Hindawi Publishing Corporation
Abstract and Applied Analysis
Volume 2013, Article ID 912846, 6 pages
http://dx.doi.org/10.1155/2013/912846

Research Article

Convergence Analysis of the Relaxed Proximal Point Algorithm

Min Li[1] and Yanfei You[2]

[1] School of Economics and Management, Southeast University, Nanjing 210096, China
[2] Department of Mathematics, Nanjing University, Nanjing 210093, China

Correspondence should be addressed to Min Li; liminnju@yahoo.com

Received 3 May 2013; Accepted 9 June 2013

Academic Editor: Abdellah Bnouhachem

Recently, a worst-case $O(1/t)$ convergence rate was established for the Douglas-Rachford alternating direction method of multipliers (ADMM) in an ergodic sense. The relaxed proximal point algorithm (PPA) is a generalization of the original PPA which includes the Douglas-Rachford ADMM as a special case. In this paper, we provide a simple proof for the same convergence rate of the relaxed PPA in both ergodic and nonergodic senses.

1. Introduction

The finite-dimensional variational inequality (VI), denoted by $\mathrm{VI}(\Omega, F)$, is to find a vector $w^* \in \Omega$ such that

$$\left(w - w^*\right)^T F\left(w^*\right) \geq 0, \quad \forall w \in \Omega, \tag{1}$$

where Ω is a nonempty closed convex set in $\mathfrak{R}^n$ and F is a monotone mapping from $\mathfrak{R}^n$ into itself. The solution set, denoted by Ω^* is assumed to be nonempty. We refer to [1–4] for the pivotal roles of VIs in various fields such as economics, transportation, and engineering.

As is well known, proximal point algorithm (PPA), which was presented originally in [5] and mainly developed in [6, 7], is a well-developed approach to solving $\mathrm{VI}(\Omega, F)$. Let w^k be the current approximation of a solution of (1); then PPA generates the new iterate $w^{k+1} \in \Omega$ by solving the following auxiliary VI:

$$\left(w - w^{k+1}\right)^T \left[F\left(w^{k+1}\right) + \frac{1}{\beta}\left(w^{k+1} - w^k\right)\right] \geq 0, \tag{2}$$

where β is a positive constant. Compared to the monotone VI (1), (2) is easier to handle since it is a strongly monotone VI. In this paper, we focus on the relaxed proximal point algorithm (PPA) proposed by Gol'shtein and Tret'yakov in [8], which combines the PPA step (3a) with a relaxation step (3b) as follows:

$$\widetilde{w}^k \in \Omega, \quad \left(w - \widetilde{w}^k\right)^T \left[F\left(\widetilde{w}^k\right) + G\left(\widetilde{w}^k - w^k\right)\right] \geq 0, \quad \forall w \in \Omega, \tag{3a}$$

$$w^{k+1} := w^k - \gamma\left(w^k - \widetilde{w}^k\right), \tag{3b}$$

where $\gamma \in (0, 2)$ is a relaxation factor and G is a symmetric positive semidefinite matrix. In particular, γ is called an under-relaxation factor when $\gamma \in (0, 1)$ or an over-relaxation factor when $\gamma \in (1, 2)$, and the relaxed PPA reduces to the original PPA (2) when $\gamma = 1$ and $G = (1/\beta)I$. For convenience, we still use the notation $\|w\|_G^2$ to represent the nonnegative number $w^T G w$ in our analysis.

The Douglas-Rachford alternating direction methods of multipliers (ADMM) scheme proposed by Glowinski and Marrocco in [9] (see also [10]) is a commonplace tool to solve the convex minimization problem with linear constraints and a separable objective function as follows:

$$\min\left\{\theta_1(x) + \theta_2(y) \mid Ax + By = b,\ x \in \mathcal{X},\ y \in \mathcal{Y}\right\}, \tag{4}$$

where $A \in \mathfrak{R}^{m\times n_1}$, $B \in \mathfrak{R}^{m\times n_2}$, $b \in \mathfrak{R}^m$, $\mathcal{X} \subseteq \mathfrak{R}^{n_1}$, and $\mathcal{Y} \subseteq \mathfrak{R}^{n_2}$ are closed convex sets and $\theta_1\colon \mathfrak{R}^{n_1} \to \mathfrak{R}$ and

