用于求解多块可分凸优化问题的惯性临近严格 收缩PRSM

王丽敏, 蒋 君, 邓 钊, 冯育强, 侯聪雅

武汉科技大学理学院, 湖北 武汉

收稿日期: 2025年1月10日; 录用日期: 2025年3月8日; 发布日期: 2025年3月18日

摘要

近年来,PRSM成为了处理具有线性约束的两部分可分凸优化问题的一个热门研究方向。本研究聚焦于 目标函数由三个解耦变量函数之和构成的可分凸优化问题。单纯地运用PRSM可能无法保证其收敛性。 因此,我们引入了一种带有惯性项和临近项的严格收缩PRSM。借助变分不等式、临近点算法以及基本不 等式,我们对所提出的方法进行了全局收敛性的分析。此外,我们将这种新方法应用于鲁棒主成分分析 (PCA)问题的求解,并提供了一些初步的数值结果,用以展示该方法的可行性和有效性。

关键词

凸优化,Peaceman-Rachford Splitting Method,惯性临近点方法,变分不等式,全局收敛性

Inertial Proximal Strictly Contractive PRSM for Solving Multi-Block Separable Convex Optimization

Limin Wang, Jun Jiang, Zhao Deng, Yuqiang Feng, Congya Hou

School of Science, Wuhan University of Science and Technology, Wuhan Hubei

Received: Jan. 10th, 2025; accepted: Mar. 8th, 2025; published: Mar. 18th, 2025

Abstract

In recent years, PRSM has become a popular research direction for dealing with two-block separable convex optimization problems with linear constraints. This study focuses on separable convex optimization problems where the objective function is composed of the sum of three decoupled variable functions. Simply applying the PRSM may not guarantee its convergence. Therefore, we introduce a strictly contractive PRSM with inertial and proximal terms. By means of variational inequalities, the proximal point algorithm, and fundamental inequalities, we have analyzed the global convergence of the proposed method. In addition, we applied this new method to the solution of the Robust Principal Component Analysis (PCA) problem and provided some preliminary numerical results to demonstrate the feasibility and effectiveness of the method.

Keywords

Convex Optimization, Peaceman-Rachford Splitting Method, Inertial Proximal Point Method, Variational Inequality, Global Convergence

Copyright © 2025 by author(s) and Hans Publishers Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). http://creativecommons.org/licenses/by/4.0/

CC Open Access

1. 介绍

我们研究以下的可分凸规划模型

$$\min\left\{\theta_{1}\left(x_{1}\right)+\theta_{2}\left(x_{2}\right)+\theta_{3}\left(x_{3}\right)\right|A_{1}x_{1}+A_{2}x_{2}+A_{3}x_{3}=b, x_{i}\in\mathcal{X}_{i}, i=1,2,3\right\},$$
(1)

其中, $\theta_i : \mathbb{R}^{n_i} \to \mathbb{R} \cup \{+\infty\} (i = 1, 2, 3)$ 是闭合凸函数(可能是非光滑的), $\mathcal{X}_i \subset \mathbb{R}^{n_i} (i = 1, 2, 3)$ 是非空闭凸集, $A_i \in \mathbb{R}^{l \times n_i}$ 是线性算子, $b \in \mathbb{R}^l$ 是一个给定的向量。可分离问题(1)在许多领域有着广泛的应用,例如图像 处理[1][2]、统计学习[3][4]以及矩阵分解[5]。

方程(1)的增广拉格朗日函数是

$$L_{\beta}(x_{1}, x_{2}, x_{3}, \lambda) = \theta_{1}(x_{1}) + \theta_{2}(x_{2}) + \theta_{3}(x_{3}) - \lambda^{T} (A_{1}x_{1} + A_{2}x_{2} + A_{3}x_{3} - b) + \frac{\beta}{2} \|A_{1}x_{1} + A_{2}x_{2} + A_{3}x_{3} - b\|^{2},$$
(2)

其中 $\lambda \in \mathbb{R}^{l}$ 是拉格朗日乘数, $\beta > 0$ 是惩罚参数。

问题(1)可以通过使用许多经典方法有效解决,例如 Hestenes [6]和 Powell [7]的乘子法,特别是增广 拉格朗日方法(ALM) [8], Douglas-Rachford 分裂方法(DRSM) [9],交替方向乘子法(ADMM) [10] [11], PRSM [12]以及前向 - 后向分裂方法[13]。在本文中,我们将专注于 PRSM。

为了确保 PRSM 生成的迭代序列具有严格的收缩性, He 等人[14]开发了一种严格收缩的 Peaceman-Rachford 分裂方法(SCPRSM),并建立了全局收敛性。惯性可以提高算法的计算效率,惯性起源于动态系统,可以追溯到 20 世纪 60 年代的重球方法[15]。多块可分离凸优化问题是一个热门话题。Wu 等人[16] 提出了一个临近 PRSM 来解决多块可分离凸最小化问题,其中一些子问题可以并行求解。对于其他多块可分离凸优化问题的研究,我们请读者参考[17]。

在本文中,提出了一种带有不确定项的惯性临近严格收缩的 PRSM,用于解决多块可分离凸优化问题。

(4)

$$\begin{cases} \left(\overline{x}_{1}^{k}, \overline{x}_{2}^{k}, \overline{x}_{3}^{k}, \overline{\lambda}^{k}\right) = \left(x_{1}^{k}, x_{2}^{k}, x_{3}^{k}, \lambda^{k}\right) + \rho_{k}\left(x_{1}^{k} - x_{1}^{k-1}, x_{2}^{k} - x_{2}^{k-1}, x_{3}^{k} - x_{3}^{k-1}, \lambda^{k} - \lambda^{k-1}\right), \\ x_{1}^{k+1} = \operatorname*{arg\,min}_{x_{1} \in \chi_{1}} \left\{ \theta_{1}\left(x_{1}\right) - \left(\overline{\lambda}^{k}\right)^{\mathrm{T}} A_{1}x_{1} + \frac{\beta}{2} \left\| A_{1}x_{1} + A_{2}\overline{x}_{2}^{k} + A_{3}\overline{x}_{3}^{k} - b \right\|^{2} + \frac{1}{2} \left\| x_{1} - \overline{x}_{1}^{k} \right\|_{B}^{2} \right\}, \\ x_{2}^{k+1} = \operatorname*{arg\,min}_{x_{2} \in \chi_{2}} \left\{ \theta_{2}\left(x_{2}\right) - \left(\overline{\lambda}^{k}\right)^{\mathrm{T}} A_{2}x_{2} + \frac{\beta}{2} \left\| A_{1}x_{1}^{k+1} + A_{2}x_{2} + A_{3}\overline{x}_{3}^{k} - b \right\|^{2} + \frac{1}{2} \left\| x_{2} - \overline{x}_{2}^{k} \right\|_{C}^{2} \right\}, \\ \lambda^{k+\frac{1}{2}} = \overline{\lambda}^{k} - \alpha\beta\left(A_{1}x_{1}^{k+1} + A_{2}x_{2}^{k+1} + A_{3}\overline{x}_{3}^{k} - b\right), \\ x_{3}^{k+1} = \operatorname*{arg\,min}_{x_{3} \in \chi_{3}} \left\{ \theta_{3}\left(x_{3}\right) - \left(\lambda^{k+\frac{1}{2}}\right)^{\mathrm{T}} A_{3}x_{3} + \frac{\beta}{2} \left\| A_{1}x_{1}^{k+1} + A_{2}x_{2}^{k+1} + A_{3}x_{3} - b \right\|^{2} + \frac{1}{2} \left\| x_{3} - \overline{x}_{3}^{k} \right\|_{D}^{2} \right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha\beta\left(A_{1}x_{1}^{k+1} + A_{2}x_{2}^{k+1} + A_{3}x_{3}^{k+1} - b\right), \end{cases}$$

$$(3)$$

其中 $\rho_k \in [0,1/3)$ 是一个惯性参数, $\alpha \in (0,(1+\sqrt{17})/8)$ 是一个步长参数, $\beta > 0$ 是一个惩罚参数和 $B = r_1 I_n - \beta A_1^T A_1, C = r_2 I_n - \beta A_2^T A_2, D = \tau r_3 I_n - \beta A_3^T A_3,$

 $\mathbb{K} \stackrel{\text{\tiny Therefore}}{=} r_1 \geq \beta \left\| A_1^{\mathsf{T}} A_1 \right\|, r_2 \geq \beta \left\| A_2^{\mathsf{T}} A_2 \right\|, r_3 > \beta \left\| A_3^{\mathsf{T}} A_3 \right\| \circ$

