Vol. 39 (2019)
No. 6

数 学 杂 志
J. of Math. (PRC)

一种求解单调包含问题的惯性混合邻近外梯度算法

何明明, 彭建文
(重庆师范大学数学科学学院, 重庆 401331)

摘 要: 本文研究了求解单调包含问题的一种新的惯性混合邻近外梯度算法. 利用 Opial 定理, 获得了惯性混合邻近外梯度算法的弱收敛性和非渐近全局收敛率. 在惯性混合邻近外梯度算法的框架下, 本文提出并分析了惯性 Tseng's 向前向后算法和惯性非精确 Spingarn's 部分逆算法的收敛性和非渐近全局收敛率.

关键词: 混合邻近外梯度算法; 惯性; Tseng's 向前向后算法; Spingarn's 部分逆算法
MR(2010) **主题分类号**: 47J05; 90C60; 90C30; 65K10　　**中图分类号**: O221.2
文献标识码: A　　**文章编号**: 0255-7797(2019)06-0931-15

1 引言

对于单调包含问题, 即: 找到 $x^* \in \mathbb{X}$ 使得

$$0 \in B\,(x^*), \tag{1.1}$$

其中 $B:\mathbb{X} \rightrightarrows \mathbb{X}$ 是极大单调算子, $\mathbb{X}$ 是实 Hilbert 空间. 凸优化问题, 平衡问题和单调变分不等式等许多问题都可以归结为单调包含问题 (1.1). 单调包含问题在图像处理和聚类等实际问题中有着广泛的应用 [1].

邻近点算法是解单调包含问题的一类经典算法, 它由 Martinet[2] 提出, Rockafellar[3] 进一步发展而成. 其迭代格式为

$$x_k \in (Id + \lambda_k B)^{-1}(x_{k-1}), \tag{1.2}$$

其中正则参数 $\lambda_k > 0$, Id 是单位算子. Alvarez 和 Attouch 于 2001 年在文献 [4] 中提出解含有单调算子的广义方程的惯性邻近点算法, 并证明了该算法的弱收敛性.

混合邻近外梯度算法是由 Solodov 和 Svaiter[5] 在 1999 年针对单调包含问题提出的另一类算法. 即

初始化 任取初始点 $x_0 \in X$, 任意给定 $\overline{\sigma} \in [0,1)$, 令 $k=1$.

步骤 1 选择 $\lambda_k \geq \underline{\lambda}$, 容差参数 $\sigma_k \in [0,\overline{\sigma})$, 并找到 y_k, $v_k \in X$, $\varepsilon_k \geq 0$ 使其满足

$$v_k \in T^{[\varepsilon_k]}\left(y_k\right),\ \left\|\lambda_k v_k + y_k - x_k\right\|^2 + 2\lambda_k\varepsilon_k \leq \sigma_k^2 \left\|y_k - x_k\right\|^2;$$

步骤 2 执行一个外梯度步: $x_{k+1} = x_k - \lambda_k v_k$. 令 $k := k+1$, 返回步骤 1.

*收稿日期: 2019-05-19　　接收日期: 2019-09-05
作者简介: 何明明 (1995–), 男, 安徽怀宁, 硕士, 主要研究方向: 最优化理论及其应用.
通讯作者: 彭建文

实质上, 混合邻近外梯度算法是邻近点算法的一种不精确版本, 在每一次迭代中, 对应的近似子问题都应在一个相对误差准则内求解, 这与 Rockafellar 提出的邻近点算法的可加误差准则不同. 混合邻近外梯度算法的另一个特点是它还允许通过文献 [6] 中提出的极大单调算子的 ε-enlargement 来松弛单调算子.

混合邻近外梯度算法作为设计新方法和分析现有方法的复杂性的框架特别有用. 例如 Monteiro 和 Svaiter[7] 在 2013 年提出块状分解混合邻近外梯度算法, 并证明了该算法的迭代复杂性, 且在该算法的框架下证明了交替方向乘子法的遍历迭代复杂度; Goncalves, Melo 和 Monteiro[8] 在 2017 年提出一种新的正则化混合邻近外梯度算法, 并建立了该算法的迭代复杂性, 且证明了交替方向乘子法的变体是该算法的特殊实例; Alves 和 Svaiter[9] 在 2016 年提出了一种新的求解强单调包含问题的混合邻近外梯度算法:

初始化 任取初始点 $x_0 \in \mathbb{X}$, 任意给定 $\sigma \in [0,1)$, $k=1$.

步骤 1 选定 $0 < \lambda_k \le \lambda_{k-1}$, 并找到 y_k, $v_k \in \mathbb{X}$, $\varepsilon_k \ge 0$, 使得

$$v_k \in A\left(y_k\right) + B^{[\varepsilon_k]}\left(y_k\right),\ \frac{\left\|\lambda_k v_k + y_k - x_{k-1}\right\|^2}{1+2\lambda_k\mu} + 2\lambda_k\varepsilon_k \le \sigma^2\|y_k - x_{k-1}\|^2;$$

步骤 2 计算 x_k:

$$x_k = \frac{(x_{k-1} - \lambda_k v_k) + 2\lambda_k\mu y_k}{1+2\lambda_k\mu}, \text{令} k := k+1, \text{返回步骤 1}.$$

并用混合邻近外梯度算法的框架证明了 Korpelevich 外梯度算法的迭代复杂性.

近年来, 惯性邻近点类算法被大量用于设计和分析各种具有惯性的一阶邻近算法. 例如 Chen[10] 等人提出求解具有可分结构的线性约束凸优化问题的惯性交替方向乘子法, 并建立了渐近收敛率和非渐近收敛率; Bot, Csetnek 和 Hendrich 在文献 [11] 中提出了求解单调包含问题的惯性 Douglas-Rachford 分裂算法, 并证明了该算法在希尔伯特空间中的收敛性; Bot 和 Csetnek[12] 在 2015 年提出求解单调包含问题的惯性混合邻近外梯度算法, 并建立了该算法的弱收敛性和渐近收敛性; Alves 和 Marcavillaca 于 2018 年在文献 [1] 中提出一种求解单调包含问题的新的惯性混合邻近外梯度算法, 并得到了它的渐近收敛率和非渐近全局收敛率.

受上述文献的启发, 本文提出了一种新的求解单调包含问题的惯性混合邻近外梯度算法, 并得到了它的收敛性和非渐近全局收敛率. 不同的是, 本文中的惯性系数的范围得到了扩大, 其系数区间为 $[0\,,1+2\underline{\lambda}\mu]$. 特别的, 当 $\mu=0$ 时, 惯性区间为 $[0\,,1]$, 正好与文献 [12] 中的区间一致.

本文结构如下: 首先是预备知识, 对概念和符号进行说明, 并对证明过程中所需的定理进行说明; 然后是本文的主要内容, 本文提出了一种用于求解单调包含问题的惯性混合邻近外梯度算法, 并证明了它的非渐近全局收敛率. 最后本文提出了惯性混合邻近外梯度算法的两种特殊情况: 求解含有 Lipschitz 连续算子的强单调包含问题的惯性 Tseng's 向前向后算法, 并建立了该算法的弱收敛性和非渐近全局收敛率; 求解含有强单调算子和 Lipschitz 连续算子的原始－对偶问题的惯性非精确 Spingarn's 部分逆算法, 并得到了该算法的弱收敛性和非渐近全局收敛率.

2 预备知识

令 $\mathbb{X}$ 是内积为 $\langle\cdot\ \cdot\rangle$, 范数为 $\|\cdot\|:=\sqrt{\langle\cdot\ \cdot\rangle}$ 的实 Hilbert 空间. 对于集值算子 $S:\mathbb{X}\rightrightarrows\mathbb{X}$, 它的图和定义域分别为

$$G_r(S):=\{(x\,,\,v)\in\mathbb{X}\times\mathbb{X}:v\in S(x)\},\ \mathrm{Dom}(S):=\{x\in\mathbb{X}:S(x)\neq\emptyset\}.$$

若 $S^{-1}(v)=\{x:v\in S(x)\}$, 则称 $S^{-1}:\mathbb{X}\rightrightarrows\mathbb{X}$ 是集值算子 S 的逆.

本文主要研究单调包含问题的求解算法, 即含有 (极大) 单调算子的包含问题. 在本节接下来的内容中将回顾一些基本概念和结论.

定义 2.1 [12] 若集值算子 $A:\mathbb{X}\rightrightarrows\mathbb{X}$ 满足

$$\langle v-v',x-x'\rangle\geq\mu\|x-x'\|^2,\ \forall\,v\in A(x),\ v'\in A(x'),\tag{2.1}$$

则称算子 A 是 μ - 强单调的. 特别地, 如果 $\mu=0$ 则称 A 是单调的.

