第44卷 第7期
2017年7月

计 算 机 科 学
COMPUTER SCIENCE

Vol. 44 No. 7
July 2017

一种非精确求解结构型变分不等式的渐近点算法

陈小彪[1] 李耿华[2] 梁 娟[1] 王建军[1]

(太原工业学院理学系 太原 030008)[1] (重庆大学数学与统计学院 重庆 401331)[2]

摘 要 近来,交替方向法成为了学者们研究的热点。对于一类子问题能够精确求解的变分不等式,该算法是有效的。然而,在实际问题中,变分不等式的子问题是非常困难甚至是不可能精确求解的。在渐近点算法的基础上得到一种非精确的渐近点算法,使得变分不等式子问题具有显式解,通过简单的预测校正步得到子问题的解。在合理的假设下,算法的收敛性得到了证明,一些数值实验表明了所提算法的有效性。

关键词 结构型变分不等式,交替方向法,渐近点算法,预测-校正步法

中图法分类号 O221.2 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.07.047

Inexact Proximal Point Algorithm for Structured Variational Inequalities

CHEN Xiao-biao[1] LI Geng-hua[2] LIANG Juan[1] WANG Jian-jun[1]

(Department of Science, Taiyuan Institute of Technology, Taiyuan 030008, China)[1]

(College of Mathematics and Statistics, Chongqing University, Chongqing 401331, China)[2]

Abstract Recently, the alternating direction method of multipliers has attracted great attention. For a class of variational inequalities, this method is efficient, when the subproblems can be solved exactly. However, the subproblems could be too difficult or impossible to be solved exactly in many practical applications. In this paper, we proposed an inexact proxi-mal point method based on proximal point method. The subproblem is simple to have a closed form solution. Instead of solving the subproblems exactly, we used the simple projection-correction method to approximate the subproblems' real solutions. Convergence of the proposed method is proved under mild assumptions and its efficiency is also verified by some numerical experiments.

Keywords Structured variational inequality, Alternating direction method, Proximal point method, Prediction-correction method

1 引言

自20世纪60年代以来,有限维变分不等式的理论和算法得到了迅速的发展,并且被广泛应用于计算机科学、经济平衡理论、交通运输、社会经济模型、凸优化等领域,详细可参考文献[1-5],因此变分不等式问题的研究和应用已经成为数学中的一个热点。一个可分离结构的变分不等式问题(记作 $VI(\Omega,F)$)即为找到一个向量 $u\in\Omega$,使得:

$$(u'-u)^{\mathrm{T}}F(u)\geqslant 0,\ \forall u'\in\Omega \tag{1}$$

其中

$$u=\begin{pmatrix}x\\y\end{pmatrix},F(u)=\begin{pmatrix}f(x)\\g(y)\end{pmatrix} \tag{2}$$

且

$$\Omega:=\{(x,y):Ax+By=b,x\in X,y\in Y\} \tag{3}$$

其中,$X\subset R^{n_1}$ 和 $Y\subset R^{n_2}$ 是非空闭凸集;$A\in R^{m\times n_1}$ 和 $B\in R^{m\times n_2}$ 是给定的矩阵;$b\in R^m$ 是给定的向量;$f:X\to R^{n_1}$ 和 $g:Y\to R^{n_2}$ 是给定的单调映射,且 $n_1+n_2=n$。

一种求解 $VI(\Omega,F)$ 的基准算法是交替方向法,最初由文献[4]给出证明,且已经得到广泛的研究,可参考文献[5-10]等。对线性约束 $Ax+By=b$ 引入拉格朗日乘子 $\lambda\in R^m$,采用交替方向通过如下的迭代序列法求解问题(2)和问题(3):对于给定的点 $w^k=(x^k,y^k,\lambda^k)$,找到 $\tilde{w}^k=(\tilde{x}^k,\tilde{y}^k,\tilde{\lambda}^k)$,使得

$$\begin{cases}(x'-\tilde{x}^k)^{\mathrm{T}}\{f(\tilde{x}^k)-A^{\mathrm{T}}[\lambda^k-\beta(A\tilde{x}^k+By^k-b)]\}\geqslant 0,\\ \qquad\qquad\qquad\qquad\forall x'\in X\\ (y'-\tilde{y}^k)^{\mathrm{T}}\{g(\tilde{y}^k)-B^{\mathrm{T}}[\lambda^k-\beta(A\tilde{x}^k+B\tilde{y}^k-b)]\}\geqslant 0,\\ \qquad\qquad\qquad\qquad\forall y'\in Y\\ \tilde{\lambda}^k=\lambda^k-\beta(A\tilde{x}^k+B\tilde{y}^k-b)\end{cases}$$

