广西科学 Guangxi Sciences 2007,14(4):371～373

求解极大单调算子零点的一个近似邻近点算法*

An Approximate Proximal Point Algorithm for Maximal Monotone Operators

唐国吉
TANG Guo-ji

(广西民族大学数学与计算机科学学院,广西南宁 530006)
(Mathematics and Computer Science College, Guangxi University for Nationalities, Nanning, Guangxi, 530006, China)

摘要:在 $T^{-1}(0)\cap C\neq\varnothing$ 的条件下,结合文献[5]的思想给出一个求解 $\bar{x}\in T^{-1}(0)\cap C$ 的近似邻近点算法,并证明新算法的收敛性.该算法的误差准则比较宽松.

关键词:极大单调算子 邻近点算法 收敛性
中图法分类号:O224 **文献标识码**:A **文章编号**:1005-9164(2007)04-0371-03

Abstract: Suppose $T^{-1}(0)\cap C\neq\varnothing$. In this paper, we construct an approximate proximal point algorithm for solving $\bar{x}\in T^{-1}(0)\cap C$ based on the work of Solodov and Svaiter[5] and prove the convergence of the algorithm. The error criterion in this paper is less restrictive.

Key words: maximal monotone operators, proximal point algorithm, convergence

设 T 是 R^n 空间到其自身的一个集值映射.称 T 是单调的,如果对任意的 $x,y\in D(T)$,任意的 $u\in T(x)$,$v\in T(y)$,总有

$$<u-v,\ x-y>\geqslant 0,$$

其中 $<\cdot,\ \cdot>$ 表示 R^n 中的内积,$D(T)$ 表示 T 的定义域.称 T 是极大单调的,如果 T 是单调的且它的图象 grh (T) 不真包含于其它单调算子的图象中.极大单调算子理论是研究凸优化问题,极小-极大问题,互补问题及变分不等式的重要工具[1~3].

关于极大单调算子理论的中心问题是

$$\text{找 } x\in R^n \text{ 使得 } 0\in T(x). \tag{1}$$

求解问题(1)的一个经典方法是邻近点算法(PPA)[4]:设 $x^i\in R^n$ 是该问题的当前近似解,新的迭代点 x^{i+1} 通过解集值映射方程

$$0\in T(x)+\mu_i(x-x^i) \tag{2}$$

得到,其中 $\mu_i>0$.由于在许多情形下求解问题(2)是不可能的,或者其困难程度与求解原问题(1)几乎是相当的,因此人们转而考虑近似地求解邻近子问题,即找 $x^{i+1}\in R^n$,$v^{i+1}\in T(x^{i+1})$ 使得

$$0=v^{i+1}+\mu_i(x^{i+1}-x^i)+\varepsilon^i, \tag{3}$$

其中 ε^i 表示问题(2)不精确解的误差.在文献[4]中,Rockafellar 给出误差准则

$$\|\varepsilon^i\|\leqslant\sigma_i\mu_i,\ \sum_{i=0}^{\infty}\sigma_i<\infty \tag{4}$$

和

$$\|\varepsilon^i\|\leqslant\sigma_i\mu_i\|x^{i+1}-x^i\|,\ \sum_{i=0}^{\infty}\sigma_i<\infty. \tag{5}$$

其中准则(4)保证了邻近点算法的全局收敛性,准则(5)还附加其它一些条件就保证算法的局部线性收敛率.

1999年,Solodov 和 Svaiter[5] 通过构造一个严格分离当前迭代点 x^i 与解集的超平面,再把当前迭代点 x^i 投影到该超平面上获得新的迭代点 x^{i+1}.文献[5]中不要求松弛因子 $\sigma_i\to 0$,因此其误差准则比文献[4]中的更为宽松,从而算法更便于执行.

然而,人们常常要在一个限定的区域内求极大单调算子的零点.设 C 是 R^n 空间中的一个非空闭凸集,

$$\text{找 } x\in C \text{ 使得 } 0\in T(x). \tag{6}$$

问题(1)是问题(6)中 C 取 R^n 的特殊情形,因此问题(6)是比问题(1)更一般的问题.最近,Yang 和 He[6] 在文献[7,8]的基础上给出了求解问题(6)的一个近似邻近点算法.

本文基于文献[6]的思想,通过修改文献[5]的算

收稿日期:2006-12-28
作者简介:唐国吉(1979-),男,讲师,主要从事泛函分析研究工作.
*广西民族大学青年科学基金项目资助.

法，建立了一个求解问题(6)的近似邻近点算法.

和文献[6]一样，我们假设在闭凸集 C 中 T 的零点集(简记为 $T^{-1}(0)\cap C$) 非空.

