方程求根

方程求根与二分法

方程

在本章中主要讨论的方程是 **单变量非线性方程** ，即

$$f(x)=0$$

其中 $x \in \mathbb{R}, f(x) \in C[a,b]$

这样的方程包括多项式方程与超越方程

方程 $f(x)=0$ 的根 x^* 满足 $f(x^*)=0$ ，若方程可分解为

$$f(x)=(x-x^*)^m g(x)$$

其中 m 为正整数且 $g(x^*) \neq 0$ ，则称 x^* 为方程的 m 重根

若 x^* 是 $f(x)$ 的 m 重根，且 $g(x)$ 充分光滑，则有

$$f\left(x^*\right)=f'\left(x^*\right)=\cdots=f^{(m-1)}\left(x^*\right)=0, f^{(m)}\left(x^*\right) \neq 0$$

且 n 次代数方程在复数域只有 n 个根

求根

若 $f(x) \in C[a,b]$ 且 $f(a)f(b)<0$ ，根据连续函数的介值定理可知方程在区间 (a,b) 至少有一个实根，称 $[a,b]$ 为方程的**有根区间**，通常可使用描图法或逐次搜索法求得方程的有根区间

描图法：可通过作草图以 $f(x)$ 和横轴的交点位置确定有根区间，而 $f(x)$ 比较复杂时也可以将其化为等价方程 $\varphi(x)=\psi(x)$ 并求 $\varphi(x), \psi(x)$ 的交点

逐步搜索法：从区间 $[a,b]$ 的左端点 a 出发，以步长 h 向右搜索，直至

$$f(a+jh)f(a+(j+1)h)<0 \quad (j=0,1,2,\cdots)$$

则区间 $[a+jh, a+(j+1)h]$ 即为有根区间

二分法

设 $[a,b]$ 为 $f(x)$ 的有根区间，取中点

$$x_0=\frac{1}{2}(a+b)$$

若 $f(x_0)=0$ 则 x_0 是方程的根，否则判定 x^* 在 x_0 的左侧或右侧

- 若 $f(a)f(x_0)<0$ ，则 $x^*\in(a,x_0)$ ，令 $a_1=a,b_1=x_0$
- 若 $f(x_0)f(b)<0$ ，则 $x^*\in(x_0,b)$ ，令 $a_1=x_0,b_1=b$

则新的有根区间为 (a_1,b_1) ，且长度只有原先有根区间的一半，如此反复可以得到一系列有根区间套

$$[a,b]\supset[a_1,b_1]\supset\cdots\supset[a_n,b_n]\supset\cdots$$

且易得

$$b_n-a_n=\frac{1}{2^n}(b-a)$$

当 $n\to\infty$ 时区间收缩为一点，即所求的根。在实际应用中，区间的长度小于一个给定的精度 ε 或 $f(x_n)$ 小于一个给定的精度 δ 即可停止。

若使用区间 $[a_n,b_n]$ 的中点

$$x_n=\frac{1}{2}(a_n+b_n)$$

作为 x^* 的近似值，则误差估计为

$$|x^*-x_n|\leqslant\frac{1}{2}(b_n-a_n)=\frac{1}{2^{n+1}}(b-a)$$

只要 n 足够大，误差就可以足够小。对于给定精度 ε ，二分法所需步骤大约为

$$\left\lceil\log_2\frac{b-a}{\varepsilon}\right\rceil$$

二分法的优点在于简单易用，且对 $f(x)$ 的要求不高，只要连续即可保证收敛，但是二分法的收敛速度慢，并且无法求偶重根和复根

迭代法及其收敛性

不动点迭代法

将方程 $f(x)=0$ 改写为等价的方程

$$x=\varphi(x)$$

则满足 $f(x^*)=0$ 的 x^* 满足 $x^*=\varphi(x^*)$，称 x^* 为 $\varphi(x)$ 的一个不动点，求 $f(x)$ 的零点等价于求 $\varphi(x)$ 的不动点

不动点迭代法的思路是，首先选一个初值 x_0 代入 $\varphi(x)$，得到

$$x_1=\varphi(x_0)$$

然后继续计算，得到

$$\begin{gathered}x_2=\varphi(x_1)\\x_3=\varphi(x_2)\\\cdots\\x_{k+1}=\varphi(x_k)\end{gathered}$$