$\theta_2: \mathfrak{R}^{n_2} \rightarrow \mathfrak{R}$ are convex smooth functions. The iterative scheme of ADMM for solving (4) at the k-th iteration runs as

$$x^{k+1} \in \mathcal{X}, \quad (x - x^{k+1})^T \{\nabla\theta_1 (x^{k+1}) - A^T [\lambda^k - H (Ax^{k+1} + By^k - b)]\} \ge 0, \quad \forall x \in \mathcal{X}, \tag{5a}$$

$$y^{k+1} \in \mathcal{Y}, \quad (y - y^{k+1})^T \{\nabla\theta_2 (y^{k+1}) - B^T [\lambda^k - H (Ax^{k+1} + By^{k+1} - b)]\} \ge 0, \quad \forall y \in \mathcal{Y}, \tag{5b}$$

$$\lambda^{k+1} := \lambda^k - H (Ax^{k+1} + By^{k+1} - b), \tag{5c}$$

where $H := hI$ and h is a positive constant. As shown in [11], ADMM can be regarded as an application of the relaxed PPA with $\gamma = 1$ (i.e., the original PPA (2)) and

$$G = \begin{pmatrix} 0 & 0 & 0 \\ 0 & B^T HB & -B^T \\ 0 & -B & H^{-1} \end{pmatrix}. \tag{6}$$

Without further assumption on B, the matrix G defined previously can be guaranteed as a symmetric and positive semidefinite matrix. Recently, He and Yuan in [12] have shown a worst-case $O(1/t)$ convergence rate of the ADMM scheme (5a), (5b), and (5c) in an ergodic sense. You et al. in [13] have proved the same convergence rate of the Lagrangian PPA-based contraction methods with nonsymmetric linear proximal term in an ergodic sense. The purpose of this paper is to establish the $O(1/t)$ convergence rate of the relaxed PPA (3a) and (3b) in both ergodic and nonergodic senses.

2. Preliminaries

In this section, we review some preliminaries which are useful for further discussions. More specially, we recall a useful characterization on Ω^*, the variational reformulation of (4), the relationship of the ADMM in [9, 10], and the relaxed PPA in [8] for solving this variational reformulation.

First, we provide a useful characterization on Ω^* as Theorem 2.3.5 in [14] and Theorem 2.1 in [12].

Theorem 1. *The solution set of* VI(Ω, F) *is convex, and it can be characterized as*

$$\Omega^* = \bigcap_{w \in \Omega} \{\widetilde{w} \in \Omega : (w - \widetilde{w})^T F(w) \ge 0\}. \tag{7}$$

Based on Theorem 1, $\widetilde{w} \in \Omega$ can be regarded as an ε-approximation solution of VI(Ω, F) if it satisfies

$$\sup_{w \in \mathcal{D}} \{(\widetilde{w} - w)^T F(w)\} \le \varepsilon, \tag{8}$$

where $\mathcal{D} \subseteq \Omega$ is some compact set. As Definition 1 in [15], we can take

$$\mathcal{D} = \mathcal{B}_\Omega (\widetilde{w}) := \{w \in \Omega \mid \|w - \widetilde{w}\| \le 1\}. \tag{9}$$

In the following, we will give a variational reformulation of (4). It is easy to see that the model (4) can be characterized by a variational inequality problem: find $w^* = (x^*, y^*, \lambda^*) \in \Omega := \mathcal{X} \times \mathcal{Y} \times \mathfrak{R}^m$ such that

$$\mathrm{VI}(\Omega, F) : (w - w^*)^T F(w^*) \ge 0, \quad \forall w \in \Omega, \tag{10a}$$

where

$$w = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}, \quad F(w) = \begin{pmatrix} \nabla\theta_1 (x) - A^T\lambda \\ \nabla\theta_2 (y) - B^T\lambda \\ Ax + By - b \end{pmatrix}. \tag{10b}$$

