Applied Mathematics & Information Sciences
An International Journal

New Multiplier Algorithm for Nonlinear Programming with Inequality Constraints

Jinchuan Zhou[1], *Xiuhua Xu*[2], *Jingyong Tang*[3]

[1] Department of Mathematics, School of Science, Shandong University of Technology, Zibo 255049, P.R.China
[2] Shandong Zibo Experimental High School, Zibo 255090, Shandong Province, P.R.China
[3] College of Mathematics and Information Science, Xinyang Normal University, Xinyang 464000, P.R.China

Received: Jul 24, 2011; Revised Oct. 4, 2011; Accepted Jan. 10, 2012
Published online: 1 May 2012

Abstract: We introduce a new class of augmented Lagrangian function, which includes the well-known essential quadratic augmented Lagrangian as special cases. Based on this new function, we propose a multiplier algorithm, whose main feature is that the multiplier sequence does not require to be bounded. Global convergence to optimal solutions and KKT points are established, respectively.

Keywords: Augmented Lagrangian; Multiplier Algorithms; KKT points.

1. Introduction.

This paper is concerned with the following nonlinear programming problem

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

$\mathrm{s.t.}\ \ g_i(x) \leq 0, \quad i = 1, \cdots, m$, where f and $g_i : R^n \rightarrow$ for $i = 1, \cdots, m$ are all continuously differentiable functions. Denoted by X the feasible region and by X^* the solution set. The *classical Lagrangian function* is

$$L(x) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x),$$

where $\lambda_i \in$ for $i = 1, \cdots, m.$

A main drawback involved in the above Lagrangian function is that for nonconvex programming problems a nonzero duality gap maybe arisen, which leads to the failure of using dual methods to find the solution of the primal problem. To overcome this difficulty, Henstence [7] and Powell [11] proposed independently the first *augmented Lagrangian function* by adding a second-order penalty term to the classical Lagrangian function. It was extended significantly by Rockafellar [12] and established its augmented Lagrangian dual theory, including the zero duality gap property and the existence of global/local saddle points. Since then, various augmented Lagrangian functions were proposed by many authors according to different requirement on theory analysis or algorithm designs; for example exponential augmented Lagrangian [15], modified barrier functions [10], nonlinear Lagrangian [17] etc. At the same time, the convergence properties of augmented Lagrangian methods have been developed; see [2,5,6,8,13] for the details.

However, an essential assumption imposed in the above algorithms is the boundedness of the Lagrangian multiplier sequence. This undoubtedly limits the applications of augmented Lagrangian methods in practice. More recently, this question attracts much attention of many scholars, and the important step in this direction includes [1, 3,4,9]. Nevertheless, it should be noted that the preformation of these algorithms are all restricted the iterative sequence $\{x_k\}$ to be convergent in advanced. Hence, our main aim in this paper is to establish the convergence property of augmented Lagrangian methods without requiring the boundedness of Lagrangian multipliers, and moreover to study the case even when the iterative sequence is divergent. More specially, we first introduce a new class of augmented Lagrangian functions, which including the essential quadratic augmented as special cases. The corresponding multiplier algorithms based on this class of augmented Lagrangian is proposed. The global convergence property is established without requiring the boundedness of Lagrangian multiplier sequence; for example, every accumulation points of iterative sequence $\{x_k\}$ is a global

* Corresponding author: e-mail: jinchuanzhou@163.com

optimal solution. Even if $\{x_k\}$ is divergent, we further develop the necessary and sufficient conditions for the convergence of $f(x_k)$ to the optimal value. Finally, under the Mangasarian-Fromovitz constraint qualification, we show that $\{x^k\}$ converges to a KKT point of primal problem.

The paper is organized as follows. In section 2, we introduce a new class of augmented Lagrangian function and propose the multiplier algorithm. Section 3 is devoted to the convergence property of our algorithms. Numerical reports are given in Section 4.

2. Multiplier Algorithm

In this section, we first introduce a new *generalized essentially quadratic augmented Lagrangian function* for (P), defined as,

$$L(x,\lambda,c) := f(x)+\frac{1}{2c}\sum_{i=1}^{m}\{\max{}^2\{0,\phi(cg_i(x))+\lambda_i\}-\lambda_i^2\}\ (1)$$

where $(x,\lambda,c) \in R^n \times R^m \times R_{++}$ and $_{++}$ denotes the all positive real scalars, i.e., $_{++} = \{a \in |a > 0\}$. The function $\phi : R \to R$ involved in (1) satisfies the following properties:

(A_1)continuously differentiable and strictly increasing on R with $\phi(0) = 0$ and $\phi(\alpha) \geq \alpha$ for $\alpha \geq 0$.

