要求

1. 熟练掌握求解非线性方程的几种基本迭代法：二分法、简单迭代法、牛顿法、弦割法
2. 迭代法的收敛性
3. 掌握求解非线性方程组的几种迭代法：简单迭代法、牛顿法
4. 了解求解非线性方程组的弦割法、Broyden法

非线性方程

Definition

若数x^*满足$f(x^*) = 0$，则称x^*是方程$f(x) = 0$的解或根，或者称x^*是函数$f(x)$的零点．若$f(x)$是n次多项式，即

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$

其中$a_n \neq 0$，则称方程为n次多项式方程或代数方程；若$f(x)$为超越函数，则称方程为超越方程．$n > 1$的代数方程和超越方程统称为非线性方程.

非线性方程

Definition

若数x^*满足$f(x^*)=0$，则称x^*是方程$f(x)=0$的解或根，或者称x^*是函数$f(x)$的零点．若$f(x)$是n次多项式，即

$$f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n,$$

其中$a_n\neq 0$，则称方程为n次多项式方程或代数方程；若$f(x)$为超越函数，则称方程为超越方程．$n>1$的代数方程和超越方程统称为非线性方程．

对于代数方程，$n\leq 4$时，它的根可用求根公式表示．而当次数$n\geq 5$时，用迭代解法进行数值求解．

非线性方程

Definition

若$f(x)=(x-x^*)^m g(x)$满足$g(x^*)\neq 0$，则称x^*是方程$f(x)=0$的m重根.

Theorem

若函数$f(x)$在区间$[a,b]$上有m阶连续导数，则x^*是$f(x)$的m重零点，$\Leftrightarrow$ $f(x^*)=f'(x^*)=\cdots=f^{(m-1)}(x^*)=0, f^{(m)}(x^*)\neq 0.$"

介绍求非线性方程实单根x^*的迭代法. 若要求多个根，则可根据已求得的根$x_1^*,\cdots,x_m^*$，由函数$f_m(x)=\dfrac{f(x)}{\prod_{i=1}^{m}(x-x_i^*)}$求方程的第$m+1$个根$x_{m+1}^*$.

非线性方程

方程$f(x) = 0$的求根步骤：

- 根的存在性：是否有根，有几个？
- 根的隔离：分成小区间，每个区间一个根
- 根的精确化：迭代求解，满足一定精度

方程$f(x) = 0$根的搜索方法：

- 图解法
- 解析法
- 近似方程法
- 迭代法

Theorem

零点存在定理："如果函数$f(x)$在区间$[a,b]$上连续，且$f(a)f(b)<0$，则在区间(a,b)内必定存在一点x^*，使$f(x^*)=0$."

二分法，其计算过程如下.

1. 给定 $\varepsilon_1>0,\ \varepsilon_2>0,\ a_0:=a,\ b_0:=b,\ k:=0.$
2. 标号1 $x_k:=\frac{1}{2}(a_k+b_k)$，做
   若 $|f(x_k)|<\varepsilon_1$ 或 $|b_k-a_k|<\varepsilon_2$ 则取 $x^*\approx x_k$，停.
   否则 若$f(a_k)f(x_k)<0$，则$a_{k+1}:=a_k$，$b_{k+1}:=x_k$；否则
   令$a_{k+1}:=x_k$，$b_{k+1}:=b_k$.
   $k:=k+1$，转标号1

二分法

由此过程可得到一系列的含根区间

$$[a,b]\supset[a_1,b_1]\supset[a_2,b_2]\supset\cdots\supset[a_n,b_n]\supset\cdots,$$

根x^*的近似值x_k的误差估计式如下：

$$|x^*-x_k|=|x^*-\frac{a_k+b_k}{2}|\le\frac{b_k-a_k}{2}=\frac{1}{2}\frac{b_{k-1}-a_{k-1}}{2}=\cdots=\frac{b-a}{2^{k+1}}.$$

显然，当$k\to\infty$时，$x_k\to x^*$.

给定$\epsilon>0$, 要使$|x_n-\alpha|<\epsilon$, i.e. $\frac{b-a}{2^{n+1}}<\epsilon$, $\therefore n=[\frac{(\ln(b-a)-\ln\epsilon)}{\ln 2}]$.

优点 过程简单、方法可靠且总是收敛，只要求f连续；

缺点 运算量大，需要多次计算函数$f(x)$的值；不能求偶重根，不能求复根。

二分法常用于求根的大体范围或用于求其他快速收敛的迭代法所需的一个初始点.

简单迭代法

将方程$f(x)=0$改写为同解方程$x=\phi(x)$.

设$\phi(x)$连续，构造迭代格式

$$x_{k+1}=\phi(x_k),\quad k=0,1,2,\cdots.$$

其中$\phi(x)$ 称为**迭代函数**.

给定x^*的初值x_0，按迭代格式逐次迭代，得到一迭代点列$\{x_k\}$.

若$\lim\limits_{k\to\infty} x_k=x^*$，得

$$x^*=\lim_{k\to\infty}x_{k+1}=\lim_{k\to\infty}\phi(x_k)=\phi(\lim_{k\to\infty}x_k)=\phi(x^*).$$

即x^*是$\phi(x)$的不动点，也即方程的根.

在实际计算时，若$|f(x_{k+1})|<\varepsilon_1$ 或者 $|x_{k+1}-x_k|<\varepsilon_2$，则取$x_{k+1}$作为$x^*$的近似值.

例7.1 用简单迭代法求区间(2,3)内方程$x^3-2x-5=0$的根.

解法1 原方程变为$x^3=2x+5$，得$x=\sqrt[3]{2x+5}$，作迭代格式

$$x_{k+1}=\sqrt[3]{2x_k+5}, \quad k=0,1,2,\cdots$$

取初始点$x_0=2.5$，按上式迭代得

$$\begin{aligned} &x_1=2.154\,434\,690, \quad x_2=2.103\,612\,029, \quad x_3=2.095\,927\,410, \\ &x_{10}=2.094\,551\,484, \quad x_{11}=2.094\,551\,482=x_{12} \end{aligned}$$

解法2 两边同加$2x^3+5$, 再同除$3x^2-2$得同解方程$x=(2x^3+5)/(3x^2-2)$, 作迭代格式

$$x_{k+1}=(2x_k^3+5)/(3x_k^2-2), \quad k=0,1,2,\cdots$$

取初始点$x_0=2.5$，按上式迭代得$x_1=2.164\,179\,104$, $x_2=2.097\,135\,356$, $x_3=2.094\,555\,232$, $x_4=2.094\,551\,482=x_5$.

解法3 $x=(x^3-5)/2$, 作迭代格式

$$x_{k+1}=(x_k^3-5)/2, \quad k=0,1,2,\cdots$$

取初始点$x_0=2.5$，按上式迭代得 $x_1=5.312\,5$,
$x_2=72.466\,430\,66$, $x_3=190\,272.011\,8$, $x_4=3.444\,250\,536\times10^{16}$,
$x_5=2.042\,933\,398\times10^{46}$，计算$x_6$时溢出.

从以上三种解法可见，迭代点列是否收敛以及收敛的快慢，同迭代函数$\phi(x)$的选取有关.

几何意义：α是$y=x$与$y=\phi(x)$交点的横坐标。
折线法。$|\phi'(x)|<1$时收敛。

牛顿法

设$f(x)$在含根区间$[a,b]$上二阶连续可微且$f'(x)\neq 0$. x_k是方程$f(x)=0$的根x^*的一个近似值. 则由泰勒公式得

$$0=f(x^*)=f(x_k)+f'(x_k)(x^*-x_k)+\frac{1}{2!}f''(\xi_k)(x^*-x_k)^2.$$

略去二次项，则得近似线性方程

$$f(x_k)+f'(x_k)(x^*-x_k)\approx 0.$$

由此解出

$$x^*\approx x_k-\frac{f(x_k)}{f'(x_k)},$$

显然，$x_k-\frac{f(x_k)}{f'(x_k)}$是$x^*$的一个更好的近似值，故令

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

称为牛顿（**Newton**）法.

