第 38 卷第 1 期 西南大学学报(自然科学版) 2016 年 1 月
Vol. 38 No. 1 Journal of Southwest University (Natural Science Edition) Jan. 2016

DOI：10.13718/j.cnki.xdzk.2016.01.014

基于投影交替方向法求解结构型单调变分不等式①

许 微， 彭建文

重庆师范大学 数学学院，重庆 401331

摘要： 通过构造新的下降方向对孙敏等人给出的投影型交替方向法进行改进和推广，提出了改进投影型交替方向法. 与前者相比较，该方法具有收敛速度快，迭代次数少的特点. 在相同的假设条件下，证明了新方法的全局收敛性，并通过数值试验初步验证了该方法的有效性.

关 键 词： 单调变分不等式；可分离结构；投影收缩；交替方向法；全局收敛性

中图分类号： O221.2 **文献标志码：** A **文章编号：** 1673-9868(2016)01-0090-08

对于一个凸优化问题：

$$\begin{aligned} &\min f(\boldsymbol{x}) \\ &\text{s. t. } \boldsymbol{x} \in X \end{aligned} \tag{1}$$

其中：X 是 $\mathbb{R}^n$ 中的一个非空闭凸子集，$f(\cdot)$ 是从 $\mathbb{R}^n$ 到 $\mathbb{R}$ 的可微凸函数. 记

$$\boldsymbol{F}(\boldsymbol{x}) = \nabla f(\boldsymbol{x})$$

若 $\boldsymbol{x}^*$ 是问题(1) 的解，则该凸优化问题可以转化为一个变分不等式的形式：

$$(\boldsymbol{x} - \boldsymbol{x}^*)^{\mathrm{T}} \boldsymbol{F}(\boldsymbol{x}^*) \geqslant 0 \qquad \forall \boldsymbol{x} \in X \tag{2}$$

本文中，我们考虑的是一种带线性约束的结构型凸优化问题：

$$\min \{ f(\boldsymbol{x}) + g(\boldsymbol{y}) \,|\, \boldsymbol{Ax} + \boldsymbol{By} = \boldsymbol{b},\, \boldsymbol{x} \in X,\, \boldsymbol{y} \in Y \} \tag{3}$$

其中：$f: \mathbb{R}^{n_1} \longrightarrow \mathbb{R}$，$g: \mathbb{R}^{n_2} \longrightarrow \mathbb{R}$ 为可微凸函数；$\boldsymbol{A} \in \mathbb{R}^{m \times n_1}$；$\boldsymbol{B} \in \mathbb{R}^{m \times n_2}$；$\boldsymbol{b} \in \mathbb{R}^m$；$X,Y$ 分别为 $\mathbb{R}^{n_1}$ 和 $\mathbb{R}^{n_2}$ 中的一个非空闭凸子集.

在问题(3) 中，对线性约束 $\boldsymbol{Ax} + \boldsymbol{By} = \boldsymbol{b}$ 引入 Lagrange 乘子 $\boldsymbol{\lambda} \in \mathbb{R}^m$，即可得到一个定义在 $W = X \times Y \times \mathbb{R}^m$ 上的 Lagrange 函数：

$$L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + g(\boldsymbol{y}) - \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{Ax} + \boldsymbol{By} - \boldsymbol{b}) \tag{4}$$

则对于问题(3) 的求解，我们可以转化为找 Lagrange 函数的鞍点$(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\lambda}^*)$，满足：

$$L(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\lambda}) \leqslant L(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\lambda}^*) \leqslant L(\boldsymbol{x}, \boldsymbol{y}^*, \boldsymbol{\lambda}^*)$$

$$L(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\lambda}) \leqslant L(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\lambda}^*) \leqslant L(\boldsymbol{x}^*, \boldsymbol{y}, \boldsymbol{\lambda}^*)$$

为了方便，记

$$\boldsymbol{F}(\boldsymbol{x}) = \nabla f(\boldsymbol{x}),\ \boldsymbol{G}(\boldsymbol{y}) = \nabla g(\boldsymbol{y})$$

利用一阶最优性条件，上述鞍点问题等价于：找$(\boldsymbol{x}^*, \boldsymbol{y}^*, \boldsymbol{\lambda}^*) \in W$，使得对 $\forall (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda}) \in W$，有

① 收稿日期：2015-03-07
基金项目：国家自然科学基金项目(11171363)；重庆市基础科学与前沿技术研究重点项目(cstc2015jcyjBX0029).
作者简介：许 微(1989-)，男，江苏宿迁人，硕士，主要从事最优化算法和最优化控制研究.
通信作者：彭建文，教授.

$$\begin{cases}(\boldsymbol{x}-\boldsymbol{x}^{*})^{\mathrm{T}}(\boldsymbol{F}(\boldsymbol{x}^{*})-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}^{*})\geqslant 0\\(\boldsymbol{y}-\boldsymbol{y}^{*})^{\mathrm{T}}(\boldsymbol{G}(\boldsymbol{y}^{*})-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\lambda}^{*})\geqslant 0\\(\boldsymbol{\lambda}-\boldsymbol{\lambda}^{*})^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x}^{*}+\boldsymbol{B}\boldsymbol{y}^{*}-\boldsymbol{b})\geqslant 0\end{cases}$$

令

$$\boldsymbol{w}=\begin{pmatrix}\boldsymbol{x}\\\boldsymbol{y}\\\boldsymbol{\lambda}\end{pmatrix}$$

$$\boldsymbol{Q}(\boldsymbol{w})=\begin{pmatrix}\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}\\\boldsymbol{G}(\boldsymbol{y})-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\lambda}\\\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{y}-\boldsymbol{b}\end{pmatrix}$$

则问题(3)等价于如下结构型变分不等式问题的形式：找一点 $\boldsymbol{w}^{*}\in W$，使得

$$(\boldsymbol{w}-\boldsymbol{w}^{*})^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{w}^{*})\geqslant 0\qquad\forall\,\boldsymbol{w}\in W\tag{5}$$

对于结构型变分不等式(5)的求解，Gabay[1] 最早提出了一种非常有效的求解方法 — 交替方向法(ADM)，该方法将一个大型原问题分解为一系列低维子问题，在每步迭代过程中只需要求解一些子变分不等式. 近年来，许多学者在如何有效提高交替方向法的计算能力方面提出了很多改进的策略. 如 Ye[2] 通过对下降方向的步长进行优化，提出了一种预测 — 校正交替方向法；由于变分不等式求解的困难性，Han[3]，He[4] 等给出了非精确求解的正交投影交替方向法；Sun[5-6] 在此基础上提出新的投影型交替方向法，在每步迭代过程中只需进行一次正交投影.