定义 2.2 [12] 设算子 $A:\mathbb{X}\rightrightarrows\mathbb{X}$ 是单调的, 如果不存在其他单调算子 B 真包含于 A, 则称算子 A 是极大单调的. 即如果单调算子 $B:\mathbb{X}\rightrightarrows\mathbb{X}$ 满足 $Gr(A)\subset Gr(B)$, 则 $A=B$.

集值算子 $S,\,S':\mathbb{X}\rightrightarrows\mathbb{X}$ 的和 $S+S':\mathbb{X}\rightrightarrows\mathbb{X}$ 定义为

$$(S+S')(x):=\{s+s'\in\mathbb{X}:s\in S(x),\ s'\in S'(x)\}.$$

显然, 如果 $A:\mathbb{X}\rightrightarrows\mathbb{X}$ 是 μ - 强单调算子, $B:\mathbb{X}\rightrightarrows\mathbb{X}$ 是单调算子, 则 $A+B$ 也是 μ - 强单调算子. 特别的, 两个单调算子的和也是单调算子.

定义 2.3 [6] 设算子 $T:\mathbb{X}\rightrightarrows\mathbb{X}$ 是极大单调的, 如果 $T^{[\varepsilon]}:\mathbb{X}\rightrightarrows\mathbb{X}$ 满足

$$T^{[\varepsilon]}(x):=\{v\in\mathbb{X}:\langle v-v',x-x'\rangle\geq-\varepsilon\ ,\forall x'\in T(v')\},\tag{2.2}$$

则称算子 $T^{[\varepsilon]}$ 是极大单调算子 T 的 ε-enlargement.

注意到 $T\subset T^{[\varepsilon]}$, $\forall x\in\mathbb{X}$. 从而可以将 $T^{[\varepsilon]}$ 视为 T 的外延或近似. 下面是关于 $T^{[\varepsilon]}$ 的性质的结论.

命题 2.1 [13] 令 $T,\,S:\mathbb{X}\rightrightarrows\mathbb{X}$ 是集值映射, 则下列论述成立

(i) 如果 $\varepsilon\leq\varepsilon'$, 则 $T^{[\varepsilon]}(x)\subset T^{[\varepsilon']}(x)$, $\forall x\in\mathbb{X}$.

(ii) $T^{[\varepsilon]}(x)+S^{[\varepsilon']}(x)\subset(T+S)^{[\varepsilon+\varepsilon']}(x)$, $\forall x\in\mathbb{X}$ 和 $\varepsilon,\,\varepsilon'\geq0$.

(iii) T 是单调算子当且仅当 $T(x)\subseteq T^0(x)$.

(iv) T 是极大单调算子当且仅当 $T(x)=T^0(x)$.

(v) 如果 T 是极大单调算子, 点列 $\{(y_k,\,v_k,\,\varepsilon_k)\}$ 使得对于任意的 $k\geq1$ 有 $v_k\in T^{[\varepsilon_k]}(y_k)$, 点列 $\{y_k\}$ 弱收敛于点 x, 点列 $\{v_k\}$ 收敛于点 v 和点列 $\{\varepsilon_k\}$ 收敛于点 ε, 则 $v\in T^{[\varepsilon]}(x)$.

定理 2.2 [12] (Opial 定理) 令 $\emptyset\neq\Omega\subset\mathbb{X}$, 若对于任意的 $x^*\in\Omega$ 和 $\mathbb{X}$ 中的数列 $\{x_k\}$ 使得, $\lim\limits_{k\to\infty}\|x_k-x^*\|$ 存在. 如果 $\{x_k\}$ 的每个弱聚点都属于 Ω, 则数列 $\{x_k\}$ 弱收敛到 Ω 中的点.

引理 2.3 [14] 令 $c,\,d\in\mathbb{X}$, $p\in\mathbb{R}$, 则

$$\|pc+(1-p)\,d\|^2=p\|c\|^2+(1-p)\,\|d\|^2-p\,(1-p)\,\|c-d\|^2.\tag{2.3}$$

3 惯性混合邻近外梯度算法

一般来说, 问题 (1.1) 多半是不适定的, 而对于含有强单调算子的单调包含问题总是适定的. 因此在实际应用中, 通过正则化将问题 (1.1) 近似为强单调包含问题, 这一技术可追溯到 Tikhonov[15] 的工作中. 下面考虑强单调包含问题

$$0 \in A(x) + B(x), \tag{3.1}$$

其中 $A : \mathbb{X} \rightrightarrows \mathbb{X}$ 是极大 μ - 强单调算子, $B : \mathbb{X} \rightrightarrows \mathbb{X}$ 是极大单调算子. 假设问题 (3.1) 有解.

下面给出求解问题 (3.1) 的惯性混合邻近外梯度算法.

算法 1

步骤 0 任取初始点 $x_0 = x_{-1} \in \mathbb{X}$, 任意给定 $\alpha \in [0, 1)$, $\alpha_0 \in [0, \alpha]$, $\sigma \in [0, 1)$, $k = 1$;

步骤 1 选定 $\alpha_k \in [\alpha_{k-1},\ \alpha]$, 计算

$$w_{k-1} := x_{k-1} + \alpha_{k-1}(x_{k-1} - x_{k-2}); \tag{3.2}$$

步骤 2 选定 $0 < \lambda_k \leq \lambda_{k-1}$, 并找到 y_k, $v_k \in \mathbb{X}$, $\varepsilon_k \geq 0$, 使得

$$v_k \in A\left(y_k\right) + B^{[\varepsilon_k]}\left(y_k\right),\ \frac{\left\|\lambda_k v_k + y_k - w_{k-1}\right\|^2}{1+2\lambda_k\mu} + 2\lambda_k\varepsilon_k \leq \sigma^2\|y_k - w_{k-1}\|^2; \tag{3.3}$$

步骤 3 计算 x_k

$$x_k = \frac{(w_{k-1} - \lambda_k v_k) + 2\lambda_k\mu y_k}{1 + 2\lambda_k\mu}, \tag{3.4}$$

令 $k := k + 1$, 返回步骤 1.

注 3.1 (i) 对单调包含问题 (1.1) 进行正则化处理

$$0 \in B\ (x) + \mu(x - x_0), \tag{3.5}$$

其中 $\mu > 0$, x_0 是初始点. 令 $A(x) = \mu(x - x_0)$, 则算子 A 是极大 μ - 强单调算子. 此时问题 (3.5) 是问题 (3.1) 的特殊情况. A 是 μ - 强单调算子, 根据 Minty 定理 [16] 可知它的解集是单点集, 记 x^*_μ 是它的解, 即 $x^*_\mu \in (\mu^{-1}B + Id)^{-1}(x_0)$. 当 μ 接近于 0 时, 它近似于问题 (1.1) 的解.

(ii) 针对问题 (3.1) 也可以用文献 [12] 中的惯性混合邻近外梯度算法, 此时惯性系数的范围从 $[0, 1[$; 而算法 1 通过正则化的技巧将惯性系数的范围从 $[0, 1]$ 扩大为 $[0, 1 + 2\underline{\lambda}\mu]$, 加快了迭代速度.

(iii) 当 $\alpha_k \equiv 0$ 时, 算法 1 退化为文献 [9] 中的求解问题 (3.1) 的算法 1.

(iv) 当 $\mu = 0$ 且 $\alpha_k \equiv 0$ 时, 算法 1 退化为文献 [5] 中的求解问题 (1.1) 的混合邻近外梯度算法.

(v) 当 $\alpha_k \equiv 0$ 且 $\sigma_k \equiv 0$ 时, $x_{k-1} = w_{k-1}$ 由 (3.3) 式可知 $\varepsilon_k \equiv 0$ 和 $x_{k-1} - \lambda_k v_k = y_k$, 从而由 (3.4) 式可知 $x_k = y_k = x_{k-1} - \lambda_k v_k$. 此时, 算法 1 退化为经典的邻近点算法 [3].

(vi) 当 $\sigma_k \equiv 0$ 时, 算法 1 退化为文献 [4] 中的惯性邻近点算法.

引理 3.1 令点列 $\{x_k\}$, $\{y_k\}$, $\{w_k\}$, $\{\alpha_k\}$ 由算法 1 生成, 取 $x \in \mathbb{X}$, 则 $\forall k \geq 1$, 有

$$\begin{aligned}\left\|w_{k-1} - x\right\|^2 = &\left(1 + \alpha_{k-1}\right)\left\|x_{k-1} - x\right\|^2 - \alpha_{k-1}\|x_{k-2} - x\|^2 \\ &+ \alpha_{k-1}\left(1 + \alpha_{k-1}\right)\left\|x_{k-1} - x_{k-2}\right\|^2.\end{aligned} \tag{3.6}$$

证 由 (3.2) 式可得 $w_{k-1}-x=(1+\alpha_{k-1})(x_{k-1}-x)-\alpha_{k-1}(x_{k-2}-x)$. 将 $p=\alpha_{k-1}$, $c=x_{k-2}-x$, $d=x_{k-1}-x$ 代入 (2.3) 式, 即得 (3.6) 式.

