六阶收敛的非线性方程组求根方法



课程:数值分析 任课教师:黎卫兵/张雨浓

年级	2022	专业	计算机科学与技术
学号	22214373	姓名	林泽佳
班级	学硕 2 班	日期	2022年12月24日

摘要

本文以张老师上课所介绍的构造 8 阶收敛的牛顿法, 和 PPT 第 63 页简要介绍 Steffenson 和 Halley 方法为基础和启发,构造了 6 阶收敛的非线性方程组求根方法,并使用前向差商代替二阶 导数,从而实现只需求一阶导数。本文使用泰勒展开证明了其收敛阶,充分的数值试验对比了本 方法与牛顿法、Halley 法和若干现有的 6 阶方法。

引言 1

本节主要对张老师上课所介绍的构造8阶收敛的牛顿法,和PPT第63页简要介绍Steffensen 和 Halley 方法做一个简要的回顾, 作为本文的基础。

1.1 8 阶收敛的牛顿法

使用三次牛顿法反复代入,如公式1所示,可以得到序列 $\{y_n\}_{n=0\infty}$ 是 2 阶收敛, $\{z_n\}_{n=0\infty}$ 是 4 阶收敛,因此最终 $\{x_n\}_{n=0}^{\infty}$ 是 8 阶收敛的,这是一种简单的提高迭代公式阶数的方法。

$$y_n = x_n - \frac{f(x_n)}{f'(x_n)}$$

$$z_n = y_n - \frac{f(y_n)}{f'(y_n)}$$

$$x_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)}$$

$$(1)$$

1.2 Steffenson 方法

由于在零点附近 $\lim_{n\to\infty} f(x_n) = 0$, 因此可以使用前向差商可以将零点附近的导数近似为:

$$f'(x) \approx \frac{f(x_n + f(x_n))}{f(x_n)} \tag{2}$$

将公式2代入牛顿迭代公式,可以得到:

$$x_{n+1} = x_n - \frac{f(x_n)^2}{f(x + f(x_n)) - f(x_n)}$$
(3)

这就是 PPT 第 63 页所提到的 Steffensen 公式,它是也二阶收敛的 [1],同时避免了求导。

1.3 Halley 方法

PPT 第 63 页同样提到了三阶收敛的 Halley 方法和 Chebyshev 方法,查阅文献 [2] 后发现,它们属于同一族方法,具有如下形式:

$$x_{n+1} = x_n - \left(1 + \frac{1}{2} \frac{L_f(x_n)}{1 - \beta L_f(x_n)}\right) \frac{f(x_n)}{f'(x_n)},\tag{4}$$

其中

$$L_f(x_n) = \frac{1}{2} \frac{f''(x_n)f(x_n)}{[f'(x_n)]^2}$$
 (5)

特别的,当 $\beta=0$ 时为 Chebyshev 方法, $\beta=\frac{1}{2}$ 时为 Halley 方法, $\beta=1$ 时为超 Halley 方法。不少研究对这族方法进行改进,但是它们都具有**较为复杂的迭代格式**,如 [3] 提出的一种无需求二阶导的三阶收敛的方法:

$$\begin{cases} y_n = x_n - \theta \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} = x_n - \frac{\theta^2 f(x_n)}{(\theta^2 - \theta + 1) f(x_n) - f(y_n)} \frac{f(x_n)}{f'(x_n)}, & (\theta \in \mathbb{R}, \theta \neq 0), \end{cases}$$
(6)

又如 [4] 提出的另一种三阶收敛的方法:

$$\begin{cases} w_n &= x_n - \frac{f(x_n)}{f'(x_n)}, \\ K_{\alpha}(x_n) &= \frac{2f(x_n)f(w_n)(1 + \alpha f'^2(x_n))}{f^2(x_n) + \alpha f'^2(x_n)[f(w_n) - f(x_n)]^2}, \quad (\alpha \in \mathbb{R}), \\ x_n + 1 &= x_n - \left(1 + \frac{1}{2} \frac{K_{\alpha}(x_n)}{1 - \beta K_{\alpha}(x_n)}\right) \frac{f(x_n)}{f'(x_n)}, \quad (\beta \in \mathbb{R}), \end{cases}$$
(7)

这些方法与同为三阶收敛的其它方法相比,或多或少能减少一些迭代次数,但代价是较为复杂的迭代格式。

1.4 六阶收敛的方法

在三阶收敛方法的基础上,再使用一次牛顿迭代或其它二阶迭代法,即可实现六阶收敛的方法。如最早由[5]提出的六阶收敛方法:

$$\begin{cases} w_n &= x_n - \frac{f(x_n)}{f'(x_n)}, \\ z_n &= w_n - \frac{f(w_n)}{f'(x_n)} \cdot \frac{f(x_n) - \frac{1}{2}f(w_n)}{f(x_n) - \frac{5}{2}f(w_n)}, \\ x_{n+1} &= z_n - \frac{f(z_n)}{f'(x_n)} \cdot \frac{f(x_n) - f(w_n)}{f(x_n) - 3f(w_n)} \end{cases}$$
(8)

又如 [6] 中对 Ostrowski 方法 [7] 改进提出的六阶收敛的方法

$$\begin{cases} y_n = x_n - \frac{f(x_n)}{f'(x_n)}, \\ z_n = y_n - \frac{y_n - x_n}{2f(y_n) - f(x_n)} f(y_n), \\ x_{n+1} = z_n - \frac{y_n - x_n}{2f(y_n) - f(x_n)} f(z_n), \end{cases}$$
(9)

又如[8] 中提出的用前向差商代替导数的六阶收敛的方法:

$$\begin{cases} w_n &= x_n + f(x_n), \\ y_n &= x_n - \frac{f(x_n)}{f[x_n, w_n]}, \\ z_n &= x_n - \frac{f(x_n)}{f[x_n, w_n]} \left(1 + \frac{f(y_n)}{f(x_n)} \left(1 + 2\frac{f(y_n)}{f(x_n)} \right) \right), \\ x_{n+1} &= z_n - \frac{f(z_n)}{f[y_n, z_n]} \left(1 - \frac{1 + f[x_n, w_n] f(z_n)}{f[x_n, w_n] f(w_n)} \right) \end{cases}$$

$$(10)$$

除此之外,[9][10][11] 等均提出了不同形式的六阶收敛的方法,虽然它们的计算效率较高,但 均具有较为复杂的迭代格式。本文希望能建立一种较为简单的迭代格式对 Halley 方法进行改造, 在避免求二阶导数的基础上实现六阶收敛的方法。诚信声明,在我所了解到的文献里没有与本文 方法相同的,虽然推导和证明过程均有参考它们,也有相应的引用,但皆为原创和独立完成;如 有雷同,是我查找资料不周全,没有恶意抄袭的意思。

