参赛队员姓名：殷泽锐

中学：南京师范大学附属中学

省份：江苏省

国家/地区：中国

指导教师姓名：邓卫兵

指导老师单位：南京大学

论文题目：

Chebyshev-Halley 迭代法探究

Chebyshev-Halley 迭代法探究

殷泽锐

摘要

本文介绍了 Halley 迭代法的一个新的推导方法, 以及 Halley 迭代法跟经典迭代方法的比较

关键词: 数值求解，收敛阶，方法效率

介绍

如今，数学里的最优化问题无疑是最和现实世界相关联，以及最先运用新的数学成果的主流研究方向. 在确立了目标函数之后，以合适的工具求得目标函数的解就是一个重要的问题. 现在的迭代方法，大多建立在一套复杂的迭代公式之上，有时候是用空间换了时间.

我们的方法基于一个很简单的思路，就是对目标函数进行改变，使其变为一个与原函数拥有相同解的函数，且其相比原函数拥有更好的收敛性质，比如收敛域更大，使用 Newton 法（或者其他方法）拥有更高的收敛阶. 其最终形式与 Chebyshev 迭代法的形式相同，所以其成为了一个 Halley-Chebyshev 的推导方法. 我们将它在收敛阶，收敛格式，运行速度，以及一些特殊情形下的表现（比如有重根的情形）与一些经典方法做比较，进行了比较深入的研究.

1 Chebyshev-Halley 迭代法的推导

Newton 迭代法的收敛阶是 2，Halley 迭代法的收敛阶是 3，要想将收敛阶提升，需要在运用 Newton 法的原函数上进行改变.

设原函数方程为 $f(x)=0$，Newton 迭代法公式为 $x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}$，将误差设为 e_n，精确解设为 x^*，则 $x_n=x^*+e_n$，注意到 $f(x^*)=0$，且根

据泰勒展开, 有

$$f(x^*+e^n)=f(x^*)+\frac{e_n}{1!}f'(x^*)+\frac{e_n^2}{2!}f''(x^*)+...=\sum_{i=0}^{\infty}\frac{e_n^i}{i!}f^{(i)}(x^*)$$

所以有以下推导过程

$$\begin{aligned}
x_{n+1}=&x_n-\frac{f(x_n)}{f'(x_n)}\\
e_{n+1}=&e_n-\frac{f(x^*+e_n)}{f'(x^*+e_n)}\\
=&\frac{(e_nf'(x^*)+\frac{e_n^2}{1!}f''(x^*)+...)-(f(x^*)+\frac{e_n}{1!}f'(x^*)+\frac{e_n^2}{2!}f''(x^*)+...)}{f'(x_n)}\\
=&\frac{e_n^2f''(x^*)+o(e_n^2)}{2f'(x_n)}
\end{aligned}$$

可以得出 Newton 法是二阶收敛，现引入一个辅助函数 $g(x)$, 对 $f(x)g(x)$ 运用求根公式，易知当 $g(x)\neq 0$ 时，$f(x)g(x)=0$ 与原函数方程拥有相同的根。对 $f(x)g(x)$ 运用 Newton 迭代法

$$\begin{aligned}
x_{n+1}&=x_n-\frac{f(x_n)g(x_n)}{f'(x_n)g(x_n)+f(x_n)g'(x_n)}\\
e_{n+1}&=e_n-\frac{f(x_n)g(x_n)}{f'(x_n)g(x_n)+f(x_n)g'(x_n)}\\
&=\frac{e_n(f'(x_n)g(x_n)+f(x_n)g'(x_n))-f(x_n)g(x_n)}{f'(x_n)g(x_n)+f(x_n)g'(x_n)}\\
&=\frac{1}{f'(x_n)g(x_n)+f(x_n)g'(x_n)}\cdot\Big[e_n(f'(x^*)+e_nf''(x^*)+\frac{e_n^2}{2!}f'''(x^*))g(x_n)+e_n(f(x^*)+e_n\\
&\quad f'(x^*)+\frac{e_n^2}{2!}f''(x^*))g'(x_n)-(f(x^*)+e_nf'(x^*)+\frac{e_n^2}{2!}f''(x_n)+\frac{e_n^3}{3!}f'''(x^*))g(x_n)+o(e_n^3)\Big]\\
&=\frac{1}{f'(x_n)g(x_n)+f(x_n)g'(x_n)}\cdot\Big[(e_n^2f''(x^*)+\frac{e_n^3}{2}f'''(x^*))g(x_n)+(e_n^2\\
&\quad f'(x^*)+\frac{e_n^3}{2}f''(x^*))g'(x_n)-(\frac{e_n^2}{2}f''(x^*)+\frac{e_n^3}{3!}f'''(x^*))g(x^*)+o(e_n^3)\Big]\\
&=\frac{e_n^2(\frac{1}{2}f''(x^*)g(x_n)+f'(x^*)g'(x_n))+e_n^3(\frac{1}{6}f'''(x^*)g(x_n)+f''(x*)g'(x_n))+o(e_n^3)}{f(x_n)g'(x_n)+f'(x_n)g(x_n)}
\end{aligned}$$