其中,$\beta(\beta>0)$ 是给定的对线性约束的罚参数。

交替方向法能够使大规模的变分不等式问题转化为两个小规模的子变分不等式问题。然而,在许多实际问题中,如果子变分不等式问题难以精确求解,交替方向法或许会失效,很多学者提出了一些策略来解决该问题。文献[11]提出了一种非精确并行分裂算法;文献[12-13]提出了一种非精确交替方向法;文献[14]将非精确交替方向法推广到具有多个算子的凸优化问题中。以上文献都说明了非精确算法在某些问题中的应用。本文提出了一种非精确的渐近点算法,使得子变分不等式问题具有显式解,从而容易求解。

到稿日期:2016-02-25 返修日期:2016-05-16 本文受太原工业学院院青年基金(2015LQ16)资助。

陈小彪(1987-),男,硕士,主要研究方向为最优化理论与算法,E-mail:imchxb@126.com;**李耿华**(1988-),男,博士,主要研究方向为最优化理论;**梁 娟**(1987-),女,硕士,主要研究方向为模型的分岔理论;**王建军**(1968-),男,硕士,教授,主要研究方向为非线性动力系统的稳定性。

2 预备知识

使 $\|\cdot\|:\|\cdot\|_2$ 表示欧氏范数且对 $\forall x\in R^n$，$\|x\|=\sqrt{x^{\mathrm{T}}x}$。对非空闭凸集 Ω，采用 $P_\Omega(\cdot)$ 表示在欧氏范数下向 Ω 的投影：

$P_\Omega(x)=\arg\min\{\|x-y\|\,|y\in\Omega\}$

引理 1 设 $\Omega\subset R^n$ 是非空闭凸集，$P_\Omega(\cdot)$ 为上述所定义的投影，则有：

$(u-P_\Omega(u))^{\mathrm{T}}(P_\Omega(u)-w)\geqslant 0,\forall u\in R^n,\forall w\in\Omega$

本文做如下假设：

假设 1 在欧氏范数意义下，易计算向集合 X 和集合 Y 的投影。

假设 2 算子 $f(x)$ 和 $g(y)$ 在集合 X 和集合 Y 上是单调且 Lipschitz 连续的，不必知道 $f(x)$ 和 $g(y)$ 的 Lipschitz 常数。

假设 3 变分不等式问题(1)－问题(3)的解集非空。

3 求解问题(1)－问题(3)的非精确渐近点算法

为便于分析，记

$$G=\begin{pmatrix}\beta^{-1} & 0 & 0\\ 0 & \beta^{-1}+B^{\mathrm{T}}HB & -B\\ 0 & -B & H^{-1}\end{pmatrix}$$

$$G_1=\begin{pmatrix}(1-v)\beta^{-1} & 0 & 0\\ 0 & (1-v)\beta^{-1}+B^{\mathrm{T}}HB & -B\\ 0 & -B & H^{-1}\end{pmatrix}$$

其中，$\beta>0$，对于给定的正定矩阵 $H\in R^{m\times n}$，一般选择 $H=tI,t>0$；$0<v<1$，G 和 G_1 是正定对阵矩阵，记矩阵 G_1 的最小特征值为 δ_1，则 $G_1=\lambda_{\min}(G_1)=\delta_1>0$。

算法 1 求解问题(1)－问题(3)的非精确渐近点算法

Step1 设 $\varepsilon>0$，$w^0=(x^0,y^0,\lambda^0)\in R^{n_1}\times R^{n_2}\times R^m$