牛顿法几何意义

在几何上，Newton法所得x_{k+1}，是曲线$y=f(x)$在点(x_k,y_k)处的**切线**

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

与x轴的交点. 继续在点$(x_{k+1},f(x_{k+1}))$处作切线，该切线与x轴的交点为x_{k+2},它比x_{k+1}更靠近x^*. 如此继续迭代下去. 若初始点x_0取得充分靠近x^*，则由牛顿法产生的迭代点列$\{x_k\}$收敛于方程的解x^*. 因此，牛顿法又称为**切线法**

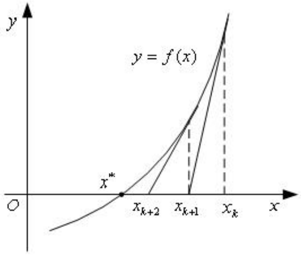


图7.1 牛顿法几何意义

(2) 改进牛顿法1

设$f(x)$在含根区间$[a,b]$上三阶连续可微且$f'(x)\neq 0$, $f''(x^*)\neq 0$.
则由泰勒公式略去三次项，则得近似二次方程

$$f''(x_k)(x^*-x_k)^2+2f'(x_k)(x^*-x_k)+2f(x_k)\approx 0.$$

由此解出

$$x^*\approx x_k+\frac{-f'(x_k)\pm\sqrt{f'^2(x_k)-2f(x_k)f''(x_k)}}{f''(x_k)}.$$

记

$$\tilde{x}_{k+1}=x_k-\frac{f'(x_k)+sign(f'(x_k))\sqrt{f'^2(x_k)-2f(x_k)f''(x_k)}}{f''(x_k)}.$$

$$\bar{x}_{k+1}=x_k-\frac{2f(x_k)}{f'(x_k)+sign(f'(x_k))\sqrt{f'^2(x_k)-2f(x_k)f''(x_k)}}.$$

取$\tilde{x}_{k+1},\bar{x}_{k+1}$中最靠近$x_k$者，作为$x_{k+1}$.

(3) 改进牛顿法2

设函数$y=f(x)$的反函数$x=g(y)$三阶连续可微，将$g(y)$在y_k处三阶泰勒展开，并

由$x^*=g(0)$，$y_k=f(x_k)$，$x_k=g(y_k)$，$g'(y_k)=\frac{1}{f'(x_k)}$， 及

$$g''(y_k)=\frac{\mathrm{d}}{\mathrm{dy}}\left(\frac{\mathrm{dx}}{\mathrm{dy}}\right)\Big|_{y=y_k}=\frac{\mathrm{d}}{\mathrm{dx}}\left(\frac{1}{f'(x)}\right)\frac{\mathrm{dx}}{\mathrm{dy}}\Big|_{y=y_k}=-\frac{f''(x_k)}{f'^3(x_k)}.$$

得

$$x^*=g(0)\approx x_k-\frac{f(x_k)}{f'(x_k)}-\frac{f''(x_k)}{2f'^3(x_k)}f^2(x_k).$$

令

$$x_{k+1}=x_k-\frac{f(x_k)}{f'(x_k)}-\frac{f''(x_k)}{2f'^3(x_k)}f^2(x_k),\quad k=0,1,2,\cdots$$

即得**改进牛顿法2**的迭代格式.

简化牛顿法

实际问题中若导数$f'(x)$难以计算或计算量较大，

用$f'(x_0)$或常数c替代牛顿法中的$f'(x_k)$，则得简化牛顿法

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_0)}, \quad k = 0, 1, \cdots$$

$$\text{或} \quad x_{k+1} = x_k - \frac{f(x_k)}{c}, \quad k = 0, 1, \cdots$$

其几何意义是用斜率为$f'(x_0)$或c的平行弦与x轴的交点作为x^*的新近似点.

牛顿下山法

为了放宽牛顿法的初始点x_0的选择范围，修改为

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

称为**牛顿下山法**，λ称为**下山因子**. 选取λ应使

$$|f(x_{k+1})| < |f(x_k)|, \quad k = 0, 1, 2, \cdots$$

成立. 实际计算时，选取λ采用试算的办法，即从$\lambda = 1$开始，逐次减半进行试算，直到上式成立为止.

弦割法

在牛顿迭代法中，$f'(x_k)$用两点$(x_{k-1}, f(x_{k-1}))$, $(x_k, f(x_k))$或$(x_0, f(x_0))$, $(x_k, f(x_k))$ 连线（弦）的斜率

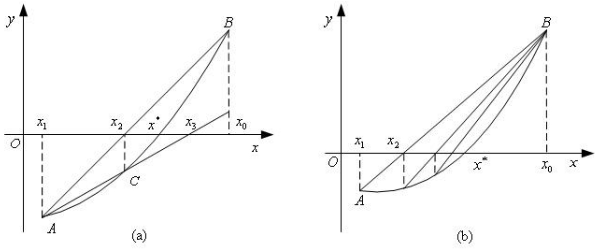
$$\frac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}}, \quad \text{或} \quad \frac{f(x_k)-f(x_0)}{x_k-x_0}$$

来替代，则得两点弦割法和单点弦割法的迭代格式

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

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

从几何上看弦割法是取过两点$(x_{k-1}, f(x_{k-1})), (x_k, f(x_k))$ 或$(x_0, f(x_0)), (x_k, f(x_k))$ 弦线与x轴的交点作为x^*的新近似点x_{k+1}.其几何意义如图7.2所示.



弦割法优点：不计算导数
弦割法缺点：需高精度运算，因为$f(x_k), f(x_{k-1})$很接近

改进弦割法

设函数$y=f(x)$的反函数为$x=g(y)$，作$g(y)$的二次牛顿插值多项式$N_2(y)$，则

$$x=g(y)\approx N_2(y)=g(y_k)+g[y_k,y_{k-1}](y-y_k)+g[y_k,y_{k-1},y_{k-2}](y-y_k)($$

取$y=0$，则得

$$x^*=g(0)\approx g(y_k)-g[y_k,y_{k-1}]\,y_k+g[y_k,y_{k-1},y_{k-2}]\,y_k\,y_{k-1}.$$

得**改进弦割法**的迭代格式

$$\begin{aligned}x_{k+1}&=x_k-\frac{x_k-x_{k-1}}{f(x_k)-f(x_{k-1})}f(x_k)\\&+\Big(\frac{x_k-x_{k-1}}{f(x_k)-f(x_{k-1})}-\frac{x_{k-1}-x_{k-2}}{f(x_{k-1})-f(x_{k-2})}\Big)\frac{f(x_k)f(x_{k-1})}{f(x_k)-f(x_{k-2})}.\end{aligned}$$

迭代法的收敛性

Definition

设x^*为方程$f(x)=0$的根，若存在x^*的一个邻域$N_\delta(x^*)=\{x\big||x-x^*|\le\delta,\ \delta>0\}$，对任意的初始点$x_0\in N_\delta(x^*)$，由迭代法产生的迭代点列$\{x_k\}$收敛于$x^*$，则称该迭代法局部收敛；若对含根区间$[a,b]$内任意的初始点$x_0$，迭代点列$\{x_k\}$都收敛于$x^*$，则称该迭代法全局收敛.

