Lect 13 Newton's method

一、引言：

1. 基本方法：

直接法：不可能；迭代法：$x_0, x_1, \cdots, x_n \cdots$ 序列

2. 收敛阶：

设 $\lim_{n\to\infty} x_n = \alpha$，$\varepsilon_n = |x_n - \alpha|$

A. 线性收敛： 典例：

$\lim_{n\to\infty} \varepsilon_{n+1}/\varepsilon_n = C$，$0<C<1$ q^n，$|q|<1$

B. 超线性收敛

$\lim_{n\to\infty} \varepsilon_{n+1}/\varepsilon_n = 0$

次线性： $\lim \varepsilon_{n+1}/\varepsilon_n = 1$ $\frac{1}{n^\alpha}$，$\alpha>0$

C. p阶收敛：

$\lim_{n\to\infty} \varepsilon_{n+1}/\varepsilon_n^p = C > 0$；$p>1$ q^{2^n}，$p=2$，$q<1$

超p阶 $\lim_{n\to\infty} \varepsilon_{n+1}/\varepsilon_n^p = 0$

3. 有效位数：

n充分大时，有 $\varepsilon_{n+1} \approx C \cdot \varepsilon_n^p$ $n=n_0, n_0+1, \cdots$

$$\varepsilon_{n_0+k} = C^{\frac{p^k-1}{p-1}} \varepsilon_{n_0}^{p^k} \qquad k=0,1,2,\cdots, \quad p>1$$

$$\varepsilon_{n_0+k} = C^k \varepsilon_{n_0} \qquad p=1$$

设 $\varepsilon_{n_0+k} = 10^{-\delta_k} \varepsilon_{n_0}$，则 δ_k 表征自 x_{n_0} ~~开~~始增加的有效位数，则

$$k>1, \quad \delta_k = \begin{cases} k \lg 1/C & p=1 \\ p^k \cdot \tilde{C} & p>1 \end{cases}$$

则

$$\delta_k \approx \begin{cases} C_1 k & p=1 \\ C_p p^k & p>1 \end{cases}$$

即 $p=1$ 时，有效位数线性增加；

$p>1$ ，〃 指数增加.

二. 非线性方程组的条件数：

设方程组 $F(\vec{x}, \vec{d}) = 0 \qquad \vec{d} \in R^n, \vec{x} \in R^m$

解 $\vec{x} = G(\vec{d})$

定义相对条件数：

$$K(\vec{d}) = \sup_{\delta\vec{d} \in D} \frac{\|\delta\vec{x}\| / \|\vec{x}\|}{\|\delta\vec{d}\| / \|\vec{d}\|}$$

13-03

$K \gg 1$，坏条件；$K \sim O(1)$ 好条件.

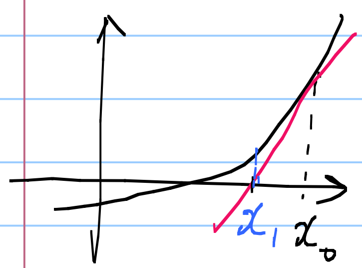
由 $\vec{x}+\delta\vec{x} = G(\vec{d}+\delta\vec{d})$

$$\approx G(\vec{d}) + G'(\vec{d})\cdot\delta\vec{d}$$

$$\Rightarrow K(\vec{d}) \approx \|G'(\vec{d})\| \cdot \frac{\|\vec{d}\|}{\|\vec{x}\|}$$

三、牛顿法：

1. 1D情形：　$f(x)=0$



基本思想：局部线性近似：

切线：　$y = f'(x_k)(x-x_k)+f(x_k)$

$y=0 \Rightarrow \quad \tilde{x} = x_k - f(x_k)/f'(x_k)$

Newton法：

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

定理：如 $f'(x^*)\neq 0$, $f''(x^*)\neq 0$, Newton法局部2阶收敛.

注：1) 局部性要求 $|x_0-x^*|<\delta$，δ充分小.

2) $f(x^*)=0$ 重要，否则不收敛或收敛极慢！

3) 证明：压缩映像原理.

2.高维: $F(X)=0$, $X,F\in\mathbb{R}^n$

$$X_{k+1}=X_k-[DF(X_k)]^{-1}\cdot F(X_k)$$

$DF(X_k)$为Jacobian矩阵.

3.收敛性:

定理1: 设X^*为$X=\varphi(X)$的解。$\varphi: D\subset\mathbb{R}^n\to\mathbb{R}^n$. 如存在$X^*$的邻域$S=S(X^*,\delta)=\{X \mid \|X-X^*\|<\delta, \delta>0\}\subset D$及$\alpha\in(0,1)$ s.t. $\forall X\in S$

$$\|\varphi(X)-\varphi(X^*)\|\le\alpha\|X-X^*\|$$

则$\forall X_0\in S$, $X_k\longrightarrow X^*$.

