doi:10.3969/j.issn.1001-4616.2020.01.002

算子分裂法求解一类变分不等式问题的收敛率分析

葛志利[1],蔡邢菊[2],张　欣[3]

(1.南京科技职业学院基础科学部,江苏 南京 210048)
(2.南京师范大学数学科学学院,江苏 南京 210023)
(3.宿迁学院文理学院,江苏 宿迁 223800)

[摘要]　考虑一类变分不等式问题:寻找 $\boldsymbol{x}^*\in\Omega$,满足 $\boldsymbol{F}(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^*)\geqslant 0,\ \forall\boldsymbol{x}\in\Omega$,其中 Ω 是 R^n 上的闭凸子集,$\boldsymbol{F}=\boldsymbol{f}+\boldsymbol{g}$ 是 R^n 到 R^n 的连续算子,$\boldsymbol{f}$ 和 $\boldsymbol{g}$ 单调但 $\boldsymbol{f}$ 的表达式未知. 针对此类应用较广的问题,本文研究了一种新的算子分裂法. 根据已有的收敛性结果,进一步分析了该方法在非遍历意义下 $O(1/k)$ 和 $o(1/k)$ 的次线性收敛率,其中 k 表示迭代步数.最后,通过数值实验展示了算法的有效性.

[关键词]　部分算子未知,单调变分不等式,算子分裂法,次线性收敛率
[中图分类号]O221.4　[文献标志码]A　[文章编号]1001-4616(2020)01-0005-08

Convergence Rate Analysis of An Operator Splitting Method for Solving a Class of Variational Inequality Problems

Ge Zhili[1],Cai Xingju[2],Zhang Xin[3]

(1.Basic Sciences Department,Nanjing Polytechnic Institute,Nanjing 210048,China)
(2.School of Mathematical Sciences,Nanjing Normal University,Nanjing 210023,China)
(3.School of Arts and Science,Suqian College,Suqian 223800,China)

Abstract:Consider a class of variational inequality problems:finding $\boldsymbol{x}^*\in\Omega$, such that $\boldsymbol{F}(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^*)\geqslant 0,\ \forall\boldsymbol{x}\in\Omega$, where $\Omega\subseteq R^n$ is nonempty,closed and convex,$\boldsymbol{F}=\boldsymbol{f}+\boldsymbol{g}$ is a continuous mapping from R^n to R^n,$\boldsymbol{f}$ and $\boldsymbol{g}$ are monotone but $\boldsymbol{f}$ is unknown. We study an operator splitting method for this class of problems with a variety of applications. Based on the previous convergence results,we further analyze the $O(1/k)$ and $o(1/k)$ sublinear convergence rate in non-ergodic sense for this operator splitting method,where k counts the iteration number. Finally,numerical results demonstrate the efficiency of the algorithm.

Key words:partially unknown mappings,monotone variational inequalities,operator splitting method,sublinear convergence rate

考虑变分不等式问题:寻找 $\boldsymbol{x}^*\in\Omega$,满足

$$\boldsymbol{F}(\boldsymbol{x}^*)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^*)\geqslant 0,\quad \forall\boldsymbol{x}\in\Omega, \tag{1}$$

其中 Ω 是 R^n 上的闭凸子集,$\boldsymbol{F}$ 是 R^n 到 R^n 的连续算子.假设问题(1)的解集 Ω^* 非空. 上述变分不等式模型在交通、经济以及图像处理等领域有着广泛的应用[1-2]. 当(1)中的 $\boldsymbol{F}$ 具有显式表达式时,学者们提出了很多有效的求解方法,如:邻近点方法、算子分裂法和投影梯度法等[3]. 其中算子分裂法因具有降低子问题难度、充分利用问题结构等特性而广受欢迎. 该方法最初是为了更好地求解如下包含式问题

$$0\in\boldsymbol{T}(\boldsymbol{x}), \tag{2}$$

这里的 $\boldsymbol{T}$ 是一个极大单调算子,针对这一问题,早在上个世纪 70 年代,Martinet[4] 和 Rockfellar[5] 就提出了经典的邻近点方法,其迭代格式为

$$\boldsymbol{x}^{k+1}=(\boldsymbol{I}+\beta T)^{-1}(\boldsymbol{x}^k),$$

收稿日期:2019-03-31.
基金项目:国家自然科学基金项目(11401315、11871279)、江苏省高校自然科学研究面上项目(17KJD110003)、江苏省高职院校教师专业带头人高端研修项目和江苏省青蓝工程资助项目.
通讯作者:葛志利,博士,研究方向:运筹学与控制论. E-mail:gezhilyly66@126.com

其中 $\beta>0$,$\boldsymbol{I}$ 是恒等算子,$(\boldsymbol{I}+\beta\boldsymbol{T})^{-1}$称为 $\boldsymbol{T}$ 的预解式. 然而,在实际应用中,由于 $\boldsymbol{T}$ 的复杂结构,$\boldsymbol{T}$ 的预解式通常难以计算. 如求解变分不等式问题(1)等价于求解 $\boldsymbol{T}=\boldsymbol{F}+\boldsymbol{N}_\Omega$ 的包含式问题,其中 $\boldsymbol{N}_\Omega(\cdot)$ 表示在某点处的法锥,具体定义如下:

$$N_\Omega(\boldsymbol{x}):=\begin{cases}\{\boldsymbol{d}\mid(\boldsymbol{y}-\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}\leqslant 0,\ \forall\,\boldsymbol{y}\in\Omega, & \text{如果 } \boldsymbol{x}\in\Omega\},\\ \boldsymbol{\phi}, & \text{否则.}\end{cases}$$

