Iteration complexity on the Generalized Peaceman-Rachford splitting method for separable convex programming *

Xue-Qing Zhang† · Jian-Wen Peng ‡

Abstract. Recently, a generalized version of Peaceman-Rachford splitting method (GPRSM) for solving a convex minmization model with a general separable structure has been proposed by **Sun** et al and its global convergence has been proved. In this paper, we further study theoretical aspects of the generalized Peaceman-Rachford splitting method. We first establish the worst-case $\mathcal{O}(1/t)$ convergence rate for the proposed GPRSM in both the ergodic and a nonergodic senses, then we give some numerical results to demonstrate the convergence rate of the algorithm.

Key Words: Convex minimization problem; Generalized Peaceman-Rachford splitting method; Iteration complexity; LASSO problem; Matrix optimization.

1 Introduction

We consider the following convex programming with linear constraints and an objective function which is sum of two functions without coupled variables

$$\begin{array}{ll} \min & f(x)+g(y) \\ \text{s.t.} & Ax+By=b. \\ & x\in\mathcal{X}, y\in\mathcal{Y}, \end{array} \tag{1.1}$$

where $f:\mathcal{R}^n\to\mathcal{R}$ and $g:\mathcal{R}^m\to\mathcal{R}$ are closed proper convex but not necessarily smooth functions, $A\in\mathcal{R}^{l\times n}$, $B\in\mathcal{R}^{l\times m}$, $b\in\mathcal{R}^l$, and $\mathcal{X}\subseteq\mathcal{R}^n$ and $\mathcal{Y}\subseteq\mathcal{R}^m$ are two nonempty closed

*Research of this author was supported by the National Natural Science Foundation of China (11431004), the Chongqing Research Program of Basic Research and Frontier Technology (cstc2015jcyjBX0029), the Education Committee Project Foundation of Chongqing for "Bayu scholar" Distinguished Professor, Key Laboratory of optimization and control of Ministry of Education in China.

†School of Mathematical Sciences, Chongqing Normal University, Chongqing 400047, P.R.China. E-mail: zxqcqspb@163.com

‡Corresponding author. School of Mathematical Science, Chongqing Normal University, Chongqing, 400047, P.R. China; E-mail: jwpeng168@hotmail.com

convex sets. In this paper, the solution set of (1.1) is assumed to be nonempty, and the matrix B are assumed to have full column rank.

In fact, the convex optimization model (1.1) contains many problems arising from compressed sensing, image deblurring, and statistical learning. There are many efficient methods to solve the model (1.1), such as the classical alternating direction method of multipliers (ADMM) [1,2], the generalized version of the alternating direction method of multipliers (GADMM) [3,4], the Peaceman-Rachford splitting method (PRSM) [5-9], and so on. Based on their high efficiency, many researchers have been studying these methods and show that the ADMM, the PRSM and their variants are quite efficient for solving some application problems. As elaborated in [16], applying PRSM to the problem (1.1), we obtain the following scheme of PRSM [6],

$$\begin{cases} x^{k+1} = \text{argmin}\{f(x) - (\lambda^k)^\top Ax + \frac{\rho}{2}\|Ax + By^k - b\|^2 | x \in \mathcal{X}\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \rho(Ax^{k+1} + By^k - b), \\ y^{k+1} = \text{argmin}\{f(y) - (\lambda^{k+\frac{1}{2}})^\top By + \frac{\rho}{2}\|Ax^{k+1} + By - b\|^2 | y \in \mathcal{Y}\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \rho(Ax^{k+1} + By^{k+1} - b). \end{cases} \tag{1.2}$$

Where $\lambda \in \mathcal{R}^l$ is the Lagrange multiplier of the linear constraints in (1.1) and $\rho > 0$ is a penalty parameter. The PRSM is always efficient if it is convergent (see [8, 17]). He et al. [7] proposed a strictly contractive PRSM (SCPRSM) by attaching an relaxation factor γ to the penalty parameter ρ in the steps of Lagrange multiplier updating in (1.2) and established a worst-case $\mathcal{O}(1/t)$ convergence rate of the iterative scheme (1.2) in the ergodic sense and a worst-case $\mathcal{O}(1/t)$ convergence rate of the SCPRSM in a nonergodic sense.

Based on the strictly contractive PRSM [7], Sun et al. [10] developed a generalized version of the PRSM (GPRSM). In this paper, we focus on the GPRSM, which was original proposed in [10] as follows

$$\begin{cases} x^{k+1} = \text{argmin}\{f(x) - (\lambda^k)^\top Ax + \frac{\rho}{2}\|Ax + By^k - b\|^2 + \frac{1}{2}\|x - x^k\|^2_{G_1} | x \in \mathcal{X}\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \gamma\rho(Ax^{k+1} + By^k - b), \\ y^{k+1} = \text{argmin}\{f(y) - (\lambda^{k+\frac{1}{2}})^\top By + \frac{\rho}{2}\|\alpha Ax^{k+1} - (1-\alpha)(By^k - b) + By - b\|^2 \\ \qquad\qquad + \frac{1}{2}\|y - y^k\|^2_{G_2} | y \in \mathcal{Y}\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \rho(\alpha Ax^{k+1} - (1-\alpha)(By^k - b) + By^{k+1} - b), \end{cases} \tag{1.3}$$

where $\alpha \in (0, 2), \gamma \in (0, 2-\alpha)$ are two relaxation factors, and $G_1 \succ \mathbf{0}, G_2 \succeq \mathbf{0}$ or $G_1 \succeq \mathbf{0}, G_2 \succ \mathbf{0}$. Here for a matrix $G \succ \mathbf{0}$ (resp., $\succeq \mathbf{0}$) means that G is positive definite (resp., semidefinite). And $\lambda \in \mathcal{R}^l$ is the Lagrange multiplier associated with the linear constraints in (1.1) and $\rho > 0$ is a penalty parameter. It is worthy noting that the GPRSM (1.3) is different from the SCPRSM [7] for the Lagrangian multiplier. Since it is not attached any relaxation factor before ρ in the second update of Lagrangian multiplier. Referring to [10], the reasons are: (i) If we attach a factor, then the matrix $\mathcal{T}$ (defined in (2.3)) will be nonsymmetric; (ii) As pointed out in [7], aggressive values of the relaxation factor γ are preferred.

The numerical results given in [10] indicate that GPRSM (1.3) is quite efficient for solving some optimization problems, such as statistical learning and image processing. We refer the reader to [9, 10, 13] for some applications and numerical verification of the efficiency of GPRSM (1.3). However, the convergence rate of the generalized version of the PRSM (1.3) is still in its infancy. In this paper, we establish the worst-case $\mathcal{O}(1/t)$ convergence rate for the proposed GPRSM (1.3) in both the ergodic and a nonergodic senses. And we also give some numerical results to further show the efficiency of the GPRSM (1.3) and show some analysis of the convergence rate of the algorithm (1.3).

The rest of this paper is organized as follows. In Sect.2, we summarize some preliminaries which are useful for further analysis. Then, we establish the worst-case convergence rate for the GPRSM (1.3) in the ergodic and a nonergodic senses in sect.3. In Sect.4, we show some numerical results. Finally, we draw some conclusions in Sect.5.

2 Preliminaries

In this section, we give some preliminaries which are useful for the following discussions.

