解可分离结构变分不等式的一种新的交替方向法

张敏^①, 韩德仁^{①②*}, 何洪津^①, 陈艳男^③

① 南京师范大学数学科学学院, 南京 210046;

② 江苏省大规模复杂系统数值模拟重点实验室, 南京 210046;

③ 南京林业大学理学院, 南京 210037

E-mail: zhangdamin56@163.com, handeren@njnu.edu.cn, hehj2003@163.com, yncen@njfu.edu.cn

收稿日期: 2011-01-17; 接受日期: 2011-12-15; * 通信作者

国家自然科学基金 (批准号: 11171159, 11071122) 资助项目

摘要 交替方向法是求解可分离结构变分不等式问题的经典方法之一, 它将一个大型的变分不等式问题分解成若干个小规模的变分不等式问题进行迭代求解. 但每步迭代过程中求解的子问题仍然摆脱不了求解变分不等式子问题的瓶颈. 从数值计算上来说, 求解一个变分不等式并不是一件容易的事情. 因此, 本文提出一种新的交替方向法, 每步迭代只需要求解一个变分不等式子问题和一个强单调的非线性方程组子问题. 相对变分不等式问题而言, 我们更容易、且有更多的有效算法求解一个非线性方程组问题. 在与经典的交替方向法相同的假设条件下, 我们证明了新算法的全局收敛性. 进一步的数值试验也验证了新算法的有效性.

关键词 变分不等式 交替方向法 全局收敛性 非线性方程组

MSC (2010) 主题分类 65K03

1 引言

一个变分不等式问题可描述为: 寻找一点 $u^* \in \Omega$, 使得如下不等式成立,

$$(u - u^*)^T F(u^*) \geq 0, \quad \forall u \in \Omega, \quad (1.1)$$

其中 Ω 是 $\mathbb{R}^n$ 中的一个非空闭凸子集, 函数 $F(\cdot)$ 是从 $\mathbb{R}^n$ 到其自身的映射. 本文中, 我们考虑变分不等式的一种特殊情形, 即考虑的变分不等式具有如下的结构:

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, \quad F(u) = \begin{pmatrix} f(x) \\ g(y) \end{pmatrix}, \quad (1.2)$$

以及

$$\Omega = \{(x, y) \mid x \in \mathcal{X}, \quad y \in \mathcal{Y}, \quad Ax + By = b\}, \quad (1.3)$$

这里的 $\mathcal{X}$ 和 $\mathcal{Y}$ 分别是 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 中给定的非空闭凸子集, $A \in \mathbb{R}^{l \times n}$ 和 $B \in \mathbb{R}^{l \times m}$ 是给定的矩阵, $b \in \mathbb{R}^l$ 是给定的向量, 且 $f: \mathcal{X} \rightarrow \mathbb{R}^n$ 和 $g: \mathcal{Y} \rightarrow \mathbb{R}^m$ 是给定的单调映射. 由于映射 f 和 g 分别只与变量 x 和 y 有关, 我们称这类变分不等式为可分离结构变分不等式. 虽然这类变分不等式有着特殊的结

英文引用格式: Zhang M, Han D R, He H J, et al. A new alternating direction method for solving separable variational inequality problems (in Chinese). Sci Sin Math, 2012, 42(2): 133–149, doi: 10.1360/012011-42

构, 但是我们不乏在很多的实际问题中找到它的应用, 尤其是在数学规划及经济和交通领域中, 如参见文献 [1–7]. 在求解这类有着特殊结构的变分不等式问题 (1.1)–(1.3) 上, 已经有很多的分解方法被提出来, 见文献 [3, 8–14].

考虑到问题 (1.1)–(1.3) 有一个线性等式约束, 很自然地通过引入一个 Lagrange 乘子 $\lambda \in \mathbb{R}^l$, 我们可以得到变分不等式问题 (1.1)–(1.3) 的等价形式:

$$z^* \in \mathcal{Z}, \quad (z - z^*)^T Q(z^*) \geq 0, \quad \forall z \in \mathcal{Z}, \quad (1.4)$$

其中

$$z = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}, \quad Q(z) = \begin{pmatrix} f(x) - A^T \lambda \\ g(y) - B^T \lambda \\ Ax + By - b \end{pmatrix}, \quad \mathcal{Z} = \mathcal{X} \times \mathcal{Y} \times \mathcal{R}^l. \quad (1.5)$$

为方便起见, 我们记变分不等式问题 (1.4)–(1.5) 为 $\text{MVI}(\mathcal{Z}, Q)$. 对于求解变分不等式问题 $\text{MVI}(\mathcal{Z}, Q)$, Gabay 等 [15, 16] 最早提出一种交替方向法, 该算法的迭代过程可简单描述如下:

首先, 给定点 $(y, \lambda) \in \mathcal{Y} \times \mathbb{R}^l$ 后, 找一新点 $\tilde{x} \in \mathcal{X}$ 使得如下不等式成立:

$$(x' - \tilde{x})^T \{f(\tilde{x}) - A^T[\lambda - \beta(A\tilde{x} + B\tilde{y} - b)]\} \geq 0, \quad \forall x' \in \mathcal{X}; \quad (1.6)$$

然后, 利用 $(\tilde{x}, \lambda)$ 的信息找一新点 $\tilde{y} \in \mathcal{Y}$ 满足如下不等式:

$$(y' - \tilde{y})^T \{g(\tilde{y}) - B^T[\lambda - \beta(A\tilde{x} + B\tilde{y} - b)]\} \geq 0, \quad \forall y' \in \mathcal{Y}; \quad (1.7)$$

最后, 利用 $(\tilde{x}, \tilde{y})$ 的信息更新 Lagrange 乘子 $\tilde{\lambda}$:

$$\tilde{\lambda} = \lambda - \beta(A\tilde{x} + B\tilde{y} - b), \quad (1.8)$$

这里的 $\beta > 0$ 是事先设定的正的罚参数.

从上述交替方向法的框架易见, 该方法将一个大型变分不等式原问题分解成一系列低维变分不等式子问题进行迭代求解. 由于其充分利用了问题的可分结构, 且有良好的数值表现, 该方法在求解大规模的可分离结构变分不等式问题上倍受青睐. 截止目前, 交替方向法的理论研究在 Lagrange 函数 [17, 18] 和极大单调算子 [9, 15] 的理论框架下已经取得了很大的研究进展. 从算法上具体分析, 交替方向法每步迭代过程中的主要工作量在于求解两个变分不等式子问题 (1.6) 和 (1.7). 因此, 算法的有效性与如何快速、有效地求解两个变分不等式子问题紧密相连. 近年来, 已有很多研究者致力于如何有效地求解子问题, 并提出了很多不同的策略. 如其中一种经典方法是: 在子问题上添加一个临近点项, 将原来单调的子问题变成一个强单调的子问题. 相对而言, 一个具有强单调性的子问题要比一个单调的子问题容易求解得多. 另外, 从数值计算角度分析, 精确求解一个子问题往往不太现实, 或者代价很大. 因此, 另一种策略是对子问题进行非精确求解. 相比精确求解子问题, 非精确求解要更可行、快速. 有关这两类策略的详细描述, 感兴趣的读者可参见文献 [8, 9, 19–26].

有趣的是, 文献 [3, 27–33] 中的大量数值结果表明, 当我们用交替方向法求解各类问题时, 罚参数 β 对算法的数值表现影响较大: 较小的 β 值虽然能保证算法的收敛性, 但是会导致较慢的收敛速度; 而一个较大的 β 值则可能会导致算法不收敛. 因此, 另一类改进策略是克服固定罚参数 β 带来的弊端, 即允许 β 在迭代过程中单调递增, 或者单调下降, 亦或自适应地调节, 感兴趣的读者可参见文献 [24, 32–34].