Note that the mapping F is monotone since θ_1 and θ_2 are convex. As shown in [11], the ADMM scheme (5a), (5b), and (5c) is identical with the following iterative scheme in a cyclical sense:

$$\widetilde{x}^k \in \mathcal{X}, \quad (x - \widetilde{x}^k)^T \{\nabla\theta_1 (\widetilde{x}^k) - A^T [\lambda^k - H (A\widetilde{x}^k + By^k - b)]\} \ge 0, \quad \forall x \in \mathcal{X}, \tag{11a}$$

$$\widetilde{\lambda}^k := \lambda^k - H (A\widetilde{x}^k + By^k - b), \tag{11b}$$

$$\widetilde{y}^k \in \mathcal{Y}, \quad (y - \widetilde{y}^k)^T \{\nabla\theta_2 (\widetilde{y}^k) - B^T [\widetilde{\lambda}^k - H (A\widetilde{x}^k + B\widetilde{y}^k - b)]\} \ge 0, \quad \forall y \in \mathcal{Y}, \tag{11c}$$

$$w^{k+1} = w^k - (w^k - \widetilde{w}^k). \tag{12}$$

Based on the definition (6) of the matrix G, we can rewrite (11a), (11b), (11c), and (12) as a special case of the relaxed PPA with $\gamma = 1$ immediately.

Lemma 2. *For given w^k, let $\widetilde{w}^k$ be generated by the ADMM scheme* (11a), (11b), *and* (11c). *Then, one has*

$$\widetilde{w}^k \in \Omega, \quad (w - \widetilde{w}^k)^T \{F(\widetilde{w}^k) + G(\widetilde{w}^k - w^k)\} \ge 0, \quad \forall w \in \Omega, \tag{13}$$

where F and G are defined by (10b) *and* (6), *respectively.*

3. The Contraction of the Relaxed Proximal Point Algorithm

In this section, we prove the contraction of the relaxed PPA. First, we give an important lemma.

Lemma 3. *Let the sequences* $\{w^k\}$ *and* $\{\widetilde{w}^k\}$ *be generated by the relaxed PPA* (3a) *and* (3b), *and let G be a symmetric positive semidefinite matrix. Then, one has*

$$\begin{aligned}(w-\widetilde{w}^k)^T F(\widetilde{w}^k) &\geq \frac{1}{2\gamma}\left(\left\|w-w^{k+1}\right\|_G^2-\left\|w-w^k\right\|_G^2\right)\\ &\quad+\left(1-\frac{\gamma}{2}\right)\left\|w^k-\widetilde{w}^k\right\|_G^2, \quad \forall w\in\Omega.\end{aligned} \tag{14}$$

Proof. First, using (3a), we have

$$(w-\widetilde{w}^k)^T F(\widetilde{w}^k) \geq (w-\widetilde{w}^k)^T G(w^k-\widetilde{w}^k), \quad \forall w\in\Omega. \tag{15}$$

Since $w^k-\widetilde{w}^k=(w^k-w^{k+1})/\gamma$ (see (3b)), we have

$$(w-\widetilde{w}^k)^T G(w^k-\widetilde{w}^k)=\frac{1}{\gamma}(w-\widetilde{w}^k)^T G(w^k-w^{k+1}). \tag{16}$$

Thus, it suffices to show that

$$\begin{aligned}(w-\widetilde{w}^k)^T G(w^k-w^{k+1}) &= \frac{1}{2}\left(\left\|w-w^{k+1}\right\|_G^2-\left\|w-w^k\right\|_G^2\right)\\ &\quad+\gamma\left(1-\frac{\gamma}{2}\right)\left\|w^k-\widetilde{w}^k\right\|_G^2.\end{aligned} \tag{17}$$

