2005 年 3 月 重庆师范大学学报（自然科学版） Mar. 2005
第 22 卷 第 1 期 Journal of Chongqing Normal University（Natural Science Edition） Vol. 22 No. 1

一个求解变分不等式问题的投影算法*

王 传 伟
（曲阜师范大学 运筹与管理学院，山东 日照 276826）

摘 要：基于 D. Han 提出的求解变分不等式问题的推广的近似点算法（generalized proximal method），提出了一个新的改进算法，该算法的最大特点是在每一步只需要近似求解一个线性方程组系统。并在适当条件下证明了算法的全局收敛性。

关键词：变分不等式；线性搜索；投影；全局收敛
中图分类号：O221.2 **文献标识码**：A **文章编号**：1672-6693（2005）01-0006-05

A Projection Algorithm for Solving Variational Inequalities

WANG Chuan-wei
（College of Operation Research and Management Science，Qufu Normal University，Rizhao Shandong 276826，China）

Abstract：Based on the generalized proximal algorithm developed by D. Han，we present an improved algorithm for solving variational inequalities，which only needs to solve a linear system of equations approximately at each step. Under mild assumptions，we show the global convergence of the algorithm.

Key words：variational inequalities；linear search；projection；global convergence

设 C 是 $\mathbf{R}^n$ 中非空闭凸子集，$F:\mathbf{R}^n\to\mathbf{R}^n$ 是连续映射。考虑变分不等式问题，求 $x^*\in C$，使得

$$\langle F(x^*),x-x^*\rangle\geqslant 0,\quad \forall x\in C$$

其中 $<\cdot,\cdot>$ 表示 $\mathbf{R}^n$ 中向量的内积。该问题简记为 VIP（F，C）。在数学规划、经济金融、交通等领域有重要的应用，是近几十年运筹学研究的一个热点[1]。

在众多求解变分不等式问题的算法中，投影算法是最重要的一种，因其在计算过程中小的计算量及存储而引起人们的兴趣。投影算法的主要迭代步骤为 $x^{k+1}=P_c[x^k-\alpha F(x^k)]$，其中 $\alpha>0$，P_C 是 $\mathbf{R}^n$ 到 C 上的投影算子。该算法的收敛性需要较强的假设条件，映射函数 F 强单调和 Lipschitz 连续，且步长满足某一条件[2]。这限制了投影算法的应用。后来 Korpelevich 创立了外梯度投影算法[3]，它的收敛性仅需要映射函数 Lipschitz 连续。当映射函数非 Lispschitz 连续，或 Lipschitz 常数未知时，人们借助 Armijo 步长搜索提出了外梯度投影算法的多种变形[4~6]。但是，外梯度投影算法及其变形一般在每一步迭代中需要执行多次投影，从而引起计算量的增大。为克服这一缺点，人们提出了二次梯度投影算法[7,8]。与上述方法不同，D. Han 将 proximal 方法与投影算法相结合提出一个混合的推广近似点算法（generalized proximal algorithm）[9]，每一步迭代只需要执行一次到可行集 C 上的投影。在该算法的每一迭代步，先求解如下关于 $y\in\mathbf{R}^n$ 的 proximal 子问题的近似解 $y^k\in\operatorname{int}(C)$：

$$c_k(F(y)+\nabla h(y)-\nabla h(x^k))+(y-x^k)=r^k \tag{1}$$

满足

$$\|r^k\|\leqslant\sigma\|x^k-y^k\| \tag{2}$$

其中 $h:\mathbf{R}^n\to\mathbf{R}$ 是一个 Bregman 函数[9,10]。下一个迭代点由 $x^{k+1}=(1-t)P_C[x^k-c_k\alpha_kF(y^k)]+tx^k$ 计算得到，其中

$$\alpha_k=\frac{\langle F(y^k),x^k-y^k\rangle}{c_k\|F(y^k)\|^2} \tag{3}$$

在适当条件下，该算法产生的迭代点列 $\{x^k\}$ 全局收敛到 VIP（F，C）的一个解。但是该算法在每一

* 收稿日期：2004-10-13
资助项目：国家自然科学基金（NO. 10171055）；山东省自然科学基金（NO. Y2003A02）
作者简介：王传伟（1968-），男，山东泰安人，硕士研究生，主要从事变分不等式问题方面的研究。