本文改进了文献[6]中的方法，通过有效利用每步迭代过程所产生的预测点，提高了正交投影交替方向法的计算能力. 相对于文献[6]中的方法，改进后的投影型交替方法具有迭代次数少，收敛速度快的特点.

1　预备知识

本文主要在 Euclidean 范数下，给出一些基本定义和相关性质，记

$$\|\boldsymbol{x}\|=\sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$$

定义 1　设 Ω 是 $\mathbb{R}^n$ 中的一个非空闭凸子集. 用 $\boldsymbol{P}_{\Omega}(\boldsymbol{v})$ 表示点 $\boldsymbol{v}$ 从 $\mathbb{R}^n$ 到 Ω 上的投影，记为：

$$\boldsymbol{P}_{\Omega}(\boldsymbol{v})=\operatorname{argmin}\{\|\boldsymbol{v}-\boldsymbol{u}\|\;\big|\;\boldsymbol{v}\in\mathbb{R}^n,\ \forall\,\boldsymbol{u}\in\Omega\}$$

引理 1　对任意的一个非空闭凸子集 $\Omega\subseteq\mathbb{R}^n$，下式成立：

$$(\boldsymbol{v}-\boldsymbol{P}_{\Omega}(\boldsymbol{v}))^{\mathrm{T}}(\boldsymbol{u}-\boldsymbol{P}_{\Omega}(\boldsymbol{v}))\leqslant 0\qquad\forall\,\boldsymbol{v}\in\mathbb{R}^n,\ \forall\,\boldsymbol{u}\in\Omega\tag{6}$$

$$(\boldsymbol{v}-\boldsymbol{w})^{\mathrm{T}}[\boldsymbol{P}_{\Omega}(\boldsymbol{v})-\boldsymbol{P}_{\Omega}(\boldsymbol{w})]\geqslant\|\boldsymbol{P}_{\Omega}(\boldsymbol{v})-\boldsymbol{P}_{\Omega}(\boldsymbol{w})\|^2$$

通过引理1，利用 Cauchy-Schwarz 不等式，我们可以得到投影的非扩张性：

$$\|\boldsymbol{P}_{\Omega}(\boldsymbol{v})-\boldsymbol{P}_{\Omega}(\boldsymbol{w})\|\leqslant\|\boldsymbol{v}-\boldsymbol{w}\|\qquad\forall\,\boldsymbol{v},\boldsymbol{w}\in\mathbb{R}^n\tag{7}$$

引理 2　对任意的 $\beta>0$，$\boldsymbol{x}\in\mathbb{R}^n$，记

$$\boldsymbol{e}(\boldsymbol{x},\beta)=\boldsymbol{x}-\boldsymbol{P}_{\Omega}(\boldsymbol{x}-\beta\boldsymbol{F}(\boldsymbol{x}))$$

为 $\boldsymbol{F}(\boldsymbol{x})$ 在 Ω 上投影的余差函数，则 $\boldsymbol{x}^{*}$ 是变分不等式(2)的解，当且仅当 $\boldsymbol{e}(\boldsymbol{x}^{*},\beta)=\boldsymbol{0}$.

引理 3[7-8]　对任意的 $\boldsymbol{x}\in\mathbb{R}^n$，若 $\tilde{\beta}\geqslant\beta\geqslant 0$，则：

$$\|\boldsymbol{e}(\boldsymbol{x},\tilde{\beta})\|\geqslant\|\boldsymbol{e}(\boldsymbol{x},\beta)\|\tag{8}$$

$$\frac{\|\boldsymbol{e}(\boldsymbol{x},\tilde{\beta})}{\tilde{\beta}}\|\leqslant\frac{\|\boldsymbol{e}(\boldsymbol{x},\beta)\|}{\beta}\tag{9}$$

引理3表明，在设计算法时若以 $\|\boldsymbol{e}(\boldsymbol{x},\beta)\|$ 作为停机准则，则 $\beta>0$ 不宜太大也不宜太小，否则会出现算法不收敛或收敛速度太慢的情况.

定义 2　设映射 $\boldsymbol{F}:\Omega\longrightarrow\mathbb{R}^n$，若满足：

$$(\boldsymbol{u}-\boldsymbol{v})^{\mathrm{T}}[\boldsymbol{F}(\boldsymbol{u})-\boldsymbol{F}(\boldsymbol{v})]\geqslant 0\qquad\forall\,\boldsymbol{u},\boldsymbol{v}\in\Omega$$

则称映射 $\boldsymbol{F}$ 在 Ω 上是单调的.

在后面对于结构型变分不等式(5)的相关性质进行讨论时，我们均假设 $\boldsymbol{F}$,$\boldsymbol{G}$ 是连续单调的.

2　算法及其收敛性分析

为了求解结构型变分不等式(5)，本节将给出改进的投影型交替方向算法.

2.1　算法的描述

步骤 0　任意给定 $\varepsilon > 0$，任取初始点$(\boldsymbol{x}^0, \boldsymbol{y}^0, \boldsymbol{\lambda}^0) \in W$，$\beta > 0$，松弛因子 $\gamma \in (0, 2)$，$0 < \delta < 1$，$0 < \nu < 1$，令 $k = 0$.