$H\in R^{m\times m}$ 是正定矩阵，$0<v<1,\gamma\in[1,2)$，另 $k=0$

Step2 预测步

对给定的 $w^k=(x^k,y^k,\lambda^k)$，通过如下迭代找到 $\tilde{w}^k=(\tilde{x}^k,\tilde{y}^k,\tilde{\lambda}^k)$。

Step2.1 $\tilde{x}^k=P_X(x^k-\beta[f(x^k)-A^{\mathrm{T}}(\lambda^k-H(Ax^k+By^k-b)])$ (4)

其中，$\beta>0$ 是给定的合适的参数，且满足

$\beta\|\xi_x^k\|\leqslant\theta\|x^k-\tilde{x}^k\|$，$\xi_x^k=f(x^k)-f(\tilde{x}^k)+A^{\mathrm{T}}HA(x^k-\tilde{x}^k)$

Step2.2 $\tilde{\lambda}^k=\lambda^k-H(A\tilde{x}^k+By^k-b)$ (5)

Step2.3 $\tilde{y}^k=P_Y(y^k-\beta[g(y^k)-B^{\mathrm{T}}(\tilde{\lambda}^k-H(A\tilde{x}^k+By^k-b)])$ (6)

其中，$\beta>0$ 是给定的合适的参数，且满足

$\beta\|\xi_y^k\|\leqslant\theta\|y^k-\tilde{y}^k\|$，$\xi_y^k=g(y^k)-g(\tilde{y}^k)+B^{\mathrm{T}}HB(y^k-\tilde{y}^k)$

Step3 校正步

$w^{k+1}=w^k-\alpha d(w^k,\tilde{w}^k,\xi^k)$

其中

$d(w^k,\tilde{w}^k,\xi^k)=G(w^k-\tilde{w}^k)-\xi^k$，$\xi^k=(\xi_x^k,\xi_y^k,0)^{\mathrm{T}}$

步长 α 由如下方法得到：

$$\alpha=\gamma\alpha^*,\alpha^*=\frac{\varphi(w^k,\tilde{w}^k,\xi^k)}{\|d(w^k,\tilde{w}^k,\xi^k)\|^2}$$

和

$$\varphi(w^k,\tilde{w}^k,\xi^k)=\|w^k-\tilde{w}^k\|_G^2-(w^k-\tilde{w}^k)^{\mathrm{T}}\xi^k \tag{7}$$

4 非精确渐近点算法的收敛性分析

引理 2 设 $\varphi(w^k,\tilde{w}^k,\xi^k)$ 由式(7)定义，$\lambda_{\min}(G_1)=\delta_1$，则对任意的 k 可得：

$\varphi(w^k,\tilde{w}^k,\xi^k)\geqslant\delta_1\|w^k-\tilde{w}^k\|^2$

证明：利用 Cauchy-Schwarz 不等式和 ξ^k 的定义可得：

$$-(w^k-\tilde{w}^k)^{\mathrm{T}}\xi^k\geqslant -v\begin{pmatrix}x^k-\tilde{x}^k\\ y^k-\tilde{y}^k\end{pmatrix}^{\mathrm{T}}\begin{pmatrix}\beta^{-1} & 0\\ 0 & \beta^{-1}\end{pmatrix}\begin{pmatrix}x^k-\tilde{x}^k\\ y^k-\tilde{y}^k\end{pmatrix}$$

然后，利用 $\varphi(w^k,\tilde{w}^k,\xi^k)$ 的定义易得：

$\varphi(w^k,\tilde{w}^k,\xi^k)\geqslant\|w^k-\tilde{w}^k\|_{G_1}^2\geqslant\delta_1\|w^k-\tilde{w}^k\|^2$

引理 3 设 $\{w^k\}$ 是由算法 1 得到的序列，且 $w^*\in W^*$，则有：

$(w^k-w^*)^{\mathrm{T}}d(w^k,\tilde{w}^k,\xi^k)\geqslant\varphi(w^k,\tilde{w}^k,\xi^k)$

