

科学与工程计算基础

任课教师：黄忠亿

清华大学数学科学系



目录

- 1 引言
- 2 线性方程组的直接解法
- 3 线性方程组的迭代解法
- 4 非线性方程(组)迭代解法
 - 非线性方程求根
 - Newton迭代法
 - 拟牛顿迭代法
 - 非线性方程组求解
 - 多项式求根
 - 同伦算法



非线性方程求根

即求解 $f(x) = 0$, 其中 $f : [a, b] \rightarrow \mathbb{R}$ 为非线性函数.

我们知道, 即便 $f(x)$ 是多项式函数 $p_n(x)$, 由 Abel 定理, 当 $n \geq 5$ 时,
也没有一般的求根公式.

因此我们需要研究迭代法来求解上述非线性方程.

设 $f(x) \in C[a, b]$, 如果 $f(a) \cdot f(b) < 0$, 那么 f 在区间 $[a, b]$ 上必然有一个根.

最简单的办法是用二分法来求解.



非线性方程求根——二分法

算法 4.1 (二分法)

令 $a_0 = a, b_0 = b, x_0 = \frac{a_0+b_0}{2}, k = 0$

① 若 $f(x_k) = 0$, 则输出 x_k , 停止迭代;

否则

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

② $k = k + 1$, 若 $|b_k - a_k| < \varepsilon$ 则令 $x_k = \frac{a_k+b_k}{2}$, 输出并停止迭代;

否则转1.

这样我们得到 $[a, b] = [a_0, b_0] \supset [a_1, b_1] \supset \cdots \supset [a_n, b_n] \supset \cdots$ 每次区间长度减半, 故称为**二分法**.

闭区间套定理



非线性方程求根—二分法

例 4.1

求 $f(x) = x^3 - x - 1$ 在 $[1, 1.5]$ 上的根.

解: $a_0 = 1, b_0 = 1.5. f(a_0) = -1 < 0, f(b_0) = 0.875 > 0.$

因此区间 $[1, 1.5]$ 上 f 必有根.

用二分法迭代到 $k = 6$:

$$a_6 = 1.3203125, b_6 = 1.328125, x_6 = \underline{1.324}21875.$$

可保证有三位有效数字: $\frac{b_6 - a_6}{2} \approx 4 \times 10^{-3} < 5 \times 10^{-3}.$

实际上 $x^* = \underline{1.324717957244} \dots, x_6$ 有四位有效数字. \square

这种迭代只是线性收敛, 收敛比较慢. 因此希望构造更快收敛格式:

$x = \varphi(x)$ 形式的迭代法—**不动点迭代**.

非线性方程求根—不动点迭代法

我们将 $f(x) = 0$ 转化为与其等价的形式 $x = \varphi(x)$:

$$\text{即 } f(\alpha) = 0 \iff \alpha = \varphi(\alpha).$$

显然有无穷多种 φ . 我们需要研究: 是否任取 x_0 , 迭代 $x_{n+1} = \varphi(x_n)$ 都能收敛到根 α ? 此外收敛速度如何? (我们当然希望越快越好!)

定义 4.1 (不动点)

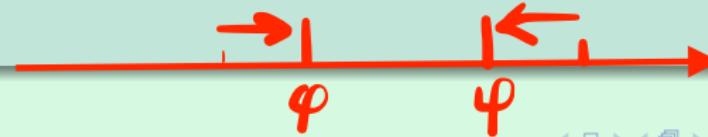
设 $\varphi : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ (称为 $D \rightarrow \mathbb{R}^n$ 的映射), 如果存在 $x^* \in D$, s.t.

$x^* = \varphi(x^*)$, 则称 x^* 为映射 φ 的不动点.

定义 4.2 (压缩映射)

设 $\varphi : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, 如果存在 $C \in (0, 1)$, s.t. $\forall x, y \in D_0 \subset D$, 有

$\|\varphi(x) - \varphi(y)\| \leq C \|x - y\|$, 则称 φ 为在 D_0 上为压缩映射, C 称为压缩系数.



非线性方程求根—不动点迭代法

显然压缩性与所取范数有关.

例如, 设 $A = \begin{pmatrix} 0.1 & 0.9 \\ 0.0 & 0.1 \end{pmatrix}$, 如果取无穷范数或1-范数, 都有

$\|A\|_\infty = \|A\|_1 = 1$, 但是 $\|A\|_2 \approx 0.912 < 1$. 即若令 $\varphi(\mathbf{x}) = A\mathbf{x}$, 那么 φ 在 $\|\cdot\|_2$ 范数下为压缩的, 在 $\|\cdot\|_\infty$ 及 $\|\cdot\|_1$ 下都不压缩.

定理 4.1 (压缩映射原理)

设 $\varphi : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ 在闭集 $D_0 \subset D$ 中是压缩映射, 且 $\varphi(D_0) \subset D_0$, 则 $\varphi(x)$ 在 D_0 上存在唯一不动点.

△ 先证明 φ 在 D_0 上有不动点: 任取 $x_0 \in D_0$, 令 $x_{k+1} = \varphi(x_k)$.



非线性方程求根—不动点迭代法

由 $\varphi(D_0) \subset D_0$ 知, $\{x_k\}_{k=0}^{+\infty} \subset D_0$, 且

$$\|x_{k+1} - x_k\| = \|\varphi(x_k) - \varphi(x_{k-1})\| \leq C \|x_k - x_{k-1}\| \leq \dots \leq C^k \|x_1 - x_0\|.$$

这样 $\forall k, l \in \mathbb{N}$, 有

$$\begin{aligned} \|x_{k+l} - x_k\| &\leq \sum_{j=1}^l \|x_{k+j} - x_{k+j-1}\| \leq \sum_{j=1}^l C^{j-1} \|x_{k+1} - x_k\| \\ (4.1) \quad &\leq \frac{1}{1-C} \|x_{k+1} - x_k\| \leq \frac{C^k}{1-C} \|x_1 - x_0\| \end{aligned}$$

由于 $0 < C < 1 \Rightarrow \{x_k\}_{k \geq 0}$ 为所谓 Cauchy 列, 即必然收敛.

又 $D_0 \subset \mathbb{R}^n$ 为闭集, 因而 $\{x_k\}_{k \geq 0} \subset D_0$, 在 D_0 中有极限点 x^* : 即 $\exists x^* \in D_0$, s.t. $x_k \rightarrow x^*$, 当 $k \rightarrow \infty$.



非线性方程求根—不动点迭代法

又由于 φ 为压缩映射, 自然也连续, 因而有

$$\left[x^* = \lim_{k \rightarrow \infty} x_k = \lim_{k \rightarrow \infty} \varphi(x_{k-1}) = \varphi(x^*) \right]$$

即 x^* 确实为 φ 的不动点. 这样证明了不动点的存在性.

再看唯一性: 假设 $\varphi(x^*) = x^*, \varphi(y^*) = y^*$, 那么

$$\|x^* - y^*\| = \|\varphi(x^*) - \varphi(y^*)\| \leq C \|x^* - y^*\|,$$

由于 $0 < C < 1 \implies x^* = y^*$. 唯一性证毕. \triangleright

推论 4.1

$$\|x_k - x^*\| \leq \frac{C}{1-C} \|x_k - x_{k-1}\| \leq \frac{C^k}{1-C} \|x_1 - x_0\|.$$

\triangleleft 在 (4.1) 中令 $l \rightarrow +\infty$ 即得. \triangleright



非线性方程求根—不动点迭代法

上面定理表明, φ 只需对一种范数压缩, 且 $\varphi(D_0) \subset D_0$, 就能保证在 D_0 中有唯一不动点.

