数值计算方法

非线性方程(组)的数值解法

信息管理与工程学院
冯银波

非线性方程的数值解法

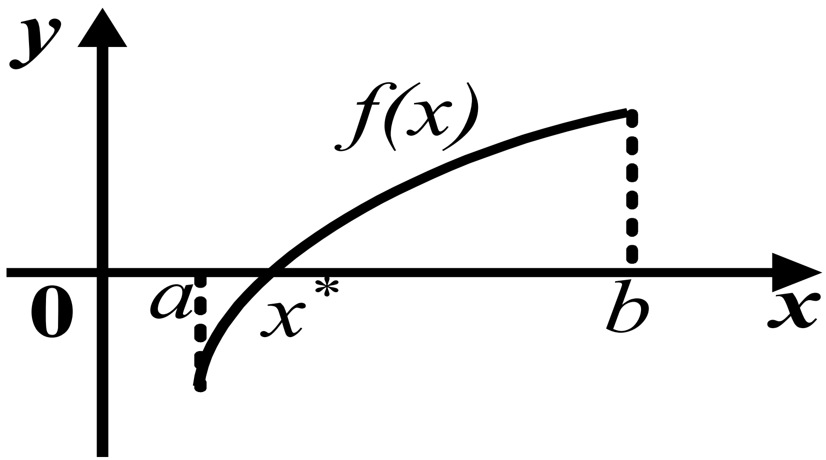
- 初始近似值的搜索
- 迭代法
- 牛顿迭代法（切线法）
- 弦截法（割线法）

初始近似值的搜索

- **方程的根**
  - 非线性方程：高次代数方程和超越方程。
  - **定义：** 非线性方程$f(x)=0$，存在x^*使$f(x^*)=0$，则称x^*为方程的根，又称为函数$f(x)$的零点。
  - 方程的根：实根、虚根。全局的根、局部的根。单根、重根。

有根区间

介值定理 若函数 $f(x)$ 在 $[a,b]$ 连续，且 $f(a)f(b)<0$，则方程 $f(x)=0$ 在 (a,b) 内至少有一个实根。将 $[a,b]$ 称为 $f(x)$ 的有根区间。



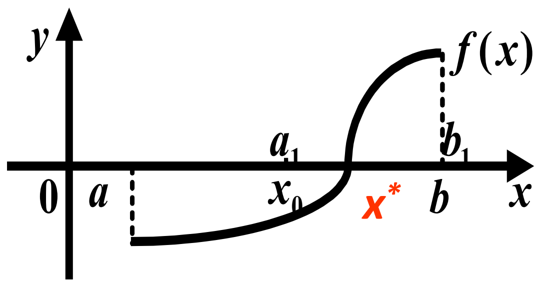
区间二分法

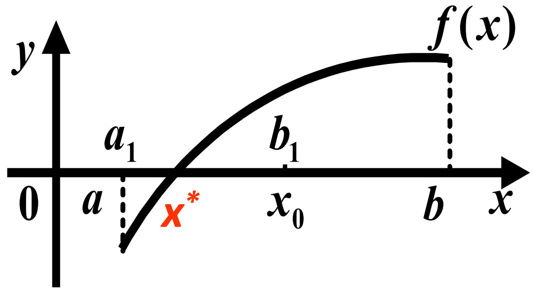
- **定理** 函数$f(x)$在$[a,b]$上单调连续，且$f(a)f(b) < 0$，则方程$f(x) = 0$在区间$[a,b]$上有且仅有一个实根x^*。
  - 二分法的基本思想
  - 将有根的区间二分为两个小区间，然后判断根在那个小区间，舍去无根的小区间，而把有根的小区间再一分为二，再判断根属于哪个更小的区间，如此反复 ，直到求出满足精度要求的近似根。

