Hans 汉斯

惯性对称交替方向乘子算法

田明珠[1]，文　萌[1]，李　军[2]

[1]西安工程大学，理学院，陕西 西安
[2]琼台师范学院，体育学院，海南 海口

摘　要

在本文中提出一种惯性对称交替方向乘子法求解两分块凸极小化优化问题，文章中证明了所提算法收敛到原问题的最优解，最后通过数据分析，验证所提算法的有效性和优越性。

关键词

交替方向乘子法，收敛性，线性约束凸优化问题，惯性对称交替方向乘子法

Inertial Symmetric Alternating Direction Multiplier Algorithm

Mingzhu Tian[1], Meng Wen[1], Jun Li[2]

[1]School of Science, Xi'an Polytechnic University, Xi'an Shaanxi
[2]College of Physical Education, Qiongtai Normal University, Haikou Hainan

Abstract

In this paper, an inertial symmetric alternating direction multiplier method is proposed to solve the two-partite convex miniaturization optimization problem, in which it is proved that the proposed algorithm converges to the optimal solution of the original problem, and finally, the effectiveness and superiority of the proposed algorithm is verified through the data analysis.

Keywords

Alternating Direction Multiplier Method, Convergence, Linearly Constrained Convex Optimization Problems, Inertially Symmetric Alternating Direction Multiplier Method

1. 引言

在本文中主要考虑以下复合优化问题：

$$\min\{f(x)+g(y)\}$$
$$\text{s.t. } Ax+By=b \tag{1}$$

其中 $f:F\to(-\infty,+\infty]$ 和 $g:G\to(-\infty,+\infty]$ 是正则下半连续凸函数，$A:F\to H$ 和 $B:G\to H$ 是非零有界线性算子，F,G,H 是实 Hilbert 空间，$b\in H$。

交替方向乘子法(ADMM)是求解问题(1)的一种常用算法，ADMM 最早是由 Gabay [1]、Glowinski [2] 提出，ADMM 算法的迭代步骤为

$$\begin{cases} x^{k+1}=\arg\min\limits_{x}\left\{f(x)+\left\langle\lambda^k,Ax+By^k-b\right\rangle+\dfrac{\beta}{2}\left\|Ax+By^k-b\right\|^2\right\} \\ y^{k+1}=\arg\min\limits_{y}\left\{g(y)+\left\langle\lambda^k,Ax^{k+1}+By-b\right\rangle+\dfrac{\beta}{2}\left\|Ax^{k+1}+By-b\right\|^2\right\} \\ \lambda^{k+1}=\lambda^k-\beta\left(Ax^{k+1}+By^{k+1}-b\right) \end{cases} \tag{2}$$

Lorenz [3]提出了向前向后分裂算法，随后在文献[4] [5] [6]的基础上结合 Nesterov 加速方法对算法进行改进，求解两分块问题的方法还有惯性 Douglas-Rachford 算子分裂算法[7]、惯性 ADMM [8]、惯性三算子分裂算法[9]等算法。Chen [10]等人通过结合邻近 ADMM [11]算法提出求解带有线性约束可分离凸优化问题的惯性邻近 ADMM 算法，并证明了算法的收敛性，薛中会等人[12]对可分离凸优化问题采用惯性技术，同时引入随机加速的随即变量更新步长，提出了惯性近似松弛交替方向乘子法。

Tseng [13]中提出了交替极小化算法(AMA)，迭代步骤如下

$$\begin{cases} x^{k+1}=\arg\min\limits_{x}\left\{f(x)-\left\langle\lambda^k,Ax\right\rangle\right\} \\ y^{k+1}=\arg\min\limits_{y}\left\{g(y)-\left\langle\lambda^k,By\right\rangle+\dfrac{\beta}{2}\left\|Ax^{k+1}+By-b\right\|^2\right\} \\ \lambda^{k+1}=\lambda^k-\beta\left(Ax^{k+1}+By^{k+1}-b\right) \end{cases} \tag{3}$$

为了加速(3)，Goldstein 在 FISTA 的基础上求解(1)的对偶问题，得到加速 AMA，迭代步骤如下

$$\begin{cases} x^{k+1}=\arg\min\limits_{x}\left\{f(x)-\left\langle\tilde{\lambda}^k,Ax\right\rangle\right\} \\ y^{k+1}=\arg\min\limits_{y}\left\{g(y)-\left\langle\tilde{\lambda}^k,By\right\rangle+\dfrac{\beta}{2}\left\|Ax^{k+1}+By-b\right\|^2\right\} \\ \lambda^{k+1}=\tilde{\lambda}^k-\beta\left(Ax^{k+1}+By^{k+1}-b\right) \\ \alpha_{k+1}=\dfrac{1+\sqrt{1+4\alpha_k^2}}{2} \\ \tilde{\lambda}^{k+1}=\lambda^k+\dfrac{\alpha_k-1}{\alpha_{k+1}}\left(\lambda^k-\lambda^{k-1}\right) \end{cases} \tag{4}$$

其中 Chambolle [14]证明了迭代序列$\{\lambda^k\}$的收敛性。吴双双在[15]中提出了一种惯性的交替极小化算法用来解决问题(1)算法迭代步骤如下

$$\begin{cases} \tilde{\lambda}^k = \lambda^k + \alpha_k\left(\lambda^k - \lambda^{k-1}\right) \\ x^{k+1} = \arg\min_x\left\{f(x) - \langle\tilde{\lambda}^k, Ax\rangle\right\} \\ y^{k+1} = \arg\min_y\left\{g(y) - \langle\tilde{\lambda}^k, By\rangle + \frac{\beta}{2}\left\|Ax^{k+1} + By - b\right\|^2\right\} \\ \lambda^{k+1} = \tilde{\lambda}^k - \beta\left(Ax^{k+1} + By^{k+1} - b\right)n \end{cases} \tag{5}$$

其中$\{\beta\} \subseteq \left(0, \dfrac{2\sigma}{\|A\|^2}\right)$。

问题(1)的特例有许多，例如财务和统计问题[16]，资源分配问题[17]，闭凸集交集上的投影问题[18] [19]，具有约束的全变分图像去噪问题[20] [21] [22]等。下面介绍一下具有约束的全变分图像去噪问题：

$$\min_x \frac{1}{2}\|x - u\|^2 + \lambda\|x\|_{TV} \tag{6}$$
$$\text{s.t. } x \in C$$

(6)式是有约束的，通过加入示性函数$\zeta_C(x)$，由于$\|x\|_{TV} = \varphi(Lx)$，(6)可以表示为

$$\min_x \frac{1}{2}\|x - u\|^2 + \lambda\|x\|_{TV} + \zeta_C(x) \tag{7}$$

设$f(x) = \dfrac{1}{2}\|x - u\|^2 + \zeta_C(x)$，$y = Lx$，因此(7)可以表示为

$$\min_{x,y} f(x) + \lambda\varphi(y) \tag{8}$$
$$\text{s.t. } y = Lx$$

上述问题没有引入对称步，因此本文考虑，在吴[5]的基础上，引入一步对称步骤求解问题(1)，本文证明所提算法的收敛到问题(1)的最优解，通过数据分析，与已有的 ADMM 算法比较，验证新算法的有效性和优越性。

下面给出文章中会用到的一些符号，如表 1 所示。

Table 1. Symbol cross-reference
表 1. 符号对照表

符号	符号名称
$\langle\cdot,\cdot\rangle$	实希尔伯特空间中的内积
$\|\cdot\|$	元素的范数
argmin	最小值点
$\in$	属于
R^n	n 维欧氏空间

2. 预备知识

定义 1. 给定一个正常凸函数 $f : H \to \bar{R}$，它在点 x 处的次微分定义为：

$$\partial f(x) = \left\{ g \in H : f(y) \geq f(x) + \langle g, y - x \rangle, \forall y \in H \right\}$$

其中 $\partial f(x)$ 里的元素称为函数 f 在点 x 处的次梯度，如果 $\partial f(x) \neq \varnothing$，称函数 f 在点 x 处是可微的。

定义 2. 设凸函数 $f : \varepsilon \to (-\infty, +\infty]$，定义 f 在点 $y \in \varepsilon$ 的 Fenchel 共轭函数：

$$f^*(y) := \sup_x \left\{ \langle x, y \rangle - f(x) \right\} = -\inf_x \left\{ f(x) - \langle x, y \rangle \right\}, \quad \forall y \in \varepsilon$$

可知共轭函数 $f^*(y)$ 是闭凸函数。

定义 3. 对任意 $y \in \varepsilon$，令 $f : \varepsilon \to (-\infty, +\infty]$ 是闭凸函数，若满足

$$f(y) \geq f(x) + \langle x^*, y - x \rangle$$

则称 x^* 为 f 在点 $x \in \varepsilon$ 处的次梯度，f 在 x 处次梯度的全体称为 f 在 x 处的次微分，记为 $\partial f(x)$。