上述定理条件并非迭代收敛的充要条件. 实际上有以下条件更宽松的定理:

定理 4.2 (Brouwer 不动点定理)

设 $\varphi : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ 在 有界闭凸集 $D_0 \subset D$ 上连续且 $\varphi(D_0) \subset D_0$, 则 φ 在 D_0 上有不动点 (注意, 此时不一定唯一).

注意, 条件 $\varphi(D_0) \subset D_0$ 不容易验证, 因此一般只要求 φ 在 x_0 的邻域 $S(x_0, r) = \{|x - x_0| \leq r\}$ 上为压缩映射, 且 $\|x_0 - \varphi(x_0)\| \leq (1 - C)r$, 其中 C 为压缩系数.



非线性方程求根—不动点迭代法

这样, $\forall x \in S(x_0, r)$,

$$\begin{aligned}\|x_0 - \varphi(x)\| &\leq \|x_0 - \varphi(x_0)\| + \|\varphi(x_0) - \varphi(x)\| \\ &\leq (1 - C)r + C\|x_0 - x\| \leq r.\end{aligned}$$

即有 $\boxed{\varphi(S(x_0, r)) \subset S(x_0, r)}$.

例 4.2

仍然求 $f(x) = x^3 - x - 1$ 在 $[1, 1.5]$ 上的根.

解: 将方程改写为 $x = \varphi(x) \equiv \sqrt[3]{x+1}$. 仍取 $x_0 = 1.25$ 有

$x_1 = 1.310371, x_2 = 1.321987, x_3 = \underline{1.324199}$ 有四位有效数字, \dots ,
 $x_6 = \underline{1.324714}$, 具有六位有效数字. 这比前面二分法收敛快. \square



不动点迭代法收敛性分析

我们下面来看不动点迭代的局部收敛性及收敛速度等分析.

定义 4.3 (局部收敛性)

假设 $\exists \delta > 0$ s.t. 在 $x = \varphi(x)$ 的根 α 的一个闭邻域 $B \equiv [\alpha - \delta, \alpha + \delta]$ 上, 迭代 $[x_{k+1} = \varphi(x_k)]$ 对任意 $x_0 \in B$ 都收敛. 则称该迭代在 α 的邻域有局部收敛性.

定义 4.4 (收敛阶)

设 $\{x_k\} \rightarrow \alpha$, 令 $\underline{\varepsilon_k = x_k - \alpha}$. 若存在 $C \neq 0$, s.t. $\frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^p} \rightarrow C$, 则称 $\{x_k\}$ 是 p 阶收敛到 α . 迭代函数 φ 称为 p 阶迭代函数.

我们有以下局部收敛性和收敛阶的判别定理.



不动点迭代法收敛性分析

定理 4.3 (局部收敛性)

设 α 为 $x = \varphi(x)$ 的不动点, $\varphi'(x)$ 在 α 邻域内连续且 $|\varphi'(\alpha)| < 1$, 则迭代 $x_{k+1} = \varphi(x_k)$ 在 α 的邻域内具有局部收敛性. 绝对值

△ 即需要证明 $\exists \delta > 0$, 在 $[\alpha - \delta, \alpha + \delta]$ 上迭代收敛.

因 $A = |\varphi'(\alpha)| < 1$, 且 φ' 在 α 邻域内连续, 故 $\exists \delta > 0$ 及 $C \in (0, 1)$ s.t.

当 $x \in B = [\alpha - \delta, \alpha + \delta]$ 时, $|\varphi'(x)| \leq C < 1$.

这时 $\forall x, y \in B$, $|\varphi(x) - \varphi(y)| = |\varphi'(\xi)| \cdot |x - y| \leq C|x - y|$, 即压缩.

且 $|\varphi(x) - \alpha| = |\varphi(x) - \varphi(\alpha)| \leq C|x - \alpha| \leq C\delta < \delta$,

即有 $\forall x \in B$, $\varphi(x) \in B$.

这样 $\varphi(B) \subset B$, 且 φ 在 B 上压缩 $\Rightarrow \{x_k\} \rightarrow \alpha$. ▷



非线性方程求根—不动点迭代法

定理 4.4 (高阶迭代函数的充要条件)

设 φ 为迭代函数, φ 的 p ($p > 1$) 阶导数 $\varphi^{(p)}$ 在 $x = \varphi(x)$ 的根 α 邻域内连续 (即 $\varphi(\alpha) = \alpha$). 若

$$(4.2) \quad \varphi^{(k)}(\alpha) = 0, \quad k = 1, \dots, p-1, \text{ 且 } \varphi^{(p)}(\alpha) \neq 0,$$

则 φ 为 p 阶迭代函数, 且 $\frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^p} \rightarrow \frac{|\varphi^{(p)}(\alpha)|}{p!}$. 反之, 若 φ 为 p 阶迭代函数, 则 (4.2) 成立.

△ “ \Rightarrow ”: 假设 (4.2) 成立, 即有 $\varphi'(\alpha) = 0$, 由前一定理 4.3 结论知该迭代有局部收敛性. 即存在 $\delta, C > 0$ s.t.

$$\forall x \in B = [\alpha - \delta, \alpha + \delta], \quad |\varphi'(x)| \leq C < 1.$$



非线性方程求根—不动点迭代法

这样, 如果取初值 $x_0 \in B$, 有 $x_{k+1} = \varphi(x_k) \in B$ 且 $\{x_k\} \rightarrow \alpha$.

把 $x_{k+1} = \varphi(x_k)$ 在 α 处 Taylor 展开:

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

其中 $\xi_k \in (x_k, \alpha)$, 可写成 $\xi_k = \alpha + \theta(x_k - \alpha)$, $\theta \in (0, 1)$.

$$\Rightarrow \varepsilon_{k+1} = x_{k+1} - \alpha = \varphi^{(p)}(\xi_k) \frac{\varepsilon_k^p}{p!}$$

$$\Rightarrow \frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^p} = \frac{|\varphi^{(p)}(\xi_k)|}{p!} \rightarrow \frac{|\varphi^{(p)}(\alpha)|}{p!} \neq 0 \text{ (因为 } \xi_k \rightarrow \alpha\text{).}$$

因此根据定义知 φ 为 p 阶迭代函数.



非线性方程求根—不动点迭代法

“ \Leftarrow ”: 反之, 设 φ 为 p 阶迭代函数, 即 $\frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^p} \rightarrow C \neq 0$

可以假设存在 $p_0 \geq 1$ s.t.

$$\varphi^{(k)}(\alpha) = 0, \quad k = 1, \dots, p_0 - 1, \text{ 且 } \varphi^{(p_0)}(\alpha) \neq 0.$$

重复上面证明过程, 我们有 $\frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^{p_0}} \rightarrow \frac{|\varphi^{(p_0)}(\alpha)|}{p_0!} \neq 0$

这样 $\frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^p} = \frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^{p_0}} \frac{1}{|\varepsilon_k|^{p-p_0}}$. 另外由于迭代收敛, 有 $\varepsilon_k \rightarrow 0$.

若 $p > p_0$, 那么 $|\varepsilon_k|^{p-p_0} \rightarrow 0 \Rightarrow \frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^p} \rightarrow \infty$, 矛盾!

若 $p < p_0$, 那么 $|\varepsilon_k|^{p-p_0} \rightarrow +\infty \Rightarrow \frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^p} \rightarrow 0$, 也矛盾!