证明：式(4)－式(6)分别等价于：

$$(x-\tilde{x}^k)^{\mathrm{T}}(f(\tilde{x}^k)-A^{\mathrm{T}}\tilde{\lambda}^k+\beta^{-1}(\tilde{x}^k-x^k)+\xi_x^k)\geqslant 0 \tag{8}$$

$$(\lambda-\tilde{\lambda}^k)^{\mathrm{T}}((A\tilde{x}^k+B\tilde{y}^k-b)-B(\tilde{y}^k-y^k)+H^{-1}(\tilde{\lambda}^k-\lambda^k))\geqslant 0 \tag{9}$$

$$(y-\tilde{y}^k)^{\mathrm{T}}(g(\tilde{y}^k)-B^{\mathrm{T}}\tilde{\lambda}^k+(\beta^{-1}+B^{\mathrm{T}}HB)(\tilde{y}^k-y^k)-B^{\mathrm{T}}(\tilde{\lambda}^k-\lambda^k)+\xi_y^k)\geqslant 0 \tag{10}$$

设 $w^*=(x^*,y^*,\lambda^*)$ 为变分不等式问题(1)－问题(3)的解，$\tilde{x}^k\in X,\tilde{y}^k\in Y$，得到：

$$(\tilde{x}^k-x^*)^{\mathrm{T}}(f(x^*)-A^{\mathrm{T}}\lambda^*)\geqslant 0 \tag{11}$$

$$(\tilde{y}^k-y^*)^{\mathrm{T}}(g(y^*)-B^{\mathrm{T}}\lambda^*)\geqslant 0 \tag{12}$$

将式(8)和式(11)相加，并利用 $f(x)$ 的单调性可得：

$$(\tilde{x}^k-x^*)^{\mathrm{T}}(\beta^{-1}(x^k-\tilde{x}^k)-\xi_x^k)+(\tilde{\lambda}^k-\lambda^*)^{\mathrm{T}}A(\tilde{x}^k-x^*)\geqslant 0 \tag{13}$$

将式(10)和式(12)相加，并利用 $g(y)$ 的单调性可得：

$$(\tilde{y}^k-y^*)^{\mathrm{T}}((\beta^{-1}+B^{\mathrm{T}}HB)(y^k-\tilde{y}^k)-B^{\mathrm{T}}(\lambda^k-\tilde{\lambda}^k)-\xi_y^k)+(\tilde{\lambda}^k-\lambda^*)^{\mathrm{T}}B(\tilde{y}^k-y^*)\geqslant 0 \tag{14}$$

式(9)可改写为：

$$(\tilde{\lambda}^k-\lambda^*)^{\mathrm{T}}(-(A\tilde{x}^k+B\tilde{y}^k-b)-B(y^k-\tilde{y}^k)+H^{-1}(\lambda^k-\tilde{\lambda}^k))\geqslant 0 \tag{15}$$

可将式(13)－式(15)写成如下紧凑的形式：

$$(\tilde{w}^k-w^*)^{\mathrm{T}}(G(w^k-\tilde{w}^k)-\xi^k)\geqslant 0 \tag{16}$$

利用 $d(w^k,\tilde{w}^k,\xi^k)$ 和 $\varphi(w^k,\tilde{w}^k,\xi^k)$ 的定义可得：

$(w^k-w^*)^{\mathrm{T}}d(w^k,\tilde{w}^k,\xi^k)\geqslant\varphi(w^k,\tilde{w}^k,\xi^k)$

此式即为所证。

为便于叙述，记：

$\theta(\alpha)=\|w^k-w^*\|^2-\|w^{k+1}(\alpha)-w^*\|^2$

引理 4 设 $\{w^k\}$ 是由算法 1 得到的序列，且 $w^*\in W^*$，则有：

$\theta(\alpha)\geqslant 2\alpha\varphi(w^k,\tilde{w}^k,\xi^k)-\alpha^2\|d(w^k,\tilde{w}^k,\xi^k)\|_G^2$