First, it is useful to characterize of the model (1.1) by a variational inequality. More specially, as well known in the literature (see [11, 12]), solving (1.1) is equivalent to solving the following variational inequality (VI) problem: Finding $w^* \in \Omega$ such that

$$\Phi(u) - \Phi(u^*) + (w - w^*)^\top \mathcal{F}(w^*) \geq 0, \qquad \forall w \in \Omega, \tag{2.1}$$

where

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, \quad w = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}, \quad \mathcal{F}(w) = \begin{pmatrix} -A^\top \lambda \\ -B^\top \lambda \\ Ax + By - b \end{pmatrix}, \tag{2.2}$$

and $\Phi(u) = f(x) + g(y)$, $\Omega = \mathcal{X} \times \mathcal{Y} \times \Re^l$. We denote by VI$(\Omega, \mathcal{F}, \Phi)$ the problem (2.1-2.2). It is easy to see that the mapping $\mathcal{F}(w)$ defined in (2.2) is affine with a skew-symmetric matrix; it is thus monotone:

$$(w_1 - w_2)^\top (\mathcal{F}(w_1) - \mathcal{F}(w_2)) \geq 0, \qquad \forall \;\; w_1, w_2 \in \Omega.$$

Under the assumption that the solution set of (1.1) is nonempty, we denote the solution set of (2.1) by Ω^*.

We recall that a multifunction $\Psi : \mathcal{R}^n \to 2^{\mathcal{R}^n}$ is called monotone if

$$(x-y)^\top (\xi - \zeta) \geq 0 \quad \forall \xi \in \Psi(x), \forall \zeta \in \Psi(y).$$

It is well known that ∂f is a monotone multifunction (see [18]) for any convex function f, that is

$$(\xi - \zeta)^\top (x-y) \geq 0 \quad \forall \xi \in \partial f(x), \forall \zeta \in \partial f(y),$$

where ∂f is a subdifferential of f.

To present our analysis in a compact way, we thus define some matrices in the following analysis.

$$\mathcal{T} = \begin{pmatrix} G_1 & 0 & 0 \\ 0 & G_2 + \frac{1}{\gamma+\alpha}\rho B^\top B & \frac{1-\gamma-\alpha}{\gamma+\alpha} B^\top \\ 0 & \frac{1-\gamma-\alpha}{\gamma+\alpha} B & \frac{1}{\rho(\gamma+\alpha)} I_l \end{pmatrix}, \quad \mathcal{T}_1 = \begin{pmatrix} G_2 + \frac{1}{\gamma+\alpha}\rho B^\top B & \frac{1-\gamma-\alpha}{\gamma+\alpha} B^\top \\ \frac{1-\gamma-\alpha}{\gamma+\alpha} B & \frac{1}{\rho(\gamma+\alpha)} I_l \end{pmatrix} \tag{2.3}$$

and

$$\mathcal{M} = \begin{pmatrix} I_n & 0 & 0 \\ 0 & I_m & 0 \\ 0 & -\rho B & (\gamma+\alpha) I_l \end{pmatrix}, \quad \mathcal{Q} = \begin{pmatrix} G_1 & 0 & 0 \\ 0 & G_2 + \rho B^\top B & (1-\gamma-\alpha) B^\top \\ 0 & -B & \frac{1}{\rho} I_l \end{pmatrix}. \tag{2.4}$$

The matrices $\mathcal{M}, \mathcal{Q}, \mathcal{T}, \mathcal{T}_1$ just defined have some nice properties, we elaborate in the following lemma.

Lemma 2.1 (see [9, Lem. 2.1]) The matrices $\mathcal{M}, \mathcal{Q}, \mathcal{T}$ the above defined, respectively, in (2.3), (2.4). If $\alpha \in (0, 2), \gamma + \alpha < 2$ and $G_1 \succ \mathbf{0}$, $G_2 \succeq \mathbf{0}$ or $G_1 \succeq \mathbf{0}$, $G_2 \succ \mathbf{0}$, then we have

(I). $\mathcal{Q} = \mathcal{T}\mathcal{M}$.

(II). If $G_1 \succ \mathbf{0}$, $G_2 \succeq \mathbf{0}$, then $\mathcal{T} \succ 0$; If $G_1 \succeq \mathbf{0}$, $G_2 \succ \mathbf{0}$, then $\mathcal{T}_1 \succ \mathbf{0}$.

(III). $\mathcal{N} = \mathcal{Q}^\top + \mathcal{Q} - \mathcal{M}^\top \mathcal{T} \mathcal{M} \succeq \mathbf{0}$.

Now, we define an auxiliary sequences for the further analysis. More sepecially, for the sequence $\{w^k\}$ generated by the GPRSM (1.3), let

$$\bar{w}^k = \begin{pmatrix} \bar{x}^k \\ \bar{y}^k \\ \bar{\lambda}^k \end{pmatrix} = \begin{pmatrix} x^{k+1} \\ y^{k+1} \\ \lambda^k - \rho(Ax^{k+1} + By^k - b) \end{pmatrix}. \tag{2.5}$$

By using (1.3) and (2.6), we can obtain

$$\lambda^{k+\frac{1}{2}} = \lambda^k - \gamma(\lambda^k - \bar{\lambda}^k), \tag{2.6}$$

and

$$\lambda^{k+1} = \lambda^k - ((\gamma+\alpha)(\lambda^k - \bar{\lambda}^k) - \rho B(y^k - \bar{y}^k)), \tag{2.7}$$

combining (2.6) with (2.8), we have

$$w^{k+1} = w^k - \mathcal{M}(w^k - \bar{w}^k), \tag{2.8}$$

where $\mathcal{M}$ is defined in (2.4).

3 Worst-case iteration complexity of the GPRSM (1.3)

In this section, we establish a worst-case iteration complexity for the GRPSM (1.3) in both the ergodic and a nonergodic senses.

By using the first-order optimality condition and the proof of Lemma 3.1 in [9], it is easy to see that the following Lemma holds.

Lemma 3.1 Let the sequence $\{w^k\}$ be generated by the GPRSM (1.3) and the associated sequence $\{\bar{w}^k\}$ be defined in (2.3). Then we have

$$\Phi(u) - \Phi(\bar{u}^k) + (w - \bar{w}^k)^\top \mathcal{F}(\bar{w}^k) \geq (w - \bar{w}^k)^\top \mathcal{Q}(w^k - \bar{w}^k), \quad \forall w \in \Omega, \tag{3.1}$$

where $\mathcal{Q}$ is defined in (2.4), $\mathcal{F}$ is defined in (2.2).

3.1 A worst-case $\mathcal{O}(1/t)$ iteration complexity for GPRSM (1.3) in a ergodic sense

We will show the worst-case $\mathcal{O}(1/t)$ convergence rate for the GPRSM (1.3) in a ergodic sense in the subsection. In order to establish this iteration complexity, we first give the following lemma.

Lemma 3.2 Let the sequence $\{w^k\}$ be generated by the GPRSM (1.3). Then, we have

$$\Phi(u) - \Phi(\bar{u}^k) + (w - \bar{w}^k)^\top \mathcal{F}(\bar{w}^k) \geq \frac{1}{2}(\|w - w^{k+1}\|_{\mathcal{T}}^2 - \|w - w^k\|_{\mathcal{T}}^2) + \frac{1}{2}\delta\|w^k - \bar{w}^k\|^2, \tag{3.2}$$

where $\delta = \min\{\lambda_{\max}(G_1), \lambda_{\max}(G_2), \frac{2-\gamma-\alpha}{\rho}\} > 0$.