在上述描述中, $r_1 \ge \beta \|A_1^T A_1\|$ 意味着矩阵 B 是半正定的, $r_2 \ge \beta \|A_2^T A_2\|$ 意味着矩阵 C 是半正定的, 而 $\tau \in ((1+\alpha)/2, 1)$ 表示矩阵 D 的不定性。

本文的其余部分组织如下。在第2节中,我们总结了一些为进一步分析所需的预备知识。然后,在 第3节中,我们对惯性临近严格收缩的 Peaceman-Rachford 分裂方法(IPSCPRSM)的预测 - 校正进行重新 表述,以通过变分不等式建立全局收敛性。在第4节中,我们进行了一些数值模拟来说明所提出方法的 有效性。第5节得出了一些结论。

2. 预备知识

在这一部分,我们提供一些预备知识和简单的结果,以便进行进一步的分析。首先,我们用 \mathbb{R} 来表示通常的欧几里得空间,并用 $\|.\|$ 和 $\langle .,. \rangle$ 来表示 \mathbb{R} 中的欧几里得范数和内积。如果 $x \in \mathbb{R}^n$ 中的一个向量, D 是一个 $n \times n$ 的对称矩阵,我们使用记号 $\|x\|_{D}^{2} = x^{T}Dx$ 来表示 D-范数。

接下来,我们引入一些辅助序列和矩阵:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, w = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda \end{pmatrix}, F(w) = \begin{pmatrix} -A_1^T \lambda \\ -A_2^T \lambda \\ -A_3^T \lambda \\ r \end{pmatrix}, r = A_1 x_1 + A_2 x_2 + A_3 x_3 - b,$$
(5)

和

$$\tilde{w}^{k} = \begin{pmatrix} \tilde{x}_{1}^{k} \\ \tilde{x}_{2}^{k} \\ \tilde{x}_{3}^{k} \\ \tilde{\lambda}^{k} \end{pmatrix} = \begin{pmatrix} x_{1}^{k+1} \\ x_{2}^{k+1} \\ x_{3}^{k+1} \\ \overline{\lambda}^{k} - \beta \left(A_{1} x_{1}^{k+1} + A_{2} \overline{x}_{2}^{k} + A_{3} \overline{x}_{3}^{k} - b \right) \end{pmatrix}.$$
(6)

Ŷ

$$Q = \begin{pmatrix} B & 0 & 0 & 0 \\ 0 & \beta A_2^{\mathrm{T}} A_2 + C & 0 & 0 \\ 0 & (1+\alpha)\beta A_2 A_3^{\mathrm{T}} & \beta A_3^{\mathrm{T}} A_3 + D & 0 \\ 0 & -A_2 & -A_3 & \frac{1}{\beta} I_l \end{pmatrix},$$
(7)

$$M = \begin{pmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & I_{n_2} & 0 & 0 \\ 0 & 0 & I_{n_3} & 0 \\ 0 & -2\alpha\beta A_2 & -\alpha\beta A_3 & 2\alpha I_l \end{pmatrix},$$
(8)

$$H = \begin{pmatrix} B & 0 & 0 & 0 \\ 0 & r_2 I_{n_2} & 0 & 0 \\ 0 & (1+\alpha)\beta A_2 A_3^{\mathrm{T}} & \tau r_3 I_{n_3} & 0 \\ 0 & 0 & -\frac{1}{2}A_3 & \frac{1}{2\alpha\beta}I_l \end{pmatrix},$$
(9)

$$G = \begin{pmatrix} B & 0 & 0 & 0 \\ 0 & (1-2\alpha)\beta A_2^{\mathrm{T}}A_2 + C & (1-\alpha)\beta A_2 A_3^{\mathrm{T}} & (2\alpha-1)A_2 \\ 0 & -\alpha\beta A_2 A_3^{\mathrm{T}} & (1-\alpha)\beta A_3^{\mathrm{T}}A_3 + D & (\alpha-1)A_3 \\ 0 & (2\alpha-1)A_2 & (2\alpha-1)A_3 & \frac{(2-2\alpha)}{\beta}I_1 \end{pmatrix}.$$
 (10)

定义

从问题(4)中,我们可以得出结论,矩阵
$$D_0$$
是正定的,并且
 $D = \tau D_0 - (1 - \tau) \beta A_3^T A_3.$ (11)

变分不等式特征描述

设方程(1)的拉格朗日函数为

$$L(x_{1}, x_{2}, x_{3}, \lambda) = \theta_{1}(x_{1}) + \theta_{2}(x_{2}) + \theta_{3}(x_{3}) - \lambda^{T}(A_{1}x_{1} + A_{2}x_{2} + A_{3}x_{3} - b),$$

 $D_0 = r_3 I_{n_3} - \beta A_3^{\rm T} A_3,$

其定义域为 $\Omega := \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \times R^l$ 。如果它满足条件,我们称 $w^* = (x_1^*, x_2^*, x_3^*, \lambda^*) \in \Omega$ 为拉格朗日函数的 鞍点。

$$L(x_1^*, x_2^*, x_3^*, \lambda) \le L(x_1^*, x_2^*, x_3^*, \lambda^*) \le L_{x_i \in \chi_i, (i=1,2,3)}(x_1, x_2, x_3, \lambda^*).$$

相应地,我们推导出以下不等式:

$$\begin{cases} \theta_{1}(x_{1}) - \theta_{1}(x_{1}^{*}) + (x_{1} - x_{1}^{*})^{\mathrm{T}} (-A_{1}^{\mathrm{T}}\lambda^{*}) \geq 0, \quad \forall x_{1} \in \mathcal{X}_{1}, \\ \theta_{2}(x_{2}) - \theta_{2}(x_{2}^{*}) + (x_{2} - x_{2}^{*})^{\mathrm{T}} (-A_{2}^{\mathrm{T}}\lambda^{*}) \geq 0, \quad \forall x_{2} \in \mathcal{X}_{2}, \\ \theta_{3}(x_{3}) - \theta_{3}(x_{3}^{*}) + (x_{3} - x_{3}^{*})^{\mathrm{T}} (-A_{3}^{\mathrm{T}}\lambda^{*}) \geq 0, \quad \forall x_{3} \in \mathcal{X}_{3}, \\ (\lambda - \lambda^{*})^{\mathrm{T}} (A_{1}x_{1}^{*} + A_{2}x_{2}^{*} + A_{3}x_{3}^{*} - b) \geq 0, \quad \forall \lambda \in \mathbb{R}^{l}. \end{cases}$$

$$(12)$$

变分不等式(12)可以通过引用(5)以以下紧凑的形式表达出来:

$$\operatorname{VI}(\Omega, F, \theta) : \theta(x) - \theta(x^*) + (w - w^*)^{\mathrm{T}} F(w^*) \ge 0, \forall w \in \Omega,$$
(13)

其中

 $\theta(x) = \theta(x_1) + \theta(x_2) + \theta(x_3).$

根据映射 F 在(5)中的定义,我们有:

$$\left(w_{1}-w_{2}\right)^{\mathrm{T}}F\left(w_{1}\right)\geq\left(w_{1}-w_{2}\right)^{\mathrm{T}}F\left(w_{2}\right),\forall w\in\Omega.$$
(14)

我们用 Ω^* 表示变分不等式 VI(Ω, F, θ)的解集,这个解集是非空的。如果 $w^* = (x_1^*, x_2^*, x_3^*, \lambda^*) \in \Omega^*$,那 么 (x_1^*, x_2^*, x_3^*) 就是问题(1)的一个解。

我们在下面的引理中总结了上述定义的矩阵的一些性质。

引理 2.1. 设 $\alpha \in (0,(1+\sqrt{17})/8)$, $\tau \in ((1+\alpha)/2,1)$,则分别在(7)、(8)和(9)中定义的矩阵Q, M和H满足:

$$H \succeq 0, HM = Q, \tag{15}$$

和

 $G \succeq 0, G = Q^{\mathrm{T}} + Q - M^{\mathrm{T}} H M.$ ⁽¹⁶⁾

证明。利用 $B \succeq 0$, $C \succeq 0$ 和 $r_3 > \beta \| A_3^{\mathsf{T}} A_3 \|$ 。

$$\tilde{H} = \begin{pmatrix} B & 0 & 0 & 0 \\ 0 & \beta A_2^{\mathrm{T}} A_2 & 0 & 0 \\ 0 & (1+\alpha) \beta A_2 A_3^{\mathrm{T}} & \tau \beta A_3^{\mathrm{T}} A_3 & 0 \\ 0 & 0 & -\frac{1}{2} A_3 & \frac{1}{2\alpha\beta} I_l \end{pmatrix},$$