1 算法

步骤1:取 $x^0\in C$ 为初始值，$\sigma\in[0,1)$；

步骤2:已知 $x^i\in C,\mu_i>0$，找 $y^i\in R^n$ 使

$$0=v^i+\mu_i(y^i-x^i)+\varepsilon^i,v^i\in T(y^i),\tag{7}$$

其中

$$\|\varepsilon^i\|\leqslant\sigma\max\{\|v^i\|,\mu_i\|y^i-x^i\|\};\tag{8}$$

步骤 3:若 $y^i=x^i$，停止；否则，计算

$$x^{i+1}=P_C[x^i-\frac{<v^i,x^i-y^i>}{\|v^i\|^2}v^i],\tag{9}$$

其中 $\mu_i\leqslant\bar{\mu}<\infty$，$P_C(\cdot)$ 表示闭凸集 C 上的投影算子.

注 假设算法在第 i 步迭代时停止，即 $y^i=x^i$，由(7)式可得 $v^i=-\varepsilon^i$，从而 $\|v^i\|=\|\varepsilon^i\|$，另一方面，由(8)式知 $\|\varepsilon^i\|\leqslant\sigma\|v^i\|$，因此 $(1-\sigma)\|v^i\|\leqslant 0$，由 $\sigma\in[0,1)$ 我们可以推出 $v^i=0$，因此有 $0\in T(x^i)$，而 $x^i\in C$，这时 x^i 是问题(6)的一个解.

以下总假设算法产生无穷序列 $\{x^i\}$.

2 定义及引理

在算法中，我们首先关心算法的可定义性，其一是(7)式中的 y^i 是否存在. 我们可以把(7)式等价地写成

$$x^i-\frac{\varepsilon^i}{\mu_i}\in y^i+\frac{1}{\mu_i}T(y^i).\tag{10}$$

引理2.1对 y^i 的存在性给出了肯定的回答.

引理2.1[9] 设 β 为一正数，R^n 上的算子是单调的，当且仅当其预解式 $J_\beta=(I+\beta T)^{-1}$ 是严格非扩张的. 进一步，T 是极大单调的，当且仅当 J_β 是严格非扩张的，而且 $Dom(J_\beta)=R^n$.

其二，由于 C 是闭凸集，则(9)式中的 x^{i+1} 是可定义的.

定义2.1[10] 设 C 是 R^n 空间中的非空闭凸集，投影算子 $P_C(\cdot)$ 定义为

$$P_C(z)=\operatorname{argmin}\{\|z-x\|:x\in C\},$$

其中 $\|\cdot\|$ 表示 Euclid 范数.

引理2.2[10] 设 C 是 R^n 空间中的非空闭凸集，投影算子 $P_C(\cdot)$ 有如下性质：

(i) $<z-P_C(z),\ P_C(z)-y>\geqslant 0,\forall\ z\in R^n,\forall\ y\in C$；

(ii) $\|P_C(y)-P_C(z)\|\leqslant\|y-z\|,\forall\ y,z\in R^n$；

(iii) $\|P_C(y)-x\|^2\leqslant\|y-x\|^2-\|y-P_C(y)\|^2,\forall\ x\in C$.

3 算法的收敛性

定理3.1 设 $\{x^i\},\{y^i\},\{\varepsilon^i\}$ 是由算法产生的序列，则

(a) $\{x^i\}$ 为有界序列；

(b) $\lim\limits_{i\to\infty}\|x^i-y^i\|=0$；

(c) $\lim\limits_{i\to\infty}v^i=0$.

证明 (a)记

$$H_i:=\{x\in R^n\mid<v^i,x-y^i>=0\},$$
$$H_i^*:=\{x\in R^n\mid<v^i,x-y^i>\leqslant 0\},$$
$$z^i:=x^i-\frac{<v^i,x^i-y^i>}{\|v^i\|^2}v^i=P_{H_i}(x^i).$$

我们先证明 H_i 严格分离当前迭代点 x^i 和问题(6)的解集 S. 对任意的 $x^*\in S=T^{-1}(0)\cap C$，由 T 的单调性有

$$<v^i,x^*-y^i>\leqslant 0.\tag{11}$$

还需要证明

$$<v^i,x^i-y^i>>0.\tag{12}$$

为此，我们对误差准则分两种情形讨论.