步骤 1　找预测点 $\tilde{\boldsymbol{x}}^k$，使得

$$\tilde{\boldsymbol{x}}^k = \boldsymbol{P}_X(\boldsymbol{x}^k - \beta(\boldsymbol{F}(\boldsymbol{x}^k) - \boldsymbol{A}^{\mathrm{T}}(\boldsymbol{\lambda}^k - \beta(\boldsymbol{A}\boldsymbol{x}^k + \boldsymbol{B}\boldsymbol{y}^k - \boldsymbol{b})))) \tag{10}$$

步骤 2　找预测点 $\tilde{\boldsymbol{y}}^k$，使得

$$\tilde{\boldsymbol{y}}^k = \boldsymbol{P}_Y(\boldsymbol{y}^k - \beta(\boldsymbol{G}(\boldsymbol{y}^k) - \boldsymbol{B}^{\mathrm{T}}(\boldsymbol{\lambda}^k - \beta(\boldsymbol{A}\tilde{\boldsymbol{x}}^k + \boldsymbol{B}\boldsymbol{y}^k - \boldsymbol{b})))) \tag{11}$$

步骤 3　找预测点 $\tilde{\boldsymbol{\lambda}}^k$，使得

$$\tilde{\boldsymbol{\lambda}}^k = \boldsymbol{\lambda}^k - \beta(\boldsymbol{A}\tilde{\boldsymbol{x}}^k + \boldsymbol{B}\tilde{\boldsymbol{y}}^k - \boldsymbol{b}) \tag{12}$$

步骤 4　令

$$\boldsymbol{e}(\boldsymbol{w}^k, \beta) = \begin{pmatrix} \boldsymbol{e}_1(\boldsymbol{w}^k, \beta) \\ \boldsymbol{e}_2(\boldsymbol{w}^k, \beta) \\ \boldsymbol{e}_3(\boldsymbol{w}^k, \beta) \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}^k - \tilde{\boldsymbol{x}}^k \\ \boldsymbol{y}^k - \tilde{\boldsymbol{y}}^k \\ \boldsymbol{\lambda}^k - \tilde{\boldsymbol{\lambda}}^k \end{pmatrix}$$

$$\boldsymbol{w}^k = \begin{pmatrix} \boldsymbol{x}^k \\ \boldsymbol{y}^k \\ \boldsymbol{\lambda}^k \end{pmatrix} \in X \times Y \times \mathbb{R}^m$$

为了书写方便，下面简记 $\boldsymbol{e}(\boldsymbol{w}^k, \beta)$ 为 $\boldsymbol{e}^k$. 若 $\|\boldsymbol{e}^k\|_\infty \leqslant \varepsilon$，则算法终止；否则，计算

$$a = \min\left\{1 - \delta - \frac{(5 + 2\sqrt{2})\beta^2}{2\sqrt{2}}\|\boldsymbol{A}\|^2,\ 1 - \delta - \frac{(4 + 2\sqrt{2})\beta^2}{2\sqrt{2}}\|\boldsymbol{B}\|^2,\ 1 - \delta\right\}$$

若 $a \geqslant 0$，且满足：

$$\beta\{(\boldsymbol{e}_1{}^k)^{\mathrm{T}}[\boldsymbol{F}(\boldsymbol{x}^k) - \boldsymbol{F}(\boldsymbol{x}^k - \boldsymbol{e}_1{}^k)] + (\boldsymbol{e}_2^k)^{\mathrm{T}}[\boldsymbol{G}(\boldsymbol{y}^k) - \boldsymbol{G}(\boldsymbol{y}^k - \boldsymbol{e}_2{}^k)]\} \leqslant \delta\|\boldsymbol{e}^k\|^2$$

则通过计算有利方向 $\boldsymbol{d}(\boldsymbol{w}^k, \beta)$，简记为 $\boldsymbol{d}^k$，使得

$$\boldsymbol{d}^k = \begin{pmatrix} \boldsymbol{d}_1{}^k \\ \boldsymbol{d}_2{}^k \\ \boldsymbol{d}_3{}^k \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}_1{}^k - \beta\boldsymbol{F}(\boldsymbol{x}^k) + \beta\boldsymbol{F}(\boldsymbol{x}^k - \boldsymbol{e}_1{}^k) - \beta^2\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{e}_1{}^k + \boldsymbol{B}\boldsymbol{e}_2{}^k) \\ \boldsymbol{e}_2{}^k - \beta\boldsymbol{G}(\boldsymbol{y}^k) + \beta\boldsymbol{G}(\boldsymbol{y}^k - \boldsymbol{e}_2{}^k) - \beta^2\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{e}_2{}^k \\ \beta(\boldsymbol{A}\boldsymbol{x}^k + \boldsymbol{B}\boldsymbol{y}^k - \boldsymbol{b}) - \beta\boldsymbol{A}\boldsymbol{e}_1{}^k - \beta\boldsymbol{B}\boldsymbol{e}_2{}^k \end{pmatrix}$$

确定最优步长

$$\alpha_k = \frac{a\|\boldsymbol{e}^k\|^2}{\|\boldsymbol{d}^k\|^2} \tag{13}$$

否则，令

$$\beta = \frac{2}{3}\beta \cdot \min\left\{1,\ \frac{\|\boldsymbol{e}^k\|}{\|\boldsymbol{F}(\boldsymbol{x}^k) - \boldsymbol{F}(\boldsymbol{x}^k - \boldsymbol{e}_1{}^k)\| + \|\boldsymbol{G}(\boldsymbol{y}^k) - \boldsymbol{G}(\boldsymbol{y}^k - \boldsymbol{e}_2{}^k)\|}\right\}$$

转步骤 1.

步骤 5　通过有利方向和最优步长得到校正迭代点

$$\boldsymbol{w}^{k+1} = (\boldsymbol{x}^{k+1}, \boldsymbol{y}^{k+1}, \boldsymbol{\lambda}^{k+1})$$

其中，

$$\boldsymbol{x}^{k+1} = \boldsymbol{P}_X(\boldsymbol{x}^k - \gamma\alpha_k\boldsymbol{d}_1{}^k)$$

$$\boldsymbol{y}^{k+1} = \boldsymbol{P}_Y(\boldsymbol{y}^k - \gamma\alpha_k\boldsymbol{d}_2{}^k)$$

$$\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k - \gamma\alpha_k\boldsymbol{d}_3{}^k$$

为了防止 β 过小，我们做如下处理：若

$$\beta\{(\boldsymbol{e}_1{}^k)^{\mathrm{T}}[\boldsymbol{F}(\boldsymbol{x}^k)-\boldsymbol{F}(\boldsymbol{x}^k-\boldsymbol{e}_1{}^k)]+(\boldsymbol{e}_2{}^k)^{\mathrm{T}}[\boldsymbol{G}(\boldsymbol{y}^k)-\boldsymbol{G}(\boldsymbol{y}^k-\boldsymbol{e}_2{}^k)]\}\leqslant\nu\|\boldsymbol{e}^k\|^2$$

则 $\beta=\dfrac{3}{2}\beta$，令 $k=k+1$，转步骤 1.

