北京航空航天大學
BEIHANG UNIVERSITY

数值分析

主讲教师：贺慧霞

北京航空航天大学数学科学学院

$f(x)=0$

第四章 非线性方程（组）的迭代解法

4.2 非线性方程组迭代法

简代迭代法

牛顿法

一、非线性方程组的解法

含有n个方程的n元非线性方程组的一般形式为

$$\begin{cases} f_1(x_1,x_2,\cdots,x_n)=0, \\ f_2(x_1,x_2,\cdots,x_n)=0, \\ \quad\vdots \\ f_n(x_1,x_2,\cdots,x_n)=0, \end{cases} \Longrightarrow \text{化为：} F(X)=\vec{0}$$

$$X=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

其中$f_i(i=1,2,\cdots,n)$是定义在$D\subset R^n$上的n元实值函数，且f_i至少有一个是非线性的．令$X=(x_1,x_2,\cdots,x_n)^T$，

$F(X)=(f_1(X),f_2(X),\cdots,f_n(X))^T$．

二、简单迭代法

基本思想： 将方程组$F(X)=0$,写成与之等价的形式：$X=G(X)$,

然后再利用$X^{(k+1)}=G(X^{(k)})$, $k=1,2,\ldots$,求解原方程的根。

简单迭代法的收敛性

定理4.12（局部收敛性定理） 设$G:D\subset R^n\to R^n$, X^*是方程组$X=G(X)$的解,G在X^*处可微.若$G'(X^*)$的谱半径$\rho(G'(X^*))=\sigma<1$，则存在开球$S=S(X^*,\delta)\subset D$，对$\forall X^{(0)}\in S$，迭代序列$\{X^{(k)}\}$收敛于X^*.

定理2.8 对任意的向量d，迭代法$x^{(k+1)}=Gx^{(k)}+d$收敛的充分必要条件是$\rho(G)<1$. 线性方程组 Ax=b

非线性方程组$x=G(x)$,谱半径$\rho(G'(x^*))<1$是迭代法收敛的充分条件.

定理4.13(压缩映像原理) 设$G: D \subset \mathrm{R}^n \to \mathrm{R}^n$,在闭区域$D_0 \subset D$上满足

(1) G把D_0 映入它自身，即$G(D_0) \subset D_0$;

(2) 存在常数$L \in (0,1)$，使得对任意的$X, Y \in D_0$,

有$\| G(X) - G(Y) \| \le L \| X - Y \|$,

则有如下结论:

(1) 对任取的$X_0 \in D_0$,由$X^{k+1} = G(X^{(k)})$产生的序列$\{X^{(k)}\} \subset D_0$,

且收敛于方程组$F(X) = 0$在D_0内的唯一解X^*;

(3) 成立误差估计式$\left\| X^* - X_k \right\| \le \dfrac{L^k}{1-L} \left\| X_1 - X_0 \right\|$,

$$\left\| X^* - X_k \right\| \le \frac{L}{1-L} \left\| X_k - X_{k-1} \right\|.$$

回顾：非线性函数的简单迭代收敛定理

定理4.1：设函数$\varphi(x)\in C[a,b]$在(a,b)内可导，且满足如下条件

（1）当$x\in[a,b]$时，$\varphi(x)\in[a,b]$； 为保证 $x^{k+1}=\varphi(x^k)$ 产生的迭代序列 $\{x^k\}\subset[a,b]$

（2）当$x\in(a,b)$时，$|\varphi'(x)|\le L<1$其中L是一常数。

则有如下结论：

（1）方程$x=\varphi(x)$在区间$[a,b]$上有唯一的根s；

（2）对任取的$x_0\in[a,b]$，简单迭代法$x_{k+1}=\varphi(x_k)$产生的序列$\{x_k\}\subset[a,b]$且收敛于s；

（3）成立误差估计式$|s-x_k|\le\dfrac{L^k}{1-L}|x_1-x_0|$，$|s-x_k|\le\dfrac{L}{1-L}|x_k-x_{k-1}|$

注记：不满足定理的条件，也可能有不动点.