If, in particular, $\phi(\alpha) = \alpha$ for all $\alpha \in$, then L reduces to the essential quadratic augmented Lagrangian function introduced by Rockafellar; see [12] for more information. Compared with [9,14,16], an important point made above is that ϕ is not required to be convex. Hence the augmented Lagrangians we introduce here is more general.

Given (x,λ,c), the *Lagrangian relaxation problem* associated with the augmented Lagrangian L is defined as

$(\mathrm{L}_{\lambda,c})$ $\min\ L(x,\lambda,c)$

s.t. $x \in R^n$. Its solution set is denoted by $S^*(\lambda,c)$.

Recall that a vector x is said to be a KKT point of (P) if there exist $\lambda_i \in_+$ for all $i = 1,\cdots,m$ such that the following system hold

$$\nabla f(x) + \sum_{i=1}^{m}\lambda_i\nabla g_i(x) = 0, \quad (2)$$

$$\lambda_i g_i(x) = 0, \quad \text{for all } i = 1,\cdots m, \quad (3)$$

where the second condition is referred to as the well-known *complementarity condition.* For notational simplification, the collect set of multipliers satisfying (2) and (3) is denoted by $\Lambda(x)$.

Throughout this paper we always assume that f is bounded from below, i.e.,

$$f_* := \inf_{x\in R^n} f(x) > -\infty.$$

This assumption is rather mild in optimization problem, because otherwise the objective function f can be replaced by $e^{f(x)}$. The multiplier algorithm based on the generalized essential quadric augmented Lagrangian L is proposed below. One of its main feature is that the Lagrangian multipliers associated are not restricted to be bounded, which make the algorithm applicable for many problems in practice. Let us denote

$$\xi = \lim_{s\to-\infty}\phi(s), \quad \text{where } \xi \in [-\infty, 0). \quad (4)$$

According to the monotonicity of ϕ by property (A_1), we know that $\xi < 0$. The case of $\xi = -\infty$ corresponds to that ϕ is unbounded from below.

(Multiplier algorithm based on L):

Step 0.Select an initial point $x^0 \in R^n$, $\lambda_i^0 \in [0, -\xi/2)$ for $i = 1,\cdots,m$, and $c_0 > 0$. Set $k := 0$,

Step 1.Compute

$$\lambda_i^{k+1} = \max\{0, \phi(c_k g_i(x^k)) + \lambda_i^k\}\phi'(c_k g_i(x^k)) \quad (5)$$

$$c_{k+1} \geq (k+1)\max\{1, \sum_{i=1}^{m}(\lambda_i^{k+1})^2\}, \quad (6)$$

Step 2.Find $x^{k+1} \in S^*(\lambda^{k+1}, c_{k+1})$,

Step 3.If $x^{k+1} \in X$ and $\lambda^{k+1} \in \Lambda(x^{k+1})$, then STOP; otherwise, let $k := k + 1$ and go back to Step 1.

The following lemma gives the relationship between the penalty parameter c_k and the multipliers λ^k.

Lemma 1.*Let (λ^k, c_k) be given as in Algorithm 2. Then the following terms*

$$\frac{\lambda^k}{c_k}, \quad \frac{(\lambda^k)^2}{c_k},$$

are all approaches to zero as $k \to \infty$.

*Proof.*This is clear from (6).

3. Convergence Analysis

For establishing the convergence property of Algorithm 2, we first consider the perturbation analysis of (P). Given $\alpha \geq 0$, define the perturbation of feasible region as

$$X(\alpha) = \{x \in R^n | g_i(x) \leq \alpha, i = 1,\cdots,m\},$$

and the perturbation of level set as

$$L(\alpha) = \{x \in R^n | f(x) \leq v(0) + \alpha\}.$$

It is clear that $X(0)$ coincides with the feasible set of (P). The corresponding perturbation function is given as:

$$v(\alpha) = \inf\{f(x) | x \in X(\alpha)\}.$$

The following result shows that the perturbation value function is upper semi-continuous at zero.