情形1 若 $\mu_i\|y^i-x^i\|\geqslant\|v^i\|$，则有 $\|\varepsilon^i\|\leqslant\sigma\mu_i\|y^i-x^i\|$. 由(7)式和 Cauchy-Swarz 不等式可得

$$<v^i,x^i-y^i>=<\mu_i(x^i-y^i)-\varepsilon^i,x^i-y^i>=\mu_i\|x^i-y^i\|^2-<\varepsilon^i,x^i-y^i>\geqslant\mu_i\|x^i-y^i\|^2-\|\varepsilon^i\|\|x^i-y^i\|\geqslant\mu_i(1-\sigma)\|x^i-y^i\|^2>0;\tag{13}$$

情形2 若 $\mu_i\|y^i-x^i\|<\|v^i\|$，则有 $\|\varepsilon^i\|\leqslant\sigma\|v^i\|$. 由(7)式和 Cauchy-Swarz 不等式可得

$$<v^i,x^i-y^i>=\frac{1}{\mu_i}<v^i,v^i+\varepsilon^i>\geqslant\frac{1}{\mu_i}(\|v^i\|^2-\|v^i\|\|\varepsilon^i\|)\geqslant\frac{1-\sigma}{\mu_i}\|v^i\|^2>0.\tag{14}$$

这样我们证明了超平面 H_i 是严格分离当前迭代点 x^i 和问题(6)的解集 S 的. 因此 $z^i=P_{H_i}(x^i)=P_{H_i^*}(x^i)$. 又因为 $x^*\in H_i^*$，在引理2.2的性质(i)中令 $z=x^i,C=H_i^*,y=x^*$ 可得 $<x^i-z^i,z^i-x^*>\geqslant 0$. 再注意到 $x^{i+1}=P_C(z^i),x^*\in C$，在引理2.2的性质(iii)中令 $y=z^i,x=x^*$ 可得

$$\|x^{i+1}-x^*\|^2\leqslant\|z^i-x^*\|^2-\|z^i-x^{i+1}\|^2\leqslant\|z^i-x^*\|^2,\tag{15}$$

从而

$$\|x^i-x^*\|^2=\|x^i-z^i\|^2+\|z^i-x^*\|^2+2<x^i-z^i,z^i-x^*>\geqslant\|x^i-z^i\|^2+\|z^i-x^*\|^2=(\frac{<v^i,x^i-y^i>}{\|v^i\|})^2+\|z^i-x^*\|^2\geqslant(\frac{<v^i,x^i-y^i>}{\|v^i\|})^2+\|x^{i+1}-x^*\|^2.\tag{16}$$

因此序列 $\{\|x^i-x^*\|\}$ 是单调递减的. 由单调有界原

理知 $\{\|x^i - x^*\|\}$ 收敛,从而得到 $\{x^i\}$ 是有界的.

(b)先证

$$\frac{<v^i, x^i - y^i>}{\|v^i\|} \geqslant (1-\sigma)\|x^i - y^i\|. \quad (17)$$

若 $\mu_i\|y^i - x^i\| \geqslant \|v^i\|$, 由(13)式可得

$$\frac{<v^i, x^i - y^i>}{\|v^i\|} \geqslant \frac{\mu_i(1-\sigma)\|x^i - y^i\|^2}{\|v^i\|} \geqslant \frac{\mu_i(1-\sigma)\|x^i - y^i\|^2}{\mu_i\|y^i - x^i\|} = (1-\sigma)\|x^i - y^i\|;$$

若 $\mu_i\|y^i - x^i\| < \|v^i\|$,由(14)式可得

$$\frac{<v^i, x^i - y^i>}{\|v^i\|} \geqslant \frac{1-\sigma}{\mu_i}\|v^i\| > (1-\sigma)\|x^i - y^i\|.$$

这样,就证明了(17)式成立. 把(17)式代入(16)式并整理可得

$$(1-\sigma)^2\|x^i - y^i\|^2 \leqslant \|x^i - x^*\|^2 - \|x^{i+1} - x^*\|^2,$$

从而

$$(1-\sigma)^2\sum_{i=0}^{n}\|x^i - y^i\|^2 \leqslant \sum_{i=0}^{n}(\|x^i - x^*\|^2 - \|x^{i+1} - x^*\|^2) = \|x^0 - x^*\|^2 - \|x^{n+1} - x^*\|^2 \leqslant \|x^0 - x^*\|^2.$$

令 $n \to \infty$ 得到 $(1-\sigma)^2\sum_{i=0}^{\infty}\|x^i - y^i\|^2 \leqslant \|x^0 - x^*\|^2$,因此 $\lim\limits_{i\to\infty}\|x^i - y^i\| = 0$.