2 方法设计

2.1 方法推导

在三阶的 Halley 方法的基础上再进行一次牛顿法迭代得到的序列,显然是六阶收敛的:

$$\begin{cases}
z_n = x_n - \frac{2f(x_n)f'(x_n)}{2[f'(x_n)]^2 - f(x_n)f''(x_n)}, \\
x_{n+1} = x_n - \frac{f(z_n)}{f'(z_n)}
\end{cases} (11)$$

参考 [12] 中的方法,希望消去 $\{z_n\}_{n=0}^{\infty}$ 中的二阶导数,如果记 $y_n = x_n - \frac{f(x)}{f'(x)}$,则有 $\lim_{n \to \infty} \frac{y_n}{x_n} = 1$,因此二阶导数可以使用一阶导计算前向差商近似为:

$$f''(x_n) = \lim_{n \to \infty} \frac{f'(y_n) - f'(x_n)}{y_n - x_n}$$

$$\approx \frac{f'(y_n) - f'(x_n)}{y_n - x_n},$$
(12)

将公式12代入公式11, 使用 sympy1 (代码见附录 1) 化简得:

$$\begin{cases} y_n = x_n - \frac{f(x)}{f'(x)} \\ z_n = x_n - \frac{2f(x_n)}{f'(x_n) + f'(y_n)} \\ x_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)} \end{cases}$$
(13)

公式13即为本文所提出的方法,由于它使用了差商代替二阶导数,因此需要对收敛阶进行证明。

2.2 收敛阶证明

我参考了 [3] 和 [12] 使用的技巧对公式13的收敛阶进行证明。由于涉及到 **6 阶泰勒展开**的计算较为复杂,因此我使用了 sympy¹进行符号计算和化简,**代码在附录 2 和 code/convergence.py** 中,本部分所推导的结果皆由该代码产生,其正确性保证了本文推导结果的正确。由于篇幅有限,部分泰勒展开用省略号表示。

假设函数 f(x) 的单根是 α 且 $f'(\alpha) \neq 0$,记截断误差 $e_n = \alpha - x_n$,对 $f(x_n)$ 在 α 处进行泰勒展开得:

$$f(x_n) = f(\alpha + e_n)$$

$$= f(\alpha) + f'(\alpha)e_n + \frac{1}{2}f''(\alpha)e_n^2 + \frac{1}{3!}f'''(\alpha)e_n^3 + \dots + O(e_n^7)$$

$$= f'(\alpha) \left[e_n + \frac{f''(\alpha)e_n^2}{2f'(\alpha)} + \frac{f'''(\alpha)e_n^3}{3!f'(\alpha)} + \dots + O(e_n^7) \right],$$
(14)

注意到 $f(\alpha) = 0$,所以公式中将它消去了,记 $c_k = \frac{f^{(k)}(\alpha)}{f'(\alpha)k!}$,则有

$$f(x_n) = f'(\alpha) \left[c_1 e_n + c_2 e_n^2 + c_3 e_n^3 + c_4 e_n^4 + c_5 e_n^5 + c_6 e_n^6 + O(e_n^7) \right]$$

$$= f'(\alpha) \left[\sum_{k=1}^6 c_k e_n^k + O(e_n^7) \right]$$
(15)

¹sympy 是 Python 的符号计算库, https://www.sympy.org/en/index.html

同理对 $f'(x_n)$ 在 α 处进行泰勒展开得:

$$f'(x_n) = f'(\alpha + e_n)$$

$$= f'(\alpha) + f''(\alpha)e_n + \frac{1}{2}f'''(\alpha)e_n^2 + \dots + O(e_n^6)$$

$$= f'(\alpha)[1 + 2c_2e_n + 3c_3e_n^2 + 4c_4e_n^3 + 5c_5e_n^4 + 6c_6e_n^6 + O(e_n^6)]$$

$$= f'(\alpha)\left[1 + \sum_{k=2}^6 kc_ke_n^{k-1} + O(e_n^6)\right]$$
(16)

希望得到 y_n 的泰勒展开形式,因此首先需要计算:

$$\frac{f(x_n)}{f'(x_n)} = \frac{\sum_{k=1}^{6} c_k e_n^k + O(e_n^7)}{1 + \sum_{k=2}^{6} k c_k e_n^{k-1} + O(e_n^6)},\tag{17}$$

记 $\theta = 1 + \sum_{k=2}^{6} kc_k e_n^{k-1} + O(e_n^6)$,显然 $\lim_{n \to \infty} \theta = 0$,因此在 $\theta = 0$ 处对分母进行泰勒展开有(代码第 61 行):

$$\frac{1}{1+\theta} = 1 - \theta - 2c_2e + \theta^2 - \theta^3 + \theta^4 - \theta^5 + O(\theta^6)$$

$$= 1 + e_n^2 \cdot (4c_2^2 - 3c_3) + e_n^3 \left(-8c_2^3 + 12c_2c_3 - 4c_4 \right)$$

$$+ e_n^4 \cdot \left(16c_2^4 - 36c_2^2c_3 + 16c_2c_4 + 9c_3^2 - 5c_5 \right)$$

$$+ e_n^5 \left(-32c_2^5 + 96c_2^3c_3 - 48c_2^2c_4 - 54c_2c_3^2 + 20c_2c_5 + 24c_3c_4 - 6c_6 \right) + O\left(e_n^6\right),$$
(18)