因此必有 $p = p_0$. \triangleright



非线性方程求根—迭代函数构造

下面一个主要问题就是如何构造高阶迭代函数？即，我们有了 $f(x) = 0$ 的表达式，如何转化为 $x = \varphi(x)$ 形式，且使得 φ 为 p 阶迭代函数（自然希望 p 越大越好），即希望有上面(4.2)式成立。

例 4.3 (还是先看前面的例子 $f(x) = x^3 - x - 1 = 0$)

前面写的迭代函数 $\varphi_1(x) = \sqrt[3]{x+1}$. 计算得到 $\varphi'(x) = \frac{1}{3}(x+1)^{-\frac{2}{3}}$, 显然在 $\alpha = 1.324717957244\cdots$ 处不为零，即此迭代为线性收敛。

若将 $f(x) = 0$ 变形为 $3x^3 - x - 2x^3 + 1 = 0$, 取 $\varphi_2(x) = \frac{2x^3+1}{3x^2-1}$.

可以验证 $\varphi'_2(\alpha) = 0$. 仍取 $x_0 = 1.25$ 计算可得

$x_2 = \underline{1.324749}\cdots$, $x_3 = \underline{1.324717958}\cdots$ 九位有效数字了。

确实**二次收敛**方法比**线性收敛**方法明显快很多！



非线性方程求根—迭代函数构造

下面先设 $f \in C^p[a, b]$ 有 p 阶连续导数, 且设 $\forall x \in [a, b], f'(x) \neq 0$ (即 f 在 $[a, b]$ 上为单射, 不会有重根).

因为 f 为单射, 即 $y = f(x)$ 存在反函数 $x = \mathcal{F}(y)$, 且 \mathcal{F} 也有 p 阶连续导数: $f(\alpha) = 0 \iff \alpha = \mathcal{F}(0)$. 给了 $x \in [a, b]$, 我们可以计算 $y = f(x)$, 亦即 $x = \mathcal{F}(y)$.

任给一个 $z = \mathcal{F}(\eta) = \mathcal{F}(y + \eta - y)$ 在 y 处 Taylor 展开

$$z = \mathcal{F}(y + \eta - y) = \mathcal{F}(y) + \mathcal{F}'(y)(\eta - y) + \cdots + \frac{\mathcal{F}^{(p)}(\xi)}{p!}(\eta - y)^p.$$

把 $z = \mathcal{F}(\eta)$ 换成 $\alpha = \mathcal{F}(0)$, 即 $\eta = 0$ (再注意 $y = f(x)$) 有

$$(4.3) \quad \alpha = x + \sum_{j=1}^{p-1} \frac{[-f(x)]^j}{j!} \mathcal{F}^{(j)}(f(x)) + \frac{\mathcal{F}^{(p)}(\xi)}{p!} [-f(x)]^p.$$



非线性方程求根—迭代函数构造

如果我们令 $\gamma_j(x) = \mathcal{F}^{(j)}(f(x))$, 可以定义

$$(4.4) \quad \varphi_p(x) = x + \sum_{j=1}^{p-1} \frac{[-f(x)]^j}{j!} \gamma_j(x).$$

显然有 $\varphi(\alpha) = \alpha$ (因为 $f(\alpha) = 0$). 且我们有以下定理

定理 4.5

设 $f(x)$ 在含有根 α 的区间 $[a, b]$ 上有 p 阶连续导数, 且 $\forall x \in [a, b]$, $f'(x) \neq 0$. 那么上面 (4.4) 式定义的迭代函数至少为 p 阶迭代函数.

△ 对 $x_{k+1} = \varphi_p(x_k)$, 令 $\varepsilon_{k+1} = x_{k+1} - \alpha = \varphi_p(x_k) - \alpha$

在 (4.3) 中令 $x = x_k$ 有 $\alpha = \varphi_p(x_k) + \frac{\mathcal{F}^{(p)}(\xi)}{p!} [-f(x_k)]^p$.



非线性方程求根—迭代函数构造

$$\text{即: } \varepsilon_{k+1} = -\frac{\mathcal{F}^{(p)}(\xi)}{p!} [-f(x_k)]^p$$

$$\begin{aligned} &= (-1)^{p+1} \frac{\mathcal{F}^{(p)}(\xi)}{p!} [\underbrace{f(\alpha) + f'(\eta)(x_k - \alpha)}_{{\color{red}\text{中值定理}}}]^p \\ &= (-1)^{p+1} \frac{\mathcal{F}^{(p)}(\xi)}{p!} [f'(\eta)]^p \varepsilon_k^p, \quad \text{其中 } \eta \in (x_k, \alpha). \end{aligned}$$

$$\implies \frac{\varepsilon_{k+1}}{\varepsilon_k^p} = \frac{(-1)^{p+1} \mathcal{F}^{(p)}(\xi)}{p!} [f'(\eta)]^p,$$

由 f, \mathcal{F} 光滑 $\implies |\mathcal{F}^{(p)}(\xi)| \leq C_1, |f'(\eta)| \leq C_2$

$$\implies \frac{|\varepsilon_{k+1}|}{|\varepsilon_k|^p} \leq \frac{C_1 C_2}{p!} \leq C \text{ 为常数. 即 } \varphi_p(x) \text{ 至少为 } p \text{ 阶迭代函数. } \triangleright$$



非线性方程求根—迭代函数构造

下面看如何计算 $\gamma_j(x) = \mathcal{F}^{(j)}(f(x))$?

由 \mathcal{F} 为 f 之反函数, 有: $x = \mathcal{F}(f(x))$.

对 x 求导(复合函数求导): $1 = \mathcal{F}'(f(x))f'(x)$

继续对 x 求导得：

$$0 = \mathcal{F}''(f(x))[f'(x)]^2 + \mathcal{F}'(f(x))f''(x)$$

$$0 = \mathcal{F}'''(f(x))[f'(x)]^3 + 3\mathcal{F}''(f(x))f''(x)f'(x) + \mathcal{F}'(f(x))f'''(x)$$

•
•
•

$$\text{即有: } \gamma_1(x) = \frac{1}{f'(x)}, \quad \gamma_2(x) = -\frac{f''(x)\gamma_1(x)}{[f'(x)]^2},$$

$$\gamma_3(x) = -\frac{3f'(x)f''(x)\gamma_2(x) + f'''(x)\gamma_1(x)}{[f'(x)]^3}, \dots$$

非线性方程求根—迭代函数构造

例 4.4

仍回到上一个例子 $x^3 - x - 1 = 0$. 我们来看之前的迭代公式:

$\varphi_1(x) = \sqrt[3]{x + 1}$, $\varphi'_1(\alpha) \neq 0$, 这显然是个一阶收敛格式.

$\varphi_2(x) = \frac{2x^3 + 1}{3x^2 - 1}$, 因为 $\varphi'_2(\alpha) = 0$, 这是个二阶收敛格式.

如果我们令 $\varphi_3(x) = \varphi_2(x) + \frac{(x^3 - x - 1)^2}{2!} \gamma_2(x)$,

仍取 $x_0 = 1.25$, 计算有 $x_1 = \underline{1.3239\dots}$, $x_2 = \underline{1.3247179565\dots}$

已相当于 $\varphi_2(x)$ 计算的第三步结果.

所以三阶格式确实收敛更快.

非线性方程求根—迭代方法的加速

前面小节中介绍的迭代函数的构造方法, 需要函数 f 足够光滑, 另外可以看到计算高阶导数也很复杂.

那么如果本身一种迭代方法收敛很慢(甚至不收敛)的话怎么办呢?
如何对其进行改造(加速)?

Aitken 加速思想:

假设由迭代函数 $\varphi(x)$ 产生的序列 $\{x_k\}_{k=0}^{+\infty} \rightarrow \alpha = \varphi(\alpha)$

利用微分中值定理:

$$x_1 - \alpha = \varphi(x_0) - \varphi(\alpha) = \varphi'(\xi)(x_0 - \alpha), \quad \text{其中 } \xi \in (x_0, \alpha)$$

假设 $\varphi(x)$ 在 α 附近变化不大, 即 $\varphi'(x)$ 在 α 附近约为一个常数 L



非线性方程求根—Aitken加速

则 $x_1 - \alpha \approx L(x_0 - \alpha)$; 类似地, 再迭代一次 $x_2 - \alpha \approx L(x_1 - \alpha)$.