Lemma 2.*The perturbation function v is upper semi-continuous at zero from right.*

*Proof.*It only needs to show that

$$\limsup_{\alpha\to 0^+} v(\alpha) \leq v(0),$$

which is followed by the fact $v(\alpha) \leq v(0)$, since $X(0) \subseteq X(\alpha)$ for all $\alpha \geq 0$.

Lemma 3. *For any $\lambda \in^m$ and $c > 0$, one has*

$$S^*(\lambda, c) \subseteq \{x \in R^n | L(x, \lambda, c) \leq v(0)\}.$$

Proof. For any $\bar{x} \in S^*(\lambda, c)$, we have

$$\begin{aligned} L(\bar{x}, \lambda, c) &= \inf\{L(x, \lambda, c) | x \in R^n\} \\ &\leq \inf\{L(x, \lambda, c) | x \in X(0)\} \\ &\leq \inf\{f(x) | x \in X(0)\} \\ &= v(0), \end{aligned}$$

where the second inequality uses the fact $\phi(cg_i(x)) \leq 0$ for all $x \in X(0)$, since ϕ is nondecreasing by the property (A_1).

Lemma 4. *Let (λ^k, c_k) be given as in Algorithm 2. For any $\epsilon > 0$, one has*

$$\{x \in R^n | L(x, \lambda^k, c_k) \leq v(0) + \epsilon\} \subseteq X(\epsilon).$$

whenever k is sufficiently large.

Proof. We prove this result by contradiction. Suppose that we can find an $\epsilon_0 > 0$ and an infinite subsequence $K \subseteq \{1, 2, \cdots\}$ such that

$$z^k \in \{x \in R^n | L(x, \lambda^k, c_k) \leq v(0) + \epsilon\}, \quad \forall k \in K, \quad (7)$$

but

$$z^k \notin X(\epsilon_0), \quad \forall k \in K.$$

It follows from (7) that

$$v(0) + \epsilon \geq L(z^k, \lambda^k, c_k). \quad (8)$$

Since $z^k \notin X(\epsilon_0)$, then there exists an index i_0 and an infinite subsequence $K_0 \subseteq K$ such that $g_{i_0}(z^k) > \epsilon_0$. It follows from (8) that

$$\begin{aligned} &v(0) + \epsilon \\ &\geq f_* + \frac{1}{2c_k} \sum_{i=1}^{m} \{\max{}^2\{0, \phi(c_k g_i(z^k)) + \lambda_i^k\} - (\lambda_i^k)^2\} \\ &= f_* + \frac{1}{2c_k} \{\max{}^2\{0, \phi(c_k g_{i_0}(z^k)) + \lambda_{i_0}^k\} - (\lambda_{i_0}^k)^2\} \\ &\quad + \frac{1}{2c_k} \sum_{i \neq i_0} \{\max{}^2\{0, \phi(c_k g_i(z^k)) + \lambda_i^k\} - (\lambda_i^k)^2\} \\ &\geq f_* + \frac{1}{2c_k} \{\max{}^2\{0, c_k g_{i_0}(z^k)) + \lambda_{i_0}^k\} - (\lambda_{i_0}^k)^2\} \\ &\quad + \frac{1}{2c_k} \sum_{i \neq i_0} \{\max{}^2\{0, \phi(c_k g_i(z^k)) + \lambda_i^k\} - (\lambda_i^k)^2\} \\ &\geq f_* + \frac{c_k}{2} g_{i_0}^2(z^k) + g_{i_0}(z^k)\lambda_{i_0}^k - \frac{1}{2c_k} \sum_{i \neq i_0} (\lambda_i^k)^2 \\ &\geq f_* + \frac{c_k}{2} \epsilon_0^2 + \epsilon_0 \lambda_{i_0}^k - \frac{1}{2c_k} \sum_{i=1}^{m} (\lambda_i^k)^2 \\ &\geq f_* + \frac{c_k}{2} \epsilon_0^2 - \frac{1}{2c_k} \sum_{i=1}^{m} (\lambda_i^k)^2, \end{aligned}$$

where the second inequality comes from the fact $\phi(a) \geq a$ for all $a \geq 0$ by property (A_1), and the last inequality follows from the nonnegativity of λ_i^k by (5) (noting that $\phi'(\alpha) \geq 0$ for all $\alpha \in$ since ϕ is nondecreasing). Taking limits in the above inequality yields $v(0) = +\infty$, which is a contradiction. The proof is complete.