将公式18代入公式17并化简有(代码第65行):

$$\frac{f(x_n)}{f'(x_n)} = e - c_2 e_n^2 + e_n^3 \cdot \left(2c_2^2 - 2c_3\right) + e_n^4 \cdot \left(-4c_2^3 + 7c_2c_3 - 3c_4\right)
+ e_n^5 \cdot \left(8c_2^4 - 20c_2^2c_3 + 10c_2c_4 + 6c_3^2 - 4c_5\right)
+ e_n^6 \cdot \left(-16c_2^5 + 52c_2^3c_3 - 28c_2^2c_4 - 33c_2c_3^2 + 13c_2c_5 + 17c_3c_4 - 5c_6\right) + O\left(e_n^7\right),$$
(19)

因此求得 y_n 的泰勒展开 (代码第 70 行):

$$y_{n} = x_{n} - \frac{f(x_{n})}{f'(x_{n})}$$

$$= \alpha + c_{2}e_{n}^{2} + e_{n}^{3} \cdot \left(-2c_{2}^{2} + 2c_{3}\right) + e_{n}^{4} \cdot \left(4c_{2}^{3} - 7c_{2}c_{3} + 3c_{4}\right)$$

$$+ e_{n}^{5} \cdot \left(-8c_{2}^{4} + 20c_{2}^{2}c_{3} - 10c_{2}c_{4} - 6c_{3}^{2} + 4c_{5}\right)$$

$$+ e_{n}^{6} \cdot \left(16c_{2}^{5} - 52c_{2}^{3}c_{3} + 28c_{2}^{2}c_{4} + 33c_{2}c_{3}^{2} - 13c_{2}c_{5} - 17c_{3}c_{4} + 5c_{6}\right) + O\left(e_{n}^{7}\right),$$

$$(20)$$

公式20同时也证明了牛顿法是二阶收敛的,截断误差为 $\frac{f''(\alpha)}{2f'(\alpha)}e_n^2 + O(e_n^3)$ 。继续对 $f'(y_n)$ 在 α

处做泰勒展开有(代码第77行):

$$f'(y_n) = f'(\alpha) + \left[c_2 e_n^2 + (2c_2^2 - 2c_3)e_n^3 + O(e_n^4)\right] f''(\alpha) + \dots + O(e_n^6)$$

$$= 1 + 2c_2^2 e_n^2 + e_n^3 \left(-4c_2^3 + 4c_2c_3\right) + e_n^4 \cdot \left(8c_2^4 - 11c_2^2c_3 + 6c_2c_4\right)$$

$$+ e_n^5 \cdot \left(-16c_2^5 + 28c_2^3c_3 - 20c_2^2c_4 + 8c_2c_5\right) + O\left(e_n^6\right),$$
(21)

将公式16与公式21相加得到 z_n 的分母部分的泰勒展开:

$$f'(x_n) + f'(y_n) = 2f'(\alpha) \left[1 + c_2 e_n + \left(c_2^2 + \frac{3}{2} c_3 \right) e_n^2 + \dots + O(e_n^6) \right], \tag{22}$$

因此将公式22与公式14相除得到 z_n 的泰勒展开:

$$x_n - \frac{2f(x_n)}{f'(x_n) + f'(y_n)} = \alpha + e_n - \frac{2\left[\sum_{k=1}^6 c_k e_n^k + O(e_n^7)\right]}{1 + c_2 e_n + \left(c_2^2 + \frac{3}{2}c_3\right)e_n^2 + \dots + O(e_n^6)},$$
 (23)

使用与公式18相似的技巧,对分母部分进行泰勒展开得(代码第82行):

$$x_{n} - \frac{2f(x_{n})}{f'(x_{n}) + f'(y_{n})} = \alpha + e_{n}^{3} \cdot \left(c_{2}^{2} + \frac{c_{3}}{2}\right) + e_{n}^{4} \cdot \left(-3c_{2}^{3} + \frac{3c_{2}c_{3}}{2} + c_{4}\right)$$

$$+ e_{n}^{5} \cdot \left(6c_{2}^{4} - 9c_{2}^{2}c_{3} + 2c_{2}c_{4} - \frac{3c_{3}^{2}}{4} + \frac{3c_{5}}{2}\right)$$

$$+ e_{n}^{6} \cdot \left(-9c_{2}^{5} + 25c_{2}^{3}c_{3} - 15c_{2}^{2}c_{4} - \frac{5c_{2}c_{3}^{2}}{2} + \frac{5c_{2}c_{5}}{2} - \frac{5c_{3}c_{4}}{2} + 2c_{6}\right)$$

$$+ O\left(e_{n}^{7}\right),$$

$$(24)$$

公式24同时也证明了序列 $\{z_n\}_{n=0}^{\infty}$ 是三阶收敛的,即使用一阶导数的前向差商来计算二阶导数的 Halley 方法是仍是三阶收敛的。继续计算 $f(z_n)$ 和 $f'(z_n)$ 的泰勒展开,使用与公式14-20相同的 方法,可以得到 x_{n+1} 的泰勒展开(代码第 88 行):

$$x_{n+1} = \alpha + e^6 \cdot \left(c_2^5 + c_2^3 c_3 + \frac{c_2 c_3^2}{4}\right) + O\left(e^7\right), \tag{25}$$

因此证明了本方法是六阶收敛的,截断误差为 $e^6 \cdot \left(c_2^5 + c_2^3 c_3 + \frac{c_2 c_3^2}{4}\right) + O\left(e^7\right)$

2.3 一些有趣的发现

由于使用 sympy 可以很方便地进行泰勒展开来对收敛阶进行推导,我又尝试了一下修改迭代格式,使用与 $\{z_n\}_{n=0}^{\infty}$ 相似的格式来代替最后一步的牛顿法:

$$\begin{cases} y_n = x_n - \frac{f(x)}{f'(x)} \\ z_n = x_n - \frac{2f(x_n)}{f'(x_n) + f'(y_n)} , \\ x_{n+1} = z_n - \frac{2f(z_n)}{f'(y_n) + f'(z_n)} \end{cases}$$
(26)

附录 2 的代码第 94 行展示了收敛性证明的计算过程,使用与公式21-25类似的方法,可以得到 x_{n+1} 的泰勒展开:

$$x_{n+1} = e^5 \cdot \left(c_2^4 + \frac{c_2^2 c_3}{2}\right) + e^6 \cdot \left(-5c_2^5 + \frac{5c_2^3 c_3}{2} + c_2^2 c_4 + c_2 c_3^2\right) + \alpha + O\left(e^7\right),\tag{27}$$

