.
16. Teo, K.L., Goh, C.J. and Wong, K.H. (1991), A Unified Computational Approach to Optimal Control Problems, Longman Scientific and Technical.
17. Wright, S.J. (1990), Implementing proximal point methods for linear programming, *Journal of Optimization Theory and Applications*, 65, 531–551.
18. Yang, X.Q. and Huang, X.X. (2001), A nonlinear Lagrangian approach to constrained optimization problems, *SIAM Journal on Optimization*, 11, 1119–1144.
19. Yang, X.Q. (1996), On second-order directional derivatives, *Nonlinear Analysis, Theory, Methods and Applications*, 26, 55–66 .
20. Yang, X.Q. (1998), Second-order global optimality conditions for convex composite optimization, *Mathematical Programming*, 81, 327–347.