虽然上面提到的各种改进策略都能很有效地提高交替方向法的计算能力, 但值得指出的是, 上述改进的方法的子问题除了提升交替方向法的计算表现外, 仍然没有突破避免求解变分不等式子问题的瓶颈. 因为从计算角度上讲, 在大部分情况下, 求解变分不等式问题并非一件简单的事情. 为此, 本文提出一种新的交替方向法: 每步迭代过程中, 我们保留原交替方向法的变分不等式子问题 (1.6), 而另一个变分不等式子问题 (1.7) 则由下面的非线性方程组替代:

$$\tilde{y} + \beta g(\tilde{y}) = y + \beta g(y) - e_y(y, \beta), \quad (1.9)$$

其中

$$e_y(y, \beta) := y - P_Y[y - \beta(g(y) - B^T \bar{\lambda})],$$

且

$$\bar{\lambda} := \lambda - \beta(Ax + By - b).$$

此时, 相比原交替方向法中的变分不等式子问题 (1.7), 这里的子问题 (1.9) 在绝大部分情况下要容易求解得多. 而且, 若映射 g 是单调的, 则映射 $y + g(y)$ 是一个强单调的映射, 这比一个单调的算子具有更多的优良性质可以利用.

本文结构如下: 第 2 节我们主要总结一些有关变分不等式问题的预备知识; 第 3 节给出一种新的交替方向法, 并分析新方法的一些性质; 新方法的全局收敛性将在第 4 节给出; 同时, 在第 5 节我们给出一些初步的数值试验结果来说明新方法的有效性; 最后, 我们给出本文的一些总结和展望.

2 预备知识

首先, 我们用 $\|x\| = \sqrt{x^T x}$ 表示 Euclidean 范数. 令 Ω 是 $\mathbb{R}^n$ 中的一个非空闭凸子集, 用 $P_\Omega[\cdot]$ 表示从 $\mathbb{R}^n$ 到 Ω 的投影, 定义如下:

$$P_\Omega[x] := \arg \min \{ \|x - y\| \mid y \in \Omega \}.$$

那么, 根据投影的定义, 我们有如下性质.

引理 1 ^[35] 令 Ω 为 $\mathbb{R}^n$ 中的非空闭凸子集, 则对任意 $u, v \in \mathbb{R}^n$ 和任意 $w \in \Omega$, 下列不等式成立:

$$(u - P_\Omega[u])^T (w - P_\Omega[u]) \leq 0, \quad (2.1)$$

和

$$\|P_\Omega[u] - P_\Omega[v]\|^2 \leq \|u - v\|^2 - \|P_\Omega[u] - u + v - P_\Omega[v]\|^2.$$

对任意的正实数 β 和向量 $x \in \mathbb{R}^n$, 用 $e(x, \beta)$ 表示关于 $F(\cdot)$ 的余差函数, 即,

$$e(x, \beta) = x - P_\Omega[x - \beta F(x)],$$

则有如下引理, 其相关证明见文献 [36, 37].

引理 2 向量 x^* 是变分不等式问题 (1.1) 的解, 当且仅当 $e(x^*, \beta) = 0$ 对任意的 $\beta > 0$ 都成立.

上面的引理告诉我们, 求解变分不等式问题 (1.1) 等价于找余差函数 $e(x, \beta)$ 的一个零点, 这也为我们设计算法提供了一个停机准则.

定义 1 对于映射 $F: \Omega \rightarrow \mathbb{R}^n$, 若不等式

$$(u - v)^T [F(u) - F(v)] \geq 0, \quad \forall u, v \in \Omega$$

成立, 则称 F 在 Ω 上是单调的; 若存在正常数 η 使得

$$(u - v)^T [F(u) - F(v)] \geq \eta \|u - v\|^2, \quad \forall u, v \in \Omega$$

成立, 则称 F 在 Ω 上是强单调的.

整篇文章中, 我们作如下假设:

假设 1 集合 $\mathcal{X}$ 和 $\mathcal{Y}$ 分别是 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 中的非空闭凸子集, 且映射 $f(x)$ 和 $g(y)$ 分别在 $\mathcal{X}$ 和 $\mathcal{Y}$ 上是连续单调的.

假设 2 变分不等式问题 $\text{MVI}(\mathcal{Z}, Q)$ 的解集 $\mathcal{Z}^*$ 是非空的.

假设 3 对任意给定的 (y, λ) , 子问题 (1.6) 总是可解的.

以上假设也是传统交替方向法的基本假设, 如文献 [3, 24, 32, 33]. 因此, 新方法的假设并没有缩小其应用范围. 另外, 由于 f 和 g 是单调的, 且约束集合 $\mathcal{X}$ 和 $\mathcal{Y}$ 是闭凸集, 则变分不等式问题 $\text{MVI}(\mathcal{Z}, Q)$ 的解集 $\mathcal{Z}^*$ 也是闭凸集.

3 算法及其主要性质

基于第 1 节的讨论, 我们现在正式提出一种新的交替方向法求解变分不等式问题 $\text{MVI}(\mathcal{Z}, Q)$.

算法 1 新的交替方向法

步 0 给定 $\gamma \in (0, 2)$, $\beta \in (0, 1/\|B\|)$, $\epsilon > 0$; 任取初始对 $(x^0, y^0, \lambda^0) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^L$; 令 $k := 0$.

步 1 找一点 $\tilde{x}^k \in \mathcal{X}$ 使得如下不等式成立:

$$(x' - \tilde{x}^k)^T \{f(\tilde{x}^k) - A^T[\lambda^k - \beta(A\tilde{x}^k + By^k - b)]\} \geq 0, \quad \forall x' \in \mathcal{X}. \quad (3.1)$$

步 2 令

$$e_y(y^k, \beta) := y^k - P_{\mathcal{Y}}[y^k - \beta(g(y^k) - B^T \bar{\lambda}^k)], \quad (3.2)$$

及

$$\bar{\lambda}^k := \lambda^k - \beta(Ax^k + By^k - b). \quad (3.3)$$

定义步长 α_k 为

$$\alpha_k := \frac{\varphi(z^k, \beta)}{\psi(z^k, \beta)}, \quad (3.4)$$

这里 $\varphi(z^k, \beta)$ 与 $\psi(z^k, \beta)$ 分别为

$$\begin{aligned} \varphi(z^k, \beta) &:= \|e_y(y^k, \beta)\|^2 + \beta^2 \|A(x^k - \tilde{x}^k)\|^2 + \|\lambda^k - \bar{\lambda}^k\|^2 \\ &\quad - \beta(\lambda^k - \bar{\lambda}^k)^T [A(x^k - \tilde{x}^k) + Be_y(y^k, \beta)], \end{aligned} \quad (3.5)$$

和

$$\psi(z^k, \beta) := \|e_y(y^k, \beta)\|^2 + \beta^2 \|A\tilde{x}^k + By^k - b - Be_y(y^k, \beta)\|^2 + \beta^4 \|A(x^k - \tilde{x}^k)\|^2. \quad (3.6)$$

步 3 如果 $\max\{\|A(x^k - \tilde{x}^k)\|, \|e_y(y^k, \beta)\|, \|A\tilde{x}^k + By^k - b\|\} \leq \epsilon$, 则算法终止; 否则, 令 $k := k + 1$, 转步 4.

步 4 通过下式更新 x^{k+1} :

$$x^{k+1} = x^k - \gamma \alpha_k \beta^2 (x^k - \tilde{x}^k). \quad (3.7)$$

步 5 求解下列非线性方程组得到 y^{k+1} :

$$y^{k+1} + \beta g(y^{k+1}) = y^k + \beta g(y^k) - \gamma \alpha_k e_y(y^k, \beta). \quad (3.8)$$

步 6 最后, 更新乘子

$$\lambda^{k+1} = \lambda^k - \gamma \alpha_k \beta [A\tilde{x}^k + By^k - b - Be_y(y^k, \beta)]. \quad (3.9)$$

说明 1 由式 (3.4)–(3.6) 定义的步长 α_k 是用来更新 x^{k+1} , y^{k+1} 与 λ^{k+1} 的共同步长, 它保证了新算法的全局收敛性.

说明 2 上述算法在每次迭代中, 主要有两个步骤: 步 1 和步 5. 由假设 3, 子问题 (3.1) 与原始的交替方向法是相同的, 也是可求解的. 在步 5 中, 我们需要解一个非线性方程组. 由于子问题中的映射 $I + \beta g$ 是强单调的, 从而这个子问题也是易于求解的.