步迭代中都需要解一个(1)式这样的非线性方程组系统，计算量比较大，并且在解方程组过程中不能保证 Newton 步中的逆矩阵非奇异。

为此本文提出的改进算法首先解一个线性方程组系统取代非线性方程组系统(1)式，接着通过 Armijo 线搜索获取步长，最后通过到可行集 C 上的投影产生新的迭代点。且仍然允许在精度标准(2)式下近似求解线性方程组系统。此外，在算法中，还引入较(3)式更长的步长规则[4]，保证每一步产生的迭代点距 VIP(F,C) 的解集更近，将证明，若映射函数 F 单调，且 VIP(F,C) 的解集非空，则算法产生的迭代点列收敛到 VIP(F,C) 的一个解。

1　引理与假设

假设 Ω 是 $\mathbf{R}^n$ 中的一个非空闭凸子集，$x\in\mathbf{R}^n$ 到 Ω 上的投影定义为 $P_\Omega[x]=\arg\min\{\|y-x\| \mid y\in\Omega\}$，其中 $\|\cdot\|$ 是 $\mathbf{R}^n$ 中的 l_2-范数。映射 $P_\Omega:\mathbf{R}^n\to\Omega$ 称为投影算子。投影算子有以下重要性质[11]。

引理 1　假设 Ω 是 $\mathbf{R}^n$ 中的一个非空闭凸子集，对 $\forall x,y\in\mathbf{R}^n$ 和 $z\in\Omega$，成立

1) $\|P_\Omega[x]-P_\Omega[y]\|\leqslant\|x-y\|$；

2) $\|P_\Omega[x]-z\|^2\leqslant\|x-z\|^2-\|P_\Omega[x]-x\|^2$

引理 2　假设 Ω 是 $\mathbf{R}^n$ 中的一个非空闭凸子集，对 $\forall x,d\in\mathbf{R}^n$ 及 $\alpha\geqslant0$，定义 $x(\alpha):=P_\Omega[x-\alpha d]$。则 $\langle d,x-x(\alpha)\rangle$ 关于 $\alpha\geqslant0$ 单调非减。

引理 3　假设 Ω 是 $\mathbf{R}^n$ 中一个非空闭凸子集，对 $\forall x\in\Omega,d\in\mathbf{R}^n$ 及 $\alpha\geqslant0$，定义 $\Psi(\alpha):=\min\{\|y-x+\alpha d\|^2 \mid y\in\Omega\}$ 则 $\Psi'(\alpha)=2\langle d,y-x+\alpha d\rangle$。

在本文以下的讨论过程中，总是假设

1) VIP(F,C) 的解集，记为 S，非空；

2) F 是单调的，即：

$$\langle F(x)-F(y),x-y\rangle\geqslant0,\quad \forall x,y\in\mathbf{R}^n；$$

3) 存在正常数序列 $\{\mu_k\}$ 和 $\mathbf{R}^{n\times n}$ 中半正定矩阵序列 $\{G_k\}$，使得下面线性系统在 C 上有解。

$$F(x)+(G_k+\mu_kI)\cdot(-x)=0\quad \forall x\in C$$

其中 I 为单位矩阵。

2　不精确型投影算法

算法 1

步 0　取参数 $\sigma,\beta,\lambda\in(0,1)$，$\bar{\mu}\in[1/(\lambda(1-\sigma)),\infty)$。任取一个初始点 $x^0\in C$，令 $k:=0$。

步 1　选取 $\mu_k\in[1/(\lambda(1-\sigma)),\bar{\mu}]$ 和半正定矩阵 $G_k\in\mathbf{R}^{n\times n}$ 满足假设 3)，求解下列关于 x 的线性方程组得到试探点 $\bar{x}^k\in C$。

$$F(x^k)+(G_k+\mu_kI)(x-x^k)=r^k \tag{4}$$

满足

$$\|r^k\|\leqslant\sigma\|x^k-\bar{x}^k\| \tag{5}$$

步 2　求 $y^k=x^k+t_k(\bar{x}^k-x^k)$，使得

$$\langle F(y^k),x^k-\bar{x}^k\rangle\geqslant\lambda(1-\sigma)\mu_k\|x^k-\bar{x}^k\|^2 \tag{6}$$