By setting $a=w$, $b=\widetilde{w}^k$, $c=w^k$, and $d=w^{k+1}$ in the identity

$$\begin{aligned}(a-b)^T G(c-d) &= \frac{1}{2}\left(\|a-d\|_G^2-\|a-c\|_G^2\right)\\ &\quad+\frac{1}{2}\left(\|c-b\|_G^2-\|d-b\|_G^2\right),\end{aligned} \tag{18}$$

we derive that

$$\begin{aligned}(w-\widetilde{w}^k)^T G(w^k-w^{k+1}) &= \frac{1}{2}\left(\left\|w-w^{k+1}\right\|_G^2-\left\|w-w^k\right\|_G^2\right)\\ &\quad+\frac{1}{2}\left(\left\|w^k-\widetilde{w}^k\right\|_G^2-\left\|w^{k+1}-\widetilde{w}^k\right\|_G^2\right).\end{aligned} \tag{19}$$

On the other hand, using (3b), we have

$$\begin{aligned}&\left\|w^k-\widetilde{w}^k\right\|_G^2-\left\|w^{k+1}-\widetilde{w}^k\right\|_G^2\\ &\quad=\left\|w^k-\widetilde{w}^k\right\|_G^2-\left\|(w^k-\widetilde{w}^k)-(w^k-w^{k+1})\right\|_G^2\\ &\quad=\left\|w^k-\widetilde{w}^k\right\|_G^2-\left\|(w^k-\widetilde{w}^k)-\gamma(w^k-\widetilde{w}^k)\right\|_G^2\\ &\quad=\gamma(2-\gamma)\left\|w^k-\widetilde{w}^k\right\|_G^2.\end{aligned} \tag{20}$$

Combining the last two equations, we obtain (17). The assertion (14) follows immediately. The proof is completed. □

With the proved lemma, we are now ready to show the contraction of the relaxed PPA (3a) and (3b).

Theorem 4. *Let the sequences* $\{w^k\}$ *and* $\{\widetilde{w}^k\}$ *be generated by the relaxed PPA* (3a) *and* (3b), *and let G be a symmetric positive semidefinite matrix. Then, for any* $k\geq 0$, *one has*

$$\begin{aligned}&\left\|w^{k+1}-w^*\right\|_G^2\\ &\quad\leq\left\|w^k-w^*\right\|_G^2-\gamma(2-\gamma)\left\|w^k-\widetilde{w}^k\right\|_G^2, \quad \forall w^*\in\Omega^*.\end{aligned} \tag{21}$$

Proof. Setting $w=w^*$ in (14), we get

$$\begin{aligned}2\gamma(w^*-\widetilde{w}^k)^T F(\widetilde{w}^k) &\geq \left\|w^*-w^{k+1}\right\|_G^2-\left\|w^*-w^k\right\|_G^2\\ &\quad+\gamma(2-\gamma)\left\|w^k-\widetilde{w}^k\right\|_G^2.\end{aligned} \tag{22}$$

On the other hand, since F is monotone and $w^*\in\Omega^*$, we have

$$0\geq(w^*-\widetilde{w}^k)^T F(w^*)\geq(w^*-\widetilde{w}^k)^T F(\widetilde{w}^k). \tag{23}$$

It follows from the previous two inequalities that

$$\left\|w^{k+1}-w^*\right\|_G^2\leq\left\|w^k-w^*\right\|_G^2-\gamma(2-\gamma)\left\|w^k-\widetilde{w}^k\right\|_G^2. \tag{24}$$

The proof is completed. □

4. Ergodic Worst-Case $O(1/t)$ Convergence Rate

In this section, we will establish an ergodic worst-case $O(1/t)$ convergence rate for the relaxed PPA in the sense that after t iterations of such an algorithm, we can find $\widetilde{w}\in\Omega$ such that

$$(\widetilde{w}-w)^T F(w)\leq\varepsilon, \quad \forall w\in\mathscr{B}_\Omega(\widetilde{w}), \tag{25}$$

with $\varepsilon=O(1/t)$ and $\mathscr{B}_\Omega(\widetilde{w}):=\{w\in\Omega\mid\|w-\widetilde{w}\|_G\leq 1\}$.