若 $\{x_k\}_{k=0}^{\infty}$ 收敛，即存在 x^* 使得

$$\lim_{k\to\infty}x_k=x^*$$

由于 $\varphi(x)$ 连续，故

$$\lim_{\kappa\to\infty}x_{k+1}=\lim_{k\to\infty}\varphi\left(x_k\right)$$

可得 $x^*=\varphi(x^*)$，即求得不动点

若对于任何 $x_0\in[a,b]$，由

$$x_{k+1}=\varphi\left(x_k\right)\quad(k=0,1,2,\cdots)$$

得到的序列 $\{x_k\}$ 均有极限

$$\lim_{k\to\infty}x_k=x^*$$

则称上述迭代方程收敛，且 x^* 为 $\varphi(x)$ 的不动点

不动点的存在性与迭代法的收敛性

首先考察 $\varphi(x)$ 在 $[a,b]$ 上不动点的存在唯一性

> **定理 1** 设 $\varphi(x)\in C[a,b]$ 满足以下两个条件
>
> - 对任意 $x\in[a,b]$ 有 $a\leqslant\varphi(x)\leqslant b$

> - 存在正数 $L<1$ ，使得对于任意 $x,y\in[a,b]$ 有
>
> $$|\varphi(x)-\varphi(y)|\leqslant L|x-y|$$
>
> 则 $\varphi(x)$ 在 $[a,b]$ 上有唯一的不动点

由不动点的存在唯一性，可以得到一个迭代法收敛的充分条件

> **定理 2** 设 $\varphi(x)\in C[a,b]$ 满足定理 1 的两个条件，则对任意 $x_0\in[a,b]$ ，由迭代方程得到的序列 $\{x_k\}$ 均收敛到不动点 x^* ，且有误差估计式
>
> $$|x^*-x_k|\leqslant\frac{L^k}{1-L}|x_1-x_0|$$
> $$|x^*-x_k|\leqslant\frac{L}{1-L}|x_{k+1}-x_k|$$

原则上根据第一个误差估计式控制精度，但由于含有未知信息 L 而不便于实际应用，故一般用第二个误差估计式，只要相邻两次计算结果的偏差足够小即可保证近似值 x_k 有较好的精度，即若给定精度 ε 要求 $|x^*-x_k|<\varepsilon$ ，只需

$$\frac{L}{1-L}|x_k-x_{k-1}|<\varepsilon$$
$$|x_k-x_{k-1}|<\frac{1-L}{L}\varepsilon\approx\varepsilon$$

迭代即可终止

且对于定理中的条件

> - 存在正数 $L<1$ ，使得对于任意 $x,y\in[a,b]$ 有
>
> $$|\varphi(x)-\varphi(y)|\leqslant L|x-y|$$

可改为导数，即若 $\varphi(x)\in C[a,b]$ 且对任意 $x\in[a,b]$ 有

$$|\varphi'(x)|\leqslant L<1$$

则根据微分中值定理有

$$|\varphi(x)-\varphi(y)|=|\varphi'(\xi)(x-y)|\leqslant L|(x-y)|,\quad\xi\in(a,b)$$

条件同样成立

局部收敛性与收敛阶

上述给出迭代序列 $\{x_k\}$ 在区间 $[a,b]$ 上的收敛性通常称为**全局收敛性**，而在不易检验定理条件的情况下，可只考察不动点 x^* 附近的收敛性，称为**局部收敛性**

> **定义 1** 设 $\varphi(x)$ 有不动点 x^* ，若存在某个 x^* 的邻域 $R: |x-x^*| \leqslant \delta$ ，对任意 $x_0 \in R$ ，迭代公式产生的序列 $\{x_k\}$ 收敛，且收敛到 x^* ，则称迭代法**局部收敛**

可类比得出局部收敛的充分条件

> **定理 3** 设 x^* 为 $\varphi(x)$ 的不动点，$\varphi'(x)$ 在 x^* 的某个邻域连续，且 $|\varphi'(x)| < 1$ ，则迭代法局部收敛

同时，迭代法的收敛快慢根据迭代方程的选取也有不同，为了衡量迭代法的收敛速度，给出 p 阶收敛的定义

> **定义 2** 设迭代过程 $x_{k+1} = \varphi(x_k)$ 收敛于方程 $x = \varphi(x)$ 的根 x^* ，若迭代误差 $e_k = x_k - x^*$ 在 $k \to \infty$ 时满足下列渐进关系式
>
> $$\frac{e_{k+1}}{e_k^p} \to C$$
>
> 其中 C 为常数且 $C \neq 0$ ，则称该迭代法是 p 阶收敛的

判定一个收敛法的收敛阶基于其导数

> **定理 4** 对于迭代过程 $x_{k+1} = \varphi(x_k)$ ，若 $\varphi^{(p)}(x)$ 在 x^* 附近连续且
>
> $$\varphi'\left(x^*\right) = \varphi''\left(x^*\right) = \cdots = \varphi^{(p-1)}\left(x^*\right) = 0, \varphi^{(p)}\left(x^*\right) \neq 0$$
>
> 则该迭代过程在 x^* 附近 p 阶收敛

