

Hilbert空间中增量型不精确拟牛顿算法的局部收敛性分析

郑 滢, 张 鹏*

牡丹江师范学院数学科学学院, 黑龙江 牡丹江

收稿日期: 2025年2月17日; 录用日期: 2025年3月9日; 发布日期: 2025年3月18日

摘 要

在图像处理, 机器学习和工程学等应用领域中, 通常需要处理一些定义在Hilbert空间中的大规模算子方程。为求解这类算子方程及最值问题, 构造了一类增量型不精确Broyden方法并证明了该算法的线性收敛和局部超线性收敛性。该算法降低了在处理大规模问题中所产生的储存成本, 并通过应用证明了该算法的有效性。

关键词

Hilbert空间, 增量型不精确拟牛顿算法, Broyden方法

Local Convergence Analysis of Incremental Inexact Quasi-Newton Algorithm in Hilbert Space

Xi Zheng, Peng Zhang*

College of Mathematical Sciences, Mudanjiang Normal University, Mudanjiang Heilongjiang

Received: Feb. 17th, 2025; accepted: Mar. 9th, 2025; published: Mar. 18th, 2025

Abstract

In application fields such as image processing, machine learning, and engineering, it is often necessary to solve large-scale operator equations defined in Hilbert spaces. To address such operator equations and optimization problems, a class of incremental inexact Broyden methods has been

*通讯作者。

文章引用: 郑滢, 张鹏. Hilbert 空间中增量型不精确拟牛顿算法的局部收敛性分析[J]. 应用数学进展, 2025, 14(3): 243-257. DOI: 10.12677/aam.2025.143111

developed, and the linear convergence as well as local superlinear convergence of this algorithm has been proven. This algorithm reduces the storage costs associated with handling large-scale problems, and its effectiveness has been demonstrated through practical applications.

Keywords

Hilbert Space, Incremental Inexact Quasi-Newton Algorithm, Broyden Method

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着大数据时代的到来,其深刻改变着人类的生活和工作方式。数据量的急剧增长是大数据时代最直观的表现。从 GB 到 TB,再到 PB、EB,数据量的爆炸式增长对存储和处理能力提出了前所未有的挑战,处理速度的快慢直接关系到大数据的时效性。因此,在大数据时代需要不断创新数据处理和分析技术,才能更好地挖掘和利用大数据的潜力,推动社会进步和经济发展。

在图像处理[1]-[3],机器学习[4] [5]和工程学[6] [7]等应用领域中通常是将实际问题转换为算子方程进行求解,考虑抽象空间中算子方程的一般形式为

$$F(x) = \frac{1}{n} \sum_{i=1}^n f_i(x) = 0 \quad (1)$$

其中 $f_i(x): X \rightarrow Y$ 是非线性算子, X, Y 是 Hilbert 空间。

关于求解非线性方程组的研究已经有很长的历史。在有限维空间中,当 F 是个非线性算子时,牛顿法较为经典,其具有二阶的收敛速度,但需要精确计算导算子,使得牛顿法在实际的应用受到限制,特别在大规模问题及其不适用。为了解决上述问题,[8]提出了一种新的迭代算法——Broyden 迭代算法,其核心优势在于避免了传统导算子的复杂计算过程,并保持了优于常规算法的收敛效率。该算法在非线性的最小范数问题和非光滑优化等应用领域中表现出较为明显的优势。然而,当涉及需要离散化处理的无穷维空间问题时,经典 Broyden 方法受限于有限维近似带来的误差积累,其收敛速率明显降低。为解决该问题,[9]通过构建 Hilbert 空间的理论框架,将 Broyden 方法扩展至无穷维情形下的非线性算子方程求解,并证明了算法的线性收敛性和超线性收敛,使得 Broyden 方法更具一般性。

在无限维空间中,关于求解非线性方程组和无约束优化问题的不精确牛顿类方法的研究很多。针对 Banach 空间中半光滑算子方程的求解问题,[10]提出了两种不精确牛顿法,并分别证明了两种方法的全局收敛性。[11]通过对非线性最小二乘问题的高斯牛顿法展开理论分析,基于 Hölder 连续性的假设框架,并在仿射不变条件下建立了不精确高斯牛顿法的局部收敛理论,证明了算法在临界点邻域内具有确定的收敛半径和 $1+p$ 阶超线性收敛速率。为求解无约束极大极小问题,[12]通过引入光滑凝聚函数对非光滑极大值函数进行可微逼近,将原问题转化为光滑无约束优化形式,再使用不精确牛顿法进行求解,并证明了算法的全局收敛性。值得注意的是,[13]基于[9]的理论基础上,进一步研究了 Hilbert 空间上算子方程的不精确拟牛顿法,成功将有限维不精确拟牛顿法拓展至 Hilbert 空间中用于解决算子方程求解问题,并证明了其线性收敛性和超线性收敛性。

在 Hilbert 空间中,关于增量型不精确拟牛顿法的研究很少。当面对大规模问题时,面对大量的数据

时, 现有的计算机只能存储有限个数据和做有限次运算, 所以如何降低存储成本和提高收敛速度是有必要研究的。[14]对有限维空间中的增量拟牛顿法(IQN)的收敛性和存储成本进行了研究, 为将不精确拟牛顿法与增量思想相结合, 并推广至无限维空间中作了准备。

本文在[13]的基础上, 将 Hilbert 空间中的不精确拟牛顿法与增量思想相结合, 构造了增量型不精确拟牛顿方法用于解决大规模的抽象算子方程问题中出现的存储成本高昂的问题, 并证明了该方法的局部收敛性。

2. 预备知识

本研究建立的理论框架基于以下基本假设:

假设 1 (解的存在性): 非线性算子 $f_i(x): X \rightarrow Y$, 存在最优解 $x^* \in X$, 使得 $F'(x^*) = \frac{1}{n} \sum_{i=1}^n f'_i(x^*) = 0$ 。

假设 2: 非线性算子 f_i 在解邻域 $B_\delta(x^*)$ 内 Fréchet 可微, 且在解邻域内 f'_i 是局部 Lipschitz 连续的。

假设 3: $f'_i(x^*)$ 是非奇异的或存在 $\gamma > 0$, 使得对任意的 $x \in X$ 有 $\|f'_i(x^*)(x)\| \geq \gamma \|x\|$ 。

为证明超线性收敛, 需要下面的定义及引理:

定义 2.1 [9]: 设 X, Y 是 Hilbert 空间。算子 $D \in B(X, Y)$ 称为 Hilbert-Schmidt 算子, 若存 X 的某个标准正交基 $\{a_\alpha: \alpha \in A\}$ 使得 $\sum_{\alpha \in A} \|Da_\alpha\|_Y^2 < \infty$ 。

引理 2.1 [9]: 设 X, Y 是 Hilbert 空间, $x_i \in X$, $y_i \in Y$, $i=1, 2$, $T \in B(X, Y)$ 则以下性质成立, 其中 T^* 表示伴随算子:

