AN ACCELERATED AUGMENTED LAGRANGIAN METHOD FOR MULTI-CRITERIA OPTIMIZATION PROBLEM

XUEYONG WANG\textsuperscript{1}, YIJU WANG AND GANG WANG

School of Management Science, Qufu Normal University
Rizhao Shandong, 276800, China

(Communicated by Shengjie Li)

Abstract. By virtue of the Nesterov’s acceleration technique, we establish an accelerated augmented Lagrangian method for solving linearly constrained multi-criteria optimization problem. For this method, we establish its global convergence under suitable condition. Further, we show that its iteration-complexity is $O(1/k^2)$ which improves the original ALM whose iteration-complexity is $O(1/k)$.

1. Introduction. Let $\Omega$ be a nonempty, compact and convex set in $\mathbb{R}^n$, $A \in \mathbb{R}^{l \times n}, b \in \mathbb{R}^l$, and $f = (f_1, f_2, \cdots, f_m)^T : \Omega \to \mathbb{R}^m$, where $f_i$ be a proper lower semi-continuous and convex function, $i \in I = \{1, 2, \cdots, m\}$. Consider the following linearly constrained multi-criteria optimization problem

$$\min_{x \in \Omega} f(x), \text{ s.t. } Ax = b, x \in \Omega.$$  \hspace{1cm} (1.1)

It is well known that multiobjective optimization is a special case of vector optimization problems which has a large number of real life applications, such as signal and image processing, see [2, 3, 4, 5, 12, 13, 14, 22, 28, 29, 31, 33, 37, 43] and references therein.

It is well known that the scalarization approach is one of the most popular technique for solving multiobjective optimization problems [21] which is characterized by substituting more than one scalar valued problems for the vector valued problem so that all optima of the scalar valued problems are solutions of the original problem. Different to the above, Fliege and Svaiter [11] propose a descent method for solving multiobjective optimization problems. Compared with the scalarization approach, this method do not need a priori-weighting factors, ranking of the objectives according to their importance, nor any other ordering a priori information of the objectives. Up to now, various of descent method [7, 10, 17, 18, 26, 30, 35, 40], the projected gradient method [1, 27, 36, 41] and the Newton/Newton-like method [20, 24], for more information, see e.g. [5, 6, 8, 9, 15, 16, 19, 24, 32, 34, 38, 39, 42]. Note that the augmented Lagrangian method (ALM) can also be applied to the

\textsuperscript{1}This research was done during his postdoctoral period in Qufu Normal University.

2010 Mathematics Subject Classification. 15A18, 15A69, 65F15, 65F10.

Key words and phrases. Iteration-complexity, augmented Lagrangian method, multi-criteria optimization problem.

This work was supported by the Natural Science Foundation of China (11671228, 11801309), Shandong Provincial Natural Science Foundation (ZR2016AM10), and Science & Technology Planning Project of Qufu Normal University (XKJ201623).
multiobjective optimization [28], and the iteration complexity $O(1/k)$ is illustrated via an numerical example [5].

In this paper, we further investigate the augmented Lagrangian method for solving linearly constrained multi-criteria optimization problem. For this method, we would introduce the Nesterov’s acceleration technique [23] into the method to established an accelerated solution method (AALM) for problem (1.1). To show the acceleration of the new method, we first show that the iteration-complexity of the ALM is $O(1/k)$, and then show that the iteration-complexity of the AALM is $O(1/k^2)$.

The remainder of the paper is organized as follows. In Section 2, gives some preliminaries and notations used for subsequent analysis. In Section 3, we establish the iteration-complexity of the ALM, and in Section 4 we establish the iteration-complexity of the AALM.

