非线性方程组的迭代解法

包承龙

丘成桐数学科学中心

本章研究对象及目标

一般的非线性方程组

$$\mathbf{F}(\mathbf{x}) = \mathbf{0}$$

代数方程

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0, \quad a_0 \neq 0,$$

当$n \geq 5$时，上述方程的根只能通过数值求解。
超越方程： $2x^2 - e^x = 0$, $x = \frac{1}{2} + \sin x$

设$f \in C[a,b]$, $x^* \in [a,b]$且$f(x^*) = 0$,

$$f(x) = (x - x^*)^m g(x), \quad g(x^*) \neq 0$$

称x^*为f的m重零根。若g充分光滑，x^*为m重零根，则有

$$f(x^*) = f'(x^*) = \cdots = f^{(m-1)}(x^*) = 0, \quad f^{(m)}(x^*) \neq 0$$

目录

二分法

设f在区间$[a, b]$上有且仅有一个根，满足

$$f(a)f(b) < 0$$

可以用二分法形成有根区间的序列$\{[a_n, b_n]\}$，使得

$$b_n - a_n = \frac{b_{n-1} - a_{n-1}}{2} = \frac{b-a}{2^n}, \quad [a_n, b_n] \subset [a_{n-1}, b_{n-1}]$$

具体而言，设$a_0 = a, b_0 = b$，且

$$x_n = \frac{b_{n-1} - a_{n-1}}{2}$$

若$f(a_{n-1})f(x_n) < 0$，则设$a_n = a_{n-1}$，$b_n = x_n$；若$f(b_{n-1})f(x_n) < 0$，则设$a_n = x_n$，$b_n = b_{n-1}$，则有$x^* \in [a_n, b_n]$，从而$|x_n - x^*| \leq (b-a)/2^n$

目录

不动点迭代

设$x \in \mathbb{R}$，有

$$f(x) = 0 \quad \Leftrightarrow \quad x = \varphi(x) \Rightarrow x_{k+1} = \varphi(x_k)$$

称$\{x_k\}$为φ的不动点的迭代序列。
目标： $\varphi_k \to x^*$，满足$x^* = \varphi(x^*)$

例： 设$f(x) = x^3 + 4x^2 - 10$，$x \in [1, 2]$

- $\varphi_1(x) = x - f(x)$
- $f(x) = 0 \Leftrightarrow 4x^2 = 10 - x^3 \Leftrightarrow x = \frac{1}{2}(10 - x^3)^{\frac{1}{2}} = \varphi_2(x)$
- $f(x) = 0 \Leftrightarrow x^2 = \frac{10}{x} - 4x \Leftrightarrow x = \left(\frac{10}{x} - 4x\right)^{\frac{1}{2}} = \varphi_3(x)$
- $f(x) = 0 \Leftrightarrow x^2 = \frac{10}{4+x} \Leftrightarrow x = \left(\frac{10}{4+x}\right)^{\frac{1}{2}} = \varphi_4(x)$

重要定理

定理

设$\varphi \in C[a,b]$, 满足

$$\varphi(x) \in [a,b], \forall x \in [a,b], \tag{1}$$

则φ在$[a,b]$上存在不动点。若φ满足(1)且有

$$\exists L \in (0,1) \Rightarrow |\varphi(x) - \varphi(y)| \le L|x-y|, \quad \forall x, y \in [a,b], \tag{2}$$

则φ在$[a,b]$上不动点唯一。

若$\varphi(a) = a$或$\varphi(b) = b$，则φ存在不动点。考虑$\varphi(a) > a, \varphi(b) < b$, 令

$$\psi(x) = \varphi(x) - x \Rightarrow \psi \in C[a,b], \psi(a) > 0, \psi(b) < 0$$

存在$x^* \in (a,b)$，使得$0 = \psi(x^*) = \varphi(x^*) - x^*$. (2)称为利普希茨条件。

设函数φ满足(1)，又$\varphi \in C^1[a,b]$，且存在常数$L \in (0,1)$，满足

$$|\varphi'(x)| \leq L, \quad \forall x \in (a,b)$$

则φ在$[a,b]$上存在唯一的不动点。

设$\varphi \in C[a,b]$满足(1)与(2)，则不动点迭代序列满足

$$|x_{k+p} - x_k| \leq \frac{1}{1-L}|x_{k+1} - x_k|$$
$$|x^* - x_k| \leq \frac{L}{1-L}|x_k - x_{k-1}|$$
$$|x^* - x_k| \leq \frac{L^k}{1-L}|x_1 - x_0|$$