Lemma 5. *Let (λ^k, c_k) be given as in Algorithm 2. For any $\epsilon > 0$, one has*

$$\{x \in R^n | L(x, \lambda^k, c_k) \leq v(0) + \frac{\epsilon}{2}\} \subseteq L(\epsilon).$$

whenever k is sufficiently large.

Proof. For arbitrarily $\bar{x} \in \{x \in R^n | L(x, \lambda^k, c_k) \leq v(0) + \frac{\epsilon}{2}\}$, it follows from (1) that

$$f(\bar{x}) \leq v(0) + \frac{\epsilon}{2} + \frac{1}{2c_k} \sum_{i=1}^{m} (\lambda_i^k)^2. \quad (9)$$

As k is large enough, Lemma 1 ensures that

$$\frac{1}{c_k} \sum_{i=1}^{m} (\lambda_i^k)^2 \leq \epsilon,$$

which together with (9) justifies the desired result.

With these preparation, the global convergence property of Algorithm 2 can be given, which shows that if the algorithm terminates in finite steps, then we obtain a KKT point of (P); otherwise every limit point of $\{x^k\}$ would be the optimal solution of (P).

Theorem 1. *Let x^k be the iterative sequence generated by Algorithm 2. Then If $\{x^k\}$ is terminated in finite steps, then we get a KKT point of (P); otherwise, every limit point of x^k belongs to X^*.*

Proof. According to the termination criterion of Algorithm 2, the first part is clear. Let us consider the second part. For any $\epsilon > 0$, it follows from Lemmas 3-5 that when k is large enough we have

$$\begin{aligned} S^*(\lambda^k, c_k) &\subseteq \{x \in R^n | L(x, \lambda^k, c_k) \leq v(0)\} \\ &\subseteq \{x \in R^n | L(x, \lambda^k, c_k) \leq v(0) + \frac{\epsilon}{2}\} \\ &\subseteq X(\epsilon) \cap L(\epsilon). \end{aligned}$$

Thus,

$$x^k \in X(\epsilon) \cap L(\epsilon). \quad (10)$$

Note that $X(\epsilon)$ and $L(\epsilon)$ are closed, due to the continuity of f and g_i for all $i = 1, \cdots, m$. Taking the limit in (10) yields $x^* \in X(\epsilon) \cap L(\epsilon)$, which further shows that $x^* \in X(0) \cap L(0)$, since $\epsilon > 0$ is arbitrary, i.e., $x^* \in X^*$. The proof is complete.

The foregoing result is applicable to the case when $\{x^k\}$ at least has an accumulation point. However, a natural question arises: *how does the algorithm perform as $\{x^k\}$ is divergent?* The following theorem gives an answer.

Theorem 2.*Let $\{x^k\}$ be an iterative sequence generalized by Algorithm 2. Then*

$$\lim_{k\to\infty} f(x^k) = v(0) \tag{11}$$

if and only if $v(\alpha)$ is lower semi-continuous at $\alpha = 0$ from right.

*Proof.*Sufficiency. According to the proof of Theorem 1 [cf. (10)], we know that

$$v(\epsilon) \le f(x^k) \le v(0) + \epsilon, \tag{12}$$

whenever k is sufficiently large. Since $v(\alpha)$ is lower semi-continuous at $\alpha = 0$, taking the lower limitation in (12) yields

$$v(0) \le \liminf_{\epsilon\to 0^+} v(\epsilon) \le \liminf_{k\to\infty} f(x^k)$$
$$\le \limsup_{k\to\infty} f(x^k) \le v(0) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, the above estimate gives us

$$\lim_{k\to\infty} f(x^k) = v(0).$$

Necessity. Suppose on the contrary that v is not lower semicontinuous at zero from right, then there must exist $\delta_0 > 0$ and $\epsilon_j \to 0^+$(as $j \to \infty$) such that

$$v(\epsilon_j) \le v(0) - \delta_0, \quad \forall j. \tag{13}$$

Let k be fixed. Since $\epsilon_j \to 0$ we can choose an subsequence j_k satisfying $\epsilon_{j_k} c_k \le 1/k$. Hence

$$\epsilon_{j_k} c_k \to 0 \quad \text{as } k \to \infty,$$

which, together with the continuity of ϕ, further implies that

$$\phi(\epsilon_{j_k} c_k) \to 0 \quad \text{as } k \to \infty. \tag{14}$$

In addition, let $z^k \in X(\epsilon_{j_k})$ with $f(z^k) \le v(\epsilon_{j_k}) + \frac{\delta_0}{2}$, which further implies $f(z^k) \le v(0) - \frac{\delta_0}{2}$ by (13). Therefore,

