DOI：10.13718/j.cnki.xdzk.2018.03.015

带非正定临近项的乘子交替方向法的收敛速率①

王逸云[1]，　欧小庆[2]，　李高西[3]

1. 西南大学 数学与统计学院，重庆 400715；2. 重庆人文科技学院 管理学院，重庆 合川 401524；
3. 重庆工商大学 数学与统计学院，重庆 400067

摘要：研究了带非正定临近正则项的乘子交替方向法(ADMM)的收敛速度. 通过引入松弛因子改进拉格朗日乘子的迭代步长，并在适当的参数条件下建立了带非正定临近正则项的 ADMM 在遍历意义下的收敛速率.

关　键　词：凸规划问题；交替方向法；非正定临近项；收敛速率

中图分类号：O224　　**文献标志码：A**　　**文章编号：**1673-9868(2018)03-0101-08

本文考虑解决以下的一类线性约束凸优化问题

$$\min\{\theta_1(\boldsymbol{x})+\theta_2(\boldsymbol{y}) \mid \boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{y}=\boldsymbol{b},\ \boldsymbol{x}\in\mathcal{X},\ \boldsymbol{y}\in\mathcal{Y}\} \tag{1}$$

其中：$\theta_1(\boldsymbol{x})$：$\mathbb{R}^{n_1}\longrightarrow\mathbb{R}$，$\theta_2(\boldsymbol{y})$：$\mathbb{R}^{n_2}\longrightarrow\mathbb{R}$ 是凸函数(但不一定光滑)；矩阵 $\boldsymbol{A}\in\mathbb{R}^{m\times n_1}$，$\boldsymbol{B}\in\mathbb{R}^{m\times n_2}$；向量 $\boldsymbol{b}\in\mathbb{R}^m$；$\mathcal{X}\subset\mathbb{R}^{n_1}$ 与 $\mathcal{Y}\subset\mathbb{R}^{n_2}$ 是给定的闭凸集. 在本文中，假设问题(1) 的解集是非空的. 将问题(1) 的增广 Lagrange 函数定义为

$$L_\beta^2(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda})=\theta_1(\boldsymbol{x})+\theta_2(\boldsymbol{y})-\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{y}-\boldsymbol{b})+\frac{\beta}{2}\|\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{y}-\boldsymbol{b}\|^2$$

其中：$\boldsymbol{\lambda}\in\mathbb{R}^m$ 为 Lagrange 乘子，$\beta>0$ 为罚因子.

经典的乘子交替方向法(ADMM) 主要是为了解决带约束的凸优化问题，最先在 20 世纪 70 年代中期由文献[1-2] 提出. ADMM 被广泛地应用于各个领域，如机器学习、图像处理、统计学习、无线网络等方面. 原始的 ADMM 形式如下：

$$(\text{ADMM})\begin{cases}\boldsymbol{x}^{k+1}=\operatorname{argmin}\{L_\beta^2(\boldsymbol{x},\boldsymbol{y}^k,\boldsymbol{\lambda}^k)\mid \boldsymbol{x}\in\mathcal{X}\}\\ \boldsymbol{y}^{k+1}=\operatorname{argmin}\{L_\beta^2(\boldsymbol{x}^{k+1},\boldsymbol{y},\boldsymbol{\lambda}^k)\mid \boldsymbol{y}\in\mathcal{Y}\}\\ \boldsymbol{\lambda}^{k+1}=\boldsymbol{\lambda}^k-\beta(\boldsymbol{A}\boldsymbol{x}^{k+1}+\boldsymbol{B}\boldsymbol{y}^{k+1}-\boldsymbol{b})\end{cases} \tag{2}$$

运行 ADMM(2) 时，为了能够求得唯一的点$(\boldsymbol{x}^{k+1},\boldsymbol{y}^{k+1})$，通常通过增加一个平方临近项来正则化(2) 式的子问题. 不失一般性，可以只讨论正则化 ADMM(2) 第二步子问题的情形. 文献[3-4] 将算法进一步改进成下面的形式：

$$(\text{PD-ADMM})\begin{cases}\boldsymbol{x}^{k+1}=\operatorname{argmin}\{L_\beta^2(\boldsymbol{x},\boldsymbol{y}^k,\boldsymbol{\lambda}^k)\mid \boldsymbol{x}\in\mathcal{X}\}\\ \boldsymbol{y}^{k+1}=\operatorname{argmin}\{L_\beta^2(\boldsymbol{x}^{k+1},\boldsymbol{y},\boldsymbol{\lambda}^k)+\frac{1}{2}\|\boldsymbol{y}-\boldsymbol{y}^k\|_{\boldsymbol{D}}^2\mid \boldsymbol{y}\in\mathcal{Y}\}\\ \boldsymbol{\lambda}^{k+1}=\boldsymbol{\lambda}^k-\beta(\boldsymbol{A}\boldsymbol{x}^{k+1}+\boldsymbol{B}\boldsymbol{y}^{k+1}-\boldsymbol{b})\end{cases} \tag{3}$$

---

① 收稿日期：2017-05-30
基金项目：重庆市基础与前沿研究项目(cstc2016jcyjA0239).
作者简介：王逸云(1993-)，女，湖北洪湖人，硕士研究生，主要从事最优化理论、算法及应用的研究.