将上面两式相除得到 $\frac{x_2 - \alpha}{x_1 - \alpha} \approx \frac{x_1 - \alpha}{x_0 - \alpha}$

$$\implies x_2 x_0 - (x_2 + x_0)\alpha + \alpha^2 = x_1^2 - 2\alpha x_1 + \alpha^2$$

$$\text{由此解出 } \alpha \approx \frac{x_0 x_2 - x_1^2}{x_2 - 2x_1 + x_0} \equiv x_0 - \frac{(x_1 - x_0)^2}{x_2 - 2x_1 + x_0} \equiv \tilde{x}_0.$$

这表明, 利用 x_1, x_2 可以进一步修正近似 x_0 .

利用两步迭代值修正



非线性方程求根—Aitken加速

可以证明, 按照公式

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

修正后得到的 $\{\tilde{x}_k\}_{k=0}^{+\infty}$ 更快地收敛到 α :

$$\text{即 } \lim_{k \rightarrow \infty} \frac{\tilde{x}_k - \alpha}{x_k - \alpha} = 0.$$

例 4.5 (对于前面 $x^3 - x - 1 = 0$ 的线性收敛迭代函数)

$\varphi_1(x) = \sqrt[3]{x + 1}$, 如果使用 *Aitken* 加速, 可以得到

$\tilde{x}_0 = \underline{1.32475\dots}$, $\tilde{x}_1 = \underline{1.324719\dots}$ 与二阶方法差不多



非线性方程求根—Steffensen加速

Steffensen加速方法:

将 Aitken 加速与不动点迭代相结合, 便得到 Steffensen 加速方法:

令 $y_k = \varphi(x_k)$, $z_k = \varphi(y_k)$, 定义

$$x_{k+1} = x_k - \frac{(y_k - x_k)^2}{z_k - 2y_k + x_k}, \quad k = 0, 1, \dots$$

加速后的值送进phi继续迭代和修正

这相当于定义了一个新的迭代函数 ψ :

$$\psi(x) = x - \frac{(\varphi(x) - x)^2}{\varphi(\varphi(x)) - 2\varphi(x) + x}.$$

也可以看成: 想求 $\varphi(x) = x$ 的根 α , 即 $\varphi(\alpha) - \alpha = 0$.

令 $\varepsilon(x) = \varphi(x) - x$, 自然有 $\varepsilon(\alpha) = 0$.



非线性方程求根—Steffensen加速

已经有 $\varepsilon(x_k) = \varphi(x_k) - x_k = y_k - x_k$, $\varepsilon(y_k) = \varphi(y_k) - y_k = z_k - y_k$.

过 $(x_k, \varepsilon(x_k)), (y_k, \varepsilon(y_k))$ 两点的直线, 与 x -轴交点为

$$\varepsilon(x_k) + \frac{\varepsilon(y_k) - \varepsilon(x_k)}{y_k - x_k}(x - x_k) = 0 \implies x = x_k - \frac{\varepsilon(x_k)}{\varepsilon(y_k) - \varepsilon(x_k)}(y_k - x_k)$$

记此交点为 x_{k+1} , 即 $x_{k+1} = \psi(x_k) \equiv x_k - \frac{(\varphi(x_k) - x_k)^2}{\varphi(\varphi(x_k)) - 2\varphi(x_k) + x_k}$.

例 4.6 (对于前面 $x^3 - x - 1 = 0$ 的线性收敛迭代函数)

$\varphi_1(x) = \sqrt[3]{x + 1}$, 如果使用 Steffensen 加速方法, 可以得到

$$\bar{x}_0 = 1.25, \quad \bar{x}_1 = \underline{1.32475} \cdots, \quad \bar{x}_2 = \underline{1.32471795725} \cdots$$

与三阶方法差不多.



非线性方程求根—Steffensen加速

定理 4.6 (Steffensen加速)

设 $\varphi'(\alpha) \neq 1$ (即 α 为单根), 则 Steffensen 加速法至少是二阶收敛的.

不妨设 $\alpha = 0$ (否则下面可以用 $x - \alpha$ 代替 x), 设 φ 为 p 阶迭代函数, 即 $\varphi'(0) = \dots = \varphi^{(p-1)}(0) = 0, \varphi^{(p)}(0) = p!A \neq 0$, 那么在 $x = 0$ 附近有 $\varphi(x) = Ax^p + \frac{\varphi^{(p+1)}(\theta x)}{(p+1)!}x^{p+1}$, 其中 $\theta \in (0, 1)$.

因而 $[\varphi(x)]^2 = [Ax^p + \mathcal{O}(x^{p+1})]^2 = A^2x^{2p} + \mathcal{O}(x^{2p+1})$.

又 $\varphi(\varphi(x)) = A(Ax^p + \mathcal{O}(x^{p+1}))^p + \mathcal{O}(Ax^p + \mathcal{O}(x^{p+1}))^{p+1}$

$$= \begin{cases} \mathcal{O}(x^{p^2}), & p > 1, \\ A^2x + \mathcal{O}(x^2), & p = 1. \end{cases} \quad \text{考虑到 } \psi(x) = x - \frac{(\varphi(x)-x)^2}{\varphi(\varphi(x))-2\varphi(x)+x}$$

这样 $p > 1$ 时, $\psi(x) = -A^2x^{2p-1} + \mathcal{O}(x^{2p})$;

$p = 1$ 时, 此时 $A = \varphi'(0) \neq 1$, 因而 $\psi(x) = \mathcal{O}(x^2)$. \triangleright



非线性方程求根—Steffensen加速

注 4.1

如果 $p = 1$ 且 $\varphi'(\alpha) = 1$, 表明 α 为 $x - \varphi(x) = 0$ 的 $m (\geq 2)$ 重根, 此时 Steffensen 加速法为一阶收敛.

从上面分析还可看出, 即便 $|\varphi'(\alpha)| > 1$ ($p = 1$), 此时 φ 迭代不收敛, 但仍有 Steffensen 加速法为二阶收敛的.

例 4.7 (对于前面 $x^3 - x - 1 = 0$ 的迭代函数 $\varphi_0(x) = x^3 - 1$)

显然它不收敛: $x_0 = 1.25, x_1 = 0.953125, x_2 = -0.134136, \dots$

但经过 Steffensen 加速之后:

$$\bar{x}_0 = 1.25, \bar{x}_1 = 1.3615, \bar{x}_2 = 1.3306, \bar{x}_3 = \underline{1.3249}, \bar{x}_4 = \underline{1.32471809}.$$

Steffensen 加速最大的好处是不用计算导数也可以得到二阶以上的收敛格式 (通常二阶格式也就够了).



目录

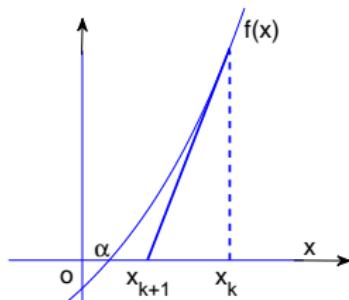
- 1 引言
- 2 线性方程组的直接解法
- 3 线性方程组的迭代解法
- 4 非线性方程(组)迭代解法

- 非线性方程求根
- **Newton迭代法**
- 拟牛顿迭代法
- 非线性方程组求解
- 多项式求根
- 同伦算法



牛顿迭代法

通常来说,二阶格式已经够用.因而工程上使用较多的是前面用到 $f'(x)$ 构造出来的Newton迭代法:



$$\varphi_2(x) = x - [f'(x)]^{-1}f(x).$$

这其实就是一种将 $f(x)$ 线性化的方法.如左图所示,

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

可理解为,过 $(x_k, f(x_k))$ 做 $f(x)$ 的切线,交 x -轴于 x_{k+1} 点.我们知道局部上该点的

切线是曲线在该点附近最好的直线近似.