是半正定的。注意, *Ĥ* 可以表示为:

$$\tilde{H} = \frac{1}{2} \begin{pmatrix} I_{n_{1}} & 0 & 0 & 0 \\ 0 & A_{2}^{\mathrm{T}} & 0 & 0 \\ 0 & A_{3}^{\mathrm{T}} & A_{3}^{\mathrm{T}} & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \begin{pmatrix} 2B & 0 & 0 & 0 \\ 0 & 2\beta & 0 & 0 \\ 0 & 2\alpha\beta & 2\tau\beta & 0 \\ 0 & 0 & -1 & \frac{1}{\alpha\beta} \end{pmatrix} \begin{pmatrix} I_{n_{1}} & 0 & 0 & 0 \\ 0 & A_{2} & 0 & 0 \\ 0 & 0 & A_{3} & 0 \\ 0 & 0 & 0 & I \end{pmatrix}$$

由于 $\alpha \in \left(0, \left(1 + \sqrt{17}\right)/8\right)$,我们只需要检查

$$\begin{pmatrix} 2B & 0 & 0 & 0 \\ 0 & 2\beta & 0 & 0 \\ 0 & 2\alpha\beta & 2\tau\beta & 0 \\ 0 & 0 & -1 & \frac{1}{\alpha\beta} \end{pmatrix} \succeq \begin{pmatrix} 2B & 0 & 0 & 0 \\ 0 & 2\beta & 0 & 0 \\ 0 & 2\alpha\beta & 2\tau\beta & 0 \\ 0 & 0 & -1 & \frac{1}{\beta} \end{pmatrix} \succeq 0,$$

回顾 $\tau \in ((1+\alpha)/2, 1)$,我们随后明确得到了H的半正定性。

因此,我们将证明在方程(10)中定义的矩阵 *G* 是一个半正定矩阵。同样,通过使用 $\tau > (1+\alpha)/2, r_2 \ge \beta \|A_2^{\mathsf{T}}A_2\| \Pi r_3 > \beta \|A_3^{\mathsf{T}}A_3\|$,我们知道,如果以下矩阵是半正定的,那么*G* \succeq 0 是有保证的。

$$\tilde{G} = \begin{pmatrix} B & 0 & 0 & 0 \\ 0 & (1-2\alpha)\beta A_2^{\mathrm{T}}A_2 & (1-\alpha)\beta A_2 A_3^{\mathrm{T}} & (2\alpha-1)A_2 \\ 0 & -\alpha\beta A_2 A_3^{\mathrm{T}} & \frac{(1-\alpha)}{2}\beta A_3^{\mathrm{T}}A_3 & (\alpha-1)A_3 \\ 0 & (2\alpha-1)A_2 & (2\alpha-1)A_3 & \frac{(2-2\alpha)}{\beta}I_1 \end{pmatrix} \succeq 0.$$

并等价地证明以下方程

$$f(\alpha) = -2\alpha^2 + \frac{1}{2}\alpha + \frac{1}{2},$$

因此,我们得到对于任意 $\alpha \in (0, (1+\sqrt{17})/8)$,矩阵G的正半定性自然得以确立。

3. 收敛性分析

在本节中,我们将基于变分不等式建立 IPSCPRSM(3)的全局收敛性,并展示它如何被重写为 He 等 人[18] [19]提出的预测 - 校正框架。最终,我们将展示 IPSCPRSM(3)的全局收敛性。

3.1. 预测 - 校正重构

假设 3.1. 设 { ρ_k }[∞]_{k=0} 被选择使得: (i) 对于所有 $k \ge 0, 0 \le \rho_k \le \rho$ 其中 $\rho \in [0, 1/3)$, (ii) 由 IPSCPRSM(3) 生成的序列 { w^k } 满足

$$\sum_{k=0}^{\infty} \rho_k \left\| w^k - w^{k-1} \right\|_H^2 < \infty.$$
(17)

确保假设 3.1 的一种方式是选择

$$\rho_k = \min\left\{\rho, 1/(k \|w^k - w^{k-1}\|_H)^2\right\}$$

假设 3.2. 设问题(1)的解集非空。

辅助变量 $\{\tilde{w}^k\}$ 表示预测变量,而 $\{w^{k+1}\}$ 表示校正变量。下面,我们展示 IPSCPRSM (3)的预测步骤的结果。

$$\theta(x) - \theta(\tilde{x}^{k}) + (w - \tilde{w}^{k})^{\mathrm{T}} F(\tilde{w}^{k}) \ge (w - \tilde{w}^{k})^{\mathrm{T}} Q(\overline{w}^{k} - \tilde{w}^{k}), \forall w \in \Omega.$$
(18)

其中 Q 在(7)中定义。

证明。基于(3)中x₁子问题的一阶最优性条件,我们得到

$$\theta_{1}\left(x_{1}\right)-\theta_{1}\left(\tilde{x}_{1}^{k}\right)+\left(x_{1}-\tilde{x}_{1}^{k}\right)^{\mathrm{T}}\left[-A_{1}^{\mathrm{T}}\tilde{\lambda}^{k}+B\left(\tilde{x}_{1}^{k}-\overline{x}_{1}^{k}\right)\right]\geq0,\forall x_{1}\in\mathcal{X}_{1}.$$
(19)

类似地,结合最优性条件和(5),我们得到 $\tilde{x}_2^k \in \mathcal{X}_2$ 和

$$\theta_{2}(x_{2}) - \theta_{2}(\tilde{x}_{2}^{k}) + (x_{2} - \tilde{x}_{2}^{k})^{\mathrm{T}} \Big[-A_{2}^{\mathrm{T}} \overline{\lambda}^{k} + \beta A_{2}^{\mathrm{T}} \Big(A_{1} x_{1}^{k+1} + A_{2} x_{2}^{k+1} + A_{3} \overline{x}_{3}^{k} - b \Big) + C \Big(x_{2}^{k+1} - \overline{x}_{2}^{k} \Big) \Big] \ge 0, \forall x_{2} \in \mathcal{X}_{2}.$$

从方程(6)我们推导出

$$\beta \left(A_1 x_1^{k+1} + A_2 x_2^{k+1} + A_3 \overline{x}_3^k - b \right) = \left(\overline{\lambda}^k - \widetilde{\lambda}^k \right) + \beta A_2 \left(x_2^{k+1} - \overline{x}_2^k \right).$$
(20)

将方程(20)代入上述不等式中,我们得到

$$\theta_2(x_2) - \theta_2(\tilde{x}_2^k) + (x_2 - \tilde{x}_2^k)^{\mathrm{T}} \Big[-A_2^{\mathrm{T}} \tilde{\lambda}^k + (\beta A_2^{\mathrm{T}} A_2 + C) (\tilde{x}_2^k - \overline{x}_2^k) \Big] \ge 0, \forall x_2 \in \chi_2.$$

$$(21)$$

类似地,我们有

$$\theta_{3}(x_{3}) - \theta_{3}(\tilde{x}_{3}^{k}) + (x_{3} - \tilde{x}_{3}^{k})^{\mathrm{T}} \left[-A_{3}^{\mathrm{T}} \lambda^{k+\frac{1}{2}} + \beta A_{3}^{\mathrm{T}} \left(A_{1} \tilde{x}_{1}^{k} + A_{2} \tilde{x}_{2}^{k} + A_{3} \tilde{x}_{3}^{k} - b \right) + D\left(\tilde{x}_{3}^{k} - \overline{x}_{3}^{k} \right) \right] \ge 0, \forall x_{3} \in \chi_{3}.$$

根据(6)中对 w^k 的定义和(3)中的方案,我们推导出

$$\lambda^{k+\frac{1}{2}} = \overline{\lambda}^{k} + \alpha \left(\tilde{\lambda}^{k} - \overline{\lambda}^{k} \right) - \alpha \beta A_{2} \left(x_{2}^{k+1} - \overline{x}_{2}^{k} \right).$$
⁽²²⁾