命题 3.2 令点列 $\{x_k\}$, $\{y_k\}$, $\{w_k\}$ 由算法 1 生成, 则 $\forall x^*\in(A+B)^{-1}(0)$, 有

$$(1+2\lambda_k\mu)\left\|x_k-x^*\right\|^2+\left(1-\sigma^2\right)\left\|y_k-w_{k-1}\right\|^2\le\left\|w_{k-1}-x^*\right\|^2. \tag{3.7}$$

证 由 (3.4) 式可知 $w_{k-1}=(1+2\lambda_k\mu)x_k-2\lambda_k\mu y_k+\lambda_k v_k$, 事实上, $\forall a,b,c\in\mathbb{X}$, 有 $2\langle a-b,a-c\rangle=\|a-b\|^2+\|a-c\|^2-\|b-c\|^2$, 从而可知

$$\begin{aligned}\left\|w_{k-1}-x\right\|^2=&\left\|w_{k-1}-x_k\right\|^2+\left\|x_k-x\right\|^2+2\left\langle w_{k-1}-x_k,\ x_k-x\right\rangle\\=&\left\|w_{k-1}-x_k\right\|^2+\left\|x_k-x\right\|^2+2\left\langle 2\lambda_k\mu\left(x_k-y_k\right)+\lambda_k v_k,\ x_k-x\right\rangle\\=&\left\|w_{k-1}-x_k\right\|^2+(1+2\lambda_k\mu)\left\|x_k-x\right\|^2\\&+2\left[\lambda_k\mu\left(\left\|x_k-y_k\right\|^2-\left\|x-y_k\right\|^2\right)+\left\langle\lambda_k v_k,\ x_k-x\right\rangle\right]\\=&\left\|w_{k-1}-x_k\right\|^2+2\left(\lambda_k\mu\|x_k-y_k\|^2+\left\langle\lambda_k v_k,\ x_k-y_k\right\rangle-\lambda_k\varepsilon_k\right)\\&-2\left(\lambda_k\mu\|x-y_k\|^2+\left\langle\lambda_k v_k,\ x-y_k\right\rangle-\lambda_k\varepsilon_k\right)+(1+2\lambda_k\mu)\left\|x_k-x\right\|^2.\end{aligned} \tag{3.8}$$

令 $\rho_k=\frac{1}{1+2\lambda_k\mu}$, 由 (3.4) 式可知 $w_{k-1}-x_k=(1-\rho_k)(w_{k-1}-y_k)-\rho_k\lambda_k v_k$, 故由 (2.3) 式可得

$$\left\|w_{k-1}-x_k\right\|^2=\rho_k\|\lambda_k v_k\|^2+(1-\rho_k)\left\|y_k-w_{k-1}\right\|^2-\rho_k(1-\rho_k)\left\|\lambda_k v_k+y_k-w_{k-1}\right\|^2. \tag{3.9}$$

由 (3.4) 式可知

$$x_k-y_k=\rho_k\left(w_{k-1}-\lambda_k v_k-y_k\right), \tag{3.10}$$

从而由 (3.10) 式有

$$\begin{aligned}2\left\langle\lambda_k v_k,x_k-y_k\right\rangle&=2\rho_k\left\langle\lambda_k v_k,w_{k-1}-\lambda_k v_k-y_k\right\rangle\\&=\rho_k\left(2\left\langle\lambda_k v_k,w_{k-1}-y_k\right\rangle-2\|\lambda_k v_k\|^2\right)\\&=\rho_k\left(\|\lambda_k v_k\|^2+\|w_{k-1}-y_k\|^2-\|\lambda_k v_k+y_k-w_{k-1}\|^2-2\|\lambda_k v_k\|^2\right)\\&=\rho_k\left(\|w_{k-1}-y_k\|^2-\|\lambda_k v_k+y_k-w_{k-1}\|^2-\|\lambda_k v_k\|^2\right).\end{aligned} \tag{3.11}$$

事实上, $(1-\rho_k)=2\lambda_k\mu\rho_k$, 从而由 (3.10) 式可得

$$2\lambda_k\mu\|x_k-y_k\|^2=\rho_k(1-\rho_k)\left\|\lambda_k v_k+y_k-w_{k-1}\right\|^2, \tag{3.12}$$

由 (3.9), (3.11) 和 (3.12) 式可得

$$\begin{aligned}&\left\|w_{k-1}-x_k\right\|^2+2\left(\lambda_k\mu\|x_k-y_k\|^2+\left\langle\lambda_k v_k,\ x_k-y_k\right\rangle-\lambda_k\varepsilon_k\right)\\=&\left\|y_k-w_{k-1}\right\|^2-\left(\rho_k\left\|\lambda_k v_k+y_k-w_{k-1}\right\|^2+2\lambda_k\varepsilon_k\right),\end{aligned}$$

并联立 (3.3) 和 (3.8) 式可得

$$\begin{aligned}\|w_{k-1}-x\|^2 \geq & (1+2\lambda_k\mu)\|x_k-x\|^2+\left(1-\sigma^2\right)\|y_k-w_{k-1}\|^2 \\ & -2\lambda_k\left(\mu\|x-y_k\|^2+\langle v_k,\ x-y_k\rangle-\varepsilon_k\right).\end{aligned} \tag{3.13}$$

若 $\forall x^* \in (A+B)^{-1}(0)$, 则存在 $a^* \in A(x^*)$, $-a^* \in B^{[\varepsilon_k]}(x^*)$. 由 (3.3) 式可知: 存在 $a_k \in A(y_k)$, $b_k \in B^{[\varepsilon_k]}(y_k)$, 使得 $v_k = a_k + b_k$. 由 (2.1) 与 (2.2) 式可知

$$\langle a^*-a_k, x^*-y_k\rangle \geq \mu\|x^*-y_k\|^2, \quad \langle b_k+a^*, y_k-x^*\rangle \geq -\varepsilon_k\,.$$

将两式相加可得 $\langle v_k\,, y_k-x^*\rangle \geq \mu\|x^*-y_k\|^2-\varepsilon_k$, 即 $\mu\|x^*-y_k\|^2+\langle v_k,\ x^*-y_k\rangle-\varepsilon_k \leq 0$. 并将 $x=x^*$ 代入 (3.13) 式即可得到 (3.7) 式.

命题 3.3 令点列 $\{x_k\}$, $\{y_k\}$ 和 $\{w_k\}$ 由算法 1 生成, 令 $\rho_k := \frac{1}{1+2\lambda_k\mu}$, 且定义 s_k 为

$$s_k := \max\left\{\eta\|x_k-w_{k-1}\|^2\,,\ \left(1-\sigma^2\right)\|y_k-w_{k-1}\|^2\right\}, \tag{3.14}$$

其中 $\overline{\rho} = \frac{1}{1+2\underline{\lambda}\mu}$, $\eta := \frac{1-\sigma^2}{\left(1+\sigma\sqrt{\overline{\rho}}\right)^2} \in (0,1]$. 则 $\forall x^* \in (A+B)^{-1}(0)$, 有

$$\|x_k-x^*\|^2+\rho_k s_k \leq \rho_k\|w_{k-1}-x^*\|^2. \tag{3.15}$$

证 由 (3.4) 式及 ρ_k 的定义可知 $x_k = \rho_k\left(w_{k-1}-\lambda_k v_k\right)+\left(1-\rho_k\right)y_k$, 故

$$x_k-w_{k-1} = \left(y_k-w_{k-1}\right)-\rho_k\left(\lambda_k v_k+y_k-w_{k-1}\right),$$

从而由三角不等式可知

$$\|x_k-w_{k-1}\| \leq \|y_k-w_{k-1}\|+\rho_k\|\lambda_k v_k+y_k-w_{k-1}\|. \tag{3.16}$$

由 (3.3) 式可知

$$\|\lambda_k v_k+y_k-w_{k-1}\| \leq \frac{\sigma}{\sqrt{\rho_k}}\|y_k-w_{k-1}\|, \tag{3.17}$$

其中 $\rho_k \in [\underline{\rho}, \overline{\rho}]$, $\underline{\rho} = \frac{1}{1+2\overline{\lambda}\mu}, \overline{\rho} = \frac{1}{1+2\underline{\lambda}\mu}$. 从而由 (3.16) 和 (3.17) 式可知

$$\|x_k-w_{k-1}\| \leq (1+\sigma\sqrt{\rho_k})\|y_k-w_{k-1}\| \leq \left(1+\sigma\sqrt{\overline{\rho}}\right)\|y_k-w_{k-1}\|,$$

从而

$$\left(1-\sigma^2\right)\|y_k-w_{k-1}\|^2 \geq \frac{1-\sigma^2}{\left(1+\sigma\sqrt{\overline{\rho}}\right)^2}\|x_k-w_{k-1}\|^2.$$

并由 (3.7) 和 (3.14) 式可得到 (3.15) 式.