此时我们发现, 当 e_n 的二次项系数为 0, 或为 $ke_n+o(e_n)$ 的形式时, $f(x)g(x)$ 三阶收敛.

$$\frac{1}{2}f''(x^*)g(x_n)+f'(x^*)g'(x_n)=ke_n+o(e_n)$$

函数 $g(x)$ 需要满足在任何情况下上式成立，由此可知，一个令

$$\frac{1}{2}f''(x)g(x)+f'(x)g'(x)=0$$

恒成立的函数 $g(x)$ 是符合要求的，于是

$$\frac{f''(x)}{2f'(x)}=-\frac{g'(x)}{g(x)}$$

两边积分，整理得

$$g(x)=\frac{1}{\sqrt{f'(x)}}$$

令 $F(x)=\frac{f(x)}{\sqrt{f'(x)}}$,Halley 迭代法可表示为

$$x_{n+1}=x_n-\frac{F(x_n)}{F'(x_n)}$$

2　与其他经典算法的比较

我们下面来比较一下 Chebyshev-Halley 方法与经典 Newton 法、Steffensen 迭代法、用 Aitken 加速的 Newton 迭代法. 它们的阶数如下（关于 Aitken 加速的 Newton 法是 3 阶格式的推导，可以参见附录）：

方法	格式	收敛阶
经典 Newton 法	$x_{n+1}=x_n-\dfrac{f(x_n)}{f'(x_n)}$	2
Steffensen 法	$y_n-x_n=f(x_n)$ $z_n-y_n=f(y_n)=f(x_n+f(x_n))$, $x_{n+1}=x_n-\dfrac{[f(x_n)]^2}{f(x_n+f(x_n))-f(x_n)}$	2
Aitken 加速的 Newton 法	$g(x)=x-\dfrac{f(x)}{f'(x)}$ $y_n=g(x_n)$, $z_n=g(y_n)$, $x_{n+1}=x_n-\dfrac{(y_n-x_n)^2}{z_n-2y_n+x_n}$	3
Chebyshev-Halley 法	$x_{n+1}=x_n-\dfrac{f(x_n)f'(x_n)}{[f'(x_n)]^2-\frac{1}{2}f(x_n)f''(x_n)}$	3

2.1 问题设置

考虑用上述四种方法数值求解函数

$$f(x) = x^5 + x^4 + 4x^3 + 2x^2 + x - 2$$

的根, 其中 $f(x)$ 的根的精确值为 $x^* = 0.554702\cdots$. 对于每一种方法，我们取初始值为 $x_0 = 0.25$.

这里需要说明的是，matlab 中浮点数的最大精度为 15 位小数，三阶收敛的迭代方法在 4 次迭代之后就会超过机器精度，二阶收敛的迭代方法在 5-6 次后就会超过机器精度.