2. Preliminaries. For any $x, y \in \mathbb{R}^n$, define $x \preceq y \iff y_i - x_i \geq 0, i \in I$ and $x \not\preceq y \iff y_i - x_i > 0, i \in I$. The indicator function of a set $\Omega \subset \mathbb{R}^n$ is defined as

$$
\sigma(x, \Omega) = \begin{cases} 0, & \text{if } x \in \Omega, \\ +\infty, & \text{if } x \notin \Omega. \end{cases}
$$

The normal cone to a convex set $\Omega$ at $x \in \Omega$ is defined as

$$
N_{\Omega}(x) = \partial \sigma(x, \Omega) = \{x^* \in \mathbb{R}^n | (x^*, y - x) \leq 0, \forall y \in \Omega\}.
$$

Let $f : \Omega \rightarrow \mathbb{R}^n \cup \{+\infty\}$ be a proper lower semi-continuous and convex function, its domain is defined as $\text{dom}(f) = \{x \in \Omega | f(x) < +\infty\}$, the sub-differential of $f$ at $x \in \text{dom}(f)$ is defined as

$$
\partial f(x) = \{v \in \mathbb{R}^n | f(y) \geq f(x) + \langle v, y - x \rangle, \forall y \in \Omega\}.
$$

Without loss of generality, we assume that each function $f_i$ in the objectives is sub-differentiable in its domain $\text{dom}f_i, \forall i \in I$.

For set $X = \{x \in \mathbb{R}^n | Ax = b, x \in \Omega\}$, point $x^* \in X$ is said a strong Pareto optimality (denoted by SPO), if there does not exist $x \in X$ such that $f(x) \preceq f(x^*)$. Similarly, point $x^* \in X$ is said a weak Pareto optimality (denoted by WPO), if there does not exist $x \in X$ such that $f(x) \prec f(x^*)$.

Based on the solution definitions of WPO and SPO, we give the optimality conditions for (1.1).

Lemma 2.1. [25] Suppose that function $f_j$ is proper lower semi-continuous and convex on $\Omega$ for $j \in I$, and $\Omega \subset \mathbb{R}^n$ is a compact and convex set. Then

1. point $x^* \in \Omega$ is a WPO to (1.1) if there exist $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)^T \in \mathbb{R}^m_+$ with $\sum_{j \in I} \lambda_j = 1$ and $\gamma \in \mathbb{R}^I$ such that

$$
0 \in \sum_{j \in I} \lambda_j \partial f_j(x^*) - A^T \lambda + N_{\Omega}(x^*), \quad (2.1)
$$

where $N_{\Omega}(x^*)$ is the normal cone to the convex set $\Omega$ at $x^* \in \Omega$. Further, if $0 \prec \lambda$, then for any $x^* \in \Omega$, condition (2.1) is sufficient for it to be an SPO.

2. point $x^* \in \Omega$ is an SPO to (1.1), if there exist $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)^T \in \mathbb{R}^m_+$ with $\sum_{j \in I} \lambda_j = 1$ and $\gamma \in \mathbb{R}^I$ satisfying (2.1).
Lemma 2.2. [25] A WPO to problem (1.1) is equivalent to finding an optimal solution to the following problem

$$\min_{x \in \mathbb{R}^n} \max_{i \in I} f_i(x), \text{ s.t. } Ax = b, x \in \Omega,$$

Further, if each function $f_j$, $\forall j \in I$, is strictly convex, then an SPO to problem (1.1) is equivalent to finding an optimal solution problem (2.1).

To end this section, we summarize some relevant existing results on the Lagrangian function of constrained optimization problems.

It is well known that the Lagrange function of (1.1) is

$$L(x, \lambda, \gamma) = \sum_{i \in I} \lambda_i f_i(x) - \gamma^T (Ax - b), \quad \forall (x, \lambda, \gamma) \in \tilde{\Omega},$$

where $\tilde{\Omega} = \{(x, \lambda, \gamma) \in \Omega \times \mathbb{R}^m_+ \times \mathbb{R}^l \mid \sum_{j \in I} \lambda_j = 1\}$, and $\gamma$ is the Lagrange multiplier.