上述方程给出了迭代法的全局收敛性，但在多数情况下，全局收敛性不容易检验。

局部收敛性与收敛阶

定义： 设函数φ在区间I上存在不动点x^*。如果存在x^*的一个邻域$S \subset I$，对于任意$x_0 \in S$，迭代法$x_k = \varphi(x_{k-1})$满足

$$\{x_k\} \subset S, \quad x_k \to x^*,\ k \to \infty$$

定理： 设φ'在x^*的某个邻域S上存在，连续且满足$|\varphi'(x^*)| < 1$，则该迭代法局部收敛。

由φ' 存在且连续，则存在$\delta > 0$，使得

$$N(x^*) = [x^* - \delta, x^* + \delta] \subset S,$$

有$\varphi'(x) = L < 1, \forall x \in N(x^*)$。则对任意$x \in N(x^*)$，有

$$|\varphi(x) - x^*| = |\varphi(x) - \varphi(x^*)| \leq L|x - x^*| < \delta \Rightarrow \varphi(x) \in N(x^*)$$

设迭代序列$\{x_k\}$收敛到x^*，记$e_k = x_k - x^*$

- p阶收敛：存在$p \geq 1$且$C \neq 0$，使得$\lim\limits_{k\to\infty} \frac{|e_{k+1}|}{|e_k|^p} = C$
- 至少p阶收敛：存在存在$p \geq 1$且$C > 0$及K，使得

$$|e_{k+1}| \leq C|e_k|^p, \quad \forall k \geq K$$

注：若$p = 1$，则需要$C \in (0, 1)$.

假设$\{x_k\}$$p$阶收敛，则有

- $p = 1$，线性收敛
- $p = 2$，平方收敛
- $p > 1$，超线性收敛，$\lim\limits_{k\to\infty} \frac{|e_{k+1}|}{|e_k|} = 0$

定理：设x^*为φ的不动点，$\varphi^{(p)}$在x^*的邻域上连续，且满足

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

则不动点迭代所产生的序列在x^*的邻域是p阶收敛，且有

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k^p} = \frac{\varphi^{(p)}(x^*)}{p!}$$

由泰勒展式与中值定理可知，

$$\begin{aligned}\varphi(x_k) =& \varphi(x^*) + \varphi'(x^*)(x_k - x^*) + \cdots + \\ & \frac{\varphi^{(p-1)}(x^*)}{(p-1)!}(x_k - x^*)^{p-1} + \frac{\varphi^{(p)}(\xi)}{p!}(x_k - x^*)^p\end{aligned}$$

目录

Aitken加速方法

设迭代序列$\{x_k\}$线性收敛到x^*，当k充分大时，有

$$\frac{|x_{k+1}-x^*|}{|x_k-x^*|}=\frac{|e_{k+1}|}{|e_k|}\approx C,\quad C\in(0,1)$$

则有

$$\frac{x_{k+2}-x^*}{x_{k+1}-x^*}\approx\frac{x_{k+1}-x^*}{x_k-x^*}\Rightarrow x^*\approx\frac{x_k x_{k+2}-x_{k+1}^2}{x_{k+2}-2x_{k+1}+x_k},$$

对每步的x_k进行修正，得到序列

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

其中 $\Delta x_k=x_{k+1}-x_k$, $\Delta^2 x_k=x_{k+2}-2x_{k+1}+x_k$

$\{\bar{x}_k\}$称为Aitken加速方法。

定理

设$\{\bar{x}_k\}$满足$x_k \neq 0$, 且存在非零常数$|\lambda| < 1$, 使

$$x_{k+1} - x^* = (\lambda + \delta_k)(x_k - x^*)$$

其中$\delta_k \to 0, k \to \infty$, 则对充分大的$k$, 有

$$\lim_{k\to\infty} \frac{\bar{x}_k - x^*}{x_k - x^*} = 0$$

Steffensen迭代法

给定x_k，有如下迭代公式

$$
\begin{aligned}
y_k &= \varphi(x_k), \\
z_k &= \varphi(y_k), \\
x_{k+1} &= x_k - \frac{(y_k - x_k)^2}{z_k - 2y_k + x_k}
\end{aligned}
$$

上述公式产生迭代公式

$$
x_{k+1} = \psi(x_k), \quad \psi(x) = \frac{x\varphi(\varphi(x)) - [\varphi(x)]^2}{\varphi(\varphi(x)) - 2\varphi(x) + x}
$$