因此公式26只有 5 阶收敛,但它的计算量比牛顿法还多(同样需要计算 $f(z_n)$ 和 $f'(z_n)$,且需要额外计算一次加法和乘法),进一步让我感受到了牛顿法无可比拟的优势。

2.4 计算效率讨论

在[13] 中定义了迭代法的效率指数(Efficiency index)为:

$$EI = q^{1/\omega},\tag{28}$$

其中 q 是收敛阶, ω 是每次迭代需要计算的函数的次数(包括其导数)。例如牛顿法二阶收敛,需要计算一次 $f(x_n)$ 和一次 $f'(x_n)$,因此效率指数为 $EI=2^{1/2}$ 。

表1列举了本方法和部分方法的效率指数,和方法中额外引入的乘除法次数,这些方法也将会在实验中进行验证。可以看到,本方法具有相对简单的迭代格式,尽管效率指数略低于现有方法,但引入额外的乘除法次数较少,且减少了一次计算 f(x),仍然具有一定的实用价值。

农工 二起代码的开放工									
二 分十	1/4 A47A	计算次数			 _ 效率指数	宛 周			
方法	收敛阶	f(x)	f'(x)	f''(x)		额外乘除法次数			
牛顿法	2	1	1	0	$2^{1/2} \approx 1.414$	1			
Halley 法(公式4)	3	1	1	1	$3^{1/3} \approx 1.442$	4			
NM[5] (公式8)	6	3	1	0	$6^{1/4} \approx 1.567$	7			
GM[6] (公式9)	6	3	1	0	$6^{1/4} \approx 1.567$	5			
本方法(公式13)	6	2	3	0	$6^{1/5} \approx 1.431$	3			

表 1: 一些迭代法的计算效率

3 实验

3.1 测试方法和指标

测试代码在附录 3 和 code/compute.py 中,本文复现了表1中的 5 种方法,表2列举了本文用于测试的函数和零点(大部分来源于 [3] 和 [9]),图1展示了这些函数在零点附近的图像。使用 IEEE 64 位浮点数(具有约 15 位十进制的规格化有效数字)进行计算,迭代终止条件是 $|x_n-x_{n-1}|+|f(x_n)|<10^{-12}$,评价指标为迭代次数。

表 2: 本文用于测试的函数和零点

记号	函数	零点
f_1	$x^3 + 4x^2 - 15$	$x^* \approx 1.6319808055660636$
f_2	$x^2 - e^x - 3x + 2$	$x^* \approx 0.2575302854398608$
f_3	$xe^{x^2} - \sin^2(x) + 3\cos(x) + 5$	$x^* \approx -1.207647827130919$
f_4	$\sin^2(x) - x^2 + 1$	$x^* \approx 1.4044916482153411$
f_5	$\ln(x^2 + 7x + 14) - x - 2$	$x^* \approx 1.1525907367571583$
f_6	$\frac{(x-4)(x+1)^4}{e^x}$	$x^* = -1$ (四重根)
f_7	$e^{x^2 + 11x - 12} - 1$	$x^* = 1$
f_8	$(x-1)^2\arctan(e^{x+3}-1)$	$x^* = 1$ (二重根)

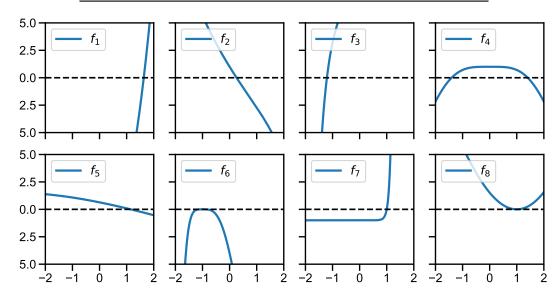


图 1: 测试函数在零点附近的图像

3.2 实验结果

表3展示表1的方法对表2的函数进行测试的实验结果,对于每个函数都使用了 2 个分布在根两端的初值进行实验。迭代次数中的 (U) 表示下溢(接近 0 而超过了精度范围,导致变成 0 了),(F) 表示上溢(绝对值太大),(N) 表示迭代发散。注意 IEEE 64 位浮点数具有大约 15 位的规格化有效数字。

表 3: 迭代次数和误差

函数	初值	方法	迭代 次数	与 <i>x</i> * 误差	函数	初值	 方法 	迭代 次数	与 <i>x</i> * 误差
		牛顿	6	$-2.22 \cdot 10^{-16}$			牛顿	4	$2.22 \cdot 10^{-16}$
		Halley	4	$-2.22 \cdot 10^{-16}$			Halley	3	$-2.22 \cdot 10^{-16}$
	1	NM	6	$-4.44 \cdot 10^{-16}$		1	NM	4	$< 10^{-16}$
		GM	3	$-2.22 \cdot 10^{-16}$			GM	3(U)	-
f		本文	3	$< 10^{-16}$			本文	2	$-2.22 \cdot 10^{-16}$
f_1		牛顿	5	$< 10^{-16}$	f_5		牛顿	5	$-2.22 \cdot 10^{-16}$
		Halley	4	$< 10^{-16}$			Halley	4	$< 10^{-16}$
	2	NM	5	$-4.44 \cdot 10^{-16}$		2	NM	5(U)	-
		GM	3	$-2.22 \cdot 10^{-16}$			GM	3(U)	-
		本文	3	$-2.22 \cdot 10^{-16}$			本文	3	$-4.44 \cdot 10^{-16}$
		牛顿	5	$< 10^{-16}$			牛顿	91	$-2.6 \cdot 10^{-12}$
		Halley	4	$< 10^{-16}$	f		Halley	53	$-1.01 \cdot 10^{-12}$
	0	NM	4	$< 10^{-16}$		-1.5	NM	22	$-1.71 \cdot 10^{-13}$
		GM	3(U)	-			GM	37	$-8.79 \cdot 10^{-13}$
£		本文	3	$< 10^{-16}$			本文	38	$-7.74 \cdot 10^{-13}$
f_2		牛顿	5	$< 10^{-16}$	f_6		牛顿	90	$2.32 \cdot 10^{-12}$
		Halley	4	$< 10^{-16}$			Halley	52	$1.25\cdot10^{-12}$
	1	NM	5	$-2.22 \cdot 10^{-16}$		-0.5	NM	22	$1.78\cdot10^{-13}$
		GM	3(U)	-			GM	37	$6.25\cdot10^{-13}$
		本文	3	$< 10^{-16}$			本文	38	$5.08 \cdot 10^{-13}$
ſ	9	牛顿	9	$2.22 \cdot 10^{-16}$	r	0.5	牛顿	2(F)	-
f_3	-2	Halley	5	$2.22\cdot10^{-16}$	f_7	0.5	Halley	7	$< 10^{-16}$
		NM	2(F)	-			NM	2(F)	-
	-2	GM	4	$< 10^{-16}$		0.5	GM	2(F)	-
		本文	4	$2.22 \cdot 10^{-16}$			本文	2(F)	-
£		牛顿	6	$2.22 \cdot 10^{-16}$	f f		牛顿	12	$-1.11 \cdot 10^{-16}$
f_3		Halley	4	$2.22\cdot10^{-16}$	f_7		Halley	7	$-1.11 \cdot 10^{-16}$
	-1	NM	7	$-1.02 \cdot 10^{-14}$		1.5	NM	2(F)	-