根据输出的结果, 我们可以观察到四种方法的收敛阶. 假设 $e_n = x_n - x^*$, 那么当一个数值算法是 k 阶收敛时,

$$e_{n+1} \approx Ce_n^k.$$

我们取对数,

$$\log e_{n+1} \approx k \log e_n + C,$$

则只需要画出横坐标为 $\log e_n$, 纵坐标为 $\log e_{n+1}$ 的散点图的拟合直线, 观察这条直线的斜率, 即可得到收敛阶. 这个数值算法的收敛阶就是这条直线的斜率. 得到的结果如图 1~4

迭代步数，运行时间（运行出 10e5 次结果），收敛阶（拟合直线斜率）结果如下:

方法	迭代步数	迭代时间	收敛阶
经典 Newton 法	6	18.470s	1.969048
Steffensen 法	5	14.580s	2.006157
Aitken 加速的 Newton 法	4	8.554s	2.956898
Chebyshev-Halley 法	4	7.553s	3.020506

2.2 结论

通过数据比对, 我们发现对于目标函数 $f(x) = x^5+x^4+4x^3+2x^2+x-2$ 的求根，Halley 迭代法在效率上占优，而且其收敛速度比其他方法稍快。但是注意到该目标函数是多项式函数，比较方便进行求导，Steffensen 方法是不需要求导的，Newtoin 法和 Aitken 加速的牛顿法只需要求一阶导数，而 Halley 迭代法需要求二阶导数，对不方便求导的函数运用 Halley 迭代法可能是一个问题.