若直接利用邻近点方法求解 $\boldsymbol{T}=\boldsymbol{F}+\boldsymbol{N}_\Omega$ 的包含式问题,$(\boldsymbol{I}+\beta\boldsymbol{T})^{-1}$将难以得到. 但幸运的是,$\boldsymbol{F}$ 和 $\boldsymbol{N}_\Omega$ 各自的预解式容易求解,如$(\boldsymbol{I}+\beta\boldsymbol{N}_\Omega)^{-1}=\boldsymbol{P}_\Omega$,其中 $\boldsymbol{P}_\Omega(\cdot)$ 表示某一向量到 Ω 上的投影. 因此,很多学者尝试在设计算法时不涉及 $\boldsymbol{T}$ 的预解式的计算,而仅涉及 $\boldsymbol{F}$ 或 $\boldsymbol{N}_\Omega$ 的预解式计算. 通过这一思想得到的算法通常称为算子分裂方法. 经典的算子分裂法[6]包括:向前向后分裂法;双向后算子分裂法;Peaceman-Rachford 分裂法和 Douglas-Rachford 分裂法. 在求解变分不等式问题(1)时,常用的是 Peaceman-Rachford 分裂法和 Douglas-Rachford 分裂法. 后来,Varga[7]将这两种方法结合起来,提出如下的一种改进的算子分裂法:

$$\boldsymbol{x}^{k+1}+\beta\boldsymbol{F}(\boldsymbol{x}^{k+1})=\boldsymbol{x}^k+\beta\boldsymbol{F}(\boldsymbol{x}^k)-\gamma\boldsymbol{e}(\boldsymbol{x}^k,\beta),\quad \gamma\in(0,2],\tag{3}$$

其中 $\boldsymbol{e}(\boldsymbol{x}^k,\beta):=\boldsymbol{x}^k-\boldsymbol{P}_\Omega(\boldsymbol{x}^k-\beta\boldsymbol{F}(\boldsymbol{x}^k))$. 在迭代序列(3)中取 $\gamma=1$ 时,算法退化到 Peaceman-Rachford 算子分裂法;取 $\gamma=2$ 时退化到 Douglas-Rachford 算子分裂法. 实际上,实现算法(3)时需要先计算 $\boldsymbol{x}^k-\beta\boldsymbol{F}(\boldsymbol{x}^k)$ 到 Ω 上的投影,再求解非线性方程组. 因此,这一方法可以看成由两个邻近步组成[4-5].

从(3)可以看出,以上的算子分裂法只能求解算子 $\boldsymbol{F}$ 有具体表达式的变分不等式问题.然而,在很多实际应用中,算子 $\boldsymbol{F}$ 通常包含部分未知的信息,也就是说,$\boldsymbol{F}$ 能表示成 $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{f}(\boldsymbol{x})+\boldsymbol{g}(\boldsymbol{x})$,其中 $\boldsymbol{f}$ 和 $\boldsymbol{g}$ 连续单调,但 $\boldsymbol{f}$ 的表达式未知. 比如,交通管理中的拥堵道路收费问题可以建模成变分不等式问题(1),但其中 $\boldsymbol{F}$ 包含的需求函数是未知的[8-9]. 如果不考虑其他信息,问题很难被求解. 幸运的是,对给定的 $\boldsymbol{g}(\boldsymbol{x}^k)$,可以通过实际观测得到如下变分不等式问题的解 $\bar{\boldsymbol{x}}^k$.

$$(\boldsymbol{f}(\bar{\boldsymbol{x}}^k)+\boldsymbol{g}(\boldsymbol{x}^k))^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}^k)\geqslant 0,\quad \forall\,\boldsymbol{x}\in\Omega.\tag{4}$$

即在拥堵道路收费问题中,对于给定的收费 $\boldsymbol{g}(\boldsymbol{x}^k)$,(4)中的路段流量 $\bar{\boldsymbol{x}}^k$ 可以通过观测路网中用户的反应得到.也就是说,对于给定 $\boldsymbol{g}(\boldsymbol{x}^k)$后,我们可以把(4)看成是一个黑匣子. 利用这些信息,Han 等人[8]提出了一种算子分裂法.他们首先从黑匣子中获得(4)的解 $\bar{\boldsymbol{x}}^k$,再求解关于 $\boldsymbol{g}$ 的非线性方程组

$$\boldsymbol{\theta}_k(\boldsymbol{x})=0$$

来获得新的迭代点 $\boldsymbol{x}^{k+1}$,其中 $\boldsymbol{\theta}_k(\boldsymbol{x})=\boldsymbol{x}+\beta\boldsymbol{g}(\boldsymbol{x})-\boldsymbol{x}^k-\beta\boldsymbol{g}(\boldsymbol{x}^k)+\alpha(\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k)$,参数 $\alpha,\beta>0$,$\boldsymbol{x}^k$ 是当前迭代点.

算法 1　算子分裂法

步 0. 任取初始点 $\boldsymbol{x}^0\in\Omega$,给定参数 $\beta,\varepsilon>0$.令 $k:=0$.

步 1. 计算 $\boldsymbol{g}(\boldsymbol{x}^k)$,观测系统达到平衡状态来获得满足如下变分不等式问题的解 $\bar{\boldsymbol{x}}^k$:

$$\left(\boldsymbol{f}(\bar{\boldsymbol{x}}^k)+\boldsymbol{g}(\boldsymbol{x}^k)+\frac{1}{\beta}(\bar{\boldsymbol{x}}^k-\boldsymbol{x}^k)\right)^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}^k)\geqslant 0,\quad \forall\,\boldsymbol{x}\in\Omega.\tag{5}$$

步 2. 任选 $0<\alpha<2$,求解非线性方程组

$$\boldsymbol{x}^{k+1}+\beta\boldsymbol{g}(\boldsymbol{x}^{k+1})=\boldsymbol{x}^{k+1}+\beta\boldsymbol{g}(\boldsymbol{x}^k)-\alpha(\boldsymbol{x}-\bar{\boldsymbol{x}}^k)\geqslant 0,\tag{6}$$

得到新的迭代点 $\boldsymbol{x}^{k+1}$.

步 3. 如果 $\|\boldsymbol{x}^{k+1}-\boldsymbol{x}^k\|\leqslant\varepsilon$,停止;否则,令 $k:=k+1$,转步 2.