Theorem

迭代法收敛定理　设迭代函数$\phi(x)$一阶连续可微且满足条件：

1) 当$x \in [a,b]$时，$\phi(x) \in [a,b]$;
2) 存在常数$L(0 < L < 1)$，使得对任意$x \in [a,b]$,
   $|\phi'(x)| \le L < 1$;

则对任意初点$x_0 \in [a,b]$，由迭代格式$x_{k+1} = \phi(x_k)$产生的迭代点列$\{x_k\}$收敛于$\phi(x)$在区间$[a,b]$的唯一不动点x^*，且有误差估计式

$$|x^* - x_k| \le \frac{1}{1-L}|x_{k+1} - x_k| \le \frac{L}{1-L}|x_k - x_{k-1}|,$$

$$|x^* - x_k| \le \frac{L^k}{1-L}|x_1 - x_0|.$$

证明 1) **证不动点的存在性**. 定义$g(x)=x-\phi(x)$,
则$g(x)$在$[a,b]$上连续, 由条件1)知$g(a)=a-\phi(a)\le 0$,
$g(b)=b-\phi(b)\ge 0$, 由零点存在定理知, 存在$x^*\in[a,b]$,
使$g(x^*)=0$, 即$x^*=\phi(x^*)$.
2)**证唯一性**. 假定还有另一个不动点$\bar{x}^*\neq x^*$, 由条件2)得

$$0<|\bar{x}^*-x^*|=|\phi(\bar{x}^*)-\phi(x^*)|=|\phi'(\xi)(\bar{x}^*-x^*)|\le L|\bar{x}^*-x^*|<|\bar{x}^*-x^*|.$$

得到矛盾, 所以必有$\bar{x}^*=x^*$.
3)**证迭代点列的收敛性**. 由条件2)知

$$|x^*-x_k|=|\phi(x^*)-\phi(x_{k-1})|=|\phi'(\xi_{k-1})(x^*-x_{k-1})|$$
$$\le L|x^*-x_{k-1}|\le L^2|x^*-x_{k-2}|\le\cdots\le L^k|x^*-x_0|.$$

因$0 < L < 1$，所以 $\lim\limits_{k\to\infty} |x^* - x_k| = 0$. 即 $\lim\limits_{k\to\infty} x_k = x^*$.

4) 证误差估计式.

$$\begin{aligned} |x_{k+n} - x_k| &= |x_{k+n} - x_{k+n-1} + x_{k+n-1} - x_{k+n-2} + \cdots + x_{k+1} - x_k| \\ &= \sum_{j=1}^{n} |x_{k+j} - x_{k+j-1}|. \end{aligned}$$

而

$$\begin{aligned} |x_{k+j} - x_{k+j-1}| &= |\phi(x_{k+j-1}) - \phi(x_{k+j-2})| = |\phi'(\xi)(x_{k+j-1} - x_{k+j-2})| \\ &\le L|x_{k+j-1} - x_{k+j-2}| \le \cdots \le L^{j-1}|x_{k+1} - x_k|. \end{aligned}$$

$$|x_{k+n} - x_k| \le \sum_{j=1}^{n} L^{j-1}|x_{k+1} - x_k| = \frac{1 - L^n}{1 - L}|x_{k+1} - x_k|.$$

简单迭代法的收敛性

两端令$k \to \infty$取极限，注意到$0 < L < 1$，则得

$$|x^* - x_k| \le \frac{1}{1-L}|x_{k+1} - x_k| \le \frac{L}{1-L}|x_k - x_{k-1}|.$$

由$|x_{k+1} - x_k| \le L^k |x_1 - x_0|$. 则得

$$|x^* - x_k| \le \frac{L^k}{1-L}|x_1 - x_0|.$$

由定理的证明过程可见L越小，迭代点列$\{x_k\}$也就收敛的越快.

由式$|x^* - x_k| \le \frac{L^k}{1-L}|x_1 - x_0| < \varepsilon$ 两端取对数，得

$$k > \ln\left(\frac{\varepsilon(1-L)}{|x_1 - x_0|}\right) / \ln L.$$

上式给出了在给定精度要求下迭代次数k的一个大致估计.

Theorem

(迭代法的局部收敛性定理) 设迭代函数$\phi(x)$在其不动点x^*的某邻域$[x^*-\delta,\, x^*+\delta]$内连续可微，且存在常数$L(0<L<1)$，使

$$|\phi'(x)| \le L < 1, \quad \forall x \in [x^*-\delta,\, x^*+\delta]$$

成立，则对$\forall x_0 \in [x^*-\delta,\, x^*+\delta]$，由迭代公式$x_{k+1}=\phi(x_k)$产生的迭代点列$\{x_k\}$ 收敛于函数$\phi(x)$的唯一不动点.

事实上,

$$|\phi(x)-x^*| = |\phi(x)-\phi(x^*)| = |\phi'(\xi)|\,|x-x^*| \le L|x-x^*| < |x-x^*| \le \delta,$$

即$\forall x \in [x^*-\delta,\, x^*+\delta]$，$\phi(x) \in [x^*-\delta,\, x^*+\delta]$,即条件1)成立.

牛顿法的收敛性

Theorem

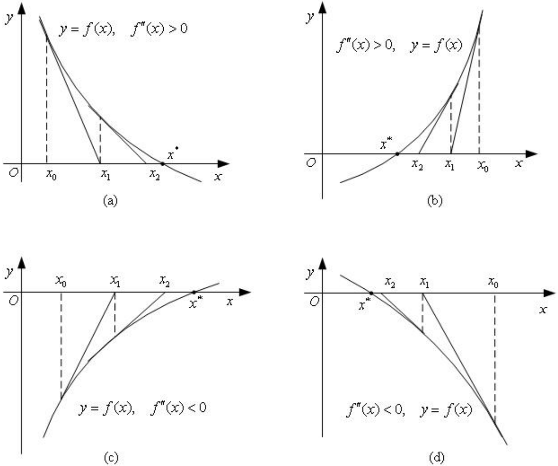
(牛顿法的局部收敛性定理) 设$f(x)$在其零点x^*附近二阶连续可微且$f'(x^*) \neq 0$，则当初始点x_0充分靠近x^*时，由牛顿法产生的迭代点列$\{x_k\}$ 收敛于x^*.

Theorem

(牛顿法的全局收敛性定理) 设$f(x)$区间$[a, b]$上二阶连续可微，且满足：*(1)* $f(a)f(b) < 0$; *(2)* $f'(x) \neq 0, x \in [a, b]$;
(3) $f''(x)$不变号，$x \in [a, b]$; *(4)* 初始点$x_0 \in [a, b]$
且$f(x_0)f''(x_0) > 0$, 则由牛顿法产生的迭代点列$\{x_k\}$单调地收敛于方程$f(x) = 0$在$[a, b]$上的唯一根x^*.

牛顿法的收敛性

证明　定理中条件(1)保证了根的存在性，条件(2)表明函数$f(x)$的严格单调性，从而保证了即根的唯一性.



牛顿法的收敛性

下面证明收敛性. 为叙述方便，不妨假定$f'(x) > 0$, $f''(x) > 0$(其他情况可类似证明). 由条件(4)有

$$0 < f(x_0) = f(x_0) - f(x^*) = f'(\xi)(x_0 - x^*),$$

可知$x_0 - x^* > 0$，即$x_0 > x^*$. 从而

$$\begin{aligned} x_1 - x_0 &= -\frac{f(x_0)}{f'(x_0)} < 0, \Rightarrow x_1 < x_0. \\ x_1 - x^* &= x_0 - x^* - \frac{f(x_0) - f(x^*)}{f'(x_0)} \\ &= \frac{f'(x_0) - f'(\xi)}{f'(x_0)}(x_0 - x^*) = \frac{f''(\xi^*)(x_0 - \xi)(x_0 - x^*)}{f'(x_0)} > 0 \end{aligned}$$

故知$x^* < x_1 < x_0$. 同理可证$x^* < x_k < x_{k-1} < \cdots < x_1 < x_0$. 这表明序列$\{x_k\}$单调减下有界，必然收敛.