首先, 我们给出下面一个引理来解释选取步 3 中停机准则的理由.

引理 3 若 $A\tilde{x}^k + By^k - b = 0$, $A\tilde{x}^k = Ax^k$ 及 $e_y(y^k, \beta) = 0$, 则 $(\tilde{x}^k, y^k, \lambda^k)$ 是 $\text{MVI}(\mathcal{Z}, Q)$ 的一个解.

证明 由命题中的前两个等式可知: $Ax^k + By^k - b = 0$. 再由 (3.3) 式有 $\lambda^k = \bar{\lambda}^k$. 根据式 (3.2) 及假设 $e_y(y^k, \beta) = 0$ 可得

$$y^k = P_{\mathcal{Y}}[y^k - \beta(g(y^k) - B^T \lambda^k)] \in \mathcal{Y}.$$

由此得到 $(\tilde{x}^k, y^k, \lambda^k) \in \mathcal{Z}$. 令 $\Omega := \mathcal{Y}$ 和 $u := y^k - \beta(g(y^k) - B^T \lambda^k)$, 并代入 (2.1) 中, 整理可得

$$(y' - y^k)^T (g(y^k) - B^T \lambda^k) \geq 0, \quad \forall y' \in \mathcal{Y}.$$

综上, 可以得到如下结论:

$$\begin{cases} Ax^k = A\tilde{x}^k, \\ e_y(y^k, \beta) = 0, \\ A\tilde{x}^k + By^k = b, \end{cases} \implies \begin{cases} (x' - \tilde{x}^k)^T (f(\tilde{x}^k) - A^T \lambda^k) \geq 0, \quad \forall x' \in \mathcal{X}, \\ (y' - y^k)^T (g(y^k) - B^T \lambda^k) \geq 0, \quad \forall y' \in \mathcal{Y}, \\ A\tilde{x}^k + By^k = b. \end{cases}$$

从而问题 $\text{MVI}(\mathcal{Z}, Q)$ 中所有条件都满足. 此即说明, $(\tilde{x}^k, y^k, \lambda^k)$ 是问题 $\text{MVI}(\mathcal{Z}, Q)$ 的一个解. 证毕. $\square$

在算法步 2 中参数 α_k 可以看作迭代步长, 因此在每次迭代中必须保证 α_k 是恒正的. 显然, 式 (3.5) 中定义的 $\varphi(z^k, \beta)$ 也应该满足非负的条件. 下面的引理即可保证此条件的成立.

引理 4 在每次迭代中, 式 (3.5) 中所定义的函数 $\varphi(z^k, \beta)$ 满足

$$\varphi(z^k, \beta) \geq c(\|e_y(y^k, \beta)\|^2 + \beta^2 \|A(x^k - \tilde{x}^k)\|^2 + \beta^2 \|Ax^k + By^k - b\|^2),$$

其中

$$c := 1 - \frac{\sqrt{2}}{2}. \quad (3.10)$$

证明 利用 Cauchy-Schwarz 不等式, 我们有

$$-[A(x^k - \tilde{x}^k)]^T (Ax^k + By^k - b) \geq -\frac{1}{2} \left(\sqrt{2} \|A(x^k - \tilde{x}^k)\|^2 + \frac{1}{\sqrt{2}} \|Ax^k + By^k - b\|^2 \right)$$

及

$$\begin{aligned} -(Ax^k + By^k - b)^T Be_y(y^k, \beta) &\geq -\frac{1}{2} \left(\sqrt{2} \|Be_y(y^k, \beta)\|^2 + \frac{1}{\sqrt{2}} \|Ax^k + By^k - b\|^2 \right) \\ &\geq -\frac{1}{2} \left(\sqrt{2} \|B\|^2 \|e_y(y^k, \beta)\|^2 + \frac{1}{\sqrt{2}} \|Ax^k + By^k - b\|^2 \right). \end{aligned}$$

将上面的两个不等式代入式 (3.5) 中, 利用步 0 中的假设 $\beta\|B\| \leq 1$, 我们可以推得

$$\begin{aligned} \varphi(z^k, \beta) &\geq \left(1 - \frac{\sqrt{2}}{2} \beta^2 \|B\|^2\right) \|e_y(y^k, \beta)\|^2 + \left(1 - \frac{\sqrt{2}}{2}\right) \beta^2 (\|A(x^k - \tilde{x}^k)\|^2 + \|Ax^k + By^k - b\|^2) \\ &\geq \left(1 - \frac{\sqrt{2}}{2}\right) (\|e_y(y^k, \beta)\|^2 + \beta^2 \|A(x^k - \tilde{x}^k)\|^2 + \beta^2 \|Ax^k + By^k - b\|^2). \end{aligned}$$

证毕. □

对于 $\mathbb{R}^n$ 中的任意向量 a 和 b , 都有 $2a^T b \leq \|a\|^2 + \|b\|^2$. 因此,

$$\begin{aligned} &\|A\tilde{x}^k + By^k - b - Be_y(y^k, \beta)\|^2 \\ &= \|Ax^k + By^k - b - A(x^k - \tilde{x}^k) - Be_y(y^k, \beta)\|^2 \\ &= \|Ax^k + By^k - b\|^2 + \|A(x^k - \tilde{x}^k)\|^2 + \|Be_y(y^k, \beta)\|^2 + 2[A(x^k - \tilde{x}^k)]^T Be_y(y^k, \beta) \\ &\quad - 2(Ax^k + By^k - b)^T [A(x^k - \tilde{x}^k)] - 2(Ax^k + By^k - b)^T Be_y(y^k, \beta) \\ &\leq 3(\|Ax^k + By^k - b\|^2 + \|A(x^k - \tilde{x}^k)\|^2 + \|Be_y(y^k, \beta)\|^2). \end{aligned}$$

结合 (3.6) 式, 有

$$\begin{aligned} \psi(z^k, \beta) &= \beta^2 \|A\tilde{x}^k + By^k - b - Be_y(y^k, \beta)\|^2 + \|e_y(y^k, \beta)\|^2 + \beta^4 \|A(x^k - \tilde{x}^k)\|^2 \\ &\leq 3\beta^2 \|Ax^k + By^k - b\|^2 + (3\beta^2 + \beta^4) \|A(x^k - \tilde{x}^k)\|^2 + (3\beta^2 \|B\|^2 + 1) \|e_y(y^k, \beta)\|^2 \\ &\leq C(\beta^2 \|Ax^k + By^k - b\|^2 + \beta^2 \|A(x^k - \tilde{x}^k)\|^2 + \|e_y(y^k, \beta)\|^2), \end{aligned} \quad (3.11)$$

其中

$$C := \max \{3 + \beta^2, 3\beta^2 \|B\|^2 + 1\}. \quad (3.12)$$

联立 (3.11) 和引理 4, 我们有如下结论:

引理 5 在任意的第 k ($k > 0$) 次迭代中, 有 $\alpha_k \geq c/C > 0$, 其中常数 c 与 C 分别由 (3.10) 式和 (3.12) 式所定义.