- (i) $(y_1 \otimes x_1)^* = x_1 \otimes y_1 \in B(Y, X)$,
- (ii) $(y_1 \otimes x_1)(x_2 \otimes y_2) = \langle x_1, x_2 \rangle (y_1 \otimes y_2) \in B(Y, Y)$,
- (iii) $T(x_1 \otimes y_1) = Tx_1 \otimes y_1 \in B(Y, Y)$,
- (iv) $(y_1 \otimes x_1)T^* = y_1 \otimes Tx_1 \in B(Y, Y)$,
- (v) $\|y_1 \otimes x_1\| = \|y_1\| \|x_1\|$ 。

引理 2.2 [9]: 设 X, Y 是 Hilbert 空间, $x_1, x_2 \in X$, D_1 是 Hilbert-Schmidt 算子。则对于定义 2.1 中的标准正交基 $\{a_\alpha: \alpha \in A\}$, 对每个算子 $T \in B(X, X)$ 进行定义

$$\text{tr} T = \sum_{\alpha \in A} \langle a_\alpha, Ta_\alpha \rangle$$

只要和是有限的, 则以下结论成立:

- (i) 对于 $i=1, 2$ 有 $\text{tr} T_i < \infty$, 则 $\text{tr}(T_1 + T_2) = \text{tr} T_1 + \text{tr} T_2$,
- (ii) $\text{tr} x_1 \otimes x_2 = \langle x_1, x_2 \rangle$ 。

3. 增量型不精确拟牛顿型的收敛性分析

为本节主要构造了增量型不精确拟牛顿算法, 并对其收敛性进行了分析和证明, 为证明第四节 Hilbert 空间增量型不精确拟牛顿算法的收敛性做好准备。

增量型不精确拟牛顿算法

为解决大规模抽象算子问题, 根据不精确拟牛顿算法和[14]中所提出的以循环方式选择进行更新的增量拟牛顿方法, 构造的增量型不精确拟牛顿算法框架如下:

算法 3.1 增量型不精确拟牛顿算法

步 0: 设 x^0 , $\{B_i^0\}_{i=1}^n$ 分别为初始迭代点和初始迭代算子;

步 1: 设 $z_1^0 = z_2^0 = \cdots = z_n^0 = x^0$;

步 2: 设 $g^0 = \frac{1}{n} \sum_{i=1}^n f_i(x^0)$, $(\tilde{B}^0) = \frac{1}{n} \sum_{i=1}^n B_i^0$;

步 3: 对于 $k=0,1,\cdots$ 做循环更新:

3.1 设 $i_k = (k \bmod n) + 1$, 是对于所有 k 做循环选择需要满足的重要索引,

3.2 计算 $x^{k+1} = x^k + \Delta x^k$, 其中 Δx^k 为非精确求解得到并满足

$$r^k = \tilde{B}^k \Delta x^k + g^k, \quad \frac{\|r^k\|}{\|g^k\|} \leq \eta^k, \quad \eta^k \in [0,1)$$

3.3 当 $i = i_k$ 时, 计算 $s_{i_k}^{k+1}, y_{i_k}^{k+1}$ 和其中根据 $B_i^k \in B(X, Y)$ 的迭代方法计算 $B_{i_k}^{k+1}$,

$$s_i^k = z_i^{k+1} - z_i^k, \quad y_i^k = f_i(z_i^{k+1}) - f_i(z_i^k)$$

3.4 更新 $\tilde{B}^{k+1}$ 和有关函数信息表, 其中

$$\begin{cases} z_{i_k}^{k+1} = x^{k+1} & i = i_k \\ z_i^{k+1} = z_i^k & i \neq i_k \end{cases} \quad \begin{cases} f_{i_k}(z_i^{k+1}) = f_{i_k}(x^{k+1}) & i = i_k \\ f_i(z_i^{k+1}) = f_i(z_i^k) & i \neq i_k \end{cases} \quad \begin{cases} \tilde{B}^{k+1} = \tilde{B}^k + \frac{1}{n}(B_{i_k}^{k+1} - B_{i_k}^k) & i = i_k \\ g^{k+1} = g^k + \frac{1}{n}(f_{i_k}(z_{i_k}^{k+1}) - f_{i_k}(z_{i_k}^k)) & i = i_k \\ B_i^{k+1} = B_i^k & i \neq i_k \end{cases}$$

基于该算法, 有如下的线性收敛性定理:

定理 3.1: 设 X, Y 是 Hilbert 空间, $f_i: X \rightarrow Y$ 是非线性算子, 假设在假设均成立的情况下, 算法 3.1

产生的迭代序列 $\{x^k\}$, 存在 $\varepsilon > 0$ 和 $\eta > 0$ 满足 $\varepsilon < \frac{1}{\|f_i'(x^*)^{-1}\|}$ 和 $\left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \left(\frac{L(\eta+1)}{2} \varepsilon + \eta \|f_i'(x^*)\| + \varepsilon \right) \right\| < 1$,

其中 $\eta^k \in [0, \eta]$ 且 $0 < \eta < 1$, L 为 Lipschitz 常数, 当首次迭代点 x^0 与 B_i^k 对任意 $\forall k \in N$, $i = 1, 2, \dots, n$ 满足 $\|x^0 - x^*\| \leq \varepsilon$ 且 $\|B_i^k - f_i'(x^*)\| \leq \varepsilon$, 则序列 $\{x^k\}$ 满足

$$\|x^k - x^*\| \leq r^{\left\lfloor \frac{k-1}{n} \right\rfloor + 1} \|x^0 - x^*\|. \quad (2)$$

证明: 由算法 3.1 可知

$$\begin{aligned} x^{k+1} - x^* &= x^k + \Delta x^k - x^* \\ &= z_i^k + (r^k - g^k)(\tilde{B}^k)^{-1} - x^* \\ &= z_i^k + \left(r^k - \frac{1}{n} \sum_{i=1}^n f_i(z_i^k) \right) \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} - x^* \\ &= \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \left[\frac{1}{n} \sum_{i=1}^n B_i^k z_i^k + r^k - \frac{1}{n} \sum_{i=1}^n f_i(z_i^k) - \frac{1}{n} \sum_{i=1}^n B_i^k x^* \right] \\ &= \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} r^k - \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \left[\frac{1}{n} \sum_{i=1}^n f_i(z_i^k) - \frac{1}{n} \sum_{i=1}^n B_i^k (z_i^k - x^*) \right] \end{aligned} \quad (3)$$

对式(3)同时加上并减去 $\frac{1}{n} \sum_{i=1}^n f_i'(x^*)(z_i^k - x^*)$, 并重新排列项得到

$$\begin{aligned}
& \|x^{k+1} - x^*\| \\
&= \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} r^k - \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \left[\frac{1}{n} \sum_{i=1}^n f_i(z_i^k) - \frac{1}{n} \sum_{i=1}^n f_i'(x^*)(z_i^k - x^*) \right. \right. \\
&\quad \left. \left. + \frac{1}{n} \sum_{i=1}^n f_i'(x^*)(z_i^k - x^*) - \frac{1}{n} \sum_{i=1}^n B_i^k(z_i^k - x^*) \right] \right\| \\
&= \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} r^k - \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \left[\frac{1}{n} \sum_{i=1}^n f_i(z_i^k) - \frac{1}{n} \sum_{i=1}^n f_i'(x^*)(z_i^k - x^*) + \frac{1}{n} \sum_{i=1}^n (f_i'(x^*) - B_i^k)(z_i^k - x^*) \right] \right\| \\
&\leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} r^k \right\| + \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \left[\frac{1}{n} \sum_{i=1}^n f_i(z_i^k) - \frac{1}{n} \sum_{i=1}^n f_i'(x^*)(z_i^k - x^*) + \frac{1}{n} \sum_{i=1}^n (f_i'(x^*) - B_i^k)(z_i^k - x^*) \right] \right\|
\end{aligned}$$