区间二分法

- 令$(a,b)=(a_0,b_0)$，中点$x_0=\frac{1}{2}(a_0+b_0)$
- 这时有三种情况
  - $f(x_0)=0$，x_0为所求的根
  - $f(x_0)$和$f(a_0)$同号，取$a_1=x_0, b_1=b_0$
  - $f(x_0)$和$f(b_0)$同号，取$a_1=a_0, b_1=x_0$
- 新的有根区间为(a_1,b_1)，长度是原来的一半
- 近似根x_k的误差估计

$$|x^*-x_0|\le\frac{1}{2}(b_0-a_0)$$



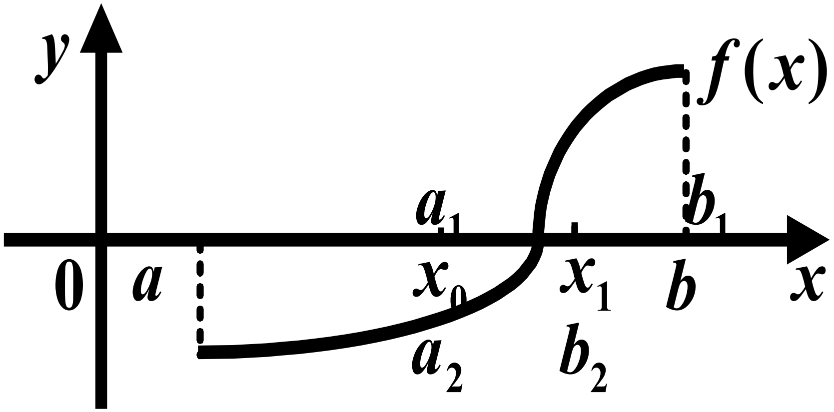


区间二分法

- 第二次二分，取中点$x_1 = \frac{1}{2}(a_1 + b_1)$
  - 若$f(a_1)f(x_1) < 0$, 则$x^* \in (a_1, x_1)$, 令$a_2 = a_1, b_2 = x_1$

    否则令$a_2 = x_1, b_2 = b_1$

新的有根区间为(a_2, b_2)



区间二分法

- 如此反复，有$(a,b)=(a_0,b_0)\supset(a_1,b_1)\supset(a_2,b_2)\supset\cdots\supset(a_k,b_k)\supset\cdots$

$$b_1-a_1=\frac{1}{2}(b_0-a_0),b_2-a_2=\frac{1}{2}(b_1-a_1)=\frac{1}{2^2}(b-a)$$

$$b_k-a_k=\frac{1}{2^k}(b-a)$$

- 近似根x_k的误差估计
- $x_k=\frac{1}{2}(a_k+b_k)\in(a_k,b_k),k=0,1,2,\cdots$
- $|x^*-x_k|\le\frac{1}{2}(b_k-a_k)=\frac{1}{2^{k+1}}(b-a)$

区间二分法

- 由此得二分过程结束的原则：
- 先给定精度要求ε(绝对误差限)，
  - 事先由ε估计出二分的最小次数k，取$x^* = x_k$

由$|x^* - x_k| \le \frac{1}{2^{k+1}}(b-a) < \varepsilon$得$2^{k+1} > \frac{b-a}{\varepsilon}$,

$$k > \frac{\lg(b-a) - \lg\varepsilon}{\lg 2} - 1$$

  - 当$|b_{k+1} - a_{k+1}| < \varepsilon$时结束二分计算，取$x^* = x_k$

区间二分法

计算步骤：

（1）找出 $f(x)=0$ 的有根区间 (a,b)，即确定 a，b 使得 $f(a)f(b)<0$，并计算 $f(a)$，$f(b)$。

（2）计算 $f(\frac{a+b}{2})$。

（3）若 $f(\frac{a+b}{2})=0$，计算停止；若 $f(\frac{a+b}{2})f(a)<0$，用 $\frac{a+b}{2}$ 代替 b；若 $f(\frac{a+b}{2})f(b)<0$，以 $\frac{a+b}{2}$ 代替 a。

（4）反复执行第二步与第三步，直到区间长缩小到允许误差范围之内，此时区间中点即可作为所**求**的近似解。

区间二分法

证明方程 $x^3-3x^2+6x-1=0$ 在区间（0，1）内有唯一的实根，并用二分法求这个根的近似值，使误差不超过 0.01。

解　设 $f(x)=x^3-3x^2+6x-1, x$□$[0,1]$。因为 $f(x)$ 在[0，1]连续，且 $f(0)=-1<0, f(1)=3>0$，由介值定理知 $f(x)=0$ 在区间[0，1]上至少有一个实根 x^*。又因为 $f'(x)=3x^2-6x+6=3(x-1)^2+3>0$，所以 $f(x)$ 在 $[0,1]$ 上单调增加，因而 $x^3-3x^2+6x-1=0$ 在区间 $(0,1)$ 内有唯一实根 x^*。