3. 惯性对称 ADMM 算法收敛性分析

在本节中，给出具体求解问题(1)的新算法，具体格式如下：

算法 1. 惯性对称 ADMM 算法

设 $\beta \subseteq \left(0, \dfrac{2\sigma}{\|A\|^2} \right)$，$\alpha_k \subseteq (0,1)$

当 $k = 0, 1, 2, \cdots$,

1) $\tilde{y}^k = y^k + \alpha_k \left(y^k - y^{k-1} \right)$

2) $\tilde{\lambda}^k = \lambda^k + \alpha_k \left(\lambda^k - \lambda^{k-1} \right)$

3) $x^{k+1} = \arg\min \left\{ f(x) - \langle \tilde{\lambda}^k, Ax \rangle \right\}$

4) $\lambda^{k+\frac{1}{2}} = \tilde{\lambda}^k + \beta \left(Ax^{k+1} + B\tilde{y}^k - b \right)$

5) $y^{k+1} = \arg\min \left\{ g(y) - \langle \lambda^{k+\frac{1}{2}}, By \rangle + \dfrac{\beta}{2} \left\| Ax^{k+1} + By - b \right\|^2 \right\}$

6) $\lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \beta \left(Ax^{k+1} + By^{k+1} - b \right)$

当满足一定条件时，停止迭代

定理 1 假设拉格朗日函数的鞍点是非空的，且 $\{x^k\}$，$\{y^k\}$，$\{\lambda^k\}$ 是由算法 1 生成的序列，如果 β 满足下列条件，存在 $\gamma > 0$，使得 $\beta \in \left(\gamma, \dfrac{2\sigma}{\|A\|^2} - \gamma \right)$，$\alpha < 1$，$\alpha_k \in [0, \alpha]$ 是非递减序列且满足

$$\sum_{k=0}^{+\infty} \alpha_k \left\| \lambda^k - \lambda^{k-1} \right\|^2 < +\infty$$

则存在拉格朗日函数的一组鞍点 (x^*, y^*, λ^*)，使得

① $x^{k+1} \to x^*$

② $y^{k+1} \to y^*$

③ $\lambda^{k+1} \to \lambda^*$

④ $\lim\limits_{k \to +\infty} f(x^{k+1}) + g(y^{k+1}) = f(x^*) + g(y^*)$

证明

问题(1)的拉格朗日函数为

$$L(x, y, \lambda) = f(x) + g(y) - \lambda^{\mathrm{T}}(Ax + By - b) \tag{9}$$

其中 $\lambda \in R^l$ 是拉格朗日乘子。

问题(1)的增广拉格朗日函数为

$$L(x, y, \lambda) = f(x) + g(y) - \lambda^{\mathrm{T}}(Ax + By - b) + \frac{\beta}{2}\|Ax + By - b\|^2 \tag{10}$$

问题(1)的对偶问题为

$$\max_{\lambda} q(\lambda) = -f^*(A^*\lambda) - g^*(B^*\lambda) + \langle \lambda, b \rangle \tag{11}$$

其中 f^*, g^* 分别是 f 和 g 的 Fenchel 共轭。

应用惯性邻近梯度算法求解(4)

$$\begin{cases} \tilde{\lambda}^k = \lambda^k + \alpha_k(\lambda^k - \lambda^{k-1}) \\ \lambda^{k+1} = prox_{\beta h}\left(\tilde{\lambda}^k - \beta A \nabla f^*(A^*\tilde{\lambda}^k)\right) \end{cases} \tag{12}$$

其中，$h(\lambda) = g^*(B^*\lambda) - \langle \lambda, b \rangle$

设 $x^{k+1} = \nabla f^*(A^*\tilde{\lambda}^k)$，则有

$$A^*\tilde{\lambda}^k \in \partial f(x^{k+1})$$

$$0 \in \partial f(x^{k+1}) - A^*\tilde{\lambda}^k$$

等价于

$$x^{k+1} = \arg\min_x \left\{ f(x) - \langle \tilde{\lambda}^k, Ax \rangle \right\} \tag{13}$$

将 $h(\lambda) = g^*(B^*\lambda) - \langle \lambda, b \rangle$ 代入(12)且有邻近算子定义可得

$$\begin{aligned} \lambda^{k+1} &= prox_{\beta h}\left(\tilde{\lambda}^k - \beta A \nabla f^*(A^*\tilde{\lambda}^k)\right) \\ &= \arg\min_{\lambda}\left\{\beta h(\lambda) + \frac{1}{2}\left\|\lambda - \left(\tilde{\lambda}^k - \beta A \nabla f^*(A^*\tilde{\lambda}^k)\right)\right\|^2\right\} \\ &= \arg\min_{\lambda}\left\{\beta\left(g^*(B^*\lambda) - \langle \lambda, b \rangle\right) + \frac{1}{2}\left\|\lambda - \tilde{\lambda}^k + \beta A \nabla f^*(A^*\tilde{\lambda}^k)\right\|^2\right\} \\ &= \arg\min_{\lambda}\left\{\beta g^*(B^*\lambda) - \beta\langle \lambda, b \rangle + \frac{1}{2}\left\|\lambda - \tilde{\lambda}^k + \beta A \nabla f^*(A^*\tilde{\lambda}^k)\right\|^2\right\} \end{aligned} \tag{14}$$

又 $x^{k+1} = \nabla f^*(A^*\tilde{\lambda}^k)$，所以(14)可以化为

$$\begin{aligned} &\arg\min_{\lambda}\left\{\beta g^*(B^*\lambda) - \beta\langle \lambda, b \rangle + \frac{1}{2}\left\|\lambda - \tilde{\lambda}^k + \beta A \nabla f^*(A^*\tilde{\lambda}^k)\right\|^2\right\} \\ &= \arg\min_{\lambda}\left\{\beta g^*(B^*\lambda) - \beta\langle \lambda, b \rangle + \frac{1}{2}\left\|\lambda - \tilde{\lambda}^k + \beta A x^{k+1}\right\|^2\right\} \end{aligned} \tag{15}$$

上述一阶优化条件可得

$$0 \in \beta \partial g^* \left(B^* \lambda^{k+1} \right) - \beta b + \left(\lambda^{k+1} - \tilde{\lambda}^k + \beta A x^{k+1} \right) \tag{16}$$

设 $y^{k+1} \in \beta \partial g^* \left(B^* \lambda^{k+1} \right)$，可以得到

$$0 \in \partial g \left(y^{k+1} \right) - B^* \lambda^{k+1} \tag{17}$$

对(17)进行变形可以得到

$$0 \in \partial g \left(y^{k+1} \right) - B^* \lambda^{k+1} - B^* \beta \left(A x^{k+1} + B y^{k+1} - b \right) + B^* \beta \left(A x^{k+1} + B y^{k+1} - b \right)$$
$$0 \in \partial g \left(y^{k+1} \right) - B^* \left(\lambda^{k+1} + \beta \left(A x^{k+1} + B y^{k+1} - b \right) \right) + B^* \beta \left(A x^{k+1} + B y^{k+1} - b \right) \tag{18}$$

由算法第 6 步可以得到(17)等价于

$$0 \in \partial g \left(y^{k+1} \right) - B^* \lambda^{k+\frac{1}{2}} + B^* \beta \left(A x^{k+1} + B y^{k+1} - b \right) \tag{19}$$

因此

$$y^{k+1} = \arg \min_y \left\{ g(y) - \left\langle \lambda^{k+\frac{1}{2}}, By \right\rangle + \frac{\beta}{2} \left\| A x^{k+1} + B y - b \right\|^2 \right\} \tag{20}$$

这是新算法的第 5 步。

新算法的第 3 步和第 5 步由一阶优化条件可以得到

$$A^* \tilde{\lambda}^k \in \partial f \left(x^{k+1} \right) \tag{21}$$

$$B^* \lambda^{k+\frac{1}{2}} - B^* \beta \left(A x^{k+1} + B y^{k+1} - b \right) \in \partial g \left(y^{k+1} \right) \tag{22}$$

由 f 和 g 单调有

$$\left\langle x^{k+1} - x^*, A^* \left(\tilde{\lambda}^k - \lambda^* \right) \right\rangle \geq \sigma \left\| x^{k+1} - x^* \right\|^2 \tag{23}$$

$$\left\langle y^{k+1} - y^*, B^* \left(\lambda^{k+\frac{1}{2}} - \lambda^* \right) - \beta B^* \left(A x^{k+1} + B y^{k+1} - b \right) \right\rangle \geq 0 \tag{24}$$