其中 $\boldsymbol{D} \in \mathbb{R}^{n_2 \times n_2}$ 为正定矩阵. 通常，在很多文献中都会要求矩阵 $\boldsymbol{D}$ 必须要满足正定性，在(3)式中，$\frac{1}{2}\|\boldsymbol{y}-\boldsymbol{y}^k\|_{\boldsymbol{D}}^2$ 就是添加的平方临近项. 另外，也可以考虑对ADMM(2)的第一、二步都进行正则化的情形，可以参考文献[5]. 进一步，文献[6]提出，PD-ADMM(3)中的矩阵 $\boldsymbol{D}$ 可以不满足正定性，若取 $\boldsymbol{D}=\tau\gamma\boldsymbol{I}_{n_2}-\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$，其中 $\gamma>\beta\|\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}\|$，$\tau\in(0,1)$，显然 $\boldsymbol{D}$ 此时不一定正定. 此时，平方正则项 $\frac{1}{2}\|\boldsymbol{y}-\boldsymbol{y}^k\|_{\boldsymbol{D}}^2$ 在目标函数中所占比重减少，$\boldsymbol{y}$ 的迭代步长变大. 特别地，文献[6]说明了当取 $\tau=0.8$ 时为最佳值，并且证明了 $\tau=0.8$ 时算法的收敛性及其在遍历意义下的收敛速率. 另一方面，文献[7]对原始的增广拉格朗日算法(ALM)进行了一些改进，扩大了生成Lagrange乘子 $\boldsymbol{\lambda}$ 的步长，从而加快收敛速度. 与文献[6]类似，文献[7]中还提出了平方临近项中的矩阵 $\boldsymbol{D}$ 可以为非正定的矩阵，即把原始的ALM改进为

$$\text{(PIDP-ALM)}\begin{cases}\boldsymbol{x}^{k+1}=\operatorname{argmin}\{L(\boldsymbol{x},\boldsymbol{\lambda}^k)+\frac{1}{2}\|\boldsymbol{x}-\boldsymbol{x}^k\|_{\boldsymbol{D}_0}^2 \mid \boldsymbol{x}\in\mathcal{X}\} \\ \boldsymbol{\lambda}^{k+1}=\boldsymbol{\lambda}^k-\gamma\beta(\boldsymbol{A}\boldsymbol{x}^{k+1}-\boldsymbol{b}),\ \gamma\in(0,2)\end{cases} \tag{4}$$

其中：$\boldsymbol{D}_0=\boldsymbol{D}-(1-\tau)\beta\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$，$\tau\in(0,1)$，$\boldsymbol{D}\in\mathbb{R}^l$ 是任意给定的正定矩阵；Lagrange函数 $L(\boldsymbol{x},\boldsymbol{\lambda})=\theta(\boldsymbol{x})-\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})$. 文献[1]中确定当系数 τ 和 γ 满足数量关系 $\tau>\frac{2+\gamma}{4}$ 时，算法PIDP-ALM(4)仍然保持收敛性.

受文献[6-7]的启发，本文考虑将文献[6]中带有非正定正则项的ADMM算法进行改进，引入松弛变量 γ，对生成Lagrange乘子 $\boldsymbol{\lambda}$ 的子问题步长进行调整. 本文将讨论如下算法1：

$$\begin{cases}\boldsymbol{x}^{k+1}=\operatorname{argmin}\{L_\beta^2(\boldsymbol{x},\boldsymbol{y}^k,\boldsymbol{\lambda}^k) \mid \boldsymbol{x}\in\mathcal{X}\} \\ \boldsymbol{\lambda}^{k+1}=\boldsymbol{\lambda}^k-\gamma\beta(\boldsymbol{A}\boldsymbol{x}^{k+1}+\boldsymbol{B}\boldsymbol{y}^k-\boldsymbol{b}) \\ \boldsymbol{y}^{k+1}=\operatorname{argmin}\{L_\beta^2(\boldsymbol{x}^{k+1},\boldsymbol{y},\boldsymbol{\lambda}^{k+1})+\frac{1}{2}\|\boldsymbol{y}-\boldsymbol{y}^k\|_{\boldsymbol{D}_0}^2 \mid \boldsymbol{y}\in\mathcal{Y}\}\end{cases} \tag{5}$$

其中：$\boldsymbol{D}_0=\boldsymbol{D}-(1-\tau)\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$，$\boldsymbol{D}\in\mathbb{R}^l$ 是任意给定的正定矩阵，$\tau\in\left[\frac{1+\beta}{2\beta},1\right)$；松弛变量 $\gamma\in\left(0,\frac{2}{1+\beta}\right]$. 另外在本文中设罚参数 $\beta>1$ 且为一固定值. 相对于文献[6]，本文中对生成Lagrange乘子 $\boldsymbol{\lambda}$ 的子问题引入松弛变量 γ，使得对偶变量的迭代步长范围变大，更具有一般性，在理论上有更好的计算性能，并且非正定矩阵 $\boldsymbol{D}_0$ 中的系数 τ，松弛变量 γ 与罚参数 β 之间产生了相互约束的作用.

1　预备知识

本节将回顾一些本文所涉及的主要理论基础知识.

在 $\mathcal{X}\times\mathcal{Y}\times\mathbb{R}^m$ 上定义问题(1)的Lagrange函数

$$L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda})=\theta_1(\boldsymbol{x})+\theta_2(\boldsymbol{y})-\boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{y}-\boldsymbol{b})$$

其中 $\boldsymbol{\lambda}\in\mathbb{R}^m$ 为Lagrange乘子. 点 $((\boldsymbol{x}^*,\boldsymbol{y}^*),\boldsymbol{\lambda}^*)$ 若满足

$$L_{\boldsymbol{\lambda}\in\mathbb{R}^m}(\boldsymbol{x}^*,\boldsymbol{y}^*,\boldsymbol{\lambda})\leqslant L(\boldsymbol{x}^*,\boldsymbol{y}^*,\boldsymbol{\lambda}^*)\leqslant L_{\boldsymbol{x}\in\mathcal{X},\,\boldsymbol{y}\in\mathcal{Y}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda}^*)$$