例　用简单迭代法求解下列方程组

$x=G(x)$

$$\begin{cases} 3x_1 - \cos x_1 - \sin x_2 = 0 \\ 4x_2 - \sin x_1 - \cos x_2 = 0 \end{cases}$$

要求满足精度 $\dfrac{\| \boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k-1)} \|_\infty}{\| \boldsymbol{x}^{(k)} \|_\infty} \leqslant 10^{-12}$。

解　把方程组写成等价形式：

$$\begin{cases} x_1 = \dfrac{1}{3}(\cos x_1 + \sin x_2) \\ x_2 = \dfrac{1}{4}(\sin x_1 + \cos x_2) \end{cases}$$

记 $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$，$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$，$\boldsymbol{G}(\boldsymbol{x}) = \begin{bmatrix} \dfrac{1}{3}(\cos x_1 + \sin x_2) \\ \dfrac{1}{4}(\sin x_1 + \cos x_2) \end{bmatrix}$

$f(x)-f(y)=f'(\xi)(x-y)$

则有 $\underline{G(\mathbf{R}^2)\subset\mathbf{R}^2}$，且对任意的

$\Uparrow$

$G(D_0)\subset D_0$

11. 设 $\mathbf{A}=[a_{ij}]\in\mathbf{R}^{n\times n}$，记

$$\|\mathbf{A}\|_F=\sqrt{\sum_{i,j=1}^{n}a_{ij}^2}$$

试证：$\|\cdot\|_F$ 是与向量范数 $\|\cdot\|_2$ 相容的矩阵范数。

$$G(x)-G(y)=\left[\begin{matrix}\cdots\\ \cdots\end{matrix}\right]\begin{pmatrix}x_1-y_1\\ x_2-y_2\end{pmatrix}$$

$$\|A\|_F=\sqrt{\frac{1}{9}(\sin^2\xi_1+\cos^2\xi_2)+\frac{1}{16}(\cos^2\eta_1+\sin^2\eta_2)}\le\sqrt{\frac{2}{9}+\frac{2}{16}}=\sqrt{\frac{2}{9}+\frac{1}{8}}=\sqrt{\frac{25}{72}}$$

其中 $\mathbf{A}=\begin{bmatrix}-\frac{1}{3}\sin\xi_1 & \frac{1}{3}\cos\xi_2\\ \frac{1}{4}\cos\eta_1 & -\frac{1}{4}\sin\eta_2\end{bmatrix}$；$\xi_1,\eta_1$ 在 x_1 与 y_1 之间；ξ_2,η_2 在 x_2 与 y_2 之间。于是有

$$\underline{\|\mathbf{G}(\mathbf{x})-\mathbf{G}(\mathbf{y})\|_2}\leqslant\|\mathbf{A}\|_F\|\mathbf{x}-\mathbf{y}\|_2\leqslant\sqrt{\frac{25}{72}}\|\mathbf{x}-\mathbf{y}\|_2$$

$\leqslant L\|x-y\|_2$ $\quad$ $=\|A(x-y)\|_2\le\|A\|_F\|x-y\|_2$

$L<1$

因此，对任取的 $\boldsymbol{x}^{(0)} \in \mathbf{R}^2$，迭代公式

$$\begin{cases} x_1^{(k+1)} = \dfrac{1}{3}(\cos x_1^{(k)} + \sin x_2^{(k)}) \\ x_2^{(k+1)} = \dfrac{1}{4}(\sin x_1^{(k)} + \cos x_2^{(k)}) \\ \qquad (k = 0,1,\cdots) \end{cases}$$

产生的序列 $\{\boldsymbol{x}^{(k)}\}$ 必收敛于所给方程组在 $\mathbf{R}^2$ 中的唯一解 $\boldsymbol{x}^*$。

取 $x_1^{(0)} = x_2^{(0)} = 1$，则当 $k = 28$ 时 满足精度要求，得

$$x_1 \approx 0.415\ 169\ 427\ 139$$
$$x_2 \approx 0.336\ 791\ 217\ 025$$

向量值函数的偏导数

设$D \subset R^n$，向量值函数$f: D \to R^m, f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{pmatrix}$.