4 收敛性分析

本节中, 在第 2 节中的假设条件下, 我们证明所提出的新算法是全局收敛的. 为了方便后面的分析, 我们记

$$\Phi(y, \beta) = y + \beta g(y).$$

引理 6 对给定的 (y^k, λ^k) , 若 $\tilde{x}^k$ 是变分不等式问题 (3.1) 的解, 则对问题 $\text{MVI}(\mathcal{Z}, Q)$ 的任意解点 (x^*, y^*, λ^*) , 都有

$$(\lambda^k - \lambda^*)^T (A\tilde{x}^k + By^k - b) \geq \beta (A\tilde{x}^k - Ax^*)^T (A\tilde{x}^k + By^k - b) + (\lambda^k - \lambda^*)^T (By^k - By^*).$$

证明 由 (x^*, y^*, λ^*) 是变分不等式问题 (1.4)–(1.5) 的解, 以及 $\tilde{x}^k \in \mathcal{X}$, 可得

$$(\tilde{x}^k - x^*)^T (f(x^*) - A^T \lambda^*) \geq 0.$$

另一方面, 由于 $\tilde{x}^k$ 满足 (3.1) 式, 并且 $x^* \in \mathcal{X}$, 可知

$$(x^* - \tilde{x}^k)^T \{f(\tilde{x}^k) - A^T [\lambda^k - \beta(A\tilde{x}^k + By^k - b)]\} \geq 0.$$

将上述两个不等式相加, 我们推得

$$(\tilde{x}^k - x^*)^T \{f(x^*) - f(\tilde{x}^k)\} - A^T (\lambda^* - \lambda^k) - \beta A^T (A\tilde{x}^k + By^k - b) \geq 0.$$

利用 f 的单调性, 将上式重新整理得

$$(\lambda^k - \lambda^*)^T (A\tilde{x}^k - Ax^*) \geq \beta(A\tilde{x}^k - Ax^*)^T (A\tilde{x}^k + By^k - b).$$

结合条件 $Ax^* + By^* = b$, 及上述不等式, 即得

$$(\lambda^k - \lambda^*)^T (A\tilde{x}^k + By^k - b) \geq \beta(A\tilde{x}^k - Ax^*)^T (A\tilde{x}^k + By^k - b) + (\lambda^k - \lambda^*)^T (By^k - By^*).$$

证毕. □

引理 7 若 $\{(x^k, y^k, \lambda^k)\}$ 是由算法 1 产生的迭代序列, 且 (x^*, y^*, λ^*) 是问题 $\text{MVI}(\mathcal{Z}, Q)$ 的任一解点, 则

$$\begin{aligned} & [\Phi(y^k, \beta) - \Phi(y^*, \beta)]^T e_y(y^k, \beta) \\ & \geq \|e_y(y^k, \beta)\|^2 - \beta(\lambda^k - \lambda^*)^T (By^k - By^*) + \beta(\lambda^k - \lambda^*)^T B e_y(y^k, \beta) \\ & \quad + \beta(\lambda^k - \bar{\lambda}^k)^T (By^k - By^*) - \beta(\lambda^k - \bar{\lambda}^k)^T B e_y(y^k, \beta). \end{aligned}$$

证明 在不等式 (2.1) 中, 令 $v := y^k - \beta[g(y^k) - B^T \bar{\lambda}^k]$, 及 $u := y^* \in \mathcal{Y}$, 整理可得

$$\{e_y(y^k, \beta) - \beta[g(y^k) - B^T \bar{\lambda}^k]\}^T [y^k - y^* - e_y(y^k, \beta)] \geq 0.$$

另一方面, 由于 (x^*, y^*, λ^*) 是问题 $\text{MVI}(\mathcal{Z}, Q)$ 的解, 以及 $P_{\mathcal{Y}}[y^k - \beta(g(y^k) - B^T \bar{\lambda}^k)] \in \mathcal{Y}$, 可推得

$$\beta[g(y^*) - B^T \lambda^*]^T [y^k - y^* - e_y(y^k, \beta)] \geq 0.$$

将以上两式相加, 利用 g 的单调性整理得

$$\begin{aligned} & [\Phi(y^k, \beta) - \Phi(y^*, \beta)]^T e_y(y^k, \beta) \\ & \geq \|e_y(y^k, \beta)\|^2 - \beta(\lambda^k - \lambda^*)^T (By^k - By^*) + \beta(y^k - y^*)^T [g(y^k) - g(y^*)] \\ & \quad + \beta(\lambda^k - \lambda^*)^T B e_y(y^k, \beta) + \beta(\lambda^k - \bar{\lambda}^k)^T (By^k - By^*) - \beta(\lambda^k - \bar{\lambda}^k)^T B e_y(y^k, \beta) \\ & \geq \|e_y(y^k, \beta)\|^2 - \beta(\lambda^k - \lambda^*)^T (By^k - By^*) + \beta(\lambda^k - \lambda^*)^T B e_y(y^k, \beta) \\ & \quad + \beta(\lambda^k - \bar{\lambda}^k)^T (By^k - By^*) - \beta(\lambda^k - \bar{\lambda}^k)^T B e_y(y^k, \beta), \end{aligned}$$

证毕. □

引理 8 对问题 $\text{MVI}(\mathcal{Z}, Q)$ 的任一解 (x^*, y^*, λ^*) , 下列不等式成立:

$$\begin{aligned} & \|\lambda^{k+1} - \lambda^*\|^2 + \|\Phi(y^{k+1}, \beta) - \Phi(y^*, \beta)\|^2 + \|A(x^{k+1} - x^*)\|^2 \\ & \leq \|\lambda^k - \lambda^*\|^2 + \|\Phi(y^k, \beta) - \Phi(y^*, \beta)\|^2 + \|A(x^k - x^*)\|^2 - \gamma(2 - \gamma)\alpha_k\varphi(z^k, \beta). \end{aligned} \quad (4.1)$$

证明 首先, 由 (3.9) 式和引理 6 可推得

$$\begin{aligned} \|\lambda^{k+1} - \lambda^*\|^2 &= \|\lambda^k - \lambda^* - \gamma\alpha_k\beta(A\tilde{x}^k + By^k - b - Be_y(y^k, \beta))\|^2 \\ &= \|\lambda^k - \lambda^*\|^2 + \gamma^2\alpha_k^2\beta^2\|A\tilde{x}^k + By^k - b - Be_y(y^k, \beta)\|^2 \\ &\quad - 2\gamma\alpha_k\beta(\lambda^k - \lambda^*)^T[A\tilde{x}^k + By^k - b - Be_y(y^k, \beta)] \\ &\leq \|\lambda^k - \lambda^*\|^2 + \gamma^2\alpha_k^2\beta^2\|A\tilde{x}^k + By^k - b - Be_y(y^k, \beta)\|^2 - 2\gamma\alpha_k\beta(\lambda^k - \lambda^*)^T(By^k - By^*) \\ &\quad + 2\gamma\alpha_k\beta(\lambda^k - \lambda^*)^TBe_y(y^k, \beta) - 2\gamma\alpha_k\beta^2(A\tilde{x}^k - Ax^*)^T(A\tilde{x}^k + By^k - b). \end{aligned} \quad (4.2)$$

类似地, 利用 (3.8) 式和引理 7 可以得到

$$\begin{aligned} & \|\Phi(y^{k+1}, \beta) - \Phi(y^*, \beta)\|^2 \\ &= \|\Phi(y^k, \beta) - \Phi(y^*, \beta) - \gamma\alpha_k e_y(y^k, \beta)\|^2 \\ &= \|\Phi(y^k, \beta) - \Phi(y^*, \beta)\|^2 - 2\gamma\alpha_k[\Phi(y^k, \beta) - \Phi(y^*, \beta)]^T e_y(y^k, \beta) + \gamma^2\alpha_k^2\|e_y(y^k, \beta)\|^2 \\ &\leq \|\Phi(y^k, \beta) - \Phi(y^*, \beta)\|^2 + \gamma^2\alpha_k^2\|e_y(y^k, \beta)\|^2 - 2\gamma\alpha_k\|e_y(y^k, \beta)\|^2 \\ &\quad + 2\gamma\alpha_k\beta(\lambda^k - \lambda^*)^T(By^k - By^*) - 2\gamma\alpha_k\beta(\lambda^k - \lambda^*)^TBe_y(y^k, \beta) \\ &\quad - 2\gamma\alpha_k\beta(\lambda^k - \bar{\lambda}^k)^T(By^k - By^*) + 2\gamma\alpha_k\beta(\lambda^k - \bar{\lambda}^k)^TBe_y(y^k, \beta). \end{aligned} \quad (4.3)$$