将方程(22)和 $\tilde{\lambda}^{k} = \overline{\lambda}^{k} - \beta \left(A_{1} \tilde{x}_{1}^{k} + A_{2} \overline{x}_{2}^{k} + A_{3} \overline{x}_{3}^{k} - b \right)$ 代入上述不等式中,我们得到

$$\theta_{3}(x_{3}) - \theta_{3}\left(\tilde{x}_{3}^{k}\right) + \left(x_{3} - \tilde{x}_{3}^{k}\right)^{\mathrm{T}} \left[-A_{3}^{\mathrm{T}}\tilde{\lambda}^{k} + (1+\alpha)\beta A_{2}A_{3}^{\mathrm{T}}\left(\tilde{x}_{2}^{k} - \overline{x}_{2}^{k}\right) + \left(\beta A_{3}^{\mathrm{T}}A_{3} + D\right)\left(\tilde{x}_{3}^{k} - \overline{x}_{3}^{k}\right) - A_{3}^{\mathrm{T}}\alpha\left(\tilde{\lambda}^{k} - \overline{\lambda}^{k}\right)\right] \geq 0,$$

$$\forall x_{3} \in \chi_{3}.$$

$$(23)$$

最后,我们可以将(6)中的关系式 $\tilde{\lambda}^{k} = \overline{\lambda}^{k} - \beta \left(A_{1} x_{1}^{k+1} + A_{2} \overline{x}_{2}^{k} + A_{3} \overline{x}_{3}^{k} - b \right)$ 重写为

$$\left(\lambda - \tilde{\lambda}^{k}\right)^{\mathrm{T}} \left[\tilde{r}^{k} - A_{2}\left(\tilde{x}_{2}^{k} - \overline{x}_{2}^{k}\right) - A_{3}\left(\tilde{x}_{3}^{k} - \overline{x}_{3}^{k}\right) + \frac{1}{\beta}\left(\tilde{\lambda}^{k} - \overline{\lambda}^{k}\right)\right] \ge 0, \forall \lambda \in \mathbb{R}^{l}.$$

$$(24)$$

结合(19)、(21)、(23)和(24),我们可以直接得到断言(18)。 现在我们来说明如何使用预测变量 \tilde{w}^k 和前一个点 \bar{w}^k 来构造校正点 w^{k+1} 。 **引理 3.2.** 设 { w^k } 由 IPSCPRSM(3)生成, { \tilde{w}^k } 在(6)中定义。那么,我们有

$$\overline{w}^{k} - w^{k+1} = M\left(\overline{w}^{k} - \widetilde{w}^{k}\right), \tag{25}$$

其中 M 在(8)中定义。

证明。根据(6)

$$\overline{\lambda}^k - \widetilde{\lambda}^k = \beta \Big(A_1 x_1^{k+1} + A_2 \overline{x}_2^k + A_3 \overline{x}_3^k - b \Big).$$

利用 w^k 的表示,以及(3)、(5)和(22),我们得出结论

$$\lambda^{k+1} = \overline{\lambda}^{k} - \left[2\alpha \left(\overline{\lambda}^{k} - \widetilde{\lambda}^{k} \right) - 2\alpha\beta A_{2} \left(\overline{x}_{2}^{k} - \widetilde{x}_{2}^{k} \right) - \alpha\beta A_{3} \left(\overline{x}_{3}^{k} - \widetilde{x}_{3}^{k} \right) \right]$$

因此,我们有以下关系

$$\begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ x_3^{k+1} \\ \lambda^{k+1} \end{pmatrix} = \begin{pmatrix} \overline{x}_1^k \\ \overline{x}_2^k \\ \overline{x}_3^k \\ \overline{\lambda}^k \end{pmatrix} - \begin{pmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & I_{n_2} & 0 & 0 \\ 0 & 0 & I_{n_3} & 0 \\ 0 & -2\alpha\beta A_2 & -\alpha\beta A_3 & 2\alpha I_l \end{pmatrix} \begin{pmatrix} \overline{x}_1^k - \tilde{x}_1^k \\ \overline{x}_2^k - \tilde{x}_2^k \\ \overline{x}_3^k - \tilde{x}_3^k \\ \overline{\lambda}^k - \tilde{\lambda}^k \end{pmatrix}$$

它可以被改写成一个简洁的形式

$$w^{k+1} = \overline{w}^k - M\left(\overline{w}^k - \widetilde{w}^k\right),$$

其中 M 在式(8)中定义。

3.2. 全局收敛性

在本节中,我们展示了 IPSCPRSM (3)的全局收敛性,并考察了其迭代复杂度。 **引理 3.3.** 设 $\{w^k\}$ 由 IPSCPRSM (3)生成, $\{\tilde{w}^k\}$ 在(6)中定义。那么对于任何 $\tilde{w}^k \in \Omega$,我们有

$$\left(w - \tilde{w}^{k}\right)^{\mathrm{T}} \mathcal{Q}\left(\overline{w}^{k} - \tilde{w}^{k}\right) = \frac{1}{2} \left(\left\|w - w^{k+1}\right\|_{H}^{2} - \left\|w - \overline{w}^{k}\right\|_{H}^{2}\right) + \frac{1}{2} \left\|\overline{w}^{k} - \tilde{w}^{k}\right\|_{G}^{2}.$$
(26)

证明。结合
$$Q = HM$$
 (15) 和 $\left(\overline{w}^{k} - w^{k+1}\right) = M\left(\overline{w}^{k} - \tilde{w}^{k}\right)$ (25), 我们推导出
 $\left(w - \tilde{w}^{k}\right)^{\mathrm{T}} Q\left(\overline{w}^{k} - \tilde{w}^{k}\right) = \left(w - \tilde{w}^{k}\right)^{\mathrm{T}} HM\left(\overline{w}^{k} - \tilde{w}^{k}\right) = \left(w - \tilde{w}^{k}\right)^{\mathrm{T}} H\left(\overline{w}^{k} - w^{k+1}\right).$

对于一个对称矩阵 H 和向量 $a, b, c, d \in R'$, 它遵循以下恒等式

$$(a-b)^{\mathrm{T}} H(c-d) = \frac{1}{2} \left(\left\| a - d \right\|_{H}^{2} - \left\| a - c \right\|_{H}^{2} \right) + \frac{1}{2} \left(\left\| c - b \right\|_{H}^{2} - \left\| d - b \right\|_{H}^{2} \right).$$

应用上述恒等式, 令 $a = w, b = \tilde{w}^k, c = \overline{w}^k, d = w^{k+1}$, 得到

$$\left(w - \tilde{w}^{k}\right)^{\mathrm{T}} H\left(\overline{w}^{k} - w^{k+1}\right) = \frac{1}{2} \left(\left\|w - w^{k+1}\right\|_{H}^{2} - \left\|w - \overline{w}^{k}\right\|_{H}^{2}\right) + \frac{1}{2} \left(\left\|\overline{w}^{k} - \tilde{w}^{k}\right\|_{H}^{2} - \left\|w^{k+1} - \tilde{w}^{k}\right\|_{H}^{2}\right).$$

$$(27)$$

对于(27)中的最后一项,我们可以推导出

$$\begin{aligned} \left\|\overline{w}^{k} - \widetilde{w}^{k}\right\|_{H}^{2} &- \left\|w^{k+1} - \widetilde{w}^{k}\right\|_{H}^{2} \\ &= \left\|\overline{w}^{k} - \widetilde{w}^{k}\right\|_{H}^{2} - \left\|\left(\overline{w}^{k} - \widetilde{w}^{k}\right) - M\left(\overline{w}^{k} - \widetilde{w}^{k}\right)\right\|_{H}^{2} \\ &= \left\|\overline{w}^{k} - \widetilde{w}^{k}\right\|_{G}^{2}. \end{aligned}$$

$$(28)$$

现在,结合(10)、(27)和(28),我们得到断言(26)。

有了 3.1 和 3.3 两个引理,我们可以得到以下结果。

引理 3.4. 设 $\{w^{\epsilon}\}$ 由 IPSCPRSM (3)生成, $\{\tilde{w}^{\epsilon}\}$ 在(6)中定义。那么,我们有

$$\theta(x) - \theta(\tilde{x}^{k}) + (w - \tilde{w}^{k})^{\mathrm{T}} F(\tilde{w}^{k}) \geq \frac{1}{2} \left(\left\| w - w^{k+1} \right\|_{H}^{2} - \left\| w - \overline{w}^{k} \right\|_{H}^{2} \right) + \frac{1}{2} \left\| \overline{w}^{k} - \tilde{w}^{k} \right\|_{G}^{2}, \forall w \in \Omega.$$

$$(29)$$

证明。通过关联(18)和(26)来验证恒等式(29)。

引理 3.5. 设 $\{w^k\}$ 由 IPSCPRSM (3)生成,那么对于任何 $w^{k+1} \in \Omega$,我们有

$$\theta(x) - \theta(x^{k+1}) + (w - w^{k+1})^{\mathrm{T}} \left[F(w^{k+1}) + H(w^{k+1} - \overline{w}^{k}) \right]$$

$$\geq \beta(r^{k+1})^{\mathrm{T}} A_{2}(\overline{x}_{2}^{k} - x_{2}^{k+1}) + (1 - \alpha)\beta(r^{k+1})^{\mathrm{T}} A_{3}(\overline{x}_{3}^{k} - x_{3}^{k+1}) + (1 - 2\alpha)\beta \left\| r^{k+1} \right\|^{2}.$$

$$(30)$$

其中r在(5)中定义。

借助引理 3.5,我们使用基本不等式推导出下一个引理。

引理 3.6. 设 $\{w^k\}$ 由 IPSCPRSM (3)生成,那么对于任何 $w^{k+1} \in \Omega$,我们有

$$-\beta (r^{k+1})^{\mathrm{T}} A_{2} (\overline{x}_{2}^{k} - x_{2}^{k+1}) - (1 - \alpha) \beta (r^{k+1})^{\mathrm{T}} A_{3} (\overline{x}_{3}^{k} - x_{3}^{k+1}) - (1 - 2\alpha) \beta \|r^{k+1}\|^{2}$$

$$\leq \frac{1}{2} \|w^{k+1} - \overline{w}^{k}\|_{M_{0}}^{2}.$$
(31)

