7 非线性方程和非线性方程组的数值解
Solutions of Nonlinear Equations and System of Linear Equations

王家兵
jbwang@scut.edu.cn

背景

- 求方程$f(x) = 0$的根是最常见的数学问题之一
  - 当$f(x)$是一次多项式时，称$f(x) = 0$为线性方程
  - 否则称之为非线性方程.
- 当$f(x) = 0$是非线性方程时，由于$f(x)$的多样性，求方程$f(x) = 0$的根尚无一般的解析解法可用
  - 在满足一定的精度要求下，求出方程的近似根
- 可根据物理背景或者$y = f(x)$的草图等方法确定求根区间$[a, b]$或给出某根的近似值x_0

背景

- 一元二次、三次、四次方程的求根公式
- 一元五次方程的求根公式?
  - 伽罗瓦(Évariste Galois， 1811.10.25-1832.5.31)

7.1 对分法

- 对分法适用于求有根区间的单实根或奇重实根
  - 设$f(x)$在$[a,b]$上连续，$f(a)\cdot f(b)<0$，即$f(a)>0,f(b)<0$或者$f(a)<0,f(b)>0$，则根据连续函数的介值定理，在(a,b)内至少存在一点ξ，使$f(\xi)=0$.
- 对分法：
  - 首先取区间$[a,b]$的中点$x_1=\frac{a+b}{2}$，把区间$[a,b]$分为两个小区间$[a,x_1],[x_1,b]$
    - 如果$f(a)\cdot f(x_1)<0$，令$a_1=a,b_1=x_1$
    - 如果$f(x_1)\cdot f(b)<0$，令$a_1=x_1,b_1=b$
  - 重复此过程，得到长度每次减半的有根区间序列$[a_k,b_k]$

□ 收敛性分析

- 当k足够大时，有根区间$[a_k, b_k]$长度趋于零，区间任一点都可以作为根的近似值.

$$|x_k - \xi| \le \frac{1}{2}(b_k - a_k) = \frac{1}{2^2}(b_{k-1} - d_{k-1}) = \cdots$$

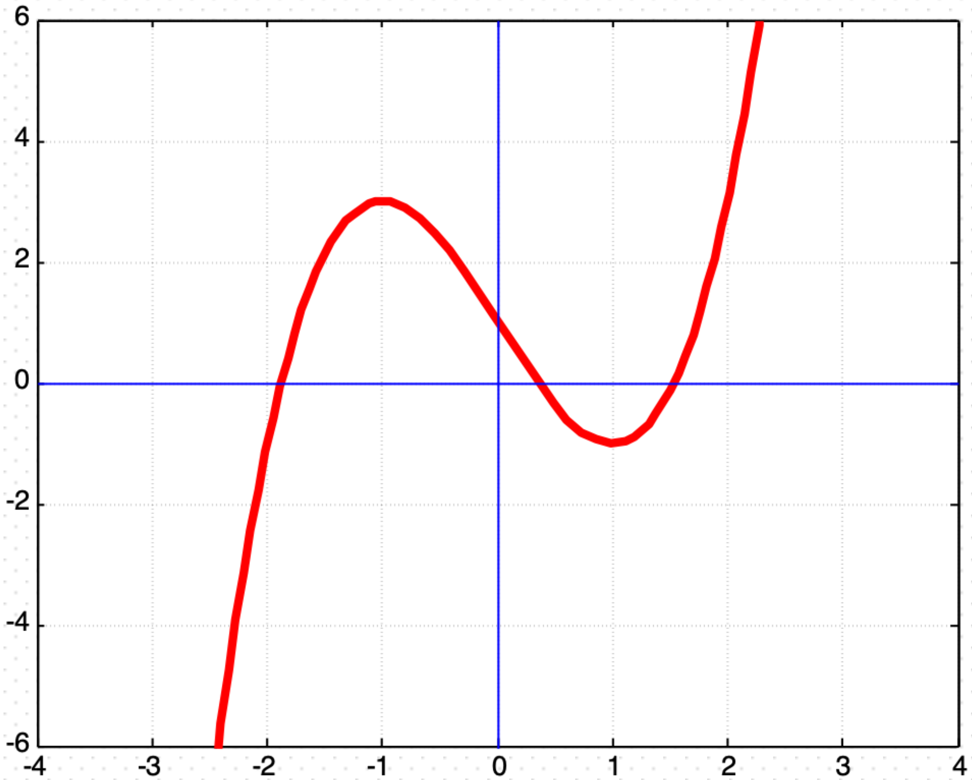
$$= \frac{1}{2^{k+1}}(b - a)$$

  - □ 当$k \to \infty$时，x_k收敛于ξ

- 对分法的收敛速度较慢，常用来试探实根的分布区间或者求根的初始近似值.
  - □ 对分法的收敛速度与公比为$\frac{1}{2}$的等比级数相同
  - □ 大约每对分10次，近似根精度可提高三位小数位

7.1 对分法

例：求$x^3 - 3x + 1 = 0$的实根分布情况，并求[0,1]中的实根近似值．要求实根近似值精确到三位小数.



7.1 对分法

- 对分法优点：

1. 对函数的要求低（只要求$f(x)$连续），方法简单、可靠，程序设计容易
2. 可事先估计计算次数

$$|x_k - \xi| \le \frac{1}{2^{k+1}}(b-a) < \epsilon \Rightarrow k >?$$

- 对分法缺点：
- 收敛速度较慢，不能求出偶重根
  - $f(x)$在偶重根两边附近函数值同号

7.1 对分法

- 对分法算法

①给出精度δ, ϵ令$a_0 = a, b_0 = b, k = 0$

②令$x_k = (a_k + b_k)/2$,计算$f(x_k)$

③若$|f(x_k)| < \delta$，则x_k是$f(x) = 0$的根，停止计算，输出结果$\xi = x_k$

若$f(a_k) \cdot f(x_k) < 0$，则令$a_{k+1} = a_k, b_{k+1} = x_k$;

否则令$a_{k+1} = x_k, b_{k+1} = b_k$

④若$b_{k+1} - a_{k+1} \le \epsilon$，退出计算，输出结果$\xi = (a_{k+1} + b_{k+1})/2$;反之，令$k = k + 1$，返回②

7.2 迭代法

- 给定方程$f(x)=0$，可用多种方法构造等价方程$x=\varphi(x)$，取定根的一个近似值x_0，构造序列

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

1. 若$\{x_k\}$收敛，即$\lim\limits_{k\to\infty} x_k=x^*$，并且$\varphi(x)$连续，则有：

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

即$\{x_k\}$的极限x^*是方程$x=\varphi(x)$的根，也就是方程$f(x)=0$的根.