$$\begin{aligned}
& f(x^k) \\
&= L(x^k, \lambda^k, c_k) - \frac{1}{2c_k}\sum_{i=1}^m \{\max{}^2\{0, \phi(c_k g_i(x^k)) + \lambda_i^k\} \\
&\quad -(\lambda_i^k)^2\} \\
&= \inf_{x\in R^n} L(x, \lambda^k, c_k) - \frac{1}{2c_k}\sum_{i=1}^m \{\max{}^2\{0, \phi(c_k g_i(x^k)) + \lambda_i^k\} \\
&\quad -(\lambda_i^k)^2\} \\
&\le \inf_{x\in R^n} L(x, \lambda^k, c_k) + \frac{1}{2c_k}\sum_{i=1}^m (\lambda_i^k)^2 \\
&\le f(z^k) + \frac{1}{2c_k}\sum_{i=1}^m \{\max{}^2\{0, \phi(c_k g_i(z^k) + \lambda_i^k)\} - (\lambda_i^k)^2\} \\
&\quad + \frac{1}{2c_k}\sum_{i=1}^m (\lambda_i^k)^2 \\
&\le v(0) - \frac{\delta_0}{2} + \frac{1}{2c_k}\sum_{i=1}^m \left(\phi(c_k \epsilon_{j_k}) + \lambda_i^k\right)^2,
\end{aligned} \tag{15}$$

where the last step is due to the fact $g_i(z^k) \le \epsilon_{j_k}$ since $z^k \in X(\epsilon_{j_k})$ and ϕ is nondecreasing by the property (A_1). Taking the limits in both sides of (15) and using Lemma 1 and (14), it follows from (11) that

$$v(0) = \lim_{k\to\infty} f(x^k) \le v(0) - \frac{\delta_0}{2},$$

which leads to a contradiction. The proof is complete.

We conclude this paper by establishing the convergence of the lagrangian multiplier sequence $\{\lambda^k\}$ in the presence of Mangasarian-Fromovitz constraint qualification. Let us first recall that M.F. constraint qualification is said to be satisfied at x^*, if there exists $h_0 \in^n$ such that

$$\langle \nabla g_i(x^*), h_0\rangle < 0, \quad \forall i \in I(x^*),$$

where $I(x^*) = \{i | g_i(x^*) = 0, i = 1, \cdots, m\}$. Here we further assume that ϕ satisfies

$(A_2)$$\phi'(s) \le 1$ whenever s is sufficiently small, i.e., there exists $s_0 < 0$ such that $\phi'(s) \le 1$ for all $s \le s_0$.

Clearly, this assumption holds automatically when $\phi(\alpha) = \alpha$.

Theorem 3.*Let $\{x^k\}$ be the iterative sequence generated by Algorithm 2 and x^* is one of its limit points. Then*

(a).*If M.F. constraint qualification is satisfied at x^*, then λ^k is bounded and any of its limit points, say λ^*, satisfies (x^*, λ^*) is a KKT point of (P).*
(b).*If the linearly independent constraint qualification holds at x^*, we further obtain that the multiplier sequence $\{\lambda_i^k\}$ for $i = 1, \cdots, m$ are convergent.*

*Proof.*We assume without loss of generality that $\lim_{k\to\infty} x^k = x^*$. It follows from Theorem 1 that $x^* \in X^*$. If $i \notin I(x^*)$, then $g_i(x^k) < 0$ as k large enough. This means the existence of $\epsilon_0 > 0$ such that $g_i(x^k) \le -\epsilon_0$ whenever k is sufficiently large. Therefore,

$$\lim_{k\to\infty} c_k g_i(x^k) = -\infty.$$

Hence, taking into account of the properties $(A1)$, (A_2) and Step 1 in Algorithm 2, we obtain

$$\begin{aligned}
\lambda_i^{k+1} &= \max\{0, \phi(c_k g_i(x^k)) + \lambda_i^k\}\phi'(c_k g_i(x^k)) \\
&\le \max\{0, \lambda_i^k\} = \lambda_i^k \le \cdots \le \lambda_i^0,
\end{aligned} \tag{16}$$