命题 3.4 令点列 $\{x_k\}$, $\{y_k\}$, $\{w_k\}$ 由算法 1 生成, s_k 由 (3.15) 式所定义, 并且令

$$\varphi_k := \|x_k-x^*\|^2,\ \forall k \geq -1 \quad \delta_k := \alpha_{k-1}\left(1+\alpha_{k-1}\right)\rho_k\|x_{k-1}-x_{k-2}\|^2,\ \forall k \geq 1, \tag{3.18}$$

且 $\rho_k := \frac{1}{1+2\lambda_k\mu}$, $\underline{\rho} = \frac{1}{1+2\overline{\lambda}\mu}, \overline{\rho} = \frac{1}{1+2\underline{\lambda}\mu}$, $x^* \in (A+B)^{-1}(0)$. 则有 $\varphi_0 = \varphi_{-1}$, 而且

$$\varphi_k - \varphi_{k-1} + \rho_k s_k \leq \rho_k \alpha_{k-1}\left(\varphi_{k-1} - \varphi_{k-2}\right) + \delta_k, \quad \forall k \geq 1. \tag{3.19}$$

证 将 $x = x^*$ 代入 (3.6) 式并由 (3.18) 式可得

$$\rho_k \|w_{k-1} - x^*\|^2 = \left(1 + \alpha_{k-1}\right)\rho_k \varphi_{k-1} - \alpha_{k-1}\rho_k\varphi_{k-2} + \delta_k,$$

并由 (3.15) 式可知

$$\left(\varphi_k - \rho_k\varphi_{k-1}\right) + \rho_k s_k \leq \alpha_{k-1}\rho_k\left(\varphi_{k-1} - \varphi_{k-2}\right) + \delta_k,$$

注意到, $0 < \rho_k \leq 1$, 有 $\varphi_k - \varphi_{k-1} \leq \varphi_k - \rho_k\varphi_{k-1}$, 故 (3.19) 式成立.

引理 3.5 [4] 令点列 $\{\varphi_k\}$, $\{\rho_k s_k\}$, $\{\rho_k\alpha_k\} \in [0, +\infty[$ 使得 $\varphi_0 = \varphi_{-1}$, $0 \leq \rho_k \leq \overline{\rho}$, $0 \leq \alpha_k \leq \alpha$, $\alpha\overline{\rho} < 1$, 其中 $\overline{\rho} = \frac{1}{1+2\underline{\lambda}\mu}$, s_k 如 (3.15) 式所定义, 且满足 (3.19) 式. 则 $\forall k \geq 1$, 有

$$\varphi_k + \sum_{j=1}^{k} \rho_j s_j \leq \varphi_0 + \frac{1}{1-\alpha\overline{\rho}} \sum_{j=1}^{k} \delta_j; \tag{3.20}$$

进一步, 如果 $\sum\limits_{j=1}^{\infty} \delta_k < +\infty$, 则 $\lim\limits_{k\to\infty} \varphi_k$ 存在, 即数列 $\{\varphi_k\}$ 收敛到某一非负实数.

命题 3.6 令点列 $\{x_k\}$, $\{\lambda_k\}$ 和 $\{\alpha_k\}$ 由算法 1 产生, 如果对于任意的 $k \geq 1$, $\lambda_k \geq \underline{\lambda} > 0$, 满足下面的条件

$$\sum_{k=0}^{\infty} \alpha_k \|x_k - x_{k-1}\|^2 < +\infty, \tag{3.21}$$

则有下面的结论成立

(i) 点列 $\{x_k - w_{k-1}\}$, $\{y_k - w_{k-1}\}$, $\{x_k - y_k\}$, $\{v_k\}$ 和 $\{\varepsilon_k\}$ 强收敛到零;

(ii) 点列 $\{x_k\}$, $\{w_{k-1}\}$ 和 $\{y_k\}$ 弱收敛到单调包含问题 (3.1) 的解.

证 由于 $\lambda_k \geq \underline{\lambda} > 0$, 则 $\rho_k \leq \overline{\rho} = \frac{1}{1+2\underline{\lambda}\mu}$, 并由命题 3.4 和引理 3.5 可知 $\forall x^* \in (A+B)^{-1}(0)$, $\lim\limits_{k\to\infty} \|x_k - x^*\|$ 存在, 故点列 $\{x_k\}$ 有界, 而且 $\sum\limits_{j=1}^{\infty} s_k < +\infty$, 故 $\lim\limits_{k\to\infty} s_k = 0$. 又由 (3.2), (3.3) 和 (3.14) 式可知

$$\lim_{k\to\infty} \|x_k - w_{k-1}\| = \lim_{k\to\infty} \|y_k - w_{k-1}\| = \lim_{k\to\infty} \|v_k\| = \lim_{k\to\infty} \varepsilon_k = 0, \tag{3.22}$$

即点列 $\{x_k - w_{k-1}\}$, $\{y_k - w_{k-1}\}$, $\{v_k\}$ 和 $\{\varepsilon_k\}$ 强收敛到零, 从而 $\{x_k - y_k\}$ 也强收敛到零.

令 $x^\infty \in \mathbb{X}$ 是点列 $\{x_k\}$ 的弱聚点, 并令子列 $\left\{x_{k_j}\right\}$ 使得 $x_{k_j} \rightharpoonup x^\infty$. 又由 (3.2) 和 (3.22) 式可得

$$v_{k_j} \in A\left(y_{k_j}\right) + B^{[\varepsilon_{k_j}]}\left(y_{k_j}\right), \ \lim_{j\to\infty} \left\|v_{k_j}\right\| = 0, \ \lim_{j\to\infty} \varepsilon_{k_j} = 0, \ \forall j \geq 1$$

故由命题 2.1(v) 可知, $x^\infty \in (A+B)^{-1}(0)$, 故根据定理 2.2 可知点列 $\{x_k\}$ 弱收敛到单调包含问题 (3.5) 的解. 由于 $\{x_k - w_{k-1}\}$ 和 $\{x_k - y_k\}$ 强收敛到零, 因此 $\{w_{k-1}\}$ 和 $\{y_k\}$ 也弱收敛到单调包含问题 (3.1) 的解.

注 3.2 由于惯性邻近点算法是算法 1 的特殊情况, 因此命题 3.6 是文献 [4] 中定理 2.1 的推广.

命题 3.7 令点列 $\{x_k\}$, $\{y_k\}$, $\{w_k\}$ 由算法 1 生成, 令 $\rho_k := \frac{1}{1+2\lambda_k\mu}$, 且

$$\beta := \begin{cases} \frac{(1+2\eta)\overline{\rho}-\sqrt{(1+2\eta)^2\overline{\rho}^2-4\eta(\eta-1)\overline{\rho}\underline{\rho}}}{2(\eta-1)\overline{\rho}}, & \sigma \in (0\ ,\ 1) \\ \frac{\underline{\rho}}{3\overline{\rho}}, & \sigma=0, \end{cases} \tag{3.23}$$

其中 $\underline{\rho} = \frac{1}{1+2\overline{\lambda}\mu}, \overline{\rho} = \frac{1}{1+2\underline{\lambda}\mu}$. 假设 $\alpha_k \le \alpha_{k+1} \le \alpha < \beta$, 定义实函数 $q(\alpha')$ 为

$$q(\alpha') := (\eta-1)\overline{\rho}\alpha'^2 - (1+2\eta)\overline{\rho}\alpha' + \underline{\rho}\eta, \tag{3.24}$$

则有 $q(\alpha) > 0$, 且 $\forall x^* \in (A+B)^{-1}(0)$, 有

$$\sum_{j=1}^{k} \|x_j - x_{j-1}\|^2 \le \frac{2\|x_0 - x^*\|^2}{(1-\alpha\overline{\rho})\,q(\alpha)}, \tag{3.25}$$

因此点列 $\{x_k\}$ 弱收敛到单调包含问题 (3.1) 的解.