定理

若x^*是ψ的不动点，则x^*是φ的不动点。反之，若x^*是φ的不动点，并设φ'存在且连续，$\varphi'(x^*) \neq 1$, 则x^*是函数ψ的不动点。

定理

设x^*是φ的不动点，在x^*邻域φ有$p+1$阶导数存在且连续。对$p=1$, 若$\varphi'(x^*) \neq 1$, 则*Steffensen*迭代法是二阶收敛的。若原始迭代$x_{k+1} = \varphi(x_k)$是$p(p>1)$阶收敛的，则*Steffensen*迭代是$2p-1$阶收敛的。

目录

牛顿法

求解$f(x)=0$，考虑$f(x)$在x_k的二阶泰勒展式

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

则有

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

几何解释： 曲线$y=f(x)$在$(x_k,f(x_k))$处做切线

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

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

定理

设$f(x^*) = 0, f'(x^*) \neq 0$, 且$f$在包含$x^*$的一个区间上有二阶连续导数，则*Newton*迭代法局部收敛到x^*, 且至少二阶收敛，并有

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

扩展: 若$f(x) = (x - x^*)^m g(x), m > 1$且$g$有二阶导数$g'(x^*) \neq 0$, 有

$$\varphi'(x^*) = 1 - \frac{1}{m} \neq 0$$

则Newton迭代局部收敛，但只是线性收敛。

修改方法一：

$$\varphi(x) = x - \frac{mf(x)}{f'(x)} \quad \Rightarrow \quad \varphi(x^*) = x^*, \varphi'(x^*) = 0$$

修改方法二：$\mu(x) = \frac{f(x)}{f'(x)}$

$$\mu(x) = (x - x^*)\left[\frac{g(x)}{mg(x) + (x - x^*)g'(x)}\right] \Rightarrow \varphi(x) = x - \frac{\mu(x)}{\mu'(x)}$$

割线法

近似导数：

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

迭代公式：

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

定理

设$f(x^*) = 0$, 在区间$\Delta = [x^* - \delta, x + \delta]$上$f'(x) \neq 0$且$f \in C^2(\Delta)$, 又设$M\delta < 1$, 其中

$$M = \frac{\max_{x\in\Delta} |f''(x)|}{2 \min_{x\in\Delta} |f'(x)|},$$

则当$x_0, x_1 \in \Delta$时，割线法产生的序列$\{x_k\} \subset \Delta$收敛到x^*，并且收敛阶为$p = \frac{1+\sqrt{5}}{2}$。

目录

向量值函数及导数

考虑$\mathbf{F}:\mathcal{D}\subset\mathbb{R}^n\to\mathbb{R}^m$, 则$\mathbf{F}$有如下定义

$$\mathbf{F}(\mathbf{x})=\begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix}=\begin{bmatrix} f_1(x_1,x_2,\ldots,x_n) \\ f_2(x_1,x_2,\ldots,x_n) \\ \vdots \\ f_m(x_1,x_2,\ldots,x_n) \end{bmatrix}$$

导数：对于$\mathcal{D}$的内点$\mathbf{x}$，若存在矩阵$\mathbf{A}\in\mathbb{R}^{m\times n}$，有

$$\lim_{\mathbf{h}\to\mathbf{0}}\frac{\|\mathbf{F}(\mathbf{x}+\mathbf{h})-\mathbf{F}(\mathbf{x})-\mathbf{A}\mathbf{h}\|}{\|\mathbf{h}\|}=0,$$

则称$\mathbf{F}$在$\mathbf{x}$处可导，记$\mathbf{F}'(\mathbf{x})=\mathbf{A}$

设$f : \mathbb{R}^n \to \mathbb{R}$, 则有$\mathbf{A} \in \mathbb{R}^{1\times n}$, 记为$\nabla f(\mathbf{x})^\top$, 则有

$$\nabla f(x) = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \dots, \frac{\partial f(\mathbf{x})}{\partial x_n}\right)^\top$$

其中,

$$\frac{\partial f(\mathbf{x})}{\partial x_j} = \lim_{h_j\to 0} \frac{f(\mathbf{x} + h_j\mathbf{e}_j) - f(\mathbf{x})}{h_j}$$