注 我们与文献[6] 的不同之处在于本文充分利用预测点 $\tilde{\boldsymbol{x}}$ 的有效信息来寻找预测点 $\tilde{\boldsymbol{y}}$，从而实现预测点的交替寻找并使得算法更快收敛到最优解. 相应地，在有利方向 $\boldsymbol{d}(\boldsymbol{w},\beta)$ 和最优步长 α 的构造方面我们也给出了新的表达形式，并在下一小节中具体分析. 下面在没有特殊说明的情况下，我们均分别简记 $\boldsymbol{e}(\boldsymbol{w},\beta)$ 和 $\boldsymbol{d}(\boldsymbol{w},\beta)$ 为 $\boldsymbol{e}$ 和 $\boldsymbol{d}$.

2.2 算法的收敛性分析

在改进的投影型交替方向算法中，我们首先给出 $\|\boldsymbol{e}\|_\infty\leqslant\varepsilon$ 作为终止准则的原因：

定理 1 若 $\boldsymbol{e}_1=\boldsymbol{e}_2=\boldsymbol{e}_3=\boldsymbol{0}$，则$(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda})\in W$ 为变分不等式(5) 的一个解.

证 当 $\boldsymbol{e}_1=\boldsymbol{e}_2=\boldsymbol{e}_3=\boldsymbol{0}$ 时，我们有 $\boldsymbol{x}=\tilde{\boldsymbol{x}}$，$\boldsymbol{y}=\tilde{\boldsymbol{y}}$，$\boldsymbol{\lambda}=\tilde{\boldsymbol{\lambda}}$.

由式(12) 知

$$\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{y}=\boldsymbol{A}\tilde{x}+\boldsymbol{B}\boldsymbol{y}=\boldsymbol{b}$$

代入式(10),(11) 即得：

$$\boldsymbol{x}=\boldsymbol{P}_X(\boldsymbol{x}-\beta[\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}])$$
$$\boldsymbol{y}=\boldsymbol{P}_Y(\boldsymbol{y}-\beta[\boldsymbol{G}(\boldsymbol{y})-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\lambda}])$$

由引理 2 可知，

$$(\boldsymbol{u}-\boldsymbol{x})^{\mathrm{T}}[\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}]\geqslant 0\qquad\forall\,\boldsymbol{u}\in X$$
$$(\boldsymbol{v}-\boldsymbol{y})^{\mathrm{T}}[\boldsymbol{G}(\boldsymbol{y})-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\lambda}]\geqslant 0\qquad\forall\,\boldsymbol{v}\in Y$$

所以，$(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda})$ 为变分不等式(5) 的一个解. 证毕.

我们利用文献[3] 中的相关方法得到如下引理：

引理 4 若 $\boldsymbol{w}\in W$ 不是变分不等式(5) 的解，则当 $\delta\in(0,1)$ 时，对任意的 $\tilde{\beta}\geqslant 1$，存在 $\beta\in(0,\tilde{\beta}]$，使得

$$\beta(\|\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1(\boldsymbol{w},\beta))\|+\|\boldsymbol{G}(\boldsymbol{y})-\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2(\boldsymbol{w},\beta))\|)\leqslant\delta\|\boldsymbol{e}(\boldsymbol{w},\beta)\|$$

证 反证法. 假设对任意的 $\beta\in(0,\tilde{\beta}]$，

$$\beta(\|\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1(\boldsymbol{w},\beta))\|+\|\boldsymbol{G}(\boldsymbol{y})-\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2(\boldsymbol{w},\beta))\|)>\delta\|\boldsymbol{e}(\boldsymbol{w},\beta)\|$$

因为 $\boldsymbol{F}$,$\boldsymbol{G}$ 均为连续函数，所以 $\boldsymbol{e}_1(\boldsymbol{w},\beta)$,$\boldsymbol{e}_2(\boldsymbol{w},\beta)$ 连续. 由引理 3 中的式(8) 可知，当 $\beta\to 0$ 时，

$$\boldsymbol{e}_1(\boldsymbol{w},\beta)\to\boldsymbol{0},\ \boldsymbol{e}_2(\boldsymbol{w},\beta)\to\boldsymbol{0}$$

所以

$$\|\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1(\boldsymbol{w},\beta))\|+\|\boldsymbol{G}(\boldsymbol{y})-\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2(\boldsymbol{w},\beta))\|\to 0$$

则

$$0\geqslant\lim_{\beta\to 0}\frac{\delta\|\boldsymbol{e}(\boldsymbol{w},\beta)\|}{\beta}$$

又因为 β,$\delta>0$，$\tilde{\beta}\geqslant 1$，所以由引理 3 中的式(9) 可知，

$$\lim_{\beta\to 0}\frac{\delta\|\boldsymbol{e}(\boldsymbol{w},\beta)\|}{\beta}\geqslant\delta\|\boldsymbol{e}(\boldsymbol{w},1)\|\geqslant 0$$

所以，

$$\|\boldsymbol{e}(\boldsymbol{w},1)\|=0$$

则

$$\boldsymbol{e}_1(\boldsymbol{w},1)=\boldsymbol{e}_2(\boldsymbol{w},1)=\boldsymbol{0}$$

取 $\beta=1$，则有

$$\|\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1(\boldsymbol{w},1))\|+\|\boldsymbol{G}(\boldsymbol{y})-\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2(\boldsymbol{w},1))\|=0$$

$0>0$ 矛盾，假设不成立. 证毕.

下面给出定理说明 $-\boldsymbol{d}$ 是距离函数 $\frac{1}{2}\|\boldsymbol{w}-\boldsymbol{w}^*\|^2$ 在 $\boldsymbol{w}$ 处的一个下降方向.