现在将(12)和(13)两式相加可以得到

$$\left\langle x^{k+1} - x^*, A^* \left(\tilde{\lambda}^k - \lambda^* \right) \right\rangle + \left\langle y^{k+1} - y^*, B^* \left(\lambda^{k+\frac{1}{2}} - \lambda^* \right) - \beta B^* \left(A x^{k+1} + B y^{k+1} - b \right) \right\rangle \geq \sigma \left\| x^{k+1} - x^* \right\|^2 \tag{25}$$

整理可得

$$\left\langle x^{k+1} - x^*, A^* \left(\tilde{\lambda}^k - \lambda^* \right) \right\rangle + \left\langle y^{k+1} - y^*, B^* \left(\lambda^{k+\frac{1}{2}} - \lambda^* \right) - \beta B^* \left(A x^{k+1} + B y^{k+1} - b \right) \right\rangle \geq \sigma \left\| x^{k+1} - x^* \right\|^2$$
$$\left\langle \tilde{\lambda}^k - \lambda^*, A \left(x^{k+1} - x^* \right) \right\rangle + \left\langle \lambda^{k+\frac{1}{2}} - \lambda^*, B \left(y^{k+1} - y^* \right) \right\rangle - \left\langle \beta \left(A x^{k+1} + B y^{k+1} - b \right), B \left(y^{k+1} - y^* \right) \right\rangle \geq \sigma \left\| x^{k+1} - x^* \right\|^2 \tag{26}$$

由算法第 4 步、算法第 6 步和 $A x^* + B y^* = b$ 可以得到

$$\left\langle \tilde{\lambda}^k - \lambda^*, A\left(x^{k+1} - x^*\right)\right\rangle + \left\langle \lambda^{k+\frac{1}{2}} - \lambda^*, B\left(y^{k+1} - y^*\right)\right\rangle$$

$$= \left\langle \lambda^{k+\frac{1}{2}} - \beta\left(Ax^{k+1} + B\tilde{y}^k - b\right) - \lambda^*, A\left(x^{k+1} - x^*\right)\right\rangle + \left\langle \lambda^{k+\frac{1}{2}} - \lambda^*, B\left(y^{k+1} - y^*\right)\right\rangle$$

$$= \left\langle \lambda^{k+\frac{1}{2}} - \lambda^* - \beta\left(Ax^{k+1} + B\tilde{y}^k - b\right), A\left(x^{k+1} - x^*\right)\right\rangle + \left\langle \lambda^{k+\frac{1}{2}} - \lambda^*, B\left(y^{k+1} - y^*\right)\right\rangle$$

$$= \left\langle \lambda^{k+\frac{1}{2}} - \lambda^*, A\left(x^{k+1} - x^*\right) + B\left(y^{k+1} - y^*\right)\right\rangle - \beta\left\langle A\left(x^{k+1} - x^*\right), \left(Ax^{k+1} + B\tilde{y}^k - b\right)\right\rangle \qquad (27)$$

$$= \left\langle \lambda^{k+\frac{1}{2}} - \lambda^*, Ax^{k+1} + By^{k+1} - b\right\rangle - \beta\left\langle A\left(x^{k+1} - x^*\right), \left(Ax^{k+1} + B\tilde{y}^k - b\right)\right\rangle$$

$$= \frac{1}{\beta}\left\langle \lambda^{k+\frac{1}{2}} - \lambda^*, \lambda^{k+\frac{1}{2}} - \lambda^{k+1}\right\rangle - \beta\left\langle A\left(x^{k+1} - x^*\right), \left(Ax^{k+1} + B\tilde{y}^k - b\right)\right\rangle$$

将(26)代入(25)可以得到

$$\frac{1}{\beta}\left\langle \lambda^{k+\frac{1}{2}} - \lambda^*, \lambda^{k+\frac{1}{2}} - \lambda^{k+1}\right\rangle - \beta\left\langle A\left(x^{k+1} - x^*\right), \left(Ax^{k+1} + B\tilde{y}^k - b\right)\right\rangle$$

$$- \left\langle \beta\left(Ax^{k+1} + By^{k+1} - b\right), B\left(y^{k+1} - y^*\right)\right\rangle \geq \sigma\left\|x^{k+1} - x^*\right\|^2 \qquad (28)$$

根据余弦等式，可以将(27)化简为

$$\frac{1}{\beta}\left\langle \lambda^{k+\frac{1}{2}} - \lambda^*, \lambda^{k+\frac{1}{2}} - \lambda^{k+1}\right\rangle - \beta\left\langle A\left(x^{k+1} - x^*\right), \left(Ax^{k+1} + B\tilde{y}^k - b\right)\right\rangle$$

$$- \beta\left\langle \left(Ax^{k+1} + By^{k+1} - b\right), B\left(y^{k+1} - y^*\right)\right\rangle \geq \sigma\left\|x^{k+1} - x^*\right\|^2 \qquad (29)$$

现在分开算(28)左边第一项等于

$$\frac{1}{\beta}\left\langle \lambda^{k+\frac{1}{2}} - \lambda^*, \lambda^{k+\frac{1}{2}} - \lambda^{k+1}\right\rangle = -\frac{1}{\beta}\left\langle \lambda^{k+\frac{1}{2}} - \lambda^*, \lambda^{k+1} - \lambda^{k+\frac{1}{2}}\right\rangle$$

$$= -\frac{1}{2\beta}\left(\left\|\lambda^* - \lambda^{k+1}\right\|^2 - \left\|\lambda^{k+\frac{1}{2}} - \lambda^*\right\|^2 - \left\|\lambda^{k+\frac{1}{2}} - \lambda^{k+1}\right\|^2\right) \qquad (30)$$

左边第二项等于

$$-\beta\left\langle A\left(x^{k+1} - x^*\right), \left(Ax^{k+1} + B\tilde{y}^k - b\right)\right\rangle = \beta\left\langle Ax^{k+1} - Ax^*, \left(b - B\tilde{y}^k\right) - Ax^{k+1}\right\rangle$$

$$= \frac{\beta}{2}\left(\left\|Ax^* + B\tilde{y}^k - b\right\|^2 - \left\|Ax^{k+1} - Ax^*\right\|^2 - \left\|Ax^{k+1} + B\tilde{y}^k - b\right\|^2\right) \qquad (31)$$

左边第三项等于

$$-\left\langle \beta\left(Ax^{k+1} + By^{k+1} - b\right), B\left(y^{k+1} - y^*\right)\right\rangle = \beta\left\langle By^{k+1} - \left(b - Ax^{k+1}\right), By^* - By^{k+1}\right\rangle$$

$$= \frac{\beta}{2}\left(\left\|b - Ax^{k+1} - By^*\right\|^2 - \left\|Ax^{k+1} + By^{k+1} - b\right\|^2 - \left\|By^{k+1} - By^*\right\|^2\right) \qquad (32)$$

将(29) (30) (31)相加并代入(28)可以得到

$$-\frac{1}{2\beta}\left(\left\|\lambda^*-\lambda^{k+1}\right\|^2-\left\|\lambda^{k+\frac{1}{2}}-\lambda^*\right\|^2-\left\|\lambda^{k+\frac{1}{2}}-\lambda^{k+1}\right\|^2\right)$$

$$+\frac{\beta}{2}\left(\left\|Ax^*+B\tilde{y}^k-b\right\|^2-\left\|Ax^{k+1}-Ax^*\right\|^2-\left\|Ax^{k+1}+B\tilde{y}^k-b\right\|^2\right) \tag{33}$$

$$+\frac{\beta}{2}\left(\left\|b-Ax^{k+1}-By^*\right\|^2-\left\|Ax^{k+1}+By^{k+1}-b\right\|^2-\left\|By^{k+1}-By^*\right\|^2\right)\geq\sigma\left\|x^{k+1}-x^*\right\|^2$$