雅可比矩阵$\mathbf{F}'(\mathbf{x})$:

$$\mathbf{F}'(\mathbf{x}) = [\nabla f_1(\mathbf{x}), \dots, \nabla f_n(\mathbf{x})]^\top = \begin{bmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial f_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \frac{\partial f_m(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

设$\mathbf{F}: D \subset \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x} \in D$

- 若$\mathbf{F}$ 在x出可导，则$\mathbf{F}$在x处连续
- 若D是一个凸域，$\mathbf{F}$在D内可导，则存在$\xi_i \in (0,1), i = 1, \ldots, m$，使

$$\mathbf{F}(\mathbf{x}+\mathbf{h}) - \mathbf{F}(\mathbf{x}) = \begin{bmatrix} \nabla f_1(\mathbf{x}+\xi_1\mathbf{h})^\top \\ \nabla f_2(\mathbf{x}+\xi_2\mathbf{h})^\top \\ \vdots \\ \nabla f_m(\mathbf{x}+\xi_m\mathbf{h})^\top \end{bmatrix} \mathbf{h}$$

设$\mathbf{G}: [a,b] \subset \mathbb{R} \to \mathbb{R}^m$, $\mathbf{G}$是连续的，则

$$\left\| \int_a^b \mathbf{G}(t)dt \right\| \leq \int_a^b \|\mathbf{G}(t)\| dt$$

定理

设D是$\mathbb{R}^n$中的凸域，$\mathbf{F}: D \subset \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{F}$在D内可导，且存在$\gamma > 0$,使

$$\|\mathbf{F}'(\mathbf{x}) - \mathbf{F}'(\mathbf{y})\| \leq \gamma \|\mathbf{x} - \mathbf{y}\|, \ \forall \mathbf{x}, \mathbf{y} \in D,$$

则有

$$\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y}) - \mathbf{F}'(\mathbf{y})(\mathbf{x} - \mathbf{y})\| \leq \frac{\gamma \|\mathbf{x} - \mathbf{y}\|^2}{2}, \ \forall \mathbf{x}, \mathbf{y} \in D.$$

压缩映射和不动点迭代

设$\mathbf{F}: D \subset \mathbb{R}^n \to \mathbb{R}^n$

$$\mathbf{F}(\mathbf{x}) = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{x} = \Phi(\mathbf{x})$$

不动点迭代序列：$\mathbf{x}^{k+1} = \Phi(\mathbf{x}^k)$

压缩映射: 存在$L \in (0,1)$, 使$\|\Phi(\mathbf{y}) - \Phi(\mathbf{x})\| \leq L\|\mathbf{y} - \mathbf{x}\|$, $\forall \mathbf{x}, \mathbf{y} \in D$

定理

设$\Phi : D \subset \mathbb{R}^n \to \mathbb{R}^n$, 如果$\Phi$在凸域$D_0 \subset D$内可导，且$\Phi(\mathbf{x})$满足：

$$\left|\frac{\partial \varphi_i(\mathbf{x})}{\partial x_j}\right| \leq \frac{L}{n}, \ \forall x \in D_0, \quad \Phi(\mathbf{x}) = (\varphi_1(\mathbf{x}), \varphi_2(\mathbf{x}), \ldots, \varphi_n(\mathbf{x}))^\top,$$

则Φ在D_0中对于∞-范数使压缩的。

定理

设$\Phi : D \subset \mathbb{R}^n \to \mathbb{R}^n$且$D_0 \subset D$是一个闭集，$\Phi$在$D_0$上是压缩的，且有

$$\Phi(\mathbf{x}) \in D_0, \quad \mathbf{x} \in D_0,$$

则Φ在D_0中存在唯一的不动点$\mathbf{x}^*$，且对任意的$\mathbf{x}^0 \in D_0$, 迭代法产生的序列$\{\mathbf{x}^k\}$收敛到$\mathbf{x}^*$，有

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

定理

设$\Phi : D \subset \mathbb{R}^n \to \mathbb{R}^n$, Φ在有界闭凸集$D_0 \subset D$上连续，且对一切$\mathbf{x} \in D_0$, 有$\Phi(\mathbf{x}) \in D_0$, 则Φ在D_0中一定存在不动点。