证明：

$$\begin{aligned}\theta(\alpha)&=\|w^k-w^*\|^2-\|w^{k+1}(\alpha)-w^*\|^2\\&=\|w^k-w^*\|^2-\|w^k-w^*-\alpha d(w^k,\tilde{w}^k,\xi^k)\|^2\\&=2\alpha(w^k-w^*)^{\mathrm{T}}d(w^k,\tilde{w}^k,\xi^k)-\alpha^2\|d(w^k,\tilde{w}^k,\xi^k)\|^2\\&\geqslant 2\alpha\varphi(w^k,\tilde{w}^k,\xi^k)-\alpha^2\|d(w^k,\tilde{w}^k,\xi^k)\|^2\end{aligned}$$

注：$q(\alpha)=2\alpha\varphi(w^k,\tilde{w}^k,\xi^k)-\alpha^2\|d(w^k,\tilde{w}^k,\xi^k)\|_G^2$ 是 α 的二次函数，当 $\alpha=\alpha^*$ 时，$q(\alpha)$ 取得极大值。若想极大化 $\theta(\alpha)$，由于它含有未知的 w^*，因此需极大化它的下界函数 $q(\alpha)$，这也是选择 α^* 为最优步长的原因所在。

推论 1 设 $\{w^k\}$ 是由算法 1 得到的序列，且 $w^*\in W^*$，则有：

$\|w^{k+1}-w^*\|^2\leqslant\|w^k-w^*\|^2-\gamma(2-\gamma)\alpha^*\delta_1\|w^k-\tilde{w}^k\|^2$

证明：

$$\|w^{k+1}(\alpha_k^*)-w^*\|^2$$
$$\leqslant \|w^k-w^*\|^2-(2\gamma\alpha^*\varphi(w^k,\tilde{w}^k,\xi^k)-\gamma^2(\alpha^*)^2\|d(w^k,\tilde{w}^k,\xi^k)\|^2)$$
$$\leqslant \|w^k-w^*\|^2-\gamma(2-\gamma)\alpha^*\varphi(w^k,\tilde{w}^k,\xi^k)$$
$$\leqslant \|w^k-w^*\|^2-\gamma(2-\gamma)\alpha^*\delta_1\|w^k-\tilde{w}^k\|^2$$

从推论 1 可以看出，$\{w^k\}$是 Feje'r 单调的。与文献[11]中的证明方法类似，可以很容易地得到定理 1。

定理 1　由算法 1 产生的序列$\{w^k\}$收敛于变分不等式问题(1)—问题(3)的解。

5　非精确渐进点算法的数值实验检验

本节通过数值实验来验证算法 1 的有效性，将所提算法与文献[11]中的 IPSALM 算法和文献[12]中的 IADM 算法进行比较。取相关参数为：

$\varepsilon=10^{-5},\beta=3^{-1},H=I,\gamma=1.2$

考虑问题：

$$\min \frac{1}{2}\|X-C\|^2+\frac{1}{2}\|Y-C\|^2$$
$$\text{s. t. } X-Y=0$$
$$X\in\{H\in R^{n\times n}\mid H^{\mathrm{T}}=H,H\geqslant 0\}$$
$$Y\in\{H\in R^{n\times n}\mid H^{\mathrm{T}}=H,H_L\leqslant H\leqslant H_U\}$$

实验结果如表 1 所列。

表 1　3 种算法迭代次数和时间的对比

n	IPSALM		IADM		算法 1	
	迭代次数	迭代时间/s	迭代次数	迭代时间/s	迭代次数	迭代时间/s
100	81	0.84	74	0.63	59	0.59
200	105	5.61	109	4.81	73	3.54
300	124	17.28	121	13.80	79	10.79
400	133	43.10	136	34.65	95	23.51
500	148	89.55	151	72.24	97	48.83
800	192	412.25	173	311.52	111	212.55

结束语　本文在渐近点算法的基础上得到了一种非精确的渐近点算法，使得子变分不等式问题具有显式解，因而容易求解，数值实验也表明了算法的有效性。参数的选取对算法也有一定的影响，如何选取参数能加快算法的收敛速度还需进一步的讨论。

参考文献

[1] BOYD S,PARIKH N,CHU E,et al. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. Foun Trends Mach Learn,2011,3(1):1-122.