定理 2　令

$$a=\min\left\{1-\delta-\frac{(5+2\sqrt{2})\beta^2}{2\sqrt{2}}\|\boldsymbol{A}\|^2,\ 1-\delta-\frac{(4+2\sqrt{2})\beta^2}{2\sqrt{2}}\|\boldsymbol{B}\|^2,\ 1-\delta\right\}$$

若存在 $\beta>0$，$\delta\in(0,1)$，使得 $a\geqslant 0$，则有：

$$(\boldsymbol{w}-\boldsymbol{w}^*)^{\mathrm{T}}\boldsymbol{d}\geqslant a\|\boldsymbol{e}\|^2\geqslant 0$$

证　记

$$\boldsymbol{K}=\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{A}^{\mathrm{T}}[\boldsymbol{\lambda}-\beta(\boldsymbol{Ax}+\boldsymbol{By}-\boldsymbol{b})]$$

$$\boldsymbol{H}=\boldsymbol{G}(\boldsymbol{y})-\boldsymbol{B}^{\mathrm{T}}[\boldsymbol{\lambda}-\beta(\boldsymbol{A}\tilde{\boldsymbol{x}}+\boldsymbol{By}-\boldsymbol{b})]$$

令 $\boldsymbol{w}^*=(\boldsymbol{x}^*,\boldsymbol{y}^*,\boldsymbol{\lambda}^*)$ 为变分不等式(5) 的一个解，显然 $\boldsymbol{w}^*\in W$.

由引理 1 中的式(6) 我们可以得到：

$$(\boldsymbol{x}-\beta\boldsymbol{K}-\tilde{\boldsymbol{x}})^{\mathrm{T}}(\tilde{\boldsymbol{x}}-\boldsymbol{x}^*)\geqslant 0$$

$$(\boldsymbol{y}-\beta\boldsymbol{H}-\tilde{\boldsymbol{y}})^{\mathrm{T}}(\tilde{\boldsymbol{y}}-\boldsymbol{y}^*)\geqslant 0$$

经化简变形后得：

$$(\boldsymbol{x}-\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{e}_1-\beta\boldsymbol{K})\geqslant\|\boldsymbol{e}_1\|^2-\beta\boldsymbol{e}_1{}^{\mathrm{T}}\boldsymbol{K}\tag{14}$$

$$(\boldsymbol{y}-\boldsymbol{y}^*)^{\mathrm{T}}(\boldsymbol{e}_2-\beta\boldsymbol{H})\geqslant\|\boldsymbol{e}_2\|^2-\beta\boldsymbol{e}_2{}^{\mathrm{T}}\boldsymbol{H}\tag{15}$$

由于 $\boldsymbol{F}$,$\boldsymbol{G}$ 均为单调的，所以

$$\beta[\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1)-\boldsymbol{F}(\boldsymbol{x}^*)]^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{e}_1-\boldsymbol{x}^*)\geqslant 0$$

$$\beta[\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2)-\boldsymbol{G}(\boldsymbol{y}^*)]^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{e}_2-\boldsymbol{y}^*)\geqslant 0$$

变形后可得：

$$\beta[\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1)-\boldsymbol{F}(\boldsymbol{x}^*)]^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^*)\geqslant\beta[\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1)-\boldsymbol{F}(\boldsymbol{x}^*)]^{\mathrm{T}}\boldsymbol{e}_1\tag{16}$$

$$\beta[\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2)-\boldsymbol{G}(\boldsymbol{y}^*)]^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{y}^*)\geqslant\beta[\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2)-\boldsymbol{G}(\boldsymbol{y}^*)]^{\mathrm{T}}\boldsymbol{e}_2\tag{17}$$

因为 $\tilde{\boldsymbol{x}}=\boldsymbol{x}-\boldsymbol{e}_1\in X$，$\tilde{\boldsymbol{y}}=\boldsymbol{y}-\boldsymbol{e}_2\in Y$，所以在式(5) 中分别将 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 替换为 $\boldsymbol{x}-\boldsymbol{e}_1$ 和 $\boldsymbol{y}-\boldsymbol{e}_2$，即有：

$$\begin{aligned}&\beta(\boldsymbol{x}-\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{F}(\boldsymbol{x}^*)-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}^*)-\beta\boldsymbol{e}_1{}^{\mathrm{T}}(\boldsymbol{F}(\boldsymbol{x}^*)-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\lambda}^*)+\\&\beta(\boldsymbol{y}-\boldsymbol{y}^*)^{\mathrm{T}}(\boldsymbol{G}(\boldsymbol{y}^*)-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\lambda}^*)-\beta\boldsymbol{e}_2{}^{\mathrm{T}}(\boldsymbol{G}(\boldsymbol{y}^*)-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{\lambda}^*)\geqslant 0\end{aligned}\tag{18}$$

将式(14)—(18) 相加合并，利用 $\boldsymbol{Ax}^*+\boldsymbol{By}^*=\boldsymbol{b}$，化简变形之后有：

$$\begin{aligned}&(\boldsymbol{x}-\boldsymbol{x}^*)^{\mathrm{T}}[\boldsymbol{e}_1-\beta\boldsymbol{F}(\boldsymbol{x})-\beta^2\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{Ax}+\boldsymbol{By}-\boldsymbol{b})+\beta\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1)+\beta^2\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x}-\tilde{\boldsymbol{x}})+\\&\beta^2\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{y}-\tilde{\boldsymbol{y}})]+(\boldsymbol{y}-\boldsymbol{y}^*)^{\mathrm{T}}[\boldsymbol{e}_2-\beta\boldsymbol{G}(\boldsymbol{y})-\beta^2\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{A}\tilde{\boldsymbol{x}}+\boldsymbol{By}-\boldsymbol{b})+\beta\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2)+\\&\beta^2\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x}-\tilde{\boldsymbol{x}})+\beta^2\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{y}-\tilde{\boldsymbol{y}})]+(\boldsymbol{\lambda}-\boldsymbol{\lambda}^*)^{\mathrm{T}}[\beta(\boldsymbol{Ax}+\boldsymbol{By}-\boldsymbol{b})-\beta\boldsymbol{Ae}_1-\beta\boldsymbol{Be}_2]\geqslant\\&\|\boldsymbol{e}_1\|^2+\|\boldsymbol{e}_2\|^2-\beta\boldsymbol{e}_1{}^{\mathrm{T}}(\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1))-\beta\boldsymbol{e}_2{}^{\mathrm{T}}(\boldsymbol{G}(\boldsymbol{y})-\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2))-\\&\beta^2(\boldsymbol{y}-\tilde{\boldsymbol{y}})\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}(\tilde{\boldsymbol{x}}-\boldsymbol{x})\end{aligned}\tag{19}$$