where the second equation comes from the nonnegativity of λ_i according to the construction in Algorithm 2. This justifies the boundedness $\{\lambda_i^k\}$ for $i \notin I(x^*)$. Since $\lambda_i^k \le \lambda_i^0 < -\frac{3}{4}\xi$ (see Step 0 in Algorithm 2) and $\lim_{k\to\infty} \phi(c_k g_i(x^k)) = \xi$ by (4), then

$$\lim_{k\to\infty} \phi(c_k g_i(x^k)) + \lambda_i^k \le \xi/4 < 0.$$

Therefore, taking limit in (16) yields

$$\lim_{k\to\infty} \lambda_i^k = 0, \quad \forall i \notin I(x^*). \tag{17}$$

We now show that λ_i^k for $i \in I(x^*)$ are bounded as well. If this is not true, then we can find an infinite subsequence $K \subseteq \{1, 2, \cdots, \}$ such that

$$T_k := \sum_{i \in I(x^*)} \lambda_i^k \to +\infty, \quad \text{as } k \to \infty. \tag{18}$$

Since $x^{k-1} \in S^*(\lambda^{k-1}, c_{k-1})$ by Algorithm 2, according to the well-known of optimality conditions for unconstrained optimization problem we must have

$$\nabla_x L(x^{k-1}, \lambda^{k-1}, c_{k-1}) = 0,$$

which together with (5) means that

$$\nabla f(x^{k-1}) + \sum_{i=1}^{m} \lambda_i^k \nabla g_i(x^{k-1}) = 0. \tag{19}$$

Since $\frac{\lambda_i^k}{T_k}$ is bounded, we can assume without loss of generality that

$$\frac{\lambda_i^k}{T_k} \to \tilde{\lambda}_i^*, \quad \forall i \in I(x^*).$$

Since

$$\sum_{i \in I(x^*)} \frac{\lambda_i^k}{T_k} = 1,$$

taking limit with respect to $k \in K$ in the above equation gives us

$$\sum_{i \in I(x^*)} \tilde{\lambda}_i^* = 1,$$

which implies that $\tilde{\lambda}_i^*$ for $i \in I(x^*)$ are not all zero. Dividing on both sides of (19) by T_k, taking limit with respect to $k \in K$, and using (18), we get

$$\sum_{i \in I(x^*)} \tilde{\lambda}_i^* \nabla g_i(x^*) = 0.$$

Thus

$$0 = \langle \sum_{i \in I(x^*)} \tilde{\lambda}_i^* \nabla g_i(x^*), h_0 \rangle = \sum_{i \in I(x^*)} \tilde{\lambda}_i^* \langle \nabla g_i(x^*), h_0 \rangle < 0,$$

where the last step is due to the fact that at least one of $\tilde{\lambda}_i^*$ is not zero. This leads to a contradiction. Therefore, we establish the boundedness of λ^k. Let λ^* be a limit point of λ^k. It follows from (17) that $\lambda^* = 0$ for $i \notin I(x^*)$ and from (19) that

$$\nabla f(x^*) + \sum_{i \in I(x^*)} \lambda_i^* \nabla g_i(x^*) = 0,$$

i.e., $\lambda^* \in \Lambda(x^*)$. This establishes Part (a). Part (b) can be proved in the same vein, just noting that in the presence of linearly independence constraint qualification the Lagrangian multiplier is unique. This together with the boundedness of λ_i^k for $i = 1, \cdots, m$ ensure the convergence to the unique accumulation point.

4. Numerical Results

To give some insight into the behavior of our proposed algorithm, we solve the following nonconvex programming problems by letting ϕ take the following different functions:

1. $\phi_1(\alpha) = \alpha$,
2. $\phi_2(\alpha) = (1 + \frac{1}{3}\alpha)^3 - 1$,
3. $\phi_3(\alpha) = \alpha\big(\ln(1 + \alpha^2) + 1\big)$,
4. $\phi_4(\alpha) = \alpha + \alpha^3$.

The test was done at a PC of Pentium 4 with 2.8GHz CPU and 1.99GB memory. The computer codes were written in Matlab 7.0. Numerical results are reported in the following Table, where k is the number of iterations, c_k is the penalty parameter, λ^k is multipliers, and $f(x^k)$ is the objective value.