同样, 由 (3.7) 式可得

$$\begin{aligned} \|A(x^{k+1} - x^*)\|^2 &= \|A(x^k - x^*) - \gamma\alpha_k\beta^2 A(x^k - \tilde{x}^k)\|^2 \\ &= \|A(x^k - x^*)\|^2 - 2\gamma\alpha_k\beta^2[A(x^k - x^*)]^T[A(x^k - \tilde{x}^k)] + \gamma^2\alpha_k^2\beta^4\|A(x^k - \tilde{x}^k)\|^2 \\ &= \|A(x^k - x^*)\|^2 - 2\gamma\alpha_k\beta^2[A(\tilde{x}^k - x^*)]^T[A(x^k - \tilde{x}^k)] + \gamma^2\alpha_k^2\beta^4\|A(x^k - \tilde{x}^k)\|^2 \\ &\quad - 2\gamma\alpha_k\beta^2\|A(x^k - \tilde{x}^k)\|^2. \end{aligned} \quad (4.4)$$

将式 (4.2)–(4.4) 相加, 整理得

$$\begin{aligned} & \|\lambda^{k+1} - \lambda^*\|^2 + \|\Phi(y^{k+1}, \beta) - \Phi(y^*, \beta)\|^2 + \|A(x^{k+1} - x^*)\|^2 \\ & \leq \|\lambda^k - \lambda^*\|^2 + \|\Phi(y^k, \beta) - \Phi(y^*, \beta)\|^2 + \|A(x^k - x^*)\|^2 \\ & \quad + \gamma^2\alpha_k^2\beta^2\|A\tilde{x}^k + By^k - b - Be_y(y^k, \beta)\|^2 + \gamma^2\alpha_k^2\|e_y(y^k, \beta)\|^2 + \gamma^2\alpha_k^2\beta^4\|A(x^k - \tilde{x}^k)\|^2 \\ & \quad - 2\gamma\alpha_k\|e_y(y^k, \beta)\|^2 - 2\gamma\alpha_k\beta^2\|Ax^k - A\tilde{x}^k\|^2 - 2\gamma\alpha_k\beta(A\tilde{x}^k - Ax^*)^T(\lambda^k - \bar{\lambda}^k) \\ & \quad - 2\gamma\alpha_k\beta(\lambda^k - \bar{\lambda}^k)^T(By^k - By^*) + 2\gamma\alpha_k\beta(\lambda^k - \bar{\lambda}^k)^TBe_y(y^k, \beta) \\ & = \|\lambda^k - \lambda^*\|^2 + \|\Phi(y^k, \beta) - \Phi(y^*, \beta)\|^2 + \|A(x^k - x^*)\|^2 \\ & \quad + \gamma^2\alpha_k^2\beta^2\|A\tilde{x}^k + By^k - b - Be_y(y^k, \beta)\|^2 + \gamma^2\alpha_k^2\|e_y(y^k, \beta)\|^2 + \gamma^2\alpha_k^2\beta^4\|A(x^k - \tilde{x}^k)\|^2 \\ & \quad - 2\gamma\alpha_k\|e_y(y^k, \beta)\|^2 - 2\gamma\alpha_k\beta^2\|Ax^k - A\tilde{x}^k\|^2 - 2\gamma\alpha_k\beta(\lambda^k - \bar{\lambda}^k)^T(A\tilde{x}^k + By^k - b) \\ & \quad + 2\gamma\alpha_k\beta(\lambda^k - \bar{\lambda}^k)^TBe_y(y^k, \beta) \end{aligned}$$

$$\begin{aligned}
&= \|\lambda^k - \lambda^*\|^2 + \|\Phi(y^k, \beta) - \Phi(y^*, \beta)\|^2 + \|A(x^k - x^*)\|^2 - 2\gamma\alpha_k\varphi(z^k, \beta) + \gamma^2\alpha_k^2\psi(z^k, \beta) \\
&= \|\lambda^k - \lambda^*\|^2 + \|\Phi(y^k, \beta) - \Phi(y^*, \beta)\|^2 + \|A(x^k - x^*)\|^2 - \gamma(2 - \gamma)\alpha_k\varphi(z^k, \beta),
\end{aligned}$$

上面的第一个等式根据条件 $Ax^* + By^* = b$ 得到, 最后一个等式利用 α_k 的定义得到. 证毕. $\square$

接下来给出本文方法的收敛性定理.

定理 1 假设算法 1 产生的无穷迭代序列为 $\{(x^k, y^k, \lambda^k)\}$, 则整个序列必定收敛到变分不等式问题 $\text{MVI}(\mathcal{Z}, Q)$ 的一个解.

证明 在引理 8 的不等式 (4.1) 中, 由 $\varphi(z^k, \beta)$ 和 α_k 的非负性及 $\gamma \in (0, 2)$ 可以推出

$$\begin{aligned}
&\|\lambda^{k+1} - \lambda^*\|^2 + \|\Phi(y^{k+1}, \beta) - \Phi(y^*, \beta)\|^2 + \|A(x^{k+1} - x^*)\|^2 \\
&\leq \|\lambda^k - \lambda^*\|^2 + \|\Phi(y^k, \beta) - \Phi(y^*, \beta)\|^2 + \|A(x^k - x^*)\|^2 \\
&\leq \dots \\
&< +\infty.
\end{aligned} \tag{4.5}$$

因此, 序列 $\{(Ax^k, y^k, \lambda^k)\}$ 是有界的. 对 (4.1) 式进行移项整理, 有

$$\begin{aligned}
\gamma(2 - \gamma)\alpha_k\varphi(z^k, \beta) &\leq \|\lambda^k - \lambda^*\|^2 + \|\Phi(y^k, \beta) - \Phi(y^*, \beta)\|^2 + \|A(x^k - x^*)\|^2 \\
&\quad - \|\lambda^{k+1} - \lambda^*\|^2 - \|\Phi(y^{k+1}, \beta) - \Phi(y^*, \beta)\|^2 - \|A(x^{k+1} - x^*)\|^2.
\end{aligned}$$

对上述不等式两边进行累加, 可得

$$\sum_{k=0}^{\infty} \gamma(2 - \gamma)\alpha_k\varphi(z^k, \beta) \leq \|\lambda^0 - \lambda^*\|^2 + \|\Phi(y^0, \beta) - \Phi(y^*, \beta)\|^2 + \|A(x^0 - x^*)\|^2 < +\infty.$$

结合 $\alpha_k \geq c/C > 0$, 由上面的不等式可知 $\lim_{k \rightarrow \infty} \varphi(z^k, \beta) = 0$ 成立. 从而, 根据引理 4 可得下列结论,

$$\lim_{k \rightarrow \infty} \|e_y(y^k, \beta)\| = 0, \quad \lim_{k \rightarrow \infty} \|Ax^k - A\tilde{x}^k\| = 0 \quad \text{及} \quad \lim_{k \rightarrow \infty} \|Ax^k + By^k - b\| = 0. \tag{4.6}$$

由于 $\{(Ax^k, y^k, \lambda^k)\}$ 是有界的, 则至少有一个聚点, 记为 $(Ax^\infty, y^\infty, \lambda^\infty)$. 令其一子列 $\{(Ax^{k_j}, y^{k_j}, \lambda^{k_j})\}$ 收敛于 $\{(Ax^\infty, y^\infty, \lambda^\infty)\}$. 从而, 对 (4.6) 式中子列取极限可得

$$\begin{aligned}
\|e_y(y^\infty, \beta)\| &= \left\| e_y\left(\lim_{j \rightarrow \infty} y^{k_j}, \beta\right) \right\| = \lim_{j \rightarrow \infty} \|e_y(y^{k_j}, \beta)\| = 0, \\
\|Ax^\infty - A\tilde{x}^\infty\| &= 0 \quad \text{及} \quad \|A\tilde{x}^\infty + By^\infty - b\| = 0,
\end{aligned}$$