其中 $t_k=\beta^{m_k}$，m_k 是使(6)式成立的最小非负整数。

步 3　计算

$$x^{k+1}=P_C[x^k-\alpha_kF(y^k)] \tag{7}$$

其中步长 α_k 满足

$$\alpha_k\geqslant\alpha_k^1:=\frac{\langle F(y^k),x^k-y^k\rangle}{\|F(y^k)\|^2} \tag{8}$$

并且

$$\langle F(y^k),P_C[x^k-\alpha_kF(y^k)]-y^k\rangle\geqslant0 \tag{9}$$

令 $k:=k+1$，返回步 1。

由 G_k 和 μ_k 的选法可知，方程组(4)的系数矩阵是正定的，有许多有效的算法来求解[12,13]。此外，算法 1 允许在精度标准(5)式下求它的近似解，从而 $\bar{x}^k$ 存在且易得。下面证明线搜索(6)式是可行的。

引理 4　算法 1 中线搜索(6)式是可行的。

证明　由(7)式可知 $x^{k+1}\in C$。如果 $\bar{x}^k=x^k$，由(5)式知 $r^k=0$，从而由(4)式可得 $F(x^k)=0$，这意味着 x^k 是 VIP(F,C) 的一个解，算法结束。下面，假设对任何 $k\geqslant0$ 有 $\|\bar{x}^k-x^k\|>0$。

由于 $\bar{x}^k$ 是(4)式的解，因此对任何 $k\geqslant0$，有

$$\begin{aligned}&\langle F(x^k),x^k-\bar{x}^k\rangle=\\&(x^k-\bar{x}^k)^\top(G_k+\mu_kI)(x^k-\bar{x}^k)+(x^k-\bar{x}^k)^\top r^k\geqslant\\&\mu_k\|x^k-\bar{x}^k\|^2-\|r^k\|\cdot\|x^k-\bar{x}^k\|\geqslant\\&(1-\sigma)\mu_k\|x^k-\bar{x}^k\|^2\end{aligned} \tag{10}$$

其中第一个不等式利用 Cauchy-Schwartz 不等式，第二个不等式由(5)式和 $\mu_k\geqslant1/(\lambda(1-\sigma))>1$ 得到。

假设对某个 $k_0\geqslant0$，(6)式对任何正整数 m 都不成立，即

$$\langle F(x^{k_0}+\beta^m(\bar{x}^{k_0}-x^{k_0})),x^{k_0}-\bar{x}^{k_0}\rangle<\lambda(1-\sigma)\mu_{k_0}\|x^{k_0}-\bar{x}^{k_0}\|^2,\forall m\geqslant1.$$

令 $m\to\infty$，则 $x^{k_0}+\beta^m(\bar{x}^{k_0}-x^{k_0})\to x^{k_0}$，由 F 连续性有

$$\langle F(x^{k_0}),x^{k_0}-\bar{x}^{k_0}\rangle\leqslant\lambda(1-\sigma)\mu_{k_0}\|x^{k_0}-\bar{x}^{k_0}\|^2 \tag{11}$$

由(10)、(11)式以及 $\|x^{k_0}-\bar{x}^{k_0}\|>0$，知 $\lambda\geqslant1$，与 λ 的选法矛盾。则(6)式必终止于某个正步长 t_k。

证毕

只要证明步长规则(8)、(9)式是可行的，则整个算法就是可行的。为此定义

$$x^k(\alpha):=P_C[x^k-\alpha F(y^k)]$$

$$\phi_k(\alpha):=2\alpha\langle F(y^k),x^k-y^k\rangle+\|x^k-x^k(\alpha)-\alpha F(y^k)\|^2-\alpha^2\|F(y^k)\|^2 \tag{12}$$

由引理3可得

$$\phi'_k(\alpha)=2\langle F(y^k),x^k-y^k\rangle+2\langle F(y^k),x^k(\alpha)-x^k+\alpha F(y^k)\rangle-2\alpha\|F(y^k)\|^2=2\langle F(y^k),x^k-y^k\rangle+2\langle F(y^k),x^k(\alpha)-x^k\rangle=2\langle F(y^k),x^k(\alpha)-y^k\rangle.$$