Then the dual problem of (1.1) is as follows

$$\begin{align*}
\left\{ \begin{array}{l}
\max \min_{\gamma \in \mathbb{R}^l} \min_{x \in \mathbb{R}^n} L(x, \lambda, \gamma), \\
\quad \text{s.t.} \quad (y - x)^T \mathcal{V} \geq 0, \forall y \in \Omega,
\end{array} \right.
\end{align*}$$

(2.2)

where $\mathcal{V} \in \partial L(x, \lambda, \gamma)$. Surely, the augmented Lagrangian function is constituted by combining Lagrangian function with an penalty item, and at each iteration of this method consists of minimizing the augmented Lagrangian function of (1.1) and the goal of updating the Lagrange multiplier. Specifically, starting with the initial point $\gamma^0 \in \mathbb{R}^l$, the iteration formula of the augmented Lagrangian method for (1.1) is as follows:

$$\begin{align*}
\left\{ \begin{array}{l}
x^{k+1} = \arg \min_{x \in \Omega} \sum_{i \in I} \lambda_i f_i(x) - \gamma^k (Ax - b) + \frac{\beta}{2} \|Ax - b\|^2, \\
\gamma^{k+1} = \gamma^k - \beta (Ax^{k+1} - b),
\end{array} \right.
\end{align*}$$

(2.3)

where $\beta > 0$ is the penalty parameter for the violation of the linear constraints.

3. The iteration complexity of ALM. In this section, we investigate the iteration complexity of the ALM. To proceed, we need to explore the properties of the sequence generated by ALM.

Lemma 3.1. Given $\lambda^k \in \mathbb{R}^m$, let $(x^{k+1}, \lambda_{k+1}, \gamma^{k+1})$ be generated by the iteration scheme (2.3). Then, for any feasible solution $(x, \lambda, \gamma)$ of the dual problem (2.2), it holds that

$$\beta [L(x^{k+1}, \lambda_{k+1}, \gamma^{k+1}) - L(x, \lambda, \gamma)] \geq \|\gamma^k - \gamma^{k+1}\|^2 + (\gamma - \gamma^k)^T (\gamma^k - \gamma^{k+1}).$$

Proof. By the convexity of $f_i(i \in I)$, we conclude that for any $(x, \lambda, \gamma)$

$$L(x^{k+1}, \lambda_{k+1}, \gamma^{k+1}) - L(x, \lambda, \gamma) \geq (x^{k+1} - x)^T \left( \sum_{i \in I} \lambda_i \partial f_i(x) + \gamma^T (Ax - b) - (\gamma^{k+1})^T (Ax^{k+1} - b) \right).$$

(3.1)

Since $(x, \lambda, \gamma)$ is the feasible solution of the dual problem (2.2) and $x^{k+1} \in \Omega$, one has

$$(x^{k+1} - x)^T \left( \sum_{i \in I} \lambda_i \partial f_i(x) - A^T \gamma \right) \geq 0,$$

i.e.,

$$(x^{k+1} - x)^T \sum_{i \in I} \lambda_i \partial f_i(x) \geq \gamma^T A(x^{k+1} - x).$$
Combining this with (3.1) yields that
\[
L(x^{k+1}, \lambda_{k+1}, \gamma^{k+1}) - L(x, \lambda, \gamma) = \gamma^T A(x^{k+1} - x) - (\gamma^{k+1})^T (A x^{k+1} - b) + \gamma^T (A x - b)
\]
\[
= (\gamma - \gamma^{k+1})^T (A x^{k+1} - b) - 2 \beta (\gamma^{k+1})^T (\gamma^{k+1} - \gamma^{k+1})
\]
and the desired result follows. □

**Lemma 3.2.** Suppose \( \gamma^{k+1} \) is generated by iteration scheme (2.3) for given \( \gamma^k \in R^l \).
Then
\[
\|\gamma^{k+1} - \gamma^*\| \leq \|\gamma - \gamma^*\| - \|\gamma - \gamma^{k+1}\| + 2 \beta (L(x^*, \lambda^*, \gamma^*) - L(x^{k+1}, \lambda_{k+1}, \gamma^{k+1})).
\]

**Proof.** Since \( (x^*, \lambda^*, \gamma^*) \) is dual feasible, from Lemma 3.1, it holds that for \( (x, \lambda, \gamma) = (x^*, \lambda^*, \gamma^*) \)
\[
(\gamma - \gamma^*)^T (\gamma - \gamma^{k+1}) \geq \|\gamma - \gamma^{k+1}\|^2 + 2 \beta (L(x^*, \lambda^*, \gamma^*) - L(x^{k+1}, \lambda_{k+1}, \gamma^{k+1})).
\]
Rearranging the above inequality and by a manipulation give
\[
\|\gamma^{k+1} - \gamma^*\|^2 = \|\gamma - \gamma^*\|^2 - 2 \beta (\gamma^{k+1})^T (\gamma^{k+1} - \gamma^{k+1}) + \|\gamma - \gamma^{k+1}\|^2
\]
\[
\leq \|\gamma - \gamma^*\|^2 - \|\gamma - \gamma^{k+1}\|^2 - 2 \beta (L(x^*, \lambda^*, \gamma^*) - L(x^{k+1}, \lambda_{k+1}, \gamma^{k+1})),
\]
and the desired result follows. □