继续化简(32)

$$-\frac{1}{2\beta}\left(\left\|\lambda^*-\lambda^{k+1}\right\|^2-\left\|\lambda^{k+\frac{1}{2}}-\lambda^*\right\|^2-\left\|\lambda^{k+\frac{1}{2}}-\lambda^{k+1}\right\|^2\right)-\frac{\beta}{2}\left\|Ax^{k+1}-Ax^*\right\|^2$$

$$+\frac{\beta}{2}\left(\left\|Ax^*-Ax^{k+1}\right\|^2-\left\|Ax^{k+1}+By^{k+1}-b\right\|^2-\left\|By^{k+1}-By^*\right\|^2\right)\geq\sigma\left\|x^{k+1}-x^*\right\|^2 \tag{34}$$

$$-\left(\left\|\lambda^*-\lambda^{k+1}\right\|^2-\left\|\lambda^{k+\frac{1}{2}}-\lambda^*\right\|^2-\left\|\lambda^{k+\frac{1}{2}}-\lambda^{k+1}\right\|^2\right)-\beta^2\left\|Ax^{k+1}-Ax^*\right\|^2$$

$$+\beta^2\left(\left\|Ax^*-Ax^{k+1}\right\|^2-\left\|Ax^{k+1}+By^{k+1}-b\right\|^2-\left\|By^{k+1}-By^*\right\|^2\right)\geq2\beta\sigma\left\|x^{k+1}-x^*\right\|^2$$

$$-\left\|\lambda^*-\lambda^{k+1}\right\|^2+\left\|\lambda^{k+\frac{1}{2}}-\lambda^*\right\|^2+\left\|\lambda^{k+\frac{1}{2}}-\lambda^{k+1}\right\|^2-\beta^2\left\|Ax^{k+1}-Ax^*\right\|^2$$

$$+\beta^2\left\|Ax^*-Ax^{k+1}\right\|^2-\beta^2\left\|Ax^{k+1}+By^{k+1}-b\right\|^2-\beta^2\left\|By^{k+1}-By^*\right\|^2\geq2\beta\sigma\left\|x^{k+1}-x^*\right\|^2$$

$$\left\|\lambda^{k+\frac{1}{2}}-\lambda^*\right\|^2+\left\|\lambda^{k+\frac{1}{2}}-\lambda^{k+1}\right\|^2-\beta^2\left\|Ax^{k+1}-Ax^*\right\|^2+\beta^2\left\|Ax^*-Ax^{k+1}\right\|^2 \tag{35}$$

$$-\beta^2\left\|Ax^{k+1}+By^{k+1}-b\right\|^2-\beta^2\left\|By^{k+1}-By^*\right\|^2-2\beta\sigma\left\|x^{k+1}-x^*\right\|^2\geq\left\|\lambda^*-\lambda^{k+1}\right\|^2$$

$$\left\|\lambda^*-\lambda^{k+1}\right\|^2\leq\left\|\lambda^{k+\frac{1}{2}}-\lambda^*\right\|^2+\left\|\lambda^{k+\frac{1}{2}}-\lambda^{k+1}\right\|^2-\beta^2\left\|Ax^{k+1}-Ax^*\right\|^2+\beta^2\left\|Ax^*-Ax^{k+1}\right\|^2$$

$$-\beta^2\left\|Ax^{k+1}+By^{k+1}-b\right\|^2-\beta^2\left\|By^{k+1}-By^*\right\|^2-2\beta\sigma\left\|x^{k+1}-x^*\right\|^2$$

$$\left\|\lambda^*-\lambda^{k+1}\right\|^2\leq\left\|\lambda^{k+\frac{1}{2}}-\lambda^*\right\|^2+\left\|\lambda^{k+\frac{1}{2}}-\left(\lambda^{k+\frac{1}{2}}-\beta\left(Ax^{k+1}+By^{k+1}-b\right)\right)\right\|^2$$

$$-\beta^2\left\|A\right\|^2\left\|x^{k+1}-x^*\right\|^2+\beta^2\left\|A\right\|^2\left\|x^*-x^{k+1}\right\|^2-\beta^2\left\|Ax^{k+1}+By^{k+1}-b\right\|^2$$

$$-\beta^2\left\|By^{k+1}-By^*\right\|^2-2\beta\sigma\left\|x^{k+1}-x^*\right\|^2 \tag{36}$$

$$\left\|\lambda^*-\lambda^{k+1}\right\|^2\leq\left\|\lambda^{k+\frac{1}{2}}-\lambda^*\right\|^2+\beta^2\left\|Ax^{k+1}+By^{k+1}-b\right\|^2-\beta^2\left\|A\right\|^2\left\|x^{k+1}-x^*\right\|^2$$

$$+\beta^2\left\|A\right\|^2\left\|x^*-x^{k+1}\right\|^2-\beta^2\left\|Ax^{k+1}+By^{k+1}-b\right\|^2-\beta^2\left\|By^{k+1}-By^*\right\|^2$$

$$-2\beta\sigma\left\|x^{k+1}-x^*\right\|^2$$

因此可以得到

$$\left\|\lambda^* - \lambda^{k+1}\right\|^2 \leq \left\|\lambda^{k+\frac{1}{2}} - \lambda^*\right\|^2 - 2\beta\sigma\left\|x^{k+1} - x^*\right\|^2 - \beta^2\left\|By^{k+1} - By^*\right\|^2$$

$$= \left\|\tilde{\lambda}^k + \beta\left(Ax^{k+1} + B\tilde{y}^k - b\right) - \lambda^*\right\|^2 - 2\beta\sigma\left\|x^{k+1} - x^*\right\|^2 - \beta^2\left\|By^{k+1} - By^*\right\|^2 \quad (37)$$

$$= \left\|\tilde{\lambda}^k - \lambda^*\right\|^2 + \beta^2\left\|Ax^{k+1} + B\tilde{y}^k - b\right\|^2 + 2\beta\left\langle\tilde{\lambda}^k - \lambda^*, Ax^{k+1} + B\tilde{y}^k - b\right\rangle$$

$$\quad - 2\beta\sigma\left\|x^{k+1} - x^*\right\|^2 - \beta^2\left\|By^{k+1} - By^*\right\|^2$$

根据算法第二步可以得到

$$\left\|\tilde{\lambda}^k - \lambda^*\right\|^2 + \beta^2\left\|Ax^{k+1} + B\tilde{y}^k - b\right\|^2 + 2\beta\left\langle\tilde{\lambda}^k - \lambda^*, Ax^{k+1} + B\tilde{y}^k - b\right\rangle - 2\beta\sigma\left\|x^{k+1} - x^*\right\|^2 - \beta^2\left\|By^{k+1} - By^*\right\|^2$$

$$= \left\|\lambda^k + \alpha_k\left(\lambda^k - \lambda^{k-1}\right) - \lambda^*\right\|^2 + \beta^2\left\|Ax^{k+1} + B\tilde{y}^k - b\right\|^2$$

$$\quad + 2\beta\left\langle\tilde{\lambda}^k - \lambda^*, Ax^{k+1} + B\tilde{y}^k - b\right\rangle - 2\beta\sigma\left\|x^{k+1} - x^*\right\|^2 - \beta^2\left\|By^{k+1} - By^*\right\|^2$$

$$= \left\|\lambda^k - \lambda^*\right\|^2 + \alpha_k{}^2\left\|\lambda^k - \lambda^{k-1}\right\|^2 + 2\alpha_k\left\langle\lambda^k - \lambda^*, \lambda^k - \lambda^{k-1}\right\rangle + \beta^2\left\|Ax^{k+1} + B\tilde{y}^k - b\right\|^2$$

$$\quad + 2\beta\left\langle\tilde{\lambda}^k - \lambda^*, Ax^{k+1} + B\tilde{y}^k - b\right\rangle - 2\beta\sigma\left\|x^{k+1} - x^*\right\|^2 - \beta^2\left\|By^{k+1} - By^*\right\|^2 \quad (38)$$

$$= \left\|\lambda^k - \lambda^*\right\|^2 + \alpha_k^2\left\|\lambda^k - \lambda^{k-1}\right\|^2 - 2\alpha_k\left\langle\lambda^k - \lambda^*, \lambda^{k-1} - \lambda^k\right\rangle + \beta^2\left\|Ax^{k+1} + B\tilde{y}^k - b\right\|^2$$

$$\quad + 2\beta\left\langle\tilde{\lambda}^k - \lambda^*, Ax^{k+1} + B\tilde{y}^k - b\right\rangle - 2\beta\sigma\left\|x^{k+1} - x^*\right\|^2 - \beta^2\left\|By^{k+1} - By^*\right\|^2$$

$$= \left\|\lambda^k - \lambda^*\right\|^2 + \alpha_k^2\left\|\lambda^k - \lambda^{k-1}\right\|^2 - \left(\left\|\lambda^* - \lambda^{k-1}\right\|^2 - \left\|\lambda^k - \lambda^*\right\|^2 - \left\|\lambda^k - \lambda^{k-1}\right\|^2\right) + \beta^2\left\|Ax^{k+1} + B\tilde{y}^k - b\right\|^2$$

$$\quad + 2\beta\left\langle\tilde{\lambda}^k - \lambda^*, Ax^{k+1} + B\tilde{y}^k - b\right\rangle - 2\beta\sigma\left\|x^{k+1} - x^*\right\|^2 - \beta^2\left\|By^{k+1} - By^*\right\|^2$$

$$= \left\|\lambda^k - \lambda^*\right\|^2 + \alpha_k^2\left\|\lambda^k - \lambda^{k-1}\right\|^2 - \left\|\lambda^* - \lambda^{k-1}\right\|^2 + \left\|\lambda^k - \lambda^*\right\|^2 + \left\|\lambda^k - \lambda^{k-1}\right\|^2 + \beta^2\left\|Ax^{k+1} + B\tilde{y}^k - b\right\|^2$$