2. 若$\{x_k\}$发散，迭代法失败.

7.2 迭代法

例： 求$x^3 - 3x + 1 = 0$在0.5附近的根.要求精确到第四位小数.

1. 构造等价方程$x = \varphi_1(x) = \frac{x^3+1}{3}$，迭代公式 $x_{k+1} = \frac{x_k^3+1}{3}, k = 0,1,2,\cdots$， 取$x_0 = 0.5$
2. 迭代公式$x_{k+1} = \frac{x_k^3+1}{3}$不能求出方程在1.5和−2附近的根
3. 其他的等价方程构造的迭代公式

$$x_{k+1} = \varphi_2(x_k) = \frac{1}{3 - x_k^2}(k = 0,1,2,\cdots)$$

$$\boldsymbol{x_{k+1} = \varphi_3(x_k) = \sqrt[3]{3x_k - 1} \quad k = 0, 1, 2, \cdots}$$

7.2 迭代法

- 迭代法的几何意义
  - 将$x = \varphi(x)$写成

$$\begin{cases} y = x \\ y = \varphi(x) \end{cases}$$

则求解方程$x = \varphi(x)$等价于求直线$y = x$和曲线$y = \varphi(x)$的交点的横坐标x^*

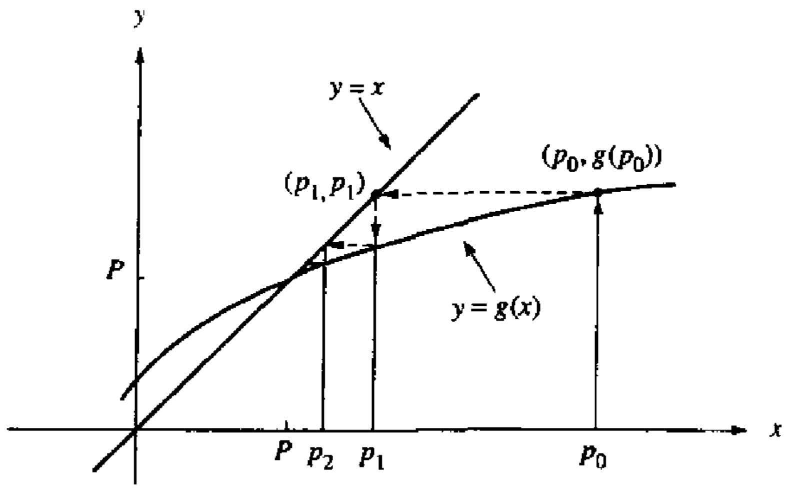


Figure 2.4 (a) Monotone convergence when $0 < g'(P) < 1$.

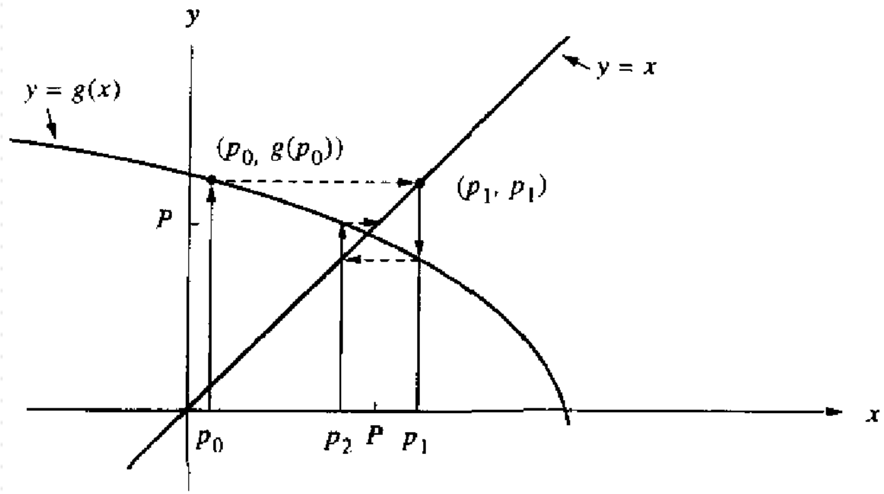


Figure 2.4 (b) Oscillating convergence when $-1 < g'(P) < 0$.

□ 迭代法的几何意义

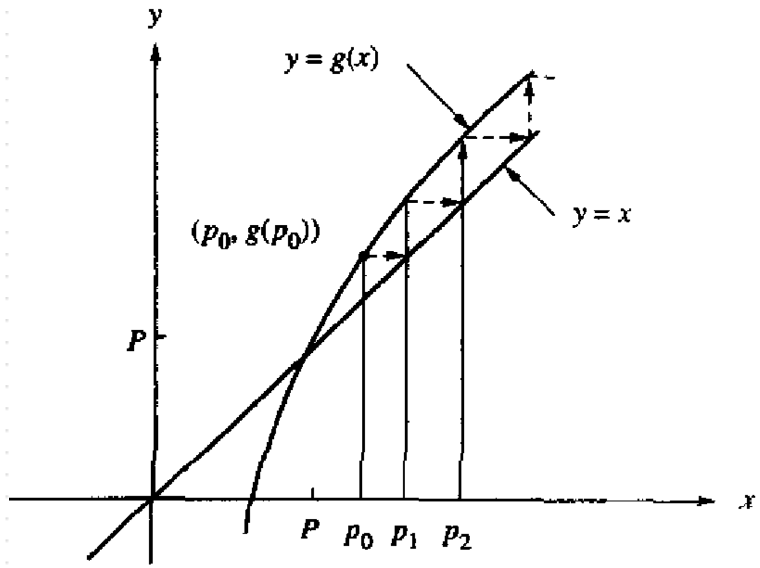


Figure 2.5 (a) Monotone divergence when $1 < g'(P)$.