迭代收敛的加速

埃特金加速收敛法

基本思想为设 x_0 是根 x^* 的某个近似值，使用迭代公式校正一次得

$$x_1 = \varphi\left(x_0\right)$$

而根据微分中值定理，有

$$x_1 - x^* = \varphi\left(x_0\right) - \varphi\left(x^*\right) = \varphi'(\xi)\left(x_0 - x^*\right)$$

其中 ξ 在 x_0 和 x^* 之间，而当 $\varphi'(x)$ 变化不大时，可近似认为

$$x_1 - x^* \approx L\left(x_0 - x^*\right)$$

将 x_1 再次校正，得到

$$x_2=\varphi\left(x_1\right)$$

由于

$$x_2-x^*\approx L\left(x_1-x^*\right)$$

联立消去未知量 L ，可得

$$\frac{x_1-x^*}{x_2-x^*}\approx\frac{x_0-x^*}{x_1-x^*}$$

即

$$x^*\approx\frac{x_0x_2-x_1^2}{x_2-2x_1+x_0}=x_0-\frac{\left(x_1-x_0\right)^2}{x_2-2x_1+x_0}$$

埃特金加速法即是基于上述思想，用上式右端作为 x^* 的新近似，记为 $\overline{x_1}$

一般是根据 x_k 计算出 x_{k+1},x_{k+2} ，然后

$$\begin{aligned}\bar{x}_{k+1}&=x_k-\frac{\left(x_{k+1}-x_k\right)^2}{x_{k+2}-2x_{k+1}+x_k}\\&=x_k-\frac{\left(\Delta x_k\right)^2}{\Delta^2x_k}\quad(k=0,1,\cdots)\end{aligned}$$

上式即为埃特金加速法，可以证明

$$\lim_{k\to\infty}\frac{\bar{x}_{k+1}-x^*}{x_k-x^*}=0$$

即序列 $\{\overline{x_k}\}$ 比 $\{x_k\}$ 收敛更快，事实上，若 $x_{k+1}=\varphi(x_k)$ 为线性收敛，埃特金加速法为平方收敛，若 $x_{k+1}=\varphi(x_k)$ 为 p 阶收敛，$\varphi(x)$ 的 p 阶导数连续，则埃特金法为 $2p-1$ 阶收敛

牛顿法

基本思想

牛顿法的思想即为将非线性函数线性化。设 $f(x)=0$ 有近似根 x_0 ，且在 x_0 附近 $f(x)$ 可用一阶泰勒多项式近似

$$f(x)\approx f\left(x_0\right)+f'\left(x_0\right)\left(x-x_0\right)$$

当 $f'(x_0)\neq 0$ ，即可用线性方程近似 $f(x)=0$

$$f\left(x_0\right)+f^{\prime}\left(x_0\right)\left(x-x_0\right)=0$$

解得

$$x=x_0-\frac{f\left(x_0\right)}{f^{\prime}\left(x_0\right)}$$

故可得到迭代公式

$$x_{k+1}=x_k-\frac{f\left(x_k\right)}{f^{\prime}\left(x_k\right)} \quad(k=0,1, \cdots)$$

上式即是牛顿迭代公式

收敛条件

设 $f(x) \in C^2[a, b]$ ，若

- $f(a) f(b)<0$
- 在整个 $[a, b]$ 上 $f^{\prime \prime}$ 不变号，并且 $f^{\prime}(x) \neq 0$
- 选取 $x_0 \in[a, b]$ 使得 $f\left(x_0\right) f^{\prime \prime}\left(x_0\right)>0$

则牛顿法产生的序列 $\left\{x_k\right\}$ 收敛到 $f(x)$ 在 $[a, b]$ 的唯一根

牛顿法的局部收敛性为，设 $f(x) \in C^2[a, b]$ ，若 x^* 为 $f(x)$ 在 $[a, b]$ 的根，且 $f^{\prime}\left(x^*\right) \neq 0$ ，则存在 x^* 的邻域 U ，使得任取初始值 $x_0 \in U$ ，牛顿法产生的序列 $\left\{x_k\right\}$ 收敛到 x^* ，并且满足

$$\lim _{k \leftarrow \infty} \frac{x_{k+1}-x^*}{\left(x_k-x^*\right)^2}=\frac{f^{\prime \prime}\left(x^*\right)}{2 f^{\prime}\left(x^*\right)}$$