由于 $r^k = \tilde{B}^k \Delta x^k + g^k$ 且满足 $\frac{\|r^k\|}{\|g^k\|} \leq \eta^k$, $\eta^k \in [0, \eta)$, 则有

$$\begin{aligned}
& \|x^{k+1} - x^*\| \\
&\leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} r^k \right\| + \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \left[\frac{1}{n} \sum_{i=1}^n f_i(z_i^k) - \frac{1}{n} \sum_{i=1}^n f_i'(x^*)(z_i^k - x^*) + \frac{1}{n} \sum_{i=1}^n (f_i'(x^*) - B_i^k)(z_i^k - x^*) \right] \right\| \\
&\leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \eta \|g^k\| + \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \left\| \frac{1}{n} \sum_{i=1}^n f_i(z_i^k) - \frac{1}{n} \sum_{i=1}^n f_i'(x^*)(z_i^k - x^*) + \frac{1}{n} \sum_{i=1}^n (f_i'(x^*) - B_i^k)(z_i^k - x^*) \right\|
\end{aligned}$$

由 $f_i'(z_i^k)$ 在 x^* 处的泰勒展开得到

$$\begin{aligned}
& \|x^{k+1} - x^*\| \\
&\leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} r^k \right\| + \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \left[\frac{1}{n} \sum_{i=1}^n f_i(z_i^k) - \frac{1}{n} \sum_{i=1}^n f_i'(x^*)(z_i^k - x^*) \right. \right. \\
&\quad \left. \left. + \frac{1}{n} \sum_{i=1}^n (f_i'(x^*) - B_i^k)(z_i^k - x^*) \right] \right\| \\
&\leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \eta \|g^k\| + \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \left\| \frac{1}{n} \sum_{i=1}^n f_i(z_i^k) - \frac{1}{n} \sum_{i=1}^n f_i'(x^*)(z_i^k - x^*) \right. \\
&\quad \left. + \frac{1}{n} \sum_{i=1}^n (f_i'(x^*) - B_i^k)(z_i^k - x^*) \right\| \\
&\leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \eta \left\| \frac{1}{n} \sum_{i=1}^n f_i(z_i^k) \right\| + \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \left(\frac{L}{2n} \sum_{i=1}^n \|z_i^k - x^*\|^2 + \frac{1}{n} \sum_{i=1}^n \|f_i'(x^*) - B_i^k\| \|z_i^k - x^*\| \right)
\end{aligned}$$

又因 $\|B_i^k - f_i'(x^*)\| \leq \varepsilon$, 最后得到

$$\begin{aligned}
\|x^{k+1} - x^*\| &\leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \eta \left(\frac{L}{2n} \sum_{i=1}^n \|z_i^k - x^*\|^2 + \frac{1}{n} \sum_{i=1}^n \|f_i'(x^*)\| \|z_i^k - x^*\| \right) \\
&\quad + \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \left(\frac{L}{2n} \sum_{i=1}^n \|z_i^k - x^*\|^2 + \frac{1}{n} \sum_{i=1}^n \|f_i'(x^*) - B_i^k\| \|z_i^k - x^*\| \right) \\
&\leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \frac{L(\eta+1)}{2n} \sum_{i=1}^n \|z_i^k - x^*\|
\end{aligned} \tag{4}$$

$$\begin{aligned}
 & + \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \left(\left\| \frac{\eta}{n} f'_i(x^*) \right\| + \frac{1}{n} \|f'_i(x^*) - B_i^k\| \right) \sum_{i=1}^n \|z_i^k - x^*\| \\
 & \leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \left(\left\| \frac{L(\eta+1)}{2n} \varepsilon + \frac{\eta}{n} f'_i(x^*) \right\| + \frac{\varepsilon}{n} \right) \sum_{i=1}^n \|z_i^k - x^*\|
 \end{aligned}$$

由式(4), 可得到首次迭代

$$\begin{aligned}
 \|x^1 - x^*\| & \leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^0 \right)^{-1} \right\| \left(\left\| \frac{L(\eta+1)}{2n} \varepsilon + \frac{\eta}{n} f'_i(x^*) \right\| + \frac{\varepsilon}{n} \right) \sum_{i=1}^n \|z_i^0 - x^*\| \\
 & \leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^0 \right)^{-1} \right\| \left(\left\| \frac{L(\eta+1)}{2} \varepsilon + \eta \|f'_i(x^*)\| + \varepsilon \right\| \right) \|x^0 - x^*\|
 \end{aligned}$$

由于 $r = \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \left(\left\| \frac{L(\eta+1)}{2} \varepsilon + \eta \|f'_i(x^*)\| + \varepsilon \right\| \right) < 1$, 则

$$\|x^1 - x^*\| \leq r \|x^0 - x^*\| < \varepsilon$$

继续进行下一次迭代, 有不等式

$$\begin{aligned}
 \|x^2 - x^*\| & \leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^1 \right)^{-1} \right\| \left(\left\| \frac{L(\eta+1)}{2n} \varepsilon + \frac{\eta}{n} f'_i(x^*) \right\| + \frac{\varepsilon}{n} \right) \sum_{i=1}^n \|z_i^1 - x^*\| \\
 & = \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^1 \right)^{-1} \right\| \left(\left\| \frac{L(\eta+1)}{2} \varepsilon + \eta \|f'_i(x^*)\| + \varepsilon \right\| \right) \left(\frac{n-1}{n} \|x^0 - x^*\| + \frac{1}{n} \|x^1 - x^*\| \right) \\
 & \leq r \left(\frac{n-1}{n} \|x^0 - x^*\| + \frac{1}{n} \|x^1 - x^*\| \right) \\
 & \leq r \|x^0 - x^*\|
 \end{aligned}$$

同理可证, 对于所有的迭代 $k=1, \dots, n$, 有 $\|x^k - x^*\| \leq \varepsilon$ 。此外, 当 $k=1, \dots, n$ 时, 有 $\|x^k - x^*\| \leq r \|x^0 - x^*\|$ 。

现在使用迭代 $k=1, \dots, n$ 作为归纳论证的基础。为了更精确, 假设对于迭代 $k=jn+1, jn+2, \dots, jn+n$, 残差的上界为 $\|x^k - x^*\| \leq r^{j+1} \|x^0 - x^*\|$ 。

下面证明, 对于迭代序列 $k=(j+1)n+1, (j+1)n+2, \dots, (j+1)n+n$, 不等式 $\|x^k - x^*\| \leq r^{j+2} \|x^0 - x^*\|$ 成立。