几何意义: 不停地做切线



牛顿迭代法

或者说, 我们可以看成在 x_k 附近将 $f(x)$ Taylor 展开 (只保留线性部分):

$$f(x) \approx f(x_k) + f'(x_k) \cdot (x - x_k).$$

这样求 $f(x) = 0$ 即为

$$\begin{aligned} 0 &\approx f(x_k) + f'(x_k) \cdot (x_{k+1} - x_k) \\ \implies x_{k+1} &= x_k - [f'(x_k)]^{-1} f(x_k). \end{aligned}$$

显然, $\varphi'(x) = 1 - \frac{[f'(x)]^2 - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2}$.

$f'(\alpha) \neq 0$ 时, 自然有 $\varphi'(\alpha) = 0$.

即单根情形牛顿法为**二阶收敛**.



牛顿迭代法

例 4.8 (求解 $xe^x - 1 = 0$. 易见 $\alpha \in (0, 1)$)

$$\varphi(x) = x - \frac{f}{f'} = x - \frac{xe^x - 1}{e^x + xe^x} = \frac{x^2 + e^{-x}}{1 + x}, \quad x^* = 0.5671432904\cdots$$

取 $x_0 = 0.5$, 有 $x_1 = 0.57102$, $x_2 = 0.\underline{5671}56$, $x_3 = 0.\underline{5671432905}$

如果我们用 $x = e^{-x}$ 来迭代, 需要15次才能得到 $x_{15} = 0.\underline{5671}57$



牛顿迭代法

关于牛顿迭代法的收敛性有以下充分性条件定理：

定理 4.7

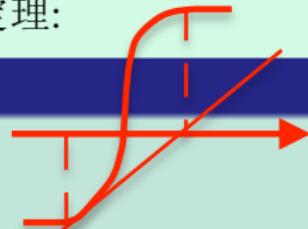
设 $f \in C[a, b]$ 并有二阶导数, 且

- $f(a)f(b) < 0$ (保证 $[a, b]$ 上有零点);
最好在收敛域内保证凹凸性,
 - $f''(x)$ 在 $[a, b]$ 上不变号, 且 $f'(x) \neq 0, \forall x \in [a, b];$
 - 令 $c = a$ 或 $b, s.t. |f'(c)| = \min\{|f'(a)|, |f'(b)|\}$, 且满足

$$\frac{|f(c)|}{|f'(c)|} \leq |f'(c)|$$
 (这为了保证 $\varphi([a, b]) \subset [a, b]$).

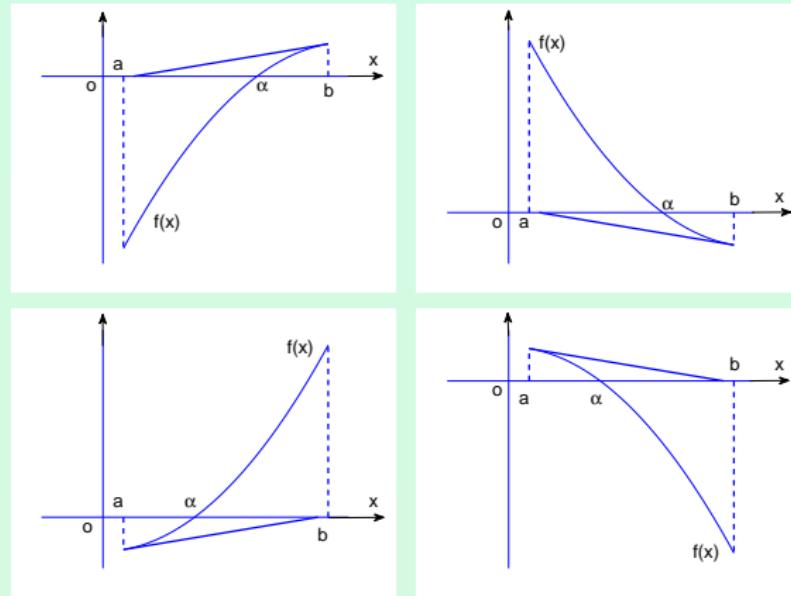
则 $\forall x_0 \in [a, b], x_{k+1} = x_k - [f'(x_k)]^{-1} f(x_k)$ 二阶收敛到 f 的根 α .

因 $f(a)f(b) < 0$, 可无妨设 $f(a) < 0, f(b) > 0$.



牛顿迭代法

又因 $f''(x)$ 不变号, 可无妨设 $f''(x) < 0$ (即 $f'(x)$ 单调下降)
 此外 $f'(x) \neq 0$, 可无妨设 $f'(x) > 0$. 实际共有如下四种情况:



牛顿迭代法

即我们只讨论上面第一种情形, 其他情形完全类似.

由 $f(a) < 0, f(b) > 0$, 可知一定有 $\alpha \in (a, b)$ s.t. $f(\alpha) = 0$.

又因为 $f'(x) > 0$, 即 $f(x)$ 单调上升, 因而只有一个零点 α .

再由假设第三条, 因为 $|f'(b)| < |f'(a)|$, 故取 $c = b$, 由假设, 应该有

$$\frac{f(b)}{b-a} \leq f'(b) \implies b - [f'(b)]^{-1}f(b) \geq a.$$
 此即 $\varphi(b) \geq a.$

下面分两种情况讨论:

1) $\forall x_0 \in [a, \alpha]$, 那么有 $f(x_0) \leq 0$, 注意到有 $f'(x_0) > 0, f''(x_0) < 0$
 $\implies x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \geq x_0$. 另外由 $\varphi'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}$
 $\implies \varphi'(x)$ 在 $[a, \alpha]$ 上为正 $\implies x_1 = \varphi(x_0) \leq \varphi(\alpha) = \alpha$.
 即若 $x_0 \in [a, \alpha] \implies \{x_k\} \subset (a, \alpha]$ 且 $\{x_k\} \uparrow \implies x_k \rightarrow x^*$.



牛顿迭代法

2) $\forall x_0 \in [\alpha, b]$, 那么有 $f(x_0) \geq 0$, 注意到有 $f'(x_0) > 0, f''(x_0) < 0$

仍然由 $\varphi'(x) = \frac{f(x)f''(x)}{[f'(x)]^2} \Rightarrow \varphi'(x)$ 在 $[\alpha, b]$ 上为负

$\Rightarrow \alpha = \varphi(\alpha) \geq x_1 = \varphi(x_0) \geq \varphi(b) \geq a$ (由假设条件3 得到).

这表明如果 $x_0 \in (\alpha, b]$, 迭代一次之后 $x_1 \in [a, \alpha]$, 即回到情形1).