例 用迭代法求方程$x^2-3=0$的根$x=\sqrt{3}$.

解:

1. $x_{k+1}=x_k^2+x_k-3$, $\varphi'(x)=2x+1$, $\varphi'(\sqrt{3})=2\sqrt{3}+1>1$, 发散
2. $x_{k+1}=\frac{3}{x_k}$, $\varphi'(x)=-\frac{3}{x^2}$, $\varphi'(x_k)=-1, 2, 1.5, 2, 1.5, \ldots$ 震荡
3. $x_{k+1}=x_k-\frac{1}{4}(x_k^2-3)$, $\varphi'(x)=1-\frac{1}{2}x$,
   $\varphi'(\sqrt{3})=1-\frac{\sqrt{3}}{2}<1$, 收敛
4. $x_{k+1}=x_k-\frac{x_k^2-3}{2x_k}=\frac{1}{2}(x_k+\frac{2}{x_k})$, $\varphi'(x)=\frac{1}{2}(1-\frac{3}{x^2})$,
   $\varphi'(\sqrt{3})=0$, 二阶收敛。

Theorem

(弦割法的局部收敛性定理) 设$f(x)$在包含$f(x)=0$的根x^*的某个邻域上二阶连续可微，且$f'(x^*)\neq 0$，取初始点x_0, x_1充分接近x^*，则由弦割法迭代格式产生的迭代点列$\{x_k\}$收敛于x^*.

证明 由牛顿插值公式有，

$$
\begin{aligned}
x^* = g(0) &= g(y_k) - g[y_k, y_{k-1}]\, y_k + \frac{1}{2} g''(\eta_k) y_k\, y_{k-1} \\
&= x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k) - \frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)} f(x_k) f(x_{k-1}) \\
&= x_{k+1} - \frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)} (f(x_k) - f(x^*))(f(x_{k-1}) - f(x^*)) \\
&= x_{k+1} - \frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)} f'(\xi_k^*)(x_k - x^*) f'(\xi_{k-1}^*)(x_{k-1} - x^*).
\end{aligned}
$$

故

$$x_{k+1} - x^* = \frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)} f'(\xi_k^*)(x_k - x^*) f'(\xi_{k-1}^*)(x_{k-1} - x^*).$$

由此可见当x_{k-1}, x_k充分靠近x^*时，

$$|x_{k+1} - x^*| < |x_k - x^*|, \quad \text{由此可知} \quad \lim_{k\to\infty} x_k = x^* \ .$$

Theorem

(弦割法的全局收敛性定理) 在牛顿法的全局收敛性定理的条件下，取$x_1 \in [a, b]$使$f(x_1)f''(x_1) < 0$，则由弦割法迭代格式产生的迭代点列$\{x_k\}$收敛于方程$f(x) = 0$在$[a, b]$上的唯一根x^*.

通常弦割法仅有局部收敛性。初值选得不好，方法不收敛。

每次迭代前，先判断有根区间，确定包含根的两点，然后将这两点代入弦割法。

找x_1, x_2，使得$f(x_1)f(x_2) < 0$,

用弦割法得x_3，再选$x_3, x_i (i = 2, 1, -1)$，使得

$$f(x_i)f(x_3) < 0$$

用弦割法得x_4. 再选$x_4, x_i (i = 3, 2, 1, -1)$，使得

$$f(x_i)f(x_4) < 0$$

$\cdots$

迭代法的收敛速度

Definition

设迭代点列$\{x_k\}$收敛于x^*，若存在常数$p \geq 1$和$c > 0$，使

$$\lim_{k\to\infty} \frac{|x^* - x_{k+1}|}{|x^* - x_k|^p} = c \neq 0 \ ,$$

则称该迭代点列是p阶收敛的. c称为渐近误差常数. 特别地，当$p = 1$ $(0 < c < 1)$时称为线性收敛；$p > 1$时称为超线性收敛；$p = 2$称为平方收敛或二阶收敛.

如果迭代法产生迭代点列$\{x_k\}$是p阶收敛的，则称该迭代法具有p阶收敛速度.

迭代法的收敛速度

Theorem

（**收敛阶定理**）　设迭代函数$\phi(x)$在其不动点x^*的邻域内具有连续的$p(p>1)$阶导数，则由迭代格式产生的迭代点列$\{x_k\}$ p阶收敛的充要条件是

$$\phi'(x^*)=\phi''(x^*)=\cdots=\phi^{(p-1)}(x^*)=0,\quad \phi^p(x^*)\neq 0.$$

证明 充分性　由泰勒公式及定理条件知

$$\begin{aligned}
x_{k+1} &= \phi(x_k)=\phi(x^*)+\phi'(x^*)(x_k-x^*)+\cdots \\
&+ \frac{1}{(p-1)!}\phi^{(p-1)}(x^*)(x_k-x^*)^{p-1}+\frac{1}{p!}\phi^{(p)}(\xi_k)(x_k-x^*)^p \\
&= \phi(x^*)+\frac{1}{p!}\phi^{(p)}(\xi_k)(x_k-x^*)^p,
\end{aligned}$$

（其中ξ_k介于x_k与x^*之间.）

由此可得

$$|x^* - x_{k+1}| = \frac{1}{p!}|\phi^{(p)}(\xi_k)(x_k - x^*)^p|.$$

所以,

$$\lim_{k\to\infty} \frac{|x^* - x_{k+1}|}{|x^* - x_k|^p} = \lim_{k\to\infty} \frac{1}{p!}|\phi^{(p)}(\xi_k)| = \frac{1}{p!}|\phi^{(p)}(x^*)| \neq 0.$$

即迭代点列$\{x_k\}$ p 阶收敛.

证必要性　反证法　设迭代点列$\{x_k\}$ p 阶收敛. 假定(1)式不成立，则必有最小正整数$p_0 < p$，使得

$$\phi^{(j)}(x^*) = 0 \ (j = 1, 2, \cdots, p_0 - 1), \qquad \phi^{(p_0)}(x^*) \neq 0.$$

由充分性证明知点列$\{x_k\}$ p_0 阶收敛，得到矛盾. 故(1)成立.

1. 简单迭代法线性收敛

简单迭代法$x_{k+1} = \phi(x_k)$当$0 < |\phi'(x^*)| < 1$ 时是线性收敛的.

事实上，设x^*是$x = \phi(x)$的不动点，即$x^* = \phi(x^*)$. 该式减去$x_{k+1} = \phi(x_k)$, 则得

$$x^* - x_{k+1} = \phi(x^*) - \phi(x_k) = \phi'(\xi_k)(x^* - x_k).$$

其中ξ_k介于x^*与x_k之间. 从而

$$\lim_{k\to\infty} \frac{|x^* - x_{k+1}|}{|x^* - x_k|} = \lim_{k\to\infty} |\phi'(\xi_k)| = |\phi'(x^*)|.$$

2. 牛顿法二阶收敛

设$f(x)$在含根区间$[a,b]$上二阶连续可微且$f'(x)\neq 0$. x_k是方程$f(x)=0$的根x^*的一个近似值. 则由泰勒公式有

$$0=f(x^*)=f(x_k)+f'(x_k)(x^*-x_k)+\frac{1}{2!}f''(\xi_k)(x^*-x_k)^2.$$

上式两端同除以$f'(x_k)$得

$$\begin{aligned}&x^*-x_k+\frac{f(x_k)}{f'(x_k)}+\frac{1}{2}\frac{f''(\xi_k)}{f'(x_k)}(x^*-x_k)^2\\&=x^*-x_{k+1}+\frac{1}{2}\frac{f''(\xi_k)}{f'(x_k)}(x^*-x_k)^2=0.\end{aligned}$$