Han 等人[8]的上述方法仅包含一个邻近步,并在 $\boldsymbol{f}$ 强单调,$\boldsymbol{g}$ 单调的假设条件下证明了该方法的全局收敛性.

但是,$\boldsymbol{f}$ 强单调的假设限制了该算法的应用范围,实际应用中常出现 $\boldsymbol{f}$ 和 $\boldsymbol{g}$ 不是强单调,而仅是单调的情况[10]. 为了减弱假设扩大应用范围,Ge 等人[11]进一步改进了 Han 等人[8]的算法,通过引入一种扰动策略,提出了一种新的算子分裂方法. 具体的算法框架在算法 1 中给出. 在 $\boldsymbol{f}$ 和 $\boldsymbol{g}$ 仅为单调的假设下,Ge 等人[11]证明新的算子分裂方法的全局收敛性.

针对特殊的包含式问题,早期便有学者[12-13]给出了经典邻近点算法在最差情况下的收敛率分析. 对更一般包含问题(2),He 等人[14]通过利用反射算子分析了 Douglas-Rachford 算子分裂法的收敛率. 进一步,在变分不等式的框架下,对带有线性约束的可分凸优化问题,He 等人[15-16]分别给出了 Douglas-Rachford 交替方向法在遍历和非遍历意义下的收敛率.对本文考虑的变分不等式问题(1)($\boldsymbol{f}$ 未知,$\boldsymbol{g}$ 已

知),Kou 等人[17]给出了 Han 等人[8]所提算子分裂法的收敛率分析,但需要假设 $\boldsymbol{f}$ 是强单调. 而对 $\boldsymbol{f}$ 和 $\boldsymbol{g}$ 均为单调的算子分裂法[11],其收敛率一直未被研究. 本文基于 Ge 等人的收敛性结果,依赖于变分不等式框架,进一步分析了 $\boldsymbol{f}$ 和 $\boldsymbol{g}$ 仅为单调情况下的算子分裂法在非遍历意义下 $O(1/k)$ 和 $o(1/k)$ 的次线性收敛率,其中 k 表示迭代步数.最后,数值实验展示了算法的有效性.

1 预备知识

本节主要给出一些文中涉及到的定义、符号以及与变分不等式和投影算子有关的性质.

对 R^n 中的任意两个向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$,用 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$ 表示 Euclidean 空间中的内积, $\|\boldsymbol{x}\|=\sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$ 为向量 $\boldsymbol{x}\in R^n$ 的 Euclidean 范数. 令 Ω 是 R^n 上的一闭凸集,则某点到集合 Ω 上的投影的定义如下:

定义 1 对给定的 $\boldsymbol{y}\in R^n$ 在 Euclidean 范数下到集合 Ω 上的投影记为 $\boldsymbol{P}_\Omega(\boldsymbol{y})$,且

$$\boldsymbol{P}_\Omega(\boldsymbol{y}):=\operatorname{argmin}\{\|\boldsymbol{x}-\boldsymbol{y}\|,\boldsymbol{x}\in\Omega\}.$$

根据投影算子的定义,有如下性质.

引理 1[3,18] 令 $\Omega\subset R^n$ 为一非空闭凸集,则有

$$(\boldsymbol{u}-\boldsymbol{P}_\Omega(\boldsymbol{u}))^{\mathrm{T}}(\boldsymbol{v}-\boldsymbol{P}_\Omega(\boldsymbol{u}))\leqslant 0,\quad \forall\,\boldsymbol{u}\in R^n,\quad \forall\,\boldsymbol{v}\in\Omega;$$

$$\|\boldsymbol{P}_\Omega(\boldsymbol{u})-\boldsymbol{P}_\Omega(\boldsymbol{v})\|\leqslant\|\boldsymbol{u}-\boldsymbol{v}\|,\quad \forall\,\boldsymbol{u},\boldsymbol{v}\in R^n;$$

$$\|\boldsymbol{P}_\Omega(\boldsymbol{u})-\boldsymbol{v}\|^2\leqslant\|\boldsymbol{u}-\boldsymbol{v}\|^2-\|\boldsymbol{u}-\boldsymbol{P}_\Omega(\boldsymbol{u})\|^2,\quad \forall\,\boldsymbol{u}\in R^n,\boldsymbol{v}\in\Omega.$$

结合定义 1,求解问题(1)等价于求解一个投影方程[5]为

$$\boldsymbol{x}=\boldsymbol{P}_\Omega(\boldsymbol{x}-\beta\boldsymbol{F}(\boldsymbol{x})),\quad \beta>0.$$

同样根据定义 $\boldsymbol{e}(\boldsymbol{x},\beta):=\boldsymbol{x}-\boldsymbol{P}_\Omega(\boldsymbol{x}-\beta\boldsymbol{F}(\boldsymbol{x}))$,可知求解(1)与求解非光滑方程 $\boldsymbol{e}(\boldsymbol{x},\beta)=0$ 等价.

对于给定的 $\boldsymbol{x}\in R^n$,下面的引理给出了 $\|\boldsymbol{e}(\boldsymbol{x},\beta)\|$ 是关于 β 的非减函数,而 $\|\boldsymbol{e}(\boldsymbol{x},\beta)\|/\beta$ 是关于 β 的非增函数.

引理 2[18] 对任意 $\boldsymbol{x}\in R^n$ 和 $\tilde{\beta}\geqslant\beta>0$,则有

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

和

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

接下来,给出算子单调、强单调和 Lipschitz 连续的定义.

定义 2 给定 $\Omega\subset R^n$ 和算子 $\boldsymbol{F}:\Omega\to R^n$,

(1)如果 $(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{F}(\boldsymbol{y}))\geqslant 0,\quad \forall\,\boldsymbol{x},\boldsymbol{y}\in\Omega$ 成立,则称 $\boldsymbol{F}(\cdot)$ 是单调的.

(2)如果存在常数 $\mu>0$,使得

$$(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{F}(\boldsymbol{y}))\geqslant 0,\quad \forall\,\boldsymbol{x},\boldsymbol{y}\in\Omega$$

成立,则称 $\boldsymbol{F}(\cdot)$ 是强单调的.