下面用二分法求这个根的近似值。

区间二分法

k	a_n	b_n	x_k	$f(x_k)$
0	0	1	0.500	+
1	0	0.500	0.250	+
2	0	0.250	0.125	-
3	0.125	0.250	0.188	+
4	0.125	0.188	0.157	-
5	0.157	0.188	0.173	-
6	0.173	0.188	0.267	+
7	0.173	0.267	0.220	+
8	0.173	0.220	0.196	+
9	0.173	0.196	0.185	+

区间二分法

k	a_n	b_n	x_k	$f(x_k)$
10	0.173	0.185	0.179	-
11	0.179	0.185	0.182	-
12	0.182	0.185	0.183	+
13	0.182	0.183	0.183	

所以 $x^* = 0.183$ 。

求多个根

- 二分法一般用于求某区间内的一个实根

- 如果一个方程在某区间内有多个根，如何把所有的根都求出来？
  - 对于多项式方程来说，问题相对简单
  - 对一般的方程来说，很难做到

普通不动点迭代法

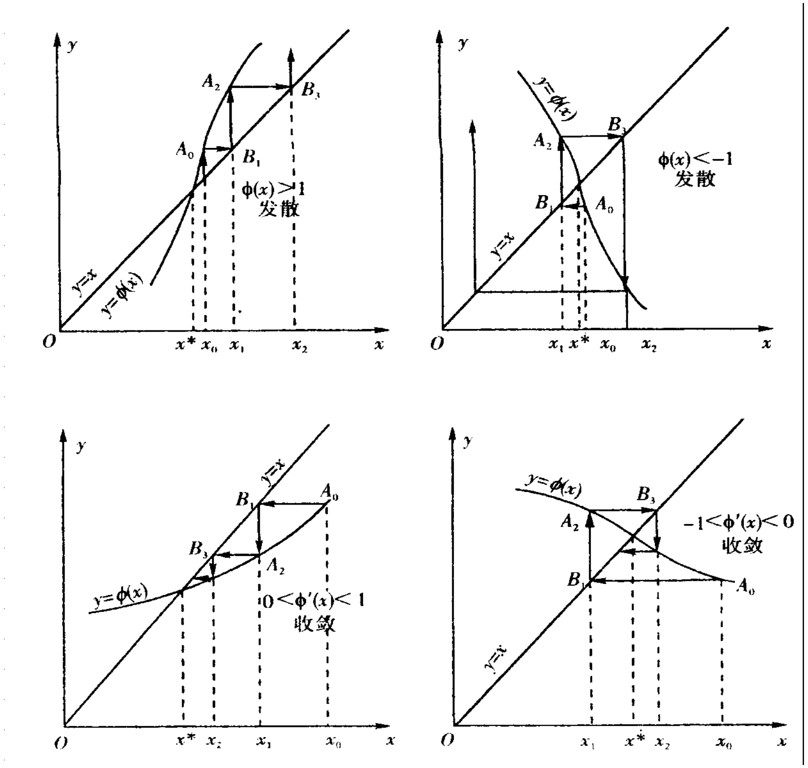
- 考虑如下不动点问题（等价于求根问题）

$$x = \varphi(x)$$

若 x^* 满足 $x^* = \varphi(x^*)$, 则称x^*是φ的一个不动点。构造迭代序列

$$x_{k+1} = \varphi(x_k), k = 0,1,2,\ldots$$

- 注意 $\varphi(x_k)$不一定收敛到x^*。



普通不动点迭代法

定理：如果$\varphi(x)$满足以下条件：

(1) 当$x \in [a,b]$时，$\varphi(x) \in [a,b]$,

(2) 对任意的$x \in [a,b]$时，存在$0 < L < 1$, 使得

$$|\varphi'(x)| \leq L < 1,$$

那么，方程$x = \varphi(x)$在上$[a,b]$有唯一的根x^*, 且对任意初值$x_0 \in [a,b]$,
迭代序列$x_{k+1} = \varphi(x_k)(k = 0,1,2,\ldots)$都收敛于$x^*$

普通不动点迭代法

证明：设 x^*是方程 f(x)=0 的根，即 $x^*=\varphi(x^*)$，由拉格朗日定理

$$x^* - x_{k+1} = \varphi(x^*) - \varphi(x_k) = \varphi'(\xi)(x^* - x_k)$$

其中 ξ 在 x^*与 x_k之间，由（2.7）

$$\begin{aligned}\left|x^* - x_{k+1}\right| &= \left|\varphi(x^*) - \varphi(x_k)\right| = \left|\varphi'(\xi)\right|\left|x^* - x_k\right| \\ &\le L\left|x^* - x_k\right| \le L^2\left|x^* - x_{k-1}\right| \le \cdots \\ &\le L^{k+1}\left|x^* - x_0\right|\end{aligned}$$