其中

$$M_{0} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \beta A_{2}^{\mathrm{T}} A_{2} + C & 0 & 0 \\ 0 & (1+\alpha)\beta A_{2} A_{3}^{\mathrm{T}} & \beta A_{3}^{\mathrm{T}} A_{3} + D & 0 \\ 0 & 0 & -\frac{1}{2} A_{3} & \frac{1}{2\alpha\beta} I_{l} \end{pmatrix},$$
(32)

和

$$H_{0} = \begin{pmatrix} \beta A_{2}^{\mathrm{T}} A_{2} & 0 & 0\\ (1+\alpha)\beta A_{2} A_{3}^{\mathrm{T}} & \beta A_{3}^{\mathrm{T}} A_{3} & 0\\ 0 & -\frac{1}{2} A_{3} & \frac{1}{2\alpha\beta} I_{l} \end{pmatrix}.$$
 (33)

证明。根据(32)的定义,我们有

$$\frac{1}{2} \left\| w^{k+1} - \overline{w}^k \right\|_{M_0}^2 = \frac{1}{2} \left\| v^{k+1} - \overline{v}^k \right\|_{H_0}^2 + \frac{1}{2} \left\| x_2^{k+1} - \overline{x}_2^k \right\|_C^2 + \frac{1}{2} \left\| x_3^{k+1} - \overline{x}_3^k \right\|_D^2.$$

类似地,根据(33)中H₀的关系,我们得到

$$\begin{aligned} &\frac{1}{2} \left\| \boldsymbol{v}^{k+1} - \overline{\boldsymbol{v}}^{k} \right\|_{H_{0}}^{2} \\ &= \frac{1}{2} \beta A_{2}^{2} \left(\boldsymbol{x}_{2}^{k+1} - \overline{\boldsymbol{x}}_{2}^{k} \right)^{2} + \frac{1}{2} (1 + \alpha) \beta A_{2} A_{3}^{\mathsf{T}} \left(\boldsymbol{x}_{2}^{k+1} - \overline{\boldsymbol{x}}_{2}^{k} \right)^{\mathsf{T}} \left(\boldsymbol{x}_{3}^{k+1} - \overline{\boldsymbol{x}}_{3}^{k} \right) \\ &+ \frac{1}{2} \beta A_{3}^{2} \left(\boldsymbol{x}_{3}^{k+1} - \overline{\boldsymbol{x}}_{3}^{k} \right)^{2} - \frac{1}{4} A_{3} \left(\boldsymbol{x}_{3}^{k+1} - \overline{\boldsymbol{x}}_{3}^{k} \right)^{\mathsf{T}} \left(\boldsymbol{\lambda}^{k+1} - \overline{\boldsymbol{\lambda}}^{k} \right) + \frac{1}{4 \alpha \beta} \left(\boldsymbol{\lambda}^{k+1} - \overline{\boldsymbol{\lambda}}^{k} \right)^{2}. \end{aligned}$$

通过使用(4)中C,D的定义以及 $\tau > (1+\alpha)/2$, $\alpha \in (0, (1+\sqrt{17})/8)$, 我们推导出

$$\frac{1}{2} \left\| w^{k+1} - \overline{w}^{k} \right\|_{M_{0}}^{2} \\
\geq \alpha \beta \left(r^{k+1} \right)^{2} + \frac{1 - \alpha}{4} \beta A_{3}^{2} \left(x_{3}^{k+1} - \overline{x}_{3}^{k} \right)^{2} - \beta \left(r^{k+1} \right)^{\mathrm{T}} A_{2} \left(\overline{x}_{2}^{k} - x_{2}^{k+1} \right).$$
(34)

由于 $\alpha \in \left(0, \left(1 + \sqrt{17}\right)/8\right)$,应用 $a^{T}b \le a^{2} + \frac{1}{4}b^{2} = a = r^{k+1} \pi b = A_{3}\left(x_{3}^{k+1} - \overline{x}_{3}^{k}\right)$ 的不等式,我们可以得出

$$-(1-\alpha)\beta(r^{k+1})^{\mathrm{T}}A_{3}(\overline{x}_{3}^{k}-x_{3}^{k+1})-(1-2\alpha)\beta\|r^{k+1}\|^{2} \leq \alpha\beta\|r^{k+1}\|^{2}+\frac{1-\alpha}{4}\beta\|A_{3}(x_{3}^{k+1}-\overline{x}_{3}^{k})\|^{2}.$$
(35)

将(34)和(35)结合起来,我们可以直接得到断言(31)。

参考(9)中 H 的定义, (32)中 M_0 的定义, 以及(4)中矩阵 B 的正半定性, 还有矩阵 C 和 D 的定义, 我 们得到 $M_0 \preceq H$ 。

引理 3.7. 设 $\{w^k\}$ 由 IPSCPRSM (3)生成,则对于任何 $w^* \in \Omega^*$,我们有 $\sum_{k=0}^{\infty} \|w^k - w^*\|_H^2 < \infty$ 。 证明。通过将 $w = w^{k+1}$ 代入(16)中,以及将 $w = w^*$ 代入(30)中,我们得到

$$\theta(x^{k+1}) - \theta(x^*) + (w^{k+1} - w^*)^{\mathrm{T}} F(w^*) \ge 0.$$

和

$$\theta(x^{*}) - \theta(x^{k+1}) + (w^{*} - w^{k+1})^{\mathrm{T}} \Big[F(w^{k+1}) + H(w^{k+1} - \overline{w}^{k}) \Big]$$

$$\geq \beta(r^{k+1})^{\mathrm{T}} A_{2}(\overline{x}_{2}^{k} - x_{2}^{k+1}) + (1 - \alpha)\beta(r^{k+1})^{\mathrm{T}} A_{3}(\overline{x}_{3}^{k} - x_{3}^{k+1}) + (1 - 2\alpha)\beta \|r^{k+1}\|^{2}.$$

将它们相加并结合(14),我们得到

$$(w^{k+1} - w^*)^{\mathrm{T}} H(w^{k+1} - \overline{w}^k) + \beta (r^{k+1})^{\mathrm{T}} A_2(\overline{x}_2^k - x_2^{k+1}) + (1 - \alpha) \beta (r^{k+1})^{\mathrm{T}} A_3(\overline{x}_3^k - x_3^{k+1}) + (1 - 2\alpha) \beta ||r^{k+1}||^2 \le 0.$$
(36)

通过回顾(3),我们有

$$\overline{w}^{k} = w^{k} + \rho_{k} \left(w^{k} - w^{k-1} \right).$$
(37)

因此可以得出

结合(31)、(36)和(38),我们得到

$$\begin{aligned} \varphi_{k+1} - \varphi_{k} - \rho_{k} \left(\varphi_{k} - \varphi_{k-1}\right) \\ &\leq -\frac{1}{2} \left\| w^{k+1} - w^{k} \right\|_{H}^{2} + \frac{\rho_{k}}{2} \left\| w^{k} - w^{k-1} \right\|_{H}^{2} \\ &+ \rho_{k} \left(w^{k+1} - w^{k} \right)^{\mathrm{T}} H \left(w^{k} - w^{k-1} \right) + \frac{1}{2} \left\| w^{k+1} - \overline{w}^{k} \right\|_{M_{0}}^{2}. \end{aligned} \tag{39}$$

将此与关系式(37)结合起来,注意 $M_0 \leq H$,我们得出

$$\varphi_{k+1} - \varphi_k - \rho_k \left(\varphi_k - \varphi_{k-1} \right) \le \frac{\rho_k + \rho_k^2}{2} \left\| w^k - w^{k-1} \right\|_H^2$$

因为 $\rho_k \in [0,1/3)$,这意味着 $\left(\rho_k + \rho_k^2\right)/2 \le \rho_k$,那么我们有

$$\varphi_{k+1} - \varphi_k - \rho_k \left(\varphi_k - \varphi_{k-1} \right) \le \rho_k \left\| w^k - w^{k-1} \right\|_H^2.$$
(40)