(3)如果存在常数 $L>0$,满足如下不等式

$$\|\boldsymbol{F}(\boldsymbol{x})-\boldsymbol{F}(\boldsymbol{y})\|\leqslant L\|\boldsymbol{x}-\boldsymbol{y}\|,\quad \forall\,\boldsymbol{x},\boldsymbol{y}\in\Omega$$

成立,则称 $\boldsymbol{F}(\cdot)$ 在 Ω 上是 L-lipschitz 连续的.

2 收敛率分析

本节基于已有的收敛性结论,给出算法 1 的收敛率分析. 首先,针对算法 1,给出文献[11]中引理 3.2.

引理 3 假设 $\boldsymbol{f}$ 和 $\boldsymbol{g}$ 单调,则由算法 1 产生的序列 $\{\boldsymbol{x}^k\}$ 满足

$$\begin{aligned}\|\boldsymbol{x}^{k+1}+\beta\boldsymbol{g}(\boldsymbol{x}^{k+1})-\boldsymbol{x}^*-\beta\boldsymbol{g}(\boldsymbol{x}^*)\|^2&\leqslant\|\boldsymbol{x}^k+\beta\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{x}^*-\beta\boldsymbol{g}(\boldsymbol{x}^*)\|^2-\alpha(2-\alpha)\|\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k\|^2=\\ &\|\boldsymbol{x}^k+\beta\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{x}^*-\beta\boldsymbol{g}(\boldsymbol{x}^*)\|^2-\frac{2-\alpha}{\alpha}\|\boldsymbol{x}^k+\beta\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{x}^{k+1}-\beta\boldsymbol{g}(\boldsymbol{x}^{k+1})\|^2.\end{aligned}\tag{7}$$

下面,给出算法 1 的收敛性结果.

定理 1 假设 $\boldsymbol{f}$ 和 $\boldsymbol{g}$ 单调,参数满足 $0<\alpha<2$ 和 $\beta>0$. 则由算法 1 产生的序列 $\{\boldsymbol{x}^k\}$ 收敛到(1)的解点.

证明 利用(7),可得

$$
\begin{aligned}
&\|x^{k+1}-x^*+\beta(g(x^{k+1})-g(x^*))\|^2\\
\leqslant&\|x^k-x^*+\beta(g(x^k)-g(x^*))\|^2\\
&\vdots\\
\leqslant&\|x^0-x^*+\beta(g(x^0)-g(x^*))\|^2<+\infty.
\end{aligned}
$$

因为 g 单调,序列 $\{x^k\}$ 有界以及(7),得到

$$\alpha(2-\alpha)\|x^k-\bar{x}^k\|^2\leqslant\|x^k-x^*+\beta(g(x^k)-g(x^*))\|^2-\|x^{k+1}-x^*+\beta(g(x^{k+1})-g(x^*))\|^2.$$

对上面的不等式两边进行累加,可得

$$\sum_{k=0}^{\infty}\alpha(2-\alpha)\|x^k-\bar{x}^k\|^2\leqslant\sum_{k=0}^{\infty}(\|x^k-x^*+\beta(g(x^k)-g(x^*))\|^2-\|x^{k+1}-x^*+\beta(g(x^{k+1})-g(x^*))\|^2)\leqslant\|x^0-x^*+\beta(g(x^0)-g(x^*))\|^2<+\infty.\tag{8}$$

又因为 $0<\alpha<2$,根据(8),可得

$$\lim_{k\to\infty}\|x^k-\bar{x}^k\|=0.\tag{9}$$

此外,由(6),我们有

$$\lim_{k\to\infty}\|x^k-x^{k+1}\|=0.$$

因为 $\{x^k\}$ 有界,则至少有一个聚点,记为 $\tilde{x}\in\Omega$. 令其中一子列 $\{x^{k_j}\}$ 收敛到 $\tilde{x}$. 根据(9),得到序列 $\{\bar{x}^k\}$ 子列 $\{\bar{x}^{k_j}\}$ 也收敛到 $\tilde{x}$. 另一方面,(5)等价于

$$e(\bar{x}^k,\beta)=\bar{x}^k-P_\Omega(x^k-\beta(f(\bar{x}^k)+g(x^k)))=0.$$

因此,对子列 $\{\bar{x}^{k_j}\}$ 取极限,我们有

$$\|e(\tilde{x},\beta)\|=\lim_{j\to\infty}\|e(\bar{x}^{k_j},\beta)\|=0,$$

这就意味着 $\tilde{x}\in\Omega$ 是变分不等式问题(1)的解. 不等式(7)表明整个序列 $\{x^k\}$ 仅有一个聚点. 因此,序列 $\{x^k\}$ 收敛到(1)的解点 $\tilde{x}$.

接下来,证明序列 $\{\|x^k-\bar{x}^k\|^2\}$ 和 $\{\|x^k+\beta g(x^k)-x^{k+1}-\beta g(x^{k+1})\|^2\}$ 的单调性.

定理 2 假设 f 和 g 单调,参数满足 $0<\alpha<2$ 和 $\beta>0$. 则对算法 1 产生的序列 $\{x^k\}$ 和 $\{\bar{x}^k\}$,有

$$\|x^{k+1}-\bar{x}^{k+1}\|^2\leqslant\|x^k-\bar{x}^k\|^2-\frac{(2-\alpha)}{\alpha}\|(x^k-\bar{x}^k)-(x^{k+1}-\bar{x}^{k+1})\|^2,\tag{10}$$

和

$$\|x^{k+1}-x^{k+2}+\beta(g(x^{k+1})-g(x^{k+2}))\|^2\leqslant\|x^k-x^{k+1}+\beta(g(x^k)-g(x^{k+1}))\|^2-\frac{(2-\alpha)}{\alpha}\|(x^k+\beta g(x^k)-x^{k+1}-\beta g(x^{k+1}))-(x^{k+1}+\beta g(x^{k+1})-x^{k+2}-\beta g(x^{k+2}))\|^2.\tag{11}$$

证明 首先,证明序列 $\{\|x^k-\bar{x}^k\|^2\}$ 的单调性.