基于不等式 $\|B_i^k - f'_i(x^*)\| \leq \varepsilon$ 且 $\|f'_i(x^*)\|$ 有界, 可以证明, 对于所有的 $k=jn+1, jn+2, \dots, jn+n$, 有

$\left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\|$ 是一致有界的。对于迭代 $k=(j+1)n+1$, 有

$$\begin{aligned}
 \|x^{(j+1)n+1} - x^*\| & \leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^{(j+1)n} \right)^{-1} \right\| \left(\left\| \frac{L(\eta+1)}{2n} \varepsilon + \frac{\eta}{n} f'_i(x^*) \right\| + \frac{\varepsilon}{n} \right) \sum_{i=1}^n \|z_i^{(j+1)n} - x^*\| \\
 & \leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^{(j+1)n} \right)^{-1} \right\| \left(\left\| \frac{L(\eta+1)}{2n} \varepsilon + \frac{\eta}{n} f'_i(x^*) \right\| + \frac{\varepsilon}{n} \right) \sum_{i=1}^n \|z_i^{(j+1)n} - x^*\|
 \end{aligned} \tag{5}$$

由于变量以循环方式更新, 因此变量集合 $\{z_i^{(j+1)n}\}_{i=1}^{i=n}$ 等于集合 $\{x^{(j+1)n-i}\}_{i=0}^{i=n-1}$ 且对于所有的 $j=1, \dots, n$ 满

足 $\|x^{jn+i} - x^*\| \leq \varepsilon$, 因此可以将式(5)的右侧简化为

$$\begin{aligned} \|x^{(j+1)n+1} - x^*\| &\leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^{(j+1)n} \right)^{-1} \left(\frac{L(\eta+1)}{2} \varepsilon + \eta \|f'_i(x^*)\| + \varepsilon \right) \left(\frac{1}{n} \sum_{i=1}^n \|x^{jn+i} - x^*\| \right) \right\| \\ &\leq r \left(\frac{1}{n} \sum_{i=1}^n \|x^{jn+i} - x^*\| \right) \end{aligned} \quad (6)$$

并假设对迭代序列 $k = jn+1, jn+2, \dots, jn+n$, 有 $\|x^k - x^*\| \leq r^{j+1} \|x^0 - x^*\|$, 则有

$$\|x^{(j+1)n+1} - x^*\| \leq r^{j+2} \|x^0 - x^*\| \quad (7)$$

即证得式(2)。

因此, 算法 3.1 具有 Q 线性收敛性, 其线性收敛因子 $r < 1$ 随参数 ε 和 η 的选择而变化, 可获得多样的收敛性因子值。

定理 3.2: 设 X, Y 是 Hilbert 空间, $f_i: X \rightarrow Y$ 是非线性算子, 在假设条件均成立的情况下, $\{\tilde{B}^k\}$ 可逆。假设由算法 3.1 产生的迭代序列 $\{x^k\}$ 收敛到 x^* , 若 $\eta^k = o(1)$, 即 $\|r_i^k\| = o(\|f_i(z_i^{k+1})\|)$, 则残差序列 $\|x^k - x^*\|$ 满足:

$$\lim_{k \rightarrow \infty} \frac{\|x^k - x^*\|}{\frac{1}{n} (\|x^{k-1} - x^*\| + \dots + \|x^{k-n} - x^*\|)} = 0 \quad (8)$$

的充要条件是对于所有的 $i=1, 2, \dots, n$ 有

$$\|(B_i^k - f'_i(x^*))(z_i^k - x^*)\| = o(\|z_i^k - x^*\|)。$$

证明: 当 f_i 满足假设条件(a)~(c)时, 对任意向量 $x, \hat{x}, x^* \in X$ 有

$$\|f_i(\hat{x}) - f_i(x) - f'_i(x^*)(\hat{x} - x)\| \leq L \max\{\|\hat{x} - x^*\|, \|x - x^*\|\} \|\hat{x} - x\| \quad (9)$$

其中 L 是 Lipschitz 常数。故

$$(B_i^k - f'_i(x^*))(z_i^{k+1} - z_i^k) = f_i(z_i^{k+1}) - f_i(z_i^k) - f'_i(x^*)(z_i^{k+1} - z_i^k) + r_i^k - f_i(z_i^{k+1}) \quad (10)$$

结合式(9)与式(10)得到

$$\begin{aligned} &\frac{\|r_i^k - f_i(z_i^{k+1})\|}{\|z_i^{k+1} - z_i^k\|} - L \max\{\|z_i^k - x^*\|, \|z_i^{k+1} - x^*\|\} \\ &\leq \frac{\|(B_i^k - f'_i(x^*))(z_i^{k+1} - z_i^k)\|}{\|z_i^{k+1} - z_i^k\|} \\ &\leq L \max\{\|z_i^k - x^*\|, \|z_i^{k+1} - x^*\|\} + \frac{\|r_i^k - f_i(z_i^{k+1})\|}{\|z_i^{k+1} - z_i^k\|} \end{aligned}$$

从而

$$\lim_{k \rightarrow \infty} \frac{\|(B_i^k - f'_i(x^*))(z_i^{k+1} - z_i^k)\|}{\|z_i^{k+1} - z_i^k\|} = 0 \quad (11)$$

式(11)成立的充要条件为

$$\lim_{k \rightarrow \infty} \frac{\|r_i^k - f_i(z_i^{k+1})\|}{\|z_i^{k+1} - z_i^k\|} = 0 \quad (12)$$

下面证明式(12)为式(8)成立的充要条件:

假设 $\lim_{k \rightarrow \infty} \frac{\|x^k - x^*\|}{\frac{1}{n}(\|x^{k-1} - x^*\| + \dots + \|x^{k-n} - x^*\|)} = 0$ 成立且 $\sum_{i=1}^n \|z_i^k - x^*\| = \sum_{i=0}^{n-1} \|x^{k-i} - x^*\|$, 则有

$$\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^*\|}{\frac{1}{n} \sum_{i=1}^n \|z_i^k - x^*\|} = 0$$

表明算法 3.1 生成的迭代序列具有平均超线性收敛速度。更准确地说, 它表明在步骤 k 捕获误差的比值除以最后 n 个误差的平均值收敛于零。

由式(9)可知

$$\|f_i(z_i^{k+1}) - f_i(x^*) - f_i'(x^*)(z_i^{k+1} - x^*)\| \leq L \|z_i^{k+1} - x^*\|^2$$

可立即得到

$$\begin{aligned} \|f_i(z_i^{k+1})\| &\leq \|f_i(x^*) + f_i'(x^*)(z_i^{k+1} - x^*)\| + L \|z_i^{k+1} - x^*\|^2 \\ &= O(\|z_i^{k+1} - x^*\|) = o(\|z_i^k - x^*\|) \end{aligned}$$

同理可得 $\|f_i(z_i^k)\| = o(\|z_i^k - x^*\|)$ 。

又因为 $\|r_i^k\| \leq \eta^k \|g_i^k\| = \eta^k \|f_i(z_i^k)\|$, 其中 $\eta^k \rightarrow 0$, 从而有

$$\|r_i^k\| = o(\|z_i^k - x^*\|)$$

故

$$\|r_i^k - f_i(z_i^{k+1})\| = o(\|z_i^k - x^*\|) = o(\|z_i^{k+1} - z_i^k\|)$$

反之若

$$\lim_{k \rightarrow \infty} \frac{\|r_i^k - f_i(z_i^{k+1})\|}{\|z_i^{k+1} - z_i^k\|} = 0$$