$$\quad + 2\left\langle\tilde{\lambda}^k - \lambda^*, \lambda^{k+\frac{1}{2}} - \tilde{\lambda}^k\right\rangle - 2\beta\sigma\left\|x^{k+1} - x^*\right\|^2 - \beta^2\left\|By^{k+1} - By^*\right\|^2$$

$$\beta^2\left\|Ax^{k+1} + B\tilde{y}^k - b\right\|^2$$

$$= \beta^2\left\|Ax^{k+1} + B\tilde{y}^k - Ax^* - By^*\right\|^2 = \beta^2\left\|Ax^{k+1} - Ax^* + B\tilde{y}^k - By^*\right\|^2$$

$$= \beta^2\left\|Ax^{k+1} - Ax^*\right\|^2 + \beta^2\left\|B\tilde{y}^k - By^*\right\|^2 + \beta^2\left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^*\right\rangle$$

$$= \beta^2\left\|Ax^{k+1} - Ax^*\right\|^2 + \beta^2\left\|B\left(y^k + \alpha_k\left(y^k - y^{k-1}\right)\right) - By^*\right\|^2 + \beta^2\left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^*\right\rangle$$

$$= \beta^2\left\|Ax^{k+1} - Ax^*\right\|^2 + \beta^2\left\|By^k - By^* + B\alpha_k\left(y^k - y^{k-1}\right)\right\|^2 + \beta^2\left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^*\right\rangle$$

$$= \beta^2\left\|Ax^{k+1} - Ax^*\right\|^2 + \beta^2\left\|By^k - By^*\right\|^2 + \beta^2\alpha_k{}^2\left\|By^k - By^{k-1}\right\|^2 \quad (39)$$

$$\quad - 2\alpha_k\left\langle By^k - By^*, By^{k-1} - By^k\right\rangle + \beta^2\left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^*\right\rangle$$

$$= \beta^2\left\|Ax^{k+1} - Ax^*\right\|^2 + \beta^2\left\|By^k - By^*\right\|^2 + \beta^2\alpha_k^2\left\|By^k - By^{k-1}\right\|^2 - \alpha_k\left\|By^* - By^{k-1}\right\|^2$$

$$\quad + \alpha_k\left\|By^k - By^*\right\|^2 + \alpha_k\left\|By^k - By^{k-1}\right\|^2 + \beta^2\left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^*\right\rangle$$

将(38)代入(37)

$$\left\| \lambda^k - \lambda^* \right\|^2 + \alpha_k^2 \left\| \lambda^k - \lambda^{k-1} \right\|^2 - \left\| \lambda^* - \lambda^{k-1} \right\|^2 + \left\| \lambda^k - \lambda^* \right\|^2 + \left\| \lambda^k - \lambda^{k-1} \right\|^2$$

$$+ \beta^2 \left\| Ax^{k+1} - Ax^* \right\|^2 + \beta^2 \left\| By^k - By^* \right\|^2 + \beta^2 \alpha_k^2 \left\| By^k - By^{k-1} \right\|^2 - \alpha_k \left\| By^* - By^{k-1} \right\|^2$$

$$+ \alpha_k \left\| By^k - By^* \right\|^2 + \alpha_k \left\| By^k - By^{k-1} \right\|^2 + \beta^2 \left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^* \right\rangle$$

$$+ \alpha_k \left\| By^k - By^* \right\|^2 + \alpha_k \left\| By^k - By^{k-1} \right\|^2 + \beta^2 \left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^* \right\rangle$$

$$= \left\| \lambda^k - \lambda^* \right\|^2 + \alpha_k^2 \left\| \lambda^k - \lambda^{k-1} \right\|^2 - \left\| \lambda^* - \lambda^{k-1} \right\|^2 + \left\| \lambda^k - \lambda^* \right\|^2 + \left\| \lambda^k - \lambda^{k-1} \right\|^2$$

$$+ \beta^2 \left\| Ax^{k+1} - Ax^* \right\|^2 + \beta^2 \left\| By^k - By^* \right\|^2 + \beta^2 \alpha_k^2 \left\| By^k - By^{k-1} \right\|^2 - \alpha_k \left\| By^* - By^{k-1} \right\|^2$$

$$+ \alpha_k \left\| By^k - By^* \right\|^2 + \alpha_k \left\| By^k - By^{k-1} \right\|^2 + \beta^2 \left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^* \right\rangle$$

$$+ \left\| \lambda^* - \lambda^{k+\frac{1}{2}} \right\|^2 - \left\| \tilde{\lambda}^k - \lambda^* \right\|^2 - \left\| \tilde{\lambda}^k - \lambda^{k+\frac{1}{2}} \right\|^2 - 2\beta\sigma \left\| x^{k+1} - x^* \right\|^2 - \beta^2 \left\| By^{k+1} - By^* \right\|^2$$

$$= \left\| \lambda^k - \lambda^* \right\|^2 + \alpha_k^2 \left\| \lambda^k - \lambda^{k-1} \right\|^2 - \left\| \lambda^* - \lambda^{k-1} \right\|^2 + \left\| \lambda^k - \lambda^* \right\|^2 + \left\| \lambda^k - \lambda^{k-1} \right\|^2$$

$$+ \beta^2 \left\| Ax^{k+1} - Ax^* \right\|^2 + \beta^2 \left\| By^k - By^* \right\|^2 + \beta^2 \alpha_k^2 \left\| By^k - By^{k-1} \right\|^2 - \alpha_k \left\| By^* - By^{k-1} \right\|^2$$

$$+ \alpha_k \left\| By^k - By^* \right\|^2 + \alpha_k \left\| By^k - By^{k-1} \right\|^2 + \beta^2 \left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^* \right\rangle$$

$$+ \left\| \lambda^* - \lambda^{k+\frac{1}{2}} \right\|^2 - \left\| \lambda^k + \alpha_k \left(\lambda^k - \lambda^{k-1} \right) - \lambda^* \right\|^2 - \left\| \lambda^k + \alpha_k \left(\lambda^k - \lambda^{k-1} \right) - \lambda^{k+\frac{1}{2}} \right\|^2$$

$$- 2\beta\sigma \left\| x^{k+1} - x^* \right\|^2 - \beta^2 \left\| By^{k+1} - By^* \right\|^2$$

$$= \left\| \lambda^k - \lambda^* \right\|^2 + \alpha_k^2 \left\| \lambda^k - \lambda^{k-1} \right\|^2 - \left\| \lambda^* - \lambda^{k-1} \right\|^2 + \left\| \lambda^k - \lambda^* \right\|^2 + \left\| \lambda^k - \lambda^{k-1} \right\|^2$$

$$+ \beta^2 \left\| Ax^{k+1} - Ax^* \right\|^2 + \beta^2 \left\| By^k - By^* \right\|^2 + \beta^2 \alpha_k^2 \left\| By^k - By^{k-1} \right\|^2 - \alpha_k \left\| By^* - By^{k-1} \right\|^2$$

$$+ \alpha_k \left\| By^k - By^* \right\|^2 + \alpha_k \left\| By^k - By^{k-1} \right\|^2 + \beta^2 \left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^* \right\rangle \tag{40}$$

$$+ \left\| \lambda^* - \lambda^{k+\frac{1}{2}} \right\|^2 - \left\| \lambda^k + \alpha_k \left(\lambda^k - \lambda^{k-1} \right) - \lambda^* \right\|^2 - \left\| \lambda^k + \alpha_k \left(\lambda^k - \lambda^{k-1} \right) - \lambda^{k+\frac{1}{2}} \right\|^2$$

$$- 2\beta\sigma \left\| x^{k+1} - x^* \right\|^2 - \beta^2 \left\| By^{k+1} - By^* \right\|^2$$