令 $\theta_k = \varphi_k - \varphi_{k-1}$, $\delta_k = \rho_k \left\| w^k - w^{k-1} \right\|_H^2$, 关系(40)表明 $\theta_{k+1} \le \rho_k \theta_k + \delta_k \le \rho[\theta_k]_+ + \delta_k$, 注意 $w^0 = w^{-1}$, 那么 $\theta_0 = [\theta_0]_+ = 0$, $\delta_0 = 0$ 。通过假设 3.1, 我们得到

$$\sum_{k=0}^{\infty} \left[\theta_k\right]_+ \le \frac{1}{1-\rho} \sum_{k=0}^{\infty} \delta_k = \frac{1}{1-\rho} \sum_{k=1}^{\infty} \delta_k < \infty.$$

$$\tag{41}$$

$$\label{eq:mk} \begin{split} & \diamondsuit m_k = \varphi_k - \sum_{j=1}^k \left[\theta_j \right]_+ \,, \ \text{结合(41)} \pi \, \varphi_k \geq 0 \,, \\ & \Re(1 \oplus 2^k) = \varphi_{k+1} - \left[\theta_{k+1} \right]_+ - \sum_{j=1}^k \left[\theta_j \right]_+ \leq \varphi_{k+1} - \theta_{k+1} - \sum_{j=1}^k \left[\theta_j \right]_+ = \varphi_k - \sum_{j=1}^k \left[\theta_j \right]_+ = m_k. \end{split}$$

这意味着 $\{m_k\}$ 是一个单调非增序列,因此 $\{m_k\}$ 收敛到

$$\lim_{k\to\infty}\varphi_k=\sum_{j=1}^{\infty}\left[\theta_j\right]_++\lim_{k\to\infty}m_k.$$

因此, $\{\varphi_k\}$ 收敛,并且作为结果, $\sum_{k=0}^{\infty} \|w^k - w^*\|_H^2 < \infty$ 也成立。引理 3.7 得证。 接下来的引理总结了有助于分析式(3)收敛性的几个关键事实。 **引理 3.8.** 设 $\{w^k\}$ 由 IPSCPRSM (3)生成,则对于任何 $w^{k+1} \in \Omega$,我们有

$$\sum_{k=0}^{\infty} \left\| w^{k+1} - \overline{w}^k \right\|_G^2 < \infty, \forall w \in \Omega.$$
(42)

证明。利用 F 的单调性, 可以得到

$$\left(w-\tilde{w}^{k}\right)^{\mathrm{T}}F\left(w\right)\geq\left(w-\tilde{w}^{k}\right)^{\mathrm{T}}F\left(\tilde{w}^{k}\right).$$

将上述不等式代入方程(18),我们得到

$$\theta(x) - \theta(\tilde{x}^{k}) + (w - \tilde{w}^{k})^{\mathrm{T}} F(w) \ge (w - \tilde{w}^{k})^{\mathrm{T}} Q(\overline{w}^{k} - \tilde{w}^{k}), \forall w \in \Omega.$$

设置 w = w^{*} 和式(13),我们得到

$$\left(\tilde{w}^{k}-w^{*}\right)^{\mathrm{T}}Q\left(\overline{w}^{k}-\tilde{w}^{k}\right)\geq0,\forall w\in\Omega.$$

将 $w = w^*$ 代入(26)中,我们得到

$$\left(w^{*}-\tilde{w}^{k}\right)^{\mathrm{T}}Q\left(\bar{w}^{k}-\tilde{w}^{k}\right)=\frac{1}{2}\left(\left\|w^{*}-w^{k+1}\right\|_{H}^{2}-\left\|w^{*}-\bar{w}^{k}\right\|_{H}^{2}\right)+\frac{1}{2}\left\|\bar{w}^{k}-\tilde{w}^{k}\right\|_{G}^{2}.$$

因为 $(\tilde{w}^k - w^*)^T Q(\bar{w}^k - \tilde{w}^k) \ge 0$,我们有

$$\left\|\overline{w}^{k} - \widetilde{w}^{k}\right\|_{G}^{2} \leq -\left\|w^{k+1} - w^{*}\right\|_{H}^{2} + \left\|\overline{w}^{k} - w^{*}\right\|_{H}^{2}.$$

根据 \overline{w}^{k} 的定义以及关系式 $\|a+b\|_{H}^{2} \leq 2\|a\|_{H}^{2} + 2\|b\|_{H}^{2}$,我们有

$$\left\|\overline{w}^{k} - \widetilde{w}^{k}\right\|_{G}^{2} \leq 2\left\|w^{k} - w^{*}\right\|_{H}^{2} - \left\|w^{k+1} - w^{*}\right\|_{H}^{2} + 2\left\|\rho_{k}\left(w^{k} - w^{k-1}\right)\right\|_{H}^{2}.$$

从 k = 0 到 k = ∞ 对两边求和。回顾假设 3.1、引理 3.7 和式(41),引理 3.8 得证。

借助引理 3.7 和 3.8,以下定理证明了由 IPSCPRSM (3)生成的序列的渐近可行性以及目标函数的收 敛性定理。

定理 3.1. 设 { α_k }[∞]_{k=0} 满足假设 3.1, $\alpha \in (0, (1+\sqrt{17})/8)$, $\beta > 0$ 。由 IPSCPRSM (3)生成的无限序列是 $\{w^k\}_{\mu=0}^{\infty}$,因此我们有

(i)
$$\sum_{k=1}^{\infty} \left\| r^k \right\|^2 < \infty$$
, $\mathbb{B} \not \sqcup \lim_{k \to \infty} \left\| r^k \right\| = 0$, $\not \amalg \not \sqcap r^k = A_1 x_1^k + A_2 x_2^k + A_3 x_3^k - b$;

(ii) 当 $k \to \infty$ 时, 目标函数值 $\theta_1(x_1^k) + \theta_2(x_2^k) + \theta_3(x_3^k)$ 收敛到(1)的最优值;

(iii) 序列 $\{w^k\}$ 收敛到 Ω^* 中的点 w° 。

以下定理为所提出的 IPSCPRSM (3)提供了特定的渐近复杂度结果。我们展示了 IPSCPRSM (3)能够 找到一个近似解,该解对于变分不等式 $VI(\Omega, F, \theta)$ 的精度为 $o(1/\sqrt{k})$ 。我们给出以下定理

定理 3.2. 假设序列 $\{w^k\}$ 由 IPSCPRSM (3)生成,当 $k \to \infty$ 时。

- (i) $\min_{1 \le i \le k} \left\| w^{i} \overline{w}^{i-1} \right\|_{G}^{-1} = o\left(1/\sqrt{k}\right);$ (ii) $\min_{1 \le i \le k} \left\| r^{i} \right\| = o\left(1/\sqrt{k}\right);$
- (iii) $\min_{1 \le i \le k} \left| \theta_1 \left(x_1^i \right) + \theta_2 \left(x_2^i \right) + \theta_3 \left(x_3^i \right) \theta_1 \left(x_1^* \right) \theta_2 \left(x_2^* \right) \theta_3 \left(x_3^* \right) \right| = o\left(\frac{1}{\sqrt{k}} \right).$

4. 数值实验

在这一部分,我们通过将 IPSCPRSM 应用于一个特定用例:鲁棒 PCA,来展示其实际效果。我们展 示了 IPSCPRSM 的初始数值结果,并将其性能与 VASALM [20]和 EADMM [21]进行了评估。

我们可以使用以下鲁棒 PCA 模型来描述它。

$$\min \|X\|_{*} + \tau \|Y\|_{1} + \frac{1}{2\mu} \|P_{\Omega}(Z)\|_{F}^{2},$$
s.t. $X + Y + Z = N,$
(43)

其中 $N \in \mathbb{R}^{g \times n}$, $\| \cdot \|_{*}$ 表示核范数(一个矩阵所有奇异值的和), $\| \cdot \|_{*}$ 表示促进稀疏性的 l_{1} 范数, 定义为所有条 目的绝对值之和,并且 $\tau > 0$ 是一个平衡参数, $\mu > 0$ 是一个控制高斯噪声包含的小常数。 Ω 是集合

 $\{1,2,\cdots,g\}\times\{1,2,\cdots,n\}$ 的一个子集,而 $P_{\Omega}(\cdot):\mathbb{R}^{g\times n}\to\mathbb{R}^{g\times n}$ 是对应的投影算子。

然后,我们可以将(43)重新表述,使其适合于可分离的凸优化(1)的范畴。具体来说,设

$$(x_1, x_2, x_3) = (X, Y, Z), A_i = I_p, i = 1, 2, 3, b = N$$

和

$$\theta_{1}(x_{1}) = \|X\|_{*}, \theta_{2}(x_{2}) = \tau \|Y\|_{1}, \theta_{3}(x_{3}) = \frac{1}{2\mu} \|P_{\Omega}(Z)\|_{F}^{2}$$