From Lemmas 3.1 and 3.2, the following conclusion is obvious.

**Lemma 3.3.** Let \( (x^{k+1}, \lambda_{k+1}, \gamma^{k+1}) \) be generated by iteration scheme (2.3). Then
\[
\beta L(x^{k+1}, \lambda_{k+1}, \gamma^{k+1}) \geq \beta L(x^k, \lambda_k, \gamma^k) + \|\gamma^k - \gamma^{k+1}\|^2,
\]
\[
\|\gamma^{k+1} - \gamma^*\|^2 \leq \|\gamma - \gamma^*\|^2 - \|\gamma^{k+1} - \gamma^k\|^2.
\]

Now, we are at the position to establish the main result in this section.

**Theorem 3.1.** Let \( (x^{k+1}, \lambda^{k+1}, \gamma^{k+1}) \) be generated by iteration scheme (2.3). Then
\[
L(x^*, \lambda^*, \gamma^*) - L(x^k, \lambda_k, \gamma^k) \leq \frac{\|\gamma^0 - \gamma^*\|^2}{2 \beta k},
\]
where \( (x^*, \lambda^*, \gamma^*) \) is a solution of (2.2).

**Proof.** For any \( i \in I \), by Lemma 3.2, one has
\[
\|\gamma^{i+1} - \gamma^*\|^2 - \|\gamma^i - \gamma^*\|^2 + \|\gamma^i - \gamma^{i+1}\|^2 \leq 2 \beta (L(x^{i+1}, \lambda_{i+1}, \gamma^{i+1}) - L(x^*, \lambda^*, \gamma^*)).
\]
By summing the above inequality for \( i \) from 0 to \( k - 1 \), and making use of the fact that \( L(x^{i+1}, \lambda_{i+1}, \gamma^{i+1}) - L(x^*, \lambda^*, \gamma^*) \leq 0 \) give
\[
\|\gamma^k - \gamma^*\|^2 - \|\gamma^0 - \gamma^*\|^2 + \sum_{i=0}^{k-1} \|\gamma^i - \gamma^{i+1}\|^2
\]
\[
\leq 2 \beta \sum_{i=0}^{k-1} \left( L(x^{i+1}, \lambda_{i+1}, \gamma^{i+1}) - L(x^*, \lambda^*, \gamma^*) \right).
\]
On the other hand, by Lemma 3.1, for any \((x, \lambda, \gamma) = (x^i, \lambda_i, \gamma^i)\), it holds that
\[
\beta \left( L(x^{i+1}, \lambda_{i+1}, \gamma^{i+1}) - L(x^i, \lambda_i, \gamma^i) \right) \geq \|\gamma^i - \gamma^{i+1}\|^2.
\]
Multiplying the above inequality by \(2i\) and summing it over \(i = 0, 1, \cdots, k - 1\) gives
\[
2\beta \sum_{i=0}^{k-1} \left( (i+1)L(x^{i+1}, \lambda_{i+1}, \gamma^{i+1}) - iL(x^i, \lambda_i, \gamma^i) - L(x^{i+1}, \lambda_{i+1}, \gamma^{i+1}) \right)
\geq 2 \sum_{i=0}^{k-1} i\|\gamma^i - \gamma^{i+1}\|^2,
\]
which implies that
\[
2\beta \left( kL(x^k, \lambda_k, \gamma_k) - \sum_{i=0}^{k-1} L(x^{i+1}, \lambda_{i+1}, \gamma^{i+1}) \right) \geq \sum_{i=0}^{k-1} 2i\|\gamma^i - \gamma^{i+1}\|^2.
\]
Combining this with (3.2) yields
\[
2k\beta \left( L(x^k, \lambda_k, \gamma_k) - L(x^*, \lambda_*, \gamma^*) \right) \geq \|\gamma^k - \gamma^*\|^2 - \|\gamma^0 - \gamma^*\|^2 + \sum_{i=0}^{k-1} (2i+1)\|\gamma^i - \gamma^{i+1}\|^2,
\]
which implies that
\[
L(x^*, \lambda_*, \gamma^*) - L(x^k, \lambda_k, \gamma_k) \leq \frac{\|\gamma^0 - \gamma^*\|^2}{2k\beta}.
\]
This means that the iteration complexity of the ALM is \(O(1/k)\). \(\square\)