(c)先证

$$\frac{<v^i, x^i - y^i>}{\|v^i\|} \geqslant \frac{1-\sigma}{\bar{\mu}}\|v^i\|. \quad (18)$$

若 $\mu_i\|y^i - x^i\| \geqslant \|v^i\|$,由(13)式可得

$$\frac{<v^i, x^i - y^i>}{\|v^i\|} \geqslant \frac{\mu_i(1-\sigma)\|x^i - y^i\|^2}{\|v^i\|} \geqslant \frac{1-\sigma}{\mu_i}\|v^i\| \geqslant \frac{1-\sigma}{\bar{\mu}}\|v^i\|;$$

若 $\mu_i\|y^i - x^i\| < \|v^i\|$,由(14)式可得

$$\frac{<v^i, x^i - y^i>}{\|v^i\|} \geqslant \frac{1-\sigma}{\mu_i}\|v^i\| \geqslant \frac{1-\sigma}{\bar{\mu}}\|v^i\|.$$

这样,就证明了(18)式成立. 把(18)式代入(16)式并整理可得

$$\left(\frac{1-\sigma}{\bar{\mu}}\right)^2\|v^i\|^2 \leqslant \|x^i - x^*\|^2 - \|x^{i+1} - x^*\|^2.$$

类似于(b)的证法可得 $\lim\limits_{i\to\infty}v^i = 0$.

定理3.2 设 $\{x^i\}, \{y^i\}, \{v^i\}, \{\varepsilon^i\}$ 是由算法产生的序列,则 $\{x^i\}$ 收敛于问题(6)的一个解.

证明 由定理3.1的第一个结论知, $\{x^i\}$ 有界,因此至少存在一个聚点 $\bar{x}$,设子序列 $\{x^{i_k}\}$ 收敛于 $\bar{x}$,由定理3.1的(b)知, $\{y^{i_k}\}$ 也收敛于 $\bar{x}$, 由 T 的单调性,对任意的 $x \in R^n, u \in T(x)$,我们有 $0 \leqslant < x - y^{i_k}, u - v^{i_k} > = < x - y^{i_k}, u > - < x - y^{i_k}, v^{i_k} >$. 令 $k \to \infty$,有 $y^{i_k} \to \bar{x}, v^{i_k} \to 0$,于是 $< x - \bar{x}, u > \geqslant 0$. 由 T 的极大单调性知 $0 \in T(\bar{x})$. 由于 $\{x^i\} \subset C$,又 $x^{i_k} \to \bar{x}$,注意到 C 是闭集,因此 $\bar{x} \in C$,从而 $\bar{x} \in T^{-1}(0) \cap C$.

为了证明 $x^i \to \bar{x} (i \to \infty)$, 在(16)式中令 $x^* = \bar{x}$ 知 $\{\|x^i - \bar{x}\|\}$ 收敛,再由 $x^{i_k} \to \bar{x}$ 知 $\|x^{i_k} - \bar{x}\| \to 0$. 又因为一个收敛数列与其任一子列的极限相等,故 $\|x^i - \bar{x}\| \to 0$,即 $\lim\limits_{i\to\infty}x^i = \bar{x}$.

参考文献:

[1] BREZIS H. Operateurs maximaux monotone et semi-groups de contractions dans les espaces de Hilbert[M]. Ameterdan:North-Holland,1973.

[2] BURACHIK R S,IUSEM A N,SVAITER B F. Enlargementof monotone operators with applications to variational inequalities [J]. Set-Valued Analysis, 1997 (5):159-180.

[3] HESTENES M R. Multiplier and gradient methods[J]. Journal of Optimization Theory and Applications, 1969 (4):303-320.

[4] ROCKAFELLAR R T. Monotone operators and the proximal point algorithm [J]. SIAM J Control Optim, 1976,14:877-898.

[5] SOLODOV M V,SVAITER B F. A hybrid projection-proximal point algorithm [J]. Journal of Convex Analysis,1999(6):59-70.

[6] YANG Z H,HE B S. A relaxed approximate proximal point algorithm[J]. Annals of Operations Research,2005 (133):119-125.

[7] ECKCTEIN J,BERTSEKAS D P. On the douglas-rachford splitting method and the proximal points algorithm for maximal monotone operators [J]. Mathematical Programming,1992(55):293-318.

[8] HAN DEREN,HE BINGSHENG. A new accuracy criterion for approximate proximal point algorithms[J]. J Math Anal and Appl,2001(263):343-354.

[9] 何炳生,廖立志,杨振华. 极大单调算子的一个新的近似邻近点算法[J]. 中国科学:A 辑,2002,32(11):1026-1032.

[10] HE B S,YANG Z H,YUAN X M. An approxiamte proximal-extragradient type method for monotone variational inequalities[J]. J Math Anal and Appl,2004 (300):362-374.

(责任编辑:尹　闯　邓大玉)