因为 0< L< 1，由 $\lim\limits_{k\to\infty} L^{k+1} = 0$ 知

$$\left|x^* - x_{k+1}\right| \to 0 \qquad (k \to \infty)$$

所以 $$\lim_{k\to\infty} x_k = x^*$$

即 $x_{k+1} = \varphi(x_k)$收敛

证完

牛顿法

- 牛顿法是求解非线性方程(组)的一种方法。

- 迭代公式：要求解非线性方程$f(x)=0$，若已知第k个迭代点x_k，且$f'(x_k)\neq 0$，那么第$k+1$个迭代点x_{k+1}有以下公式给出：

$$\boxed{x_{k+1}=x_k-\frac{f(x_k)}{f'(x_k)},k=0,1,2,\cdots}$$

- 迭代公式的导出：给定当前迭代点，把原方程（非线性方程）近似为线性方程，从而通过求解线性方程来得到原方程的一个近似解，即为下一个迭代点。

$$f(x)=f(x_k)+f'(x_k)(x-x_k)+\frac{f''(x_k)}{2}(x-x_k)^2+\cdots$$

$$\approx f(x_k)+f'(x_k)(x-x_k)$$ 当x在x_k附近时

因此，得到$f(x)=0$的一个近似解：$x=x_k-\frac{f(x_k)}{f'(x_k)}$

牛顿法几何意义

- 过曲线上的点 $(x_k, f(x_k))$ 作切线，切线方程

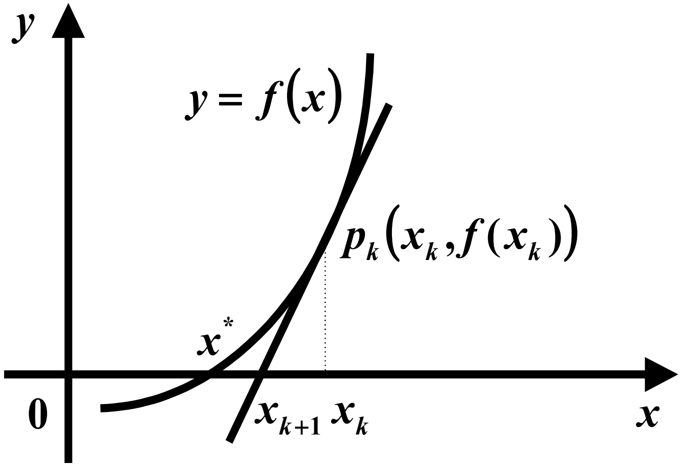
$$y = f(x_k) + f'(x_k)(x - x_k)$$

切线方程和横轴的交点$(x_{k+1}, 0)$，即

$$0 = f(x_k) + f'(x_k)(x - x_k)$$

若 $f'(x_k) \neq 0$，解出的解记为x_{k+1}，则得 Newton 迭代公式

$$\boxed{x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, k = 0,1,2,\cdots}$$



牛顿法-例

例：用牛顿迭代法求方程 $xe^x - 1 = 0$在$x = 0.5$附近的根。

解：

$$f(x) = xe^x - 1 \qquad f'(x) = e^x + xe^x$$

牛顿迭代法：

$$x_{k+1} = x_k - \frac{x_k e^{x_k} - 1}{e^{x_k} + x_k e^{x_k}} = x_k - \frac{x_k - e^{-x_k}}{1 + x_k}$$

取$x_0 = 0.5$，经计算可得

$$x_1 = 0.57102044 \qquad x_2 = 0.567165569$$

$$x_3 = 0.56714329 \qquad x_4 = 0.56714329$$

$$x^* \approx 0.567143$$

牛顿迭代法算法步骤

(1)给出 $x_0, \varepsilon, N = 0$

(2)计算 $\boldsymbol{x}_1 \Leftarrow \boldsymbol{x}_0 - \dfrac{f(\boldsymbol{x}_0)}{f'(\boldsymbol{x}_0)}$，$N \Leftarrow \mathrm{N} + 1$