Proof By using (2.3), if $\alpha \in (0,2), \gamma + \alpha < 2$ and $G_1 \succ \mathbf{0}$, $G_2 \succeq \mathbf{0}$ or $G_1 \succeq \mathbf{0}$, $G_2 \succ \mathbf{0}$, then we have $\mathcal{T} \succeq \mathbf{0}$. Then, it holds that

$$(a-b)^\top \mathcal{T}(c-d) = \frac{1}{2}(\|a-d\|_{\mathcal{T}}^2 - \|a-c\|_{\mathcal{T}}^2) + \frac{1}{2}(\|c-b\|_{\mathcal{T}}^2 - \|d-b\|_{\mathcal{T}}^2), \quad \forall\, a,b,c,d \in \Omega.$$

Then, setting $a = w, b = \bar{w}^k, c = w^k, d = w^{k+1}$ in the above identity and using (2.5) and (2.9), we have

$$
\begin{aligned}
&(w - \bar{w}^k)^\top \mathcal{Q}(w^k - \bar{w}^k) \\
=&(w - \bar{w}^k)^\top \mathcal{T}(w^k - w^{k+1}) \\
=&\frac{1}{2}(\|w - w^{k+1}\|_{\mathcal{T}}^2 - \|w - w^k\|_{\mathcal{T}}^2) \\
&\frac{1}{2}(\|w^k - \bar{w}^k\|_{\mathcal{T}}^2 - \|w^{k+1} - \bar{w}^k\|_{\mathcal{T}}^2).
\end{aligned}
\tag{3.3}
$$

It is a fact that

$$
\begin{aligned}
&\|w^k - \bar{w}^k\|_{\mathcal{T}}^2 - \|w^{k+1} - \bar{w}^k\|_{\mathcal{T}}^2 \\
&= \|w^k - \bar{w}^k\|_{\mathcal{T}}^2 - \|(w^k - \bar{w}^k) - (w^k - w^{k+1})\|_{\mathcal{T}}^2 \\
&= \|w^k - \bar{w}^k\|_{\mathcal{T}}^2 - \|(w^k - \bar{w}^k) - \mathcal{M}(w^k - \bar{w}^k)\|_{\mathcal{T}}^2 \\
&= 2(w^k - \bar{w}^k)^\top \mathcal{T}\mathcal{M}(w^k - \bar{W}^k) - (w^k - \bar{w}^k)^\top \mathcal{M}^\top \mathcal{T}\mathcal{M}(w^k - \bar{w}^k) \\
&= (w^k - \bar{w}^k)^\top (\mathcal{Q}^\top + \mathcal{Q} - \mathcal{M}^\top \mathcal{T}\mathcal{M})(w^k - \bar{w}^k) \\
&= \|w^k - \bar{w}^k\|_{\mathcal{N}}^2.
\end{aligned}
\tag{3.4}
$$

where

$$
\mathcal{N} = \mathcal{Q}^\top + \mathcal{Q} - \mathcal{M}^\top \mathcal{T}\mathcal{M} = \begin{pmatrix} G_1 & 0 & 0 \\ 0 & G_2 & 0 \\ 0 & 0 & \frac{2-\gamma-\alpha}{\rho} I_l \end{pmatrix}. \tag{3.5}
$$

substituting (3.5) into (3.4) and using (3.1), we can obtain

$$
\Phi(u) - \Phi(\bar{u}^k) + (w - \bar{w}^k)^\top \mathcal{F}(\bar{w}^k) \geq \frac{1}{2}(\|w - w^{k+1}\|_{\mathcal{T}}^2 - \|w - w^k\|_{\mathcal{T}}^2) + \frac{1}{2}\|w^k - \bar{w}^k\|_{\mathcal{N}}^2, \tag{3.6}
$$

Let $\delta = \min\{\lambda_{\max}(G_1), \lambda_{\max}(G_2), \frac{2-\gamma-\alpha}{\rho}\} > 0$, then the assertion of the lemma is proved.

The following theorem shows the worst case $\mathcal{O}(1/t)$ iteration complexity for the GPRSM in a ergodic sense.

Theorem 3.2 Let $\{w^k\}$ be the sequence generated by the GPRSM (1.3) and for any integer $t > 0$, let $w_t = \frac{1}{t+1}\sum_{k=0}^t \bar{w}^k$, then we have $w_t \in \Omega$ and

$$
\Phi(u_t) - \Phi(u) + (w_t - w)^\top \mathcal{F}(w) \leq \frac{1}{2(t+1)}\|w - w^0\|_{\mathcal{T}}^2, \quad \forall w \in \Omega. \tag{3.7}
$$

Proof By using the notation in (2.6) and $w^k \in \Omega$, then we can obtain that $\bar{w}^k \in \Omega$ for all $k \geq 0$. Due to the convexity of $\mathcal{X}, \mathcal{Y}$ and the definition of w_t, it implies that $w_t \in \Omega$. Because of the monotonicity of the $\mathcal{F}(w)$, we have

$$
(w - \bar{w}^k)^\top \mathcal{F}(w) \geq (w - \bar{w}^k)^\top \mathcal{F}(\bar{w}^k), \tag{3.8}
$$

Using the lemma 3.2 and the above inequality, we can obtain

$$\Phi(u)-\Phi(\bar{u}^k)+(w-\bar{w}^k)^\top\mathcal{F}(w)+\frac{1}{2}\|w^k-w\|_{\mathcal{T}}^2\geq\frac{1}{2}\|w^{k+1}-w\|_{\mathcal{T}}^2+\frac{1}{2}\delta\|w^k-\bar{w}^k\|^2,\ \ \forall\, w\in\Omega.$$

It is easy to see that

$$\Phi(u)-\Phi(\bar{u}^k)+(w-\bar{w}^k)^\top\mathcal{F}(w)+\frac{1}{2}(\|w^k-w\|_{\mathcal{T}}^2-\|w^{k+1}-w\|_{\mathcal{T}}^2)\geq 0,\ \ \forall\, w\in\Omega.$$

Summing the above inequality over k=0,1,...,t, we have

$$(t+1)\Phi(u)-\sum_{k=0}^{t}\Phi(\bar{u}^k)+((t+1)w-\sum_{k=0}^{t}\bar{w}^k)^\top\mathcal{F}(w)\geq-\frac{1}{2}\|w-w^0\|_{\mathcal{T}}^2,\ \ \forall w\in\Omega. \tag{3.9}$$

Due to the definition of w_t and the convexity of $\Phi(u)$, the above inequality can be rewritten as

$$\Phi(u_t)-\Phi(u)+(w_t-w)^\top\mathcal{F}(w)\leq\frac{1}{2(t+1)}\|w-w^0\|_{\mathcal{T}}^2. \tag{3.10}$$

The proof of this theorem is completed .

It is assumed that an substantial compact set $\mathcal{S}\subset\Omega$ and let

$$\mathcal{S}=\sup\{\|w-w^0\|_{\mathcal{T}}|w\in\mathcal{S}\}, \tag{3.11}$$

where $w^0=(x^0,y^0,\lambda^0)$ is the initial point. After t iterations of the GPRSM (1.3), there exists a $w_t\in\Omega$ such that

$$\sup_{w\in\mathcal{S}}\{\Phi(u_t)-\Phi(u)+(w_t-w)^\top\mathcal{F}(w)\}\leq\frac{\mathcal{S}^2}{2t}. \tag{3.12}$$

Therefore, a worst-case $\mathcal{O}(1/t)$ iteration complexity for the GPRSM in a ergodic sense is established.

3.2 A worst-case $\mathcal{O}(1/t)$ iteration complexity for GPRSM (1.3) in a nonergodic sense

Next, we establish a worst-case $\mathcal{O}(1/t)$ convergence rate for GPRSM (1.3) in a nonergodic sense. We first show that the term $\|w^k-w^{k+1}\|_{\mathcal{T}}^2$ can be used to measure the accuracy of an iterate.