证 由 (3.2) 式可知 $x_k - w_{k-1} = (x_k - x_{k-1}) - \alpha_{k-1}(x_{k-1} - x_{k-2})$, 故

$$\begin{aligned} \|x_k - w_{k-1}\|^2 =& \|x_k - x_{k-1}\|^2 + \alpha_{k-1}^2\|x_{k-1} - x_{k-2}\|^2 - 2\alpha_{k-1}\langle x_k - x_{k-1},\ x_{k-1} - x_{k-2}\rangle \\ \ge& \|x_k - x_{k-1}\|^2 + \alpha_{k-1}^2\|x_{k-1} - x_{k-2}\|^2 - \alpha_{k-1}\left(2\|x_k - x_{k-1}\|\,\|x_{k-1} - x_{k-2}\|\right) \\ \ge& (1-\alpha_{k-1})\|x_k - x_{k-1}\|^2 - \alpha_{k-1}(1-\alpha_{k-1})\|x_{k-1} - x_{k-2}\|^2. \end{aligned} \tag{3.26}$$

并联立 (3.19) 式可得

$$\begin{aligned} &\varphi_k - \varphi_{k-1} - \rho_k\alpha_{k-1}(\varphi_{k-1} - \varphi_{k-2}) \\ \le& \rho_k\left[\alpha_{k-1}(1+\alpha_{k-1})\|x_{k-1} - x_{k-2}\|^2 - \eta\|x_k - w_{k-1}\|^2\right] \\ \le& \rho_k\left[(1-\eta)\alpha_{k-1}^2 + (1+\eta)\alpha_{k-1}\right]\|x_{k-1} - x_{k-2}\|^2 - \rho_k\eta(1-\alpha_{k-1})\|x_k - x_{k-1}\|^2. \end{aligned} \tag{3.27}$$

令 $\gamma_k := (1-\eta)\rho_{k+1}\alpha_k^2 + (1+\eta)\rho_{k+1}\alpha_k$, $\forall k \ge 0$. 定义

$$\xi_0 := (1-\rho_1\alpha_0)\varphi_0,\ \ \xi_k := \varphi_k - \rho_k\alpha_{k-1}\varphi_{k-1} + \gamma_k\|x_k - x_{k-1}\|^2,\ \forall k \ge 1,$$

又由于 $\{\rho_k\}$ 和 $\{\alpha_k\}$ 是单增非负数列, 且 $\rho_k \in [\underline{\rho}, \overline{\rho}]$, 因此

$$\begin{aligned} \xi_k - \xi_{k-1} \le& \left[\varphi_k - \varphi_{k-1} - \rho_k\alpha_{k-1}(\varphi_{k-1} - \varphi_{k-2}) - \gamma_{k-1}\|x_{k-1} - x_{k-2}\|^2\right] + \gamma_k\|x_k - x_{k-1}\|^2 \\ \le& \left[\gamma_k - \rho_k\eta(1-\alpha_k)\right]\|x_k - x_{k-1}\|^2 \\ \le& -\left[(\eta-1)\overline{\rho}\alpha_k^2 - (1+2\eta)\overline{\rho}\alpha_k + \underline{\rho}\eta\right]\|x_k - x_{k-1}\|^2 \\ =& -q(\alpha_k)\|x_k - x_{k-1}\|^2. \end{aligned} \tag{3.28}$$

当 $\eta \in (0,1)$ 时, 即 $\sigma \in (0,1)$, 令 $q(\alpha') := (\eta-1)\overline{\rho}\alpha'^2 - (1+2\eta)\overline{\rho}\alpha' + \underline{\rho}\eta = 0$, 解得

$$\alpha_1' = \frac{(1+2\eta)\overline{\rho} + \sqrt{(1+2\eta)^2\overline{\rho}^2 - 4\eta(\eta-1)\overline{\rho}\underline{\rho}}}{2(\eta-1)\overline{\rho}} < 0,$$

$$\alpha_2' = \frac{(1+2\eta)\overline{\rho} - \sqrt{(1+2\eta)^2\overline{\rho}^2 - 4\eta(\eta-1)\overline{\rho}\underline{\rho}}}{2(\eta-1)\overline{\rho}} \in (0,1).$$

另一方面, 当 $\eta = 1$ 时, 即 $\sigma = 0$, 令 $q(\alpha') := -3\overline{\rho}\alpha' + \underline{\rho} = 0$, 解得 $\alpha' = \frac{\underline{\rho}}{3\overline{\rho}}$.

在上述两种情况下, 当 $\alpha_k \le \alpha_{k+1} \le \alpha < \beta$ 时都有 $q(\alpha_k) \ge q(\alpha) > q(\beta) = 0$, 从而由 (3.28) 式可知

$$\|x_k - x_{k-1}\|^2 \le \frac{1}{q(\alpha)}\left(\xi_{k-1} - \xi_k\right), \quad \forall k \ge 1, \tag{3.29}$$

由 ξ_k 的定义可得 $\xi_k \ge -\alpha\overline{\rho}\xi_k$, 从而由 (3.29) 式有

$$\sum_{j=1}^{k} \|x_j - x_{j-1}\|^2 \le \frac{1}{q(\alpha)}\left(\xi_0 - \xi_k\right) \le \frac{1}{q(\alpha)}\left(\xi_0 + \alpha\overline{\rho}\varphi_{k-1}\right), \quad \forall k \ge 1. \tag{3.30}$$

又由于 $\xi_0 \ge \ldots \ge \xi_k = \varphi_k - \rho_k\alpha_{k-1}\varphi_{k-1} + \gamma_k\|x_k - x_{k-1}\|^2 \ge \varphi_k - \alpha\overline{\rho}\varphi_{k-1}$, 所以

$$\varphi_k \le (\alpha\overline{\rho})^k\varphi_0 + \frac{\xi_0}{1-\alpha\overline{\rho}} \le \varphi_0 + \frac{\xi_0}{1-\alpha\overline{\rho}}, \quad \forall k \ge 0. \tag{3.31}$$

联立 (3.30) 和 (3.31) 式可得 (3.25) 式.

又由于 $\alpha_k \in [0,1)$, 从而由 (3.25) 式可知 (3.21) 式成立, 由命题 3.6 可知点列 $\{x_k\}$ 弱收敛到单调包含问题 (3.1) 的解.

注 3.3 (i) 当 $\sigma = 0$ 时, 算法 1 退化为惯性邻近点算法, 条件 (3.23) 式退化为 $\alpha_k \le \alpha_{k+1} \le \alpha < \frac{\underline{\rho}}{3\overline{\rho}}$; 进一步, 当 $\mu = 0$ 时, 有 $\underline{\rho} = \overline{\rho} = 1$, 则条件 (3.23) 式退化为 $\alpha_k \le \alpha_{k+1} \le \alpha < \frac{1}{3}$, 因此命题 3.7 对文献 [4] 中的命题 2.1 作了推广.

(ii) 如果 $\sigma = 1$, 则对于任意的 $\alpha' \in [0,1)$, 二次函数 $q(\alpha') = -\overline{\rho}\alpha'^2 - \overline{\rho}\alpha' \le 0$, 并不能满足算法收敛的条件, 从而说明了容差参数 σ 的最大取值范围为 $[0,1)$.

推论 3.8 令点列 $\{x_k\}$, $\{y_k\}$, $\{w_k\}$ 由算法 1 生成. 令 $\rho_k := \frac{1}{1+2\lambda_k\mu}$, $\underline{\rho} = \frac{1}{1+2\overline{\lambda}\mu}$, $\overline{\rho} = \frac{1}{1+2\underline{\lambda}\mu}$. β 如 (3.23) 式所定义, $q(\alpha')$ 如 (3.24) 式所定义, 且 $x^* \in (A+B)^{-1}(0)$, 假设 $\alpha_k \le \alpha_{k+1} \le \alpha < \beta$. 则 $\forall k \ge 1$, 有

$$\|x_k - x^*\|^2 + \sum_{j=1}^{k} \frac{1-\sigma^2}{1+2\lambda_j\mu}\|y_j - w_{j-1}\|^2 \le \left(1 + \frac{2\alpha\overline{\rho}(1+\alpha)}{(1-\alpha\overline{\rho})^2 q(\alpha)}\right)\|x_0 - x^*\|^2, \tag{3.32}$$

$$\|x_k - x^*\|^2 + \sum_{j=1}^{k} \frac{\rho_j^2(1-\sigma^2)}{\left(\sqrt{\rho_j}+\sigma\right)^2}\|\lambda_j v_j\|^2 \le \left(1 + \frac{2\alpha\overline{\rho}(1+\alpha)}{(1-\alpha\overline{\rho})^2 q(\alpha)}\right)\|x_0 - x^*\|^2. \tag{3.33}$$

证 由 (3.20) 和 (3.25) 式可得

$$\varphi_k + \sum_{j=1}^{k}\rho_j s_j \le \varphi_0 + \frac{\alpha\overline{\rho}(1+\alpha)}{1-\alpha\overline{\rho}}\sum_{j=1}^{k}\|x_j - x_{j-1}\|^2 \le \left(1 + \frac{2\alpha\overline{\rho}(1+\alpha)}{(1-\alpha\overline{\rho})^2 q(\alpha)}\right)\|x_0 - x^*\|^2,$$

从而由 s_k 的定义可知, (3.32) 式成立.