则被称为Lagrange函数的鞍点. 鞍点问题等价于下面的变分不等式组：

$$\begin{cases}\boldsymbol{x}^*\in\mathcal{X},\ \theta_1(\boldsymbol{x})-\theta_1(\boldsymbol{x}^*)+(\boldsymbol{x}-\boldsymbol{x}^*)^{\mathrm{T}}(-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}^*)\geqslant 0 & \forall\,\boldsymbol{x}\in\mathcal{X} \\ \boldsymbol{y}^*\in\mathcal{Y},\ \theta_2(\boldsymbol{y})-\theta_2(\boldsymbol{y}^*)+(\boldsymbol{y}-\boldsymbol{y}^*)^{\mathrm{T}}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\lambda}^*)\geqslant 0 & \forall\,\boldsymbol{y}\in\mathcal{Y} \\ \boldsymbol{\lambda}^*\in\mathbb{R}^m,\ (\boldsymbol{\lambda}-\boldsymbol{\lambda}^*)^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x}^*+\boldsymbol{B}\boldsymbol{y}^*-\boldsymbol{b})\geqslant 0 & \forall\,\boldsymbol{\lambda}\in\mathbb{R}^m\end{cases} \tag{6}$$

如果记

$$\Omega = \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^m$$

以及

$$\boldsymbol{u} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \qquad \theta(\boldsymbol{u}) = \theta_1(\boldsymbol{x}) + \theta_2(\boldsymbol{y}),\ \boldsymbol{\omega} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ \boldsymbol{\lambda} \end{pmatrix} \qquad F(\boldsymbol{\omega}) = \begin{pmatrix} -\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda} \\ -\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\lambda} \\ \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{y} - \boldsymbol{b} \end{pmatrix} \tag{7}$$

问题(6) 可以写成变分不等式

$$\boldsymbol{\omega}^* \in \Omega,\ \theta(\boldsymbol{u}) - \theta(\boldsymbol{u}^*) + (\boldsymbol{\omega} - \boldsymbol{\omega}^*)^{\mathrm{T}} F(\boldsymbol{\omega}^*) \geqslant 0 \qquad \forall \boldsymbol{\omega} \in \Omega \tag{8}$$

将问题(8) 的解集记作 Ω^*. 定义在(7) 式中的算子 F 是一个带有斜对称矩阵的仿射单调算子，并且对于任意的 $\boldsymbol{\omega}, \overline{\boldsymbol{\omega}} \in \Omega$ 有

$$(\boldsymbol{\omega} - \overline{\boldsymbol{\omega}})^{\mathrm{T}}(F(\boldsymbol{\omega}) - F(\overline{\boldsymbol{\omega}})) =$$

$$(\boldsymbol{\omega} - \overline{\boldsymbol{\omega}})^{\mathrm{T}} \begin{pmatrix} -\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\lambda} - \overline{\boldsymbol{\lambda}}) \\ -\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{\lambda} - \overline{\boldsymbol{\lambda}}) \\ \boldsymbol{A}(\boldsymbol{x} - \overline{\boldsymbol{x}}) + \boldsymbol{B}(\boldsymbol{y} - \overline{\boldsymbol{y}}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{x} - \overline{\boldsymbol{x}} \\ \boldsymbol{y} - \overline{\boldsymbol{y}} \\ \boldsymbol{\lambda} - \overline{\boldsymbol{\lambda}} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} -\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\lambda} - \overline{\boldsymbol{\lambda}}) \\ -\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{\lambda} - \overline{\boldsymbol{\lambda}}) \\ \boldsymbol{A}(\boldsymbol{x} - \overline{\boldsymbol{x}}) + \boldsymbol{B}(\boldsymbol{y} - \overline{\boldsymbol{y}}) \end{pmatrix} =$$

$$-(\boldsymbol{x} - \overline{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\lambda} - \overline{\boldsymbol{\lambda}}) - (\boldsymbol{y} - \overline{\boldsymbol{y}})^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{\lambda} - \overline{\boldsymbol{\lambda}}) + (\boldsymbol{\lambda} - \overline{\boldsymbol{\lambda}})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x} - \overline{\boldsymbol{x}}) + (\boldsymbol{\lambda} - \overline{\boldsymbol{\lambda}})^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{y} - \overline{\boldsymbol{y}}) = 0$$

引理 1[6]　设 $\mathcal{X} \subset \mathbb{R}^n$ 是一个闭凸集合，$\theta(\boldsymbol{x})$ 和 $f(\boldsymbol{x})$ 是定义在 $\mathcal{X}$ 上的凸函数. 若函数 f 是可微的，且问题 $\min\{\theta(\boldsymbol{x}) + f(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathcal{X}\}$ 的解集是非空的. 则有

$$\boldsymbol{x}^* \in \operatorname{argmin}\{\theta(\boldsymbol{x}) + f(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathcal{X}\}$$

当且仅当存在 $\boldsymbol{x}^* \in \mathcal{X}$, $\theta(\boldsymbol{x}) - \theta(\boldsymbol{x}^*) + (\boldsymbol{x} - \boldsymbol{x}^*)^{\mathrm{T}} \nabla f(\boldsymbol{x}^*) \geqslant 0$，$\forall \boldsymbol{x} \in \mathcal{X}$.

2　算法 1 的预估校正形式

本节将用预估校正方法解释算法 1，这样的形式有助于在后文中对算法的收敛速率进行分析. 首先定义

$$\boldsymbol{v} = \begin{pmatrix} \boldsymbol{y} \\ \boldsymbol{\lambda} \end{pmatrix},\ \mathcal{V} = \mathcal{Y} \times \mathbb{R}^m,\ \mathcal{V}^* = \{(\boldsymbol{y}^*, \boldsymbol{\lambda}^*) \mid (\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\lambda}^*) \in \Omega^*\}$$

由引理 1 知算法 1 的子问题可以分别写作

$$\boldsymbol{x}^{k+1} \in \mathcal{X}$$

$$\theta_1(\boldsymbol{x}) - \theta_1(\boldsymbol{x}^{k+1}) + (\boldsymbol{x} - \boldsymbol{x}^{k+1})^{\mathrm{T}}[-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}^k + \beta\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x}^{k+1} + \boldsymbol{B}\boldsymbol{y}^k - \boldsymbol{b})] \geqslant 0,\ \forall \boldsymbol{x} \in \mathcal{X} \tag{9}$$