Lemma 3.3 For a given sequence $\{w^k\}$, let $\{w^{k+1}\}$ be generated by the GPRSM (1.3). Then, if $\|w^k-w^{k+1}\|_{\mathcal{T}}^2=0$, we have

$$\Phi(u)-\Phi(\bar{u}^k)+(w-\bar{w}^k)^\top\mathcal{F}(\bar{w}^k)\geq 0,\qquad\forall w\in\Omega, \tag{3.13}$$

Proof Using (3.1), we can obtain

$$\Phi(u)-\Phi(\bar{u}^k)+(w-\bar{w}^k)^\top\mathcal{F}(\bar{w}^k)\geq(w-\bar{w}^k)^\top\mathcal{Q}(w^k-\bar{w}^k),\qquad\forall w\in\Omega, \tag{3.14}$$

Due to (2.5) and (2.9), we have

$$\Phi(u) - \Phi(\bar{u}^k) + (w - \bar{w}^k)^\top \mathcal{F}(\bar{w}^k) \geq (w - \bar{w}^k)^\top \mathcal{T}(w^k - w^{k+1}), \qquad \forall w \in \Omega, \tag{3.15}$$

By using the positive semidefiniteness of $\mathcal{T}$ in (2.3), the assertion of (3.14) is proved when $\|w^{k+1} - w^k\|_{\mathcal{T}}^2 = 0$. Then $\bar{w}^k$ defined in (2.6) is a solution to (2.1) if $\|w^{k+1} - w^k\|_{\mathcal{T}}^2 = 0$.

To prove a worst-case $\mathcal{O}(1/t)$ convergence rate for the iterative scheme (1.8) in a nonergodic sense. First, we prove the following two lemmas.

Lemma 3.4 Let the sequence $\{w^k\}$ be generated by the GPRSM (1.3) with $\alpha \in (0, 2)$ and $\gamma \in (2 - \alpha)$, and the associated $\{\bar{w}^k\}$ is defined in (2.3), the matrix $\mathcal{Q}$ is defined in (2.4). Then, we have

$$(\bar{w}^k - \bar{w}^{k+1})^\top \mathcal{Q} \left[(w^k - w^{k+1}) - (\bar{w}^k - \bar{w}^{k+1})\right] \geq 0. \tag{3.16}$$

Proof Setting $u = \bar{u}^{k+1}, w = \bar{w}^{k+1}$ in (3.1),then we have

$$\Phi(\bar{u}^{k+1}) - \Phi(\bar{u}^k) + (\bar{w}^{k+1} - \bar{w}^k)^\top \mathcal{F}(\bar{w}^k) \geq (\bar{w}^{k+1} - \bar{w}^k)^\top \mathcal{Q}(w^k - \bar{w}^k). \tag{3.17}$$

Clearly, it is true when $k := k + 1$. By using $u = \bar{u}^k, w = \bar{w}^k$, we get

$$\Phi(\bar{u}^k) - \Phi(\bar{u}^{k+1}) + (\bar{w}^k - \bar{w}^{k+1})^\top \mathcal{F}(\bar{w}^{k+1}) \geq (\bar{w}^k - \bar{w}^{k+1})^\top \mathcal{Q}(w^{k+1} - \bar{w}^{k+1}). \tag{3.18}$$

Adding (3.18) and (3.19) and because of the monotonicity of $\mathcal{F}$, the assertion (3.17) thus follows directly.

Lemma 3.5 Let the sequence $\{w^k\}$ be generated by the GPRSM (1.3) and the associated $\{\bar{w}^k\}$ be defined in (2.6), then we have

$$\begin{aligned} &(w^k - \bar{w}^k)^\top \mathcal{M}^\top \mathcal{T} \mathcal{M} \left[(w^k - \bar{w}^k) - (w^{k+1} - \bar{w}^{k+1})\right] \\ &\geq \frac{1}{2}\|(w^k - \bar{w}^k) - (w^{k+1} - \bar{w}^{k+1})\|_{(\mathcal{Q}^\top + \mathcal{Q})}^2, \end{aligned} \tag{3.19}$$

where the matrices $\mathcal{T}$ and $\mathcal{M}, \mathcal{Q}$ is definied in (2.3) and (2.4), respectively.

Proof Adding the equation

$$\begin{aligned} &\left[(w^k - w^{k+1}) - (\bar{w}^k - \bar{w}^{k+1})\right]^\top \mathcal{Q} \left[(w^k - w^{k+1}) - (\bar{w}^k - \bar{w}^{k+1})\right] \\ &= \frac{1}{2}\|(w^k - \bar{w}^k) - (w^{k+1} - \bar{w}^{k+1})\|_{(\mathcal{Q}^\top + \mathcal{Q})}^2 \end{aligned}$$

to both sides of (3.17), we have

$$\begin{aligned} &(w^k - w^{k+1})^\top \mathcal{Q} \left[(w^k - w^{k+1}) - (\bar{w}^k - \bar{w}^{k+1})\right] \\ &\geq \frac{1}{2}\|(w^k - \bar{w}^k) - (w^{k+1} - \bar{w}^{k+1})\|_{(\mathcal{Q}^\top + \mathcal{Q})}^2 \end{aligned} \tag{3.20}$$

Furthermore, from (2.5) and (2.9), the proof of lemma 3.5 is completed.

From the above lemmas, it is easy to see that the sequence $\{\|w^{k+1} - w^k\|_{\mathcal{T}}\}$ is monotonically non-increasing. And we have the following theorem.

Theorem 3.3 Let the sequence $\{w^k\}$ be generated by the GPRSM (1.3) and the matrix $\mathcal{T}$ be defined in (2.3). Then, we have

$$\|w^{k+1} - w^{k+2}\|_{\mathcal{T}}^2 \leq \|w^k - w^{k+1}\|_{\mathcal{T}}^2. \tag{3.21}$$

Proof Using $c = \mathcal{M}(w^k - \bar{w}^k)$ and $d = \mathcal{M}(w^{k+1} - \bar{w}^{k+1})$, we get

$$\|c\|_{\mathcal{T}}^2 - \|d\|_{\mathcal{T}}^2 = 2c^\top \mathcal{T}(c - d) - \|c - d\|_{\mathcal{T}}^2,$$

Moreover, we have

$$\begin{aligned}
&\|\mathcal{M}(w^k - \bar{w}^k)\|_{\mathcal{T}}^2 - \left\|\mathcal{M}(w^{k+1} - \bar{w}^{k+1})\right\|_{\mathcal{T}}^2 \\
&=2(w^k - \bar{w}^k)\mathcal{M}^\top \mathcal{T}\mathcal{M}\left[(w^k - \bar{w}^k) - (w^{k+1} - \bar{w}^{k+1})\right] \\
&\quad - \left\|\mathcal{M}[(w^k - \bar{w}^k) - (w^{k+1} - \bar{w}^{k+1})]\right\|_{\mathcal{T}}^2.
\end{aligned} \tag{3.22}$$

By using (3.20), we can obtain

$$\begin{aligned}
&\|\mathcal{M}(w^k - \bar{w}^k)\|_{\mathcal{T}}^2 - \left\|\mathcal{M}(w^{k+1} - \bar{w}^{k+1})\right\|_{\mathcal{T}}^2 \\
&\geq\|(w^k - \bar{w}^k) - (w^{k+1} - \bar{w}^{k+1})\|_{\mathcal{Q}^\top + \mathcal{Q})}^2 \\
&\quad - \left\|\mathcal{M}[(w^k - \bar{w}^k) - (w^{k+1} - \bar{w}^{k+1})]\right\|_{\mathcal{T}}^2 \\
&=\|(w^k - \bar{w}^k) - (w^{k+1} - \bar{w}^{k+1})\|_{(\mathcal{Q}^\top + \mathcal{Q}) - \mathcal{M}^\top \mathcal{T}\mathcal{M}}^2 \\
&=\|(w^k - \bar{w}^k) - (w^{k+1} - \bar{w}^{k+1})\|_{\mathcal{N}}^2.
\end{aligned} \tag{3.23}$$

where $\mathcal{N}$ is defined in (3.6). Due to the positive definiteness of G_i (i=1, 2), it is not difficult to see that $\mathcal{N}$ is positive definite when $\alpha \in (0, 2)$ and $\gamma \in (0, 2 - \alpha)$. We thus have

$$\left\|\mathcal{M}(w^k - \bar{w}^k)\right\|_{\mathcal{T}}^2 - \left\|\mathcal{M}(w^{k+1} - \bar{w}^{k+1})\right\|_{\mathcal{T}}^2 \geq 0.$$

Using (2.8), the assertion (3.22) follows immediately.