4. The iteration complexity of AALM. In this section, we first show that the ALM can be accelerated by introducing the Nesterov technique [23], and then establish its iteration complexity of the accelerated ALM.

Inspired by the Nesterov work [23] of finding the zeros of variational inequality, we can establish the following iterative scheme for solving problem (1.1).

Accelerated ALM (AALM).

**Initial Step.** Take \(\hat{\gamma}^0 \in R^l\), set \(\gamma^1 = \hat{\gamma}^0, \lambda_1 \in R\) and \(t_1 = 1\).

**Iterative Step.** Set \(t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}\), and compute
\[
\hat{x}^k = \arg \min_{x \in \Omega} \sum_{i \in I} \lambda_i f_i(x) - (\gamma^k)^T(Ax - b) + \frac{\beta}{2} \|Ax - b\|^2,
\]
\[
\hat{\gamma}^k = \gamma^k - \beta(A\hat{x}^k - b),
\]
\[
\gamma^{k+1} = \gamma^k + \frac{t_k-1}{t_{k+1}}(\hat{\gamma}^k - \gamma^{k-1}) + \frac{t_k}{t_{k+1}}(\hat{\gamma}^k - \gamma^k).
\]

For this method, it follows from \(t_1 = 1\), \(t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}\) that
\[
t_k \geq \frac{k+1}{2}, \quad \forall k \geq 1.
\]

To investigate the iteration complexity of the AALM, we need the following notations.
\[
p_k = L(x^*, \lambda_*, \gamma^*) - L(\hat{x}^k, \hat{\lambda}_k, \hat{\gamma}^k),
\]
\[
q^k = t_k(2\hat{\gamma}^k - \gamma^k - \gamma^{k-1}) + \hat{\gamma}^{k-1} - \gamma^*.
\]
Lemma 4.1. For sequence \( \{ \gamma^k \} \) and \( \{ \hat{\gamma}^k \} \) generated by AALM, it holds that
\[
4\beta (t_k^2 p_k - t_{k+1}^2 p_{k+1}) \geq \| q^{k+1} \|^2 - \| q^k \|^2, \quad \forall k \geq 1,
\]
where \( p_k \) and \( q^k \) are defined by (4.3) and (4.4), respectively.

Proof. By Lemma 3.1, one has that for any \((x, \lambda, \gamma) = (\hat{x}^k, \hat{\lambda}_k, \hat{\gamma}^k)\) and \((x, \lambda, \gamma) = (\hat{x}^*, \hat{\lambda}_*, \hat{\gamma}^*)\), it holds that
\[
\beta \left( L(\hat{x}^{k+1}, \hat{\lambda}_{k+1}, \hat{\gamma}^{k+1}) - L(\hat{x}, \hat{\lambda}, \hat{\gamma}) \right) \geq \| \gamma^{k+1} - \hat{\gamma}^{k+1} \|^2 + (\hat{\gamma}^k - \gamma^{k+1})^T (\gamma^{k+1} - \hat{\gamma}^{k+1}),
\]
\[
\beta \left( L(\hat{x}^{k+1}, \hat{\lambda}_{k+1}, \hat{\gamma}^{k+1}) - L(x^*, \lambda_*, \gamma^*) \right) \geq \| \gamma^{k+1} - \hat{\gamma}^{k+1} \|^2 + (\hat{\gamma}^* - \gamma^{k+1})^T (\gamma^{k+1} - \hat{\gamma}^{k+1}).
\]