设点$x = (x_1, x_2, \cdots, x_n)^T \in D$,

如果每个分量函数$f_k(x), (k = 1, 2, \cdots, n)$在$x_0$关于$x_i$可偏导，

则称$f(x)$在x_0关于x_i可偏导,并且

$$\frac{\partial f}{\partial x_i} = \left(\frac{\partial f_1}{\partial x_i}, \frac{\partial f_2}{\partial x_i}, \cdots, \frac{\partial f_n}{\partial x_i} \right)^T$$

例　球面$x^2 + y^2 + z^2 = R^2$.　球坐标表示

$$\vec{r}(u, v) = (R\cos u \cos v, R\cos u \sin v, R\sin u)^T,$$

$$\vec{r}(u,v)=(R\cos u\cos v,R\cos u\sin v,R\sin u)^T,$$

$$\vec{r}_u(u,v)=(R(-\sin u)\cos v,R(-\sin u)\sin v,R\cos u)^T,$$

$$\vec{r}_v(u,v)=(R\cos u(-\sin v),R\cos u\cos v,0)^T,$$

定义：设点$x_0=(x_1,x_2,\cdots,x_n)\in D$,如果存在$m\times n$阶矩阵$A=(a_{ij})_{m\times n}$,

使得 $f(x_0+\Delta x)-f(x_0)=A\Delta x+r(\Delta x)$, $\lim\limits_{\|\Delta x\|\to 0}\dfrac{\|r(\Delta x)\|}{\|\Delta x\|}=0$,

则称f在x_0点可微，并称$A\Delta x$为f在x_0点的微分，记做

$$\mathrm{d}f(x_0)=A\,\mathrm{d}\,x.$$

$$\mathrm{d}f(x_0)=(df_1(x_0),df_2(x_0),\cdots,df_n(x_0))^T.$$

定理

向量值函数$f:D\to R^m$在x_0可微的充要条件是它的分量函数$f_i(x)$ $(i=1,2,\cdots,m)$在x_0可微.

$df(x_0)=$

f在x_0点的Jacobi矩阵 $$\mathrm{J}f(x_0)\cancel{dx}=\begin{pmatrix}\frac{\partial f_1(x_0)}{\partial x_1} & \cdots & \frac{\partial f_1(x_0)}{\partial x_n}\\ \vdots & \ddots & \vdots\\ \frac{\partial f_m(x_0)}{\partial x_1} & \cdots & \frac{\partial f_m(x_0)}{\partial x_n}\end{pmatrix}_{mn}.\begin{pmatrix}dx_1\\ dx_2\\ \vdots\\ dx_n\end{pmatrix}$$

并且 $\mathrm{d}f(x_0)=A\mathrm{d}x=\mathrm{J}f(x_0)\mathrm{d}x.$

$=(f_{x_1}, f_{x_2} \quad f_{x_n})$ $\quad f_{x_1}=\begin{pmatrix}(f_1)_{x_1}\\ (f_2)_{x_1}\\ \vdots\\ (f_n)_{x_1}\end{pmatrix}$

例2

求向量值函数$f(x,y,z)=\begin{pmatrix} x^3+ze^y \\ y^3+z\ln x \\ z^3+x\ln y \end{pmatrix}$在(1,1,1)处的*Jacobi*阵和微分.

f在x_0点的Jacobi矩阵

$$\mathrm{J}f(1,1,1)=\begin{pmatrix} 3x^2 & ze^y & e^y \\ \dfrac{z}{x} & 3y^2 & \ln y \\ \ln y & \dfrac{x}{y} & 3z^2 \end{pmatrix}_{(1,1,1)}=\begin{pmatrix} 3 & e & e \\ 1 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix}$$

$$\mathrm{d}f(1,1,1)=\begin{pmatrix} 3 & e & e \\ 1 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix}\begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \end{pmatrix}.$$