Theorem 5. *Let $\{w^k\}$ and $\{\widetilde{w}^k\}$ be the sequences generated by the relaxed PPA* (3a) *and* (3b)*, and let G be a symmetric positive semidefinite matrix. For any integer number $t > 0$, let*

$$\widetilde{w}_t := \frac{1}{t+1}\sum_{k=0}^{t}\widetilde{w}^k. \quad (26)$$

Then, one has $\widetilde{w}_t \in \Omega$ and

$$(\widetilde{w}_t - w)^T F(w) \le \frac{1}{2\gamma(t+1)}\left\|w^0 - w\right\|_G^2, \quad \forall w \in \Omega. \quad (27)$$

Proof. From (14), we have

$$\left(w - \widetilde{w}^k\right)^T F\left(\widetilde{w}^k\right) + \frac{1}{2\gamma}\left\|w^k - w\right\|_G^2 \ge \frac{1}{2\gamma}\left\|w^{k+1} - w\right\|_G^2, \quad \forall w \in \Omega. \quad (28)$$

Since F is monotone, from the previous inequality, we have

$$\left(w - \widetilde{w}^k\right)^T F(w) + \frac{1}{2\gamma}\left\|w^k - w\right\|_G^2 \ge \frac{1}{2\gamma}\left\|w^{k+1} - w\right\|_G^2, \quad \forall w \in \Omega. \quad (29)$$

Summing the inequality (29) over $k = 0, 1, \ldots, t$, we obtain

$$\left[(t+1)w - \left(\sum_{k=0}^{t}\widetilde{w}^k\right)\right]^T F(w) + \frac{1}{2\gamma}\left\|w^0 - w\right\|_G^2 \ge \frac{1}{2\gamma}\left\|w^{t+1} - w\right\|_G^2 \ge 0, \quad \forall w \in \Omega. \quad (30)$$

Since $\sum_{k=0}^{t} 1/(t+1) = 1$, $\widetilde{w}_t$ is a convex combination of $\widetilde{w}^0, \widetilde{w}^1, \ldots, \widetilde{w}^t$ and thus $\widetilde{w}_t \in \Omega$. Using the notation of $\widetilde{w}_t$, we derive

$$(w - \widetilde{w}_t)^T F(w) + \frac{1}{2\gamma(t+1)}\left\|w^0 - w\right\|_G^2 \ge 0, \quad \forall w \in \Omega. \quad (31)$$

The assertion (27) follows from the previous inequality immediately. □

It follows from Theorem 4 that the sequence $\{\|w^k\|_G\}$ is bounded. According to (21), the sequence $\{\|\widetilde{w}^k\|_G\}$ is also bounded. Therefore, there exists a constant $D > 0$ such that

$$\left\|w^k\right\|_G \le D, \qquad \left\|\widetilde{w}^k\right\|_G \le D, \quad \forall k \ge 0. \quad (32)$$

Recall that $\widetilde{w}_t$ is the average of $\{\widetilde{w}^0, \widetilde{w}^1, \ldots, \widetilde{w}^t\}$. Thus, we have $\|\widetilde{w}_t\|_G \le D$. For any $w \in \mathscr{B}_\Omega(\widetilde{w}_t) := \{w \in \Omega \mid \|w - \widetilde{w}_t\|_G \le 1\}$, we get

$$\begin{aligned}(\widetilde{w}_t - w)^T F(w) &\le \frac{1}{2\gamma(t+1)}\left\|w^0 - w\right\|_G^2 \\ &\le \frac{1}{2\gamma(t+1)}\left(\left\|w^0 - \widetilde{w}_t\right\|_G + \left\|\widetilde{w}_t - w\right\|_G\right)^2 \\ &\le \frac{1}{2\gamma(t+1)}\left(\left\|w^0\right\|_G + \left\|\widetilde{w}_t\right\|_G + \left\|\widetilde{w}_t - w\right\|_G\right)^2 \\ &\le \frac{(2D+1)^2}{2\gamma(t+1)}.\end{aligned} \quad (33)$$

Thus, for any given $\varepsilon > 0$, after at most $t := \lceil((2D+1)^2/2\gamma\varepsilon) - 1\rceil$ iterations, we have

$$(\widetilde{w}_t - w)^T F(w) \le \varepsilon, \quad \forall w \in \mathscr{B}_\Omega(\widetilde{w}_t), \quad (34)$$

which means that $\widetilde{w}_t$ is an approximate solution of $\mathrm{VI}(\Omega, F)$ with an accuracy of $O(1/t)$. That is, a worst-case $O(1/t)$ convergence rate of the relaxed PPA in an ergodic sense is established.