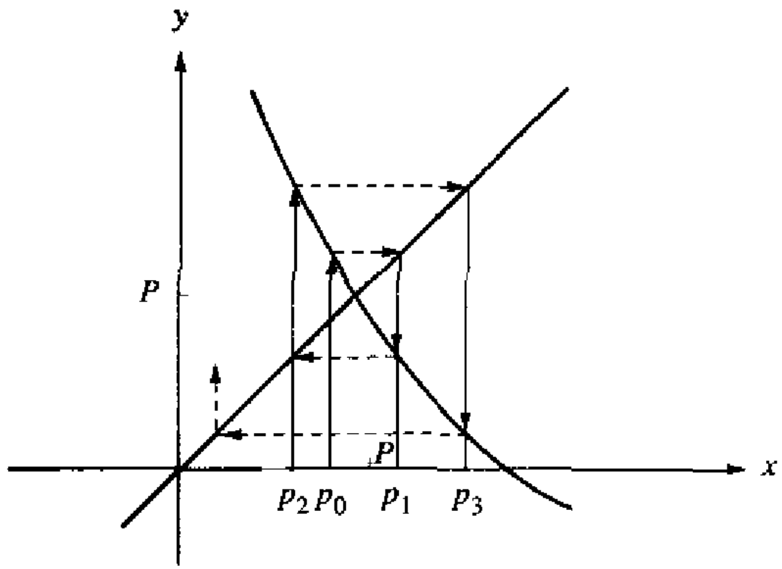


Figure 2.5 (b) Divergent oscillation when $g'(P) < -1$.

7.2 迭代法

- 迭代法收敛条件

定义：如果在根x^*的某个邻域$\boldsymbol{R}$: $|x-x^*|\le\delta$中，对任意的$x_0\in\boldsymbol{R}$，迭代公式$x_{k+1}=\varphi(x_k), k=0,1,2\cdots$收敛，则称迭代在$x^*$附近局部收敛.

定理：设$x^*=\varphi(x^*)$，$\varphi'(x)$在x^*的某个邻域$\boldsymbol{R}$内连续，并且$|\varphi'(x)|\le q, q<1$是常量，则

1. 对任意$x_0\in\boldsymbol{R}$，由迭代公式$x_{k+1}=\varphi(x_k)$决定的序列$\{x_k\}$收敛于x^*
2. $|x_k-x^*|\le\frac{1}{1-q}|x_{k+1}-x_k|$
3. $|x_k-x^*|\le\frac{q^k}{1-q}|x_1-x_0|$

- (1) 由拉格朗日中值定理有

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

- 递推有: $|x_k - x^*| \le q\,|x_{k-1} - x^*| \le q^2\,|x_{k-2} - x^*| \le \ldots \le q^k\,|x_0 - x^*|$
- 所以 $\lim\limits_{k\to\infty} x_k = x^*$
- (2) $x_k - x^* = x_k - x_{k+1} + x_{k+1} - x^* = x_k - x_{k+1} + \varphi'(\xi)(x_k - x^*)$
- 故: $|x_k - x^*| \le |x_k - x_{k+1}| + q\,|x_k - x^*|$

$$\Rightarrow (1-q)\,|x_k - x^*| \le |x_k - x_{k+1}| = |\varphi(x_k) - \varphi(x_{k-1})|$$

$$= |\varphi'(\xi)(x_k - x_{k-1})| \le q\,|x_k - x_{k-1}| \le \ldots \le q^k\,|x_1 - x_0|$$

$$\Rightarrow |x_k - x^*| \le \frac{q^k}{1-q}\,|x_1 - x_0|$$

定理：给定方程$x=\varphi(x)$，若$\varphi(x)$满足：

①对任意的$x\in[a,b]$，有$\varphi(x)\in C[a,b]$

②对任意的$x,y\in[a,b]$，有$|\varphi(x)-\varphi(y)|\leq q|x-y|$，$0\leq q<1$为常数，则对任意的$x_0\in[a,b]$，迭代公式$x_{k+1}=\varphi(x_k)$生成的序列$\{x_k\}$收敛于$x=\varphi(x)$的根$x^*$.

注：满足条件①②的函数$\varphi(x)$称为区间的压缩映射.

7.2 迭代法

- 证明:

$$|x_k - x^*| = |\varphi(x_{k-1}) - \varphi(x^*)| \le q\,|x_{k-1} - x^*|$$

$$= q\,|\varphi(x_{k-2}) - \varphi(x^*)| \le q^2\,|x_{k-2} - x^*| \le \ldots \le q^k\,|x_0 - x^*|$$

- 所以 $\lim\limits_{k\to\infty} x_k = x^*$

例：方程$x^3-3x+1=0$，三种迭代公式

$$x_{k+1}=\varphi_1(x_k)=\frac{x_k^3+1}{3}$$

$$x_{k+1}=\varphi_2(x_k)=\frac{1}{3-x_k^2}$$

$$x_{k+1}=\varphi_3(x_k)=\sqrt[3]{3x_k-1}$$

在三个根附近的收敛情况，三个根近似取$x_1=0.347, x_2=1.53, x_3=-1.88$

7.2 迭代法

迭代法的加速

- 由$f(x)=0$构造出的迭代公式$x=\varphi(x)$收敛时，收敛速度取决于$|\varphi'(x)|$的大小，当$|\varphi'(x)|$接近与1时，收敛可能很慢.
- 松弛法：已知$\varphi(x_k)$与x_k同是x^*的近似值，两个近似值的加权平均

$$x_{k+1}=(1-\omega_k)x_k+\omega_k\varphi(x_k)$$

$$x=(1-\omega)x+\omega\varphi(x)=\phi(x)$$

- 令$\phi'(x)=1-\omega+\omega\varphi'(x)=0$，当$\varphi'(x)\neq 1$时，得到

$$\omega=\frac{1}{1-\varphi'(x)}$$

7.2 迭代法

- 取$\omega_k = \frac{1}{1-\varphi'(x_k)}$, $1-\omega_k = \frac{-\varphi'(x_k)}{1-\varphi'(x_k)}$可望获得较好的加速效果.

- 松弛法迭代公式

$$\begin{cases} \omega_k = \frac{1}{1-\varphi'(x_k)} \\ x_{k+1} = (1-\omega_k)x_k + \omega_k\varphi(x_k) \end{cases}$$

- 埃特金Aitken方法
  - 松弛法要计算导数$\varphi'(x_k)$，使用中有时不方便
  - 设方程$x=\varphi(x), x^*$是它的准确解，x_0是近似根，取 $x_1=\varphi(x_0), x_2=\varphi(x_1)$

$$x^*=x_2+x^*-x_2=x_2+\varphi'(\xi)(x^*-x_1)$$

用差商

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

代替$\varphi'(\xi)$，得到:

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

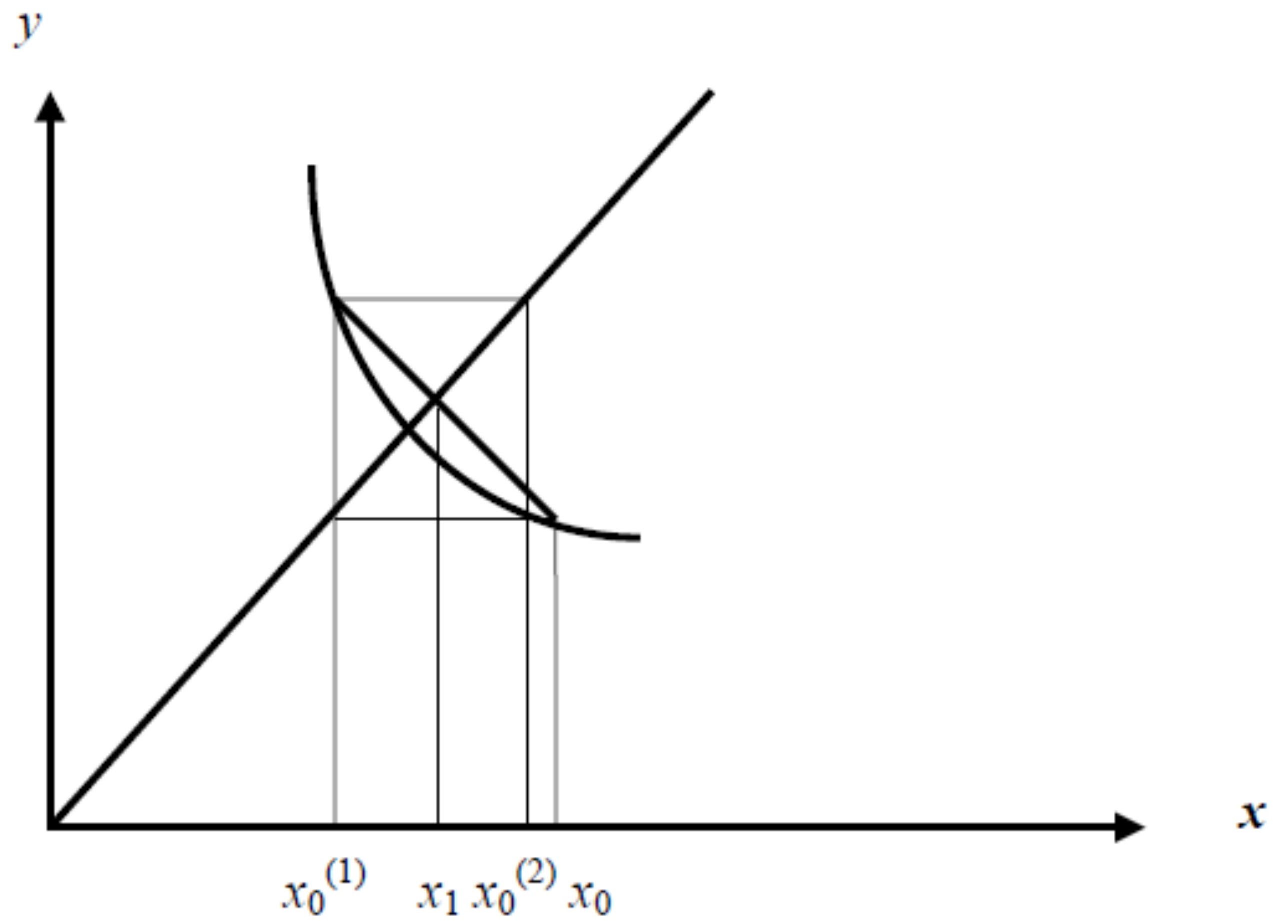
上式解得：$x^* \approx x_2 - \frac{(x_2 - x_1)^2}{x_2 - 2x_1 + x_0}$

- 埃特金迭代公式

$$
\begin{cases}
x_k^{(1)} = \varphi(x_k) \\
x_k^{(2)} = \varphi(x_k^{(1)}) \\
x_{k+1} = x_k^{(2)} - \dfrac{\left(x_k^{(2)} - x_k^{(1)}\right)^2}{x_k^{(2)} - 2x_k^{(1)} + x_k} \quad (k = 0,1,\cdots)
\end{cases}
$$

7.2 迭代法

- 埃特金迭代的几何解释: 点$(x_0, x_0^{(1)})$与点$(x_0^{(1)}, x_0^{(2)})$连线与$y = x$交点的横坐标



例：解方程$x^3-3x+1=0$，迭代公式$x_{k+1}=\frac{x_k^3+1}{3}$，用松弛法和埃特金法，取$x_0=0.5$，精确到6位小数.

松弛法：$\varphi(x)=\frac{x^3+1}{3}, \varphi'(x)=x^2, \omega_k=\frac{1}{1-x_k^2}$

$$x_{k+1}=(1-\omega_k)x_k+\omega_k\frac{x_k^3+1}{3}$$

7.2 迭代法

例：解方程$x^3-3x+1=0$，迭代公式$x_{k+1}=\frac{x_k^3+1}{3}$，用松弛法和埃特金法，取$x_0=0.5$，精确到6位小数.

埃特金法：

$$\begin{cases} x_k^{(1)}=\dfrac{x_k^3+1}{3} \\ x_k^{(2)}=\dfrac{\left(x_k^{(1)}\right)^3+1}{3} \\ x_{k+1}=x_k^{(2)}-\dfrac{\left(x_k^{(2)}-x_k^{(1)}\right)^2}{x_k^{(2)}-2x_k^{(1)}+x_k}\quad(k=0,1,2,\cdots) \end{cases}$$

7.3 牛顿法

牛顿公式

- 由于$f(x)=0$是非线性方程，一般解决非线性问题较困难，而解决线性问题较容易,可以考虑将非线性问题线性化，采取线性化方法求解.
- 将$f(x)$在x_0处泰勒展开

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(\xi)}{2!}(x-x_0)$$

- 用$p(x)=f(x_0)+f'(x_0)(x-x_0)=0$来近似$f(x)=0$，即把求曲线$f(x)=0$与$X$轴交点的横坐标近似为求直线$p(x)=0$与$X$轴交点的横坐标.

$$p(x)=f(x_0)+f'(x_0)(x-x_0)=0 \Longrightarrow$$

$$\boldsymbol{x=x_0-\frac{f(x_0)}{f'(x_0)}} \quad \text{设}\boldsymbol{f'(x_0)\neq 0}$$