关于 $\phi'_k(\alpha)$,类似于文献[4],有以下引理。

引理5 方程 $\phi'_k(\alpha)=0$ 在 $\alpha>0$ 上是可解的。

证明 由假设3)知 $\bar{x}^k\in C$。当 $x^k\notin S$ 时,有

$$\langle F(y^k),\bar{x}^k-y^k\rangle=(1-t_k)\langle F(y^k),\bar{x}^k-x^k\rangle\leqslant-1(1-t_k)\lambda(1-\sigma)\mu_k\|\bar{x}^k-x^k\|^2<0$$

由 $x^k-y^k=t_k(x^k-\bar{x}^k)$ 和(6)式得

$$\langle F(y^k),x^k-y^k\rangle=t_k\langle F(y^k),x^k-\bar{x}^k\rangle\geqslant t_k\lambda(1-\alpha)\mu_k\|x^k-\bar{x}^k\|^2=\frac{\lambda(1-\sigma)\mu_k}{t_k}\|x^k-y^k\|^2\geqslant\|x^k-y^k\|^2>0 \tag{13}$$

其中第二个不等式根据 $\mu_k\geqslant\dfrac{1}{\lambda(1-\sigma)}$ 和 $t_k\leqslant1$。从而 $\{x\in\mathbf{R}^n\mid\langle F(y^k),x-y^k\rangle<0\}\cap C$,$\{x\in\mathbf{R}^n\mid\langle F(y^k),x-y^k\rangle>0\}\cap C$,$H_k\cap C$ 都是非空的,其中 $H_k:=\{x\in\mathbf{R}^n\mid\langle F(y^k),x-y^k\rangle=0\}$ 是一个超平面。

余下的证明与文献[4]中引理3.1的证明类似,其简要说明如下。

由 $x^k-\alpha F(y^k)\in\{x\in\mathbf{R}^n\mid\langle F(y^k),x-y^k\rangle<0\}$,$\alpha>\alpha_k^1$ 和投影的定义知,存在 $\alpha'_k\geqslant\alpha_k^1>0$,使得

$$P_C[x^k-\alpha'_kF(y^k)]\in\{x\in\mathbf{R}^n\mid\langle F(y^k),x-y^k\rangle<0\}$$

另一方面,由(13)式当 $x^k\notin S$ 时,有

$$P_C[x^k-0\cdot F(y^k)]\in\{x\in\mathbf{R}^n\mid\langle F(y^k),x-y^k\rangle>0\}$$

由投影算子的连续性知,存在 $\alpha_k^2\in(0,\alpha'_k)$,使得

$$x^k(\alpha_k^2)=P_C[x^k-\alpha_k^2F(y^k)]\in H_k\cap C$$

这表明 $\phi'_k(\alpha_k^2)=0$。 证毕

考虑最优化问题 $\max\{\phi_k(\alpha)\mid\alpha\geqslant0\}$,由引理2知,$\phi'_k(\alpha)$ 关于 $\alpha>0$ 单调不增,并且当 x^k 不是 VIP(F,C) 的解时,有

$$\phi'_k(0)=\langle F(y^k),x^k-y^k\rangle\geqslant\|x^k-y^k\|^2>0$$

因此方程 $\phi'_k(\alpha)=0$ 在 $\alpha>0$ 上的任意一个解,都是上述最优化问题的解。又由于

$$\langle x^k(\alpha)-x^k,F(y^k)\rangle\leqslant\alpha\|F(y^k)\|^2$$

所以如果 $\alpha<\alpha_k^1$,那么 $\phi'_k(\alpha)>0$。这就是说,方程 $\phi'_k(\alpha)=0$ 的最小正解,设为 α_k^2,满足 $\alpha_k^2\geqslant\alpha_k^1$,显然有 $\langle F(y^k),x^k(\alpha_k^2)-y^k\rangle\geqslant0$,即 α_k^2 满足(9)式。