因此结合上面两种情形的讨论可以得到

$$\forall x_0 \in [a, b], x_{k+1} = \varphi(x_k) \xrightarrow{k \rightarrow \infty} \alpha.$$

$\{x_k\}$ 的二次收敛性由 $\varphi'(\alpha) = 0$ 可以得到. \triangleright

牛顿迭代法

从上面的定理可以看到，一般来说牛顿迭代只能保证局部收敛性。

例 4.9 (特殊情形有全局收敛的格式)

求解 $x^2 = C > 0$ 的根 $\sqrt{C} > 0$ 的牛顿迭代函数为：

$$\varphi(x) = x - \frac{f(x)}{f'(x)} = x - \frac{x^2 - C}{2x} = \frac{x^2 + C}{2x}.$$

可以证明此格式 $\forall x_0 > 0, x_{k+1} \rightarrow \sqrt{C}$.

$$\triangleleft x_{k+1} = \frac{x_k^2 + C}{2x_k} = \frac{(x_k \pm \sqrt{C})^2}{2x_k} \mp \sqrt{C}$$

$$\implies \frac{x_{k+1} - \sqrt{C}}{x_{k+1} + \sqrt{C}} = \left(\frac{x_k - \sqrt{C}}{x_k + \sqrt{C}} \right)^2 = \left(\frac{x_0 - \sqrt{C}}{x_0 + \sqrt{C}} \right)^{2^{k+1}} \rightarrow 0$$

即 $x_k \rightarrow \sqrt{C}$. \triangleright



重根的处理办法

前面都假设 $f'(x) \neq 0$, 即单根情形. 但我们总会遇到**重根情况**.

下面设 α 为 $f(x)$ 的 $m \in \mathbb{N}$ 重根, 即

$$f(x) = (x - \alpha)^m g(x), g(\alpha) \neq 0.$$

当 $m \geq 2$ 时, 牛顿迭代格式**不再具有二阶收敛性**, 计算可得

$$\varphi'(x) = \frac{f(x)f''(x)}{[f'(x)]^2} \implies \varphi'(\alpha) = 1 - \frac{1}{m} \neq 0.$$

即牛顿迭代此时只能是线性收敛.

如果我们知道 m 的值, 可以如下修改迭代函数

$$(4.5) \quad \psi(x) = x - m \frac{f(x)}{f'(x)} \implies \psi'(\alpha) = 0.$$

仍可以有二阶收敛性. 但通常 m 的值不知道, 我们可以如下处理.



重根的处理办法

假设 f 充分光滑, 令 $\mu(x) = \frac{f(x)}{f'(x)}$, 显然有 α 为 $\mu(x)$ 的单根.

这样我们对 $\mu(x)$ 使用牛顿迭代便可以得到二阶收敛格式:

$$(4.6) \quad \varphi(x) = x - \frac{\mu(x)}{\mu'(x)} = x - \frac{f(x)f'(x)}{[f'(x)]^2 - f(x)f''(x)}.$$

当然上面需要计算 f 的二阶导数.

如果不想计算高阶导数, 还可以如下考虑:(在 α 附近展开)

$$f(x) = \sum_{j=0}^{m-1} \frac{f^{(j)}(\alpha)}{j!} (x-\alpha)^j + \frac{(x-\alpha)^m f^{(m)}(\xi)}{m!} \equiv \frac{(x-\alpha)^m f^{(m)}(\xi)}{m!}$$

同理, 将 $f'(x)$ 也在 α 附近展开得 $f'(x) = \frac{(x-\alpha)^{m-1} f^{(m)}(\eta)}{(m-1)!}$

$$\Rightarrow \mu(x) = \frac{f(x)}{f'(x)} = \frac{x-\alpha}{m} \left(1 + \mathcal{O}(x-\alpha)\right).$$



重根的处理办法

$$\text{令 } h(x) = \frac{\ln |f(x)|}{\ln |\mu(x)|} = \frac{m \ln |x - \alpha| + \ln \left| \frac{f^{(m)}(\xi)}{m!} \right|}{\ln |x - \alpha| + \ln \left| \frac{1}{m} (1 + \mathcal{O}(x - \alpha)) \right|} \xrightarrow{x \rightarrow \alpha} m$$

因此利用上式和(4.5)式, 可以定义迭代函数为

$$(4.7) \quad \varphi(x) = x - h(x) \frac{f(x)}{f'(x)} \equiv x - \frac{f(x) \ln |f(x)|}{f'(x) (\ln |f(x)| - \ln |f'(x)|)}$$

上式也是渐近二阶格式.

例 4.10 (求 $(x^2 - \frac{1}{8})^2 = 0$ 的二重根 $x = \frac{\sqrt{2}}{4} \approx 0.3535533905 \dots$)

牛顿法 $\varphi(x) = x - \frac{f}{f'} = \frac{3x^2 + \frac{1}{8}}{4x} : x_0 = 0.3, x_3 = 0.348, x_{14} = 0.35355$

(4.6)式 $\varphi(y) = y - \frac{\mu}{\mu'} = \frac{y}{4y^2 + \frac{1}{2}} : y_0 = 0.3, y_3 = 0.353553389 \dots$

(4.7)式: $z_0 = 0.3, z_3 = 0.353556 \dots, z_5 = 0.353553392 \dots$

可以看到 (4.7)式确实是超线性收敛的 (接近于二阶收敛).

目录

- 1 引言
- 2 线性方程组的直接解法
- 3 线性方程组的迭代解法
- 4 非线性方程(组)迭代解法

- 非线性方程求根
- Newton迭代法
- 拟牛顿迭代法
- 非线性方程组求解
- 多项式求根
- 同伦算法

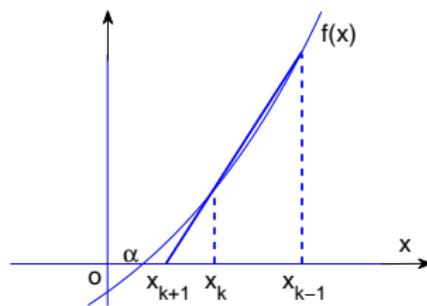


拟牛顿迭代法—弦截法

有时候函数表达式过于复杂(或者没有明确表达式), 为避免求导数, 则需要用别的简单函数来近似 f 或者说 f' :

1) 弦截法(即用线性函数来近似 f):

假设已经计算出来 x_{k-1}, x_k , 过 $(x_{k-1}, f(x_{k-1}))$ 、 $(x_k, f(x_k))$ 做一条直线 $P_1(x) = f(x_k) + \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}(x - x_k)$.



再求 $P_1(x) = 0$ 的点有

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

这相当于用 $\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$ 来近似 $f'(x)$.

几何解释见左图, 故名“弦截法”.



拟牛顿迭代法—弦截法

例 4.11 (求 $xe^x - 1 = 0$ 的零点 $\alpha \approx 0.5671432904 \dots$)

若取 $x_0 = 0.5, x_1 = 0.6$ 用弦截法:

$$x_2 = 0.565315 \dots, x_3 = 0.567095 \dots, x_4 = 0.56714336 \dots$$

可以看到它快于线性收敛, 但慢于牛顿法.

事实上有以下收敛性定理:

定理 4.8

设 f 在 $\Delta = [\alpha - \delta, \alpha + \delta]$ 上有二阶连续导数, $f(\alpha) = 0, f''(\alpha) \neq 0,$

$f'(x) \neq 0$ (即单根), 且 $x_0, x_1 \in \Delta$. 设 $M = \frac{\max_{x \in \Delta} |f''(x)|}{2 \min_{x \in \Delta} |f'(x)|}$.

若 $\delta > 0$ 充分小, s.t. $M\delta < 1$, 则弦截法收敛
 收敛阶为 $\frac{\sqrt{5} + 1}{2}$

(为方程 $p^2 - p - 1 = 0$ 的根)

拟牛顿迭代法—弦截法

△ 记 $P_1(x) = f(x_k) + \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}(x - x_k)$ 为过前两次迭代点 $(x_{k-1}, f(x_{k-1}))$ 、 $(x_k, f(x_k))$ 的直线. 我们有

$$\begin{aligned} f(x) - P_1(x) &= \cancel{f(x_k) + f'(x_k)(x - x_k)} + \frac{f''(\xi)}{2}(x - x_k)^2 \\ &\quad - [\cancel{f(x_k) + f'(x_k)(x - x_k)} - \frac{f''(\eta)}{2}(x_k - x_{k-1})(x - x_k)] \\ &= \frac{f''(\zeta)}{2}(x - x_k)(x - x_{k-1}). \end{aligned}$$

f(a)=0 上式中令 $x = \alpha$, 即 $P_1(\alpha) = -\frac{f''(\zeta)}{2}(\alpha - x_k)(\alpha - x_{k-1})$.