7.3 牛顿法

- 牛顿法实质是一般迭代法用松弛法加速

$$f(x) = 0 \Leftrightarrow x = x + f(x) = \varphi(x)$$

  - 使用松弛法

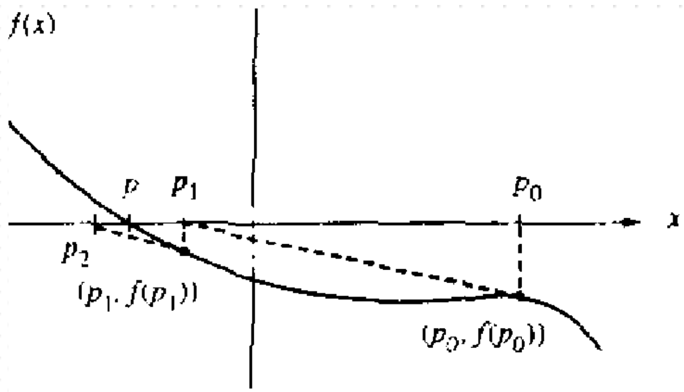
$$\varphi'(x) = 1 + f'(x)$$

$$\omega_k = \frac{1}{1-\varphi'(x_k)} = -\frac{1}{f'(x_k)}$$

  - 即得牛顿法公式

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

- 几何意义



7.3 牛顿法

- 牛顿法公式

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

$$\vdots$$

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

例：用牛顿法计算$x^3 - 3x + 1 = 0$在0.5和−2附近的两个根.

$$x_{k+1} = x_k - \frac{x_k^3 - 3x_k + 1}{3x_k^2 - 3}$$

牛顿法的收敛速度

定义： 设序列$\{x_k\}$收敛于x^*，令$\epsilon_k = x^* - x_k$，设$k \to \infty$时，有

$$\frac{|\epsilon_{k+1}|}{|\epsilon_k|^p} \to c \ (c > 0\text{为常数})$$

则称序列$\{x_k\}$是p阶收敛.

- 当$p = 1$时，称为线性收敛
- 当$p = 2$时，称为二阶收敛（几何收敛）
- 当$1 < p < 2$时，称为超线性收敛

7.3 牛顿法

定理：设$x^* = \varphi(x^*)$，在x^*的某个邻域$\boldsymbol{R}$内$\varphi^{(p)}(x)$连续，$p > 1$是常量，并且满足

$$\varphi^{(l)}(x^*) = 0 \quad (l = 1,2,\cdots,p-1)$$

$$\varphi^{(p)}(x^*) \neq 0$$

则由$x_{k+1} = \varphi(x_k)$生成的序列$\{x_k\}$收敛于x^*，并且序列$\{x_k\}$是p阶收敛.

7.3 牛顿法

- 因 $\varphi'(x^*)=0$，所以序列$\{x_k\}$收敛于x*。
- 再由泰勒展开，其中ξ_k在x_k和x*之间:

$$x_{k+1}=\varphi(x_k)=\varphi(x^*)+\varphi'(x^*)(x_k-x^*)$$
$$+\ldots+\frac{\varphi^{(p-1)}(x^*)}{(p-1)!}(x_k-x^*)^{p-1}+\frac{\varphi^{(p)}(\xi_k)}{p!}(x_k-x^*)^p$$

- 所以有

$$\frac{e_{k+1}}{e_k^p}=\frac{\varphi^{(p)}(\xi_k)}{p!}$$

- 当$k\to\infty$, $x_k\to x^*$，从而$\xi_k\to x^*$，定理成立。

牛顿法的示例

- 试用牛顿迭代法导出下列各式的迭代格式，并讨论其收敛性：
  - (1) 不使用除法运算 $\frac{1}{c}$；
  - (2) 不使用开方运算求 $\sqrt{c}$