综上所述,可知步长规则(8)、(9)式不仅可行,而且具有弹性。即存在 $\alpha_k^2\geqslant\alpha_k^1$,对任意的 $\alpha_k\in[\alpha_k^1,\alpha_k^2]$ 都满足(8)和(9)式。从而算法1可行。

3 算法的全局收敛性

若算法1在有限步内终止,则 x^k 是 VIP(F,C) 的一个解。以下假设算法1产生无穷点列 $\{x^k\}$。

引理6 假设1)~3)成立,$\{x^k\}$ 和 $\{y^k\}$ 是由算法1产生的两个无穷点列,并且对任意 $k\geqslant0$,$x^k\notin S$。则 $-F(y^k)$ 是价值函数 $\frac{1}{2}\|x-x^*\|^2$ 在 x^k 处的下降方向,其中 $x^*\in S$。

证明 由假设3)知 $y^k\in C$。再由 F 的单调性和 $x^*\in S$,有

$$\langle F(y^k),x^k-x^*\rangle=\langle F(y^k),x^k-y^k\rangle+\langle F(y^k),y^k-x^*\rangle\geqslant\langle F(y^k),x^k-y^k\rangle+\langle F(x^*),y^k-x^*\rangle\geqslant\langle F(y^k),x^k-y^k\rangle \tag{14}$$

由(13)、(14)式及 $x^k\notin S$ 可知结论成立。 证毕

关于由(12)式定义的 $\phi_k(\alpha)$,有下面引理。

引理7 1) $\phi_k(\alpha)\geqslant0$,$\alpha\in[0,\alpha_k^2]$;2) $\phi_k(\alpha_k)\geqslant\phi_k(\alpha_k^1)$,$\alpha_k\in[\alpha_k^1,\alpha_k^2]$。

证明 由 $\phi'_k(\alpha)$ 在 $\alpha\geqslant0$ 时连续,单调不增,并且 $\phi'_k(0)=2\langle F(y^k),x^k-y^k\rangle>0$ 和 $\phi'_k(\alpha_k^2)=0$。当 $\alpha\in[0,\alpha_k^2]$ 时,$\phi'_k(\alpha)>0$,又 $\phi_k(0)=0$,从而第一个结论得证。

由结论2)的证明过程知,在 $[0,\alpha_k^2]$ 上,$\phi_k(\alpha)$ 单调不减,又 $\alpha_k\geqslant\alpha_k^1$,所以 $\phi_k(\alpha_k)\geqslant\phi_k(\alpha_k^1)$。 证毕

引理8 假设1)~3)成立,$\{x^k\}$ 和 $\{y^k\}$ 是由算法1产生的两个无穷点列,则1) $\{x^k\}$ 和 $\{y^k\}$ 都有界;2) $\lim\limits_{k\to\infty}(x^k-y^k)=0$。

证明 1)任取 $x^*\in S$。由(7)式得

$$\|x^{k+1}-x^*\|^2=\|P_C[x^k-\alpha_kF(y^k)]-x^*\|^2\leqslant\|x^k-x^*-\alpha_kF(y^k)\|^2-\|x^k-x^{k+1}-\alpha_kF(y^k)\|^2\leqslant\|x^k-x^*\|^2-\phi_k(\alpha_k)\leqslant\|x^k-x^*\|^2-\phi_k(\alpha_k^1) \tag{15}$$

其中第一个不等式根据引理1,第二个不等式根据(14)式和 $\phi_k(\alpha)$ 的定义,第三个不等式根据引理7。

由 $\phi_k(\alpha)$ 和 α_k^1 的定义,得到

$$\phi_k(\alpha_k^1)=(\alpha_k^1)^2\|F(y^k)\|^2+\|x^k-x^k(\alpha_k^1)-\alpha_k^1F(y^k)\|^2\geqslant(\alpha_k^1)^2\|F(y^k)\|^2$$

由(13)、(15)、(16)式和 α_k^1 的定义得

$$\|x^{k+1}-x^*\|^2\leqslant\|x^k-x^*\|^2-\frac{\|x^k-y^k\|^4}{\|F(y^k)\|^2} \tag{17}$$

于是得到

$$\| x^{k+1}-x^* \|^2 \leqslant \| x^k - x^* \|^2, \forall k \geqslant 0 \quad (18)$$