续下页

函数	初值	方法	迭代 次数	与 <i>x</i> * 误差	函数	初值	方法	迭代 次数	与 <i>x</i> * 误差
		GM	3	$< 10^{-16}$			GM	5	$< 10^{-16}$
		本文	3	$2.22\cdot10^{-16}$			本文	6	$< 10^{-16}$
		牛顿	6	$< 10^{-16}$			牛顿	39	$-9.01 \cdot 10^{-13}$
	1	Halley	4	$2.22\cdot10^{-16}$			Halley	26	$-1.96 \cdot 10^{-13}$
		NM	7	$2.22\cdot10^{-16}$		0.5	NM	27(N)	-4.0
		GM	3	$< 10^{-16}$			GM	17	$-2.18 \cdot 10^{-13}$
£		本文	3	$< 10^{-16}$			本文	16	$-1.75 \cdot 10^{-13}$
f_4		牛顿	6	$< 10^{-16}$	f_8		牛顿	39	$9.14 \cdot 10^{-13}$
	2	Halley	4	$2.22\cdot10^{-16}$			Halley	26	$1.97\cdot10^{-13}$
		NM	7	$2.22\cdot10^{-16}$		1.5	NM	14(N)	-4.0
		GM	3	$< 10^{-16}$			GM	17	$2.19\cdot10^{-13}$
		本文	3	$< 10^{-16}$			本文	16	$1.78\cdot10^{-13}$

表 3 迭代次数和误差(接上页)

3.3 结果分析

表4总结了这 5 种方法的成功次数和失败原因。Halley 法在所有测试中均成功得到结果,这可能是因为它使用了二阶导数,数值稳定性非常好。牛顿法和本方法除了有一次上溢以外也都成功了。GM 法出现了 4 次下溢错误,这是因为它的公式中包含"小数减小数"、"小数除小数"等计算,因此非常容易产生下溢,数值稳定性不好。NM 法中出现了 3 次上溢和 2 次发散,这可能是因为它每代额外使用 7 次乘除法,加速了误差积累。本方法作为六阶收敛的方法,数值稳定性与牛顿法相当,高于另外两种方法,且计算效率(公式28)高于牛顿法,具有良好的实用价值。

对前 5 个函数 $f_1 \sim f_5$,各方法基本有比较好的效果,并且迭代次数趋势也与收敛阶相匹配。对于第 6 个函数 f_6 ,其**零点是 4 重根**,因此牛顿法收敛速度非常慢,NM 法最快,GM 法和本方法相当。对于第 7 个函数 f_7 ,选取初值 $x_0 = 0.5$ 时,多种方法均产生了上溢,观察函数图像(图1)发现它在 (-2,1] 上非常平缓,因此**导数很小,基于牛顿法的一系列方法在计算** $\frac{f(x_n)}{f'(x_n)}$ 时**容易上溢**。但是当选取另一个初值 $x_0 = 1.5$ 时,由于导数很大,因此没有这个问题。对于第 8 个函数 f_8 ,NM 很意外的发散了,因此及时终止了迭代,由于时间有限我没有仔细探究其原因。由于**零点是 2 重根**的原因,牛顿法同样收敛的很慢,但本方法拥有最快的收敛速度,GM 法其次。

	表 4: 各万法成功次数和失败原因									
类型	牛顿	Halley	NM[5]	GM[6]	本文					
		(公式4)	(公式8)	(公式9)	(公式13)					
成功	15	16	10	12	15					
下溢	0	0	1	3	0					
上溢	1	0	3	1	1					
发散	0	0	2	0	0					

表 4: 各方法成功次数和失败原因

3.4 对重根的讨论

计算收敛阶 (Computational order of convergence) [14] 是数值方法在实际计算过程中的收敛阶,其定义如下:

$$COC = \frac{\ln\left(\left|\frac{x_{n+1} - x_n}{x_n - x_{n-1}}\right|\right)}{\ln\left(\left|\frac{x_n - x_{n-1}}{x_{n-1} - x_{n-2}}\right|\right)}$$
(29)

表5所示本文列举了 $f_6 \sim f_8$ 的计算收敛阶。对于单根的函数 f_7 ,各方法的计算收敛阶基本符合 其理论的收敛阶;对于重根函数 f_6 和 f_8 ,所有方法只有线性收敛。