和

$$\boldsymbol{y}^{k+1} \in \mathcal{Y}$$

$$\theta_2(\boldsymbol{y}) - \theta_2(\boldsymbol{y}^{k+1}) + (\boldsymbol{y} - \boldsymbol{y}^{k+1})^{\mathrm{T}}[-\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{\lambda}^{k+1} - \beta(\boldsymbol{A}\boldsymbol{x}^{k+1} + \boldsymbol{B}\boldsymbol{y}^{k+1} - \boldsymbol{b})) + \boldsymbol{D}_0(\boldsymbol{y}^{k+1} - \boldsymbol{y}^k)] \geqslant 0,\ \forall \boldsymbol{y} \in \mathcal{Y} \tag{10}$$

对于给定的$(\boldsymbol{y}^k, \boldsymbol{\lambda}^k)$算法 1 生成了$(\boldsymbol{x}^{k+1}, \boldsymbol{y}^{k+1})$，分别定义 $\widetilde{\boldsymbol{x}}^k = \boldsymbol{x}^{k+1}$，$\widetilde{\boldsymbol{y}}^k = \boldsymbol{y}^{k+1}$，以及辅助变量

$$\widetilde{\boldsymbol{\lambda}}^k := \boldsymbol{\lambda}^k - \beta(\boldsymbol{A}\boldsymbol{x}^{k+1} + \boldsymbol{B}\boldsymbol{y}^k - \boldsymbol{b}) \tag{11}$$

因此不等式(9) 可以写作

$$\widetilde{\boldsymbol{x}}^k \in \mathcal{X},\ \theta_1(\boldsymbol{x}) - \theta_1(\widetilde{\boldsymbol{x}}^k) + (\boldsymbol{x} - \widetilde{\boldsymbol{x}}^k)^{\mathrm{T}}(-\boldsymbol{A}^{\mathrm{T}}\widetilde{\boldsymbol{\lambda}}^k) \geqslant 0,\ \forall \boldsymbol{x} \in \mathcal{X} \tag{12}$$

另外，由 $\boldsymbol{\lambda}^{k+1}$，$\widetilde{\boldsymbol{\lambda}}^k$ 和 $\boldsymbol{y}^{k+1}$ 的定义，有

$$\boldsymbol{\lambda}^{k+1} - \beta(\boldsymbol{A}\boldsymbol{x}^{k+1} + \boldsymbol{B}\boldsymbol{y}^{k+1} - \boldsymbol{b}) =$$

$$\boldsymbol{\lambda}^k - \gamma\beta(\boldsymbol{A}\boldsymbol{x}^{k+1} + \boldsymbol{B}\boldsymbol{y}^k - \boldsymbol{b}) - \beta(\boldsymbol{A}\boldsymbol{x}^{k+1} + \boldsymbol{B}\boldsymbol{y}^k - \boldsymbol{b}) - \beta\boldsymbol{B}(\boldsymbol{y}^{k+1} - \boldsymbol{y}^k) =$$

$$\boldsymbol{\lambda}^k - \beta(\boldsymbol{A}\boldsymbol{x}^{k+1} + \boldsymbol{B}\boldsymbol{y}^k - b) - \gamma(\boldsymbol{\lambda}^k - \widetilde{\boldsymbol{\lambda}}^k) - \beta\boldsymbol{B}(\widetilde{\boldsymbol{y}}^k - \boldsymbol{y}^k) =$$

$$\tilde{\boldsymbol{\lambda}}^k-\gamma(\boldsymbol{\lambda}^k-\tilde{\boldsymbol{\lambda}}^k)-\beta\boldsymbol{B}(\tilde{\boldsymbol{y}}^k-\boldsymbol{y}^k)$$

则不等式(10)可以写为

$$\tilde{\boldsymbol{y}}^k\in\mathscr{Y}$$

$$\theta_2(\boldsymbol{y})-\theta_2(\tilde{\boldsymbol{y}}^k)+(\boldsymbol{y}-\tilde{\boldsymbol{y}}^k)^{\mathrm{T}}[-\boldsymbol{B}^{\mathrm{T}}\tilde{\boldsymbol{\lambda}}^k+\gamma\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{\lambda}^k-\tilde{\boldsymbol{\lambda}}^k)+(\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}+\boldsymbol{D}_0)(\tilde{\boldsymbol{y}}^k-\boldsymbol{y}^k)]\geqslant0,\ \forall\,\boldsymbol{y}\in\mathscr{Y}\tag{13}$$

由 $\tilde{\boldsymbol{\lambda}}^k$ 的定义,有

$$(\boldsymbol{A}\tilde{\boldsymbol{x}}^k+\boldsymbol{B}\tilde{\boldsymbol{y}}^k-\boldsymbol{b})-\boldsymbol{B}(\tilde{\boldsymbol{y}}^k-\boldsymbol{y}^k)+\frac{1}{\beta}(\tilde{\boldsymbol{\lambda}}^k-\boldsymbol{\lambda}^k)=0\tag{14}$$

联立(12)-(14)式,可以得到:

$$\theta(\boldsymbol{u})-\theta(\tilde{\boldsymbol{u}}^k)+\begin{pmatrix}\boldsymbol{x}-\tilde{\boldsymbol{x}}^k\\ \boldsymbol{y}-\tilde{\boldsymbol{y}}^k\\ \boldsymbol{\lambda}-\tilde{\boldsymbol{\lambda}}^k\end{pmatrix}^{\mathrm{T}}\left\{\begin{pmatrix}-\boldsymbol{A}^{\mathrm{T}}\tilde{\boldsymbol{\lambda}}^k\\ -\boldsymbol{B}^{\mathrm{T}}\tilde{\boldsymbol{\lambda}}^k\\ \boldsymbol{A}\tilde{\boldsymbol{x}}^k+\boldsymbol{B}\tilde{\boldsymbol{y}}^k-\boldsymbol{b}\end{pmatrix}\right.$$
$$\left.-\begin{pmatrix}(\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}+\boldsymbol{D}_0)(\boldsymbol{y}^k-\tilde{\boldsymbol{y}}^k)-\gamma\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{\lambda}^k-\tilde{\boldsymbol{\lambda}}^k)\\ -\boldsymbol{B}(\boldsymbol{y}^k-\tilde{\boldsymbol{y}}^k)+\frac{1}{\beta}(\boldsymbol{\lambda}^k-\tilde{\boldsymbol{\lambda}}^k)\end{pmatrix}\right\}\geqslant0\qquad\forall\,\boldsymbol{\omega}\in\Omega\tag{15}$$