利用 Cauchy-Schwarz 不等式有：

$$\begin{aligned}-\beta^2(\boldsymbol{y}-\tilde{\boldsymbol{y}})\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}(\tilde{\boldsymbol{x}}-\boldsymbol{x})&\geqslant-\frac{\beta^2}{2}\left(\sqrt{2}\|\boldsymbol{B}(\boldsymbol{y}-\tilde{\boldsymbol{y}})\|^2+\frac{1}{\sqrt{2}}\|\boldsymbol{A}(\tilde{\boldsymbol{x}}-\boldsymbol{x})\|^2\right)\geqslant\\&-\frac{\beta^2}{2}\left(\sqrt{2}\|\boldsymbol{B}\|^2\|\boldsymbol{e}_2\|^2+\frac{1}{\sqrt{2}}\|\boldsymbol{A}\|^2\|\boldsymbol{e}_1\|^2\right)\end{aligned}\tag{20}$$

将式(20) 代入到式(19)，利用引理 4 化简可得：

$$\begin{aligned}&(\boldsymbol{x}-\boldsymbol{x}^*)^{\mathrm{T}}[\boldsymbol{e}_1-\beta\boldsymbol{F}(\boldsymbol{x})-\beta^2\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{Ax}+\boldsymbol{By}-\boldsymbol{b})+\beta\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1)+\beta^2\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x}-\tilde{\boldsymbol{x}})+\\&\beta^2\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}(\boldsymbol{y}-\tilde{\boldsymbol{y}})]+(\boldsymbol{y}-\boldsymbol{y}^*)^{\mathrm{T}}[\boldsymbol{e}_2-\beta\boldsymbol{G}(\boldsymbol{y})-\beta^2\boldsymbol{B}^{\mathrm{T}}(\boldsymbol{A}\tilde{\boldsymbol{x}}+\boldsymbol{By}-\boldsymbol{b})+\beta\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2)+\end{aligned}$$

$$\beta^2\boldsymbol{B}^{\mathrm T}\boldsymbol{A}(\boldsymbol{x}-\tilde{\boldsymbol{x}})+\beta^2\boldsymbol{B}^{\mathrm T}\boldsymbol{B}(\boldsymbol{y}-\tilde{\boldsymbol{y}})]+(\boldsymbol{\lambda}-\boldsymbol{\lambda}^*)^{\mathrm T}[\beta(\boldsymbol{Ax}+\boldsymbol{By}-\boldsymbol{b})-\beta\boldsymbol{Ae}_1-\beta\boldsymbol{Be}_2]\geqslant$$

$$\|\boldsymbol{e}_1\|^2+\|\boldsymbol{e}_2\|^2-\beta\boldsymbol{e}_1{}^{\mathrm T}(\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1))-\beta\boldsymbol{e}_2{}^{\mathrm T}(\boldsymbol{G}(\boldsymbol{y})-\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2))-$$

$$\frac{\beta^2}{2}\left(\sqrt{2}\|\boldsymbol{B}\|^2\|\boldsymbol{e}_2\|^2+\frac{1}{\sqrt{2}}\|\boldsymbol{A}\|^2\|\boldsymbol{e}_1\|^2\right)\geqslant$$

$$\|\boldsymbol{e}_1\|^2+\|\boldsymbol{e}_2\|^2-\beta\|\boldsymbol{e}_1\|\,\|\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{F}(\boldsymbol{x}-\boldsymbol{e}_1)\|-\beta\|\boldsymbol{e}_2\|\,\|\boldsymbol{G}(\boldsymbol{y})-\boldsymbol{G}(\boldsymbol{y}-\boldsymbol{e}_2)\|-$$

$$\frac{\beta^2}{2}\left(\sqrt{2}\|\boldsymbol{B}\|^2\|\boldsymbol{e}_2\|^2+\frac{1}{\sqrt{2}}\|\boldsymbol{A}\|^2\|\boldsymbol{e}_1\|^2\right)\geqslant$$

$$\left(1-\delta-\frac{\beta^2\|\boldsymbol{A}\|^2}{2\sqrt{2}}\right)\|\boldsymbol{e}_1\|^2+\left(1-\delta-\frac{\sqrt{2}\beta^2\|\boldsymbol{B}\|^2}{2}\right)\|\boldsymbol{e}_2\|^2+(-\delta)\|\boldsymbol{e}_3\|^2 \tag{21}$$

另一方面，利用 $\boldsymbol{Ax}^*+\boldsymbol{By}^*=\boldsymbol{b}$ 有：

$$\|\boldsymbol{e}_3\|^2=\beta^2(\boldsymbol{A}\tilde{\boldsymbol{x}}+\boldsymbol{B}\tilde{\boldsymbol{y}}-\boldsymbol{b})^{\mathrm T}(\boldsymbol{A}\tilde{\boldsymbol{x}}+\boldsymbol{B}\tilde{\boldsymbol{y}}-\boldsymbol{b})=$$

$$\beta^2[\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{x}^*+\tilde{\boldsymbol{x}}-\boldsymbol{x})+\boldsymbol{B}(\boldsymbol{y}-\boldsymbol{y}^*+\tilde{\boldsymbol{y}}-\boldsymbol{y})]^{\mathrm T}[\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{x}^*+\tilde{\boldsymbol{x}}-\boldsymbol{x})+$$

$$\boldsymbol{B}(\boldsymbol{y}-\boldsymbol{y}^*+\tilde{\boldsymbol{y}}-\boldsymbol{y})]=$$

$$\beta^2[(\boldsymbol{x}-\boldsymbol{x}^*)^{\mathrm T}\boldsymbol{A}^{\mathrm T}(2\boldsymbol{A}\tilde{\boldsymbol{x}}+2\boldsymbol{B}\tilde{\boldsymbol{y}}-\boldsymbol{Ax}-\boldsymbol{By}-\boldsymbol{b})+(\boldsymbol{y}-\boldsymbol{y}^*)^{\mathrm T}\boldsymbol{B}^{\mathrm T}(2\boldsymbol{A}\tilde{\boldsymbol{x}}+$$

$$2\boldsymbol{B}\tilde{\boldsymbol{y}}-\boldsymbol{Ax}-\boldsymbol{By}-\boldsymbol{b})]+\beta^2(\|\boldsymbol{Ae}_1\|^2+\|\boldsymbol{Be}_2\|^2+2(\boldsymbol{Ae}_1)^{\mathrm T}(\boldsymbol{Be}_2)) \tag{22}$$