(3)若 $\left|\boldsymbol{x}_1 - \boldsymbol{x}_0\right| < \boldsymbol{\varepsilon}$ 则转(4)；否则 $\boldsymbol{x}_0 \Leftarrow \boldsymbol{x}_1$，转(2);

(4)输出x_1 ,结束。

- 初始选取x_0和ε时，要使得ε远远小于$\frac{f(x_0)}{f'(x_0)}$，以免第一次迭代就终止

牛顿法的收敛性与收敛速度

牛顿法的迭代函数为：

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

$$|\varphi'(x)| = \frac{|f''(x)|}{f'(x)^2}|f(x)|$$

根据普通迭代法的收敛性定理可知

定理： 设 $f(x)$在[a, b]上满足下列条件：

（1）$f(a) \cdot f(b) < 0$；

（2）$f'(x) \neq 0$；

（3）$f''(x)$ 存在且不变号；

（4）取 $x_0 \in [a, b]$，使得 $f''(x) \cdot f(x_0) > 0$

则牛顿迭代序列$\{x_k\}$收敛于$f(x)$在[a, b]上的唯一根 x^*。

条件(4)为了保证 $\varphi(x)$也属于$[a, b]$

牛顿法的收敛性与收敛速度

我们已知：

$$\varphi'(x^*) = \frac{f(x^*)f''(x^*)}{[f'(x^*)]^2} = 0$$

$$|\varphi''(x^*)| = \frac{f''(x^*)}{f'(x^*)}$$

根据Taylor展式：

$$x_{k+1} - x^* = \varphi(x_k) - \varphi(x^*) = \frac{1}{2}\frac{f''(x^*)}{f'(x^*)}(x_k - x^*)^2 + o(x_k - x^*)^2$$

于是：

定理：设$f(x^*) = 0, f'(x^*) \neq 0, f''(x^*)$在$x^*$的一个邻域内连续，则牛顿法在$x^*$局部收敛，且至少2阶收敛。

更详细的证明可参考《Numerical Analysis: A Graduate Course》P195

牛顿法的拓展

1. 可以考虑用Taylor展式的前三项来近似$f(x)$，其几何意义就是用抛物线来近似曲线。

2. 当$f''(x)$会变号时，如果x_0选得不好，牛顿法很容易发散。此时，可以考虑Guarded Newton Method. 其迭代公式为：$x_{k+1} = x_k - s\frac{f(x_k)}{f'(x_k)}$，$0 < s \leq 1$. 每一步迭代时通过某种搜索法来确定满足$|f(x_{k+1})| \leq \left|\left(1 - \frac{1}{2}s\right)f(x_k)\right|$的尽可能大的$s$。

3. 当$f'(x)$不存在或者不方便计算时，可以用很多方法来替代$f'(x)$，比如用某个常数c来永久代替$f'(x)$，迭代公式为：$x_{k+1} = x_k - \frac{f(x_k)}{c}$. 也可采用下面的割线法。

4. 其他。

弦截法(割线法)

- 牛顿迭代法虽然有较高的收敛速度，但要计算导数值$f'(x)$, 很多情况下导数不方便计算
- 为避免导数的计算，用平均变化率

$$\frac{f(x_k)-f(x_0)}{x_k-x_0} \text{ 和 } \frac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}}$$

分别代替$f'(x_k)$，于是得到两种弦截法：

单点弦法

$$\boxed{x_{k+1}=x_k-\frac{f(x_k)}{f(x_k)-f(x_0)}(x_k-x_0), \quad k=1,2,\cdots}$$

双点弦法

$$\boxed{x_{k+1}=x_k-\frac{f(x_k)}{f(x_k)-f(x_{k-1})}(x_k-x_{k-1}), k=1,2,\cdots}$$

单点弦截法几何意义

设非线性方程 $f(x)=0$，在区间 $[x_0, x_1]$ 内有单根 x^*，选定曲线 $y=f(x)$ 上的两个点 $P_0(x_0, f(x_0))$，$P_1(x_1, f(x_1))$，过这两点作一条直线 $\overline{P_0P_1}$，其于 x 轴的交点即为 x_2