因此(36)等价于

$$\left\| \lambda^* - \lambda^{k+1} \right\|^2 \le \left\| \lambda^k - \lambda^* \right\|^2 + \alpha_k^2 \left\| \lambda^k - \lambda^{k-1} \right\|^2 - \left\| \lambda^* - \lambda^{k-1} \right\|^2 + \left\| \lambda^k - \lambda^* \right\|^2 + \left\| \lambda^k - \lambda^{k-1} \right\|^2 + \beta^2 \left\| Ax^{k+1} - Ax^* \right\|^2$$

$$+ \beta^2 \left\| By^k - By^* \right\|^2 + \beta^2 \alpha_k^2 \left\| By^k - By^{k-1} \right\|^2 - \alpha_k \left\| By^* - By^{k-1} \right\|^2 + \alpha_k \left\| By^k - By^* \right\|^2$$

$$+ \alpha_k \left\| By^k - By^{k-1} \right\|^2 + \beta^2 \left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^* \right\rangle + \left\| \lambda^* - \lambda^{k+\frac{1}{2}} \right\|^2 - \left\| \lambda^k - \lambda^* \right\|^2$$

$$- \alpha_k^2 \left\| \lambda^k - \lambda^{k-1} \right\|^2 + \alpha_k \left\| \lambda^k - \lambda^{k-1} \right\|^2 - \alpha_k \left\| \lambda^k - \lambda^* \right\|^2 - \alpha_k \left\| \lambda^k - \lambda^{k-1} \right\|^2 - \left\| \lambda^k - \lambda^{k+\frac{1}{2}} \right\|^2 \tag{41}$$

$$- \alpha_k^2 \left\| \lambda^k - \lambda^{k-1} \right\|^2 + \alpha_k \left\| \lambda^{k-1} - \lambda^{k+\frac{1}{2}} \right\|^2 - \alpha_k \left\| \lambda^k - \lambda^{k-1} \right\|^2 - \alpha_k \left\| \lambda^k - \lambda^{k+\frac{1}{2}} \right\|^2$$

$$- 2\beta\sigma \left\| x^{k+1} - x^* \right\|^2 - \beta^2 \left\| By^{k+1} - By^* \right\|^2$$

$$\left\|\lambda^* - \lambda^{k+1}\right\|^2 \leq \left\|\lambda^k - \lambda^*\right\|^2 - \left\|\lambda^* - \lambda^{k-1}\right\|^2 + \left\|\lambda^k - \lambda^{k-1}\right\|^2 + \left\|\lambda^* - \lambda^{k+\frac{1}{2}}\right\|^2 - \alpha_k\left\|\lambda^k - \lambda^*\right\|^2 - \alpha_k^2\left\|\lambda^k - \lambda^{k-1}\right\|^2$$

$$-\alpha_k\left\|\lambda^k - \lambda^{k-1}\right\|^2 - \left\|\lambda^k - \lambda^{k+\frac{1}{2}}\right\|^2 + \beta^2\left\|Ax^{k+1} - Ax^*\right\|^2 + \beta^2\left\|By^k - By^*\right\|^2 \tag{42}$$

$$+\beta^2\alpha_k^2\left\|By^k - By^{k-1}\right\|^2 - \alpha_k\left\|By^* - By^{k-1}\right\|^2 + \alpha_k\left\|By^k - By^*\right\|^2 + \alpha_k\left\|By^k - By^{k-1}\right\|^2$$

$$-2\beta\sigma\left\|x^{k+1} - x^*\right\|^2 - \beta^2\left\|By^{k+1} - By^*\right\|^2 + \beta^2\left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^*\right\rangle$$

由 $Ax^* + By^* = b$ 以及算法的第 1 步可以得到

$$\beta^2\left\langle Ax^{k+1} - Ax^*, B\tilde{y}^k - By^*\right\rangle = \beta^2\left\langle Ax^{k+1} - \left(b - Bx^*\right), B\tilde{y}^k - By^*\right\rangle = \beta^2\left\langle Bx^* - \left(b - Ax^{k+1}\right), B\tilde{y}^k - By^*\right\rangle$$

$$= \frac{\beta^2}{2}\left(\left\|b - Ax^{k+1} - B\tilde{y}^k\right\|^2 - \left\|Bx^* - \left(b - Ax^{k+1}\right)\right\|^2 - \left\|Bx^* - B\tilde{y}^k\right\|^2\right)$$

$$= \frac{\beta^2}{2}\left\|b - Ax^{k+1} - B\tilde{y}^k\right\|^2 - \frac{\beta^2}{2}\left\|Bx^* - \left(b - Ax^{k+1}\right)\right\|^2 - \frac{\beta^2}{2}\left\|Bx^* - B\tilde{y}^k\right\|^2 \tag{43}$$

$$= \frac{\beta^2}{2}\left\|Ax^* - Ax^{k+1} - \left(B\tilde{y}^k - By^*\right)\right\|^2 - \frac{\beta^2}{2}\left\|Ax^{k+1} - Ax^*\right\|^2 - \frac{\beta^2}{2}\left\|Bx^* - B\tilde{y}^k\right\|^2$$

$$= -\beta^2\left\langle Ax^* - Ax^{k+1}, B\tilde{y}^k - By^*\right\rangle$$

所以(41)可以化为

$$\left\|\lambda^{k+1} - \lambda^*\right\|^2 \leq \left\|\lambda^k - \lambda^*\right\|^2 + \left(1 - \alpha_k^2 - \alpha_k\right)\left\|\lambda^k - \lambda^{k-1}\right\|^2 - \alpha_k\left\|\lambda^k - \lambda^*\right\|^2 + \left\|\lambda^* - \lambda^{k-1}\right\|^2$$

$$+\left\|\lambda^* - \lambda^{k+\frac{1}{2}}\right\|^2 - \left\|\lambda^k - \lambda^{k+\frac{1}{2}}\right\|^2 + \left(\beta^2 - 2\beta\sigma\right)\|A\|^2\left\|x^{k+1} - x^*\right\|^2 - \beta^2\left\|By^{k+1} - By^*\right\|^2 \tag{44}$$

$$+\left(\beta^2 + \alpha_k\right)\left\|By^k - By^*\right\|^2 - \alpha_k\left\|By^* - By^{k-1}\right\|^2 + \left(\beta^2\alpha_k^2 + \alpha_k\right)\left\|By^k - By^{k-1}\right\|^2$$

由算法的第 4 步和第 6 步可得

$$\left\|\lambda^* - \lambda^{k+\frac{1}{2}}\right\|^2 - \left\|\lambda^k - \lambda^{k+\frac{1}{2}}\right\|^2$$

$$= \left\|\lambda^* - \lambda^{k+1} - \beta\left(Ax^{k+1} + By^{k+1} - b\right)\right\|^2 - \left\|\lambda^k - \lambda^{k+1} - \beta\left(Ax^{k+1} + By^{k+1} - b\right)\right\|^2$$

$$= \left\|\lambda^* - \lambda^{k+1}\right\|^2 - \left\|\lambda^k - \lambda^{k+1}\right\|^2 + 2\left\langle\lambda^k - \lambda^*, \lambda^{k+\frac{1}{2}} - \lambda^{k+1}\right\rangle$$

$$= \left\|\lambda^* - \lambda^{k+1}\right\|^2 - \left\|\lambda^k - \lambda^{k+1}\right\|^2 + 2\left\langle\lambda^k - \lambda^*£, \tilde{\lambda}^k + \beta\left(Ax^{k+1} + B\tilde{y}^k - b\right) - \lambda^{k+1}\right\rangle$$

$$= \left\|\lambda^* - \lambda^{k+1}\right\|^2 - \left\|\lambda^k - \lambda^{k+1}\right\|^2 + 2\left\langle\lambda^k - \lambda^*, \lambda^k + \alpha_k\left(\lambda^k - \lambda^{k-1}\right) - \lambda^{k+1} + \beta\left(Ax^{k+1} + B\tilde{y}^k - b\right)\right\rangle$$

$$= \left\|\lambda^* - \lambda^{k+1}\right\|^2 - \left\|\lambda^k - \lambda^{k+1}\right\|^2 - 2\left\langle\lambda^k - \lambda^*, \left[\lambda^{k+1} - \alpha_k\left(\lambda^k - \lambda^{k-1}\right) - \beta\left(Ax^{k+1} + B\tilde{y}^k - b\right)\right] - \lambda^k\right\rangle$$

$$= \left\|\lambda^* - \lambda^{k+1}\right\|^2 - \left\|\lambda^k - \lambda^{k+1}\right\|^2 - \left\|\lambda^k - \lambda^*\right\|^2 + \left\|\lambda^* - \lambda^{k+1} + \alpha_k\left(\lambda^k - \lambda^{k-1}\right) + \beta\left(Ax^{k+1} + B\tilde{y}^k - b\right)\right\|^2$$

$$-\left\|\lambda^k - \lambda^{k+1} + \alpha_k\left(\lambda^k - \lambda^{k-1}\right) + \beta\left(Ax^{k+1} + B\tilde{y}^k - b\right)\right\|^2$$

$$= 2\left\|\lambda^* - \lambda^{k+1}\right\|^2 - 2\left\|\lambda^k - \lambda^{k+1}\right\|^2 - \left\|\lambda^k - \lambda^*\right\|^2 + \alpha_k\left\langle\lambda^* - \lambda^k, \lambda^k - \lambda^{k-1}\right\rangle + 2\beta\left\langle Ax^{k+1} + B\tilde{y}^k - b, \lambda^* - \lambda^k\right\rangle$$

$$= 2\left\|\lambda^* - \lambda^{k+1}\right\|^2 - 2\left\|\lambda^k - \lambda^{k+1}\right\|^2 - \left\|\lambda^k - \lambda^*\right\|^2 + \alpha_k\left\|\lambda^{k-1} - \lambda^*\right\|^2 - \alpha_k\left\|\lambda^k - \lambda^{k-1}\right\|^2 \tag{45}$$

$$-\alpha_k\left\|\lambda^k - \lambda^*\right\|^2 + 2\beta\left\langle Ax^{k+1} + B\tilde{y}^k - b, \lambda^* - \lambda^k\right\rangle$$