Example 1.[18] $\min\ 0.5(x_1 + x_2)^2 + 50(x_2 - x_1)^2 + \sin^2(x_1 + x_2)$
s.t. $(x_1 - 1)^2 + (x_2 - 1)^2 + (\sin(x_1 + x_2) - 1)^2 - 1.5 \le 0$

Example 2.[18] $\min\ 0.5(x_1 + x_2)^2 + 50(x_2 - x_1)^2 + x_3^2 + |x_3 - \sin(x_1 + x_2)|$
s.t. $(x_1 - 1)^2 + (x_2 - 1)^2 + (x_2 - 1)^2 - 1.5 \le 0$

Example 3.[19] $\min\ f(x) = -5(x_1 + x_2) + 7(x_4 - 3x_3) + x_1^2 + x_2^2 + 2x_3^2 + x_4^2$
s.t. $\sum_{i=1}^4 x_i^2 + x_1 - x_2 + x_3 - x_4 - 8 \le 0$
$x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 10 \le 0$
$2x_1^2 + x_2^2 + x_3^2 + 2x_1 - x_2 - x_4 - 5 \le 0$

Example 4.[19] $\min\ f(x) = (x_1 - 10)^2 + 5(x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2 + 10x_5^6 + 7x_6^2 + x_7^4 - 4x_6x_7 - 10x_6 - 8x_7$
s.t. $2x_1^2 + 3x_2^4 + x_3 + 4x_4^2 + 5x_5 - 127 \le 0$
$7x_1 + 3x_2 + 10x_3^2 + x_4 - x_5 - 282 \le 0$
$23x_1 + x_2^2 + 6x_6^2 - 8x_7 - 196 \le 0$
$4x_1^2 + x_2^2 - 3x_1x_2 + 2x_3^2 + 5x_6 - 11x_7 \le 0$

Comparing with their numerical behaviors given in Tables 1-4, it is clear that Algorithm 2 is more preferable when ϕ is nonconvex than $\phi(\alpha) = \alpha$, since the iterative step k is fewer and penalty coefficient c_k is smaller.

Acknowledgement

Research of Jinchuan Zhou was partly supported by National Natural Science Foundation of China (11101248, 11171247, 71101140), Shandong Province Natural Science Foundation (ZR2010AQ026), and Young Teacher Support Program of Shandong University of Technology. The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

Table 1 Numerical Results of Example 4.1

$\phi_i(\alpha)$	k	c_k	$f(x^k)$
$\phi_1(\alpha)$	2	2.0000	0.3141
	4	4.0000	0.3012
	6	6.0000	0.3004
$\phi_2(\alpha)$	2	2.0000	0.2739
	3	3.0000	0.2992
	4	4.0000	0.3004
$\phi_3(\alpha)$	2	2.0000	0.2331
	4	4.0000	0.2992
	5	5.0000	0.3004
$\phi_4(\alpha)$	2	2.0000	0.2845
	3	3.0000	0.2998
	4	4.0000	0.3004

Table 2 Numerical Results of Example 4.2

$\phi_i(\alpha)$	k	c_k	$f(x^k)$
$\phi_1(\alpha)$	2	2.0000	0.3052
	4	4.0000	0.3007
	6	6.0000	0.3004
$\phi_2(\alpha)$	2	2.0000	0.2739
	3	3.0000	0.2991
	4	4.0000	0.3004
$\phi_3(\alpha)$	2	2.0000	0.2629
	3	3.0000	0.2974
	5	5.0000	0.3004
$\phi_4(\alpha)$	2	2.0000	0.2759
	3	3.0000	0.2976
	4	4.0000	0.3004

Table 3 Numerical Results of Example 4.3

$\phi_i(\alpha)$	k	c_k	$f(x^k)$
$\phi_1(\alpha)$	1	10.2288	-43.9121
	4	19.4016	-44.0044
	8	37.3764	-44.0051
$\phi_2(\alpha)$	1	2.4000	-44.0926
	4	19.9795	-44.0190
	6	21.3354	-44.0098
$\phi_3(\alpha)$	1	10.0128	-44.8715
	3	19.3040	-44.0044
	4	24.9815	-44.0000
$\phi_4(\alpha)$	1	3.3715	-45.5757
	3	8.2728	-44.0347
	4	10.7013	-44.0180

Table 4 Numerical Results of Example 4.4

$\phi_i(\alpha)$	k	c_k	$f(x^k)$
$\phi_1(\alpha)$	1	1.0000	679.6946
	2	0.1637	680.5644
	5	0.1516	680.6300
$\phi_2(\alpha)$	1	1.0000	679.8201
	2	0.1420	680.5978
	4	0.1435	680.6300
$\phi_3(\alpha)$	1	1.0000	679.9210
	2	0.1421	680.5856
	4	0.1435	680.6300
$\phi_4(\alpha)$	1	1.0000	679.9520
	2	0.1421	680.5871
	4	0.1435	680.6300

References

[1] R. Andreani, E.G.Birgin, J.M.Martínez, *On Augmented Lagrangian methods with general lower-level constraints*, SIAM J. Optim. 18:1286-1309, 2007.