由式(22)，再次利用 Cauchy-Schwarz 不等式可得：

$$\beta^2[(\boldsymbol{x}-\boldsymbol{x}^*)^{\mathrm T}\boldsymbol{A}^{\mathrm T}(2\boldsymbol{A}\tilde{\boldsymbol{x}}+2\boldsymbol{B}\tilde{\boldsymbol{y}}-\boldsymbol{Ax}-\boldsymbol{By}-\boldsymbol{b})+(\boldsymbol{y}-\boldsymbol{y}^*)^{\mathrm T}\boldsymbol{B}^{\mathrm T}(2\boldsymbol{A}\tilde{\boldsymbol{x}}+$$

$$2\boldsymbol{B}\tilde{\boldsymbol{y}}-\boldsymbol{Ax}-\boldsymbol{By}-\boldsymbol{b})]=$$

$$\|\boldsymbol{e}_3\|^2-\beta^2(\|\boldsymbol{Ae}_1\|^2+\|\boldsymbol{Be}_2\|^2+2(\boldsymbol{Ae}_1)^{\mathrm T}(\boldsymbol{Be}_2))\geqslant$$

$$\|\boldsymbol{e}_3\|^2-\beta^2\|\boldsymbol{A}\|^2\|\boldsymbol{e}_1\|^2-\beta^2\|\boldsymbol{B}\|^2\|\boldsymbol{e}_2\|^2-\beta^2\left(\sqrt{2}\|\boldsymbol{Ae}_1\|^2+\frac{1}{\sqrt{2}}\|\boldsymbol{Be}_2\|^2\right)\geqslant$$

$$-(1+\sqrt{2})\beta^2\|\boldsymbol{A}\|^2\|\boldsymbol{e}_1\|^2-\left(1+\frac{1}{\sqrt{2}}\right)\beta^2\|\boldsymbol{B}\|^2\|\boldsymbol{e}_2\|^2+\|\boldsymbol{e}_3\|^2 \tag{23}$$

将式(21),(23) 相加合并后可得：

$$(\boldsymbol{w}-\boldsymbol{w}^*)^{\mathrm T}\boldsymbol{d}\geqslant a\|\boldsymbol{e}\|^2\geqslant 0 \tag{24}$$

其中 $\boldsymbol{d}$ 即为步骤 4 中所描述的形式. 证毕.

对任意给定的 $\boldsymbol{w}^*$，$(\boldsymbol{w}-\boldsymbol{w}^*)$ 表示距离函数$\frac{1}{2}\|\boldsymbol{w}-\boldsymbol{w}^*\|^2$ 在 $\boldsymbol{w}$ 处的梯度. 所以由式(24) 知，$-\boldsymbol{d}$ 是距离函数$\frac{1}{2}\|\boldsymbol{w}-\boldsymbol{w}^*\|^2$ 在 $\boldsymbol{w}$ 处的一个下降方向.

定理 3 设 $\boldsymbol{w}^*$ 是变分不等式(5) 的解，则算法产生的点列$\{\boldsymbol{w}^k\}$是有界的.

证 由引理 1 中的式(24)，我们可以得到：

$$\|\boldsymbol{w}^{k+1}-\boldsymbol{w}^*\|^2=$$

$$\|\boldsymbol{x}^{k+1}-\boldsymbol{x}^*\|^2+\|\boldsymbol{y}^{k+1}-\boldsymbol{y}^*\|^2+\|\boldsymbol{\lambda}^{k+1}-\boldsymbol{\lambda}^*\|^2\leqslant$$

$$\|\boldsymbol{x}^k-\boldsymbol{x}^*-\gamma\alpha_k\boldsymbol{d}_1{}^k\|^2+\|\boldsymbol{y}^k-\boldsymbol{y}^*-\gamma\alpha_k\boldsymbol{d}_2{}^k\|^2+\|\boldsymbol{\lambda}-\boldsymbol{\lambda}^*-\gamma\alpha_k\boldsymbol{d}_3{}^k\|^2\leqslant$$

$$\|\boldsymbol{w}^k-\boldsymbol{w}^*\|^2-2\gamma\alpha_k(\boldsymbol{w}^k-\boldsymbol{w}^*)^{\mathrm T}\boldsymbol{d}^k+\gamma^2\alpha_k{}^2\|\boldsymbol{d}^k\|^2 \tag{25}$$

然后，利用式(13) 和式(24) 即可得到：

$$\|\boldsymbol{w}^{k+1}-\boldsymbol{w}^*\|^2\leqslant\|\boldsymbol{w}^k-\boldsymbol{w}^*\|^2-a(2-\gamma)\gamma\alpha_k\|\boldsymbol{e}^k\|^2 \tag{26}$$

因为 $a\geqslant 0$，$\gamma\in(0,2)$，所以

$$\|\boldsymbol{w}^{k+1}-\boldsymbol{w}^*\|^2\leqslant\|\boldsymbol{w}^k-\boldsymbol{w}^*\|^2\leqslant\cdots\leqslant\|\boldsymbol{w}^0-\boldsymbol{w}^*\|^2$$

故对任意的 $k\in N$，点列$\{\boldsymbol{w}^k\}$ 有界. 证毕.

定理 4[3] 设 $\boldsymbol{w}^*$ 是变分不等式(5) 的解，若算法产生的点列$\{\boldsymbol{w}^k\}$是有界的，则

1) $\lim\limits_{k\to\infty}\|\boldsymbol{e}^k\|=0$；

2) $\{\boldsymbol{w}^k\}$ 收敛到变分不等式(5)的一个解.

3　数值试验

我们利用文献[9]中的例子验证本文所提出的改进交替方向算法的有效性.

例　$\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}\,|\,\boldsymbol{x}\in\Omega_1\cap\Omega_2\}$，其中：

$$\boldsymbol{c}\in\mathbb{R}^n,\ \Omega_1=\{\boldsymbol{x}\,|\,\|\boldsymbol{x}\|\leqslant r_1,\ \boldsymbol{x}\in\mathbb{R}^n\}$$

$$\Omega_2=\{\boldsymbol{x}\,|\,\|\boldsymbol{x}-\boldsymbol{b}\|\leqslant r_2,\ \boldsymbol{x}\in\mathbb{R}^n,\ \boldsymbol{b}\in\mathbb{R}^n\}$$

我们将其等价地表示成式(3)的形式：

$$\begin{aligned}&\min\ \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}\\&\text{s. t. }\ \boldsymbol{x}+\boldsymbol{y}-\boldsymbol{b}=\boldsymbol{0}\\&\boldsymbol{x}\in\odot_{r_1},\ \boldsymbol{y}\in\odot_{r_2}\end{aligned}$$

其中 $\odot_{r_i}$ 表示以原点为心 r_i 为半径的球，$i=1,2$.