Note that this convergence rate is in an ergodic sense and $\widetilde{w}_t$ is a convex combination of the previous vectors $\{\widetilde{w}^0, \widetilde{w}^1, \ldots, \widetilde{w}^t\}$ with equal weights. One may ask if we can establish the same convergence rate in a nonergodic sense directly for the sequence $\{w^k\}$ generated by the relaxed PPA (3a) and (3b), and this is the main purpose of the next section.

5. Nonergodic Worst-Case $O(1/t)$ Convergence Rate

This section shows that the relaxed PPA has a worst-case $O(1/t)$ convergence rate in a nonergodic sense. First, we establish two important inequalities in the following lemmas.

Lemma 6. *Let the sequences $\{w^k\}$ and $\{\widetilde{w}^k\}$ be generated by the relaxed PPA* (3a) *and* (3b)*, and let G be a symmetric positive semidefinite matrix. Then, one has*

$$\left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)^T G\left[\left(w^k - w^{k+1}\right) - \left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)\right] \ge 0. \quad (35)$$

Proof. Setting $w = \widetilde{w}^{k+1}$ in (3a), we have

$$\left(\widetilde{w}^{k+1} - \widetilde{w}^k\right)^T \left[F\left(\widetilde{w}^k\right) + G\left(\widetilde{w}^k - w^k\right)\right] \ge 0. \quad (36)$$

Note that (3a) is also true for $k := k+1$, and thus we have

$$\left(w - \widetilde{w}^{k+1}\right)^T \left[F\left(\widetilde{w}^{k+1}\right) + G\left(\widetilde{w}^{k+1} - w^{k+1}\right)\right] \ge 0, \quad \forall w \in \Omega. \quad (37)$$

Setting $w = \widetilde{w}^k$ in the previous inequality, we obtain

$$\left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)^T \left[F\left(\widetilde{w}^{k+1}\right) + G\left(\widetilde{w}^{k+1} - w^{k+1}\right)\right] \geq 0. \quad (38)$$

Adding (36) and (38) and using the monotonicity of F, we get (35) immediately. □

Lemma 7. *Let the sequences* $\{w^k\}$ *and* $\{\widetilde{w}^k\}$ *be generated by the relaxed PPA* (3a) *and* (3b), *and let* G *be a symmetric positive semidefinite matrix. Then, one has*

$$\begin{aligned}\left(w^k - \widetilde{w}^k\right)^T G\left\{\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\} \\ \geq \frac{1}{\gamma}\left\|\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\|_G^2.\end{aligned} \quad (39)$$

Proof. First, adding the term

$$\begin{aligned}\left\{\left(w^k - w^{k+1}\right) - \left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)\right\}^T \\ \times G\left\{\left(w^k - w^{k+1}\right) - \left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)\right\}\end{aligned} \quad (40)$$

to the both sides of (35), we get

$$\begin{aligned}\left(w^k - w^{k+1}\right)^T G\left\{\left(w^k - w^{k+1}\right) - \left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)\right\} \\ \geq \left\|\left(w^k - w^{k+1}\right) - \left(\widetilde{w}^k - \widetilde{w}^{k+1}\right)\right\|_G^2.\end{aligned} \quad (41)$$

Reordering $(w^k - w^{k+1}) - (\widetilde{w}^k - \widetilde{w}^{k+1})$ in the previous inequality to $(w^k - \widetilde{w}^k) - (w^{k+1} - \widetilde{w}^{k+1})$, we get

$$\begin{aligned}\left(w^k - w^{k+1}\right)^T G\left\{\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\} \\ \geq \left\|\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\|_G^2.\end{aligned} \quad (42)$$

Substituting the term $w^k - w^{k+1} = \gamma(w^k - \widetilde{w}^k)$ (see (3b)) into the left-hand side of the last inequality, we obtain (39). The proof is completed. □

Next, we prove that $\{\|w^k - \widetilde{w}^k\|_G\}$ is monotonically nonincreasing.