设$F: D \subset R^n \to R^n$是一个向量值函数，如果存在n维向量X*∈ D,使得F(X*)=0,则称X*是F(X)=0的根.

$$F'(X) = \left(\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \cdots, \frac{\partial F}{\partial x_n}\right)$$

$$F'(X) = \left(\frac{\partial f_i(X)}{\partial x_j}\right)_{n\times n} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

$F(X)$在X点的**Jacobi矩阵.**

例2.5.2. 求向量值函数

$$f(x,y)=\left(x^2+y\sin x, x^2\ln y, e^y\cos xy\right)$$

在$(1,\pi)$点的导数.

解2.5.1. 这时坐标分量分别为

$$f_1(x,y)=x^2+y\sin x,\quad f_2(x,y)=x^2\ln y,\quad f_3(x,y)=e^y\cos x,$$

$$J_f(1,\pi)=\begin{bmatrix}\dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y}\\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y}\\ \dfrac{\partial f_3}{\partial x} & \dfrac{\partial f_3}{\partial y}\end{bmatrix}_{(1,\pi)}=\begin{bmatrix}2x+y\cos x & \sin x\\ 2x\ln y & \dfrac{x^2}{y}\\ -e^y\sin x & e^y\cos x\end{bmatrix}_{(1,\pi)}=\begin{bmatrix}2+\pi\cos 1 & \sin 1\\ 2\ln\pi & \dfrac{1}{\pi}\\ -e^\pi\sin 1 & e^\pi\cos 1\end{bmatrix}$$

例2.5.3. 求向量值函数

$$f(x,y)=\left(x^2+y\sin x, x^2\ln y, e^y\cos xy\right)$$

在$(1,\pi)$点的微分.

解2.5.2. 由例*2.5.2*可知,

$$J_f(1,\pi)=\begin{bmatrix} 2+\pi\cos 1 & \sin 1 \\ 2\ln\pi & \dfrac{1}{\pi} \\ -e^\pi\sin 1 & e^\pi\cos 1 \end{bmatrix}.$$

所以微分为

$$df(1,\pi)=\begin{bmatrix} 2+\pi\cos 1 & \sin 1 \\ 2\ln\pi & \dfrac{1}{\pi} \\ -e^\pi\sin 1 & e^\pi\cos 1 \end{bmatrix}\begin{bmatrix} dx \\ dy \end{bmatrix}=\begin{bmatrix} (2+\pi\cos 1)dx+\sin 1dy \\ 2\ln\pi dx+\dfrac{1}{\pi}dy \\ -e^\pi\sin 1dx+e^\pi\cos 1dy \end{bmatrix}$$

定理 设$D \subset R^n, f:D \to R^m$是$D$到$R^m$上的映射.则$f$是$D$上连续映射的充要条件是$f$的每个分量$f_i(x_1,\cdots,x_n)$都在$D$上连续.

定理 设$D \subset R^n, f:D \to R^m$是$D$到$R^m$上的映射.则$f$是$D$上连续可微映射的充要条件是$f$的每个分量$f_i(x_1,\cdots,x_n)$都在$D$上连续可微.

定义：设$F : D \subset R^n \to R^n$是一个向量值函数，如果存在n维向量X*∈ D,使得F(X*)=0,则称X*是F(X)=0的**根**.

定义 设向量序列X_k收敛于X^*，并且$e_k = X^* - X_k \neq 0(k = 0,1,\ldots)$，如果存在常数$r \geq 1$和常数$C > 0$，使得极限 $\lim\limits_{k\to+\infty}\dfrac{\|e_{k+1}\|}{\|e_k\|^r} = C$成立，

或者使得当$k \geq K$(某个正数）时，$\dfrac{\|e_{k+1}\|}{\|e_k\|^r} \leq C$成立，

则称向量序列X_k收敛于X^*具有r阶收敛速度，简称X_k是r阶收敛的。C称作渐近收敛常数或者收敛因子.