则式(11)成立。

下证 $\lim_{k \rightarrow \infty} \frac{\|x^k - x^*\|}{\frac{1}{n}(\|x^{k-1} - x^*\| + \dots + \|x^{k-n} - x^*\|)} = 0$, 由定理 3.1 中的式(4)可得

$$\begin{aligned} \|x^{k+1} - x^*\| &\leq \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \left(\frac{L}{2n} \sum_{i=1}^n \|z_i^k - x^*\|^2 + \frac{1}{n} \sum_{i=1}^n \|f_i'(x^*)\| \|z_i^k - x^*\| \right) \\ &\quad + \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \left(\frac{L}{2n} \sum_{i=1}^n \|z_i^k - x^*\|^2 + \frac{1}{n} \sum_{i=1}^n \|f_i'(x^*) - B_i^k\| \|z_i^k - x^*\| \right) \end{aligned} \quad (13)$$

将式(13)的两边同时除以 $\frac{1}{n} \sum_{i=1}^n \|z_i^k - x^*\|$ 得到

$$\frac{\|x^{k+1} - x^*\|}{\frac{1}{n} \sum_{i=1}^n \|z_i^k - x^*\|} \leq \frac{\left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \left(\left(\frac{\eta L}{2} \sum_{i=1}^n \|z_i^k - x^*\|^2 + \eta \sum_{i=1}^n \|f'_i(x^*)\| \|z_i^k - x^*\| \right) \right) \right\|}{\sum_{i=1}^n \|z_i^k - x^*\|} + \frac{\left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \left(\left(\frac{L}{2} \sum_{i=1}^n \|z_i^k - x^*\|^2 + \sum_{i=1}^n \|(f'_i(x^*) - B_i^k)(z_i^k - x^*)\| \right) \right) \right\|}{\sum_{i=1}^n \|z_i^k - x^*\|}$$

由于误差 $\|z_i^k - x^*\|$ 是误差 $\sum_{i=1}^n \|z_i^k - x^*\|$ 的和下界, 因此将 $\sum_{i=1}^n \|z_i^k - x^*\|$ 替换为 $\|z_i^k - x^*\|$, 则有

$$\frac{\|x^{k+1} - x^*\|}{\frac{1}{n} \sum_{i=1}^n \|z_i^k - x^*\|} \leq \frac{\left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| L(1+\eta)}{2} \sum_{i=1}^n \|z_i^k - x^*\| + \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \eta \sum_{i=1}^n \|f'_i(x^*)\| \|z_i^k - x^*\| + \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \sum_{i=1}^n \frac{\|(f'_i(x^*) - B_i^k)(z_i^k - x^*)\|}{\|z_i^k - x^*\|} \quad (14)$$

由于 $\left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\|$ 有上界, 对式(14)求极限得到:

$$\lim_{k \rightarrow \infty} \frac{\|x^{k+1} - x^*\|}{\frac{1}{n} \sum_{i=1}^n \|z_i^k - x^*\|} = 0.$$

定理 3.2 的结果并不等价于经典的 Q-超线性收敛的拟牛顿法, 即 $\lim_{k \rightarrow \infty} \|x^{k+1} - x^*\| / \|x^k - x^*\| = 0$ 。虽然残差 $\|x^k - x^*\|$ 的 Q-超线性收敛性是不可证明的, 但可以证明序列 $\|x^k - x^*\|$ 存在一个超线性收敛于零的子序列且该子序列为 $\|x^k - x^*\|$ 的上界。

定理 3.3: 假设定理 3.1 和定理 3.2 中的条件成立, 迭代序列由算法 3.1 产生, 存在 $\|x^k - x^*\|$ 一个子序列并超线性收敛到零。即存在一个序列 ζ^k , 对于所有 $k \geq 0$ 满足 $\|x^k - x^*\| \leq \zeta^k$, 且序列 ζ^k 以超线性速率收敛到零, 即有

$$\lim_{k \rightarrow \infty} \frac{\zeta^{k+1}}{\zeta^k} = 0$$

证明: 见文[14]。

4. Broyden 修正不精确增量拟牛顿法的收敛性分析

本节基于第三节的内容, 进一步证明了 Hilbert 空间中增量型不精确拟牛顿算法的局部收敛性。

增量型不精确 Broyden 算法

若 B_i^k 的迭代公式由 Broyden 修正给出, 得到算法 4.1:

算法 4.1 增量型不精确 Broyden 算法

步 0: 给出初始迭代点 $x^0 \in X$ 和初始迭代算子 $\{B_i^0\}_{i=1}^n$, 其中 x^0 与 x^* 充分接近, 算子 $\{B_i^0\}_{i=1}^n$ 与 $\{f'_i(x^0)\}_{i=1}^n$ 也充分接近, 并给出合理的容差范围;

步 1: 设 $z_1^0 = z_2^0 = \dots = z_n^0 = x^0$;

步 2: 设 $g^0 = \frac{1}{n} \sum_{i=1}^n f_i(x^0)$, $(\tilde{B}^0) = \frac{1}{n} \sum_{i=1}^n B_i^0$;

步 3: 对于 $k=0, 1, \dots$ 做循环更新直到 $\{x^k\}$ 收敛;

3.1 设 $i_k = (k \bmod n) + 1$, 是对于所有 k 做循环选择需要满足的重要索引,

3.2 计算 $x^{k+1} = x^k + \Delta x^k$, 其中 Δx^k 为非精确求解得到并满足

$$r^k = \tilde{B}^k \Delta x^k + g^k, \quad \frac{\|r^k\|}{\|g^k\|} \leq \eta^k, \quad \eta^k \in [0, 1)$$

3.3 当 $i = i_k$ 时, 计算 $s_{i_k}^{k+1}, y_{i_k}^{k+1}, B_{i_k}^{k+1}$, 其中

$$s_i^k = z_i^{k+1} - z_i^k, \quad y_i^k = f_i(z_i^{k+1}) - f_i(z_i^k), \quad B_i^{k+1} = B_i^k + \frac{(y_i^k - B_i^k s_i^k) \otimes s_i^k}{\langle s_i^k, s_i^k \rangle}$$

3.4 更新 $\tilde{B}^{k+1}$ 和有关函数信息表, 其中

$$\begin{cases} z_{i_k}^{k+1} = x^{k+1} & i = i_k \\ z_i^{k+1} = z_i^k & i \neq i_k \end{cases} \quad \begin{cases} f_{i_k}(z_i^{k+1}) = f_{i_k}(x^{k+1}) & i = i_k \\ f_i(z_i^{k+1}) = f_i(z_i^k) & i \neq i_k \end{cases} \quad \begin{cases} \tilde{B}^{k+1} = \tilde{B}^k + \frac{1}{n} (B_{i_k}^{k+1} - B_{i_k}^k) & i = i_k \\ g^{k+1} = g^k + \frac{1}{n} (f_{i_k}(z_{i_k}^{k+1}) - f_{i_k}(z_{i_k}^k)) & i = i_k \\ B_i^{k+1} = B_i^k & i \neq i_k \end{cases}$$