由此可得当 x^* 为单根时，牛顿法在 x^* 附近是二阶收敛的

牛顿法的变体

牛顿法优点是收敛快，缺点是每步迭代时计算量大，而且初始值只有在根附近才能保证收敛。为了克服这两个缺点，通常可用下述方法

简化牛顿法

也称作平行弦法，迭代公式为

$$x_{k+1}=x_k-C f\left(x_k\right) \quad C \neq 0, k=0,1, \cdots$$

即迭代函数 $\varphi(x)=x-C f(x)$

若 $|\varphi'(x_k)| = |1 - Cf'(x)| < 1$ ，取 $0 < Cf'(x) < 2$ ，在 x^* 附近成立，则迭代法局部收敛

取 $C = \dfrac{1}{f'(x_0)}$ 称为简化牛顿法，计算量小但只有线性收敛

牛顿下山法

牛顿法收敛性依赖于初始值 x_0 的选取，为了防止迭代发散，对迭代过程再加一项要求

$$|f(x_{k+1})| < |f(x_k)|$$

即迭代过程具有单调性

将下山法和牛顿法结合，将本次计算的结果

$$\overline{x}_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

与上一步得到的近似值 x_k 作加权平均

$$x_{k+1} = \lambda \overline{x}_{k+1} + (1-\lambda) x_k$$

$\lambda (0 < \lambda \leqslant 1)$ 即为下山因子，迭代过程变为

$$x_{k+1} = x_k - \lambda \frac{f(x_k)}{f'(x_k)} (k = 0, 1, \cdots)$$

在计算时 λ 从 1 开始，若不满足则逐次折半，直至满足下山条件

重根

当 x^* 是 $f(x)$ 的 m 重根时，可表示为

$$f(x) = (x - x^*)^m g(x)$$

但是牛顿法解重根时收敛速度将大大减慢，只有线性收敛。有两种提高重根的收敛速度的方法

一是取如下迭代函数

$$\varphi(x) = x - m \frac{f(x)}{f'(x)}$$

在求 m 重根时有 2 阶收敛。有一种估算重数 m 的方法。令

$$\lambda_k = \frac{x_k - x_{k-1}}{x_{k-1} - x_{k-2}}$$

则

$$\lambda_k=\frac{e_k-e_{k-1}}{e_{k-1}-e_{k-2}}=\frac{e_k}{e_{k-1}}\cdot\frac{1-\frac{e_{k-1}}{e_k}}{1-\frac{e_{k-2}}{e_{k-1}}}$$

根据

$$\lim_{k\to\infty}\frac{e_{k+1}}{e_k}=1-\frac{1}{m}$$

有

$$\lim_{k\to\infty}\lambda_k=1-\frac{1}{m}=\frac{m-1}{m}$$

故可得根重数为

$$m\approx\frac{1}{1-\lambda_k}$$

二是将求重根的问题化为求单根的问题，对于函数

$$f(x)=\left(x-x^*\right)^m g(x), g\left(x^*\right)\neq 0$$

可以令

$$\mu(x)=\frac{f(x)}{f'(x)}=\frac{\left(x-x^*\right)g(x)}{mg(x)+\left(x-x^*\right)g'(x)}$$

化为求 $\mu(x)=0$ 的单根 x^* 的问题，对其使用牛顿法是 2 阶收敛的

迭代函数为

$$\varphi(x)=x-\frac{\mu(x)}{\mu'(x)}=x-\frac{f(x)f'(x)}{\left[f'(x)\right]^2-f(x)f''(x)}$$

于是得到迭代法为

$$x_{k+1}=x_k-\frac{f\left(x_k\right)f'\left(x_k\right)}{\left[f'\left(x_k\right)\right]^2-f\left(x_k\right)f''\left(x_k\right)}(k=0,1,\cdots)$$

弦截法与抛物线法

用牛顿法求方程 $f(x)=0$ 的根时，每步需要计算导数值 $f'(x_k)$ ，当函数较为复杂时计算导数值往往比较困难，可以利用已求函数值 $f(x_k), f(x_{k-1}), ...$ 来回避导数值的计算，这种方法是建立在插值的基础之上的

弦截法

又称为割线法，设 x_k, x_{k-1} 是 $f(x)=0$ 的近似根，利用求出的 $f(x_k), f(x_{k-1})$ 构造一次插值多项式 $p_1(x)$ ，并且用 $p_1(x)=0$ 的根作为新的近似根 x_{k+1} 。由于

$$p_1(x)=f\left(x_k\right)+\frac{f\left(x_k\right)-f\left(x_{k-1}\right)}{x_k-x_{k-1}}\left(x-x_k\right)$$