Using the notations (4.3), the above two inequality can be rewritten as
\[
\beta (p_k - p_{k+1}) \geq \| \gamma^{k+1} - \hat{\gamma}^{k+1} \|^2 + (\hat{\gamma}^k - \gamma^{k+1})^T (\gamma^{k+1} - \hat{\gamma}^{k+1}),
\]
(4.5)
\[
- \beta p_{k+1} \geq \| \gamma^{k+1} - \hat{\gamma}^{k+1} \|^2 + (\hat{\gamma}^* - \gamma^{k+1})^T (\gamma^{k+1} - \hat{\gamma}^{k+1}).
\]
(4.6)

Multiplying (4.5) by \( t_{k+1} - 1 \) and adding it to (4.6) give
\[
\beta \left( (t_{k+1} - 1)p_k - t_{k+1}p_{k+1} \right) \geq t_{k+1} \| \gamma^{k+1} - \hat{\gamma}^{k+1} \|^2 + (\hat{\gamma}^k - \gamma^{k+1})^T (t_{k+1}\gamma^{k+1} - (t_{k+1} - 1)\hat{\gamma}^k - \gamma^*).
\]

Multiplying (4.6) by \( t_{k+1} \) and using the fact that \( t_{k+1} = \frac{1+\sqrt{1+4\beta^2}}{2} \) yields
\[
\beta \left( (t_{k+1} - 1)p_k - t_{k+1}p_{k+1} \right) \geq t_{k+1} \| \gamma^{k+1} - \hat{\gamma}^{k+1} \|^2 + t_{k+1}(\hat{\gamma}^k - \gamma^{k+1})^T (t_{k+1}\gamma^{k+1} - (t_{k+1} - 1)\hat{\gamma}^k - \gamma^*).
\]

Since \( t_k^2 = t_{k+1}^2 - t_{k+1} \), from the above inequality, one has
\[
\beta \left( t_k^2 p_k - t_{k+1}^2 p_{k+1} \right) \geq t_{k+1}^2 \| \gamma^{k+1} - \hat{\gamma}^{k+1} \|^2 + t_{k+1}(\hat{\gamma}^k - \gamma^{k+1})^T (t_{k+1}\gamma^{k+1} - (t_{k+1} - 1)\hat{\gamma}^k - \gamma^*).
\]

Using the identity \( 4(b-a)^T(b-c) = (2b-a-c)^2 - (a-c)^2 \) to the right-hand side of the above inequality, we have
\[
4\beta (t_k^2 p_k - t_{k+1}^2 p_{k+1}) \geq \| t_k (2\hat{\gamma}^k - \gamma^k - \hat{\gamma}^{k-1}) + \hat{\gamma}^{k-1} - \gamma^* \|^2 - \| t_{k+1}(\gamma^{k+1} - \hat{\gamma}^k) + \hat{\gamma}^{k-1} - \gamma^* \|^2.
\]

Set \( q^k = t_k (2\hat{\gamma}^k - \gamma^k - \hat{\gamma}^{k-1}) + \hat{\gamma}^{k-1} - \gamma^* \). Then the above inequality can be written as
\[
4\beta (t_k^2 p_k - t_{k+1}^2 p_{k+1}) \geq \| q^{k+1} \|^2 - \| q_k \|^2. \quad (4.7)
\]

By (4.3), since
\[
t_{k+1}(\gamma^{k+1} - \hat{\gamma}^k) + \hat{\gamma}^k - \gamma^* = t_k (2\hat{\gamma}^k - \gamma^k - \hat{\gamma}^{k-1}) + \hat{\gamma}^{k-1} - \gamma^* = q^k,
\]
on one has from (4.7) that
\[
4\beta (t_k^2 p_k - t_{k+1}^2 p_{k+1}) \geq \| q^{k+1} \|^2 - \| q^k \|^2,
\]
and the desired conclusions follows. 

Based on the analysis above, we can establish the iteration complexity of the AALM.