由(7)式中的定义,(15)式等价于变分不等式

$$\tilde{\boldsymbol{\omega}}^k\in\Omega,\ \theta(\boldsymbol{u})-\theta(\tilde{\boldsymbol{u}}^k)+(\boldsymbol{\omega}-\tilde{\boldsymbol{\omega}}^k)^{\mathrm{T}}F(\tilde{\boldsymbol{\omega}}^k)\geqslant(\boldsymbol{v}-\tilde{\boldsymbol{v}}^k)^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{v}^k-\tilde{\boldsymbol{v}}^k)\qquad\forall\,\boldsymbol{\omega}\in\Omega\tag{16}$$

其中

$$\boldsymbol{Q}=\begin{pmatrix}\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}+\boldsymbol{D}_0 & -\gamma\boldsymbol{B}^{\mathrm{T}}\\ -\boldsymbol{B} & \frac{1}{\beta}\boldsymbol{I}_m\end{pmatrix}=\begin{pmatrix}\tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}+\boldsymbol{D} & -\gamma\boldsymbol{B}^{\mathrm{T}}\\ -\boldsymbol{B} & \frac{1}{\beta}\boldsymbol{I}_m\end{pmatrix}\tag{17}$$

另一方面,由 $\boldsymbol{\lambda}^{k+1}$,$\tilde{\boldsymbol{\lambda}}^k$ 和 $\boldsymbol{y}^{k+1}$ 的定义,有

$$\boldsymbol{y}^{k+1}=\boldsymbol{y}^k-(\boldsymbol{y}^k-\tilde{\boldsymbol{y}}^k)$$

$$\boldsymbol{\lambda}^{k+1}=\boldsymbol{\lambda}^k-\gamma(\boldsymbol{A}\boldsymbol{x}^{k+1}+\boldsymbol{B}\boldsymbol{y}^k-\boldsymbol{b})=\boldsymbol{\lambda}^k-\gamma(\boldsymbol{\lambda}^k-\tilde{\boldsymbol{\lambda}}^k)$$

因此,有

$$\boldsymbol{v}^{k+1}=\boldsymbol{v}^k-\boldsymbol{M}(\boldsymbol{v}^k-\tilde{\boldsymbol{v}}^k)\tag{18}$$

其中

$$\boldsymbol{M}=\begin{pmatrix}\boldsymbol{I}_{n_2} & 0\\ 0 & \gamma\boldsymbol{I}_m\end{pmatrix}\tag{19}$$

由上面的内容可知,算法 1 从理论上可以看作是分为两个阶段的. 第一步由变分不等式(16)生成一个预测点($\tilde{\boldsymbol{y}}^k$, $\tilde{\boldsymbol{\lambda}}^k$);第二步再通过(18)式对($\tilde{\boldsymbol{y}}^k$, $\tilde{\boldsymbol{\lambda}}^k$)进行校正,生成新的迭代点($\boldsymbol{y}^{k+1}$, $\boldsymbol{\lambda}^{k+1}$).

3 遍历意义下的收敛速率

本节将分析算法 1 在遍历意义下的收敛速率. 首先给出以下的几个引理:

引理 2 设定义在(1)式中的矩阵 $\boldsymbol{B}$ 是列满秩的,罚参数 $\beta>1$. 定义矩阵

$$\boldsymbol{H}=\begin{pmatrix}\tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}+\boldsymbol{D} & -\boldsymbol{B}^{\mathrm{T}}\\ -\boldsymbol{B} & \frac{1}{\gamma\beta}\boldsymbol{I}_m\end{pmatrix}$$

则

$$\boldsymbol{HM}=\boldsymbol{Q}\tag{20}$$

另外，当 $\tau \in \left[\frac{1+\beta}{2\beta}, 1\right)$，$\gamma \in \left(0, \frac{2}{1+\beta}\right]$时，有 $\boldsymbol{H} \succ \boldsymbol{0}$.

证　由(17)式和(19)式，有

$$\boldsymbol{HM} = \begin{pmatrix} \tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} + \boldsymbol{D} & -\boldsymbol{B}^{\mathrm{T}} \\ -\boldsymbol{B} & \frac{1}{\gamma\beta}\boldsymbol{I}_m \end{pmatrix} \begin{pmatrix} \boldsymbol{I}_{n_2} & \boldsymbol{0} \\ \boldsymbol{0} & \gamma\boldsymbol{I}_m \end{pmatrix} = \begin{pmatrix} \tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} + \boldsymbol{D} & -\gamma\boldsymbol{B}^{\mathrm{T}} \\ -\boldsymbol{B} & \frac{1}{\beta}\boldsymbol{I}_m \end{pmatrix} = \boldsymbol{Q}$$

等式(20)立即证得.