因此,所提出的 IPSCPRSM 是适用的。首先,让我们引入两个重要的算子。这些算子对于获得封闭形式的解是有用的。设c > 0为一个常数, $E \in \mathbb{R}^{g \times n}$ 为一个给定的矩阵。对于矩阵变量的收缩算子 shrink(·,·) 定义如下:

$$(\operatorname{shrink}(E,c))_{ii} := \operatorname{sign}(E_{ij}) \cdot \max\{|E_{ij}| - c, 0\}, 1 \le i \le g, 1 \le j \le n.$$

此外,我们在奇异值域中定义算子 $S_c(\cdot): \mathbb{R}^{g \times n} \to \mathbb{R}^{g \times n}$ 如下:

$$S_{c}(E) = U \operatorname{diag}(\operatorname{shrink}(\Sigma, c)) V^{\mathrm{T}}, \forall E \in \mathbb{R}^{g \times n}.$$

设Λ∈ℝ^{g×n} 是与问题(43)中的线性约束相关联的拉格朗日乘子。接下来,我们将详细说明应用 IPSCPRSM 到问题(43)时产生的子问题的显式表述。

X子问题可以通过以下方式显式求解:

$$X^{k+1} = \arg\min_{X} \left\{ \left\| X \right\|_{*} - \left(\overline{\lambda}^{k} \right)^{\mathrm{T}} X + \frac{\beta}{2} \left\| X + \overline{Y}^{k} + \overline{Z}^{k} - N \right\|^{2} + \frac{1}{2} \left\| X - \overline{X}^{k} \right\|_{B}^{2} \right\}$$
$$= S_{\frac{1}{r_{1}}} \left[\frac{1}{r_{1}} \left(\left(r_{1} - \beta \right) \overline{X}^{k} - \beta \overline{Y}^{k} - \beta \overline{Z}^{k} + \overline{\lambda}^{k} + \beta N \right) \right].$$

Y子问题可以通过以下方式求解:

$$Y^{k+1} = \arg\min_{Y} \left\{ \tau \left\| Y \right\|_{1} - \overline{\lambda}^{k} Y + \frac{\beta}{2} \left\| X^{k+1} + Y + \overline{Z}^{k} - N \right\|^{2} + \frac{1}{2} \left\| Y - \overline{Y}^{k} \right\|_{C}^{2} \right\}$$
$$= \operatorname{shrink} \left\{ \frac{1}{r_{2}} \left(\beta N + \left(r_{2} - \beta \right) \overline{Y}^{k} - \beta \overline{Z}^{k} - \beta X^{k+1} + \overline{\lambda}^{k} \right), \frac{\tau}{r_{2}} \right\}.$$

更新拉格朗日乘子 $\Lambda^{k+\frac{1}{2}}$ 通过

$$\Lambda^{k+\frac{1}{2}} = \overline{\Lambda}^k - \alpha \beta \Big(X^{k+1} + Y^{k+1} + \overline{Z}^k - N \Big).$$

Z子问题可以通过以下方式求解:

$$Z^{k+1} = \arg\min_{Z} \left\{ \frac{1}{2\mu} \| P_{\Omega}(Z) \|_{F}^{2} - \left(\lambda^{k+\frac{1}{2}}\right)^{T} Z + \frac{\beta}{2} \| X^{k+1} + Y^{k+1} + Z - N \|^{2} \right\}$$
$$+ \frac{1}{2} \| Z - \overline{Z}^{k} \|_{D}^{2}.$$

其中 $\tilde{\Lambda}^{k} = \Lambda^{k+\frac{1}{2}} - \beta \left(X^{k+1} + Y^{k+1} + \overline{Z}^{k} - N \right)$ 。那么, Z的封闭形式解描述如下:

$$Z_{ij}^{k} = \begin{cases} \frac{\mu\tau r_{3}}{1+\mu\tau r_{3}} \tilde{Z}_{ij}^{k}, \text{ if } (i,j) \in \Omega\\ \tilde{Z}_{ij}^{k}, & \text{otherwise} \end{cases}$$

其中 $\tilde{Z}^{k} = \overline{Z}^{k} + \frac{1}{\tau r_{3}} \tilde{\Lambda}^{k}$ 。 对偶变量 Λ^{k+1} 可以更新为

$$\Lambda^{k+1} = \Lambda^{k+\frac{1}{2}} - \alpha \beta \left(X^{k+1} + Y^{k+1} + Z^{k+1} - N \right).$$

为了展示所提出方法解决(43)问题的可行性和有效性。在这一部分中,我们将使用合成数据测试我们的 算法,正如[20]中所测试的那样。具体来说,设 $N = X^* + Y^*$ 为观测矩阵,其中 $X^* \pi Y^*$ 分别是N的待恢复 的低秩和稀疏分量。 X^* 是作为 $X^* = L^T R$ 生成的,其中 $L \pi R$ 是两个全行秩矩阵, $L \pi R$ 的条目是独立同 分布的高斯值,均值为零,方差为单位方差。对于稀疏分量 Y^* ,其非零条目在区间[-500,500]内均匀随机 生成。索引集 Ω 是随机确定的,有一个预设的样本比率(用 *sr* 表示),定义为 *sr* = ($|\Omega|/gn$),其中 $|\Omega|$ 是 Ω 的 基数。设 rr 为低秩比率(定义为*rr* = g/g),并且设 spr 为稀疏比率(*spr* = $||Y^*||_0/gn$,其中 $||+||_0$ 表示非零条目的 数量)。最后,观测数据被均值为零、标准差 σ = 10⁻³的高斯噪声污染。在实验中,我们设定了 $\tau = 1/\sqrt{g} \pi$ $\mu = \sqrt{g + \sqrt{8g\delta}/10}$ 。对于测试数据,在本节中,我们只考虑样本比率 sr = 80% $\pi g = n \in \{500,1000\}$ 的情况。 相应地,我们分别测试了两个场景,其中 *rr* $\in \{0.05, 0.10\}$ 和三个场景,其中 *spr* $\in \{0.05, 0.10, 0.15\}$ 。

在我们的实验中,所有变量都初始化在原点。整个实验过程中,停止准则是

$$\max\left\{\frac{\left\|X^{k+1}-X^{k}\right\|_{F}}{1+\left\|X^{k}\right\|_{F}},\frac{\left\|Y^{k+1}-Y^{k}\right\|_{F}}{1+\left\|Y^{k}\right\|_{F}}\right\} < tol.$$

其中tol > 0是一个容差参数,我们设定 $tol := 10^{-4}$ 。

在我们的实验中,所有变量都初始化为原点。在实验中,惩罚参数 β 是通过 $\beta = \hat{t} |\Omega| / \|P_{\Omega}(N)\|_{1}$ 来确定的,对于所有测试的方法,其中 \hat{t} 分别对应于 spr = 0.05,0.10 和 0.15 的情况,取值为 0.05,0.10 和 0.20。

为了说明效率,我们将所提出的 IPSCPRSM 与[20]中的 VASALM 和[21]中的 EADMM 进行了比较。 对于 VASALM,我们设定 $\eta = 2.01$ 。对于我们的方法,我们设定 $\alpha = (1+\sqrt{17})/8$,并且对于 IPSCPRSM, 我们设定 $\rho_k = 0.3$ 。数值结果在表 1 中给出,其中 Iter.、Time 和 Tol 分别代表迭代次数、计算时间(以秒 为单位)以及恢复的低秩分量的相对误差(即 $\|X^k - X^*\|_{F}/\|X^*\|_{F}$)。

表 1 中的结果表明,在大多数情况下,IPSCPRSM 在迭代次数和 CPU 时间方面都比 VASALM 更快。 IPSCPRSM 的相对误差也优于 VASALM。初看表 1 中的结果似乎表明,EADMM 在迭代次数最少方面优 于其他两种方法。然而,EADMM 得到的相对误差结果也比其他所有方法的结果更差。事实上,所提出的 IPSCPRSM 算法在所有测试案例中都获得了最佳的相对误差结果。EADMM 在最后阶段的收敛速度可能非 常慢。因此,EADMM 通常需要许多次迭代才能将相对误差结果提高到与 IPSCPRSM 方法相当的水平。

			IPSCPRSM				EADM	М	VASALM		
g	spr	rr	Iter.	Time	Tol	Iter.	Time	Tol	Iter.	Time	Tol
100	0.05	0.05	104	1.8	6.19e-04	147	1.2	6.40e-04	146	1.4	6.34e-04
	0.05	0.10	389	3.4	5.87e-03	505	4.3	6.66e-03	505	4.2	6.65e-03
	0.10	0.05	115	1.0	6.29e-04	149	1.1	6.63e-04	148	1.2	6.58e-04
	0.10	0.10	358	3.4	2.44e-03	477	4.4	3.08e-03	477	4.5	3.07e-03
	0.15	0.05	154	2.0	4.18e-04	170	1.9	5.71e-04	166	1.0	5.65e-04
	0.15	0.10	757	6.6	7.01e-02	790	6.7	7.09e-02	796	6.1	7.08e-02