由第三节内容可证得算法 4.1 的线性收敛定理如下:

定理 4.1: X, Y 是 Hilbert 空间, $f_i: X \rightarrow Y$ 是非线性算子, 在满足所有假设条件下且 $\{\tilde{B}^k\}$ 可逆. L 为 Lipschitz 常数. 控制序列 $\{\eta^k\}$ 满足 $\eta^k \in [0, \eta)$, $0 < \eta < 1$, 存在 $\varepsilon > 0$ 和 $\eta > 0$, 当

$$\varepsilon < \frac{1}{\|f'_i(x^*)^{-1}\|}$$

且

$$\left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \left(\left(\frac{L(\eta+1)}{2} \varepsilon + \eta \|f'_i(x^*)\| + \varepsilon \right) \right) \right\| < 1$$

时, 若 $x^0 \in X$ 与 B_i^0 满足 $\|x^0 - x^*\| \leq \varepsilon$ 且 $\|B_i^0 - f'_i(x^*)\| \leq \varepsilon$, 则由算法 4.1 产生的迭代点列 $\{x^k\}$ 至少 Q-线性收敛于 x^* .

证明: 由定理 3.1 证明可知, 存在 $\varepsilon^1 > 0$ 和 $\eta^1 > 0$, 使得 $\|x^0 - x^*\| \leq \varepsilon^1$, $\|B_i^k - f'_i(x^*)\| \leq \varepsilon^1$, 对任意 $k \in N$ 有 $\|x^{k+1} - x^*\| \leq r^1 \|x^k - x^*\|$, 其中 $r^1 \in (0, 1)$. 若初始迭代点 x^0 满足 $\|x^0 - x^*\| \leq \varepsilon < \varepsilon^1$ 且初始迭代算子 B_i^0 满足 $\|B_i^0 - f'_i(x^*)\| \leq \varepsilon < \varepsilon^1$, 则式子 $\|x^t - x^*\| \leq \varepsilon^1$, $\|B_i^t - f'_i(x^*)\| \leq \varepsilon^1$, 对于 $t=0$ 是成立的. 接下来, 将采用归纳法, 证明上述两个式子对任意情况的 t 均成立. 假设对于 $t=0, 1, \dots, k \in N$ 时, 有 $\|x^t - x^*\| \leq \varepsilon^1$,

$\|B_i^t - f_i'(x^*)\| \leq \varepsilon^1$ 成立, 又

$$\begin{aligned} \|B_i^{t+1} - f_i'(x^*)\| &= \left\| B_i^t + \frac{(y_i^t - B_i^t s_i^t) \otimes s_i^t}{\langle s_i^t, s_i^t \rangle} - f_i'(x^*) \right\| \\ &= \left\| (B_i^t - f_i'(x^*)) \left(I - \frac{s_i^t \otimes s_i^t}{\langle s_i^t, s_i^t \rangle} \right) + \frac{(y_i^t - f_i'(x^*) s_i^t) \otimes s_i^t}{\langle s_i^t, s_i^t \rangle} \right\| \\ &\leq \|B_i^t - f_i'(x^*)\| \left\| I - \frac{s_i^t \otimes s_i^t}{\langle s_i^t, s_i^t \rangle} \right\| + \left\| \frac{(y_i^t - f_i'(x^*) s_i^t) \otimes s_i^t}{\langle s_i^t, s_i^t \rangle} \right\| \end{aligned} \quad (15)$$

由 X 的内积的定义, 有

$$\left\| I - \frac{s_i^t \otimes s_i^t}{\langle s_i^t, s_i^t \rangle} \right\| = 1$$

利用归纳假设并由定理 3.1 的证明过程可知

$$\|x^{t+1} - x^*\| \leq r^1 \|x^t - x^*\|, t = 0, 1, \dots, k \quad (16)$$

由上式(16)以及 $f_i'(x^*)$ 的局部 Lipschitz 连续性有对于 $t = 0, 1, \dots, k$,

$$\begin{aligned} \|(y_i^t - f_i'(x^*) s_i^t) \otimes s_i^t\| &\leq \|s_i^t\| \|y_i^t - f_i'(x^*) s_i^t\| \\ &\leq \|f_i'(x^t + \tau_i^t s_i^t) - f_i'(x^*)\| \|s_i^t\|^2 \\ &\leq L(\tau_i^t \|x^{t+1} - x^*\| + (1 - \tau_i^t) \|x^t - x^*\|) \|s_i^t\|^2 \\ &\leq L \|x^t - x^*\| \|s_i^t\|^2 \end{aligned}$$

由 $f_i'(x^*)$ 的连续性和中值定理可得 $\tau_i^t \in [0, 1]$, L 为 Lipschitz 常数。

令

$$\varepsilon = \varepsilon^1 \left(1 + 2L + 2L(n-1)r^1 + 2L(r^1)^2 \right)^{-1}$$

显然 $\varepsilon < \varepsilon^1$, 对于 $t = 0, 1, \dots, k$ 有

$$\|x^{t+1} - x^*\| \leq r^1 \|x^t - x^*\| \leq \dots \leq (r^1)^{\left[\frac{t}{n}\right]+1} \|x^0 - x^*\| \leq (r^1)^{\left[\frac{t}{n}\right]+1} \varepsilon$$

故

$$\|B_i^{t+1} - f_i'(x^*)\| \leq \|B_i^t - f_i'(x^*)\| + L \|x^t - x^*\| \leq \|B_i^t - f_i'(x^*)\| + 2L \|x^t - x^*\| \quad (17)$$

对式(17)两边进行求和可得

$$\begin{aligned} \|B_i^{k+1} - f_i'(x^*)\| &\leq \|B_i^0 - f_i'(x^*)\| + \sum_{t=0}^k 2L \|x^t - x^*\| \\ &\leq \|B_i^0 - f_i'(x^*)\| + \sum_{t=0}^k 2L (r^1)^{\left[\frac{t-1}{n}\right]+1} \varepsilon \\ &\leq \varepsilon + 2L \varepsilon \left(1 + (n-1)r^1 + (r^1)^2 \right) \\ &= \varepsilon \left(1 + 2L + 2L(n-1)r^1 + 2L(r^1)^2 \right) = \varepsilon^1 \end{aligned}$$

这说明 $\|x' - x^*\| \leq \varepsilon^1$, $\|B_i^t - f'_i(x^*)\| \leq \varepsilon^1$ 对于所有的 k 都成立, 结合定理 3.1 有

$$\|x^{k+1} - x^*\| \leq (r^1)^{\left[\frac{k}{n}\right]+1} \|x^0 - x^*\| < \varepsilon^1$$