又因为 $P_1(x_{k+1}) = 0 \Rightarrow$

$$P_1(\alpha) - P_1(x_{k+1}) = P'_1(\gamma)(\alpha - x_{k+1}) = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}(\alpha - x_{k+1}).$$

这样令 $\varepsilon_k = x_k - \alpha$, 有 $\varepsilon_{k+1} = -\frac{P_1(\alpha)}{\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}} = \frac{f''(\zeta)}{\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}} \varepsilon_{k-1} \varepsilon_k.$



拟牛顿迭代法—弦截法

令 $M = \max_{\xi_1, \xi_2 \in \Delta} \frac{|f''(\xi_1)|}{2|f'(\xi_2)|}$, 且设 $M\delta < 1$, 有 $|\varepsilon_{k+1}| \leq M|\varepsilon_k| \cdot |\varepsilon_{k-1}|$ 由于

$M\delta < 1, |\varepsilon_0| \leq \delta, |\varepsilon_1| \leq \delta \implies |\varepsilon_{k+1}| \leq \delta, k = 1, 2, \dots$,

即总有 $x_k \in \Delta$. 且 $|\varepsilon_{k+1}| \leq M\delta|\varepsilon_k| \leq (M\delta)^k|\varepsilon_1| \rightarrow 0$, 即迭代收敛.

再看收敛阶: 设 $|\frac{\varepsilon_{k+1}}{\varepsilon_k^p}| \rightarrow C \neq 0$,

即 $|\varepsilon_{k+1}| \sim C|\varepsilon_k|^p, |\varepsilon_k| \sim C|\varepsilon_{k-1}|^p \implies |\varepsilon_{k-1}| \sim C^{-\frac{1}{p}}|\varepsilon_k|^{\frac{1}{p}}$

即 $|\varepsilon_{k+1}| = A|\varepsilon_k| \cdot |\varepsilon_{k-1}| = AC^{-\frac{1}{p}}|\varepsilon_k|^{\frac{1+p}{p}} = C|\varepsilon_k|^p$

即 $\frac{1+p}{p} = p \implies p^2 - p - 1 = 0 \implies p = \frac{1 \pm \sqrt{5}}{2}$.

已经知道迭代收敛, 故 $p = \frac{1+\sqrt{5}}{2} \approx 1.618$. \triangleright



拟牛顿迭代法—抛物线法

类似于前面介绍的弦截法, 如果有了 x_{k-2}, x_{k-1}, x_k , 可以构造一个二次函数 $P_2(x)$ 来近似 $f(x)$, 进而近似计算 $f'(x_k) \approx P'_2(x_k)$, 或者说求 $P_2(x_{k+1}) = 0$ 解出 x_{k+1} . 这里

$$P_2(x) = f(x_k) + f[x_k, x_{k-1}](x - x_k) + f[x_k, x_{k-1}, x_{k-2}](x - x_k)(x - x_{k-1}),$$

$$\text{其中 } f[x_k, x_{k-1}] = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}, f[x_k, x_{k-1}, x_{k-2}] = \frac{f[x_k, x_{k-1}] - f[x_{k-1}, x_{k-2}]}{x_k - x_{k-2}}.$$

解 $P_2(x) = 0$ 得

$$(4.8) \quad x_{k+1}^{\pm} = x_k - \frac{2f(x_k)}{\omega \pm \sqrt{\omega^2 - 4f(x_k)f[x_k, x_{k-1}, x_{k-2}]}}$$

其中 $\omega = f[x_k, x_{k-1}] + f[x_k, x_{k-1}, x_{k-2}](x_k - x_{k-1})$. 这里取‘±’使得 x_{k+1}^{\pm} 更靠近 x_k (因为一般认为 x_k 比之前计算的点更靠近根), 即取‘±’与 ω 符号相同. 可以看到格式 (4.8) 可计算复根.



拟牛顿迭代法—抛物线法

例 4.12 (仍考虑求 $xe^x - 1 = 0$ 的根 $\alpha \approx 0.5671432904 \dots$)

若取 $x_0 = 0.5, x_1 = 0.6, x_2 = 0.565315 \dots$, 用抛物线法:

$\Rightarrow x_3 = 0.567148 \dots$ 可以看到它快于弦截法, 但慢于牛顿法.

我们有以下收敛性定理:

定理 4.9

设 f 在 $\Delta = [\alpha - \delta, \alpha + \delta]$ 上有三阶连续导数, $f(\alpha) = 0, f'''(\alpha) \neq 0,$

$f'(x) \neq 0$ (即单根), 且 $x_0, x_1, x_2 \in \Delta$. 设 $M = \max_{\xi, \eta \in \Delta} \frac{|f'''(\xi)|}{6|f'(\eta)|}$.

若 $\delta > 0$ 充分小, s.t. $M\delta^2 < 1$, 则抛物线法收敛, 收敛阶为 $1.839 \dots$

(为方程 $p^3 - p^2 - p - 1 = 0$ 的根).



目录

- 1 引言
- 2 线性方程组的直接解法
- 3 线性方程组的迭代解法
- 4 非线性方程(组)迭代解法

- 非线性方程求根
- Newton迭代法
- 拟牛顿迭代法
- 非线性方程组求解
- 多项式求根
- 同伦算法



Newton法及其收敛性

记 $\vec{F} = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, 即 $f_i = f_i(x_1, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}$.

欲求解 $\vec{F}(\mathbf{x}) = 0$, 可将一元函数情形推广过来. 记 \vec{F} 的Jacobi矩阵为

$$\vec{F}'(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

向量形式的Newton迭代法为

$$(4.9) \quad \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - [\vec{F}'(\mathbf{x}^{(k)})]^{-1} \cdot \vec{F}(\mathbf{x}^{(k)}) \equiv \Phi(\mathbf{x}^{(k)}),$$

即 $\Phi(\mathbf{x}) = \mathbf{x} - [\vec{F}'(\mathbf{x})]^{-1} \cdot \vec{F}(\mathbf{x})$.



Newton法及其收敛性

定义 4.5 (可微的定义)

设 $\vec{F}(\mathbf{x})$ 在 \mathbf{x} 处存在, 且 $\lim_{\mathbf{h} \rightarrow 0} \frac{\|\vec{F}(\mathbf{x} + \mathbf{h}) - \vec{F}(\mathbf{x}) - \vec{F}'(\mathbf{x})\mathbf{h}\|}{\|\mathbf{h}\|} = 0$, 则称 \vec{F} 在 \mathbf{x} 处可微.

定义 4.6

设存在 \mathbf{x}^* 的一个邻域 S , 当 $\mathbf{x}^{(0)} \in S$ 时, 迭代 $\mathbf{x}^{(k+1)} = G(\mathbf{x}^{(k)})$ 有定义且 $\mathbf{x}^{(k)} \rightarrow \mathbf{x}^*$, 则称 $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 的不动点 \mathbf{x}^* 是迭代序列 $\{\mathbf{x}^{(k)}\}$ 的吸引点, 且称该迭代 局部收敛.



Newton法及其收敛性

定理 4.10 (局部收敛的充分条件)

设 $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 在其不动点 \mathbf{x}^* 处可微, 且 $\rho(G'(\mathbf{x}^*)) < 1$, 则 \mathbf{x}^* 是迭代 $\mathbf{x}^{(k+1)} = G(\mathbf{x}^{(k)})$ 的吸引点 (即该迭代局部收敛).