另一方面，

$$\boldsymbol{H} = \begin{pmatrix} \tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} + \boldsymbol{D} & -\boldsymbol{B}^{\mathrm{T}} \\ -\boldsymbol{B} & \frac{1}{\gamma\beta}\boldsymbol{I}_m \end{pmatrix} = \begin{pmatrix} \tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} & -\boldsymbol{B}^{\mathrm{T}} \\ -\boldsymbol{B} & \frac{1}{\gamma\beta}\boldsymbol{I}_m \end{pmatrix} + \begin{pmatrix} \boldsymbol{D} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}$$

并且 $\boldsymbol{D}$ 为正定矩阵. 因此，若矩阵$\begin{pmatrix} \tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} & -\boldsymbol{B}^{\mathrm{T}} \\ -\boldsymbol{B} & \frac{1}{\gamma\beta}\boldsymbol{I}_m \end{pmatrix}$为正定矩阵，则 H 一定正定. 矩阵

$$\begin{pmatrix} \tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} & -\boldsymbol{B}^{\mathrm{T}} \\ -\boldsymbol{B} & \frac{1}{\gamma\beta}\boldsymbol{I}_m \end{pmatrix} = \begin{pmatrix} \boldsymbol{B}^{\mathrm{T}} & 0 \\ \boldsymbol{0} & \boldsymbol{I}_m \end{pmatrix} \begin{pmatrix} \tau\beta\boldsymbol{I}_m & -\boldsymbol{I}_m \\ -\boldsymbol{I}_m & \frac{1}{\gamma\beta}\boldsymbol{I}_m \end{pmatrix} \begin{pmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_m \end{pmatrix}$$

由假设矩阵 $\boldsymbol{B}$ 是列满秩的，因此只需

$$\begin{pmatrix} \tau\beta\boldsymbol{I}_m & -\boldsymbol{I}_m \\ -\boldsymbol{I}_m & \frac{1}{\gamma\beta}\boldsymbol{I}_m \end{pmatrix} \succ \boldsymbol{0}$$

即

$$\frac{1}{\gamma}(\tau - \gamma) > 0$$

由于 $\gamma \in \left(0, \frac{2}{1+\beta}\right]$，只需 $\tau > \gamma$. 对于任意固定的罚参数 $\beta > 1$，当 $\tau \in \left[\frac{1+\beta}{2\beta}, 1\right)$，$\gamma \in \left(0, \frac{2}{1+\beta}\right]$时，一定有 $\tau > \gamma$，因此有 $\boldsymbol{H} \succ \boldsymbol{0}$.

定义对称矩阵 $\boldsymbol{G} = \boldsymbol{Q}^{\mathrm{T}} + \boldsymbol{Q} - \boldsymbol{M}^{\mathrm{T}}\boldsymbol{HM}$，因为 $\boldsymbol{HM} = \boldsymbol{Q}$，$\boldsymbol{M}^{\mathrm{T}}\boldsymbol{HM} = \boldsymbol{M}^{\mathrm{T}}\boldsymbol{Q}$，有

$$\boldsymbol{M}^{\mathrm{T}}\boldsymbol{HM} = \begin{pmatrix} \boldsymbol{I}_{n_2} & \boldsymbol{0} \\ \boldsymbol{0} & \gamma\boldsymbol{I}_m \end{pmatrix} \begin{pmatrix} \tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} + \boldsymbol{D} & -\gamma\boldsymbol{B}^{\mathrm{T}} \\ -\boldsymbol{B} & \frac{1}{\beta}\boldsymbol{I}_m \end{pmatrix} = \begin{pmatrix} \tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} + \boldsymbol{D} & -\gamma\boldsymbol{B}^{\mathrm{T}} \\ -\gamma\boldsymbol{B} & \frac{\gamma}{\beta}\boldsymbol{I}_m \end{pmatrix}$$

所以

$$\boldsymbol{G} = \boldsymbol{Q}^{\mathrm{T}} + \boldsymbol{Q} - \boldsymbol{M}^{\mathrm{T}}\boldsymbol{HM} = \begin{pmatrix} 2(\tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} + \boldsymbol{D}) & -\boldsymbol{B}^{\mathrm{T}} - \gamma\boldsymbol{B}^{\mathrm{T}} \\ -\boldsymbol{B} - \gamma\boldsymbol{B} & \frac{2}{\beta}\boldsymbol{I}_m \end{pmatrix} - \begin{pmatrix} \tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} + \boldsymbol{D} & -\gamma\boldsymbol{B}^{\mathrm{T}} \\ -\gamma\boldsymbol{B} & \frac{\gamma}{\beta}\boldsymbol{I}_m \end{pmatrix} = \begin{pmatrix} \tau\beta\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B} + \boldsymbol{D} & -\boldsymbol{B}^{\mathrm{T}} \\ -\boldsymbol{B} & \frac{2-\gamma}{\beta}\boldsymbol{I}_m \end{pmatrix} \tag{21}$$

引理 3 设($\boldsymbol{\omega}^k$)为算法1所生成的序列，$\widetilde{\boldsymbol{\omega}}^k$ 由(11)式定义，则有 $\widetilde{\boldsymbol{\omega}}^k \in \Omega$ 以及

$$\theta(\boldsymbol{u})-\theta(\widetilde{\boldsymbol{u}}^k)+(\boldsymbol{\omega}-\widetilde{\boldsymbol{\omega}}^k)^{\mathrm{T}}F(\boldsymbol{\omega}) \geqslant$$

$$\frac{1}{2}(\|\boldsymbol{v}-\boldsymbol{v}^{k+1}\|_{\boldsymbol{H}}^2-\|\boldsymbol{v}-\boldsymbol{v}^k\|_{\boldsymbol{H}}^2)+\frac{1}{2}(\boldsymbol{v}^k-\widetilde{\boldsymbol{v}}^k)^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{v}^k-\widetilde{\boldsymbol{v}}^k) \qquad \forall \boldsymbol{\omega} \in \Omega \tag{22}$$

其中 $\boldsymbol{G}$ 在(21)式中被定义.

由于证明过程与文献[6]中的引理2类似，故而省略.