Theorem 8. *Let the sequences* $\{w^k\}$ *and* $\{\widetilde{w}^k\}$ *be generated by the relaxed PPA* (3a) *and* (3b), *and let* G *be a symmetric positive semidefinite matrix. Then, one has*

$$\left\|w^{k+1} - \widetilde{w}^{k+1}\right\|_G \leq \left\|w^k - \widetilde{w}^k\right\|_G, \quad \forall k \geq 0. \quad (43)$$

Proof. Setting $a = w^k - \widetilde{w}^k$ and $b = w^{k+1} - \widetilde{w}^{k+1}$ in the identity

$$\|a\|_G^2 - \|b\|_G^2 = 2a^T G(a - b) - \|a - b\|_G^2, \quad (44)$$

we obtain

$$\begin{aligned}&\left\|w^k - \widetilde{w}^k\right\|_G^2 - \left\|w^{k+1} - \widetilde{w}^{k+1}\right\|_G^2 \\ &\quad = 2\left(w^k - \widetilde{w}^k\right)^T G\left\{\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\} \\ &\quad\quad - \left\|\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\|_G^2.\end{aligned} \quad (45)$$

Inserting (39) into the first term of the right-hand side of the last equality and using $\gamma \in (0, 2)$, we obtain

$$\begin{aligned}&2\left(w^k - \widetilde{w}^k\right)^T G\left\{\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\} \\ &\quad - \left\|\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\|_G^2 \\ &\geq \frac{2 - \gamma}{\gamma}\left\|\left(w^k - \widetilde{w}^k\right) - \left(w^{k+1} - \widetilde{w}^{k+1}\right)\right\|_G^2 \\ &\geq 0.\end{aligned} \quad (46)$$

The assertion (43) follows immediately. □

With Theorems 4 and 8, we can prove the worst-case $O(1/t)$ convergence rate in a nonergodic sense for the relaxed PPA.

Theorem 9. *Let the sequences* $\{w^k\}$ *and* $\{\widetilde{w}^k\}$ *be generated by the relaxed PPA* (3a) *and* (3b), *and let* G *be a symmetric positive semidefinite matrix. Then, for any integer* $t \geq 0$, *one has*

$$\left\|w^t - \widetilde{w}^t\right\|_G^2 \leq \frac{1}{\gamma(2 - \gamma)(t + 1)}\left\|w^0 - w^*\right\|_G^2, \quad \forall w^* \in \Omega^*. \quad (47)$$

Proof. Summing the inequality (21) over $k = 0, 1, \ldots, t$, we obtain

$$\begin{aligned}&\gamma(2 - \gamma)\sum_{k=0}^{t}\left\|w^k - \widetilde{w}^k\right\|_G^2 \\ &\quad \leq \left\|w^0 - w^*\right\|_G^2 - \left\|w^{t+1} - w^*\right\|_G^2 \\ &\quad \leq \left\|w^0 - w^*\right\|_G^2, \quad \forall w^* \in \Omega^*.\end{aligned} \quad (48)$$

According to Theorem 8, the sequence $\{\|w^k - \widetilde{w}^k\|_G\}$ is monotonically nonincreasing. Therefore, we have

$$(t + 1)\left\|w^t - \widetilde{w}^t\right\|_G^2 \leq \sum_{k=0}^{t}\left\|w^k - \widetilde{w}^k\right\|_G^2. \quad (49)$$