所以,

$$\lim_{k\to\infty}\frac{|x^*-x_{k+1}|}{|x^*-x_k|^2}=\lim_{k\to\infty}\frac{1}{2}\frac{|f''(\xi_k)|}{|f'(x_k)|}=\frac{1}{2}\frac{|f''(x^*)|}{|f'(x^*)|}\neq 0.$$

3. 弦割法超线性收敛

由牛顿插值公式有，

$$x_{k+1} - x^* = \frac{1}{2}\frac{f''(\xi_k)}{f'^3(\xi_k)} f'(\xi_k^*)(x_k - x^*) f'(\xi_{k-1}^*)(x_{k-1} - x^*).$$

$$\frac{|x^* - x_{k+1}|}{|x^* - x_k|^p} = \frac{1}{2}\left|\frac{f''(\xi_k) f'(\xi_k^*) f'(\xi_{k-1}^*)}{f'^3(\xi_k)}\right|$$

$$\times \left(\frac{|x_k - x^*|}{|x^* - x_{k-1}|^p}\right)^{1-p} |x^* - x_{k-1}|^{1+p(1-p)}.$$

在上式两端令$k \to \infty$，取极限得

$$c = \frac{1}{2}\left|\frac{f''(x^*)}{f'(x^*)}\right| c^{1-p} \lim_{k\to\infty} |x_{k-1} - x^*|^{1+p-p^2}.$$

得$p = \frac{1+\sqrt{5}}{2} \approx 1.618$，即弦割法超线性收敛.

4. 单点弦割法线性收敛

单点弦割法实质上是简单迭代法，其迭代函数

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

在点x^*处的导数 $\phi'(x^*) = 1 - \frac{f'(x^*)}{f'(\xi)}$, 其中$\xi$介于$x_0$与$x^*$之间，由此可见单点弦割法一般线性收敛，但当$f'(x)$变化不大时$\phi'(x^*) \approx 0$,收敛可能很快.

4. **单点弦割法线性收敛**

单点弦割法实质上是简单迭代法，其迭代函数

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

在点x^*处的导数 $\phi'(x^*) = 1 - \frac{f'(x^*)}{f'(\xi)}$, 其中$\xi$介于$x_0$与$x^*$之间，由此可见单点弦割法一般线性收敛，但当$f'(x)$变化不大时$\phi'(x^*) \approx 0$,收敛可能很快.

弦割法的收敛阶虽然低于牛顿法，但每次迭代只需计算一个函数值$f(x_k)$，不需计算导数值$f'(x_k)$，效率高，实际问题中经常采用.

1. 松弛加速法

将原来的方程$x=\phi(x)$作同解变形，在方程两端减去$\omega x(\omega\neq 1)$ (ω称为**松弛因子**)，得$x-\omega x=\phi(x)-\omega x$. 由此可得

$$x=\frac{\phi(x)-\omega x}{1-\omega}\triangleq\psi(x),$$

则有$x=\psi(x)$. 由此可得迭代格式

$$x_{k+1}=\psi(x_k)=\frac{\phi(x_k)-\omega x_k}{1-\omega}.$$

通常取$\omega=\phi'(\bar{x})$, 其中$\bar{x}$ 是x^*的一个好的近似值（例如$\bar{x}$用二分法求得. ），得迭代格式

$$x_{k+1}=\frac{\phi(x_k)-\phi'(\bar{x})\,x_k}{1-\phi'(\bar{x})}.$$

虽然该迭代格式的$\psi'(x^*)\neq 0$ ，但$|\psi'(x^*)|<|\phi'(x^*)|$. 这就大大的提高了收敛速度.

2. 艾特肯加速法

设迭代点列$\{x_k\}$线性收敛于x^*. 这时

$$\frac{x^*-x_{k+1}}{x^*-x_k}\approx\frac{x^*-x_k}{x^*-x_{k-1}}.$$

由此可以解出$x^*\approx x_{k+1}-\frac{(x_{k+1}-x_k)^2}{x_{k+1}-2x_k+x_{k-1}}$.

取上式的右端作为x^*的新近似值. 即令

$$\bar{x}_{k+1}=x_{k+1}-\frac{(x_{k+1}-x_k)^2}{x_{k+1}-2x_k+x_{k-1}}.$$

则$\bar{x}_{k+1}$比x_{k+1}更接近于x^*. 称为**艾特肯（Aitken）加速法**，它是利用迭代过程中相邻的三点x_{k-1},x_k,x_{k+1}推出x^*的更好近似值的一种方法.

由艾特肯加速法可得**斯特芬森(Steffensen)加速技术**.其计算过程如下:

给定 x_0, ε

$$for \quad k = 0 \ to \ K \ do$$

$$x_{3k+1} = \phi(x_{3k}), \quad x_{3k+2} = \phi(x_{3k+1})$$

$$\bar{x} = x_{3k+2} - \frac{(x_{3k+2} - x_{3k+1})^2}{x_{3k+2} - 2x_{3k+1} + x_{3k}}$$

如果 $|\bar{x} - x_{3k+2}| < \varepsilon$, 则取 $x^* \approx \bar{x}$, 停

否则 $x_{3(k+1)} := \bar{x}$

$$end \ do$$

$$\begin{cases} f_1(x_1, x_2, \cdots, x_n) = 0, \\ f_2(x_1, x_2, \cdots, x_n) = 0, \\ \vdots \\ f_n(x_1, x_2, \cdots, x_n) = 0. \end{cases}$$

其中$f_i(x_1, x_2, \cdots, x_n)$是定义在区域$D \subset R^n$上的实值函数,
$f_i(x_1, x_2, \cdots, x_n)$ 中至少有一个是非线性函数.
记$\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$, $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \cdots, f_n(\mathbf{x}))^T$.
则可表示成

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}.$$

求解非线性代数方程组的简单迭代法

改写为同解方程组：

$$x_i = \phi_i(x_1, x_2, \cdots, x_n), \quad i = 1, 2, \cdots, n.$$

建立迭代格式

$$x_i^{(k+1)} = \phi_i(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}), \quad i = 1, 2, \cdots, n.$$

选取初始点$\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)})^T$，令$k = 1, 2, \cdots$，可得向量迭代点列$\{\mathbf{x}^{(k)}\}$, 如果方程组有唯一解$\mathbf{x}^*$, 且$\{\mathbf{x}^{(k)}\}$收敛，则$\{\mathbf{x}^{(k)}\}$收敛于$\mathbf{x}^*$. 此方法称为**简单迭代法**.

改写为同解方程组：

$$x_i = \phi_i(x_1, x_2, \cdots, x_n), \quad i = 1, 2, \cdots, n.$$

建立迭代格式

$$x_i^{(k+1)} = \phi_i(x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)}), \quad i = 1, 2, \cdots, n.$$

选取初始点$\mathbf{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \cdots, x_n^{(0)})^T$，令$k = 1, 2, \cdots$，可得向量迭代点列$\{\mathbf{x}^{(k)}\}$，如果方程组有唯一解$\mathbf{x}^*$，且$\{\mathbf{x}^{(k)}\}$收敛，则$\{\mathbf{x}^{(k)}\}$收敛于$\mathbf{x}^*$. 此方法称为**简单迭代法**.