 Table1. Numerical comparisons of different algorithms for solving the robust PCA

 表 1. 解决鲁棒主成分分析问题的不同算法的数值比较

卖表											
	0.05	0.05	57	10.1	1.01e-04	52	9.0	1.01e-04	58	9.7	1.02e-04
	0.05	0.10	60	10.9	1.20e-04	48	8.5	1.98e-04	71	12.4	1.39e-04
500	0.10	0.05	81	15.6	1.11e-04	50	9.5	1.99e-04	62	10.9	1.89e-04
500	0.10	0.10	86	16.6	1.25e-04	62	10.3	3.66e-04	82	14.7	2.30e-04
	0.15	0.05	79	15.8	1.02e-04	38	6.7	1.82e-04	78	15.0	9.45e-05
	0.15	0.10	104	21.1	1.75e-04	78	13.7	3.27e-04	91	16.7	1.79e-04
	0.05	0.05	66	72.8	5.86e-05	46	46.2	8.85e-05	69	71.5	6.55e-05
	0.05	0.10	70	78.8	6.31e-05	48	50.4	1.26e-04	77	79.4	9.55e-05
1000	0.10	0.05	81	111.2	6.18e-05	50	76.6	1.09e-04	76	119.6	8.25e-05
1000	0.10	0.10	87	105.2	7.32e-05	59	68.4	1.45e-04	87	134.5	9.64e-05
	0.15	0.05	105	171.1	5.94e-05	45	49.3	9.61e-05	94	107.9	6.63e-05
	0.15	0.10	114	205.1	6.56e-05	63	132.2	1.21e-04	118	190.6	7.81e-05

最后,为了进一步观察所提方法的收敛性,在图 1 中,我们绘制了恢复的低秩和稀疏成分的相对误差演 变曲线、 X^k 的秩以及 Y^k 的稀疏性随迭代计数器k的变化情况,其中 $g = 500, spr \in \{0.10, 0.15\}, rr = 0.10$ 。 VASALM 和 IPSCPRSM 具有全局收敛保证,而 EADMM 则没有。图 1 中的结果表明,我们所提出的 IP-SCPRSM 方法能够以更少的迭代次数收敛到期望的最优解,从而比 VASALM 和 EADMM 得到更好的解。



Figure 1. Evolution of relative errors of the recovered low-rank and sparse components (the left four subplots), the rank of X^k and sparsity of Y^k (the right four subplots) with respect to the number of iterations
图 1. 恢复的低秩和稀疏成分的相对误差演变(左侧四个子图)、 X^k 的秩和 Y^k 的稀疏性(右侧四个子图)随迭代次数的 变化

王丽敏 等

5. 结论

在本文中,我们考虑了具有线性约束的可分离凸最小化问题,其目标函数是两个以上不含耦合变量的函数之和。为了解决这个问题,我们考虑了一种带有不定邻近项的惯性邻近严格压缩 Peaceman-Rachford 分裂方法,并且结合惯性项和不定线性化技术,提高了算法的计算效率。此外,我们证明了该方法的全局收敛性和 $o(1/\sqrt{k})$ 收敛速率。最后,我们将所提出的算法应用于鲁棒主成分分析的求解,并验证了其可行性和有效性。

基金项目

国家级大学生创新创业训练计划项目(编号: 202210488007)。

参考文献

- Peng, Y., Ganesh, A., Wright, J., et al. (2012) RASL: Robust Alignment by Sparse and Low-Rank Decomposition for Linearly Correlated Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34, 2233-2246. <u>https://doi.org/10.1109/tpami.2011.282</u>
- [2] Yang, J. and Zhang, Y. (2011) Alternating Direction Algorithms for \$\ell_1\$-Problems in Compressive Sensing. *SIAM Journal on Scientific Computing*, **33**, 250-278. <u>https://doi.org/10.1137/090777761</u>
- [3] Bouwmans, T., Aybat, N.S. and Zahzah, E.H. (2016) Handbook of Robust Low-Rank and Sparse Matrix Decomposition: Applications in Image and Video Processing. CRC Press.
- [4] Boyd, S. (2010) Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends in Machine Learning*, **3**, 1-122. <u>https://doi.org/10.1561/2200000016</u>
- [5] Padcharoen, A., Kumam, P. and Martínez-Moreno, J. (2019) Augmented Lagrangian Method for Tv-11-12 Based Colour Image Restoration. *Journal of Computational and Applied Mathematics*, **354**, 507-519. <u>https://doi.org/10.1016/j.cam.2018.09.053</u>
- [6] Hestenes, M.R. (1969) Multiplier and Gradient Methods. *Journal of Optimization Theory and Applications*, **4**, 303-320. <u>https://doi.org/10.1007/bf00927673</u>
- [7] Powell, M.J.D. (1969) A Method for Nonlinear Constraints in Minimization Problems. In: Fletcher, R., Ed., Optimization, Academic Press, 283-298.
- [8] Douglas, J. and Rachford, H.H. (1956) On the Numerical Solution of Heat Conduction Problems in Two and Three Space Variables. *Transactions of the American Mathematical Society*, 82, 421-439. https://doi.org/10.1090/s0002-9947-1956-0084194-4
- [9] Lions, P.L. and Mercier, B. (1979) Splitting Algorithms for the Sum of Two Nonlinear Operators. SIAM Journal on Numerical Analysis, 16, 964-979. <u>https://doi.org/10.1137/0716071</u>
- [10] Gabay, D. and Mercier, B. (1976) A Dual Algorithm for the Solution of Nonlinear Variational Problems via Finite Element Approximation. *Computers & Mathematics with Applications*, 2, 17-40. <u>https://doi.org/10.1016/0898-1221(76)90003-1</u>
- [11] Glowinski, R. and Marroco, A. (1975) Sur l'approximation, paréléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. *Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique*, 9, 41-76. <u>https://doi.org/10.1051/m2an/197509r200411</u>
- [12] Peaceman, D.W. and Rachford, Jr., H.H. (1955) The Numerical Solution of Parabolic and Elliptic Differential Equations. *Journal of the Society for Industrial and Applied Mathematics*, **3**, 28-41. <u>https://doi.org/10.1137/0103003</u>
- [13] Combettes, P.L. and Wajs, V.R. (2005) Signal Recovery by Proximal Forward-Backward Splitting. *Multiscale Modeling & Simulation*, 4, 1168-1200. <u>https://doi.org/10.1137/050626090</u>
- [14] He, B., Liu, H., Wang, Z. and Yuan, X. (2014) A Strictly Contractive Peaceman—Rachford Splitting Method for Convex Programming. SIAM Journal on Optimization, 24, 1011-1040. <u>https://doi.org/10.1137/13090849x</u>
- [15] Polyak, B.T. (1964) Some Methods of Speeding up the Convergence of Iteration Methods. USSR Computational Mathematics and Mathematical Physics, 4, 1-17. <u>https://doi.org/10.1016/0041-5553(64)90137-5</u>
- [16] Wu, Z., Liu, F. and Li, M. (2018) A Proximal Peaceman-Rachford Splitting Method for Solving the Multi-Block Separable Convex Minimization Problems. *International Journal of Computer Mathematics*, 96, 708-728. https://doi.org/10.1080/00207160.2018.1435864
- [17] Sun, M., Wang, Y. and Liu, J. (2016) Generalized Peaceman-Rachford Splitting Method for Multiple-Block Separable

Convex Programming with Applications to Robust PCA. *Calcolo*, **54**, 77-94. <u>https://doi.org/10.1007/s10092-016-0177-0</u>

- [18] Li, P., Chen, W.G. and Sun, Q.Y. (2023) Inertial Proximal ADMM for Separable Multi-Block Convex Optimizations and Compressive Affine Phase Retrieval. Acta Mathematica Sinica, English Series, 39, 1459-1496. https://doi.org/10.1007/s10114-023-1401-x
- [19] He, B., Tao, M. and Yuan, X. (2012) Alternating Direction Method with Gaussian Back Substitution for Separable Convex Programming. SIAM Journal on Optimization, 22, 313-340. <u>https://doi.org/10.1137/110822347</u>
- [20] Tao, M. and Yuan, X. (2011) Recovering Low-Rank and Sparse Components of Matrices from Incomplete and Noisy Observations. SIAM Journal on Optimization, 21, 57-81. <u>https://doi.org/10.1137/100781894</u>
- [21] Li, S., Li, K., Li, W. and Yang, M. (2024) Evolutionary Alternating Direction Method of Multipliers for Constrained Multi-Objective Optimization with Unknown Constraints. *IEEE Transactions on Evolutionary Computation*. <u>https://doi.org/10.1109/tevc.2024.3425629</u>