代入(1.33)可得

$$\left\| \lambda^{k+1} - \lambda^* \right\|^2 \leq \left(1 - \alpha_k^2 - 2\alpha_k\right)\left\| \lambda^k - \lambda^{k-1} \right\|^2 - 2\alpha_k \left\| \lambda^k - \lambda^* \right\|^2 + \left(1 + \alpha_k\right)\left\| \lambda^* - \lambda^{k-1} \right\|^2$$
$$- 2\left\| \lambda^k - \lambda^{k+1} \right\|^2 + \left(\beta^2 - 2\beta\sigma\right)\|A\|^2 \left\| x^{k+1} - x^* \right\|^2 - \beta^2 \left\| By^{k+1} - By^* \right\|^2 \tag{46}$$
$$+ \left(\beta^2 + \alpha_k\right)\left\| By^k - By^* \right\|^2 - \alpha_k \left\| By^* - By^{k-1} \right\|^2 + \left(\beta^2\alpha_k^2 + \alpha_k\right)\left\| By^k - By^{k-1} \right\|^2$$

由于 $\beta \in \left(0, \dfrac{2\sigma}{\|A\|^2}\right)$

所以(45)可以化为

$$\left\| \lambda^{k+1} - \lambda^* \right\|^2 \leq \left(1 - \alpha_k^2 - 2\alpha_k\right)\left\| \lambda^k - \lambda^{k-1} \right\|^2 - 2\alpha_k \left\| \lambda^k - \lambda^* \right\|^2 + \left(1 + \alpha_k\right)\left\| \lambda^* - \lambda^{k-1} \right\|^2 \tag{47}$$

设 $P^k = \left\| \lambda^k - \lambda^* \right\|^2$，所以(46)可以化为

$$P^{k+1} \leq \left(1 - \alpha_k^2 - 2\alpha_k\right)\left\| \lambda^k - \lambda^{k-1} \right\|^2 - 2\alpha_k P^k + \left(1 + \alpha_k\right)P^{k-1}$$
$$P^{k+1} + 2\alpha_k P^k - \left(1 + \alpha_k\right)P^{k-1} \leq \left(1 - \alpha_k^2 - 2\alpha_k\right)\left\| \lambda^k - \lambda^{k-1} \right\|^2$$
$$P^{k+1} + \alpha_k P^k + \alpha_k P^k - P^{k-1} - \alpha_k P^{k-1} \leq \left(1 - \alpha_k^2 - 2\alpha_k\right)\left\| \lambda^k - \lambda^{k-1} \right\|^2 \tag{48}$$
$$P^{k+1} - P^{k-1} + \alpha_k P^k + \alpha_k \left(P^k - P^{k-1}\right) \leq \left(1 - \alpha_k^2 - 2\alpha_k\right)\left\| \lambda^k - \lambda^{k-1} \right\|^2 \leq \left\| \lambda^k - \lambda^{k-1} \right\|^2$$

对(45)进行整理可得

$$- \left(\beta^2 - 2\beta\sigma\right)\|A\|^2 \left\| x^{k+1} - x^* \right\|^2 + \beta^2 \left\| By^{k+1} - By^* \right\|^2$$
$$\leq \left(1 - \alpha_k^2 - 2\alpha_k\right)\left\| \lambda^k - \lambda^{k-1} \right\|^2 - 2\alpha_k \left\| \lambda^k - \lambda^* \right\|^2 + \left(1 + \alpha_k\right)\left\| \lambda^* - \lambda^{k-1} \right\|^2 - 2\left\| \lambda^k - \lambda^{k+1} \right\|^2 \tag{49}$$
$$- \left\| \lambda^{k+1} - \lambda^* \right\|^2 + \left(\beta^2 + \alpha_k\right)\left\| By^k - By^* \right\|^2 - \alpha_k \left\| By^* - By^{k-1} \right\|^2 + \left(\beta^2\alpha_k^2 + \alpha_k\right)\left\| By^k - By^{k-1} \right\|^2$$

对(48)两边同时求和

$$\left(2\beta\sigma - \beta^2\right)\|A\|^2 \sum_{k=0}^{N} \left\| x^{k+1} - x^* \right\|^2 + \beta^2 \sum_{k=0}^{N} \left\| By^{k+1} - By^* \right\|^2$$
$$\leq \left(-1 - \alpha_k^2 - 2\alpha_k\right)\sum_{k=0}^{N}\left\| \lambda^k - \lambda^{k-1} \right\|^2 - 2\alpha_k \left\| \lambda^N - \lambda^* \right\|^2 + \left(1 + \alpha_k\right)\left\| \lambda^* - \lambda^{N-1} \right\|^2 - \left\| \lambda^{N+1} - \lambda^* \right\|^2 \tag{50}$$
$$+ \left(\beta^2 + \alpha_k\right)\sum_{k=0}^{N}\left\| By^k - By^* \right\|^2 - \alpha_k \sum_{k=0}^{N}\left\| By^* - By^{k-1} \right\|^2 + \left(\beta^2\alpha_k^2 + \alpha_k\right)\sum_{k=0}^{N}\left\| By^k - By^{k-1} \right\|^2$$

上式两边同时取极限可以得到

$$\lim_{k\to+\infty}\left\| x^{k+1} - x^* \right\|^2 = 0, \ \lim_{k\to+\infty}\left\| By^{k+1} - By^* \right\|^2 = 0 \tag{51}$$

因此

$$x^{k+1} \to x^*, \ y^{k+1} \to y^*$$

①②得证。

由于 $L\left(x^{k+1}, y^{k+1}, \lambda^{k+1}\right) \geq L\left(x^*, y^*, \lambda^*\right)$

由(12)式可得

$$f\left(x^{k+1}\right)+g\left(y^{k+1}\right)-\left\langle\lambda^{k+1},Ax^{k+1}+By^{k+1}-b\right\rangle\geq f\left(x^*\right)+g\left(y^*\right)-\left\langle\lambda^*,Ax^*+By^*-b\right\rangle$$

$$f\left(x^{k+1}\right)+g\left(y^{k+1}\right)\geq f\left(x^*\right)+g\left(y^*\right)+\left\langle\lambda^{k+1},Ax^{k+1}+By^{k+1}-b\right\rangle \tag{52}$$

又 $\lim\limits_{k\to+\infty}\left(Ax^{k+1}+By^{k+1}-b\right)=0$，所以(51)可以化为

$$f\left(x^{k+1}\right)+g\left(y^{k+1}\right)\geq f\left(x^*\right)+g\left(y^*\right) \tag{53}$$

由算法 1 第 3 步、第 5 步的一阶优化条件可得

$$0\in\partial f\left(x^{k+1}\right)-A^*\tilde{\lambda}^k$$

$$0\in\partial g\left(y^{k+1}\right)-A^*\tilde{\lambda}^k+\beta B^*\left(Ax^{k+1}+B\tilde{y}^k-b\right) \tag{54}$$

即

$$A^*\tilde{\lambda}^k\in\partial f\left(x^{k+1}\right)$$

$$A^*\tilde{\lambda}^k-\beta B^*\left(Ax^{k+1}+B\tilde{y}^k-b\right)\in\partial g\left(y^{k+1}\right) \tag{55}$$

根据次微分定义可以得到

$$f\left(x^*\right)\geq f\left(x^{k+1}\right)+\left\langle A^*\tilde{\lambda}^k,x^*-x^{k+1}\right\rangle$$

$$g\left(y^*\right)\geq g\left(y^{k+1}\right)+\left\langle B^*\tilde{\lambda}^k-\beta B^*\left(Ax^{k+1}+B\tilde{y}^k-b\right),y^*-y^{k+1}\right\rangle \tag{56}$$

现在将(55)中两式相加可得

$$f\left(x^*\right)+g\left(y^*\right)\geq f\left(x^{k+1}\right)+g\left(y^{k+1}\right)+\left\langle A^*\tilde{\lambda}^k,x^*-x^{k+1}\right\rangle+\left\langle B^*\tilde{\lambda}^k-\beta B^*\left(Ax^{k+1}+B\tilde{y}^k-b\right),y^*-y^{k+1}\right\rangle \tag{57}$$

又因为 $x^{k+1}\to x^*$，$y^{k+1}\to y^*$，所以(56)可以化为

$$f\left(x^*\right)+g\left(y^*\right)\geq f\left(x^{k+1}\right)+g\left(y^{k+1}\right) \tag{58}$$