在(5)中令 $x=\bar{x}^{k+1}$,得到

$$(\bar{x}^{k+1}-\bar{x}^k)^{\mathrm{T}}\left(f(\bar{x}^k)+g(x^k)+\frac{1}{\beta}(\bar{x}^k-x^k)\right)\geqslant0.\tag{12}$$

类似地,令 $x=\bar{x}^k$,则有

$$(\bar{x}^k-\bar{x}^{k+1})^{\mathrm{T}}\left(f(\bar{x}^{k+1})+g(x^{k+1})+\frac{1}{\beta}(\bar{x}^{k+1}-x^{k+1})\right)\geqslant0.\tag{13}$$

将(12)和(13)进行相加并重组,可得

$$((x^k-\bar{x}^k)-(x^{k+1}-\bar{x}^{k+1}))^{\mathrm{T}}(g(x^k)-g(x^{k+1}))\geqslant((x^k-x^{k+1})^{\mathrm{T}}(g(x^k)-g(x^{k+1}))+(\bar{x}^{k+1}-\bar{x}^k))^{\mathrm{T}}(f(\bar{x}^{k+1})-f(\bar{x}^k))+\frac{1}{\beta}(\bar{x}^k-\bar{x}^{k+1})^{\mathrm{T}}((\bar{x}^k-\bar{x}^{k+1})-(x^k-x^{k+1}))\geqslant\frac{1}{\beta}\|\bar{x}^k-\bar{x}^{k+1}\|^2-\frac{1}{\beta}(\bar{x}^k-\bar{x}^{k+1})^{\mathrm{T}}(x^k-x^{k+1}).\tag{14}$$

其中第二个不等式根据 f 和 g 的单调性.

根据(14),得到

$$((x^k-\bar{x}^k)-(x^{k+1}-\bar{x}^{k+1}))^{\mathrm{T}}\left(g(x^k)-g(x^{k+1})+\frac{1}{\beta}(x^k-x^{k+1})\right)\geqslant\frac{1}{\beta}\|(x^k-x^{k+1})\|^2-$$

$$\frac{2}{\beta}(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})^{\mathrm{T}}(\bar{\boldsymbol{x}}^k-\bar{\boldsymbol{x}}^{k+1})+\frac{1}{\beta}\|(\bar{\boldsymbol{x}}^k-\bar{\boldsymbol{x}}^{k+1})\|^2. \tag{15}$$

利用$-2\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}=-\|\boldsymbol{a}\|^2-\|\boldsymbol{b}\|^2+\|\boldsymbol{a}-\boldsymbol{b}\|^2$,我们获得如下的等式

$$-2(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})^{\mathrm{T}}(\bar{\boldsymbol{x}}^k-\bar{\boldsymbol{x}}^{k+1})=-\|\boldsymbol{x}^k-\boldsymbol{x}^{k+1}\|^2-\|\bar{\boldsymbol{x}}^k-\bar{\boldsymbol{x}}^{k+1}\|^2+\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})-(\bar{\boldsymbol{x}}^k-\bar{\boldsymbol{x}}^{k+1})\|^2. \tag{16}$$

将(16)代入到(15),得到

$$((\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k)-(\boldsymbol{x}^{k+1}-\bar{\boldsymbol{x}}^{k+1}))^{\mathrm{T}}\left(\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{g}(\boldsymbol{x}^{k+1})+\frac{1}{\beta}(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})\right)\geqslant\frac{1}{\beta}\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})-(\bar{\boldsymbol{x}}^k-\bar{\boldsymbol{x}}^{k+1})\|^2. \tag{17}$$

对(17)两边同乘以β,得到

$$((\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k)-(\boldsymbol{x}^{k+1}-\bar{\boldsymbol{x}}^{k+1}))^{\mathrm{T}}((\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{g}(\boldsymbol{x}^{k+1})))\geqslant\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})-(\bar{\boldsymbol{x}}^k-\bar{\boldsymbol{x}}^{k+1})\|^2. \tag{18}$$

根据$\boldsymbol{x}^{k+1}$的定义,得到

$$\alpha((\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k)-(\boldsymbol{x}^{k+1}-\bar{\boldsymbol{x}}^{k+1}))^{\mathrm{T}}(\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k)\geqslant\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})-((\bar{\boldsymbol{x}}^k-\bar{\boldsymbol{x}}^{k+1}))\|^2.$$

利用等式$2\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{a}-\boldsymbol{b})=\|\boldsymbol{a}\|^2-\|\boldsymbol{b}\|^2+\|\boldsymbol{a}-\boldsymbol{b}\|^2$,得到

$$\frac{\alpha}{2}(\|\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k\|^2-\|\boldsymbol{x}^{k+1}-\bar{\boldsymbol{x}}^{k+1}\|^2+\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})-(\bar{\boldsymbol{x}}^k-\bar{\boldsymbol{x}}^{k+1})\|^2)\geqslant\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})-(\bar{\boldsymbol{x}}^k-\bar{\boldsymbol{x}}^{k+1})\|^2. \tag{19}$$

序列(10)的单调性,可由以上不等式(19)直接得到.