由柯西不等式及 (3.3) 式可知

$$\|\lambda_j v_j\| \le \|\lambda_j v_j + y_j - w_{j-1}\| + \|y_j - w_{j-1}\| \le \left(1 + \frac{\sigma}{\sqrt{\rho_j}}\right)\|y_j - w_{j-1}\|,$$

由 (3.32) 式可知 (3.33) 式成立.

接下来, 本文给出算法 1 的非渐近性全局收敛率.

定理 3.9 令点列 $\{x_k\}$, $\{y_k\}$, $\{w_k\}$ 由算法 1 生成, 令 $\rho_k := \frac{1}{1+2\lambda_k\mu}$, $\underline{\rho} = \frac{1}{1+2\overline{\lambda}\mu}$, $\overline{\rho} = \frac{1}{1+2\underline{\lambda}\mu}$, β 如 (3.23) 式所定义, $q(\alpha')$ 如 (3.24) 式所定义, $x^* \in (A+B)^{-1}(0)$, 且 $d_0 = \|x_0 - x^*\|$, 假设 $\alpha_k \le \alpha_{k+1} \le \alpha < \beta$. 则 $\forall k \ge 1$, 存在 $i \in \{1, \dots, k\}$ 使得 $v_i \in A(y_i) + B^{[\varepsilon_i]}(y_i)$, 而且

$$\|v_i\| \le \frac{(\sigma + \sqrt{\overline{\rho}})d_0}{\underline{\rho}\underline{\lambda}\sqrt{k(1-\sigma^2)}}\sqrt{1 + \frac{2\alpha\overline{\rho}\,(1+\alpha)}{(1-\alpha\overline{\rho})^2 q\,(\alpha)}}, \tag{3.34}$$

$$\varepsilon_i \le \frac{\sigma^2 d_0^2}{2(1-\sigma^2)\underline{\lambda}k}\left(1 + \frac{2\alpha\overline{\rho}\,(1+\alpha)}{(1-\alpha\overline{\rho})^2 q\,(\alpha)}\right). \tag{3.35}$$

证 令 $x^* \in (A+B)^{-1}(0)$ 使得 $d_0 = \|x_0 - x^*\|$. 由 (3.32) 和 (3.33) 式可知, 对 $\forall k \ge 1$ 存在 $i \in \{1, \dots, k\}$ 使得

$$\frac{\rho_i^2\,(1-\sigma^2)}{\left(\sqrt{\rho_i}+\sigma\right)^2}\|\lambda_i v_i\|^2 \le \left(1 + \frac{2\alpha\bar{\rho}\,(1+\alpha)}{(1-\alpha\bar{\rho})^2 q\,(\alpha)}\right)\frac{d_0^2}{k},$$

$$\rho_i(1-\sigma^2)\|y_i - w_{i-1}\|^2 \le \left(1 + \frac{2\alpha\bar{\rho}\,(1+\alpha)}{(1-\alpha\bar{\rho})^2 q\,(\alpha)}\right)\frac{d_0^2}{k}.$$

又由于 $\lambda_i \ge \underline{\lambda} > 0$, 因此

$$\|v_i\| \le \frac{(\sigma + \sqrt{\overline{\rho}})d_0}{\underline{\rho}\underline{\lambda}\sqrt{k(1-\sigma^2)}}\sqrt{1 + \frac{2\alpha\overline{\rho}\,(1+\alpha)}{(1-\alpha\overline{\rho})^2 q\,(\alpha)}},$$

$$\|y_i - w_{i-1}\|^2 \le \frac{d_0^2}{\underline{\rho}(1-\sigma^2)k}\left(1 + \frac{2\alpha\bar{\rho}\,(1+\alpha)}{(1-\alpha\bar{\rho})^2 q\,(\alpha)}\right).$$

由 (3.3) 式可知 $\varepsilon_i \le \frac{\sigma^2}{2\lambda_i}\|y_i - w_{i-1}\|^2$, 故

$$\varepsilon_i \le \frac{\sigma^2 d_0^2}{2\underline{\rho}(1-\sigma^2)\underline{\lambda}k}\left(1 + \frac{2\alpha\bar{\rho}\,(1+\alpha)}{(1-\alpha\bar{\rho})^2 q\,(\alpha)}\right).$$

注 3.4 定理 3.9 说明了算法 1 的全局逐点收敛率为 $\mathcal{O}(\frac{1}{\sqrt{k}})$.

4 求解单调包含问题的惯性 Tseng's 向前向后算法

在这部分, 本文考虑单调包含问题

$$0 \in F(x) + T(x), \tag{4.1}$$

其中 $F: \mathbb{X} \to \mathbb{X}$ 是 (单值) 强单调和 L-Lipschitc 连续算子, 即存在 $\mu \ge 0$ 和 $L > 0$ 使得

$$\langle F(x) - F(x'), x - x' \rangle \ge \mu\|x - x'\|^2, \tag{4.2}$$

$$\|F(x) - F(x')\| \le L\,\|x - x'\|, \quad \forall x, x' \in \mathbb{X}. \tag{4.3}$$

$T:\mathbb{X}\rightrightarrows\mathbb{X}$ 是极大单调算子. 假设问题有解, 即 $(F+T)^{-1}(0)$ 非空.

下面本文给出求解问题 (4.1) 的惯性 Tseng's 向前向后算法.

算法 2

步骤 0 任取初始点 $x_0=x_{-1}\in\mathbb{X}$, 任意给定 $\alpha\in[0,1)$, $\alpha_0\in[0,\alpha]$, $\sigma<1$, $k=1$, 令 $\overline{\lambda}:=\frac{\sigma}{L^2}\left(\sigma\mu+\sqrt{\sigma^2\mu^2+L^2}\right)$, 任意给定 $\lambda_0\in[0\ ,\overline{\lambda}]$;

步骤 1 选定 $\alpha_k\in[\alpha_{k-1},\ \alpha]$, 定义

$$w_{k-1}:=x_{k-1}+\alpha_{k-1}(x_{k-1}-x_{k-2});\tag{4.4}$$

步骤 2 选定 $0<\lambda_k\le\lambda_{k-1}$, 计算 y_k

$$y_k=(\lambda_kT+Id)^{-1}\left(w_{k-1}-\lambda_kF\left(w_{k-1}\right)\right);\tag{4.5}$$

步骤 3 计算 x_k

$$x_k=y_k-\frac{\lambda_k}{1+2\lambda_k\mu}\left(F\left(y_k\right)-F\left(w_{k-1}\right)\right).\tag{4.6}$$

令 $k:=k+1$, 返回步骤 1.

注 4.1 (i) 当 $\alpha_k\equiv0$, $\lambda_k\equiv\overline{\lambda}$ 时, 算法 2 退化为文献 [12] 中的算法 2.

(ii) 算法 2 是算法 1 在选定 $\varepsilon_k\equiv0$ 时的特殊情况. 事实上, 在算法 2 中定义: $A:=F$, $B:=T$, 且

$$v_k:=F\left(y_k\right)-F\left(w_{k-1}\right)+\frac{1}{\lambda_k}(w_{k-1}-y_k),\quad\forall k\ge1.\tag{4.7}$$

由 $\overline{\lambda}$ 的定义可知 $\sigma^2=\frac{\overline{\lambda}^2L^2}{1+2\overline{\lambda}\mu}$. 由 (4.5) 式可得 $q_k:=\frac{1}{\lambda_k}(w_{k-1}-y_k)-F\left(w_{k-1}\right)\in T(y_k)$, 又由 (4.7) 式可得 $v_k=F(y_k)+q_k\in F(y_k)+T(y_k)=A\left(y_k\right)+B^{[\varepsilon_k]}\left(y_k\right)$. 由 (4.5) 式可知

$$\lambda_kv_k+y_k-w_{k-1}=\lambda_k(F\left(y_k\right)-F\left(w_{k-1}\right)).\tag{4.8}$$

从而可知

$$\frac{\left\|\lambda_kv_k+y_k-w_{k-1}\right\|^2}{1{+}2\lambda_k\mu}=\frac{\left\|\lambda_k(F\left(y_k\right)-F\left(w_{k-1}\right))\right\|^2}{1{+}2\lambda_k\mu}\le\frac{\lambda_k^2L^2}{1{+}2\lambda_k\mu}\left\|y_k-w_{k-1}\right\|^2\tag{4.9}$$

注意到 $f(\lambda)=\frac{\lambda^2L^2}{1+2\lambda\mu}$ 是单增函数, 因此 $f(\lambda_k)\le f(\overline{\lambda})$, 故 $\frac{\|\lambda_kv_k+y_k-w_{k-1}\|^2}{1+2\lambda_k\mu}\le\sigma^2\left\|y_k-w_{k-1}\right\|^2$, 即算法 1 的步骤 2 满足.