表 5. 月异里限时时月异牧蚁团								
重复度	初值	牛顿	Halley	NM	GM	本文		
单根	1.5	2.002	3.066	发散	5.990	6.053		
4 重根	-1.5	0.991	0.998	1.000	0.985	0.993		
	-0.5	0.996	0.996	0.999	1.001	1.000		
2 重根	0.5	1.000	0.998	发散	1.000	1.000		
	1.5	0.977	1.021	发散	1.000	0.999		
	重复度 单根 4 重根	重复度 初值 単根 1.5 4 重根 -1.5 -0.5 2 重根 0.5	重复度 初值 牛顿 单根 1.5 2.002 4 重根 -1.5 0.991 -0.5 0.996 2 重根 0.5 1.000	重复度 初值 牛顿 Halley 单根 1.5 2.002 3.066 4 重根 -1.5 0.991 0.998 -0.5 0.996 0.996 2 重根 0.5 1.000 0.998	重复度 初值 牛顿 Halley NM 单根 1.5 2.002 3.066 发散 4 重根 -1.5 0.991 0.998 1.000 -0.5 0.996 0.996 0.999 2 重根 0.5 1.000 0.998 发散	重复度 初值 牛顿 Halley NM GM 单根 1.5 2.002 3.066 发散 5.990 4 重根 -1.5 0.991 0.998 1.000 0.985 -0.5 0.996 0.996 0.999 1.001 2 重根 0.5 1.000 0.998 发散 1.000		

表 5: 计算重根时的计算收敛阶

3.5 迭代过程展示

作为补充材料,表6和表7对迭代过程进行了展示。

4 总结

表 6: f4 选取初值为 2 的迭代过程

\overline{n}	牛顿	Halley	NM	GM	本文
0	2	2	2	2	2
1	1.543143068960336	1.456885216221384	1.120415325387814	1.407237330215151	1.405535212978439
2	1.417094222312942	1.404562548049610	1.309731172778497	1.404491648215341	1.404491648215341
3	1.404614018363034	1.404491648215529	1.397104012247177	1.404491648215341	1.404491648215341
4	1.404491659946959	1.404491648215341	1.404448806412075	-	-
5	1.404491648215341	-	1.404491646777146	-	-
6	1.404491648215341	-	1.404491648215341	-	-

表 7: f7 选取初值为 1.5 的迭代过程

	牛顿	Halley	NM	GM	本文
0	1.5	1.5	1.5	1.5	1.5
1	1.428655062830056	1.356011165775886	0.666554094953474	1.302765996348761	1.323425736359648
2	1.356719234358469	1.211129011680508	$-1.275596982 \cdot 10^{52}$	1.109913322973212	1.147701833153800
3	1.284419811223503	1.078073976922075	-	1.002996956434495	1.017028589466088
4	1.212406308451123	1.006179477275287	-	1.000000000003765	1.000000403894250
5	1.142418159478025	1.000003327216270	-	1.00000000000000000	1.00000000000000000
6	1.078725914448773	1.00000000000000000	-	-	1.00000000000000000
7	1.029866713280862	1.00000000000000000	-	-	-
8	1.005182160439837	-	-	-	-
9	1.000172764038992	-	-	-	-
10	1.000000196158916	-	-	-	-
11	1.000000000000253	-	-	-	-
12	1.00000000000000000	-	-	-	-

参考文献

- [1] Pankaj Jain. Steffensen type methods for solving non-linear equations. Applied Mathematics and Computation, 194(2):527–533, 2007.
- [2] Jose Manuel Gutierrez and Miguel A Hernández. A family of chebyshev-halley type methods in banach spaces. Bulletin of the Australian Mathematical Society, 55(1):113–130, 1997.
- [3] Jisheng Kou, Yitian Li, and Xiuhua Wang. Modified halley's method free from second derivative. Applied Mathematics and Computation, 183(1):704–708, 2006.
- [4] Changbum Chun. Some second-derivative-free variants of chebyshev-halley methods. Applied Mathematics and Computation, 191(2):410-414, 2007.
- [5] Beny Neta. A sixth-order family of methods for nonlinear equations. Int. J. Comput. Math, 7(1997):157–161, 1979.
- [6] Miquel Grau and José Luis Díaz-Barrero. An improvement to ostrowski root-finding method. Applied Mathematics and Computation, 173(1):450–456, 2006.
- [7] AS Householder. Solution of equations and systems of equations (am ostrowski). SIAM Review, 9(3):608, 1967.
- [8] F Soleymani and V Hosseinabadi. New third-and sixth-order derivative-free techniques for nonlinear equations. *Journal of Mathematics Research*, 3(2):107, 2011.
- [9] Changbum Chun and Beny Neta. A new sixth-order scheme for nonlinear equations. Applied mathematics letters, 25(2):185–189, 2012.
- [10] Obadah Said Solaiman and Ishak Hashim. Two new efficient sixth order iterative methods for solving nonlinear equations. *Journal of King Saud University-Science*, 31(4):701–705, 2019.
- [11] Mona Narang, Saurabh Bhatia, and Vinay Kanwar. New two-parameter chebyshev-halley-like family of fourth and sixth-order methods for systems of nonlinear equations. Applied Mathematics and Computation, 275:394–403, 2016.
- [12] Tahereh Eftekhari. A new sixth-order steffensen-type iterative method for solving nonlinear equations. International Journal of Analysis, 2014, 2014.
- [13] A.M. Ostrowski. Solutions of equations and system of equations. Academic Press, 1960.
- [14] Alicia Cordero and Juan R Torregrosa. Variants of newton's method using fifth-order quadrature formulas. *Applied Mathematics and Computation*, 190(1):686–698, 2007.

附录

附录 1

对化简 Halley 方法二阶导数的代码

```
import sympy as sp
   fx = sp.symbols('fx')
                             # f(x)
3
   dx = sp.symbols('dx')
                             # f'(x)
   x = sp.symbols('x')
                               # x
   dy = sp.symbols('dy')
                               # f'(y)
   y = x - fx / dx
   ddx = (dy - dx) / (y - x) # f''(x)
   halley = x - (2 * fx * dx) / (2 * dx * dx - fx * ddx)
10
11
  print(sp.simplify(halley - x))
```