局部收敛： 设$\mathbf{x}^*$为Φ的不动点，若存在$\mathbf{x}^*$的邻域$S\subset D$, 对一切$x^0\in S$, 有$\{\mathbf{x}^k\}\subset S$, 且$\mathbf{x}^k\to\mathbf{x}^*, k\to\infty$

p阶收敛：存在常数$p\geq 1$及$C>0$，（若$p=1$,规定$C\in(0,1)$）使

$$\lim_{k\to\infty}\frac{\|\mathbf{x}^{k+1}-\mathbf{x}^*\|}{\|\mathbf{x}^k-\mathbf{x}^*\|^p}=C\Leftarrow\underbrace{\|\mathbf{x}^{k+1}-\mathbf{x}^*\|\leq C\|\mathbf{x}^k-\mathbf{x}^*\|^p}_{\text{至少}p\text{阶收敛}}$$

定理

设$\mathbf{x}^*$为Φ的不动点，若存在开球 $S\subset D$及常数$L\in(0,1)$, 使得

$$\|\Phi(\mathbf{x})-\Phi(\mathbf{x}^*)\|\leq L\|\mathbf{x}-\mathbf{x}^*\|,\ \forall\mathbf{x}\in S, \tag{3}$$

则若$\mathbf{x}^0\in S$, 迭代法产生的序列收敛到$\mathbf{x}^*$

不等式(3) 成立的一个充分条件是$\rho(\Phi'(\mathbf{x}^*))=\sigma<1$

目录

Newton迭代

设$\mathbf{F}: D \subset \mathbb{R}^n \to \mathbb{R}^n$,则

$$\mathbf{F}(\mathbf{x}) = \mathbf{0} \quad \Leftrightarrow \quad \underbrace{\mathbf{x}^{k+1} = \mathbf{x}^k - (\mathbf{F}'(\mathbf{x}^k))^{-1}\mathbf{F}(\mathbf{x}^k)}_{\text{Newton 迭代}}$$

具体迭代格式:

$$\underbrace{\mathbf{F}'(\mathbf{x}^k)\Delta\mathbf{x}^k = -\mathbf{F}(\mathbf{x}^k)}_{\text{Newton 方程}}, \quad \mathbf{x}^{k+1} = \mathbf{x}^k + \Delta\mathbf{x}^k$$

例：求解方程组$\mathbf{F}(\mathbf{x}) = \mathbf{0}$, 其中

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

定理

设$\mathbf{F}: D \subset \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{x}^* \in D$满足$\mathbf{F}(\mathbf{x}^*) = \mathbf{0}$. 在$\mathbf{x}^*$的邻域$S_0 \subset D$, $\mathbf{F}$可导且$\mathbf{F}'(\mathbf{x})$连续，并且$\mathbf{F}'(\mathbf{x}^*)$可逆，则

1. 存在开球$S = S(\mathbf{x}^*, \delta)$, 使$\Phi(\mathbf{x}) = \mathbf{x} - \mathbf{F}(\mathbf{x})$所定义的不动点迭代对所有的$\mathbf{x} \in S$有意义
2. *Newton*迭代产生的序列$\{\mathbf{x}^k\}$局部收敛到$\mathbf{x}^*$，且超线性收敛
3. 若还有$\|\mathbf{F}'(\mathbf{x}) - \mathbf{F}'(\mathbf{x}^*)\| \le \gamma \|\mathbf{x} - \mathbf{x}^*\|, \forall \mathbf{x} \in S$, 则$\{\mathbf{x}^k\}$至少平方收敛

注意：上述条件涉及$\mathbf{F}$在$\mathbf{x}^*$处的性质，无法预先知道和检验。

拟Newton方法

迭代格式

$$\mathbf{x}^{k+1} = \mathbf{x}^{k} - \mathbf{A}_k^{-1}\mathbf{F}(\mathbf{x}^{k}),$$

且$\mathbf{A}_k$是逼近$\mathbf{F}'(\mathbf{x}^k)$的。

回忆一维情形： $f'(x_{k+1}) \approx \frac{f(x_{k+1})-f(x_k)}{x_{k+1}-x_k}$

$\mathbf{A}_{k+1}$具有如下性质：

$$\underbrace{\mathbf{A}_{k+1} = \mathbf{A}_k + \Delta\mathbf{A}_k}_{\text{修正格式}}$$

$$\underbrace{\mathbf{A}_{k+1}(\mathbf{x}^{k+1} - \mathbf{x}^{k}) = \mathbf{F}(\mathbf{x}^{k+1}) - \mathbf{F}(\mathbf{x}^{k})}_{\text{割线方程}}$$