将 (4.8) 式代入 (4.6) 式可得

$$x_k=y_k-\frac{\lambda_kv_k+y_k-w_{k-1}}{1+2\lambda_k\mu}=\frac{(w_{k-1}-\lambda_kv_k)+2\lambda_k\mu y_k}{1+2\lambda_k\mu},$$

即算法 1 的步骤 3 满足. 综上可知, 算法 2 是算法 1 的特殊情形.

由于算法 2 是算法 1 的特殊情况, 因此由命题 3.6, 3.7 和定理 3.9 可得到下面的结论.

命题 4.1 令点列 $\{x_k\}$ 由算法 2 产生, 令 $\rho_k:=\frac{1}{1+2\lambda_k\mu}$, $\underline{\rho}=\frac{1}{1+2\overline{\lambda}\mu}$, $\overline{\rho}=\frac{1}{1+2\underline{\lambda}\mu}$, $\{\varepsilon_k\}$ 和 $\{v_k\}$ 如 (4.7) 式所定义, $d_0=\|x_0-x^*\|$, 其中 $x^*\in(F+T)^{-1}(0)$, β 如 (3.23) 式所定义, $q(\alpha')$ 如 (3.24) 式所定义, 假设 $\alpha_k\le\alpha_{k+1}\le\alpha<\beta$. 则有下面的结论成立

(i) 点列 $\{x_k\}$ 弱收敛到单调包含问题 (4.1) 的解.

(ii) 对任意的 $k\geq 1$ 存在 $i\in\{1,\ldots,k\}$ 使得 $v_i\in F(y_i)+B^{[\varepsilon_i]}(y_i)$, 而且

$$\|v_i\|\leq \frac{(\sigma+\sqrt{\overline{\rho}})d_0}{\underline{\rho}\underline{\lambda}\sqrt{k(1-\sigma^2)}}\sqrt{1+\frac{2\alpha\overline{\rho}\,(1+\alpha)}{(1-\alpha\overline{\rho})^2q\,(\alpha)}}.$$

注 4.2 (i) 命题 4.1 说明了算法 2 的全局逐点收敛率为 $\mathcal{O}(\frac{1}{\sqrt{k}})$.

(ii) 当 $\lambda_k\equiv\overline{\lambda}$ 时, $\frac{d_0}{\underline{\lambda}}=\frac{d_0L^2}{\sigma\left(\sigma\mu+\sqrt{\sigma^2\mu^2+L^2}\right)}$. 在这种情况下, 给定容差 $\epsilon>0$, 在算法 2 中找到有序对 (x,v) 使得 $v\in(F+T)(x)$, $\|v\|\leq\epsilon$ 至多需要 $\mathcal{O}(\lceil\frac{d_0^2L^2}{\epsilon^2}\rceil)$ 次迭代.

5 求解原始一对偶问题的惯性非精确 Spingarn's 部分逆算法

接下来, 本文考虑原始一对偶问题 [17]. 找到 x, $u\in\mathbb{X}$ 使得

$$z\in V,\quad u\in V^{\perp},\quad u\in T(z),\tag{5.1}$$

其中 $T:\mathbb{X}\rightrightarrows\mathbb{X}$ 是实 Hilbert 空间 $\mathbb{X}$ 中的极大单调算子, V 是闭子空间, $V^{\perp}$ 是 V 的正交补空间. 特别的, 当 $V=\mathbb{X}$ 时, 问题 (5.1) 退化为单调包含问题 (1.1). 假设 T 是 κ - 强单调和 L-Lipschitc 连续算子.

定义 5.1 [18] 称 $T_V:\mathbb{X}\rightrightarrows\mathbb{X}$ 是极大单调算子 $T:\mathbb{X}\rightrightarrows\mathbb{X}$ 关于 $\mathbb{X}$ 中的闭子空间 V 的 Spingarn's 部分逆算子, 如果 T_V 满足

$$G_r(T_V):=\{(x,v)\in\mathbb{X}\times\mathbb{X}|P_V(v)+P_{V^{\perp}}(x)\in T(P_V(x)+P_{V^{\perp}}(v))\},\tag{5.2}$$

其中 P_V 和 $P_{V^{\perp}}$ 分别是 V 和 $V^{\perp}$ 上的正交投影.

由定义可知, 原始一对偶问题 (5.1) 等价于单调包含问题 $0\in T_V(x)$, 而且若 x^* 是单调包含问题 $0\in T_V(x^*)$ 的解, 则 $z^*=P_V(x^*)$, $u^*=P_{V^{\perp}}(x^*)$ 是原始一对偶问题 (5.1) 的解.

引理 5.2 [19] 令 $T:\mathbb{X}\rightrightarrows\mathbb{X}$ 是 $\mathbb{X}$ 中的极大单调算子, V 是 $\mathbb{X}$ 中的闭子空间, 且 $\varepsilon>0$, 则 $(T_V)^{[\varepsilon]}=(T^{[\varepsilon]})_V$.

引理 5.3 [20] 如果极大单调算子 T 是 κ - 强单调和 L-Lipschitc 连续算子, 则它的关于 V 的部分逆 T_V 是 μ - 强单调的, 其中 $\mu=\frac{\kappa}{1+L^2}$.

下面给出求解原始一对偶问题 (5.1) 的惯性非精确 Spingarn's 部分逆算法如下

算法 3

步骤 0 任取初始点 $x_0=x_{-1}\in\mathbb{X}$, 给定 $\alpha\in[0,1)$, $\alpha_0\in[0,\alpha]$, $\sigma\in[0,1)$, $k=1$;

步骤 1 选定 $\alpha_k\in[\alpha_{k-1},\ \alpha]$, 计算

$$w_{k-1}:=x_{k-1}+\alpha_{k-1}(x_{k-1}-x_{k-2});\tag{5.3}$$

步骤 2 找到 z_k, $u_k\in\mathbb{X}$, $\varepsilon_k\geq 0$, 使得

$$u_k\in T^{[\varepsilon_k]}(z_k)\,,\ \frac{\|u_k+z_k-w_{k-1}\|^2}{1+2\mu}+2\varepsilon_k\leq\sigma^2\|P_V(z_k)+P_{V^{\perp}}(u_k)-w_{k-1}\|^2;\tag{5.4}$$

步骤 3 计算 x_k

$$x_k=\frac{w_{k-1}-u_k-z_k}{1+2\mu}+P_V(z_k)+P_{V^{\perp}}(u_k),\tag{5.5}$$

令 $k := k+1$, 返回步骤 1.

注 5.1 (i) 当 $\alpha_k \equiv 0$ 且 $\mu = 0$ 时, 算法 3 退化为文献 [21] 中的算法 2.

(ii) 算法 3 是算法 1 在选定 $\lambda_k \equiv 1$ 时的特殊情况. 事实上, 在算法 3 中定义 $A := T_V$, $B := 0$, 且

$$v_k := P_V(u_k) + P_{V^\perp}(z_k), \quad y_k := P_V(z_k) + P_{V^\perp}(u_k), \quad \forall k \geq 1. \tag{5.6}$$

由引理 5.2, (5.2) 及 (5.4) 式可知

$$P_V(u_k) + P_{V^\perp}(z_k) \in (T_V)^{[\varepsilon_k]}(P_V(z_k) + P_{V^\perp}(u_k)),$$

从而由 (5.6) 式可知 $v_k \in (T_V)^{[\varepsilon_k]}(y_k) = A^{[\varepsilon_k]}(y_k) + B(y_k)$, 且

$$v_k + y_k = P_V(u_k) + P_{V^\perp}(z_k) + P_V(z_k) + P_{V^\perp}(u_k) = u_k + z_k. \tag{5.7}$$

由 $\lambda_k \equiv 1$, (5.6), (5.7) 及 (5.4) 式可知

$$\begin{aligned}\frac{\|\lambda_k v_k + y_k - w_{k-1}\|^2}{1+2\lambda_k\mu} + 2\lambda_k\varepsilon_k &= \frac{\|u_k + z_k - w_{k-1}\|^2}{1+2\mu} + 2\varepsilon_k \\ &\leq \sigma^2\|P_V(z_k) + P_{V^\perp}(u_k) - w_{k-1}\|^2 = \sigma^2\|y_k - w_{k-1}\|^2,\end{aligned}$$

即算法 1 的步骤 2 满足. 由 $\lambda_k \equiv 1$, (5.6), (5.7) 及 (5.5) 式可知

$$\begin{aligned}x_k &= \frac{w_{k-1} - u_k - z_k}{1+2\mu} + P_V(z_k) + P_{V^\perp}(u_k) = \frac{w_{k-1} - v_k - y_k}{1+2\mu} + y_k \\ &= \frac{w_{k-1} - v_k + 2\mu y_k}{1+2\mu} = \frac{(w_{k-1} - \lambda_k v_k) + 2\lambda_k\mu y_k}{1+2\lambda_k\mu},\end{aligned}$$

即算法 1 的步骤 3 满足. 故算法 3 是算法 1 的特殊情形.