For further analysis, we show the following two lemmas to establish the worst-case $\mathcal{O}(1/t)$ convergence rate for GPRSM (1.3) in a nonergodic sense.

Lemma 3.6 Let $\{y^k\}$ be the sequence generated by the GPRSM (1.3) with $\alpha \in (0, 2)$. Then, we have

$$(y^k - y^{k+1})^\top B^\top (\lambda^k - \lambda^{k+1}) \geq \frac{1}{2}\|y^k - y^{k+1}\|_{G_2}^2 - \frac{1}{2}\|y^{k-1} - y^k\|_{G_2}^2. \tag{3.24}$$

Proof By using the optimality condition of the y-subpeoblem in (1.3), there exists $\zeta_1 \in \partial g(y^{k+1})$ such that

$$(y - y^{k+1})^\top \left[\zeta_1 - A_3^\top \lambda^{k+1} + G_2(y^{k+1} - y^k)\right] \geq 0, \quad \forall\, y \in \mathcal{Y}, \tag{3.25}$$

where $\partial g(y)$ is a subdifferential of $g(y)$. Setting $y = y^k$, then we have

$$(y^k - y^{k+1})^\top \left[\zeta_1 - B^\top \lambda^{k+1} + G_2(y^{k+1} - y^k)\right] \geq 0. \tag{3.26}$$

And using $y = y^{k+1}$, there exists $\zeta_2 \in \partial g(y^k)$ such that

$$(y^{k+1} - y^k)^\top \left[\zeta_2 - B^\top \lambda^k + G_2(y^k - y^{k-1})\right] \geq 0. \tag{3.27}$$

Adding (3.27) to (3.28) and using the monotonicity of the $\partial g(y)$, we have

$$\begin{aligned}(y^k - y^{k+1})^\top B^\top (\lambda^k - \lambda^{k+1}) &\geq (y^{k+1} - y^k)^\top G_2(y^{k+1} - y^k + y^{k-1} - y^k)\\ &= \|y^{k+1} - y^k\|_{G_2}^2 + (y^{k+1} - y^k)^\top G_2(y^{k-1} - y^k)\end{aligned}$$

Using the inequality

$$(y^{k+1} - y^k)^\top G_2(y^{k-1} - y^k) \geq -\frac{1}{2}\|y^k - y^{k+1}\|_{G_2}^2 - \frac{1}{2}\|y^{k-1} - y^k\|_{G_2}^2,$$

the assertion (3.25) is proved.

Lemma 3.7 The sequence $\{w^k\}$ generated by the GPRSM (1.3) with $\alpha \in (0,2), \gamma \in (0, 2-\alpha)$ and the associated $\{\bar{w}^k\}$ be defined in (2.6), then there exists $0 < \kappa \leq 1$ such that

$$\begin{aligned}&\|w^k - \bar{w}^k\|_{\mathcal{N}}^2\\ &\geq \kappa \left(\|x^k - \bar{x}^k\|_{G_1}^2 + \|y^k - \bar{y}^k\|_{G_2}^2 + \frac{\gamma + \alpha}{\rho}\|\lambda^k - \bar{\lambda}^k\|^2\right).\end{aligned} \tag{3.28}$$

Proof By the definition of $\mathcal{N}$ in (3.6), we can obtain

$$\begin{aligned}&\|w^k - \bar{w}^k\|_{\mathcal{N}}^2\\ =&\|x^k - \bar{x}^k\|_{G_1}^2 + \|y^k - \bar{y}^k\|_{G_2}^2 + \frac{2-\gamma-\alpha}{\rho}\|\lambda^k - \bar{\lambda}^k\|^2\\ \geq&\min\{\frac{2-\gamma-\alpha}{\gamma+\alpha}, 1\}(\|x^k - \bar{x}^k\|_{G_1}^2 + \|y^k - \bar{y}^k\|_{G_2}^2 + \frac{\gamma+\alpha}{\rho}\|\lambda^k - \bar{\lambda}^k\|^2).\end{aligned} \tag{3.29}$$

Let $\kappa = \min\left\{\frac{2-\gamma-\alpha}{\gamma+\alpha}, 1\right\} \in (0,1]$, then the proof of the assertion (3.29) is completed.

Finally, we will show the worst-case $\mathcal{O}(1/t)$ convergence rate for GPRSM (1.3) in a nonergodic sense in the following theorem.

Theorem 3.4 Let the sequence $\{w^k\}$ be generated by the iterative scheme GPRSM (1.3) with

$\alpha \in (0,2), \gamma \in (0, 2-\alpha)$. Then, we have

$$\|w^t - w^{t+1}\|_{\mathcal{T}}^2 \le \frac{1}{t}\left(\frac{1}{\kappa}\|w^0 - w^*\|_{\mathcal{T}}^2 + \|y^0 - y^1\|_{\mathcal{T}}^2\right). \tag{3.30}$$

Proof Setting $w = w^*$ in (3.7), we can obtain

$$\begin{aligned}(w^* - \bar{w}^k)^\top \mathcal{F}(w^*) \ge& \frac{1}{2}(\|w^* - w^{k+1}\|_{\mathcal{T}}^2 - \|w^* - w^k\|_{\mathcal{T}}^2) \\ &+ \frac{1}{2}\|w^k - \bar{w}^k\|_{\mathcal{N}}^2.\end{aligned}$$

where $\mathcal{N} = \mathcal{Q}^\top + \mathcal{Q} - \mathcal{M}^\top \mathcal{T} \mathcal{M}$ is defined in (3.6). Using (2.1), then we get

$$(w^* - \bar{w}^k)^\top \mathcal{F}(w^*) \le 0.$$

Therefore, we have

$$\|w^k - \bar{w}^k\|_{\mathcal{N}}^2 \le \|w^* - w^k\|_{\mathcal{T}}^2 - \|w^* - w^{k+1}\|_{\mathcal{T}}^2. \tag{3.31}$$

Because of the positive definiteness of G_i (i=1, 2), it holds that $\mathcal{N}$ is positive definite when $\alpha \in (0,2)$ and $\gamma \in (0, 2-\alpha)$. We have thus

$$\sum_{k=0}^{\infty} \|w^k - \bar{w}^k\|_{\mathcal{N}}^2 \le \|w^0 - w^*\|_{\mathcal{T}}^2. \tag{3.32}$$

Moreover, by using the definition of $\mathcal{T}$ in (2.3), we can obtain

$$\begin{aligned}&\|w^k - w^{k+1}\|_{\mathcal{T}}^2 \\ =&\|x^k - \bar{x}^k\|_{G_1}^2 + \|y^k - \bar{y}^k\|_{(G_2 + \frac{\rho}{\gamma+\alpha}B^\top B)}^2 + \frac{1}{\rho(\gamma+\alpha)}\|\lambda^k - \lambda^{k+1}\|^2 \\ &+ 2\frac{1-\gamma-\alpha}{\gamma+\alpha}(y^k - y^{k+1})^\top B^\top (\lambda^k - \lambda^{k+1}) \\ =&\|x^k - \bar{x}^k\|_{G_1}^2 + \|y^k - \bar{y}^k\|_{G_2}^2 + \frac{1}{\rho(\gamma+\alpha)}\|\rho B(y^k - \bar{y}^k) + (\lambda^k - \lambda^{k+1})\|^2 \\ &- 2(y^k - y^{k+1})^\top B^\top (\lambda^k - \lambda^{k+1}).\end{aligned}$$