记

$$\phi(\mathbf{x}) = (\phi_1(\mathbf{x}), \phi_2(\mathbf{x}), \cdots, \phi_n(\mathbf{x}))^T, \quad \mathbf{x}^{(k)} = (\mathbf{x}_1^{(k)}, \mathbf{x}_2^{(k)}, \cdots, \mathbf{x}_n^{(k)})^T.$$

$\phi(\mathbf{x})$称为迭代函数，得向量形式

$$\mathbf{x}^{(k+1)} = \phi(\mathbf{x}^{(k)}), \quad k = 0, 1, \cdots,$$

与单个方程不动点的定义相同，向量函数的不动点定义为：
若$\mathbf{x}^*$满足$\mathbf{x}^* = \phi(\mathbf{x}^*)$，则称$\mathbf{x}^*$是$\phi(\mathbf{x})$的不动点. 不动点$x^*$满足方程$\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$，即$\mathbf{x}^*$是方程组的解.
类似于求解线性方程组的高斯-赛德尔迭代法，求解非线性方程组的**高斯-赛德尔迭代格式**为

$$x_i^{(k+1)} = \phi_i(x_1^{(k+1)}, \cdots, x_{i-1}^{(k+1)}, x_i^{(k)}, \cdots, x_n^{(k)}), \quad i = 1, 2 \cdots, n.$$

例7.2 解方程组

$$\begin{cases} 4x_1 - x_2 + 0.1e^{x_1} = 1, \\ -x_1 + 4x_2 + x_1^2/8 = 0. \end{cases}$$

解　用简单迭代法作迭代格式

$$\begin{cases} x_1^{(k+1)} = \frac{1}{4}\left(1 + x_2^{(k)} - 0.1e^{x_1^{(k)}}\right), \\ x_2^{(k+1)} = \frac{1}{4}\left(x_1^{(k)} - \frac{1}{8}(x_1^{(k)})^2\right). \end{cases} \qquad k = 0, 1, 2, \cdots$$

高斯-赛德尔迭代格式

$$\begin{cases} x_1^{(k+1)} = \frac{1}{4}\left(1 + x_2^{(k)} - 0.1e^{x_1^{(k)}}\right), \\ x_2^{(k+1)} = \frac{1}{4}\left(x_1^{(k+1)} - \frac{1}{8}(x_1^{(k+1)})^2\right). \end{cases} \qquad k = 0, 1, 2, \cdots$$

取$\mathbf{x}^{(0)} = (0, 0)^T$，得P230

设向量函数$\phi(\mathbf{x})$的所有分量函数$\phi_i(\mathbf{x})$在点$\mathbf{x}$处可微，矩阵

$$J_\phi(\mathbf{x}) = \begin{pmatrix} \frac{\partial \phi_1(\mathbf{x})}{\partial x_1} & \frac{\partial \phi_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial \phi_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial \phi_2(\mathbf{x})}{\partial x_1} & \frac{\partial \phi_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial \phi_2(\mathbf{x})}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial \phi_n(\mathbf{x})}{\partial x_1} & \frac{\partial \phi_n(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial \phi_n(\mathbf{x})}{\partial x_n} \end{pmatrix}$$

称为向量函数$\phi(\mathbf{x})$在$\mathbf{x}$处的导数或在$\mathbf{x}$处的**雅克比(Jacobi)矩阵**.

Definition

设向量点列$\{\mathbf{x}^{(k)}\}$收敛于$\mathbf{x}^*$，若存在常数$p \geq 1$和$c > 0$，使

$$\lim_{k\to\infty} \frac{\|\mathbf{x}^* - \mathbf{x}^{(k+1)}\|}{\|\mathbf{x}^* - \mathbf{x}^{(k)}\|^p} = c,$$

则称$\{\mathbf{x}^{(k)}\}$ p阶收敛于$\mathbf{x}^*$.

Theorem

全局收敛性定理 设$\phi : D \subset R^n \to R^n$，在闭区域$D_0 \subset D$满足

1) $\phi(\mathbf{x}) \subset D_0, \forall\ \mathbf{x} \in D_0$;
2) 存在常数$L(0 < L < 1)$使得$\|\phi(\mathbf{x}) - \phi(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|$, $\forall\ \mathbf{x}, \mathbf{y} \in D_0$;

则以下结论成立：

1) ϕ在D_0上存在唯一的不动点$\mathbf{x}^*$;
2) 对任意的初始点$\mathbf{x}^{(0)} \in D_0$，由简单迭代法产生的迭代点列$\{\mathbf{x}^{(k)}\} \subset D_0$线性收敛于$\mathbf{x}^*$，且有误差估计式：

$$\|\mathbf{x}^* - \mathbf{x}^{(k)}\| \leq \frac{L^k}{1-L}\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|.$$

$$\|\mathbf{x}^* - \mathbf{x}^{(k)}\| \leq \frac{L}{1-L}\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|.$$

求解非线性代数方程组的简单迭代法

条件（1）表明ϕ 将区域D_0映入自身，条件（2）称为ϕ的压缩条件. 若ϕ是压缩的，则它是连续的. 该定理也称为**压缩映射原理**.

Theorem

局部收敛性定理　设映射ϕ在其定义域内有不动点$\mathbf{x}^*$, ϕ的分量有连续的偏导数且矩阵$\mathbf{J}_\phi(\mathbf{x}^*)$的谱半径

$$\rho(\mathbf{J}_\phi(\mathbf{x}^*)) < 1.$$

则存在$\mathbf{x}^*$的一个邻域$D_\delta = \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}^*\| < \delta, \delta > 0\} \subset D_0$，对任意的初始点$\mathbf{x}^{(0)} \in D_\delta$，由简单迭代法产生的迭代点列$\{\mathbf{x}^{(k)}\} \subset D_\delta$收敛于$\mathbf{x}^*$.

求解非线性代数方程组的牛顿法

仿照单个方程求根的牛顿法，将方程组的每个方程在点$\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})^T$ 处按多元函数的泰勒公式展开有

$$f_i(\mathbf{x}) = f_i(\mathbf{x}^{(k)}) + \nabla f_i(\mathbf{x}^{(k)})^T(\mathbf{x} - \mathbf{x}^{(k)}) = 0, \quad i = 1, 2, \cdots, n.$$

取线性项，由$f_i(\mathbf{x}) = 0$得近似方程组

$$f_i(\mathbf{x}^{(k)}) + \nabla f_i(\mathbf{x}^{(k)})^T(\mathbf{x} - \mathbf{x}^{(k)}) = 0, \quad i = 1, 2, \cdots, n.$$

则

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - J_f^{-1}(\mathbf{x}^{(k)})\mathbf{f}(\mathbf{x}^{(k)}).$$

称为求解非线性方程组的**牛顿法**.

求解非线性代数方程组的牛顿法

求解非线性方程组的牛顿法的计算过程如下：

1. 在$\mathbf{x}^*$的附近选取$\mathbf{x}^{(0)} \in D$，给定允许误差限ε_1，ε_2和最大迭代次数K；
2. *for* $k = 0$ *to* K *do*

   1) 计算 $\mathbf{f}(\mathbf{x}^{(k)})$，$\mathbf{J}_f(\mathbf{x}^{(k)})$；

   2) 求解关于$\Delta\mathbf{x}^{(k)}$的线性方程组

   $$\mathbf{J}_f(\mathbf{x}^{(k)})\Delta\mathbf{x}^{(k)} = -f(\mathbf{x}^{(k)}).$$

   3) 计算 $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)}$.

   4) 若 $\|\Delta\mathbf{x}^{(k)}\|/\|\mathbf{x}^{(k)}\| < \varepsilon_1$， 或者 $\|\mathbf{f}(\mathbf{x}^{(k+1)})\| < \varepsilon_2$ 则取 $\mathbf{x}^* \approx \mathbf{x}^{(k+1)}$，停止计算；

   如果$k = K$，输出K次迭代不满足精度要求的信息，停机；
3. *end do*

例7.4 用牛顿法解$\begin{cases} 4x_1 - x_2 + 0.1e^{x_1} = 1, \\ -x_1 + 4x_2 + x_1^2/8 = 0. \end{cases}$.