所以由引理 3 可知, $(\tilde{x}^\infty, y^\infty, \lambda^\infty)$ 是问题 $\text{MVI}(\mathcal{Z}, Q)$ 的解. 因为 (x^*, y^*, λ^*) 是任意的解, 故在上述分析中可用 $(\tilde{x}^\infty, y^\infty, \lambda^\infty)$ 来代替. 特别地, 根据 (4.5) 式可得

$$\begin{aligned}
&\|\lambda^{k+1} - \lambda^\infty\|^2 + \|\Phi(y^{k+1}, \beta) - \Phi(y^\infty, \beta)\|^2 + \|Ax^{k+1} - A\tilde{x}^\infty\|^2 \\
&\leq \|\lambda^k - \lambda^\infty\|^2 + \|\Phi(y^k, \beta) - \Phi(y^\infty, \beta)\|^2 + \|Ax^k - A\tilde{x}^\infty\|^2,
\end{aligned}$$

即

$$\left\| \begin{array}{c} \lambda^{k+1} - \lambda^\infty \\ \Phi(y^{k+1}, \beta) - \Phi(y^\infty, \beta) \\ Ax^{k+1} - A\tilde{x}^\infty \end{array} \right\| \leq \left\| \begin{array}{c} \lambda^k - \lambda^\infty \\ \Phi(y^k, \beta) - \Phi(y^\infty, \beta) \\ Ax^k - A\tilde{x}^\infty \end{array} \right\|.$$

下面我们将证明由算法 1 产生的迭代序列 $\{(Ax^k, y^k, \lambda^k)\}$ 收敛到唯一的聚点 $(Ax^\infty, y^\infty, \lambda^\infty)$. 假设 $(A\hat{x}, \hat{y}, \hat{\lambda})$ 是子列 $\{(Ax^{k_j}, y^{k_j}, \lambda^{k_j})\}$ 的另一聚点, 则一定存在一个 $K^* > 0$, 使得对任意的 $k \geq K^*$, 都有

$$\left\| \begin{array}{c} \lambda^{k+1} - \hat{\lambda} \\ \Phi(y^{k+1}, \beta) - \Phi(\hat{y}, \beta) \\ Ax^{k+1} - A\hat{x} \end{array} \right\| \leq \varsigma/2, \quad \text{其中} \quad \varsigma := \left\| \begin{array}{c} \lambda^\infty - \hat{\lambda} \\ \Phi(y^\infty, \beta) - \Phi(\hat{y}, \beta) \\ Ax^\infty - A\hat{x} \end{array} \right\| > 0.$$

然后利用三角不等式, 可得如下结论

$$\begin{aligned} \left\| \begin{array}{c} \lambda^{k+1} - \lambda^\infty \\ \Phi(y^{k+1}, \beta) - \Phi(y^\infty, \beta) \\ Ax^{k+1} - Ax^\infty \end{array} \right\| &\geq \left\| \begin{array}{c} \lambda^\infty - \hat{\lambda} \\ \Phi(y^\infty, \beta) - \Phi(\hat{y}, \beta) \\ Ax^\infty - A\hat{x} \end{array} \right\| - \left\| \begin{array}{c} \lambda^{k+1} - \hat{\lambda} \\ \Phi(y^{k+1}, \beta) - \Phi(\hat{y}, \beta) \\ Ax^{k+1} - A\hat{x} \end{array} \right\| \\ &\geq \varsigma - \frac{\varsigma}{2} = \frac{\varsigma}{2} > 0, \end{aligned}$$

这说明 $(Ax^\infty, y^\infty, \lambda^\infty)$ 不是序列的聚点, 这与假设矛盾. 因此, 序列 $\{(Ax^k, y^k, \lambda^k)\}$ 存在唯一聚点并且收敛到变分不等式问题 $\text{MVI}(\mathcal{Z}, Q)$ 的一个解. 证毕. $\square$

5 数值实验

本节, 为了考察新算法 (记为 “NADM”) 的数值表现, 我们用 Matlab 2008b 进行编程实现. 同时, 我们将新算法与文献 [25] 和 [32] 中的算法进行比较, 且分别将两种比较的方法简记为 “LPPA” 和 “VADM”. 所有程序均在 Windows XP 系统下, 配置为 CPU Intel-Pentium Dual-Core 2.6GHz, 内存 2GB 的 HP 台式机上运行. 在数据结果中, 我们用 “Its.” 表示迭代次数, “CPU(s)” 表示计算机的 CPU 运行时间.

例 1 我们考虑如下优化问题

$$\min \{c^T x \mid x \in \Omega_1 \cap \Omega_2\}, \quad (5.1)$$

其中

$$\Omega_1 = \{x \in \mathbb{R}^n \mid \|x\| \leq r_1\} \quad \text{和} \quad \Omega_2 = \{x \in \mathbb{R}^n \mid \|x - b\| \leq r_2\}. \quad (5.2)$$

为了保证问题的可行性, 应该满足条件 $\|b\| \leq r_1 + r_2$. 因此, 在我们的数值试验中, 我们选取 $r_1 = 0.5\|b\|$ 及 $r_2 = 0.6\|b\|$. 类似文献 [32] 中的构造方法: 向量 c 和 b 的元素分别是随机分布在 $(-K, K)$ 和 $(0, 10)$ 之间的数.

通过引入辅助向量 y , 将问题 (5.1)–(5.2) 转化为如下等价形式:

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

其中 $\odot_r$ 表示圆心在原点半径为 r 的圆.

上述凸规划问题可转化为如下变分不等式: 找一点 $u^* \in \Omega$, 使得

$$(u - u^*)^T F(u^*) \geq 0, \quad \forall u \in \Omega,$$

其中

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, \quad F(u) = \begin{pmatrix} c \\ 0 \end{pmatrix},$$

以及

$$\Omega = \{ (x, y) \mid x \in \odot_{r_1}, y \in \odot_{r_2}, x + y - b = 0 \}.$$

在新算法中, 子问题 (3.1) 等价于寻找点 $\tilde{x}^k \in \odot_{r_1}$ 使得

$$(x' - \tilde{x}^k)^T \{ \beta \tilde{x}^k - [\lambda^k - \beta(y^k - b) - c] \} \geq 0, \quad \forall x' \in \odot_{r_1}.$$

令 $q^k := \lambda^k - \beta(y^k - b) - c$, 则上面变分不等式问题的解可显式表示为

$$\tilde{x}^k = \begin{cases} q^k / \beta, & \text{若 } \|q^k\| \leq \beta r_1, \\ r_1 q^k / \|q^k\|, & \text{否则.} \end{cases}$$

另外, 由于 $g(y) = 0$, 则新算法中的子问题 (3.8) 可用如下迭代格式显示求解,

$$y^{k+1} = y^k - \gamma \alpha_k e(y^k, \beta).$$

在数值试验中, 我们取初始迭代点 (x^0, y^0, λ^0) 为全为 0 的向量, 设 $K = 10$, $\gamma = 1.8$, 取参数 $\beta = 0.8$ 以保证 $\beta\|B\| = 0.8 < 1$ 恒成立 (注意: $\|B\| = 1$). 设停机精度为 $\epsilon = 10^{-8}$.

对于例 1, 我们分别用三种方法测试了问题维数从 10 到 10^6 的不同情形. 其中, 表 1 列出了文献 [25] 的方法 “LPPA” 和新方法 “NADM” 之间的数值比较结果, 而表 2 列出了文献 [32] 的方法 “VADM” 和新方法 “NADM” 之间的数值比较结果. 数值结果表明, 对于不同维数的问题, 三种算法都是比较稳定的. 从表 1 发现, 新的算法较之于算法 “LPPA”, 大部分情况下, 迭代次数和 CPU 运行时间都要少. 而相对于方法 “VADM” 而言, 新方法只优于 “VADM” 中的策略 2 方法. 主要原因有两点: 第一, 我们测试的问题, 对于 “VADM” 都有显示解, 这样 “VADM” 也可以很快找到问题的解; 第二, 新方法使用的是固定罚参数 β , 而 “VADM” 是专门针对固定罚因子 β 提出的改进方法. 这也提醒我们在未来的研究工作中要摒弃这一缺点, 提出更有效的方法. 另外, 为了直观地了解新算法的收敛性态, 图 1 给出了当 $n = 2$ 时, 新方法的迭代点随着迭代次数变化的模拟图, 且从图中容易看出新算法的收敛速度很快, 在 50 次迭代以后变化很小.