下面证明序列$\{\|\boldsymbol{x}^k+\beta\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{x}^{k+1}-\beta\boldsymbol{g}(\boldsymbol{x}^{k+1})\|^2\}$的单调性.将(18)式两边应用等式

$$\alpha(\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k)=\boldsymbol{x}^k+\beta\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{x}^{k+1}-\beta\boldsymbol{g}(\boldsymbol{x}^{k+1}),$$

和

$$\alpha(\boldsymbol{x}^{k+1}-\bar{\boldsymbol{x}}^{k+1})=\boldsymbol{x}^{k+1}+\beta\boldsymbol{g}(\boldsymbol{x}^{k+1})-\boldsymbol{x}^{k+2}-\beta\boldsymbol{g}(\boldsymbol{x}^{k+2}),$$

得到

$$\begin{aligned}&\{[(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{g}(\boldsymbol{x}^{k+1}))]-[(\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k+2})+\beta(\boldsymbol{g}(\boldsymbol{x}^{k+1})-\boldsymbol{g}(\boldsymbol{x}^{k+2}))]\}^{\mathrm{T}}[(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\\&\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{g}(\boldsymbol{x}^{k+1}))]\geqslant\frac{1}{\alpha}\|[(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{g}(\boldsymbol{x}^{k+1}))]-\\&[(\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k+2})+\beta(\boldsymbol{g}(\boldsymbol{x}^{k+1})-\boldsymbol{g}(\boldsymbol{x}^{k+2}))]\|^2.\end{aligned} \tag{20}$$

利用$\|\boldsymbol{a}\|^2-\|\boldsymbol{b}\|^2=2\boldsymbol{a}^{\mathrm{T}}(\boldsymbol{a}-\boldsymbol{b})-\|\boldsymbol{a}-\boldsymbol{b}\|^2$,得到

$$\begin{aligned}&\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{g}(\boldsymbol{x}^{k+1}))\|^2-\|(\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k+2})+\beta(\boldsymbol{g}(\boldsymbol{x}^{k+1})-\boldsymbol{g}(\boldsymbol{x}^{k+2}))\|^2=2[(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\\&\boldsymbol{g}(\boldsymbol{x}^{k+1}))]^{\mathrm{T}}[(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{g}(\boldsymbol{x}^{k+1}))]-[(\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k+2})+\beta(\boldsymbol{g}(\boldsymbol{x}^{k+1})-\boldsymbol{g}(\boldsymbol{x}^{k+2}))]-\\&\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{g}(\boldsymbol{x}^{k+1}))-(\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k+2})-\beta(\boldsymbol{g}(\boldsymbol{x}^{k+1})-\boldsymbol{g}(\boldsymbol{x}^{k+2}))\|^2,\end{aligned} \tag{21}$$

将(20)代入到(21)的右边第一部分,可得

$$\begin{aligned}&\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{g}(\boldsymbol{x}^{k+1}))\|^2-\|(\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k+2})+\beta(\boldsymbol{g}(\boldsymbol{x}^{k+1})-\boldsymbol{g}(\boldsymbol{x}^{k+2}))\|^2\geqslant\frac{2}{\alpha}\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\\&\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{g}(\boldsymbol{x}^{k+1}))-[(\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k+2})+\beta(\boldsymbol{g}(\boldsymbol{x}^{k+1})-\boldsymbol{g}(\boldsymbol{x}^{k+2}))]\|^2-\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\\&\boldsymbol{g}(\boldsymbol{x}^{k+1}))-(\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k+2})-\beta(\boldsymbol{g}(\boldsymbol{x}^{k+1})-\boldsymbol{g}(\boldsymbol{x}^{k+2}))\|^2=\frac{2-\alpha}{\alpha}\|(\boldsymbol{x}^k-\boldsymbol{x}^{k+1})+\\&\beta(\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{g}(\boldsymbol{x}^{k+1}))-(\boldsymbol{x}^{k+1}-\boldsymbol{x}^{k+2})-\beta(\boldsymbol{g}(\boldsymbol{x}^{k+1})-\boldsymbol{g}(\boldsymbol{x}^{k+2}))\|^2.\end{aligned} \tag{22}$$

序列(11)的单调性可由(22)直接得到.

现在证明最差情况下算法 1 在非遍历意义下的$O(1/k)$的收敛率.

定理 3　假设$\boldsymbol{f}$和$\boldsymbol{g}$单调,参数满足$0<\alpha<2$和$\beta>0$.序列$\{\boldsymbol{x}^k\}$和$\{\bar{\boldsymbol{x}}^k\}$由算法 1 产生.则对任意的$\boldsymbol{x}^*\in\Omega^*$,有

$$\|\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k\|^2\leqslant\frac{1}{\alpha(2-\alpha)(k+1)}\|\boldsymbol{x}^0+\beta\boldsymbol{g}(\boldsymbol{x}^0)-\boldsymbol{x}^*-\beta\boldsymbol{g}(\boldsymbol{x}^*)\|^2, \tag{23}$$

和

$$\|\boldsymbol{x}^k+\beta\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{x}^{k+1}-\beta\boldsymbol{g}(\boldsymbol{x}^{k+1})\|^2\leqslant\frac{\alpha}{(2-\alpha)(k+1)}\|\boldsymbol{x}^0+\beta\boldsymbol{g}(\boldsymbol{x}^0)-\boldsymbol{x}^*-\beta\boldsymbol{g}(\boldsymbol{x}^*)\|^2. \tag{24}$$

证明　由(7),得到

$$\alpha(2-\alpha)\sum_{t=0}^{\infty}\|\boldsymbol{x}^t-\bar{\boldsymbol{x}}^t\|^2\leqslant\|\boldsymbol{x}^0+\beta\boldsymbol{g}(\boldsymbol{x}^0)-\boldsymbol{x}^*-\beta\boldsymbol{g}(\boldsymbol{x}^*)\|^2,\quad\forall\boldsymbol{x}^*\in\Omega^*, \tag{25}$$

和

$$\frac{2-\alpha}{\alpha}\sum_{t=0}^{\infty}\|\boldsymbol{x}^t+\beta\boldsymbol{g}(\boldsymbol{x}^t)-\boldsymbol{x}^{t+1}-\beta\boldsymbol{g}(\boldsymbol{x}^{t+1})\|^2\leqslant\|\boldsymbol{x}^0+\beta\boldsymbol{g}(\boldsymbol{x}^0)-\boldsymbol{x}^*-\beta\boldsymbol{g}(\boldsymbol{x}^*)\|^2,\quad\forall\boldsymbol{x}^*\in\Omega^*. \tag{26}$$