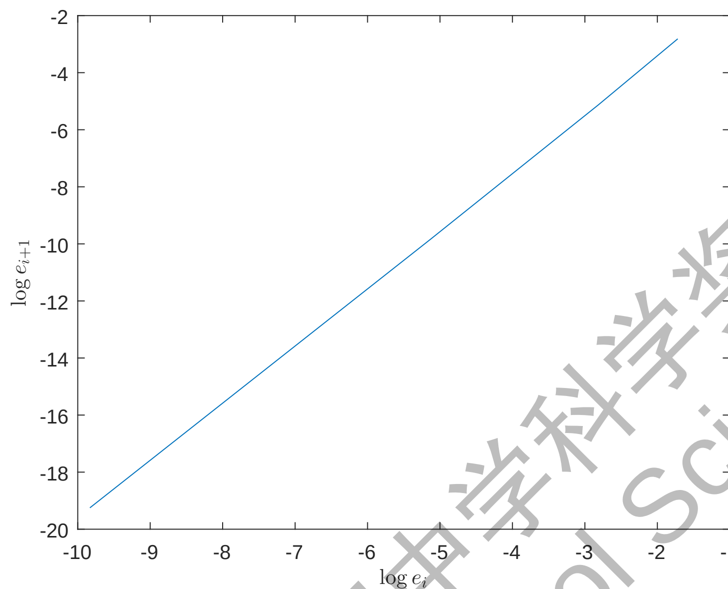

图 1: Newton 迭代法的 $\log e_n \sim \log e_{n+1}$ 曲线

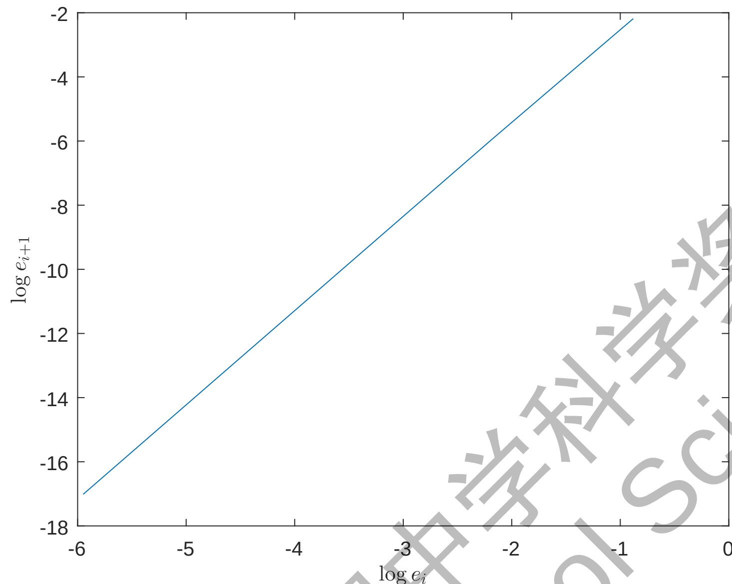

图 2: Aitken 加速的牛顿迭代法的 $\log e_n \sim \log e_{n+1}$ 曲线

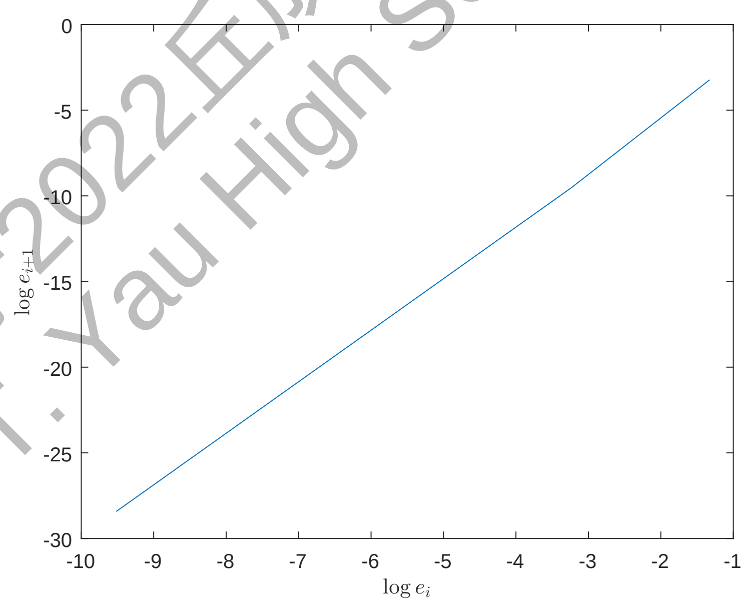

图 3: Halley 迭代法的 $\log e_n \sim \log e_{n+1}$ 曲线

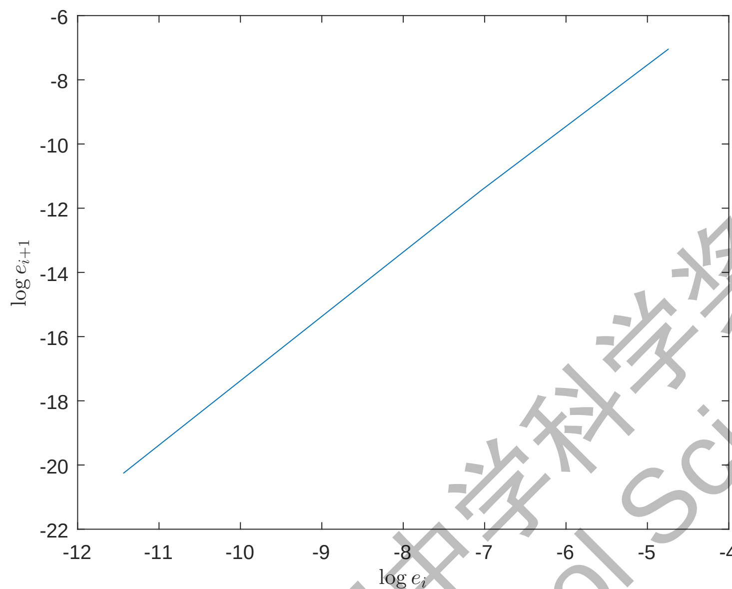

图 4: Steffensen 法的 $\log e_n \sim \log e_{n+1}$ 曲线

3 Chebyshev-Halley 方法简单的应用

在计算机中我们需要用数值方法来求平方根. 对于函数 $f(x) = x^2 - R$, 求 $f(x)$ 的根的 Chebyshev-Halley 法为

$$\begin{aligned} x_{n+1} &= x_n - \frac{f(x_n)f'(x_n)}{[f'(x_n)]^2 - (f(x_n)f''(x_n))/2} \\ &= x_n - \frac{(x_n^2 - R) \cdot 2x_n}{(2x_n)^2 - (x_n^2 - R)} \\ &= x_n - \frac{2x_n^3 - 2Rx_n}{3x_n^2 + R} \\ &= \frac{x_n(x_n^2 + 3R)}{3x_n^2 + R}. \end{aligned}$$

因此我们导出了求正实数 R 的平方根 $\sqrt{R}$ 的三阶方法:

$$x_{n+1} = \frac{x_n(x_n^2 + 3R)}{3x_n^2 + R}. \tag{1}$$

在计算机中, 如果我们需要求 R 的平方根, 只需要应用迭代格式 (1), 然后迭代几步到达机器精度即可.