And using (2.8), we have

$$\begin{aligned}\|w^k - w^{k+1}\|_{\mathcal{T}}^2 =& \|x^k - \bar{x}^k\|_{G_1}^2 + \|y^k - \bar{y}^k\|_{G_2}^2 + \frac{\gamma+\alpha}{\rho}\|\lambda^k - \bar{\lambda}^k\|^2 \\ &- 2(y^k - y^{k+1})^\top B^\top (\lambda^k - \lambda^{k+1}).\end{aligned} \tag{3.33}$$

By (3.25), (3.29), (3.33) and (3.34), we have

$$\begin{aligned}\sum_{k=1}^{t} \|w^k - w^{k+1}\|_{\mathcal{T}}^2 \le& \frac{1}{\kappa}\sum_{k=1}^{t}\|w^k - \bar{w}^k\|_{\mathcal{N}}^2 \\ &+ \sum_{k=1}^{t}\left(\|y^{k-1} - y^k\|_{G_2}^2 - \|y^k - y^{k+1}\|_{G_2}^2\right) \\ \le& \frac{1}{\kappa}\|w^0 - w^*\|_{\mathcal{T}}^2 + \|y^0 - y^1\|_{G_2}^2.\end{aligned} \tag{3.34}$$

Due to the theorem 3.3, the sequence $\{\|w^k - w^{k+1}\|_{\mathcal{T}}^2\}$ is non-increasing. Thus, we obtain

$$\begin{aligned} t\|w^t - w^{t+1}\|_{\mathcal{T}}^2 &\leq \sum_{k=1}^{t} \|w^k - w^{k+1}\|_{\mathcal{T}}^2 \\ &\leq \frac{1}{\kappa}\|w^0 - w^*\|_{\mathcal{T}}^2 + \|y^0 - y^1\|_{\mathcal{T}}^2. \end{aligned} \tag{3.35}$$

Using the above inequality, then the assertion (3.31) follows immediately.

A worst-case $\mathcal{O}(1/t)$ convergence rate for GPRSM (1.3) in a nonergodic sense is thus established.

4 Numerical experiments

In fact, in [10], the efficiency of the generalized version of the Peaceman-Rachford splitting method has been very well illustrated, but there is few people demonstrating its convergence rate by some numerical. In this section, we illustrate the convergence rate of the GPRSM (1.3) by some numerical experiments. In our experiments, we show a popular model that satisfies our assumption is the LASSO model and Calibrating the correlation matrices.

All experiments were implemented in MATLAB R2010b on a hp-notebook with an Intel Core i5-3340M CPU at 2.70 GHz and 8 GB memory.

4.1 The Lasso problem

In this subsection, we apply the GPRSM (1.3) to solve the lasso problem. To achieve the goal, we first give the following LASSO model

$$\min \ \tfrac{1}{2}\|Ex - q\|_2^2 + \tau\|x\|_1$$

We introducing an auxiliary variable $y \in \mathcal{R}^n$, then we can change the above model to the following equivalent form

$$\begin{aligned} &\min \ \tfrac{1}{2}\|Ex - q\|_2^2 + \tau\|y\|_1 \\ &\text{s.t.} \quad x - y = 0. \\ &\qquad\quad x \in \mathcal{R}^n, y \in \mathcal{R}^n. \end{aligned} \tag{4.1}$$

Clearly, the above model follows the framework of (1.1), and we can use the GPRSM scheme (1.3) to solve (4.1). And we generate a 1500×5000 random sparse matrix E together with corresponding random sparse vector $q \in \mathcal{R}^{1500}$, where we set $\tau = \frac{1}{10}\|E^\top q\|_\infty$ and $f(x) = \frac{1}{2}\|Ex - q\|_2^2$, $g(y) = \tau\|y\|_1$, $A = I_n$, $B = -I_n$.

Here, we apply the GPRSM (1.3) to solve (4.1) and we derive the subproblems.

$$\begin{cases} x^{k+1} = \text{argmin}\{\frac{1}{2}\|Ex-q\|_2^2 + \frac{\rho}{2}\|x+y^k-\frac{1}{\rho}\lambda^k\|^2 + \frac{1}{2}\|x-x^k\|_{G_1}^2 | x\in\mathcal{R}^n\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \gamma\rho(x^{k+1}-y^k), \\ y^{k+1} = \text{argmin}\{\tau\|y\|_1 + \frac{\rho}{2}\|y-(\alpha x^{k+1}+(1-\alpha)y^k - \frac{1}{\rho}\lambda^{k+\frac{1}{2}})\|^2 + \frac{1}{2}\|y-y^k\|_{G_2}^2 | y\in\mathcal{R}^n\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \rho(\alpha x^{k+1}+(1-\alpha)y^k - y^{k+1}). \end{cases} \tag{4.2}$$

We then elaborate on the subproblem in (4.2) by the algorithm (1.3). Here, we set $G_1 = \frac{1}{100}\rho I_n, G_2 = 0$ and $\gamma = \frac{1}{2} * (2-\alpha)$.

For the x-subproblem (4.2), by simple calculation, we have

$$x^{k+1} = (G_1 + \rho I_n + A^\top A)^{-1}(A^\top b + \rho(y^k + \frac{1}{\rho}\lambda^k) + G_1 x^k).$$

For the y-subproblem (4.2), let shrinkage$(\mathbf{a},\kappa)$:=sign$(\mathbf{a})\cdot$ max$(0, |\mathbf{a}|-\kappa)$ be the soft shrinkage operater, where sign$(\cdot)$ is the sign function. Then the closed-form solution of the y-subproblem in (4.2) is written as

$$y^{k+1} = \text{shrinkage}\left(\alpha x^{k+1} + (1-\alpha)y^k - \frac{1}{\rho}\lambda^{k+\frac{1}{2}}, \frac{\tau}{\rho}\right).$$

To implement Algorithm (1.3), we use the termination criterion in [15]. First, we give the following some preparations.

The primal feasibility of the necessary and sufficient optimality conditions for the problem (4.1) is showed in the following identity

$$x^* - y^* = 0, \tag{4.3}$$

and the dual feasibility is written as

$$0 \in \partial f(x^*) + \lambda^*, \tag{4.4}$$

$$0 \in \partial g(y^*) - \lambda^*, \tag{4.5}$$

where $f(x^*) = \frac{1}{2}\|Ex^*-q\|_2^2$, $g(y^*) = \tau\|y^*\|_1$. And we refer to $s^{k+1} = -\rho(y^{k+1}-y^k)$ as a dual residual for the condition (4.4) at $(k+1)$-th iteration, and to $r^{k+1} = x^{k+1} - y^{k+1}$ as the primal residual at $(k+1)$-th iteration. In [15], the suggests that a reasonable stopping criteria is that the primal and dual residuals must be small

$$\|r^k\| \le \varepsilon^{pri} \quad \text{and} \quad \|s^k\| \le \varepsilon^{dual}, \tag{4.6}$$

where $\varepsilon^{pri} > 0$ and $\varepsilon^{dual} > 0$ are feasibility tolerances

$$\varepsilon^{pri} = \sqrt{n}\varepsilon^{abs} + \varepsilon^{rel}\max\{\|x^k\|, \|-y^k\|\},$$
$$\varepsilon^{dual} = \sqrt{n}\varepsilon^{abs} + \varepsilon^{rel}\|y^k\|.$$

where $\varepsilon^{abs} > 0$ is an absolute tolerance and $\varepsilon^{rel} > 0$ is a relative tolerance.