这表明点列 $\{x^k\}$ 有界。

根据 F 的单调性和 Cauchy-Schwartz 不等式,由(13)式得到 $\| F(x^k) \| \geqslant \| x^k - y^k \|$。由 F 的连续性及 $\{x^k\}$ 的有界性立即可推得点列 $\{y^k\}$ 也有界。

2)由 F 连续性及 $\{y^k\}$ 的有界性可知,必存在常数 $M>0$,使得对 $\forall k \geqslant 0$,有 $\| F(y^k) \| \leqslant M$。由(17)式得到 $\frac{1}{M^2}\sum_{k=0}^{\infty} \| x^k - y^k \|^4 < \infty$,则结论成立。　证毕

根据以上讨论,可以建立算法的全局收敛性。

定理1　在假设1)~3)下,算法1产生的点列 $\{x^k\}$ 收敛到 VIP(F,C) 的一个解。

证明　由引理8,点列 $\{x^k\}$ 有界,因此它至少有一个聚点。设 $\hat{x}$ 是 $\{x^k\}$ 的一个聚点,并设 $\{x^{k_j}\}$ 是收敛到 $\hat{x}$ 的相应的子列。由引理8的结论2)及 $y^k = x^k + t_k(\bar{x}^k - x^k)$ 可得

$$\lim_{k\to\infty} t_k \| \bar{x}^k - x^k \| = \lim_{k\to\infty} \| x^k - y^k \| = 0 \quad (19)$$

考虑以下两种情况。首先假设 $\lim\sup_{k\to\infty} t_k > 0$。由(19)式得到 $\lim_{k\to\infty}\inf \| \bar{x}^k - x^k \| = \lim_{k\to\infty} \| x^k - y^k \| = 0$,从而由(5)式得到 $\lim_{k\to\infty}\inf r^k = 0$。

因 F 连续及 $\bar{x}^{k_j}$ 是(4)式的解,在(4)式两边令 $k_j \to\infty$ 取极限,得 $F(\hat{x})=0$,这表明 $\hat{x}$ 是 VIP(F,C) 的一个解。由于(18)式对任意 $x^* \in S$ 都成立,用 $\hat{x}$ 替换(18)式中的 x^*,得到 $\| x^{k+1} - \hat{x} \| \leqslant \| x^k - \hat{x} \|$,由此即得点列 $\{x^k\}$ 收敛到 $\hat{x}$。

其次,考虑另一种情况 $\lim_{k\to\infty} t_k = 0$。在这种情况下,将证明 $\| \bar{x}^k - x^k \|$ 也收敛到0。假设 $\{x^{k_l}\}$ 是 $\{x^k\}$ 的一个无穷子列,并且使得对任 k_l,都有 $\| \bar{x}^{k_l} - x^{k_l} \| \geqslant d > 0$。不失一般性,假设 $\{x^{k_l}\}$ 和 $\{\bar{x}^{k_l}\}$ 都是收敛子列,并令 $\hat{z} = \lim_{k\to\infty}(\bar{x}^{k_l} - x^{k_l})$,则 $\| \hat{z} \| \geqslant d$。

由 t^k 的取法知道(6)式对 $m_k - 1$ 不成立,即

$$\langle F(x^{k_l} + \beta^{m_{k_l}-1}(\bar{x}^{k_l} - x^{k_l})), x^{k_l} - \bar{x}^{k_l} \rangle < \cdot$$
$$\lambda(1-\sigma)\mu_{k_l} \| \bar{x}^{k_l} - x^{k_l} \|^2$$

由上式和(10)式得(对充分大的 k,有 $m_k > 1$)

$(1-\sigma)\mu_{k_l} \| \bar{x}^{k_l} - x^{k_l} \|^2 + \langle F(x^{k_l} + \beta^{m_{k_l}-1}(\bar{x}^{k_l} - x^{k_l})) - F(x^{k_l}), x^{k_l} - \bar{x}^{k_l} \rangle < \lambda(1-\sigma)\mu_{k_l} \| \bar{x}^{k_l} - x^{k_l} \|^2$,令 $k_l \to\infty$ 并应用 F 的连续性,得到

$$(1-\sigma) \| \hat{z} \|^2 \leqslant \lambda(1-\sigma) \| \hat{z} \|^2$$

由 $\| \hat{z} \| \geqslant d > 0$,得到 $\lambda \geqslant 1$,这与 $\lambda \in (0,1)$ 矛盾,从而得到 $\lim_{k\to\infty} \| \bar{x}^k - x^k \| = 0$。