三、Newton迭代法

多元函数$f(X)=f(x_1,x_2,\cdots,x_n)$在$X^{(0)}$点Taylor展开

$$f(X)\approx f(X^{(0)})+\sum_{j=1}^{n}\frac{\partial f(X^{(0)})}{\partial x_j}(x_j-x_j^{(0)}),$$

则$f(X)=0$可用线性方程近似逼近

$$f(X^{(k)})+\sum_{j=1}^{n}\frac{\partial f(X^{(k)})}{\partial x_j}(x_j-x_j^{(k)})=0$$

多元函数的全微分 $\Leftrightarrow \Delta f(x_1,\cdots,x_n)\approx df=\frac{\partial f}{\partial x_1}dx_1+\cdots\frac{\partial f}{\partial x_n}dx_n$

$f(x,y,z)$在$X^{(k)}=(x^{(k)},y^{(k)},z^{(k)})$点Taylor展开

$$f(X)=f(x,y,z)$$

$$\approx f(X^{(k)})+\frac{\partial f}{\partial x}\Big|_{X^{(k)}}(x-x^{(k)})+\frac{\partial f}{\partial y}\Big|_{X^{(k)}}(y-y^{(k)})+\frac{\partial f}{\partial z}\Big|_{X^{(k)}}(z-z^{(k)})$$

所以 $f(X)=0$可近似地用线性方程

$$f(X^{(k)})+\frac{\partial f}{\partial x}\Big|_{X^{(k)}}(x-x^{(k)})+\frac{\partial f}{\partial y}\Big|_{X^{(k)}}(y-y^{(k)})+\frac{\partial f}{\partial z}\Big|_{X^{(k)}}(z-z^{(k)})=0$$ 近似逼近.

向量值函数 $F(X)=(f(X),g(X),h(X)),X=(x,y,z)^T$, 则

$$g(X)\approx g(X^{(k)})+\frac{\partial g}{\partial x}\Big|_{X^{(k)}}(x-x^{(k)})+\frac{\partial g}{\partial y}\Big|_{X^{(k)}}(y-y^{(k)})+\frac{\partial g}{\partial z}\Big|_{X^{(k)}}(z-z^{(k)}).$$

$$h(X)\approx h(X^{(k)})+\frac{\partial h}{\partial x}\Big|_{X^{(k)}}(x-x^{(k)})+\frac{\partial h}{\partial y}\Big|_{X^{(k)}}(y-y^{(k)})+\frac{\partial h}{\partial z}\Big|_{X^{(k)}}(z-z^{(k)}).$$

向量值函数 $F(X)=(f(X),g(X),h(X))=0$可用线性方程组

$$\begin{cases} f(X^{(k)})+\dfrac{\partial f}{\partial x}\Big|_{X^{(k)}}(x-x^{(k)})+\dfrac{\partial f}{\partial y}\Big|_{X^{(k)}}(y-y^{(k)})+\dfrac{\partial f}{\partial z}\Big|_{X^{(k)}}(z-z^{(k)})=0 \\ g(X^{(k)})+\dfrac{\partial g}{\partial x}\Big|_{X^{(k)}}(x-x^{(k)})+\dfrac{\partial g}{\partial y}\Big|_{X^{(k)}}(y-y^{(k)})+\dfrac{\partial g}{\partial z}\Big|_{X^{(k)}}(z-z^{(k)})=0 \\ h(X^{(k)})+\dfrac{\partial h}{\partial x}\Big|_{X^{(k)}}(x-x^{(k)})+\dfrac{\partial h}{\partial y}\Big|_{X^{(k)}}(y-y^{(k)})+\dfrac{\partial h}{\partial z}\Big|_{X^{(k)}}(z-z^{(k)})=0. \end{cases}$$