The assertion (47) follows from (48) and (49) immediately. □

Note that Ω^* is convex and closed (see Theorem 1). Let $d := \inf\{\|w^0 - w^*\|_G \mid w^* \in \Omega^*\}$. Then, for any given $\varepsilon > 0$, Theorem 9 shows that the relaxed PPA (3a) and (3b) needs at most $\lceil d^2/(\varepsilon\gamma(2-\gamma)) - 1\rceil$ iterations to ensure that $\|w^t - \widetilde{w}^t\|_G^2 \leq \varepsilon$. Recall that $\widetilde{w}^t$ is a solution of VI(Ω, F) if $\|w^t - \widetilde{w}^t\|_G^2 = 0$. In other words, if $\|w^t - \widetilde{w}^t\|_G^2 = 0$, we have $G(w^t - \widetilde{w}^t) = 0$ since G is a positive semidefinite matrix. And thus from (3a), it follows that

$$\left(w - \widetilde{w}^t\right)^T F\left(\widetilde{w}^t\right) \geq 0, \quad \forall w \in \Omega, \quad (50)$$

which means that $\widetilde{w}^t$ is a solution of VI(Ω, F) according to (1). A worst-case $O(1/t)$ convergence rate in a nonergodic sense for the relaxed PPA (3a) and (3b) is thus established from Theorem 9.

6. Concluding Remarks

This paper established the worst-case $O(1/t)$ convergence rate in both ergodic and nonergodic senses for the relaxed PPA. Recall that ADMM is a primal application of the relaxed PPA with $\gamma = 1$. And thus ADMM also has the same worst-case $O(1/t)$ convergence rate in both ergodic and nonergodic senses.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 11001053), Program for New Century Excellent Talents in University (Grant no. NCET-12-0111), and Natural Science Foundation of Jiangsu Province, China (Grant no. BK2012662).

References

[1] S. Dafermos and A. Nagurney, "Supply and demand equilibration algorithms for a class of market equilibrium problems," *Transportation Science*, vol. 23, no. 2, pp. 118–124, 1989.

[2] M. C. Ferris and J. S. Pang, "Engineering and economic applications of complementarity problems," *SIAM Review*, vol. 39, no. 4, pp. 669–713, 1997.

[3] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and their Applications*, vol. 88, Academic Press, New York, NY, USA, 1980.

[4] A. Nagurney, *Network Economics: A Variational Inequality Approach*, vol. 1, Kluwer Academic, Dordrecht, The Netherlands, 1993.

[5] B. Martinet, "Régularisation D'inéquations Variationnelles par Approximations Successives," *Revue Française dInformatique et de Recherche Opérationelle*, vol. 4, pp. 154–158, 1970.

[6] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM Journal on Control and Optimization*, vol. 14, no. 5, pp. 877–898, 1976.

[7] R. T. Rockafellar, "Augmented Lagrangians and applications of the proximal point algorithm in convex programming," *Mathematics of Operations Research*, vol. 1, no. 2, pp. 97–116, 1976.

[8] E. G. Gol'shtein and N. V. Tret'yakov, "Modified Lagrangians in convex programming and their generalizations," *Mathematical Programming Study*, no. 10, pp. 86–97, 1979.

[9] R. Glowinski and A. Marrocco, "Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéaires," *RAIRO*, vol. 9, no. R-2, pp. 41–76, 1975.

[10] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite-element approximations," *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[11] X. J. Cai, G. Y. Gu, B. S. He, and X. M. Yuan, *A Proximal Point Algorithm Revisit on the Alternating Direction Method of Multipliers*, Science China Mathematics.

[12] B. He and X. Yuan, "On the $O(1/n)$ convergence rate of the Douglas-Rachford alternating direction method," *SIAM Journal on Numerical Analysis*, vol. 50, no. 2, pp. 700–709, 2012.

[13] Y. F. You, X. L. Fu, and B. S. He, *Lagrangian PPA-Based Contraction Methods for Linearly Constrained Convex Optimization*, manuscript, 2013.

[14] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems, Springer Series in Operations Research*, Springer, New York, NY, USA, 2003.

[15] Yu. Nesterov, "Gradient methods for minimizing composite functions," *Mathematical Programming*, 2012.