4 附录: Aitken 加速的 Newton 迭代法是三阶格式的证明

我们简要推导把 Aitken 加速法应用于 Newton 迭代法的收敛阶.

设 $e_n - x^* = \varepsilon_n$, $y_n - x^* = \varepsilon_n$, $z_n - x^* = \delta_n$, 则

$$0 = f(x^*) = f(x_n - e_n) = f(x_n) - e_n f'(x_n) + \frac{1}{2}e_n^2 f''(x_n) - \frac{1}{6}e_n^3 f'''(\xi_n),$$

其中 ξ_n 介于 x_n 与 x^*, 因此

$$e_n f'(x_n) - f(x_n) = \frac{1}{2}e_n^2 f''(x_n) - \frac{1}{6}e_n^3 f'''(\xi_n),$$

代入迭代格式, 可知 $\varepsilon_n = e_n - \dfrac{f(x_n)}{f'(x_n)}$, 则

$$\varepsilon_n = \frac{1}{2}e_n^2 \frac{f''(x_n)}{f'(x_n)} - \frac{1}{6}e_n^3 \frac{f'''(\xi_n)}{f'(x_n)} \approx \frac{1}{2}e_n^2 \frac{f''(x^*)}{f'(x^*)} - \frac{1}{6}e_n^3 \frac{f'''(x^*)}{f'(x^*)}.$$

同理可得

$$\delta_n \approx \frac{1}{2}\varepsilon_n^2 \frac{f''(x^*)}{f'(x^*)} - \frac{1}{6}\varepsilon_n^3 \frac{f'''(x^*)}{f'(x^*)}.$$

我们记 $\dfrac{f''(x^*)}{2f'(x^*)} = C$, $-\dfrac{f'''(x^*)}{6f'(x^*)} = D$. 于是

$$\begin{aligned}
e_{n+1} &\approx e_n - \frac{[(y_n - x^*) - (x_n - x^*)]^2}{(z_n - x^*) - 2(y_n - x^*) + (x_n - x^*)} \\
&\approx e_n - \frac{(\varepsilon_n - e_n)^2}{\delta_n - 2\varepsilon_n + e_n} \\
&\approx \frac{e_n(\delta_n - 2\varepsilon_n + e_n) - (\varepsilon_n - e_n)^2}{\delta_n - 2\varepsilon_n + e_n} = \frac{e_n\delta_n - \varepsilon_n^2}{\delta_n - 2\varepsilon_n + e_n} \\
&\approx \frac{e_n[C(Ce_n^2 + De_n^3)^2 + D(Ce_n^2 + De_n^3)^3] - (Ce_n^2 + De_n^3)^2}{C(Ce_n^2 + De_n^3)^2 + D(Ce_n^2 + De_n^3)^3 - 2(Ce_n^2 + De_n^3) + e_n} \\
&= \frac{-C^2 e_n^4 + o(e_n^4)}{e_n + o(e_n)} \\
&\approx -C^2 e_n^3.
\end{aligned}$$

所以用 Aitken 加速的 Newton 法的收敛阶是 3.

5 参考文献

[1] 林成森, 数值计算方法 (上册), 第 2 版, 科学出版社, 2005.

[2] David R. Kincaid, E. Ward Cheney, Numerical analysis, 1991, Brooks-Cole edition.

[3] 武鹏, 解非线性方程的高阶迭代算法及其收敛性分析, 浙江大学理学院博士学位论文, 2008.

6 致谢

我受到了《从 $\sqrt{2}$ 谈起》这本书的启发，研究起了数值求根方法，然后接触到了 Newton 迭代法. 在研究如何提高 Newton 迭代法收敛阶的时候，我想到了对原函数进行改进.

经过一定的研究, 我发现改进的方法有很多, 其中 Chebyshev-Halley 迭代法引起了我的注意. 我发现可以用一种新的方式推导 Chebyshev-Halley 迭代法. 这篇论文在我的指导老师和学长的帮助下完成，主要是提供了一些文章写作思路和代码上的帮助.