According the above analysis, the termination criterion is $\varepsilon^{rel} = 10^{-3}$ in our experiment. We apply the GPRSM (1.3) to solve the model (4.1) and show the convergence rate of the algorithm (1.3) by using different relaxation factors α or different penalty parameter ρ. We plot the the revolutions of objective function value with iterations for different values of α or ρ in Figure 4.1. On the one hand, from the Figure 4.1, we can see that when α is close to 2, the objective value

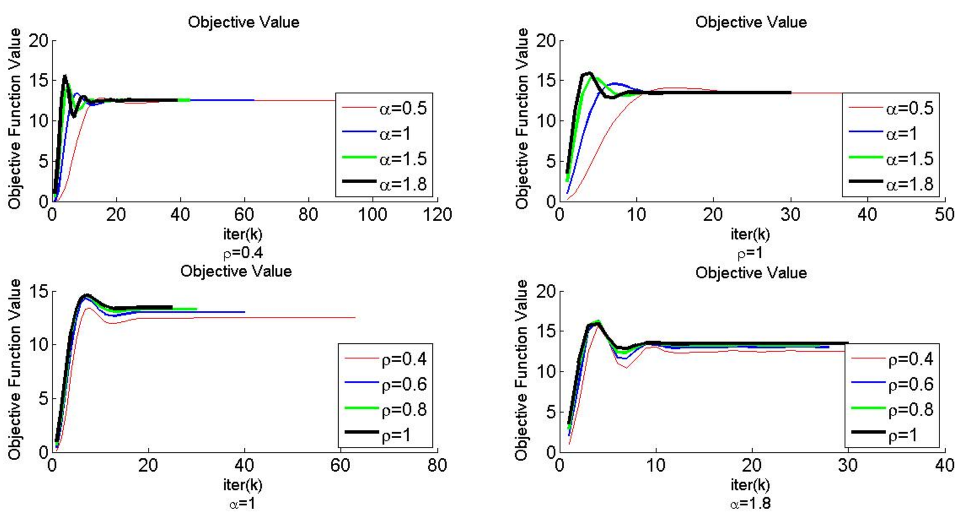

Figure 4.1: Comparison results of GPRSM on (1.2) for different α and ρ

decrease faster than the cases where α is close to 1. The penalty parameter ρ has little impact on the convergence rate when the parameter has a small change. On the other hand, we give

Figure 4.2: Evolutions of the relative error of the objective function value with respect to different values of α

some analysis of the Error of the objective function in Figure 4.2. And

$$\text{Objval}_{RelativeError} = \frac{\|\Psi(u^{k+1}) - \Psi(u^k)\|}{\|\Psi(u^k)\|},$$

where $\Psi(u) = \frac{1}{2}\|Ex - q\|_2^2 + \tau\|y\|_1$. And we also show some numerical results about the penalty parameter $\rho = 1$ in Table 4.1. Furthermore, we give the primal residual, dual residual, the CPU time and the iterations for the whole experiment process. The numerical results in Table 4.1 indicate that the convergence rate is faster when the relaxation α is close to 2.

α	No. It	$\|r\|_2$	$\|s\|_2$	CPU Sec
0.5	44	0.0080	0.0012	1.931261
1	25	0.0067	0.0022	1.212279
1.5	19	0.0060	0.0058	1.095425
1.8	30	0.0073	0.0001	1.465247

Table 4.1: Numerical comparison of the different α ($\rho = 1$).

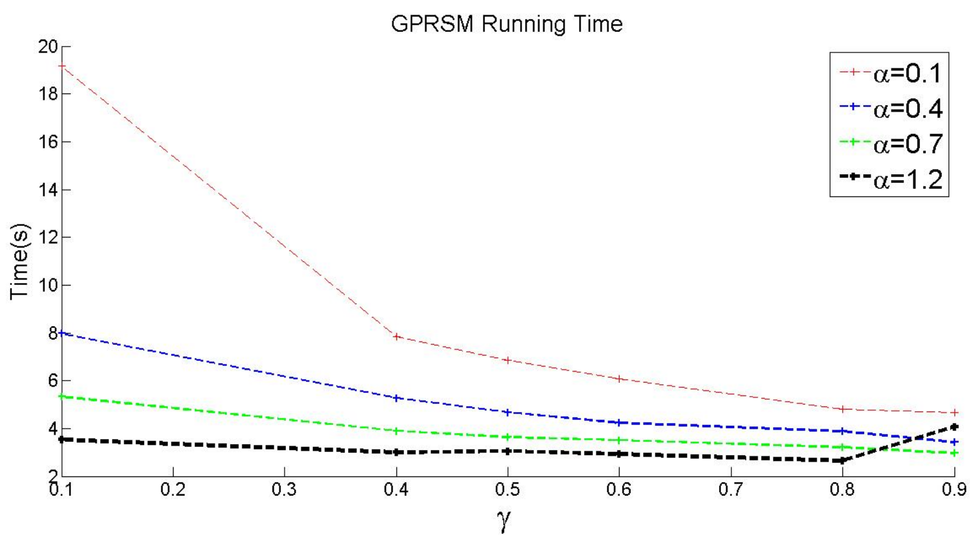

Figure 4.3: evolutions of computing time and the parameter γ in seconds with respect to different values of α

In fact, we also plot the evolutions of the parameter γ and computing time in seconds with respect to different values of α in the interval [0.1,0.9] with an equal distance of 0.1 in Fig.4.3. From the Fig.4.3, we can see that when α is close to 1.2, the iterative time is less than the cases where α is close to 0.1. However, if the α is close to 2, the convergence of the algorithm (1.2) will be poor, We thus set the $\alpha \in (0, 1.2]$. And in our experiment, the convergence rate will become faster, if γ is close to 0.9. This case where the values of γ close to 1 follows the numerical results in [7], but our experiment involves in the relaxation parameter α.

4.2 Calibrating the correlation matrices

In this subsection, we use the algorithm (1.3) to calibrate the correlation matrices.

We consider to solve the following matrix optimization problem

$$\min \{\frac{1}{2}\|X-C\|_F^2 | X \in S_+^n \cap \mathcal{S}_B\}, \tag{4.7}$$

where

$$S_+^n = \{H \in R^{n\times n} | H^\top = H, H \succeq \mathbf{0}\},$$

and

$$S_B = \{H \in R^{n\times n} | H^\top = H, H_L \leq H \leq H_U\}.$$

The above problem (4.2) could be converted to the following equivalent form

$$\begin{array}{ll} \min & \frac{1}{2}\|X-C\|^2 + \frac{1}{2}\|Y-C\|^2 \\ \text{s.t.} & X - Y = \mathbf{0}. \\ & X \in S_+^n, Y \in S_B. \end{array} \tag{4.8}$$

To solve the above matrix optimization problem, it is easy to see that the GPRSM (1.2) could be applicable and we obtain the following subproblem for model (4.8).

$$\begin{cases} \bar{X}^k = \text{argmin}\{\frac{1}{2}\|X-C\|_F^2 + \frac{\rho}{2}\|X - Y^k - \frac{1}{\rho}\lambda^k\|_F^2 + \frac{1}{2}\|X - X^k\|_{G_1}^2 | X \in \mathcal{S}_+^n\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \gamma\rho(\bar{X}^k - Y^k), \\ \bar{Y}^k = \text{argmin}\{\frac{1}{2}\|Y-C\|_F^2 + \frac{\rho}{2}\|Y - (\alpha\bar{X}^k + (1-\alpha)Y^k - \frac{1}{\rho}\lambda^{k+\frac{1}{2}})\|_F^2 + \frac{1}{2}\|Y - Y^k\|_{G_2}^2 | Y \in \mathcal{S}_B\}, \\ \bar{\lambda}^k = \lambda^{k+\frac{1}{2}} - \rho(\alpha\bar{X}^k + (1-\alpha)Y^k - \bar{Y}^k). \end{cases} \tag{4.9}$$