另一方面,根据定理 2,序列$\{\|\boldsymbol{x}^t-\bar{\boldsymbol{x}}^t\|^2\}$和$\{\|\boldsymbol{x}^t+\beta\boldsymbol{g}(\boldsymbol{x}^t)-\boldsymbol{x}^{t+1}-\beta\boldsymbol{g}(\boldsymbol{x}^{t+1})\|^2\}$单调非增.因此,我们有

$$(k+1)\|\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k\|^2\leqslant\sum_{t=0}^{k}\|\boldsymbol{x}^t-\bar{\boldsymbol{x}}^t\|^2, \tag{27}$$

和

$$(k+1)\|\boldsymbol{x}^k+\beta\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{x}^{k+1}-\beta\boldsymbol{g}(\boldsymbol{x}^{k+1})\|^2\leqslant\sum_{t=0}^{k}\|\boldsymbol{x}^t+\beta\boldsymbol{g}(\boldsymbol{x}^t)-\boldsymbol{x}^{t+1}-\beta\boldsymbol{g}(\boldsymbol{x}^{t+1})\|^2. \tag{28}$$

不等式(23)由(25)和(27)可得,不等式(24)由(26)和(28)可得.

定理 3 给出了算法在非遍历意义下 $O(1/k)$ 的收敛率.实际上,注意到解集 Ω^* 非空,设 $d:=\inf\{\|\boldsymbol{x}^0+\beta\boldsymbol{g}(\boldsymbol{x}^0)-\boldsymbol{x}^*-\beta\boldsymbol{g}(\boldsymbol{x}^*)\|$,其中 $\boldsymbol{x}^*\in\Omega^*\}$,则对任给的 $\varepsilon>0$,定理 3 表明,算法 1 至多需要 $k=\left\lfloor\frac{d^2}{\alpha(2-\alpha)\varepsilon}\right\rfloor$次迭代就能保证

$$\|\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k\|^2\leqslant\varepsilon\text{ 和 }\|\boldsymbol{x}^k+\beta\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{x}^{k+1}-\beta\boldsymbol{g}(\boldsymbol{x}^{k+1})\|^2\leqslant\alpha^2\varepsilon.$$

此外,由

$$\lim_{k\to\infty}\|\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k\|^2=0\text{ 和}\lim_{k\to\infty}\|\boldsymbol{x}^k+\beta\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{x}^{k+1}-\beta\boldsymbol{g}(\boldsymbol{x}^{k+1})\|^2=0,$$

得到 $\boldsymbol{x}^k$ 收敛到变分不等式问题(1)的解.

下面,采用与文献[19]相同的方法,通过一个基本的引理将收敛率从 $O(1/k)$ 改进到 $o(1/k)$. 直观上,序列$\left\{\frac{1}{k}\right\}$是不可和的.

引理 4[19]　如果实数集 $\mathbf{R}$ 上的一个序列$\{\alpha_k\}$满足(1)$\alpha_k\geqslant0$;(2)$\sum_{k=0}^{\infty}\alpha_k<+\infty$;(3)$\alpha_k$ 单调非增,则有 $\alpha_k=o(1/k)$.

定理 4　假设 $\boldsymbol{f}$ 和 $\boldsymbol{g}$ 单调,参数满足 $0<\alpha<2$ 和 $\beta>0$.序列$\{\boldsymbol{x}^k\}$和$\{\bar{\boldsymbol{x}}^k\}$由算法 1 产生.则有

$$\|\boldsymbol{x}^k-\bar{\boldsymbol{x}}^k\|^2=o(1/k),$$

和

$$\|\boldsymbol{x}^k+\beta\boldsymbol{g}(\boldsymbol{x}^k)-\boldsymbol{x}^{k+1}-\beta\boldsymbol{g}(\boldsymbol{x}^{k+1})\|^2=o(1/k).$$

证明　根据定理 2,定理 3 和引理 4 即可得到.

注 1　Kou 等人[17]在 $\boldsymbol{f}$ 强单调(假设单调系数为 μ)并要求参数 $\beta>\frac{1}{2\mu}$的情况下,证明了 Han 等人的算法[8]在非遍历意义下具有 $O(1/k)$ 的收敛率.事实上,当 $\boldsymbol{f}$ 是强单调时,一般能够得到线性收敛率.而我们在 $\boldsymbol{f}$ 仅是单调且参数只需满足 $\beta>0$ 的条件下,证明了 Ge 等人的算法[11]在非遍历意义下具有 $O(1/k)$ 和 $o(1/k)$ 的次线性收敛率.

3　数值实验

本节,我们测试了算法的数值结果.所有程序均用 MATLAB2009b 语言编写,且在 win7 系统下,配置为 CPU:Intel(R)-Core(TM)i5-3210M 2.5GHz,内存:4GB 的笔记本电脑上运行.

在本例中,我们仅考虑数学结构,而忽略问题的实际背景.算子 $\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{f}(\boldsymbol{x})+\boldsymbol{g}(\boldsymbol{x})$,其中 $\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{D}\boldsymbol{x}+\boldsymbol{M}\boldsymbol{x}+\boldsymbol{q}$,$\boldsymbol{g}(\boldsymbol{x})=\boldsymbol{D}\arctan(\boldsymbol{x})$以及 $\boldsymbol{D}$ 是一个 $n\times n$ 对角阵,其中的元素随机分布在区间[0,10]上.矩阵 $\boldsymbol{M}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}+\boldsymbol{B}$,其中 $\boldsymbol{A}$ 是一个 $n\times n$ 矩阵,其中元素随机分布在区间(−5,5)上.反对称矩阵 $\boldsymbol{B}$ 中元素也是以相同的方式产生.向量 $\boldsymbol{q}$ 一致分布在区间(−500,500)内. 因此,$\boldsymbol{f}$ 和 $\boldsymbol{g}$ 仅是单调的,而非强单调.现在,我们用算法 1 来求解这个问题,问题维数 100,参数 $\beta=0.01$,$\alpha=0.998$.表 1、2 给出了计算结果,其中‘ε’表示终止准则,‘Outer’表示整个的迭代次数,‘Inner’表示算法步 2 中产生的平均迭代次数,‘Total’表示 Outer 与 Inner 的乘积,‘CPU’表示运行时间.