解: $\forall X_0\in S$, $\|X_k-X^*\|\le\alpha^k\|X_0-X^*\|<\delta$

故$X_k\in S$, 由压缩映像原理, $X_k\longrightarrow X^*$.

定理2: 设$\varphi: D\subset\mathbb{R}^n\longrightarrow\mathbb{R}^n$有不动点, X^*, 且X^*局部可微, 且$\rho(D\varphi(X^*))=\sigma<1$, 则$\exists S=S(X^*,\delta)\subset D$ s.t.

$\forall X_0\in S$, $X_k\to X^*$.

解: 由$\rho(D\varphi(X^*))<1$, $\exists\|\cdot\|$ s.t. $\exists\varepsilon$ s.t.

$\sigma+2\varepsilon<1$, $\|D\varphi(X^*)\|\le\sigma+\varepsilon$,

且存在 $S(x^*,\delta)$

$\forall x\in S(x^*,\delta)$,

$$\|\varphi(x)-\varphi(x^*)-D\varphi(x^*)(x-x^*)\|\le\varepsilon\|x-x^*\|$$

$$\Rightarrow \|\varphi(x)-\varphi(x^*)\|\le\left(\|D\varphi(x^*)\|+\varepsilon\right)\|x-x^*\|$$

$$\le \underbrace{(\sigma+2\varepsilon)}_{\substack{\shortparallel\\ \alpha<1}}\|x-x^*\|$$

由定理1得证.

定理3: 如 $DF(x^*)$ 非奇异,且 $\exists r>0$, s.t.

$$\|DF(x)-DF(y)\|\le L\|x-y\|\qquad \forall x,y\in B_r(x^*)$$

则 Newton 法局部2阶收敛.

证: 对 Newton 迭代 $\varphi(x)=x-(DF(x))^{-1}\cdot F(x)$

$\Rightarrow D\varphi(x^*)=0$, 由定理2有收敛性.

又 $x_{k+1}=x_k-[DF(x_k)]^{-1}\cdot F(x_k)$

$x^*=x^*-[DF(x_k)]^{-1}\cdot F(x^*)$

$$\Rightarrow x_{k+1}-x^*=x_k-x^*-[DF(x_k)]^{-1}\cdot(F(x_k)-F(x^*))$$

$$\Rightarrow E_{k+1}=-[DF(x_k)]^{-1}\cdot[F(x_k)-F(x^*)-DF(x_k)(x_k-x^*)]$$

$$\|E_{k+1}\| \le \|[DF(X_k)]^{-1}\| \cdot \|DF(X_\theta) - DF(X_k)\| \cdot \|E_k\|$$

$$X_\theta = \theta X_k + (1-\theta) X^*$$

$$\le \|DF(X_k)^{-1}\| \cdot L(1-\theta)\|E_k\|^2$$

即 $\|E_{k+1}\| / \|E_k\|^2 \le L(1-\theta)\|DF(X_k)^{-1}\| < C$ #.

四. Newton法的变形:

秩1拟牛顿法:

$$X_{k+1} = X_k - A_k^{-1} \cdot F(X_k) \quad \text{近似 } A_k \approx DF(X_k)$$

直观上 $F(X_k) - F(X_{k-1}) \approx DF(X_k) \cdot (X_k - X_{k-1})$

对应于 $g_{k-1} = A_k \cdot y_{k-1}$

令 $A_k = A_{k-1} + \Delta A_k$, ΔA_k 为修正

取 $\Delta A_k = N \cdot W^T$ 为秩1修正.

· Broyden方法:

令 $W = X_k - X_{k-1} = y_{k-1}$

故

$$N = (g_{k-1} - A_{k-1} y_{k-1}) / y_{k-1}^T \cdot y_{k-1}$$

要让 $y_{k-1}\neq 0$，否则停机。

· Broyden 第二方法：

在优化问题中 DF 对应于 Hessian 具对称性，能否保持对称性？

$$A_{k-1}y_{k-1}+v\cdot w^T\cdot y_{k-1}=g_{k-1}$$

$$\Rightarrow v=(g_{k-1}-A_{k-1}y_{k-1})/w^T\cdot y_{k-1}$$

$$\Rightarrow v\cdot w^T=F(x_k)\cdot w^T/w^T\cdot y_{k-1}$$

取 $w=F(x_k)$，则 $v\cdot w^T$ 对称！

五、延拓法：

为求解 $F(x)=0$，取 $G(x)=0$，构造

$$H(x,\lambda)=\lambda F(x)+(1-\lambda)G(x)$$

$H(x,0)=G(x)$，$H(x,1)=F(x)$，$G(x)=0$ 的 x_0 已知。

设 $H(x,\lambda_i)=0$ 解为 x_i，用 x_i 作 x_{i+1} 的初值！

"顺藤摸瓜"