近似逼近

$$\Leftrightarrow \begin{pmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} & \dfrac{\partial f}{\partial z} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} & \dfrac{\partial g}{\partial z} \\ \dfrac{\partial h}{\partial x} & \dfrac{\partial h}{\partial y} & \dfrac{\partial h}{\partial z} \end{pmatrix}\Bigg|_{X^{(k)}} \begin{pmatrix} x-x^{(k)} \\ y-y^{(k)} \\ z-z^{(k)} \end{pmatrix} = -\begin{pmatrix} f(X^{(k)}) \\ g(X^{(k)}) \\ h(X^{(k)}) \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{pmatrix}\Bigg|_{X^{(k)}} \begin{pmatrix} x - x^{(k)} \\ y - y^{(k)} \\ z - z^{(k)} \end{pmatrix} = -\begin{pmatrix} f(X^{(k)}) \\ g(X^{(k)}) \\ h(X^{(k)}) \end{pmatrix}$$

$$\Leftrightarrow F'(X^{(k)})(X - X^{(k)}) = -F(X^{(k)})$$

$$\Leftrightarrow X = X^{(k)} - F'(X^{(k)})^{-1} F(X^{(k)})$$

F的Jacobi矩阵的逆矩阵

函数$f(x)=0$的 Newton 法:

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

非线性方程组的Newton迭代法

假设方程组$F(X)=0$的解$X^*\in \mathrm{int}(D)$，则存在$B_\delta(X^*)\subset D$，并且$F(X)$在$B_\delta(X^*)$内可微.

设$X^{(k)}\in B_\delta(X^*)$是方程组$F(X)=0$的第k个近似解.

对$F(X)$的第i个分量$f_i(X)$在$X^{(k)}$点Taylor展开

$$f_i(X)\approx f_i(X^{(k)})+\sum_{j=1}^{n}\frac{\partial f_i(X^{(k)})}{\partial x_j}(x_j-x_j^{(k)}),\quad i=1,2,\dots,n$$

则$f_i(X)=0$可用线性方程近似逼近

$$f_i(X^{(k)})+\sum_{j=1}^{n}\frac{\partial f_i(X^{(k)})}{\partial x_j}(x_j-x_j^{(k)})=0,\quad i=1,2,\dots,n$$

$$f_i(X^{(k)})+\sum_{j=1}^{n}\frac{\partial f_i(X^{(k)})}{\partial x_j}(x_j-x_j^{(k)})=0,\quad i=1,2,\dots,n$$

即： $$F'(X^{(k)})(X-X^{(k)})=-F(X^{(k)})$$

$f(x)=0$

解出来的的X作为下一步的迭代值，得

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

$$X^{(k+1)}=X^{(k)}-F'(X^{(k)})^{-1}F(X^{(k)})$$

$$F'(X)=\left(\frac{\partial f_i(X)}{\partial x_j}\right)_{n\times n}=\begin{pmatrix}\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n}\\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n}\\ \vdots & \vdots & \cdots & \vdots\\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n}\end{pmatrix}$$

$F(X)$在X点的**Jacobi矩阵.**

定理4.14(局部收敛定理)　设$F(x)$的定义域为$D \subset R^n$,设$x^* \in \mathrm{int}(D)$,满足$F(x^*)=0$, 在x^*的开邻域$S \subset D$上$F(x)$连续可微，且$F'(x^*)$非奇异，则牛顿法生成的序列$\{x^{(k)}\}$在闭域$S_0 \subset S$上超线性收敛于x^*，若还存在常数$L>0$, 使

$$\|F'(x) - F'(x^*)\| \le L\|x - x^*\|, \quad \forall x \in S,$$

则$\{x^{(k)}\}$至少平方收敛于x^*.

如果$F(x)$在S内二次连续可微，则序列$\{x_k\}$至少是平方收敛.

例3　用牛顿法求解方程组

$$\begin{cases} f_1(x_1,x_2)=x_1+2x_2-3=0, \\ f_2(x_1,x_2)=2x_1^2+x_2^2-5=0, \end{cases}$$

$$X^{(k+1)}=X^{(k)}-F'(X^{(k)})^{-1}F(X^{(k)})$$

初值$\mathbf{x}^{(0)}=(1.5,1.0)^T$.