表 1　算法 1 的数值结果

Table 1　Numerical results of algorithm 1

ε	算法 1:$x^0=(1,1,\cdots,1)$				算法 1:$x^0=(0,0,\cdots,0)$			
	Outer	Inner	Total	CPU	Outer	Inner	Total	CPU
10^{-1}	6	174	1 044	0.28	5	171.60	858	0.20
10^{-2}	8	151.88	1 215	0.27	7	145	1 018	0.27
10^{-3}	9	140.67	1 266	0.34	8	133.86	1 071	0.30
10^{-4}	11	119.55	1 315	0.35	9	122.78	1 105	0.28
10^{-5}	13	102.38	1 331	0..47	11	103.36	1 137	0.34
10^{-6}	15	89.20	1 338	0.44	12	95.33	1 144	0.42
10^{-7}	23	58.52	1 346	0.53	22	52.64	1 158	0.64
10^{-8}	57	24.210 5	1 380	1.48	52	22.85	1 188	1.28

表 2　算法 1 和改进的算法 1 的数值结果

Table 2　Numerical results of algorithm 1 and modified algorithm 1

ε	算法 1:x^0 是区间(0,1)上的随机变量				改进的算法 1:x^0 是区间(0,1)上的随机变量			
	Outer	Inner	Total	CPU	Outer	Inner	Total	CPU
10^{-1}	6	182.17	1 093	0.22	5	27	135	0.17
10^{-2}	8	160.50	1 284	0.30	5	31.40	157	0.14
10^{-3}	10	137.90	1 379	0.42	6	27.67	160	0.15
10^{-4}	11	127.27	1 400	0.37	6	29.50	177	0.20
10^{-5}	13	109.15	1 419	0.45	7	32.86	230	0.27
10^{-6}	15	95.27	1 429	0.50	7	32.14	225	0.20
10^{-7}	21	68.33	1 435	0.51	8	30.88	247	0.23
10^{-8}	55	26.71	1 469	1.50	8	38.88	307	0.25

从表 1 可以看出,对不同的精度和初始点,算法均稳定、有效.根据以往的计算经验[8,11],参数β,α的选取往往不是固定的,而是自适应选取.另外,对方程组(6)通常作非精确求解.这里,我们采用与文献[8,11]类似的自适应策略和非精确求解准则,不妨将其记为改进的算法 1. 表 2 的结果告诉我们,改进的算法 1 在数值表现上更有优势,这也进一步说明了自适应策略和非精确求解方程组能够很好地提高算法的有效性.

4　结论

本文考虑了部分算子未知的变分不等式问题中的一种算子分裂法.在算子$\boldsymbol{f}$和$\boldsymbol{g}$单调的情况下,证明了算法的全局收敛性并分析了其在非遍历意义下$O(1/k)$和$o(1/k)$的次线性收敛率.最后,通过数值实验进一步展示了算法的稳定、有效.

[参考文献]

[1]　FERRIS M C,PANG J S. Engineering and economic applications of complimentarity problems[J]. SIAM review,1997,39(4):669-713.

[2]　FISCHER A. Solution of monotone complementarity problems with locally Lipschitzian functions[J]. Mathematical programming,1997,76(3):513-532.

[3]　FACCHINEI F,PANG J S. Finite-dimensional variational inequalities and complementarity problems,Volumes I and II[M]. Berlin,GER:Springer,2003.

[4]　MARTINET B. Brève communication. Régularisation d' inéquations variationnelles par approximations successives[J]. Revue francaise D' informatique etde recherche Opérationnelle,Série Rouge,1970,4(R3):154-158.

[5]　ROCKAFELLAR R T. Monotone operators and the proximal point algorithm[J]. SIAM journal on control optimization,1976,14(5):877-898.

[6]　ECKSTEIN J,BERTSEKAS D. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators[J]. Mathematical programming,1992,55(1-3):293-318.

[7]　VARGA R S. Matrix iterative analysis[M]. New York,USA:Springer,1999.

[8] HAN D R,XU W,YANG H. An operator splitting method for variational inequalities with partially unknown mappings[J]. Numerische mathematik,2008,111(2):207-237.

[9] YANG H,MENG Q,LEE D H. Trial-and-error implementation of marginal-cost pricing on networks in the absence of demand functions[J]. Transportation research part B:methodological,2004,38(6):477-493.

[10] STARCK J L,MURTAGH F,FADILI J M. Sparse image and signal processing,wavelets,curvelets,morphological diversity[M]. Cambridge,UK:Cambridge University Press,2010.

[11] GE Z L,HAN D R,NI Q,et al. An operator splitting method for monotone variational inequalities with a new perturbation strategy[J]. Optimization letters,2018,12(1):103-122.

[12] GJLER O. On the convergence of the proximal point algorithm for convex minimization[J]. SIAM journal on control and optimization,1991,29(2):403-419.

[13] NEMIROVSKI A. Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems[J]. SIAM journal on optimization,2005,15(1):229-251.

[14] HE B S,YUAN X M. On non-ergodic convergence rate of Douglas-Rachford alternating direction method of multipliers[J]. Numerische mathematik,2015,130(3):567-577.

[15] HE B S,YUAN X M. On the $O(1/n)$ convergence rate of the Douglas-Rachford alternating direction method[J]. SIAM journal on numerical analysis,2012,50(2):700-709.

[16] HE B S,YUAN X M. On the convergence rate of Douglas-Rachford operator splitting method[J]. Mathematical programming, 2015,153(2):715-722.

[17] KOU X P,LI S J. On non-ergodic convergence rate of the operator splitting method for a class of variational inequalities[J]. Optimization letters,2017,11(1):71-80.

[18] ZHU T,YU Z G. A simple proof for some important properties of the projection mapping[J]. Mathematical inequalities and applications,2004,7(3):453-456.

[19] CAI X J,HAN D R,YUAN X M. On the convergence of the direct extension of ADMM for three-block separable convex minimization models with one strongly convex function[J]. Computational optimization and applications,2017,66(1):39-73.

[责任编辑:陆炳新]