如果矩阵 $\boldsymbol{G}$ 是正定的，则由(22)式很容易推导出算法1的收敛性以及在遍历意义下的收敛速率. 本文需要处理的难点在于定义在(21)式中的矩阵 $\boldsymbol{G}$ 不一定是正定的，为了解决这个问题，考虑寻找$(\boldsymbol{v}^k-\widetilde{\boldsymbol{v}}^k)^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{v}^k-\widetilde{\boldsymbol{v}}^k)$ 的一个下界.

引理 4 设($\boldsymbol{\omega}^k$)为算法1所生成的序列，$\widetilde{\boldsymbol{\omega}}^k$ 由(11)式定义，则有

$$(\boldsymbol{v}^k-\widetilde{\boldsymbol{v}}^k)^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{v}^k-\widetilde{\boldsymbol{v}}^k) \geqslant \|\boldsymbol{y}^k-\boldsymbol{y}^{k+1}\|_{\boldsymbol{D}}^2+(\tau\beta-\frac{1}{\gamma})\|\boldsymbol{B}(\boldsymbol{y}^k-\boldsymbol{y}^{k+1})\|^2+$$
$$\left(\frac{2-\gamma}{\gamma^2\beta}-\frac{1}{\gamma}\right)\|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2 \tag{23}$$

证 由(21)式 $\boldsymbol{G}$ 的定义和 $\widetilde{\boldsymbol{y}}^k=\boldsymbol{y}^{k+1}$，有

$$(\boldsymbol{v}^k-\widetilde{\boldsymbol{v}}^k)^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{v}^k-\widetilde{\boldsymbol{v}}^k)=\tau\beta\|\boldsymbol{B}(\boldsymbol{y}^k-\boldsymbol{y}^{k+1})\|^2+\|\boldsymbol{y}^k-\boldsymbol{y}^{k+1}\|_{\boldsymbol{D}}^2+\frac{2-\gamma}{\beta}\|\boldsymbol{\lambda}^k-\widetilde{\boldsymbol{\lambda}}^k\|^2-$$
$$2(\boldsymbol{\lambda}^k-\widetilde{\boldsymbol{\lambda}}^k)^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{y}^k-\boldsymbol{y}^{k+1}) \tag{24}$$

而从

$$\widetilde{\boldsymbol{\lambda}}^k=\boldsymbol{\lambda}^k-\beta(\boldsymbol{A}\boldsymbol{x}^{k+1}+\boldsymbol{B}\boldsymbol{y}^k-\boldsymbol{b}),\ \boldsymbol{\lambda}^{k+1}=\boldsymbol{\lambda}^k-\gamma\beta(\boldsymbol{A}\boldsymbol{x}^{k+1}+\boldsymbol{B}\boldsymbol{y}^k-\boldsymbol{b})$$

有

$$\boldsymbol{\lambda}^k-\widetilde{\boldsymbol{\lambda}}^k=\beta(\boldsymbol{A}\boldsymbol{x}^{k+1}+\boldsymbol{B}\boldsymbol{y}^k-\boldsymbol{b})=\frac{1}{\gamma}(\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1})$$

所以

$$\frac{2-\gamma}{\beta}\|\boldsymbol{\lambda}^k-\widetilde{\boldsymbol{\lambda}}^k\|^2=\frac{2-\gamma}{\gamma^2\beta}\|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2 \tag{25}$$

和

$$(\boldsymbol{\lambda}^k-\widetilde{\boldsymbol{\lambda}}^k)^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{y}^k-\boldsymbol{y}^{k+1})=\frac{1}{\gamma}(\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1})^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{y}^k-\boldsymbol{y}^{k+1}) \tag{26}$$

将(25)式和(26)式代入(24)式中，有

$$(\boldsymbol{v}^k-\widetilde{\boldsymbol{v}}^k)^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{v}^k-\widetilde{\boldsymbol{v}}^k)=\tau\beta\|\boldsymbol{B}(\boldsymbol{y}^k-\boldsymbol{y}^{k+1})\|^2+\|\boldsymbol{y}^k-\boldsymbol{y}^{k+1}\|_{\boldsymbol{D}}^2+\frac{2-\gamma}{\gamma^2\beta}\|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2-$$
$$\frac{2}{\gamma}(\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1})^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{y}^k-\boldsymbol{y}^{k+1}) \tag{27}$$

接下来处理(27)式右边的交叉项，由 Cauchy-Schwarz 不等式以及 $\gamma \in (0, 1)$ 知

$$-\frac{2}{\gamma}(\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1})^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{y}^k-\boldsymbol{y}^{k+1}) \geqslant -\frac{1}{\gamma}\|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2-\frac{1}{\gamma}\|\boldsymbol{B}(\boldsymbol{y}^k-\boldsymbol{y}^{k+1})\|^2 \tag{28}$$

结合(27)式和(28)式，即可得到(23)式.

由引理3和引理4可以得到下面的引理：

引理 5 设($\boldsymbol{\omega}^k$)为算法1所生成的序列，$\widetilde{\boldsymbol{\omega}}^k$ 由(11)式定义，罚参数 $\beta>1$，则对任意的 $\tau \in \left[\frac{1+\beta}{2\beta}, 1\right)$，$\gamma \in \left(0, \frac{2}{1+\beta}\right]$ 有

$$\theta(\boldsymbol{u})-\theta(\widetilde{\boldsymbol{u}}^k)+(\boldsymbol{\omega}-\widetilde{\boldsymbol{\omega}}^k)^{\mathrm{T}}F(\boldsymbol{\omega})\geqslant$$
$$\frac{1}{2}(\|\boldsymbol{v}-\boldsymbol{v}^{k+1}\|_{\boldsymbol{H}}^2-\|\boldsymbol{v}-\boldsymbol{v}^k\|_{\boldsymbol{H}}^2)+\frac{1}{2}\|\boldsymbol{y}^k-\boldsymbol{y}^{k+1}\|_{\boldsymbol{D}}^2+\frac{1}{2}(\tau\beta-\frac{1}{\gamma})\|\boldsymbol{B}(\boldsymbol{y}^k-\boldsymbol{y}^{k+1})\|^2+$$
$$\frac{1}{2}\left(\frac{2-\gamma}{\gamma^2\beta}-\frac{1}{\gamma}\right)\|\boldsymbol{\lambda}^k-\boldsymbol{\lambda}^{k+1}\|^2\qquad\forall\boldsymbol{\omega}\in\Omega\tag{29}$$

证　(29)式可直接由(22)式和(23)式得到.