[2] NAGURNEY A,ZHANG D. Projected Dynamical System and Variatioal Inequalities with Applications[M]. Kluwer Academic,Boston,1996.

[3] FUKUSHIMA M. Application of the alternating directions method of multipliers to separable convex programming problems [J]. Computational Optimization and Applications,1992,1(1):93-111.

[4] GLOWINSKI R,MARROCCO A. Sur l'approximation par éléments finisd' ordre un et la résolution par pénalisation-dualité d'une classe de problémes de Dirichlet non linéaires[J]. Journal of Equine Veterinary Science,1975,31(5):41-76.

[5] ECKSTEIN J,BERTSEKAS D P. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators[J]. Mathematical Programming, 1992, 55: 293-318.

[6] HE B S,LIAO L Z,HAN D R,et al. A new inexact alternating directions method for monotone variational inequalities[J]. Mathematical Programming,2002,92(1):103-118.

[7] YE C H,YUAN X M. A descent method for structured monotone variational inequalities [J]. Optimization Methods Software,2007,22(2):329-338.

[8] HE B S,TAO M,YUAN X M. Alternating direction method with Gaussian back substitution for Separable convex programming[J]. SIAM J. Optimization,2012,22(2):313-340.

[9] HE B S. Parallel splitting augmented Lagrangian methods for monotone structured variational inequalities Computational[J]. Optimization and Applications,2009,42(2):195-212.

[10] CAI X J,GU G Y,HE B S,et al. A proximal point algorithm revisit on the alternating direction method of multipliers[J]. Science China Mathematics,2013,56(10):2179-2186.

[11] TAO M,YUAN X M. An inexact parallel splitting augmented Lagrangian method for monotone variational inequalities with separable structures[J]. Computational Optimization and Applications,2012,52(2):439-461.

[12] HE B S,LIAO L Z,YUAN X M. Alternating projection based prediction-correction methods for structured variational inequalities[J]. Computational Mathematics,2006,24(6):693-710.

[13] CHEN Z M,WAN L,YANG Q Z. An Inexact Direction Method for Structured Variational Inequalities[J]. Journal of Optimization Theory & Applications,2014,163(2):439-459.

[14] GU G Y,HE B S,YANG J F. Inexact Alternating Direction-Based Contraction Methods for Separable Linearly Constrained Convex Optimization[J]. Journal of Optimization Theory and Applications,2013,163(1):105-129.

(上接第 266 页)

[8] LI W J,ZHAO H,ZHANG Y,et al. Research on massive data mining technology based on MapReduce[J]. Computer Engineering and Applications,2013(20):113-114. (in Chinese)
李伟卫,赵航,张阳,等. 基于 MapReduce 的海量数据挖掘技术研究[J]. 计算机工程与应用,2013(20):113-114.

[9] CHANG F,DEAN J,GHEMAWAT S,et al. Bigtable:A distributed storage system for structured data [J]. ACM Transactions on Computer Systems (TOCS),2008,26(2):4-10.

[10] LIU S J. Research of Frequent Itemsets Mining Alogorithm Based on MapReduce Calculation Model[D]. Harbin:Harbin University of Science and Technology,2015:39-50. (in Chinese)
刘士佳. 基于 MapReduce 框架的频繁项集挖掘算法研究[D]. 哈尔滨:哈尔滨理工大学,2015:39-50.

[11] KUANG S H,LI B. Analysis Cloud Computing Architecture and its Application[J]. Computer&Digital Engineering,2010,3(6):61-63. (in Chinese)
框胜徽,李勃. 云计算体系结构及应用实例分析[J]. 计算机与数字工程,2010,3(6):61-63.

[12] BONCHI F,LUCCHESE C. Pushing tougher constraints in frequent patternvmining[M]// Advances in Knowledge Discovery and Data Mining. Springer Berlin Heidelberg,2005:114-124.

[13] BCHEUNG D W,HAN J,NG V T,et al. Maintenance of discovered association rules in largedatabases: An incremental updating technique[C]//Proceedings of the Twelfth International Conference on Data Engineering,1996. IEEE,1996,106-114.