7.3 牛顿法

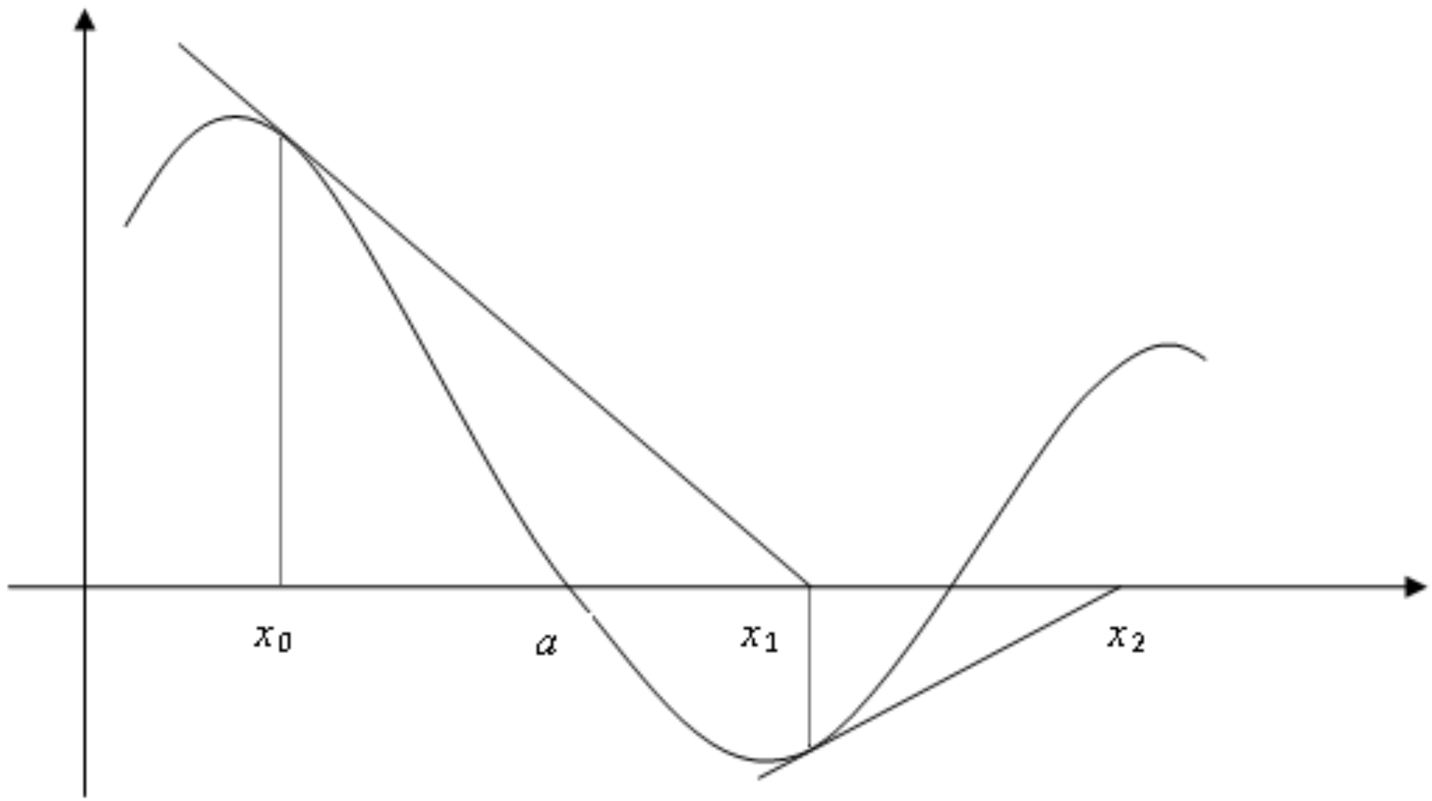
- 牛顿法的收敛速度

$$x = x - \frac{f(x)}{f'(x)} = \varphi(x),\quad \varphi'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}$$

  - 当$f'(x^*) \neq 0$时，$\varphi'(x^*) = 0$，即牛顿法在单根附近至少二阶收敛，而在重根附近，牛顿法是线性收敛.

- 牛顿法的优点：收敛快，算法简单，是常用的快速收敛法
- 缺点：对重根收敛较慢，对初值x_0要求较严，要求x_0相当接近真实解x^*
- 在实际应用中，常与其他方法联用，现用其他方法确定初值x_0，再用牛顿法提高精度.

注意问题



当初始点x_0与根$x=a$附件有极值点或拐点时，迭代过程就可能发散或收敛于另一个根。

为此，应用中常把隔根区间取得如此之小，使在其中$f'(x),f''(x)$不变号，并取x_0使得, $f'(x_0)f''(x_0)>0$。

牛顿法求重根

- 设$f(x)=(x-x^*)^m g(x), g(x^*)\neq 0$，则$x^*$是$f(x)$的m重根。此时有$f(x^*)=f'(x^*)=\ldots=f^{(m-1)}(x^*)=0$, $f^{(m)}(x^*)\neq 0$。

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

- $\varphi'(x^*)=1-1/m\neq 0$, 因此**线性收敛**。

- 若取 $\varphi(x)=x-m\dfrac{f(x)}{f'(x)}$，即 $x_{k+1}=x_k-m\dfrac{f(x_k)}{f'(x_k)}$

则具有**二阶收敛**，但要知道m。

牛顿法求重根

- 也可以：令 $\mu(x)=\dfrac{f(x)}{f'(x)}$， 则有下式成立，故x*是μ(x)=0的单根。

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

- 对μ(x)=0使用牛顿法，

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

- 得如下迭代格式，是**二阶收敛**的。

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

7.4 割线法

- 已知$f(x)=0$的两个近似根x_k, x_{k-1}，过点$(x_{k-1}, f(x_{k-1})), (x_k, f(x_k))$连一条直线，据两点式可写出直线方程

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

- 把求曲线$f(x)=0$与X轴交点的横坐标近似为求割线与X轴交点的横坐标，即将割线与X轴交点的横坐标作为x^*的近似值

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

称为割线法

7.4 割线法

- 割线法需要两个初值x_0, x_1，在单根附近是超线性收敛.
  - 通过较复杂的推导可知收敛阶p在单根附近为$1.618\cdots$
  - 收敛较快，而且不用计算导数值

定理：如$f(x)$在零点x^*附近有连续的2阶导数，$f'(x^*) \neq 0$，且初值x_0, x_1充分接近x^*，则割线法收敛，收敛速度为

$$|x_{k+1} - x^*| \approx \left|\frac{f''(x^*)}{2f'(x^*)}\right|^{0.618} |x_k - x^*|^{1.618}$$

7.4 割线法

- 割线法需要两个初值x_0, x_1，也称为双点割线法.
  - 单点割线法.

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

- 例：用双点割线法求$x^3 - 3x + 1 = 0$在0.5附近的根，精确到小数点后第六位.

7.5 解非线性方程组的迭代法和牛顿法

- 解一元非线性方程的迭代法和牛顿法可以推广到多元.
- n个未知数n个方程的非线性方程组可表示为

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

其中$\boldsymbol{X} = (x_1, x_2, \cdots, x_n)^T$为$n$维列向量； $f_i(\boldsymbol{X})(i = 1,2, \cdots, n)$中至少有一个是$\boldsymbol{X}$的非线性函数，并假设自变量和函数值都是实数.

7.5 解非线性方程组的迭代法和牛顿法

➢ 牛顿法—多元函数泰勒展开公式

$$f(x_1+\Delta x_1, x_2+\Delta x_2, \ldots, x_n+\Delta x_n)$$

$$=\sum_{i=0}^{\infty}\frac{1}{i!}(\Delta x_1\frac{\partial}{\partial x_1}+\Delta x_2\frac{\partial}{\partial x_2}+\ldots+\Delta x_n\frac{\partial}{\partial x_n})^i f(x_1, x_2, \ldots, x_n)$$

其中（二元为例）：

$$(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y})^m f(x_0, y_0)=\sum_{i=0}^{m}C_m^i\frac{\partial^m}{\partial x^i\partial y^{m-i}}h^i k^{m-i} f(x_0, y_0)$$

7.5 解非线性方程组的迭代法和牛顿法

二元函数的 Taylor 级数展开公式

$$f(x+\Delta x, y+\Delta y) = f(x,y) + \Delta x \frac{\partial f(x,y)}{\partial x} + \Delta y \frac{\partial f(x,y)}{\partial y}$$

$$+\frac{1}{2!}\left[(\Delta x)^2 \frac{\partial^2 f(x,y)}{\partial x^2} + 2\Delta x \Delta y \frac{\partial^2 f(x,y)}{\partial x \partial y} + (\Delta y)^2 \frac{\partial^2 f(x,y)}{\partial y^2}\right]$$

$$+\frac{1}{3!}\left[(\Delta x)^3 \frac{\partial^3 f(x,y)}{\partial x^3} + 3(\Delta x)^2 \Delta y \frac{\partial^3 f(x,y)}{\partial x^2 \partial y} + 3\Delta x (\Delta y)^2 \frac{\partial^3 f(x,y)}{\partial x \partial y^2} + (\Delta y)^3 \frac{\partial^3 f(x,y)}{\partial y^3}\right] + \cdots$$

$$= f(x,y) + \frac{1}{1!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right) f(x,y)$$

$$+\frac{1}{2!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^2 f(x,y) + \frac{1}{3!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^3 f(x,y) + \cdots$$

$$= \sum_{i=0}^{\infty} \frac{1}{i!}\left(\Delta x \frac{\partial}{\partial x} + \Delta y \frac{\partial}{\partial y}\right)^i f(x,y) \text{。}$$