设$m = \mathbf{rank}(\Delta\mathbf{A}_k) \geq 1$，通常情况下取$m = 1$或$2$, 称为低秩修正。

由$\mathbf{A}_{k+1}$选择的不确定性，以下给出$\mathbf{A}_{k+1}$的构造方法。考虑仿射函数

$$\mathbf{F} = \mathbf{D}\mathbf{u} + \mathbf{b}, \quad \mathbf{D} \in \mathbb{R}^{n\times n}, \mathbf{b} \in \mathbb{R}^n \tag{4}$$

定理

对任意矩阵$\mathbf{A} \in \mathbb{R}^{n\times n}$，$\mathbf{F}$如(4)所定义。设$\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$, 且$\mathbf{x} \neq \mathbf{x}^*$, 定义

$$\mathbf{p} = \mathbf{x}' - \mathbf{x}, \quad , \mathbf{q} = \mathbf{F}(\mathbf{x}') - \mathbf{F}(\mathbf{x}) = \mathbf{D}\mathbf{p},$$

则矩阵

$$\mathbf{A}' = \mathbf{A} + \frac{(\mathbf{q} - \mathbf{A}\mathbf{p})\mathbf{p}^\top}{\|\mathbf{p}\|_2^2}$$

满足

$$\mathbf{A}'\mathbf{p} = \mathbf{D}\mathbf{p} = \mathbf{q}$$
$$\|\mathbf{A}' - \mathbf{D}\|_2 \leq \|\mathbf{A} - \mathbf{D}\|_2$$

在x附近用仿射变换近似，则有Broyden秩1迭代。

$$
\begin{aligned}
\mathbf{x}^{k+1} &= \mathbf{x}^k - \mathbf{A}_k^{-1}\mathbf{F}(\mathbf{x}^k) \\
\mathbf{p}^k &= \mathbf{x}^{k+1} - \mathbf{x}^k, \quad \mathbf{q}^k = \mathbf{F}(\mathbf{x}^{k+1}) - \mathbf{F}(\mathbf{x}^k) \\
\mathbf{A}_{k+1} &= \mathbf{A}_k + \frac{(\mathbf{q}^k - \mathbf{A}_k\mathbf{p}^k)(\mathbf{p}^k)^\top}{\|\mathbf{p}^k\|_2^2}
\end{aligned}
\tag{5}
$$

迭代(5) 具有性质

$$(\mathbf{A}_{k+1} - \mathbf{A}_k)\mathbf{v} = \mathbf{0}, \quad \forall \mathbf{v} \quad 满足\ (\mathbf{p}^k, \mathbf{v}) = 0$$

同时有$\mathbf{A}_{k+1}$满足

$$\mathbf{A}_{k+1} = \arg\min\{\|\mathbf{A} - \mathbf{A}_k\|_F | \mathbf{A}\mathbf{p}^k = \mathbf{q}^k\}$$

下一步：如何减少$\mathbf{A}_k^{-1}$的计算量？

SMW公式：$\mathbf{A}$非奇异，$\mathbf{u},\mathbf{v}\in\mathbb{R}^n$，当$1+\mathbf{v}^\top\mathbf{A}^{-1}\mathbf{u}\neq 0$时，$\mathbf{A}+\mathbf{u}\mathbf{v}^\top$可逆，并且

$$(\mathbf{A}+\mathbf{u}\mathbf{v}^\top)^{-1}=\mathbf{A}^{-1}-\frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^\top\mathbf{A}^{-1}}{1+\mathbf{v}^\top\mathbf{A}^{-1}\mathbf{u}}$$

令$\mathbf{B}_k=\mathbf{A}_k^{-1}$，则迭代公式(5)等价于

$$\begin{aligned}
\mathbf{x}^{k+1}&=\mathbf{x}^k-\mathbf{B}_k\mathbf{F}(\mathbf{x}^k)\\
\mathbf{p}^k&=\mathbf{x}^{k+1}-\mathbf{x}^k,\quad \mathbf{q}^k=\mathbf{F}(\mathbf{x}^{k+1})-\mathbf{F}(\mathbf{x}^k)\\
\mathbf{B}_{k+1}&=\mathbf{B}_k+\frac{(\mathbf{q}^k-\mathbf{B}_k\mathbf{q}^k)(\mathbf{p}^k)^\top\mathbf{B}_k}{(\mathbf{p}^k)^\top\mathbf{B}_k\mathbf{q}^k}
\end{aligned}\tag{6}$$

在迭代(6)没有矩阵求逆的运算。