解： $\mathrm{F}'(\mathrm{x})=\begin{bmatrix} 1 & 2 \\ 4x_1 & 2x_2 \end{bmatrix}$, $\mathrm{F}'(\mathrm{x})^{-1}=-\dfrac{1}{2x_2-8x_1}\begin{bmatrix} 2x_2 & -2 \\ -4x_1 & 1 \end{bmatrix}$,

$$X^{(k+1)}=X^{(k)}-\frac{1}{2x_2^{(k)}-8x_1^{(k)}}\begin{pmatrix} 2x_2^{(k)} & -2 \\ -4x_1^{(k)} & 1 \end{pmatrix}\begin{pmatrix} x_1^{(k)}+2x_2^{(k)}-3 \\ 2(x_1^{(k)})^2+(x_2^{(k)})^2-5 \end{pmatrix}$$

$$X^{(k+1)} = X^{(k)} - \frac{1}{2x_2^{(k)} - 8x_1^{(k)}} \begin{pmatrix} 2x_2^{(k)} & -2 \\ -4x_1^{(k)} & 1 \end{pmatrix} \begin{pmatrix} x_1^{(k)} + 2x_2^{(k)} - 3 \\ 2(x_1^{(k)})^2 + (x_2^{(k)})^2 - 5 \end{pmatrix}$$

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} - \dfrac{(x_2^{(k)})^2 - 2(x_1^{(k)})^2 + x_1^{(k)}x_2^{(k)} - 3x_2^{(k)} + 5}{x_2^{(k)} - 4x_1^{(k)}}, \\ x_2^{(k+1)} = x_2^{(k)} - \dfrac{(x_2^{(k)})^2 - 2(x_1^{(k)})^2 - 8x_1^{(k)}x_2^{(k)} + 12x_2^{(k)} - 5}{2(x_2^{(k)} - 4x_1^{(k)})}. \end{cases}$$

k	$x^{(k)}$
0	$(1.5, 1.0)^T$
1	$(1.5, 0.75)^T$
2	$(1.488095, 0.755952)^T$
3	$(1.488034, 0.755983)^T$

逐次迭代得结果.

例4 用牛顿法解方程组 $\begin{cases} x_1^2-10x_1+x_2^2+8=0, \\ x_1x_2^2+x_1-10x_2+8=0. \end{cases}$

解 由于

$$F(x)=\begin{pmatrix} x_1^2-10x_1+x_2^2+8 \\ x_1x_2^2+x_1-10x_2+8 \end{pmatrix},\quad F'(x)=\begin{pmatrix} 2x_1-10 & 2x_2 \\ x_2^2+1 & 2x_1x_2-10 \end{pmatrix}.$$

选$x^{(0)}=(0,0)^T$,解线性方程组$F'(x^{(0)})\Delta x^{(0)}=-F(x^{(0)})$，即

$$\begin{pmatrix} -10 & 0 \\ 1 & -10 \end{pmatrix}\begin{pmatrix} \Delta x_1^{(0)} \\ \Delta x_2^{(0)} \end{pmatrix}=\begin{pmatrix} -8 \\ -8 \end{pmatrix}.$$

解得$\Delta x^{(0)}=(0.8,0.88)^T$, $\Delta x^{(1)}=x^{(0)}+\Delta x^{(0)}=(0.8,0.88)^T$,按牛顿迭代法计算结果入表.

	$x^{(0)}$	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$
$x_1^{(k)}$	0	0.80	0.9917872	0.9999752	1.0000000
$x_2^{(k)}$	0	0.88	0.9917117	0.9999685	1.0000000

小　　结

1、本章的目的是求解形如 $f(x)=0$ 的方程，而其核心方法是将所要求解的方程变形为 $x=\varphi(x)$ ，利用 $\varphi(x)$ 为压缩映射，通过迭代求出其解。

2、变形中切记要恒等变形！

3、恒等变形的一种重要格式是牛顿迭代，证明其迭代收敛阶的一个常用技巧是泰勒展开。

$x = G(x)$　　$\rho(G'(x^*)) < 1$

4、n维空间中代数方程迭代求解的收敛的充分条件是谱半径小于 1 .