解 牛顿迭代格式为

$$\begin{pmatrix} 4 + 0.1e^{x_1^{(k)}} & -1 \\ -1 + 0.25x_1^{(k)} & 4 \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x}_1^{(k)} \\ \Delta \mathbf{x}_2^{(k)} \end{pmatrix} = \begin{pmatrix} -4x_1^{(k)} - 0.1e^{x_1^{(k)}} + x_2^{(k)} + 1 \\ x_1^{(k)} - 0.125(x_1^{(k)})^2 - 4x_2^{(k)} \end{pmatrix}$$

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} + \Delta x_1^{(k)}, \\ x_2^{(k+1)} = x_2^{(k)} + \Delta x_2^{(k)}. \end{cases}$$

例7.4 用牛顿法解$\begin{cases} 4x_1 - x_2 + 0.1e^{x_1} = 1, \\ -x_1 + 4x_2 + x_1^2/8 = 0. \end{cases}$.

解 牛顿迭代格式为

$$\begin{pmatrix} 4+0.1e^{x_1^{(k)}} & -1 \\ -1+0.25x_1^{(k)} & 4 \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x}_1^{(k)} \\ \Delta \mathbf{x}_2^{(k)} \end{pmatrix} = \begin{pmatrix} -4x_1^{(k)} - 0.1e^{x_1^{(k)}} + x_2^{(k)} + 1 \\ x_1^{(k)} - 0.125(x_1^{(k)})^2 - 4x_2^{(k)} \end{pmatrix}$$

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} + \Delta x_1^{(k)}, \\ x_2^{(k+1)} = x_2^{(k)} + \Delta x_2^{(k)}. \end{cases}$$

取$\mathbf{x}^{(0)} = (0,0)^T$，得

$$\begin{aligned} \mathbf{x}^{(1)} &= (0.233\,766\,233\,8, \quad 0.058\,441\,558\,4)^T \\ \mathbf{x}^{(3)} &= (0.232\,567\,005\,1, \quad 0.056\,451\,519\,7)^T = \mathbf{x}^{(4)} \end{aligned}$$

故$\mathbf{x}^* \approx \mathbf{x}^{(3)}$.

求解非线性代数方程组的牛顿法

Theorem

牛顿法的局部收敛性定理　设$\mathbf{x}^* \in int(D)$是方程组$\mathbf{f}(\mathbf{x}) = 0$的解，$\mathbf{f}(\mathbf{x})$在包含$\mathbf{x}^*$的某个开区域$D_0 \subset D$内连续可微且$\mathbf{J}_f(\mathbf{x}^*)$ 非奇异，则存在闭球

$$\bar{B}(\mathbf{x}^*; \delta) = \{\mathbf{x} \big| \|\mathbf{x} - \mathbf{x}^*\| \le \delta,\ \delta > 0\} \subset D_0,$$

使得对任意的$\mathbf{x}^{(0)} \in \bar{B}(\mathbf{x}^*; \delta)$,由牛顿法产生的迭代点列$\{\mathbf{x}^{(k)}\} \subset \bar{B}(\mathbf{x}^*; \delta)$且超线性收敛于$\mathbf{x}^*$.
进而，若$\mathbf{f}(\mathbf{x})$在D_0上二阶连续可微，则点列$\{\mathbf{x}^{(k)}\}$至少二阶收敛.

牛顿法的优点是收敛速度快，一般能达到二阶收敛速度. 不足之处是每次迭代都要计算一个由n^2个偏导数构成的雅克比矩阵$\mathbf{J}_f(\mathbf{x}^{(k)})$，且要解一个线性方程组，当$n$较大时,计算量巨大.

求解非线性代数方程组的简化牛顿法

为了减少计算量，将牛顿迭代格式中的雅克比矩阵$\mathbf{J}_f(\mathbf{x}^{(k)})$用常数矩阵$\mathbf{J}_f(\mathbf{x}^{(0)})$替代，得**简化牛顿法**的迭代格式：

$$\begin{cases} \mathbf{J}(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(k)} = -\mathbf{f}(\mathbf{x}^{(k)}), \\ \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Delta\mathbf{x}^{(k)}. \end{cases}$$

在迭代开始前，可先将线性方程组的系数矩阵$\mathbf{J}(\mathbf{x}^{(0)})$做三角分解，以后每次迭代只需解两个三角方程组，从而可以大大地减少每次迭代的计算量. 简化牛顿法的收敛速度一般比牛顿法低.

为了避免牛顿法中计算$f_i(\mathbf{x})$ $(i=1,2,\cdots,n)$的偏导数和n^2个偏导数函数值，用常数矩阵A_k替代牛顿法中的雅可比矩阵$\mathbf{J}_f(\mathbf{x}^{(k)})$.
得到如下的**布洛顿(Broyden)**方法的迭代公式：

$$\mathbf{x}^{(k+1)}=\mathbf{x}^{(k)}-A_k^{-1}\mathbf{f}(\mathbf{x}^{(k)}).$$

1. 矩阵$\mathbf{A}_k$应满足的条件

取　$\mathbf{f}(\mathbf{x})\approx\mathbf{f}(\mathbf{x}^{(k)})+\mathbf{J}_f(\mathbf{x}^{(k)})(\mathbf{x}-\mathbf{x}^{(k)})$,则

$$\mathbf{f}(\mathbf{x}^{(k-1)})\approx\mathbf{f}(\mathbf{x}^{(k)})+\mathbf{J}_f(\mathbf{x}^{(k)})(\mathbf{x}^{(k-1)}-\mathbf{x}^{(k)}).$$

用　$\mathbf{A}_k$替代上式中的$\mathbf{J}_f(\mathbf{x}^{(k)})$，　并令

$$\mathbf{f}(\mathbf{x}^{(k-1)})=\mathbf{f}(\mathbf{x}^{(k)})+A_k(\mathbf{x}^{(k-1)}-\mathbf{x}^{(k)}).$$

由此可得

$$\mathbf{A}_k(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}) = \mathbf{f}(\mathbf{x}^{(k)}) - \mathbf{f}(\mathbf{x}^{(k-1)}).$$

记 $\mathbf{S}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}$, $\mathbf{y}^{(k)} = \mathbf{f}(\mathbf{x}^{(k)}) - \mathbf{f}(\mathbf{x}^{(k-1)})$.

则有$\mathbf{A}_k\mathbf{s}^{(k)} = \mathbf{y}^{(k)}$.

2. 构造矩阵$\mathbf{A}_k$

令 $\mathbf{A}_k = \mathbf{A}_{k-1} + \Delta\mathbf{A}_{k-1} = \mathbf{A}_{k-1} + \mathbf{u}\mathbf{s}^{(k)T}$, 其中$\mathbf{u} \in R^n$, $\mathbf{s}^{(k)} \in R^n$. $\Delta\mathbf{A}_{k-1}$称为修正矩阵. 下面讨论向量$\mathbf{u}$的取法.

$$\mathbf{y}^{(k)} = \mathbf{A}_k\mathbf{s}^{(k)} = (\mathbf{A}_{k-1} + \mathbf{u}\mathbf{s}^{(k)T})\mathbf{s}^{(k)} = \mathbf{A}_{k-1}\mathbf{s}^{(k)} + \mathbf{s}^{(k)T}\mathbf{s}^{(k)}\mathbf{u}.$$

由此可得

$$\mathbf{u} = \frac{\mathbf{y}^{(k)} - \mathbf{A}_{k-1}\mathbf{s}^{(k)}}{\mathbf{s}^{(k)T}\mathbf{s}^{(k)}}.$$

$$\mathbf{A}_k = \mathbf{A}_{k-1} + \mathbf{u}\mathbf{s}^{(k)T} = \mathbf{A}_{k-1} + \frac{(\mathbf{y}^{(k)} - \mathbf{A}_{k-1}\mathbf{s}^{(k)})\mathbf{s}^{(k)T}}{\mathbf{s}^{(k)T}\mathbf{s}^{(k)}}.$$

3. 求$\mathbf{A}_k^{-1}$

Theorem

（**谢尔曼-莫里森, Sherman-Morrison**） 设$\mathbf{A}$是n阶可逆矩阵，$\mathbf{u},\mathbf{v} \in R^n$是任意向量，若$1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u} \neq 0$，则 $\mathbf{A}+\mathbf{u}\mathbf{v}^T$ 是可逆矩阵，其逆矩阵为$(\mathbf{A}+\mathbf{u}\mathbf{v}^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1+\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}$.