根据和第一种情况类似的分析,得到点列 $\{x^k\}$ 收敛到 VIP(F,C) 的一个解。　证毕

注1　如果映射函数 F 在 C 上 Lipschitz 连续(设常数为 L),则 t_k 有一个正的下界。事实上,因为

$$\langle F(x^k), x^k - \bar{x}^k \rangle = \frac{1}{t_k}\langle F(x^k), x^k - y^k \rangle \geqslant$$
$$\frac{\lambda(1-\sigma)\mu_k}{t_k} \| x^k - \bar{x}^k \|^2$$

其中不等式根据 F 的单调性和(6)式得到,所以对任意 $k \geqslant 1$,如果 $t_k \neq 1$,有

$$\frac{L}{\beta} t_k \| x^k - \bar{x}^k \|^2 \geqslant$$
$$\langle F(x^k) - F(x^k - \frac{t_k}{\beta}(\bar{x}^k - x^k)), x^k - \bar{x}^k \rangle \geqslant$$
$$(t_k^{-1} - 1)\lambda(1-\sigma)\mu_k \| x^k - \bar{x}^k \|^2 \geqslant$$
$$(t_k^{-1} - 1) \| x^k - \bar{x}^k \|^2$$

由 $\| x^k - \bar{x}^k \| \neq 0$ 即得 $\frac{L}{\beta}(t_k)^2 + t_k - 1 \geqslant 0$。由初等代数知识可知,必存在 $t_0 > 0$,使得 $t_k \geqslant t_0$。

注2　算法1所给出的步长 α_k 是一种长步长 $(\alpha_k \geqslant \alpha_k^1)$,它保证新迭代点与 VIP$(F,C)$ 的解集的距离较 $\alpha_k = \alpha_k^1$ 时更近。但是,如果可行集 C 不具备特殊结构,很难给出 α_k 的明确的表达式。下面的引理给出了计算投影 $P_C[x^k - \alpha_k^2 F(y^k)]$ 的一个简单的方法,其证明过程类似于文献[4]中的引理3.2。

引理9　在算法1第 k 步迭代中,如果取 $\alpha_k = \alpha_k^2$,则 $P_C[x^k - \alpha_k^2 F(y^k)] = P_{C\cap H_k}[x^k - \alpha_k^1 F(y^k)]$,其中

$$H_k := \{x \in \mathbf{R}^n \mid \langle F(y^k), x - y^k \rangle = 0\} \quad (20)$$

是一个超平面。

由引理9,得到如下含最优步长的算法。

算法2　将算法1中的步3替换成

$$x^{k+1} = P_{C\cap H_k}[x^k - \alpha_k^1 F(y^k)],$$

其中 α_k^1 由(8)式定义,H_k 由(20)式定义。

显然算法2与算法1有相同的收敛结果。

4　初步数值试验

在本节,对算法1(取 $\alpha_k = \alpha_k^1$)进行初步的数值试验分析。

在下面的数值实验中,在每一迭代步,取 $G_k = F'(x^k)$,并用 MATLAB 中的左除运算来解线性方程组(4)式。这意味着将精确地解线性方程组(4)式,即 $\sigma = 0$。在算法中的其它参数分别取为 $\lambda = 0.95$,$\beta = 0.6$ 和 $\mu_k = 2.5$。取 $\| x^{k+1} - x^k \| \leqslant 10^{-6}$ 为停止标准。所有的程序都用 MATLAB 5.3 编写,并在 PIV

2.0 GHz 个人电脑上运行。

例1[9] 这是个线性互补问题(LCP),$F(x)=Mx+q$,其中

$$M=\begin{pmatrix}1&2&2&\cdots&2&2\\0&1&2&\cdots&2&2\\\square&\square&\square&\square&\square&\square\\0&0&0&\cdots&1&2\\0&0&0&\cdots&0&1\end{pmatrix},q=\begin{pmatrix}-1\\-1\\\square\\-1\end{pmatrix}$$