结合(52)和(57)可以得到

$$\lim\limits_{k\to+\infty}f\left(x^{k+1}\right)+g\left(y^{k+1}\right)=f\left(x^*\right)+g\left(y^*\right) \tag{59}$$

综上，定理 1 得证。

4. 数据分析

在本节中，我们将给出所提算法的数值分析，具体来说，我们比较了 ADMM 算法和惯性对称 ADMM 算法，我们采用以下目标函数和约束条件：

$$\min_{x,y}\left\{(x-1)^2+(y-2)^2\right\}$$

$$\text{s.t. } 0\leq x\leq 3$$

$$1\leq y\leq 4 \tag{60}$$

$$2x+3y=5$$

上式(60)的拉格朗日表达式为

$$L\left(x,y,\lambda\right)=f\left(x\right)+g\left(y\right)+\lambda^{\mathrm{T}}\left(2x+3y-5\right)+\frac{\rho}{2}\left\|2x+3y-5\right\|^2 \tag{61}$$

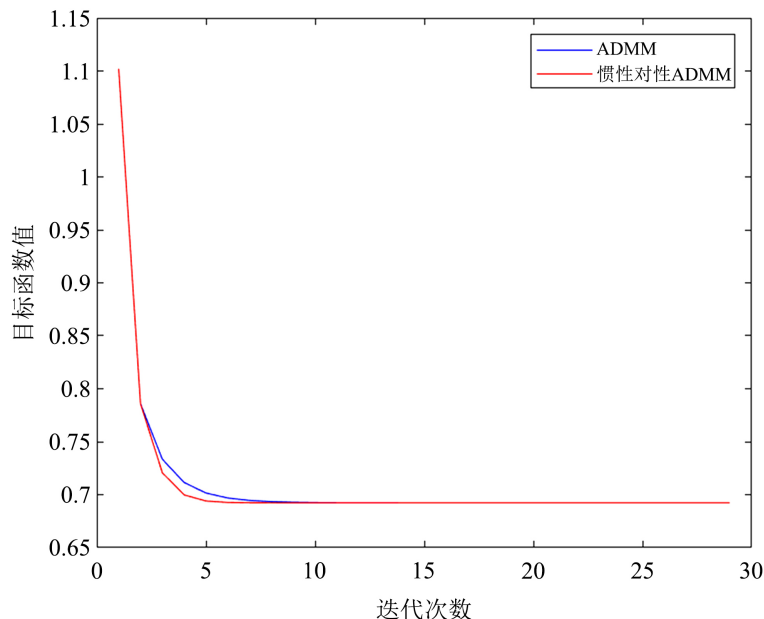在迭代次数为 30 的情况下，ADMM 和惯性对称 ADMM 算法如下图 1 所示。

Figure 1. Plot of results comparing ADMM and inertial ADMM
图 1. ADMM 和惯性 ADMM 比较结果图

注：由图可知本文新算法较 ADMM 在相同迭代次数下其目标函数值收敛速度相对较快。

5. 结论

提出求解目标函数的惯性对称 ADMM 算法，并进行了收敛性分析，最后，数值实验表明，新算法与 ADMM 算法相比较优越。

参考文献

[1] Gabay, D. (1983) Chapter IX Applications of the Method of Multipliers to Variational Inequalities. *Studies in Mathematics and Its Applications*, **15**, 299-231. https://doi.org/10.1016/S0168-2024(08)70034-1

[2] Glowinski, R. and Marroco, A. (1975) Sur l'approximation, par elements finis d'ordre un, et laresolution, par penalisation-dualite, d'une classe de problemes de dirichlet non lineares. *Revue Francaise d'Automatique, Informatique et Recherche Operationelle*, **9**, 41-76. https://doi.org/10.1051/m2an/197509R200411

[3] Lorenz, D.A. and Pock, T. (2015) An Inertial Forward-Backward Algorithm for Monotone Inclusions. *Journal of Mathematical Imaging and Vision*, **51**, 311-325. https://doi.org/10.1007/s10851-014-0523-2

[4] Moudafi, A. and Oliny, M. (2003) Convergence of a Splitting Inertial Proximal Method for Monotone Operators. *Journal of Computational and Applied Mathematics*, **155**, 447-454. https://doi.org/10.1016/S0377-0427(02)00906-8

[5] Alvarez, F. and Attouch, H. (2001) An Inertial Proximal Method for Maximal Monotone Operators via Discretization of a Nonlinear Oscillator with Damping. *Set-Valued Analysis*, **9**, 3-11. https://doi.org/10.1023/A:1011253113155

[6] Attouch, H. and Cabot, A. (2020) Convergence of a Relaxed Inertial Proximal Algorithm for Maximally Monotone Operators. *Mathematical Programming*, **184**, 243-287. https://doi.org/10.1007/s10107-019-01412-0

[7] Bot, R.I., Csetnek, E.R. and Hendrich, C. (2015) Inertial Douglas-Rachford Splitting for Monotone Inclusion Problems. *Applied Mathematics and Computation*, **256**, 472-487. https://doi.org/10.1016/j.amc.2015.01.017

[8] Cui, F.Y., Tang, Y.C. and Yang, Y. (2019) An Inertial Three-Operator Splitting Algorithm with Applications to Image in Painting. *Application of Set-Valued Analysis and Optimization*, **1**, 113-134. https://doi.org/10.23952/asvao.1.2019.2.03

[9] Jia, H., Liu, S. and Dang, Y. (2021) An Inertial Accelerated Algorithm for Solving Split Feasibility Problem with Multiple Output Sets. *Journal of Mathematics*, **2021**, 1-12. https://doi.org/10.1155/2021/6252984

[10] Chen, C.H., Chan, R.H., Ma, S.Q., *et al.* (2015) Inertial Proximal ADMM for Linearly Constrained Separable Convex Optimization. *SIAM Journal on Imaging Sciences*, **8**, 2239-2267. https://doi.org/10.1137/15100463X

[11] Attouch, H. and Soueycatt, M. (2008) Augmented Lagrangian and Proximal Alternating Direction Methods of Multipliers in Hilbert Spaces. Applications to Games, PDE's and Control. *Pacific Journal of Optimization*, **5**, 17-37.

[12] 薛中会, 殷倩雯, 党亚峥. 求解可分离凸优化问题的惯性近似松弛交替方向乘子法[J]. 上海理工大学学报, 2022, 44(2): 204-212.

[13] Tseng, P. (1991) Applications of a Splitting Algorithm to Decomposition in Convex Programming and Variational Inequalities. *SIAM Journal on Control and Optimization*, **29**, 119-138. https://doi.org/10.1137/0329006

[14] Chambolle, A. and Dossal, Ch. (2015) On the Convergence of the Iterates of the "Fast Iterative Shrinkage/Thresholding Algorithm". *Journal of Optimization Theory and Applications*, **166**, 968-982. https://doi.org/10.1007/s10957-015-0746-4

[15] 吴双双, 唐玉超. 一种惯性交替极小化算法及其应用[J]. 南昌大学学报(理科版), 2022, 46(5): 481-491.

[16] Tao, M., Yuan, X.M. and He, B.S. (2011) Solving a Class of Matrix Minimization Problems by Linear Variational Inequality Approaches. *Linear Algebra and Its Applications*, **434**, 2343-2352. https://doi.org/10.1016/j.laa.2010.11.041

[17] Beck, A. and Teboulle, M. (2014) A Fast Dual Proximal Gradient Algorithm for Convex Minimization and Applications. *Operations Research Letters*, **42**, 1-6. https://doi.org/10.1016/j.orl.2013.10.007

[18] Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. (2010) Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends in Machine Learning*, **3**, 1-122. https://doi.org/10.1561/9781601984616

[19] 宗春香, 蔡用, 唐玉超. 有限族非空闭凸集交上的投影算子迭代算法[J]. 南昌大学学报, 2018, 42(4): 327-338.

[20] Beck, A. and Teboulle, M. (2009) Fast Gradient-Based Algorithms for Constrained Total Variation Image Denoising and Deblurring Problems. *IEEE Transactions on Image Processing*, **18**, 2419-2434. https://doi.org/10.1109/TIP.2009.2028250

[21] Ng, M.K., Weiss, P. and Yuan, X.M. (2010) Solving Constrained Total-Variation Image Restoration and Reconstruction Problems via Alternating Direction Methods. *SIAM Journal on Scientific Computing*, **32**, 2710-2736. https://doi.org/10.1137/090774823

[22] Chen, B. and Tang, Y.C. (2019) Iterative Methods for Computing the Resolvent of the Sum of a Maximal Monotone Operator and Composite Operator with Applications. *Mathematical Problems in Engineering*, **2019**, Article ID: 7376263. https://doi.org/10.1155/2019/7376263