如前面引言中所述, 参数 β 对交替方向法的影响较大. 因此, 下面我们针对 $n = 10^5$ 的情形, 在 $(0, 1)$ 区间中随机产生 10 个不同的初始迭代点, 然后测试不同的罚参数 β 对新算法的影响. 相应的结果见图 2, 且从该图中我们发现新算法在迭代次数和 CPU 时间上还是相对稳定的.

例 2 考虑如下矩阵优化问题

$$\min \left\{ \frac{1}{2} \|X - C\|_F^2 \mid X \in S_+^n \cap \mathcal{B} \right\}, \quad (5.3)$$

其中 C 是给定的 $n \times n$ 阶对称矩阵, $\|\cdot\|_F$ 表示 Fröbenius 范数, 即 $\|C\|_F := \sqrt{\sum_{i,j=1}^n C_{ij}^2}$. 集合 S_+^n 表示对称半正定锥

$$S_+^n = \{H \in \mathbb{R}^{n \times n} \mid H^T = H, H \succeq 0\},$$

表 1 文献 [25] 的方法 “LPPA” 与本文方法 “NADM” 求解例 1 的数值比较结果

方法	不同迭代策略“LPPA”(l=2)						NADM	
	策略 1		策略 2		策略 3			
	n	Its.	CPU(s)	Its.	CPU(s)	Its.	CPU(s)	Its.
10 ¹	127	0.010	58	0.002	82	0.002	62	0.001
10 ²	167	0.005	82	0.003	84	0.003	76	0.001
10 ³	161	0.015	79	0.006	84	0.007	74	0.031
10 ⁴	163	0.091	79	0.041	84	0.043	74	0.047
10 ⁵	163	2.934	79	1.431	84	1.434	74	1.094
10 ⁶	163	30.266	79	14.127	84	14.384	74	10.516

表 2 文献 [32] 的方法 “VADM” 与本文方法 “NADM” 求解例 1 的数值比较结果

方法	不同罚参数调节策略的 “VADM”						NADM	
	策略 1		策略 2		策略 3			
	n	Its.	CPU(s)	Its.	CPU(s)	Its.	CPU(s)	Its.
10^1	36	0.003	103	0.004	36	0.002	62	0.001
10^2	39	0.002	137	0.006	42	0.002	76	0.001
10^3	38	0.004	141	0.013	38	0.004	74	0.031
10^4	42	0.041	149	0.119	42	0.041	74	0.047
10^5	43	0.621	156	2.247	43	0.620	74	1.094
10^6	44	6.865	164	27.792	44	7.620	74	10.516

图 1 对于例 1, 在 $n = 2$ 时, 本文算法的迭代点随着迭代次数变化的模拟图

及盒子约束

$$\mathcal{B} = \{H \in \mathbb{R}^{n \times n} \mid H^T = H, H_L \leq H \leq H_U\}.$$

这里, H_L 和 H_U 都是给定的 $n \times n$ 阶对称矩阵, $H_L \leq H_U$ 表示 $(H_L)_{ij} \leq (H_U)_{ij}$ ($i, j = 1, \dots, n$).

上述例子是一个带约束的最小二乘协方差矩阵校正问题, 即寻找一个满足约束条件的半正定矩阵

图 2 对于例 1, 在 $n = 10^4$ 时, 不同的参数 β 对迭代次数和 CPU 时间的影响

去逼近一个给定矩阵 C . 此类问题在诸如金融、保险、再保险等领域有广泛应用, 近几年也引起了人们的关注, 详见文献 [38].

类似例 1 的处理方式, 通过引入辅助变量 Y , 问题 (5.3) 可转化为如下等价的凸优化问题:

$$\begin{aligned} \min \quad & \frac{1}{2} \|X - C\|_F^2 + \frac{1}{2} \|Y - C\|_F^2 \\ \text{s.t.} \quad & X - Y = 0, \\ & X \in S_+^n, Y \in \mathcal{B}, \end{aligned}$$

易见, 其等价的变分不等式为

$$\begin{cases} \langle X - X^*, (X^* - C) - Z^* \rangle \geq 0, \\ \langle Y - Y^*, (Y^* - C) + Z^* \rangle \geq 0, \\ X^* - Y^* = 0, \end{cases} \quad \forall (X, Y, Z) \in \mathcal{Z},$$

其中 $\mathcal{Z} = S_+^n \times \mathcal{B} \times \mathbb{R}^{n \times n}$.

由新算法可知, 子问题 (3.1) 等价于寻找点 $\tilde{X}^k$ 使得

$$\langle X' - \tilde{X}^k, (\tilde{X}^k - C) - [Z^k - \beta(\tilde{X}^k - Y^k)] \rangle \geq 0, \quad \forall X' \in S_+^n.$$

易得上述变分不等式的显式解为

$$\tilde{X}^k = P_{S_+^n} \left(\frac{1}{1+\beta} (\beta Y^k + Z^k + C) \right).$$

而非线性方程组 (3.8) 等价于

$$Y^{k+1} + \beta(Y^{k+1} - C) = Y^k + \beta(Y^k - C) - \gamma \alpha_k e_Y(Y^k, \beta);$$

整理可得

$$Y^{k+1} = Y^k - \frac{\gamma \alpha_k}{1+\beta} e_Y(Y^k, \beta),$$

表 3 文献 [25] 的方法 “LPPA” 与本文方法 “NADM” 求解例 2 的数值比较结果

方法	不同迭代策略的 “LPPA” ($l = 7$)						NADM	
	策略 1		策略 2		策略 3			
	n	Its.	CPU(s)	Its.	CPU(s)	Its.	CPU(s)	Its.
10	63	0.011	19	0.003	25	0.004	16	0.003
50	147	0.36	116	0.28	70	0.17	85	0.20
100	215	2.82	215	2.82	170	2.26	146	1.88
200	265	15.06	291	16.41	249	13.90	198	10.69
500	314	204.41	355	231.83	315	204.08	240	154.38
1000	348	1902.28	399	2161.58	361	1996.81	267	1352.44

表 4 文献 [25] 的方法 “LPPA” 与本文方法 “NADM” 求解例 2 的数值比较结果

方法	不同迭代策略的 “LPPA” ($l = 2$)						NADM	
	策略 1		策略 2		策略 3			
	n	Its.	CPU(s)	Its.	CPU(s)	Its.	CPU(s)	Its.
10	63	0.011	13	0.002	29	0.005	16	0.003
50	147	0.37	69	0.17	70	0.17	85	0.20
100	215	2.81	120	1.59	117	1.54	146	1.88
200	265	14.98	157	8.88	153	8.60	198	10.69
500	314	206.26	188	123.96	184	119.41	240	154.38
1000	348	1875.03	209	1078.59	205	1051.55	267	1352.44

其中

$$\begin{aligned}
 e_Y(Y^k, \beta) &= Y^k - P_{\mathcal{B}}(Y^k - \beta[(Y^k - C) + Z^k - \beta(X^k - Y^k)]) \\
 &= Y^k - P_{\mathcal{B}}(\beta^2 X^k + (1 - \beta - \beta^2)Y^k - \beta Z^k + \beta C).
 \end{aligned}$$

在数值试验中, 矩阵 C 的对角元素均随机分布在 $(0, 2)$ 之间, 而非对角元素为 $(-1, 1)$ 之间的随机数. 矩阵 H_U 和 H_L 的对角元素都取为 1, 而它们的非对角元素分别取为 0.1 和 -0.1 . 选择初始点 (X^0, Y^0, Z^0) 都为零矩阵, 参数 $\beta = 0.6$, $\gamma = 1.95$ 以及停机准则设为

$$\epsilon = 10^{-4} \max(\|X^0 - \tilde{X}^0\|_F, \|e_Y(Y^0, \beta)\|_F, \|\tilde{X}^0 - Y^0\|_F).$$

对于例 2, 我们测试了矩阵阶数从 10 到 1000 的不同情形, 相关的数值结果见表 3 和表 4. 从表 3 中发现, 新算法的数值表现要优于 “LPPA” 在 $l = 7$ 的情形下的数值表现; 而从表 4 中可见, 当 $l = 2$ 时, “LPPA” 只有迭代策略 1 要逊色于新方法. 因此, 图 3 列出了当维数 $n = 500$ 时, 参数 l 对 “LPPA” 在迭代次数上的影响情况. 总体而言, 新算法是稳定的, 有效的.