多变量函数泰勒展开式：与一元函数一致形式

我们用如下记号：

$$\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n), \quad \beta = (\beta_1, \beta_2, \ldots, \beta_n) \qquad (\alpha_j, \beta_j \in \{0, 1, 2, \ldots\}).$$

$$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n, \qquad \alpha! = \alpha_1!\alpha_2! \cdots \alpha_n!,$$

$$\mathbf{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \text{ (where } \mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n\text{)},$$

$$\partial^\alpha f = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \cdots \partial_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}$$

EXAMPLE. With $n = 3$ and $\mathbf{x} = (x, y, z)$, we have

$$\partial^{(0,3,0)} f = \frac{\partial^3 f}{\partial y^3}, \qquad \partial^{(1,0,1)} f = \frac{\partial^2 f}{\partial x \partial z}, \qquad \mathbf{x}^{(2,1,5)} = x^2 y z^5.$$

多变量函数泰勒展开式：与一元函数一致形式

Theorem 2 (Taylor's Theorem in Several Variables). *Suppose* $f : \mathbb{R}^n \to \mathbb{R}$ *is of class* C^{k+1} *on an open convex set* S. *If* $\mathbf{a} \in S$ *and* $\mathbf{a} + \mathbf{h} \in S$, *then*

$$f(\mathbf{a}+\mathbf{h}) = \sum_{|\alpha| \le k} \frac{\partial^\alpha f(\mathbf{a})}{\alpha!}\mathbf{h}^\alpha + R_{\mathbf{a},k}(\mathbf{h}), \tag{3}$$

where the remainder is given in Lagrange's form by

$$R_{\mathbf{a},k}(\mathbf{h}) = \sum_{|\alpha|=k+1} \partial^\alpha f(\mathbf{a}+c\mathbf{h})\frac{\mathbf{h}^\alpha}{\alpha!} \text{ for some } c \in (0,1). \tag{4}$$

and in integral form by

$$R_{\mathbf{a},k}(\mathbf{h}) = (k+1)\sum_{|\alpha|=k+1} \frac{\mathbf{h}^\alpha}{\alpha!}\int_0^1 (1-t)^k \partial^\alpha f(\mathbf{a}+t\mathbf{h})\,dt. \tag{5}$$

7.5 解非线性方程组的迭代法和牛顿法

- 迭代法
  - 记$F(\boldsymbol{X})=[f_1(\boldsymbol{X}),f_2(\boldsymbol{X}),\cdots,f_n(\boldsymbol{X})]^T$，则方程组可以简写为

$$F(\boldsymbol{X})=\mathbf{0}$$

  - 将$F(\boldsymbol{X})=\mathbf{0}$转换成等价方程组

$$\boldsymbol{X}=\Phi(\boldsymbol{X})$$

其中 $\Phi(\boldsymbol{X})=[\varphi_1(\boldsymbol{X}),\varphi_2(\boldsymbol{X}),\cdots,\varphi_n(\boldsymbol{X})]^T$

  - 构造迭代公式

$$\boldsymbol{X}^{(k+1)}=\Phi\left(\boldsymbol{X}^{(k)}\right)\ (k=0,1,2,\cdots)$$

  - 对于给定的初始点$\boldsymbol{X}^{(0)}$，若由此生成的序列收敛，即$\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}=\boldsymbol{X}^*$，则$\boldsymbol{X}^*$是方程组$F(\boldsymbol{X})=\mathbf{0}$的解

例：设有非线性方程组

$$\begin{cases} x_1^2 - 10x_1 + x_2^2 + 8 = 0 \\ x_1 x_2^2 + x_1 - 10x_2 + 8 = 0 \end{cases}$$

迭代公式：

$$\begin{cases} x_1^{(k+1)} = \varphi_1\left(x_1^{(k)}, x_2^{(k)}\right) = \dfrac{1}{10}\left[\left(x_1^{(k)}\right)^2 + \left(x_2^{(k)}\right)^2 + 8\right] \\ x_2^{(k+1)} = \varphi_2\left(x_1^{(k)}, x_2^{(k)}\right) = \dfrac{1}{10}\left[x_1^{(k)}\left(x_2^{(k)}\right)^2 + x_1^{(k)} + 8\right] \end{cases}$$

定义： $\boldsymbol{R}^n$中集合M称为闭集，是指M中的任一个向量序列$\{\boldsymbol{X}^{(k)}\}$，若$\boldsymbol{X}^{(k)} \to \boldsymbol{X}^*$（即$\|\boldsymbol{X}^{(k)} - \boldsymbol{X}^*\| \to 0$），则 $\boldsymbol{X}^* \in \boldsymbol{R}^n$.

定义： 设$\Phi(\boldsymbol{X}) = [\varphi_1(\boldsymbol{X}), \varphi_2(\boldsymbol{X}), \cdots, \varphi_n(\boldsymbol{X})]^T$是一个$n$维列向量函数，$M$是$\boldsymbol{R}^n$中的子集合，若满足：

- $\boldsymbol{X} \in M \Longrightarrow \Phi(\boldsymbol{X}) \in \mathrm{M}$
- 存在常数$q, 0 \le q < 1$，使得M中任意$\boldsymbol{X}, \boldsymbol{Y}$，满足

$$\|\Phi(\boldsymbol{X}) - \Phi(\boldsymbol{Y})\| \le q\|\boldsymbol{X} - \boldsymbol{Y}\|$$

则称$\Phi(\boldsymbol{X})$为M上的压缩映射.