附录 2

对收敛性证明的代码

```
# %%
   import sympy as sp
   # %%
   c2, c3, c4, c5, c6 = sp.symbols('c2 c3 c4 c5 c6')
   c1 = 1
   alpha = sp.symbols('alpha')
   e = sp.symbols('e')
   x = alpha + e
   e1 = e
10
   e2 = e ** 2
11
   e3 = e ** 3
   e4 = e ** 4
   e5 = e ** 5
   e6 = e ** 6
15
16
   # %%
17
def taylor_1plus_theta(the, order):
```

```
inv = 1 / (1 + the)
19
        return inv.series(e, 0, order)
20
21
    def getO(the, order):
22
        return taylor_1plus_theta(the, order) - taylor_1plus_theta(the, order).removeO()
23
^{24}
    o3 = getO(e, 3)
25
    o4 = get0(e, 4)
26
    o6 = getO(e, 6)
    o7 = get0(e, 7)
28
    o6, o7
29
30
    # %%
31
    def df(err, expand=False):
        eee = lambda x: x
33
        if expand:
34
            eee = sp.expand
35
        return eee(1 + 2*c2*err + 3*c3*err**2 + 4*c4*err**3 + 5*c5*err**4 + 6*c6*err**5 + o6)
36
37
38
    def f(err):
39
        return c1*err + c2*err**2 + c3*err**3 + c4*err**4 + c5*err**5 + c6*err**6 + o7
40
41
42
    def taylor_inv(fenmu, order):
43
        inv = 1 / fenmu
44
        return 1 / inv.series(e, 0, order)
45
46
    def texify(algo):
47
        sp.print_latex(sp.collect(sp.expand(algo), e))
48
        print()
49
50
    # %%
51
    fx = f(e)
52
    fx
53
    # %%
55
    dx = df(e)
```

```
dx
57
58
    # %%
59
    print(" 公式 18")
60
    texify(1 / taylor_inv(df(e), 6))
62
    fx_div_dx = sp.collect(sp.expand(fx / taylor_inv(df(e), 6)), e)
    print(" 公式 19")
    texify(fx_div_dx)
66
67
    # %%
    y = x - fx_div_dx
    print(" 公式 20")
    texify(y)
72
    # %%
73
    dy = df(y - alpha, True)
75
    # %%
76
    print(" 公式 21")
    texify(sp.collect(dy, e))
79
80
    z = x - 2 * fx / taylor_inv(dx + dy, 6)
81
    print(" 公式 24")
    texify(sp.collect(sp.expand(z), e))
    # %%
    fz = f(z-alpha)
86
    dz = df(z-alpha)
    print(" 公式 25")
88
    texify(sp.expand(z - fz / taylor_inv(dz, 6)))
89
    # %%
   fy = f(y - alpha)
   newx = z - 2 * fz / taylor_inv(dz + dy, 6)
   print(" 公式 27")
```

```
texify(sp.collect(sp.expand(newx), e))
96
97
    附录 3
        实验代码
    # %%
1
    import numpy as np
    import matplotlib.pyplot as plt
    import seaborn as sns
    import sympy as sym
    import decimal
    from decimal import Decimal
    plt.rcParams['text.usetex'] = False
    plt.rcParams["font.family"] = "Arial Unicode MS"
11
    # %%
12
    mdl = np
13
    roots = [
14
        1.6319808055660636,
15
        0.2575302854398608,
16
        -1.207647827130919,
17
        1.4044916482153411,
18
        1.1525907367571583,
19
        -1,
20
        1,
21
22
    1
23
    funcs = [
        lambda x: x**3 + 4*x**2 - 15,
25
        lambda x: x**2 - mdl.exp(x) - 3*x + 2,
26
        lambda x: x*mdl.exp(x**2) - mdl.sin(x)**2 + 3*mdl.cos(x) + 5,
27
        lambda x: mdl.sin(x)**2 - x**2 + 1,
28
        lambda x: mdl.log(x**2 + 7*x + 14) - x - 2,
29
        lambda x: (x - 4) * (x + 1)**4 / (mdl.exp(x)),
30
        lambda x: mdl.exp(x**2 + 11*x - 12) - 1,
31
        lambda x: mdl.arctan(mdl.exp(x + 3) - 1) * (x - 1)**2
```

```
]
33
34
    # %%
35
    fig, axes = plt.subplots(2, 4, figsize=(10, 5),
36
                sharex=True, sharey=True)
37
    sns.set_context('talk')
    xmin, xmax = -2, 2
39
    mdl = np
40
    for i in range(8):
41
        ax = axes.flatten()[i]
42
        x = np.linspace(xmin, xmax, 1000)
43
        y = funcs[i](x)
44
        sns.lineplot(x=x, y=y, ax=ax, linewidth=3, label=rf'$f_{i+1}$')
45
        ax.legend(loc='upper left')
46
        ax.hlines(0, xmin, xmax, color='k', linestyles='--')
47
        ax.set_xlim(xmin, xmax)
48
        ax.set_ylim(-5, 5)
49
        # ax.xaxis.set_minor_locator(plt.MultipleLocator(0.5))
50
        ax.xaxis.set_major_locator(plt.MultipleLocator(1))
51
52
    plt.subplots_adjust(left=0.05, right=0.99, bottom=0.07, top=0.98)
    plt.savefig('../figure/eval-func.pdf')
55
    # %%
56
    mdl = sym
57
    setattr(sym, 'arctan', sym.atan)
58
    func_d1 = []
59
    func_d2 = []
    for f in funcs:
        x = sym.symbols('x')
62
        symf = f(x)
63
        func_d1.append(sym.diff(symf, x, 1))
64
        func_d2.append(sym.diff(symf, x, 2))
65
    # for f in func_d1:
          print(f"lambda x: {f},")
   # print()
69
   # for f in func_d2:
```

```
print(f"lambda x: \{f\},")
 71
 72
           # %%
 73
          mdl = np
 74
          func_d1 = [
 75
                     lambda x: 3*x**2 + 8*x,
 76
                     lambda x: 2*x - np.exp(x) - 3,
                     lambda x: 2*x**2*np.exp(x**2) + np.exp(x**2) - 2*np.sin(x)*np.cos(x) - 3*np.sin(x),
 78
                     lambda x: -2*x + 2*np.sin(x)*np.cos(x),
 79
                     lambda x: (2*x + 7)/(x**2 + 7*x + 14) - 1,
 80
                     lambda x: -(x - 4)*(x + 1)**4*np.exp(-x) + 4*(x - 4)*(x + 1)**3*np.exp(-x) + (x + 1)**4*np.exp(-x),
 81
                     lambda x: (2*x + 11)*np.exp(x**2 + 11*x - 12),
 82
                     lambda x: (x - 1)**2*np.exp(x + 3)/((np.exp(x + 3) - 1)**2 + 1) + (2*x - 2)*np.atan(np.exp(x + 3) - 1)