6 结论

本文针对可分离的结构型变分不等式提出了一种新的交替方向法. 与原始的交替方向法相比, 每

图 3 对于例 2, 当 $n = 500$ 时, 不同参数 l 对算法迭代次数的影响

步迭代过程中, 新算法只保留原交替方向法的一个变分不等式子问题, 而另一个变分不等式子问题则用一个强单调的非线性方程组子问题替代. 相比求解一个变分不等式问题, 我们有更多的有效方法求解非线性方程组问题. 这使得算法简单易行且能提高计算的效率. 在比较宽松的假设条件下, 我们证明了新算法的全局收敛性, 给出的数值试验结果也进一步表明了新算法的有效性.

事实上, 在我们的算法中, 还保留有一个变分不等式子问题. 那么, 我们自然会问, 是否可以类似地通过求解一个非线性方程组来避免求解另一个变分不等式子问题呢? 这将是我们未来的研究重点. 另外, 从数值实验中也能发现, 我们固定罚参数 β 会使得新方法逊色于变参数的交替方向法, 这也促使我们在未来的研究中去寻找一种变参数策略来提升新算法的计算能力.

致谢 作者诚挚感谢审稿人对本文的初稿提出的宝贵的修改意见.

参考文献

- 1 Bertsekas D P, Gafni E M. Projection method for variational inequalities with applications to the traffic assignment problem. *Math Programming Study*, 1987, 17: 139–159
- 2 Dafermos S. Traffic equilibrium and variational inequalities. *Transport Sci*, 1980, 14: 42–54
- 3 Fukushima M. Application of the alternating directions method of multipliers to separable convex programming problems. *Comput Optim Appl*, 1992 2: 93–111
- 4 何炳生. 一类广义线性变分不等式的求解与应用. *中国科学 A 辑*, 1995, 38: 939–945
- 5 Nagurney A. *Network Economics, A Variational Inequality Approach*. Dordrecht: Kluwer Academics Publishers, 1993
- 6 Nagurney A, Ramanujam P. Transportation network policy modeling with goal targets and generalized penalty functions. *Transport Sci*, 1996, 30: 3–13
- 7 Nagurney A, Thore S, Pan J. Spatial market policy modeling with goal targets. *Oper Res*, 1996, 44: 393–406
- 8 Chen G, Teboulle M. A proximal-based decomposition method for convex minimization problems. *Math Program*, 1994, 64: 81–101
- 9 Eckstein J, Bertsekas D P. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math Program*, 1992, 55: 293–318
- 10 Eckstein J, Fukushima M. Some reformulation and applications of the alternating directions method of multipliers. In: Hager W W, Hearn D W, Pardalos P M, eds. *Large Scale Optimization: State of the Art*. Dordrecht: Kluwer Academic Publishers, 1994, 115–134
- 11 Glowinski R. *Numerical Methods for Nonlinear Variational Problems*. New York: Springer, 1984
- 12 Glowinski R, Tallec P L. *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*. SIAM Studies in Applied Mathematics. Philadelphia: SIAM, 1989
- 13 Tseng P. Applications of a splitting algorithm to decomposition in convex programming and variational inequalities.

- SIAM J Control Optim, 1991, 29: 119–138
- 14 Tseng P. Alternating projection-proximal methods for convex programming and variational inequalities. SIAM J Optim, 1997, 7: 951–965
- 15 Gabay D. Applications of the method of multipliers to variational inequalities. In: Fortin M, Glowinski R, eds. Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems. Amsterdam: North-Holland, 1983, 299–331
- 16 Gabay D, Mercier B. A dual algorithm for the solution of nonlinear variational problems via finite element approximations. Comput Math Appl, 1976, 2: 17–40
- 17 Fortin M, Glowinski R. Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems. Amsterdam: North-Holland, 1983
- 18 Rockafellar R T. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. Math Oper Res, 1976, 1: 97–116
- 19 Han D R, Sun W Y. New decomposition methods for solving variational inequality problems. Math Comput Model, 2003, 37: 405–418
- 20 Han D R. A proximal decomposition algorithm for variational inequality problems. J Comput Appl Math, 2003, 161: 231–244.
- 21 Han D R, Lo H K. Solving variational inequality problems with linear constraints by a proximal decomposition algorithm. J Global Optim, 2004, 28: 97–113
- 22 Han D R. A hybrid entropic proximal decomposition method with self-adaptive strategy for solving variational inequality problems. Comput Math Appl, 2008, 55: 101–115
- 23 He B S. Parallel splitting augmented Lagrangian methods for monotone structured variational inequalities. Comput Optim Appl, 2009, 42: 195–212
- 24 He B S, Liao L Z, Han D R, Yang H. A new inexact alternating direction method for monotone variational inequalities. Math Program, 2002, 92: 103–118
- 25 He B S, Fu X L, Jiang Z K. Proximal-point algorithm using a linear proximal term. J Optim Theory Appl, 2009, 141: 299–319
- 26 He B S. Inexact implicit methods for monotone general variational inequalities. Math Program, 1999, 86: 199–217
- 27 Attouch H, Soueycatt M. Augmented Lagrangian and proximal alternating direction methods of multipliers in Hilbert spaces-applications to games, PDEs and control. Pacific J Optim, 2009, 5: 17–37
- 28 Douglas J, Rachford H H. On the numerical solution of the heat conduction problem in 2 and 3 space variables. Trans Amer Math Soc, 1956, 82: 421–439
- 29 Han D R. Inexact operator splitting methods with self-adaptive strategy for variational inequality problems. J Optim Theory Appl, 2007, 132: 227–243
- 30 Han D R, Xu W, Yang H. An operator splitting method for variational inequalities with partially unknown mappings. Numer Math, 2008, 111: 207–237
- 31 He B S, Liao L Z, Wang S L. Self-adaptive operator splitting methods for monotone variational inequalities. Numer Math, 2003, 94: 715–737
- 32 He B S, Yang H, Wang S L. Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities. J Optim Theory Appl, 2000, 106: 337–356
- 33 Kontogiorgis S, Meyer R R. A variable-penalty alternating directions method for convex optimization. Math Program, 1998, 83: 29–53
- 34 He B S, Yang H. Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities. Oper Res Lett, 1998, 23: 151–161
- 35 Facchinei F, Pang J S. Finite-Dimensional Variational Inequalities and Complementarity Problems. New York: Springer, 2003
- 36 Bertsekas D P, Tsitsiklis J N. Parallel and Distributed Computation: Numerical Methods. Upper Saddle River, NJ: Prentice-Hall, 1989
- 37 Eaves E C. On the basic theorem of complementarity. Math Program, 1971, 1: 68–75
- 38 Gao Y, Sun D F. Calibrating least squares semidefinite programming with equality and inequality constraints. SIAM J Matrix Anal A, 2009, 31: 1432–1457

A new alternating direction method for solving separable variational inequality problems

ZHANG Min, HAN DeRen, HE HongJin & CHEN YanNan

Abstract The alternating direction methods are attractive for solving separable variational inequality problems, which solve the original large-scale problems via solving a series of small-scale problems. However, the two subproblems solved per iteration are still variational inequality problems, which are structurally as difficult to solve as the original problems. In this paper, we propose a new alternating direction method, which, per iteration, solves a variational inequality problem and a system of nonlinear equations. The nonlinear equation is well-conditioned and many efficient methods, such as Newton-type methods, can be adopted directly to solve it. Moreover, from numerical point of view, nonlinear equations are much easier to solve than variational inequality problems in many cases. Under the same mild assumptions as those in classical alternating direction methods, we prove global convergence of the new method. We also present some preliminary numerical results, which show that our method is efficient.

Keywords variational inequality problems, alternating direction methods, global convergence, nonlinear equations

MSC(2010) 65K03

doi: 10.1360/012011-42