Here, we set $G_1 = \rho I_n$, $G_2 = \frac{1}{2}\rho I_n$, then the X-subproblem in (4.9) can be rewritten as

$$\bar{X}^k = P_{\mathcal{S}_+^n}\{\frac{1}{1+2\rho}(\rho(Y^k + X^k) + \lambda^k + C)\}, \tag{4.10}$$

where $P_{\mathcal{S}_+^n}(\mathcal{A}) = \mathcal{U}\Lambda^+\mathcal{U}^\top$ and $\Lambda^+ = \max(\Lambda, 0)$, $[\mathcal{U}, \Lambda] = \text{eig}(\mathcal{A})$. And, the solution of the Y-subproblem in (4.9) is given by

$$\bar{Y}^k = P_{\mathcal{S}_B}\{\frac{1}{1+\frac{3}{2}\rho}(\rho(\frac{1}{2}Y^k + (1-\alpha)Y^k + \alpha\bar{X}^k) - \lambda^{k+\frac{1}{2}} + C)\}, \tag{4.11}$$

where $S_B = \{H \in R^{n\times n} | H_L \leq H \leq H_U\}$ and $P_{\mathcal{S}_B}(\mathcal{A}) = \min(\max(H_L, \mathcal{A}), H_U)$.

To implement the GPRSM (1.3), we use the following matlab code to produce the matrices

C, H_L and H_U:

$rand(state, 0); C = rand(n, n); C = (C + C) - ones(n, n) + eye(n);$

%%%C is symmetric and C_{ij} is in$(-1, 1), C_{jj}$ is in $(0, 2)$%%%

$HU = ones(n, n) * 0.1; HL = -HU;\ for\ i = 1 : n\ HU(i, i) = 1; HL(i, i) = 1; end;$

And we set $n = 1000$ and $\gamma = 0.89$, then we apply the GPRSM (1.3) to solve the problem (4.1) and show the convergence rate of the algorithm (1.3) by using different penalty parameter ρ in Fig.4.4. From Fig.4.4, it is easy to see that the iterations decrease faster when the penalty parameter ρ get larger in our experiment. In fact, the value of the parameter ρ should not be too large or too small, since the convergence will be poor if the ρ is too small or too large. On the other hand, we can also observe that the selection of α can affect the convergence rate of the GPRSM (1.3) significantly. Clearly, we see that when α is close to 0.5, the number of iterations is more than the cases where α is close to 1.8. Therefore, we can obtain that the convergence rate of the algorithm (1.3) is related to the choice of the parameter ρ or α by our experiment.

5 Conclusions

In this paper, we further study the convergence rate of the Generalized version of the Peaceman-Rachford splitting method by establishing the worst-case $\mathcal{O}(1/t)$ convergence rate for the algorithm (1.3) in both the ergodic and a nonergodic senses. And we further illustrated its numerical

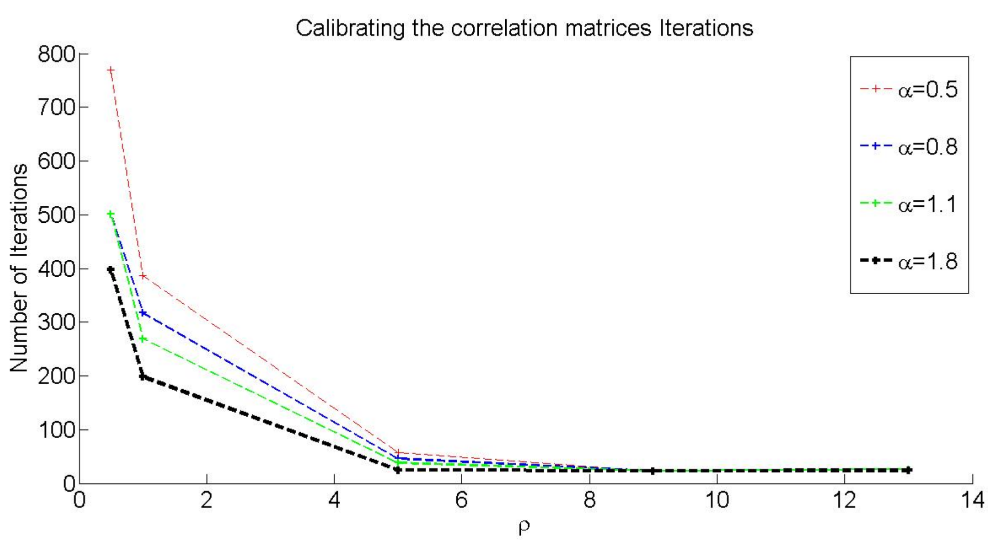

Figure 4.4: evolutions of the number of iterations and the penalty parameter ρ when the stopping criterion is achieved

efficiency and convergence rate by some numerical experiments.

References

1 Glowinski, R., Marrocco, A.: Approximation par elements finis dordre un, et la resolution, par penalisation-dualite, dune classe de problmes de Dirichlet non lineaires. RAIRO Anal. Numer. 9, 41-76 (1975)

2 D. Gabay, B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximations, Comput. Math. Appl. 2(1976) 17-40.

3 Eckstein, J., Bertsekas, D.: On theDouglas-Rachford splittingmethod and the proximal point algorithm for maximal monotone operators. Math. Program. 55, 293-318 (1992)

4 Han, D., Yuan, X.M.: Local linear convergence of the alternating direction method of multipliers for quadratic programs. SIAM J. Numer. Anal. 51(6), 3446-3457 (2013)

5 Peaceman, D.H., Rachford, H.H.: The numerical solution of parabolic elliptic differential equations. SIAM J. Appl. Math. 3, 28-41 (1955) (2012)

6 Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. 16, 964-979 (1979)

7 He, B.S., Liu, H.,Wang, Z.R., Yuan, X.M.: A strictly contractive Peaceman-Rachford splitting method for convex programming. SIAM J. Optim. 24(3), 1011-1040 (2014)

8 Bertsekas, D.P.: Constrained Optimization and Lagrange Multiplier Methods. Academic Press, New York (1982)

9 Sun, M., Liu, J.: A proximal Peaceman-Rachford splitting method for compressive sensing, J. Appl. Math. Comput., 50:349-363 (2016)

10 Sun, M., Liu, J.: Generalized Peaceman-Rachford splitting method for separable convex programming with applications to image processing, J. Appl. Math. Comput., 51:605-622 (2016)

11 He, B., Yang, H.: Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities. Oper. Res. Let. 23, 151-161 (1998)

12 He, B., Yuan, X.: On the O(1/n) convergence rate of Douglas-Rachford alternating direction method. SIAM J. Numer. Anal. 50, 700-709 (2012)

13 Sun, M., Wang,YJ., Liu, J.: Generalized Peaceman-Rachford splitting method for multiple-block separable convex programming with applications to robust PCA, Calcolo., 54:77-94 (2017)

14 Ethan X,Fang, Bingsheng He, Han Liu, Xiaoming Yuan, Generalized alternating direction method of multipliers: new theoretical insights and applications. Math. Prog. Comp.7, 149-187(2015)

15 Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. 3, 1-122 (2011)

16 D. Gabay, Applications of the method of multipliers to variational inequalities, in Augmented Lagrange Methods: Applications to the Solution of Boundary-Valued Problems, M. Fortin and R. Glowinski, eds., NorthCHolland, Amsterdam, 1983, pp. 299-331.

17 R. Glowinski and P. Le Tallec, Augmented Lagrangian and Operator Splitting Methods in Nonlinear Mechanics, SIAM Stud. Appl. Math. 9, SIAM, Philadelphia, 1989.

18 R. T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.