          ]
 84
 85
           func_d2 = [
 86
                     lambda x: 2*(3*x + 4),
 87
                     lambda x: 2 - np.exp(x),
 88
                     89
                     lambda x: 2*(-np.sin(x)**2 + np.cos(x)**2 - 1),
                     lambda x: (-(2*x + 7)**2/(x**2 + 7*x + 14) + 2)/(x**2 + 7*x + 14),
                     lambda x: (x + 1)**2*(20*x + (x - 4)*(x + 1)**2 - 8*(x - 4)*(x + 1) - 2*(x + 1)**2 - 40)*np.exp(-x),
 92
                     lambda x: ((2*x + 11)**2 + 2)*np.exp(x**2 + 11*x - 12),
 93
                     lambda x: (x - 1)**2*(np.exp(x + 3) - 2*(np.exp(x + 3) - 1)*np.exp(2*x + 6)/((np.exp(x + 3) - 1)**2 + 1))/((np.exp(x + 3) - 1)**2 + 1)/((np.exp(x + 3) - 1)/((np.exp(
 94
          1
 95
 96
           # %%
 97
          np.seterr(all='raise')
           def newton(f, fd, _, x):
100
                     return x - f(x) / fd(x)
101
102
103
           def halley(f, fd1, fd2, x):
104
                     return x - (2 * f(x) * fd1(x)) / (2 * fd1(x)**2 - f(x)*fd2(x))
105
106
107
          def neta(f, fd, _, x):
108
```

```
w = x - f(x) / fd(x)
109
         z = w - f(x) / fd(x) * (f(x) - f(w)/2) / (f(x) - 5*f(w)/2)
110
         xn = z - f(z) / fd(x) * (f(x) - f(w)) / (f(x) - 3*f(w))
111
         return xn
112
113
    def grau(f, fd, _, x):
115
         y = x - f(x) / fd(x)
116
         z = y - (y - x) / (2*f(y) - f(x)) * f(y)
117
         xn = z - (y - x) / (2*f(y) - f(x)) * f(z)
118
         return xn
119
120
121
    def linz(f, fd, _, x):
122
         y = x - f(x) / fd(x)
123
         z = x - 2 * f(x) / (fd(x) + fd(y))
124
         xn = z - f(z) / fd(z)
125
         return xn
126
127
128
    name_map = {
129
         newton.__name__: '牛顿',
130
131
         halley.__name__: 'Halley',
         neta.__name__: 'NM',
132
         grau.__name__: 'GM',
133
         linz.__name__: '\\textbf{本文}',
134
    }
135
136
137
    def coc(x0, x1, x2, x3):
138
         x0 = np.float128(x0)
139
         x1 = np.float128(x1)
140
         x2 = np.float128(x2)
141
         x3 = np.float128(x3)
142
         return np.log((x3 - x2) / (x2 - x1)) / np.log((x2 - x1) / (x1 - x0))
143
144
145
    def compute(f, fd1, fd2, x, callback, root, tol, maxiter=100):
146
```

```
x0, x1, x2, x3 = x, x, x
147
         for i in range(1, maxiter):
148
             try:
149
                 tmp = x3
150
                 x3 = callback(f, fd1, fd2, x3)
151
                 x0 = x1
                 x1 = x2
153
                 x2 = tmp
154
                  if abs(x3 - x2) + abs(f(x3)) < tol:
155
                      return i, (x0, x1, x2, x3)
156
             except FloatingPointError as e:
157
                 print(x3, x2, e, '\n')
158
                 return -i, (x0, x1, x2, x3)
159
         return -1, (x0, x1, x2, x3)
160
161
     # %%
162
    mdl = np
163
    np.seterr('ignore')
164
     init_vals = [
165
         [1, 2],
166
         [0, 1],
         [-2, -1],
168
         [1, 2],
169
         [1, 2],
170
         [-1.5, -0.5],
171
         [0.5, 1.5],
172
         [0.5, 1.5]
173
    ]
174
     # for i in [0]:# range(len(funcs)):
175
     # print('\multirow{10}{*}{\$f_'} + str(i) + '\$}', end='')
176
     for initx_idx in range(2):
177
         for method in [newton, halley, neta, grau, linz]:
178
             for i in [3, 7]:
179
                 inix = init_vals[i][initx_idx]
180
                  iters, xs = compute(funcs[i], func_d1[i], func_d2[i], inix, method, roots[i], 1e-12, 1000)
181
                  err = float(Decimal("{:.3g}".format(xs[3] - roots[i])))
                  if abs(err) < 1e-16:</pre>
183
                      err = '\$< 10^{-16}\$'
184
```

```
else:
185
                      err = f'${sym.latex(err)}$'
186
                 if i > 3:
187
                      print('&', end='\t')
188
                  if method == newton:
189
                      print(' ', '\multirow{5}{*}{{' + str(inix) + '}', name_map[method.__name__], iters, err, sep=' {
191
                  else:
                      print(' ', ' ', name_map[method.__name__], iters, err, sep=' & \t', end='')
192
             print('\\\\n')
193
             if method == linz:
194
                 print('\\cline{2-5}\\cline{7-10}')
195
    print('\\hline')
196
197
     # %%
198
    mdl = np
199
    np.seterr(all='raise')
200
     init_vals = [
201
         [1, 2],
202
         [0, 1],
203
         [-2, -1],
204
         [1, 2],
         [1, 2],
206
         [-1.5, -0.5],
207
         [0.5, 1.5],
208
         [0.5, 1.5]
209
    ]
210
    for i in [7]:
211
         for inix in init_vals[i]:
212
             for method in [newton, halley, neta, grau, linz]:
213
                  # print(inix, method.__name__)
214
                 iters, xs = compute(funcs[i], func_d1[i], func_d2[i], inix, method, roots[i], 1e-12, 1000)
215
                 err = "{:.3g}".format(xs[3] - roots[i])
216
                 print(i, inix, name_map[method.__name__], iters, err, sep='\t')
217
218
219
```