局部收敛

- 非线性方程组的数值解法
  - 方程求根与二分法
  - 简单迭代法
    - 简单迭代法的存在性与收敛性
    - 局部收敛性与收敛阶
  - 迭代收敛的加速方法
    - 斯蒂芬森迭代法
  - 牛顿法
    - 单根的牛顿迭代法
      - 牛顿下山法、简化牛顿法
    - 重根的牛顿迭代法
  - 割线法
    - 割线法及其收敛性
    - 单点割线法
  - 非线性方程组的数值解法
    - 多变量方程的简单迭代法
    - 非线性方程组的的牛顿迭代法

三、离散Newton迭代法（以下内容理解就可以）

1 拟牛顿法或弦截法

与单个方程的情形类似，牛顿法中 f 的导数的元素用合适的差商来近似，如

$$\frac{\partial f_j(x_1^{(k)}, x_2^{(k)}, \cdots x_n^{(k)})}{\partial x_i}$$

$$\approx \frac{f_j(x_1^{(k)}, \cdots, x_i^{(k)} + h_i, \cdots x_n^{(k)}) - f_j(x_1^{(k)}, \cdots, x_i^{(k)} - h_i, \cdots x_n^{(k)})}{2h_i}$$

或 $$\approx \frac{f_j(x_1^{(k)}, \cdots, x_i^{(k)}, \cdots x_n^{(k)}) - f_j(x_1^{(k)}, \cdots, x_i^{(k-1)}, \cdots x_n^{(k)})}{x_i^{(k)} - x_i^{(k-1)}}$$

2 简化牛顿法 目的是避免计算迭代公式中繁杂的导数，解决方法与一元函数牛顿法类似，即将所有导数取为固定值，如迭代初值的导数值。

$$X^{(k+1)} = X^{(k)} - F'(X^{(k)})^{-1} F(X^{(k)}) \qquad x^{(k+1)} = x^{(k)} - [f'(x^{(0)})]^{-1} f(x^{(k)})$$

$$\begin{cases} f_1(x_1^{(k)}, x_2^{(k)}, \cdots x_n^{(k)}) + \sum\limits_{i=1}^{n} \dfrac{\partial f_1(x_1^{(0)}, x_2^{(0)}, \cdots x_n^{(0)})}{\partial x_i} \Delta x_i^{(k)} = 0 \\ f_2(x_1^{(k)}, x_2^{(k)}, \cdots x_n^{(k)}) + \sum\limits_{i=1}^{n} \dfrac{\partial f_2(x_1^{(0)}, x_2^{(0)}, \cdots x_n^{(0)})}{\partial x_i} \Delta x_i^{(k)} = 0 \\ \qquad \vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\ f_n(x_1^{(k)}, x_2^{(k)}, \cdots x_n^{(k)}) + \sum\limits_{i=1}^{n} \dfrac{\partial f_n(x_1^{(0)}, x_2^{(0)}, \cdots x_n^{(0)})}{\partial x_i} \Delta x_i^{(k)} = 0 \end{cases}$$

$$x_i^{(k+1)} = x_i^{(k)} + \Delta x_i^{(k)} \ (i = 1, 2, \cdots, n)$$

如果用格式$x^{(k+1)}=x^{(k)}-\lambda_k(f'(x^{(k)}))^{-1}f(x^{(k)})$

其中下山因子$\lambda_k\in(0,1]$合适地选取使得$\left\|f(x^{(k+1)})\right\|<\left\|f(x^{(k)})\right\|$

就得到**牛顿下山法**。

作业

- 教材P93页习题10、11

本讲课程结束

北京航空航天大学数学科学学院