由(29)式，结合文献[6]中第6部分中关于IP-LADMM的收敛性证明过程，可类似推出算法1的收敛性成立，故此处省略.

下面分析算法1在遍历意义下的收敛速率.

定理1　设$(\boldsymbol{\omega}^k)$为算法1所生成的序列，$\widetilde{\boldsymbol{\omega}}^k$定义在(11)式中，罚参数$\beta>1$，则对任意的$\tau\in\left[\frac{1+\beta}{2\beta},1\right)$，$\gamma\in\left(0,\frac{2}{1+\beta}\right]$以及迭代次数$t$有

$$\theta(\widetilde{\boldsymbol{u}}_t)-\theta(\boldsymbol{u})+(\widetilde{\boldsymbol{\omega}}_t-\boldsymbol{\omega})^{\mathrm{T}}F(\boldsymbol{\omega})\leqslant\frac{1}{2t}\|\boldsymbol{v}-\boldsymbol{v}^1\|_{\boldsymbol{H}}^2\tag{30}$$

其中

$$\widetilde{\boldsymbol{\omega}}_t=\frac{1}{t}(\sum_{k=1}^{t}\widetilde{\boldsymbol{\omega}}^k)$$

证　由(29)式，且其中$\varphi(\boldsymbol{v}^k,\boldsymbol{v}^{k+1})$为非负函数，有

$$\theta(\boldsymbol{u})-\theta(\widetilde{\boldsymbol{u}}^k)+(\boldsymbol{\omega}-\widetilde{\boldsymbol{\omega}}^k)^{\mathrm{T}}F(\boldsymbol{\omega})\geqslant\frac{1}{2}\|\boldsymbol{v}-\boldsymbol{v}^{k+1}\|_{\boldsymbol{H}}^2-\frac{1}{2}\|\boldsymbol{v}-\boldsymbol{v}^k\|_{\boldsymbol{H}}^2\tag{31}$$

从$k=1,2,\cdots,t$对(31)式进行求和，有

$$\sum_{k=1}^{t}\theta(\widetilde{\boldsymbol{u}}^k)-t\theta(\boldsymbol{u})+(\sum_{k=1}^{t}\widetilde{\boldsymbol{\omega}}^k-t\omega)^{\mathrm{T}}F(\boldsymbol{\omega})\leqslant\frac{1}{2}\|\boldsymbol{v}-\boldsymbol{v}^1\|_{\boldsymbol{H}}^2\tag{32}$$

因为$\theta(\cdot)$是凸函数，并且

$$\widetilde{\boldsymbol{u}}_t=\frac{1}{t}(\sum_{k=1}^{t}\widetilde{\boldsymbol{u}}^k)$$

所以有

$$\theta(\widetilde{\boldsymbol{u}}_t)=\frac{1}{t}(\sum_{k=1}^{t}\theta(\widetilde{\boldsymbol{u}}^k))$$

结合(31)式，即可得到(30)式.

参考文献：

[1] GLOWINSKI R, MARROCCO A. Sur l'approximation, Paréléments Finis D'ordre 1, et la Résolution, par Pénalisation-Dualité d'une Classe de Problémes de Dirichlet non Linéaires [J]. Journal of Equine Veterinary Science, 1975, 2(R-2): 41-76.

[2] GABAY D, MERCIER B. A Dual Algorithm for the Solution of Nonlinear Variational Problems Via Finite Element Approximation [J]. Comput Math Appl, 1976, 2(1): 17-40.

[3] YANG J F, YUAN X M. Linearized Augmented Lagrangian and Alternating Direction Methods for Nuclear Norm Minimization [J]. Math Comp, 2013, 82(281): 301-329.

[4] ZHANG X Q, BURGER M, OSHER S. A Unified Primal-Dual Algorithm Framework Based on Bergman Iteration [J]. J Sci Comput, 2011, 46(1): 20-46.

[5] HE B S, LIAO L Z, HAN D, et al. A New Inexact Alternating Directions Method for Monotone Variational Inequalities

[J]. Math Program, 2002, 92(1): 103－118.

[6] HE B S, MA F, YUAN X M. Linearized Alternating Direction Method of Multipliers Via Positive-Indefinite Proximal Regularization for Convex Programming [DB/OL]. (2016－7－31)[2017－5－20]. http://www.optimization-online.org/DB_HTML/2016/07/5569.html.

[7] HE B S, MA F, YUAN X M. Positive-Indefinite Proximal Augmented Lagrangian Method and its Application to Full Jacobian Splitting for Multi-Block Separable Convex Minimization Problems [DB/OL]. (2016－9－15) [2017－5－20]. http://www.optimization-online.org/DB_HTML/2016/09/5631.html.

On the Convergence Rate of ADMM with a Positive-Indefinite Proximal Term

WANG Yi-yun[1], OU Xiao-qing[2], LI Gao-xi[3]

1. *School of Mathematics and Statistics, Southwest University, Chongqing 400715, China;*
2. *School of Management, Chongqing College of Humanities, Science & Technology, Hechuan Chongqing 401524, China;*
3. *School of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China*

Abstract: In this paper, we mainly investigate the convergence rate of ADMM with a positive-indefinite proximal term. We improve the iterative step size of the Lagrangian multiplier by introducing a relaxation factor, and establish the convergence rate of ADMM with a positive-indefinite proximal term in the ergodic sense under suitable assumptions on parameters.

Key words: convex programming problem; alternating direction method of multipliers; positive-indefinite proximal term; convergence rate

责任编辑 张 构