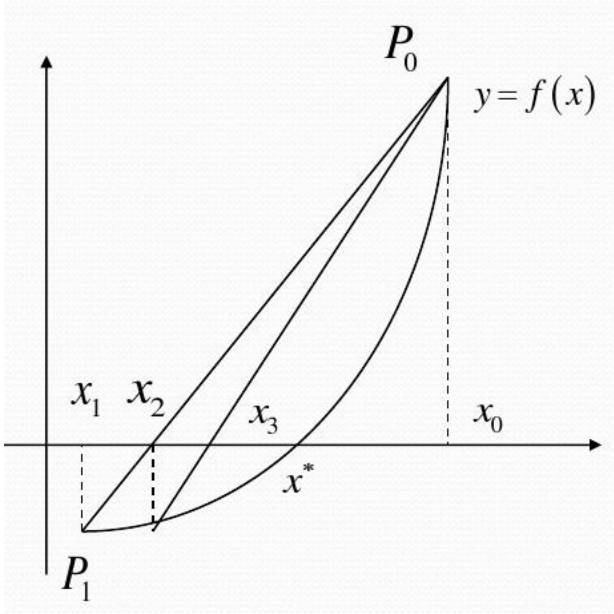
$$x_2 = x_1 - \frac{f(x_1)}{f(x_1)-f(x_0)}(x_1 - x_0)$$

写成迭代格式

$$x_{k+1} = x_k - \frac{f(x_k)}{f(x_k)-f(x_0)}(x_k - x_0) \qquad k = 1, 2, \cdots$$

这个公式的几何意义是过两个点作弦，这个弦与 x 轴的交点即是根的新的近似值，因为弦的一个端点 $P_0(x_0, f(x_0))$ 始终不变，所以这种方法成为单点弦截法

单点弦截法几何意义



单点弦截法收敛速度

迭代函数是:

$$\varphi(x) = x - \frac{f(x)}{f(x) - f(x_0)}(x - x_0)$$

$$\varphi'(x) = 1 + \frac{f'(x^*)}{f(x_0)}(x^* - x_0) = 1 - \frac{f'(x^*)}{\dfrac{f(x^*) - f(x_0)}{x^* - x_0}}$$

当x_0充分接近x^*时，$\frac{f(x^*)-f(x_0)}{x^*-x_0}$很接近$f'(x^*)$，且符号相同，所以$0 < |\varphi'(x)| < 1$，所以单点弦截法仅为线性收敛。

双点弦截法

- 迭代公式：

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

- 双点弦截法的收敛性与牛顿迭代法一样，即在根的某个邻域内$f(x)$ 有二阶连续导数，且$f'(x) \neq 0$，具有局部收敛性，同时在邻域任取初值x_0, x_1均收敛。双点弦截法具有超线性敛速，收敛的阶为$(1+\sqrt{5})/2$，证明略。

多元不动点迭代法

多元不动点问题：

$$\begin{cases} x_1 = g_1(x_1, \dots, x_n) \\ x_2 = g_2(x_1, \dots, x_n) \\ \qquad \vdots \\ x_n = g_n(x_1, \dots, x_n) \end{cases}$$

向量形式的迭代公式：$\boldsymbol{x}_{k+1} = \boldsymbol{g}(\boldsymbol{x}_k)$。

记$\nabla \boldsymbol{g}(\boldsymbol{x})$为$\boldsymbol{g}(\cdot)$在处的Jacobi矩阵（梯度）

$$\nabla \boldsymbol{g}(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial g_1}{\partial x_1} & \cdots & \dfrac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_n}{\partial x_1} & \cdots & \dfrac{\partial g_n}{\partial x_n} \end{bmatrix}$$

多元不动点迭代法

定理：设$\boldsymbol{x}^*$为$\boldsymbol{x}=\boldsymbol{g}(\boldsymbol{x})$的解。如果$\nabla\boldsymbol{g}(\boldsymbol{x})$是连续函数，且$\|\nabla\boldsymbol{g}(\boldsymbol{x}^*)\|<1$，则存在$\delta>0$使得当$\|\boldsymbol{x}_0-\boldsymbol{x}^*\|<\delta$时，迭代序列$\{\boldsymbol{x}_k\}$收敛到$\boldsymbol{x}^*$。

证明：证明思路和一元的情形类似，下面给出简要的证明过程。

根据多元中值定理知：

$$\boldsymbol{g}(\boldsymbol{v})-\boldsymbol{g}(\boldsymbol{u})=\int_0^1\nabla\boldsymbol{g}\big(\boldsymbol{u}+s(\boldsymbol{v}-\boldsymbol{u})\big)(\boldsymbol{v}-\boldsymbol{u})ds$$