解 对于这个问题,取初始点为 $x^0=(1,1,\cdots,1)$,表1列出不同维数的计算结果。

表1 例1的计算结果

维数/维	8	16	32	64	128	256
迭代次数/次	5	5	5	5	5	5
CPU 时间/s	0.06	0.10	0.29	0.90	2.82	14.82

例2[9] 约束集 C 和映射 F 分别取为

$$C=\{x\in\mathbf{R}^5\mid\sum_{i=1}^{5}x_i\geqslant 10,x_i\geqslant 0,i=1,2,\cdots,5\}$$

和 $F(x)=\rho D(x)+Mx+q$,其中 M 是 5×5 阶矩阵,其元素在$(-5,5)$内随机产生;向量 q 的元素在$(-10,10)$上一致分布;$D_i(x)=\arctan(x_i-2)$,$i=1,2,\cdots,5$. 参数 ρ 用来控制非线性程度。

与文献[9]中问题转化成一个等价的非线性互补问题不同,本文用 MATLAB 最优工具箱中的 quadratic-program solver quadprog. m 直接执行到可行集上的投影来解本例。表2和表3分别给出了 $\rho=100$ 和 $\rho=200$ 时的计算结果。

表2 $\rho=100$ 时的计算结果

初始点	迭代次数/次	CPU 时间/s
(25,0,0,0,0)	13	0.07
(10,0,10,0,10)	11	0.06
(0,2.5,2.5,2.5,2.5)	10	0.06
(10,0,0,0,0)	11	0.06

表3 $\rho=200$ 时的计算结果

初始点	迭代次数/次	CPU 时间/s
(25,0,0,0,0)	13	0.07
(10,0,10,0,10)	12	0.07
(0,2.5,2.5,2.5,2.5)	9	0.06
(10,0,0,0,0)	11	0.06

参考文献:

[1] FACCHINEI F,PANG J S. Finite-Dimensional Variational Inequalities and Complementarity Problems[M]. New York Springer,2003.

[2] BERTSEKAS D,TSITSIKLIS J N. Parallel and Distributed Computation:Numerical Methods,Prentice-Hall[M]. NJ Englewood Cliffs,1989.

[3] KORPELEVICH E N. The Extragradient Method for Finding Saddle Points and Other Problems[J]. Matecon,1976,(12):747-756.

[4] WANG Y J,XIU N H,WANG C Y. Unified Framework of Extragradient-Type Methods for Pseudomonotone Variational Inqualities[J]. Journal of Optimization Theory and Applications,2001,111:641-656.

[5] IUSEM A N,SVAITER B F. A Variant of Korpelevich´s Method for Variational Inequalities with a New Search Strategy[J]. Optimization,1997,42:309-321.

[6] HE B S. A Class of Projection and Contraction Methods for Monotone Variational Inequalities[J]. Applied Mathematics and Optimization,1997,35:69-76.

[7] WANG Y J,XIU N H,WANG C Y. A New Version of Extragradient Method for Variational Inquality Problems[J]. Computers and Mathematics with Applications,2001,42:969-979.

[8] SOLODOV M V,SVAITER B F. A New Projection Method for Variational Inquality Problems[J]. SIAM Journal on Control and Optimization,1999,37:765-776.

[9] Han D R. A New Hybrid Generalized Proximal Point Algorithm for Variational Inequality Problems[J]. Journal of Global Optimization,2003,26:125-140.

[10] SOLODOV M V,SVAITER B F. An Inexact Hybrid Generalized Proximal Point Algorithm and Some New Results on the Theory of Bregman Functions[J]. Mathematics of Operations Research,2000,25:214-230.

[11] CALAMAI P H,MORE J J. Projected Gradient Methods for Linearly Constrained Problems[J]. Mathematical Programming,1987,39:93-116.

[12] FLETCHER R. Practical Mathods of Optimization[M]. New York:Wiley,1985.

[13] GOLUB G,VAN LOAN C. Matrix Computations[M]. 2nd ed. Baltimore:Johns-Hopkins,1989.

(责任编辑 黄 颖)