定理：若$\Phi(\boldsymbol{X})$为M上的压缩映射，则$\boldsymbol{X}=\Phi(\boldsymbol{X})$在$M$上有唯一解，且对任意$\boldsymbol{X}^{(0)}\in M$，由迭代公式$\boldsymbol{X}^{(k+1)}=\Phi\left(\boldsymbol{X}^{(k)}\right)$产生的向量序列$\{\boldsymbol{X}^{(k)}\}$收敛于这个解。

例：设有非线性方程组

$$\begin{cases} x_1^2-10x_1+x_2^2+8=0 \\ x_1x_2^2+x_1-10x_2+8=0 \end{cases}$$

迭代公式：

$$\begin{cases} x_1^{(k+1)}=\varphi_1\left(x_1^{(k)},x_2^{(k)}\right)=\dfrac{1}{10}\left[\left(x_1^{(k)}\right)^2+\left(x_2^{(k)}\right)^2+8\right] \\ x_2^{(k+1)}=\varphi_2\left(x_1^{(k)},x_2^{(k)}\right)=\dfrac{1}{10}\left[x_1^{(k)}\left(x_2^{(k)}\right)^2+x_1^{(k)}+8\right] \end{cases}$$

- Hessain矩阵

$$\boldsymbol{H}(\boldsymbol{X}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

- 矩阵形式的泰勒展开式：

$$f(X) = f(X_0) + \nabla f(X_0)^T \Delta X + \frac{1}{2!} \Delta X^T H(X_0) \Delta X + \cdots$$

- 牛顿法
  - 设$\boldsymbol{X}^*$是方程组$F(\boldsymbol{X}) = \mathbf{0}$的解，$\boldsymbol{X}^{(k)}$是某个值，用点$\boldsymbol{X}^{(k)}$处的一阶泰勒展开近似每一个分量函数量$f_i(\boldsymbol{X}^*) = 0$，有

$$0 = f_i(\boldsymbol{X}^*) \approx f_i\left(\boldsymbol{X}^{(k)}\right) + \sum_{j=1}^{n} \frac{\partial f_i\left(\boldsymbol{X}^{(k)}\right)}{\partial x_j}\left(x_j^* - x_j^{(k)}\right) \quad (i = 1, \cdots, n)$$

  - 矩阵和向量形式

$$0 = F(\boldsymbol{X}^*) \approx F\left(\boldsymbol{X}^{(k)}\right) + F'\left(\boldsymbol{X}^{(k)}\right)\left(\boldsymbol{X}^* - \boldsymbol{X}^{(k)}\right)$$

- 导数$F'(\boldsymbol{X})$是函数$F(\boldsymbol{X})$的Jacobi矩阵

$$F'(\boldsymbol{X}) = \begin{bmatrix} \nabla f_1(\boldsymbol{X})^T \\ \nabla f_2(\boldsymbol{X})^T \\ \vdots \\ \nabla f_n(\boldsymbol{X})^T \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_1(\boldsymbol{X})}{\partial x_1} & \dfrac{\partial f_1(\boldsymbol{X})}{\partial x_2} & \cdots & \dfrac{\partial f_1(\boldsymbol{X})}{\partial x_n} \\ \dfrac{\partial f_2(\boldsymbol{X})}{\partial x_1} & \dfrac{\partial f_2(\boldsymbol{X})}{\partial x_2} & \cdots & \dfrac{\partial f_2(\boldsymbol{X})}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial f_n(\boldsymbol{X})}{\partial x_1} & \dfrac{\partial f_n(\boldsymbol{X})}{\partial x_2} & \cdots & \dfrac{\partial f_n(\boldsymbol{X})}{\partial x_n} \end{bmatrix}$$

7.5 解非线性方程组的迭代法和牛顿法

- 如果矩阵$F'\left(\boldsymbol{X}^{(k)}\right)$非奇异，则

$$0=F(\boldsymbol{X}^{*})\approx F\left(\boldsymbol{X}^{(k)}\right)+F'\left(\boldsymbol{X}^{(k)}\right)\left(\boldsymbol{X}^{*}-\boldsymbol{X}^{(k)}\right)$$

$$\boldsymbol{X}^{(k+1)}=\boldsymbol{X}^{(k)}-F'\left(\boldsymbol{X}^{(k)}\right)^{-1}F\left(\boldsymbol{X}^{(k)}\right)$$

其中$k=0,1,\cdots,\boldsymbol{X}^{(0)}$是给定的初始值.

- 牛顿法的迭代函数

$$\Phi(\boldsymbol{X})=\boldsymbol{X}-F'(\boldsymbol{X})^{-1}F(\boldsymbol{X})$$

- 牛顿法

定义：设序列$\{\boldsymbol{X}^{(k)}\}_{k=0}^{\infty} \subset \boldsymbol{R}^n$收敛到$\boldsymbol{X}^*$，若有常数$p \geq 1$和$c > 0$，使得

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

则称p为该序列的收敛阶.

- 当$p = 1$时称为线性收敛
- 当$p > 1$时称为超线性收敛
- 当$p = 2$时称为二次收敛或平方收敛

7.5 解非线性方程组的迭代法和牛顿法

定理：对于函数$F: D \subset \boldsymbol{R}^n \to \boldsymbol{R}^n$. 设$\boldsymbol{X}^* \in D$满足$F(\boldsymbol{X}^*) = 0.$ 若有$\boldsymbol{X}^*$的开邻域$S_0 \subset D, F'(\boldsymbol{X})$在其上存在并连续，而且$F'(\boldsymbol{X}^*)$非奇异，则存在$\boldsymbol{X}^*$的闭球$S = S(\boldsymbol{X}^*, \delta) \subset S_0$（其中$\delta > 0$）

- 牛顿迭代函数$\Phi(\boldsymbol{X})$对所有$\boldsymbol{X} \in S$有定义，并且$\Phi(\boldsymbol{X}) \in S$
- 对于任何初始值$\boldsymbol{X}^{(0)} \in S$，牛顿序列$\{\boldsymbol{X}^{(k)}\}$超线性收敛于$\boldsymbol{X}^*$
- 若有常数$q > 0$，使得：

  $\|F'(\boldsymbol{X}) - F'(\boldsymbol{X}^*)\| \le q\|\boldsymbol{X} - \boldsymbol{X}^*\| \ \forall \boldsymbol{X} \in S,$

  则牛顿序列$\{\boldsymbol{X}^{(k)}\}$至少二阶收敛于$\boldsymbol{X}^*$

例：牛顿法求解方程组

$$\begin{cases} x_1^2 - 10x_1 + x_2^2 + 8 = 0 \\ x_1x_2^2 + x_1 - 10x_2 + 8 = 0 \end{cases}$$

1. $F(\boldsymbol{X}) = \begin{bmatrix} x_1^2 - 10x_1 + x_2^2 + 8 \\ x_1x_2^2 + x_1 - 10x_2 + 8 \end{bmatrix}$

2. $F'(\boldsymbol{X}) = \begin{bmatrix} 2x_1 - 10 & 2x_2 \\ x_2^2 + 1 & 2x_1x_2 - 10 \end{bmatrix}$

3. 取初值 $\boldsymbol{X}^{(0)} = (0,0\ )^T$


Thank You!