在$\boldsymbol{x}^*$的某个邻域内有$\|\nabla\boldsymbol{g}(\boldsymbol{x})\|=L<1$，于是

$$\|\boldsymbol{x}_{k+1}-\boldsymbol{g}(\boldsymbol{x}^*)\|=\|\boldsymbol{g}(\boldsymbol{x}_k)-\boldsymbol{g}(\boldsymbol{x}^*)\|=\left\|\int_0^1\nabla\boldsymbol{g}\big(\boldsymbol{x}^*+s(\boldsymbol{x}_k-\boldsymbol{x}^*)\big)(\boldsymbol{x}_k-\boldsymbol{x}^*)ds\right\|$$

$$\le\int_0^1\left\|\nabla\boldsymbol{g}\big(\boldsymbol{x}^*+s(\boldsymbol{x}_k-\boldsymbol{x}^*)\big)(\boldsymbol{x}_k-\boldsymbol{x}^*)\right\|ds$$

$$\le\int_0^1\left\|\nabla\boldsymbol{g}\big(\boldsymbol{x}^*+s(\boldsymbol{x}_k-\boldsymbol{x}^*)\big)\right\|\|\boldsymbol{x}_k-\boldsymbol{x}^*\|ds$$

$$\le L\|\boldsymbol{x}_k-\boldsymbol{x}^*\|$$

于是$\boldsymbol{x}_k$收敛到$\boldsymbol{x}^*$。

牛顿法求解非线性方程组

考虑n元非线性方程组$\boldsymbol{f}(\boldsymbol{x})=\mathbf{0}$，$\nabla\boldsymbol{f}(\boldsymbol{x})$为Jacobi矩阵，$\nabla\boldsymbol{f}(\boldsymbol{x})^{-1}$为其逆矩阵。则牛顿法迭代公式为：

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k-\nabla\boldsymbol{f}(\boldsymbol{x}_k)^{-1}\boldsymbol{f}(\boldsymbol{x}_k)$$

牛顿法的几何意义：

- 一元情形中，给定$\boldsymbol{x}_k$，在$\boldsymbol{x}_k$附近用线性函数来近似$\boldsymbol{f}(\boldsymbol{x})$，将求解非线性方程问题$\boldsymbol{f}(\boldsymbol{x})=\mathbf{0}$近似看成是求解线性方程$0=f(x_k)+f'(x_k)(x-x_k)$，从而得到$\boldsymbol{x}_{k+1}$。
- 多元情形中，给定$\boldsymbol{x}_k$，$\boldsymbol{f}(\boldsymbol{x})$在$\boldsymbol{x}_k$处Taylor展示为：$\boldsymbol{f}(\boldsymbol{x})\approx\boldsymbol{f}(\boldsymbol{x}_k)+\nabla\boldsymbol{f}(\boldsymbol{x}_k)(\boldsymbol{x}-\boldsymbol{x}_k)$。然后将求解$\boldsymbol{f}(\boldsymbol{x})=\mathbf{0}$看成是求解线性方程组$\mathbf{0}=\boldsymbol{f}(\boldsymbol{x}_k)+\nabla\boldsymbol{f}(\boldsymbol{x}_k)(\boldsymbol{x}-\boldsymbol{x}_k)$，从而得到$\boldsymbol{x}_{k+1}$。这其实就是用一个线性方程组来近似一个非线性方程组。

牛顿法求解非线性方程组

定理：设$\boldsymbol{f}(\boldsymbol{x}^*) = \boldsymbol{0}$, $\nabla \boldsymbol{f}(\boldsymbol{x}^*)$可逆, $\boldsymbol{f}(\boldsymbol{x})$在$x^*$的一个邻域内的二阶导连续，则存在$\delta > 0$使得当$\|\boldsymbol{x}_0 - \boldsymbol{x}^*\| < \delta$时，牛顿法收敛，且至少2阶收敛。

证明可参考《Numerical Analysis: A Graduate Course》P201

牛顿法拓展：对于多元的情形，同样可以考虑用Taylor展式的前三项来近似$\boldsymbol{f}(\boldsymbol{x})$，即每一步迭代都用二次方程组来近似原方程组。当一阶导或二阶导不存在或不好计算时，可以考虑用很多近似手段或替代手段。

作业

1 / 编写牛顿法、双弦截法求平方根的程序。