[2] D.P.Bertsekas, *Constrained Optimization and Lagrange Multipliers Method*, Academic Press, Inc., 1982.

[3] E.G.Birgin, C.A.Floudas, J.M.Martínez, *Global minimization using an Augmented Lagrangian method with variable lower-level constraints*, Math. Programming, 125:139-162, 2010.

[4] E.G.Birgin, J.M.Martinez, *Augmented Lagrangian method with nonmonotone penalty parameters for constrained optimization*, Comput. Optim. Appl., doi: 10.1007/s10589-011-9396-0, 2011.

[5] R.S.Burachik, A.M.Rubinov, *Abstract convexity and augmented lagrangians*, SIAM J. Optim. 18:413- 436, 2007.

[6] A.R.Conn, N.I.M.Gould, A.Sartenaer, P.L.Toint, *Convergence properties of an augmented Lagrangian algorithm for optimization with a combination of general equality and linear constraints*, SIAM J. Optim., 6:674-703, 1996.

[7] M.R.Hestenes, *Multiplier and gradient methods*, J. Optim. Theory Appl., 4:303-320, 1969.

[8] R.M.Lewis, V.Torczon, *A globally convergent augmented Lagrangian pattern search algorithm for optimization with general constraints and simple bounds*, SIAM J. Optim., 12:1075-1089, 2002.

[9] H.Z.Luo, X.L.Sun, D.Li, *On the convergence of augmented Lagrangian methods for constrained global optimization*, SIAM J. Optim., 18:1209-1230, 2007.

[10] R.Polyak, *Modified barrier functions: Theory and mehtods*, Math. Programming, 54:177-222, 1992.

[11] M.J.D.Powell, *A Method for Nonlinear Constraints in Minimization Problems*, (R.Fletcher, ed.) Academic Press, New York, 1969.

[12] R.T.Rockafellar, *Augmented Lagrange multiplier functions and duality in nonconvex programming*, SIAM. J. Control Optim., 12:268-285, 1974.

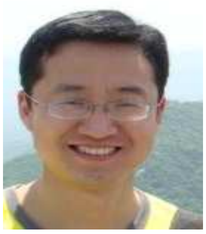
[13] A.M.Rubinov, X.Q.Yang, *Lagrange-type Functions in Constrained Non-Convex Optimization*, Kluwer Academic Publishers, Dordrecht, 2003.

[14] X.L.Sun, D.Li, K.I.M.McKinnon, *On saddle points of augmented Lagrangians for constrained nonconvex optimization*, SIAM J. Optim., 15:1128-1146, 2005.

[15] P.Tseng, *On the convergence of the exponential multiplier method for convex prgoramming*, Math. Programming, 60:1-19, 1993.
[16] C.Y.Wang, D.Li, *Unified theory of augmented Lagrangian methods for constrained global optimization*, J. Global, Optim., 44:433-458, 2009.
[17] X.Q.Yang, X.X.Huang, *A nonlinear Lagrangian approach to constrained optimization problems*, SIAM J. Optim., 11:1119-1144, 2001.
[18] Khenkin, E.I. , A Search Algorithm for General Problem of Mathematical Programming, USSR Journal of Computational Mathematics and Mathematical Physics 16, 61-71, 1976.
[19] G.D.Pillo and L.Grippo, *A New Augmented Lagrangian Function for Inequality Constraints in Nonlinear Programming Problems*, J. Optim. Theory Appl., 36:495-519, 1982.

Jinchuan Zhou is a lecturer at Shandong University of Technology. He obtained his PhD from Beijing Jiaotong University in July, 2009. His research interests include nonlinear programming, variational analysis, and cone optimizations.

Xiuhua Xu obtained her M.Sc at Qufu Normal University in 2006 and since then as a lecturer at Zibo Experimental High School. Her research fields are nonlinear programming: theory algorithm, and applications.



Jingyong Tang is a lecturer at Xinyang Normal University, China, and currently is a PhD candidate at Shanghai Jiaotong University, China. He worked in numerical experiments of various methods in optimization.