上述凸优化问题可以转化为结构型变分不等式(5)的形式：找 $\boldsymbol{w}^*\in\Omega$，使得

$$(\boldsymbol{w}-\boldsymbol{w}^*)^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{w}^*)\geqslant 0\qquad\forall\,\boldsymbol{w}\in\Omega$$

其中：

$$\boldsymbol{w}=\begin{pmatrix}\boldsymbol{x}\\\boldsymbol{y}\\\boldsymbol{\lambda}\end{pmatrix}$$

$$\boldsymbol{Q}(\boldsymbol{w})=\begin{pmatrix}\boldsymbol{c}\\\boldsymbol{0}\\\boldsymbol{x}+\boldsymbol{y}-\boldsymbol{b}\end{pmatrix}$$

$$\Omega=\{(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda})\,|\,\boldsymbol{x}\in\odot_{r_1},\ \boldsymbol{y}\in\odot_{r_2},\ \boldsymbol{\lambda}\in\mathbb{R}^n\}$$

我们记文献[6]的方法为(ADM1)，本文中所提出的方法为(ADM2)，分别从终止条件 ε 和维数 n 两个方面进行比较. 为了方便，令 $n=n_1=n_2$，取 $\beta=0.45$，$\gamma=1.9$，$\delta=0.01$，$\nu=0.001$，$\boldsymbol{b}$，$\boldsymbol{c}$ 为元素都为 1 的向量，$r_1=0.5\|\boldsymbol{b}\|$，$r_2=0.6\|\boldsymbol{b}\|$，$\boldsymbol{A}$，$\boldsymbol{B}$ 均为 n 维单位矩阵，设初始值$(\boldsymbol{x}^0,\boldsymbol{y}^0,\boldsymbol{\lambda}^0)$各元素均为 1. 两种方法的程序均用 Matlab 2012a 进行编程实现，并在 Windows 7 系统下，配置为 CPU Intel Core i5-2450M 2.5 GHz 的 Lenovo 笔记本上运行. 数值比较结果如表 1 和表 2 所示.

从收敛精度和空间维数两个方面进行比较可知：经过改进的投影型交替方向法具有迭代次数少、收敛速度快的特点，说明本文中所构造的方法有意义且计算效果更好. 出现这种现象有两个方面原因：首先，在每步迭代过程中，我们充分利用所产生的预测点的信息，提高了正交投影交替方向法的计算能力，因而造成了迭代次数少收敛速度快的特点. 另外，在数值试验中 δ 的选取也影响两种方法的运行效果. 当 δ 选取稍微小一些时，本文中所提出的方法在收敛速度方面会明显加快. 因此，在接下来的研究中，可以对 δ 进行适当优化.

表 1　$n=50$ 时(ADM1)与(ADM2)对终止条件 ε 的比较

终止条件 ε	ADM1		ADM2	
	迭代次数/次	运行时间/s	迭代次数/次	运行时间/s
10^{-3}	191	0.062 6	52	0.028 7
10^{-6}	429	0.121 5	115	0.062 5
10^{-9}	669	0.202 2	178	0.106 5
10^{-12}	909	0.258 4	241	0.138 4

表2 (ADM1)与(ADM2)基于 $\varepsilon=10^{-10}$ 对维数 n 的比较

维数 n	ADM1		ADM2	
	迭代次数/次	运行时间/s	迭代次数/次	运行时间/s
10	728	0.189 8	192	0.058 4
50	749	0.221 0	199	0.115 7
100	765	0.451 3	200	0.156 9
500	791	14.180 9	211	9.157 0

参考文献:

[1] GABAY D, MERCIER B. A Dual Algorithm for the Solution of Nonlinear Variational Problems via Finite Element Approximations [J]. Computers and Mathematics with Applications, 1976, 2(1): 17-40.

[2] YE C H, YUAN X M. A Descent Method for Structured Monotone Variational Inequalities [J]. Optimization Methods and Software, 2007, 22(2): 329-338.

[3] HAN D R. A Modified Alternating Direction Method for Variational Inequality Problems [J]. Applied Mathematics and Optimization, 2002, 45(1): 63-74.

[4] 何炳生. 一类广义线性变分不等式的求解与应用 [J]. 中国科学: A辑, 1995, 38(9): 939-945.

[5] 孙 敏, 徐健腾, 时贞军. 一种新的投影型变分不等式交替方向方法 [J]. 工程数学学报, 2006, 23(6): 1101-1104.

[6] 孙 敏. 求解结构型单调变分不等式的投影类交替方向法 [J]. 安徽大学学报(自然科学版), 2009, 33(2): 12-14.

[7] HE B S, LIAO L Z. Improvements of Some Projection Methods for Monotone Nonlinear Variational Inequalities [J]. Journal of Optimization Theory and Applications, 2002, 112(1): 111-128.

[8] GAFNI E M, BERTSEKAS D P. Two-Metric Projection Methods for Constrained Optimization [J]. SIAM Journal on Control and Optimization, 1984, 22(6): 936-964.

[9] 张 敏, 韩德仁, 何洪津, 等. 解可分离结构变分不等式的一种新的交替方向法 [J]. 中国科学: 数学, 2012, 42(2): 133-149.

Solving Structured Monotone Variational Inequalities Based on the Projection Alternating Direction Method

XU Wei, PENG Jian-wen

College of Mathematics Science, Chongqing Normal University, Chongqing 401331, China

Abstract: Alternating direction method plays an important role in solving structured monotone variational inequality problems. In this paper, we propose a modified projection alternating direction method by constructing a new descent direction, which improves and promotes the projection alternating direction method given by Sun, M. et al. Compared with Sun's method, the proposed method has greater speed of convergence and fewer iterations. Under the same assumptions, we prove the global convergence property of the method, and then test its effectiveness with a numerical example.

Key words: monotone variational inequality; separable structure; projection and contraction; alternating direction method; global convergence

责任编辑 张 枸