由于算法 3 是算法 1 的特殊情况, 因此由命题 3.6, 3.7 和定理 3.9 可得到下面的结论.

命题 5.1 令点列 $\{x_k\}$ 由算法 3 产生, 且 $\{\lambda_k\}$, $\{v_k\}$ 和 $\{y_k\}$ 如 (5.6) 式所定义, 令 $\rho := \frac{1+L^2}{1+L^2+2\kappa}$, $\eta := \frac{1-\sigma^2}{\left(1+\sigma\sqrt{\rho}\right)^2} \in (0,1]$, $d_0 = \|x_0 - x^*\|$, 其中 x^* 是单调包含问题 $0 \in T_V(x^*)$ 的解. 令

$$\beta := \begin{cases} \frac{1+2\eta-\sqrt{8\eta+1}}{2(\eta-1)}, & \sigma \in (0\ ,\ 1), \\ \frac{1}{3}, & \sigma = 0, \end{cases} \tag{5.8}$$

定义实函数 $q(\alpha')$ 为

$$q(\alpha') := (\eta - 1)\rho\alpha'^2 - (1+2\eta)\rho\alpha' + \rho\eta. \tag{5.9}$$

假设 $\alpha_k \leq \alpha_{k+1} \leq \alpha < \beta$, 则有下面的结论成立

(i) 点列 $\{x_k\}$ 弱收敛到单调包含问题 $0 \in T_V(x)$ 的解.

(ii) 对任意的 $k \geq 1$ 存在 $i \in \{1, \ldots, k\}$ 使得 $v_i \in (T_V)^{[\varepsilon_i]}(y_i)$, 而且

$$\|v_i\| \leq \frac{(\sigma + \sqrt{\rho})d_0}{\rho\sqrt{k(1-\sigma^2)}}\sqrt{1 + \frac{2\alpha\rho\,(1+\alpha)}{(1-\alpha\rho)^2 q\,(\alpha)}},$$

$$\varepsilon_i \leq \frac{\sigma^2 d_0^2}{2(1-\sigma^2)k}\left(1+\frac{2\alpha\rho\,(1+\alpha)}{(1-\alpha\rho)^2 q\,(\alpha)}\right).$$

注 5.2 有序对 $(P_V(x_k), P_{V^\perp}(x_k))$ 弱收敛到原始一对偶问题 (5.1) 的解.

6 总结

本文提出了一种新的求解单调包含问题的惯性混合邻近外梯度算法, 并得到了该算法在实希尔伯特空间中的弱收敛性和非渐近全局收敛率. 惯性混合邻近外梯度算法包含一些现有的算法, 例如 Tseng's 向前向后算法和 Spingarn's 部分逆算法等. 利用该算法的框架可以设计求解优化问题与变分不等式问题的新的算法.

参 考 文 献

[1] Boţ R I, Csetnek E R. A hybrid proximal–extragradient algorithm with inertial effects[J]. Numerical Functional Analysis and Optimization, 2015, 36(8): 13, 951–963.

[2] Martinet B. Régularisation d'inéquations variationnelles par approximations successives[J]. Rev. Française Inform. Rech. Opérationnelle 4 (Ser. R–3), 1970, 154–158.

[3] Rockafellar R T. Monotone operators and the proximal point algorithm[J]. SIAM J. Control Optim., 1976, 14(5): 877–898.

[4] Alvarez F, Attouch H. An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping[J]. Set-Valued Anal., 2001, 9(1-2): 3–11.

[5] Solodov M V, Svaiter B F. A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator[J]. Set-Valued Anal., 1999, 7(4): 323–345.

[6] Burachik R S, Iusem A, Svaiter B F. Enlargement of monotone operators with applications to variational inequalities[J]. Set-Valued Anal., 1997, 5(2): 159–180.

[7] Monteiro R D C, Svaiter B F. Iteration-complexity of block-decomposition algorithms and the alternating direction method of multipliers[J]. SIAM J. Optim., 2013, 23(1): 475–507.

[8] Gonçalves M L N, Melo J G, Monteiro R D C. Improved pointwise iteration-complexity of a regularized ADMM and of a regularized non-euclidean HPE framework[J]. SIAM J. Optim., 2017, 27(1): 379–407.

[9] Alves M M, Svaiter B F. A variant of the hybrid proximal extragradient method for solving strongly monotone inclusions and its complexity analysis[J]. J. Optim. Theory Appl., 2016, 168: 198–215.

[10] Chen Caihua, Chan Raymond H, Ma Shiqian, Yang Junfeng. Inertial proximal ADMM for linearly constrained separable convex optimization[J]. SIAM J. Imag. Sci., 2015, 8(4): 2239–2267.

[11] Boţ R I, Csetnek E R, Hendrich C. Inertial Douglas-Rachford splitting for monotone inclusion problems[J]. Appl. Math. Comput., 2015, 256: 472–487.

[12] Alves M M, Marcavillaca R T. On inexact relative-error hybrid proximal extragradient, forward-backward and Tseng's modified forward-backward methods with inertial effects[J]. Set-Valued and Variational Analysis, 2019, 1–25. Published online: 15 March 2019, https://doi.org/10.1007/s11228-019-00510-7.

[13] Burachik R S, Svaiter B F. ϵ-enlargements of maximal monotone operators in Banach spaces[J]. Set-Valued Anal., 1999, 7(2): 117–132.

[14] Bauschke H H, Combettes P L. Convex analysis and monotone operator theory in Hilbert spaces[M]. New York: Springer-Verlag, 2010.

[15] Tikhonov A N, Arsenin V Y. Solutions of ill-posed problems[M]. Washington: Winston and Sons, 1977.

[16] Minty G J. Monotone (nonlinear) operators in Hilbert spaces[J]. Duke Math.J, 1962, 29(3): 341–346.

[17] Briceño A, Luis M. Forward-Douglas-Rachford splitting and forward-partial inverse method for solving monotone inclusions[J]. Optimization, 2015, 64(5): 1239–1261.

[18] Spingarn J E. Partial inverse of a monotone operator[J]. Appl. Math. Optim., 1983, 10(3): 247–265.

[19] Burachik R S, Sagastizábal C A, Scheimberg S. An inexact method of partial inverses and a parallel bundle method[J]. Optim. Methods Softw., 2006, 21(3): 385–400.

[20] Marques A M, Costa L S. On the convergence rate of the scaled proximal decomposition on the graph of a maximal monotone operator (SPDG) algorithm[J]. Optimization, 2018, 1–10.

[21] Alves M M, Lima S C. An inexact spingarn's partial inverse method with applications to operator splitting and composite optimization[J]. J. Optim. The. Appl., 2017, 175(3): 818–847.

A INERTIAL HYBRID PROXIMAL EXTRAGRADIENT METHOD FOR SOLVING MONOTONE INCLUSIONS

HE Ming-ming , PENG Jian-wen

(*School of Mathematical Sciences, Chongqing Normal University, Chongqing 401331, China*)

Abstract: In this paper, we study a new inertial hybrid proxiaml extragradient method for solving monotone inclusion problems. By using the Opial theorem, we obtain the weak convergence and the non-asymptotic global convergence rate of the inertial hybrid proxiaml extragradient method. In the framework of inertial hybrid proxiaml extragradient method, we propose and analyze the convergemce and non-asymptotic global convergence rate of an inertial Tseng's forward-backward method and an inertial inexact Spingarn's partial inverse method.

Keywords: hybrid proximal extragradient method; inertia; Tseng's forward-backward method; Spingarn's partial inverse method

2010 MR Subject Classification: 47J05; 90C60; 90C30; 65K10