故带入牛顿法的方程，有

$$x_{k+1}=x_k-\frac{f\left(x_k\right)}{f\left(x_k\right)-f\left(x_{k-1}\right)}\left(x_k-x_{k-1}\right)$$

可以看作是用差商 $f[x_k, x_{k-1}]$ 代替了 x_k 处的导数值 $f'(x_k)$

切线法与牛顿法的不同之处在于用到了两个之前计算的值，故开始计算时需要提供两个开始值。

假设 $f(x)$ 在根 x^* 的邻域 $\Delta:|x-x^*|\leqslant\delta$ 具有二阶连续导数，且对任意 $x\in\Delta$ 有 $f'(x)\neq 0$ ，所取的开始值 $x_0, x_1\in\Delta$ ，当邻域充分小时，弦截法按阶

$$p=\frac{1+\sqrt{5}}{2}\approx 1.618$$

收敛到 x^*

这种方法用到之前两点的值，称为双点割线法，而将式中的 x_{k-1} 改为 x_0 则得到单点割线法，迭代公式为

$$x_{k+1}=x_k-\frac{f\left(x_0\right)}{f\left(x_k\right)-f\left(x_0\right)}\left(x_k-x_0\right)$$

抛物线法

设 x_k, x_{k-1}, x_{k-2} 是 $f(x)=0$ 的三个近似根，以三点构造二次插值多项式 $p_2(x)$ ，并适当选取一个零点作为新的近似根 x_{k+1} ，这种方法称为抛物线法，也称为密勒法。插值多项式为

$$\begin{aligned} p_2(x)= & f\left(x_k\right)+f\left[x_k, x_{k-1}\right]\left(x-x_k\right) \\ & +f\left[x_k, x_{k-1}, x_{k-2}\right]\left(x-x_k\right)\left(x-x_{k-1}\right) \end{aligned}$$

故零点为

$$x_{k+1}=x_k-\frac{2f\left(x_k\right)}{\omega \pm \sqrt{\omega^2-4f\left(x_k\right)f\left[x_k,x_{k-1},x_{k-2}\right]}}$$

其中 $\omega=f\left[x_k,x_{k-1}\right]+f\left[x_k,x_{k-1},x_{k-2}\right]\left(x_k-x_{k-1}\right)$ ，一般选取接近 x_k 的根作为 x_{k+1} ，故一般令分母上符号为正

抛物线法的收敛阶约为 $p=1.840$

解非线性方程的牛顿迭代法

考察非线性方程组

$$\begin{cases} f_1\left(x_1,\cdots,x_n\right)=0 \\ \cdots \\ f_n\left(x_1,\cdots,x_n\right)=0 \end{cases}$$

其中 $f_1,f_2,\ldots,f_n$ 为关于 $x_1,x_2,\ldots x_n$ 的多元函数。记

$$\begin{gathered} x=\left(x_1,\cdots,x_n\right)^{\mathrm{T}} \in R^n \\ F=\left(f_1,\cdots,f_n\right)^{\mathrm{T}} \end{gathered}$$

非线性方程组可记为

$$F(x)=0$$

非线性方程组求根是方程求根的推广。将单变量函数 $f(x)$ 看作向量函数 $F(x)$ ，向量函数的近似根为 $x^{(k)}=\left(x_1^{(k)},\cdots,x_n^{(k)}\right)^{\mathrm{T}}$ ，将 $F(x)$ 的分量 $f_i(x)$ 在 $x^{(k)}$ 处多元函数泰勒展开，取其线性部分

$$F(x)\approx F\left(x^{(k)}\right)+F^{\prime}\left(x^{(k)}\right)\left(x-x^{(k)}\right)$$

令右端为 0，可得到线性方程组

$$F^{\prime}\left(x^{(k)}\right)\left(x-x^{(k)}\right)=-F\left(x^{(k)}\right)$$

其中

$$F'(x)=\begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \cdots & \frac{\partial f_2(x)}{\partial x_n} \\ \vdots & \vdots & \vdots & \\ \frac{\partial f_n(x)}{\partial x_1} & \frac{\partial f_n(x)}{\partial x_n} & \cdots & \frac{\partial f_n(x)}{\partial x_n} \end{pmatrix}$$

称为 $F(x)$ 的 Jacobi 矩阵，求解线性方程组，记解为 $x^{(k+1)}$ ，则可得

$$x^{(k+1)}=x^{(k)}-F'\left(x^{(k)}\right)^{-1}F\left(x^{(k)}\right) \quad (k=0,1,\cdots)$$

这就是解非线性方程组的牛顿迭代法