利用谢尔曼-莫里森定理求矩阵$\mathbf{A}_k$的逆矩阵时，$\mathbf{u}$的取法同前，取$\mathbf{v} = \mathbf{s}^{(k)}$,由定理可求得矩阵$\mathbf{A}_k$的逆矩阵为

$$\mathbf{A}_k^{-1} = \mathbf{A}_{k-1}^{-1} + \frac{(\mathbf{s}^{(k)} - \mathbf{A}_{k-1}^{-1}\mathbf{y}^{(k)})\,\mathbf{s}^{(k)T}\mathbf{A}_{k-1}^{-1}}{\mathbf{s}^{(k)T}\mathbf{A}_{k-1}^{-1}\mathbf{y}^{(k)}} .$$

布洛顿算法的计算步骤

1. 给定初始点$\mathbf{x}^{(0)} \in D$，允许误差限ε_1，ε_2和最大迭代次数K；取初始矩阵$\mathbf{A}_0 = \mathbf{J}_f(\mathbf{x}^{(0)})$，计算$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - \mathbf{A}_0^{-1}\mathbf{f}(\mathbf{x}^{(0)})$;
2. *for* $k = 1$ *to* K *do*

   $\mathbf{s}^{(k)} := \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}$;

   $\mathbf{y}^{(k)} := \mathbf{f}(\mathbf{x}^{(k)}) - \mathbf{f}(\mathbf{x}^{(k-1)})$;

   $$\mathbf{A}_k^{-1} := \mathbf{A}_{k-1}^{-1} + \frac{(\mathbf{s}^{(k)} - \mathbf{A}_{k-1}^{-1}\mathbf{y}^{(k)})\,\mathbf{s}^{(k)T}\mathbf{A}_{k-1}^{-1}}{\mathbf{s}^{(k)T}\mathbf{A}_{k-1}^{-1}\mathbf{y}^{(k)}} \,.$$

   $\mathbf{x}^{(k+1)} := \mathbf{x}^{(k)} - \mathbf{A}_k^{-1}\mathbf{f}(\mathbf{x}^{(k)})$;

   如果 $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| < \varepsilon_1$ 或者 $\|\mathbf{f}(\mathbf{x}^{(k+1)})\| < \varepsilon_2$

   取$\mathbf{x}^* \approx \mathbf{x}^{(k+1)}$，停；

   如果$k = K$，输出K次迭代不满足精度要求的信息，停机；

   end do

布洛顿算法在一定的条件下具有超线性收敛速度

例： 用布洛顿法求解如下方程组的近似解，准确到两位小数.

$$x_1^2 + x_2^2 - 1 = 0, \qquad 2x_1 + x_2 - 1 = 0.$$

解： 第1步迭代

令$\mathbf{x}^{(0)} = (0.6, -0.8)^T, \mathbf{f}(\mathbf{x}^{(0)}) = (0, -1.6)^T),$

$$\mathbf{J}_f(\mathbf{x}) = \begin{pmatrix} 2x_1 & 2x_2 \\ 2 & 1 \end{pmatrix} \quad A_0 = \mathbf{J}_f(\mathbf{x}^{(0)}) = \begin{pmatrix} 1.2 & -1.6 \\ 2 & 1 \end{pmatrix}$$

$$A_0^{-1} = \mathbf{J}_f(\mathbf{x}^{(0)})^{-1} = \frac{1}{22} \begin{pmatrix} 5 & 8 \\ -10 & 6 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - A_0^{-1}\mathbf{f}(\mathbf{x}^{(0)}) = \begin{pmatrix} 0.6 \\ -0.8 \end{pmatrix} - \frac{1}{22} \begin{pmatrix} 5 & 8 \\ -10 & 6 \end{pmatrix} \begin{pmatrix} 0 \\ -0.6 \end{pmatrix}$$

$$= (\tfrac{9}{11}, -\tfrac{7}{11})^T$$

第2步迭代

$$\mathbf{s}^{(1)} = \mathbf{x}^{(1)} - \mathbf{x}^{(0)} = (\tfrac{12}{55}, \tfrac{9}{55})^T,$$

$$\mathbf{y}^{(1)} = \mathbf{f}(\mathbf{x}^{(1)}) - \mathbf{f}(\mathbf{x}^{(0)}) = (\tfrac{9}{121}, \tfrac{3}{5})^T$$

$$\mathbf{A}_1^{-1} := \mathbf{A}_0^{-1} + \frac{(\mathbf{s}^{(1)} - \mathbf{A}_0^{-1}\mathbf{y}^{(1)})\,\mathbf{s}^{(1)T}\mathbf{A}_0^{-1}}{\mathbf{s}^{(1)T}\mathbf{A}_0^{-1}\mathbf{y}^{(1)}} = \frac{1}{2596}\begin{pmatrix} 605 & 869 \\ -1210 & 858 \end{pmatrix}$$

$$\mathbf{x}^{(2)} := \mathbf{x}^{(1)} - \mathbf{A}_1^{-1}\mathbf{f}(\mathbf{x}^{(1)}) = (0.8008, -0.6017)^T$$

第2步迭代

$$\mathbf{s}^{(1)} = \mathbf{x}^{(1)} - \mathbf{x}^{(0)} = (\tfrac{12}{55}, \tfrac{9}{55})^T,$$

$$\mathbf{y}^{(1)} = \mathbf{f}(\mathbf{x}^{(1)}) - \mathbf{f}(\mathbf{x}^{(0)}) = (\tfrac{9}{121}, \tfrac{3}{5})^T$$

$$\mathbf{A}_1^{-1} := \mathbf{A}_0^{-1} + \frac{(\mathbf{s}^{(1)} - \mathbf{A}_0^{-1}\mathbf{y}^{(1)})\,\mathbf{s}^{(1)T}\mathbf{A}_0^{-1}}{\mathbf{s}^{(1)T}\mathbf{A}_0^{-1}\mathbf{y}^{(1)}} = \frac{1}{2596}\begin{pmatrix} 605 & 869 \\ -1210 & 858 \end{pmatrix}$$

$$\mathbf{x}^{(2)} := \mathbf{x}^{(1)} - \mathbf{A}_1^{-1}\mathbf{f}(\mathbf{x}^{(1)}) = (0.8008, -0.6017)^T$$

第3步迭代

$$\mathbf{s}^{(2)} = \mathbf{x}^{(2)} - \mathbf{x}^{(1)} = \cdots,$$

$$\mathbf{y}^{(2)} = \mathbf{f}(\mathbf{x}^{(2)}) - \mathbf{f}(\mathbf{x}^{(1)}) = \cdots$$

$$\mathbf{A}_2^{-1} := \mathbf{A}_1^{-1} + \frac{(\mathbf{s}^{(k)} - \mathbf{A}_{k-1}^{-1}\mathbf{y}^{(k)})\,\mathbf{s}^{(k)T}\mathbf{A}_{k-1}^{-1}}{\mathbf{s}^{(k)T}\mathbf{A}_{k-1}^{-1}\mathbf{y}^{(k)}} = \cdots$$

$$\mathbf{x}^{(3)} := \mathbf{x}^{(2)} - \mathbf{A}_2^{-1}\mathbf{f}(\mathbf{x}^{(2)}) = \cdots$$