该算法的线性收敛得到证明。

定理 4.2: 在定理 4.1 的假设条件下, 初始迭代算子 B_i^0 满足 $B_i^0 - f'_i(x^*)$ 是 Hilbert-Schmidt 算子且 $\|r_i^k\| = o(\|f_i(z_i^{k+1})\|)$ 。则存在 $\varepsilon > 0$ 使得当 x^0 和 B_i^0 满足 $\|x^0 - x^*\| \leq \varepsilon$ 和 $\|B_i^0 - f'_i(x^*)\| \leq \varepsilon$ 时, 算法 4.1 产生的残差序列 $\|x^k - x^*\|$ 满足:

$$\lim_{k \rightarrow \infty} \frac{\|x^k - x^*\|}{\frac{1}{n}(\|x^{k-1} - x^*\| + \dots + \|x^{k-n} - x^*\|)} = 0$$

证明: 由式(15)有

$$\begin{aligned} B_i^{k+1} - f'_i(x^*) &= B_i^k - f'_i(x^*) + \frac{(y_i^t - B_i^t s_i^t) \otimes s_i^t}{\langle s_i^t, s_i^t \rangle} \\ &= (B_i^k - f'_i(x^*)) \left(I - \frac{s_i^k \otimes s_i^k}{\langle s_i^k, s_i^k \rangle} \right) + \frac{(y_i^k - f'_i(x^*) s_i^k) \otimes s_i^k}{\langle s_i^k, s_i^k \rangle} \end{aligned} \quad (18)$$

令 $D_k = B_i^k - f'_i(x^*) \in B(X, Y)$, $e_k = \frac{s_i^k}{\|s_i^k\|} \in X$, $d_k = \frac{y_i^k - f'_i(x^*) s_i^k}{\|s_i^k\|} \in Y$, 则式(18)可简化为

$$D_{k+1} = D_k (I - e_k \otimes e_k) + d_k \otimes e_k$$

由引理 2.1 中的性质(i)~(iv)和 $\|e_k\| = 1$, 可以简化 $D_{k+1}^* D_{k+1} \in B(Y, Y)$ 如下:

$$\begin{aligned} D_{k+1}^* D_{k+1} &= ((I - e_k \otimes e_k) D_k^* + e_k \otimes d_k) (D_k (I - e_k \otimes e_k) + d_k \otimes e_k) \\ &= (I - e_k \otimes e_k) D_k^* D_k (I - e_k \otimes e_k) + (e_k \otimes d_k) (D_k - D_k e_k \otimes e_k) \\ &\quad + (I - e_k \otimes e_k) D_k^* d_k \otimes e_k + \|d_k\|^2 e_k \otimes e_k \\ &= D_k^* D_k - (e_k \otimes D_k^* D_k e_k + D_k^* D_k e_k \otimes e_k) + \|D_k e_k\|^2 e_k \otimes e_k \\ &\quad + e_k \otimes D_k^* d_k + D_k^* d_k \otimes e_k + (\|d_k\|^2 - 2 \langle d_k, D_k e_k \rangle) e_k \otimes e_k \end{aligned} \quad (19)$$

通过假设, D_1 是一个 Hilbert-Schmidt 算子, 即对于 X 存在 $\{e_i\}_{i \in I}$ 的标准正交基使得

$$\sum_{i \in I} \langle e_i, D_1^* D_1 e_i \rangle < \infty$$

在式(19)中从 $D_k^* D_k$ 到 $D_{k+1}^* D_{k+1}$ 的每一个修正都由秩为 1 的算子的有限和给出。因此, 根据引理 2.2 知所有算子 $D_k^* D_k$, $k \in N$, 的迹 $\text{tr}(D_k^* D_k)$ 是有限的。通过 tr 和 $\text{tr}(x_1 \otimes x_2) = \langle x_1, x_2 \rangle$ 的线性性质, 对式(19)在两侧取迹进行化简得到:

$$\text{tr}(D_{k+1}^* D_{k+1}) = \text{tr}(D_k^* D_k) - \|D_k e_k\|^2 + \|d_k\|^2 \quad (20)$$

连续使用式(20)可得

$$\text{tr}(D_{k+1}^* D_{k+1}) = \text{tr}(D_1^* D_1) + \sum_{i=1}^k (\|d_i\|^2 - \|D_i e_i\|^2) \quad (21)$$

根据定义 $\text{tr}(D_{k+1}^* D_{k+1})$ 是范数之和是非负的, 则有:

$$\sum_{i=1}^k \|D_i e_i\|^2 \leq \text{tr}(D_1^* D_1) + \sum_{i=1}^k \|d_i\|^2 \quad (22)$$

在 f_i 可微的情况下, 可以使用 Lipschitz 常数 $L > 0$ 和[15]中的中值定理知对 $\|d_k\|$ 进行估计

$$\begin{aligned} \|d_k\| &= \|f_i(x^{k+1}) - f_i(x^k) - f'_{ix^*}(x^{k+1} - x^k)\| \|x^{k+1} - x^k\|^{-1} \\ &\leq L \|x^k - x^*\| \end{aligned}$$

利用线性收敛速度, 得到对于某个 $\kappa \in (0, 1)$, 对所有的 $k \in N$ 有

$$\begin{aligned} \sum_{i=0}^k \|d_i\|^2 &\leq \sum_{i=0}^k L^2 \|x^i - x^*\|^2 \leq L^2 \|x^0 - x^*\|^2 \sum_{i=0}^{k-1} (\kappa^2)^i \\ &\leq L^2 \|x^0 - x^*\| (1 - \kappa^2)^{-1} \end{aligned}$$

因此, 式(22)中左侧不等式的和也是有限的, 即

$$\sum_{i=1}^k \|D_i e_i\|^2 < \infty$$

并且得到

$$\lim_{k \rightarrow \infty} \|D_k e_k\| = \lim_{k \rightarrow \infty} \frac{\|(B_i^k - f'_i(x^*)) s_i^k\|}{\|s_i^k\|} = 0 \quad (23)$$

式(23)为 Dennis-More 条件[16]。已知其在有限维情况下是超线性收敛的。为了达到在无限维情况下的目的有,

$$\gamma \frac{\|x^{k+1} - x^*\|}{\frac{1}{n} \sum_{i=1}^n \|z_i^k - x^*\|} \leq \frac{\|f'_i(x^*)(x^{k+1} - x^*)\|}{\frac{1}{n} \sum_{i=1}^n \|z_i^k - x^*\|} \leq n \frac{\|f'_i(x^*)(x^{k+1} - x^*)\|}{\|z_i^k - x^*\|} \leq 2n \frac{\|(B_i^k - f'_i(x^*)) s_i^k\|}{\|s_i^k\|} + L \|z_i^k - x^*\|^2 \quad (24)$$

其中第一个不等式由假设 3 得到, 第二个不等式由误差 $\|z_i^k - x^*\|$ 是误差 $\sum_{i=1}^n \|z_i^k - x^*\|$ 的和下界得到, 第三个不等式由文[9]中定理 3.1 的证明可知, 并结合式(23), 当 $k \rightarrow \infty$ 时, 有

$$\lim_{k \rightarrow \infty} \frac{\|x^k - x^*\|}{\frac{1}{n} (\|x^{k-1} - x^*\| + \dots + \|x^{k-n} - x^*\|)} = 0.$$

5. 应用

针对非线性算子方程的高效求解问题, 考虑采用具有增量型不精确 Broyden 方法。参见[17]知可将非线性积分方程如下表示:

X 为 Hilbert 空间, Y 为 Banach 空间, $G: X \rightarrow Y$ 是非线性算子, $V: Y \rightarrow X$ 是线性算子。对于给定 $h \in Y$, 有最优解 x^* 满足

$$x^* - VGx^* = h \quad (25)$$

定理 5.1: 设 X, Y 分别是 Hilbert 空间和 Banach 空间, 在 $B_\delta(x^*)$ 内, $G: X \rightarrow Y$ 是 Fréchet 可微的且导数 G'_x 满足 Lipschitz 连续性, $V \in B(Y, X)$, 且对任意的 $x \in X$, $V \circ G'_x: X \rightarrow X$ 是 Hilbert-Schmidt 算

子, 且 $V \circ G'_{x^*}$ 的特征值不是 1, $\|r_i^k\| = o(\|f_i(z_i^{k+1})\|)$ 。则存在 $\varepsilon > 0$ 和 $\eta > 0$, 当

$$\varepsilon < \frac{1}{\|f'_i(x^*)^{-1}\|} \text{ 且 } \left\| \left(\frac{1}{n} \sum_{i=1}^n B_i^k \right)^{-1} \right\| \left(\frac{L(\eta+1)}{2} \varepsilon + \eta \|f'_i(x^*)\| + \varepsilon \right) < 1,$$

若初始迭代点 $x^0 \in X$ 满足 $\|x^0 - x^*\| \leq \varepsilon$, 初始迭代算子 $B_i^0 \in B(X, X)$ 且 $B_i^0 = I - v_i g'_i(x^0)$ 时, 算法 4.1 产生的迭代点列超线性收敛到 x^* 。

证明: 定义 $F(x) = \frac{1}{n} \sum_{i=1}^n f_i(x)$ 且 $f_i(x): X \rightarrow X$ 为

$$f_i(x) = x - v_i g_i x - h$$

并假设 f_i 是 f_i -可微的, 则

$$f'_i(x) = I - v_i g'_i(x)$$

在满足定理 5.1 的条件下, 存在 $m > 0$ 对于所有的 $u \in U$ 有

$$\|f'_i(x^*)(u)\| = \|(I - v_i g'_i(x^*))u\| \geq m \|u\|.$$

由 $B_i^0 = I - v_i g'_i(x^0)$ 有

$$B_i^0 - f'_i(x^*) = I - v_i g'_i(x^0) - I + v_i g'_i(x^*) = v_i (g'_i(x^*) - g'_i(x^0)) \quad (26)$$

和对于某些常数 $\kappa > 0$ 有

$$\|B_i^0 - f'_i(x^*)\| \leq \|v_i\| \|g'_i(x^*) - g'_i(x^0)\| \leq \|v_i\| \kappa \|x^* - x^0\|,$$

因此, 若 $\|x^* - x^0\|$ 很小, 则 $B_i^0 - f'_i(x^*)$ 也很小。此外, 由假设与 Hilbert-Schmidt 算子的可加性, 证明式(26)是 Hilbert-Schmidt 算子。因此, 由定理 4.2 可得到超线性收敛速度结果。

基金项目

这项研究由黑龙江省科研业务费一般项目(项目编号: 1452MSYYB007)资助。

参考文献

- [1] Zhang, W., Zhuang, P., Sun, H., Li, G., Kwong, S. and Li, C. (2022) Underwater Image Enhancement via Minimal Color Loss and Locally Adaptive Contrast Enhancement. *IEEE Transactions on Image Processing*, **31**, 3997-4010. <https://doi.org/10.1109/tip.2022.3177129>
- [2] Xiang, H., Zou, Q., Nawaz, M.A., Huang, X., Zhang, F. and Yu, H. (2023) Deep Learning for Image Inpainting: A Survey. *Pattern Recognition*, **134**, Article ID: 109046. <https://doi.org/10.1016/j.patcog.2022.109046>
- [3] Tian, C., Zheng, M., Zuo, W., Zhang, S., Zhang, Y. and Lin, C. (2024) A Cross Transformer for Image Denoising. *Information Fusion*, **102**, Article ID: 102043. <https://doi.org/10.1016/j.inffus.2023.102043>
- [4] Hamdia, K.M., Zhuang, X. and Rabczuk, T. (2020) An Efficient Optimization Approach for Designing Machine Learning Models Based on Genetic Algorithm. *Neural Computing and Applications*, **33**, 1923-1933. <https://doi.org/10.1007/s00521-020-05035-x>
- [5] Zhang, J. (2019) Gradient Descent Based Optimization Algorithms for Deep Learning Models Training.
- [6] Yan, G., Zou, H., Wang, S., Zhao, L., Wu, Z. and Zhang, W. (2022) Bio-Inspired Toe-Like Structure for Low-Frequency Vibration Isolation. *Mechanical Systems and Signal Processing*, **162**, Article ID: 108010. <https://doi.org/10.1016/j.ymssp.2021.108010>
- [7] Li, H.N., Wang, W., Lai, S.K., Yao, L.Q. and Li, C. (2023) Nonlinear Vibration and Stability Analysis of Rotating Functionally Graded Piezoelectric Nanobeams. *International Journal of Structural Stability and Dynamics*, **24**, Article

ID: 2450103. <https://doi.org/10.1142/s0219455424501037>

- [8] Broyden, C.G. (1965) A Class of Methods for Solving Nonlinear Simultaneous Equations. *Mathematics of Computation*, **19**, 577-593. <https://doi.org/10.1090/s0025-5718-1965-0198670-6>
- [9] Sachs, E.W. (1986) Broyden's Method in Hilbert Space. *Mathematical Programming*, **35**, 71-82. <https://doi.org/10.1007/bf01589442>
- [10] 刘晶, 高岩. Banach 空间中半光滑算子方程的不精确牛顿法(英文) [J]. 运筹学学报, 2010, 14(3): 41-47.
- [11] 李丙通, 贾春霞. 不精确高斯法的局部收敛性质[J]. 上海师范大学学报(自然科学版), 2011, 40(5): 460-468.
- [12] 路云龙. 求解无约束极大极小问题的光滑化不精确牛顿算法[J]. 北华大学学报(自然科学版), 2014, 15(5): 593-595.
- [13] 王娟, 于波. 一类不精确拟牛顿型算法的局部收敛性分析[J]. 数学的实践与认识, 2017, 47(19): 237-244.
- [14] Mokhtari, A., Eisen, M. and Ribeiro, A. (2018) IQN: An Incremental Quasi-Newton Method with Local Superlinear Convergence Rate. *SIAM Journal on Optimization*, **28**, 1670-1698. <https://doi.org/10.1137/17m1122943>
- [15] 薛小平, 秦泗甜, 等. 非线性分析[M]. 北京: 科学出版社, 2017.
- [16] Dennis, J.E. and Moré, J.J. (1974) A Characterization of Superlinear Convergence and Its Application to Quasi-Newton Methods. *Mathematics of Computation*, **28**, 549-560. <https://doi.org/10.1090/s0025-5718-1974-0343581-1>
- [17] Anselone, P.M. and Davis, J. (1971) Collectively Compact Operator Approximation Theory and Applications to Integral Equations.