Theorem 4.1. Let sequence \( \{ \gamma^k \} \) and \( \{ \hat{\gamma}^k \} \) be generated by AALM. Then
\[
L(x^*, \lambda_*, \gamma^*) - L((\hat{x}^k, \hat{\lambda}_k, \hat{\gamma}^k) \leq \frac{\| \gamma^1 - \gamma^* \|^2}{2\beta(k+1)^2}.
\]
Proof. From Lemma 4.1, one has
\[4\beta t_k^2 p_k + \|q_k\|^2 \geq 4\beta t_{k+1}^2 p_{k+1} + \|q_{k+1}\|^2.\]

Since \(\{q^k\}\) is nonnegative, it holds that
\[4\beta t_k^2 p_k \geq 4\beta t_k^2 p_k + \|q_k\|^2 \geq 4\beta t_k^2 p_1 + \|q_1\|^2.\]

Combining these two inequalities, we have that
\[L(x^*, \lambda*, \gamma*) - L(\hat{x}_k, \hat{\lambda}_k, \hat{\gamma}_k) = p_k \leq \frac{4\beta t_k^2 p_1 + \|q_1\|^2}{4\beta t_k^2}.\]

It is obvious that for \(t_1 = 1\)
\[4\beta t_1^2 p_1 = 4\beta p_1 = 4\beta \left(L(x^*, \lambda*, \gamma*) - L(\hat{x}_1, \hat{\lambda}_1, \hat{\gamma}_1)\right)\]
and
\[q^1 = \|2\hat{\gamma}^1 - \gamma^1 - \gamma^*\|^2.\]

Further, Lemma 3.2 implies that
\[\|\hat{\gamma}^k - \gamma^*\|^2 \leq \|\gamma^k - \gamma^*\|^2 - \|\hat{\gamma}^k - \gamma^k\|^2 - 2\beta \left(L(x^*, \lambda*, \gamma*) - L(\hat{x}_k, \hat{\lambda}_k, \hat{\gamma}_k)\right),\]
that is
\[2\beta \left(L(x^*, \lambda*, \gamma*) - L(\hat{x}_k, \hat{\lambda}_k, \hat{\gamma}_k)\right) \leq \|\gamma^k - \gamma^*\|^2 - \|\hat{\gamma}^k - \gamma^k\|^2 - \|\hat{\gamma}^k - \gamma^*\|^2.\]

Setting \(k = 1\) in the above inequality yields
\[2\beta \left(L(x^*, \lambda*, \gamma*) - L(\hat{x}_1, \hat{\lambda}_1, \hat{\gamma}_1)\right) \leq \|\gamma^1 - \gamma^*\|^2 - \|\hat{\gamma}^1 - \gamma^1\|^2 - \|\hat{\gamma}^1 - \gamma^*\|^2.\]

By the identity \(2(a-c)^2 - 2(b-c)^2 - 2(a-b)^2 = (a-c)^2 - ((a-b) + (b-c))^2,\) the above inequality can be written as
\[4\beta p_1 = 4\beta \left(L(x^*, \lambda*, \gamma*) - L(\hat{x}_1, \hat{\lambda}_1, \hat{\gamma}_1)\right) \leq \|\gamma^1 - \gamma^*\|^2 - \|2\hat{\gamma}^1 - \gamma^1 - \gamma^*\|^2,\]
which means that
\[4\beta p_1 + \|q^1\|^2 \leq \|\gamma^1 - \gamma^*\|^2.\]

As \(t_k \geq \frac{k+1}{2}\), it is easy to verify that
\[L(x^*, \lambda*, \gamma^*) - L(\hat{x}_k, \hat{\lambda}_k, \hat{\gamma}_k) \leq \frac{\|\gamma^1 - \gamma^*\|^2}{4\beta t_k^2} \leq \frac{\|\gamma^1 - \gamma^*\|^2}{\beta(k+1)^2},\]
which means that iteration complexity of the AALM is \(O(1/k^2).\)

\[\Box\]

5. Conclusion. In this paper, by exploring of the Nesterov’s acceleration technique, we proposed the accelerated augment Lagrangian method and established the iteration complexity of the new method which shows that the new method is accelerated compared to the the ALM.

Acknowledgments. The authors wish to give their grateful thanks to two anonymous referees for their insightful comments and suggestions, which improve the paper.
REFERENCES


Received December 2016; revised November 2017.

E-mail address: yonggk@163.com
E-mail address: wang-yiju@163.com
E-mail address: wgllj1977@163.com