定理 4.11 (Newton迭代的局部收敛性)

设 $\vec{F}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 在 \mathbf{x}^* 的邻域 S 内可微, 且 $\vec{F}'(\mathbf{x})$ 在 \mathbf{x}^* 处连续, $\vec{F}'(\mathbf{x})$ 在 S 内可逆. 则牛顿迭代(4.9)局部收敛, \mathbf{x}^* 是吸引点, 且有

$$\lim_{k \rightarrow +\infty} \frac{\|\mathbf{x}^{(k+1)} - \mathbf{x}^*\|}{\|\mathbf{x}^{(k)} - \mathbf{x}^*\|} = 0, \quad \text{即超线性收敛.}$$

若存在 $\gamma > 0$, s.t. $\forall \mathbf{x} \in S$, $\|\vec{F}'(\mathbf{x}) - \vec{F}'(\mathbf{x}^*)\| \leq \gamma \|\mathbf{x} - \mathbf{x}^*\|$
(\vec{F}' 为Lipschitz连续), 则该迭代至少为二阶收敛.



Newton法及其收敛性

若再加强一些条件, 则可以得到下面的牛顿迭代收敛的充分性条件.

定理 4.12 (Newton迭代收敛的充分性条件)

设 $\vec{F}: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ 在凸集 D 上可微, 且

- ① 存在 $\gamma > 0$ s.t. $\forall \mathbf{x}, \mathbf{y} \in D$, $\| \vec{F}'(\mathbf{x}) - \vec{F}'(\mathbf{y}) \| \leq \gamma \|\mathbf{x} - \mathbf{y}\|$;
- ② 在 D 上, $\vec{F}'(\mathbf{x})$ 可逆, 且存在 $\beta > 0$, s.t. $\forall \mathbf{x} \in D$, $\|[\vec{F}'(\mathbf{x})]^{-1}\| \leq \beta$;
- ③ 对 $\mathbf{x}^{(0)} \in D$, 令 $\alpha = \|[\vec{F}'(\mathbf{x}^{(0)})]^{-1} \cdot \vec{F}'(\mathbf{x}^{(0)})\| \geq 0$, 有 $q = \alpha\beta\gamma < \frac{1}{2}$;
- ④ 记 $S = \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}^{(0)}\| \leq 2\alpha\}$, 有 $S \subset D$.

那么 \vec{F} 在 S 中有唯一零点 \mathbf{x}^* , 且牛顿迭代(4.9)有定义, $\mathbf{x}^{(k+1)} \rightarrow \mathbf{x}^*$, 且有 $\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \leq 2\alpha q^{2^k - 1}$, $k = 0, 1, \dots$.



Newton法及其收敛性

△ 我们分四步来证明该定理.

$$\begin{aligned} 1) \text{ 由 } f_j(\mathbf{x}) - f_j(\mathbf{y}) &= \int_0^1 \frac{d}{d\lambda} f_j(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) d\lambda \\ &= \sum_{k=1}^n \int_0^1 \frac{\partial f_j}{\partial x_k}(\lambda \mathbf{x} + (1-\lambda)\mathbf{y})(x_k - y_k) d\lambda \end{aligned}$$

写成向量形式: $\vec{F}(\mathbf{x}) - \vec{F}(\mathbf{y}) = \int_0^1 \vec{F}'(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) d\lambda$

因而: $\vec{F}(\mathbf{x}) - \vec{F}(\mathbf{y}) - \vec{F}'(\mathbf{z})(\mathbf{x} - \mathbf{y}) = \int_0^1 [\vec{F}'(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) - \vec{F}'(\mathbf{z})] \cdot (\mathbf{x} - \mathbf{y}) d\lambda$

取范数有

$$\| \vec{F}(\mathbf{x}) - \vec{F}(\mathbf{y}) - \vec{F}'(\mathbf{z})(\mathbf{x} - \mathbf{y}) \| \leq \int_0^1 \| \vec{F}'(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) - \vec{F}'(\mathbf{z}) \| \cdot \| \mathbf{x} - \mathbf{y} \| d\lambda$$



Newton法及其收敛性

利用定理第一条假设(关于 $F'(\mathbf{x})$ 的Lipschitz连续性)有

$$\begin{aligned} & \|\vec{F}(\mathbf{x}) - \vec{F}(\mathbf{y}) - \vec{F}'(\mathbf{z})(\mathbf{x} - \mathbf{y})\| \leq \gamma \int_0^1 \|\lambda \mathbf{x} + (1-\lambda)\mathbf{y} - \mathbf{z}\| \cdot \|\mathbf{x} - \mathbf{y}\| d\lambda \\ & \leq \frac{\gamma}{2} \|\mathbf{x} - \mathbf{y}\| \cdot \{\|\mathbf{x} - \mathbf{z}\| + \|\mathbf{y} - \mathbf{z}\|\} \quad \left(\text{三角不等式及 } \int_0^1 \lambda d\lambda = \frac{1}{2} \right) \end{aligned}$$

令 $\mathbf{z} = \mathbf{y}$ 有

$$(4.10) \quad \|\vec{F}(\mathbf{x}) - \vec{F}(\mathbf{y}) - \vec{F}'(\mathbf{y})(\mathbf{x} - \mathbf{y})\| \leq \frac{\gamma}{2} \|\mathbf{x} - \mathbf{y}\|^2.$$

取 $\mathbf{z} = \mathbf{x}^{(0)}$, 则 $\forall \mathbf{x}, \mathbf{y} \in S$, 因 $\|\mathbf{x} - \mathbf{x}^{(0)}\| \leq 2\alpha$, $\|\mathbf{y} - \mathbf{x}^{(0)}\| \leq 2\alpha$, 即:

$$\begin{aligned} \| \vec{F}(\mathbf{x}) - \vec{F}(\mathbf{y}) - \vec{F}'(\mathbf{x}^{(0)})(\mathbf{x} - \mathbf{y}) \| & \leq \frac{\gamma}{2} \|\mathbf{x} - \mathbf{y}\| (\|\mathbf{x} - \mathbf{x}^{(0)}\| + \|\mathbf{y} - \mathbf{x}^{(0)}\|) \\ (4.11) \quad & \leq 2\alpha\gamma \|\mathbf{x} - \mathbf{y}\|. \end{aligned}$$



Newton法及其收敛性

2) 归纳证明 $\|\mathbf{x}^{(k)} - \mathbf{x}^{(0)}\| \leq 2\alpha$, $\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \leq \alpha \cdot q^{2^k - 1}$.

显然 $\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\| = \left\| [\vec{F}'(\mathbf{x}^{(0)})]^{-1} \vec{F}(\mathbf{x}^{(0)}) \right\| = \alpha \leq 2\alpha$.

假设 $\|\mathbf{x}^{(k)} - \mathbf{x}^{(0)}\| \leq 2\alpha$, 自然 $\vec{F}'(\mathbf{x}^{(k)})$ 可逆, 即 $\mathbf{x}^{(k+1)}$ 有定义

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| = \left\| [\vec{F}'(\mathbf{x}^{(k)})]^{-1} \cdot \vec{F}(\mathbf{x}^{(k)}) \right\| \stackrel{\text{假设2}}{\leq} \beta \left\| \vec{F}(\mathbf{x}^{(k)}) \right\|$$

$$\text{由 } \mathbf{x}^{(k)} \text{ 的定义 } \beta \left\| \vec{F}(\mathbf{x}^{(k)}) - [\vec{F}(\mathbf{x}^{(k-1)}) + \vec{F}'(\mathbf{x}^{(k-1)})(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)})] \right\|$$

$$\begin{aligned} & \text{已证1) } \leq \frac{\beta\gamma}{2} \left\| \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)} \right\|^2 \stackrel{\text{归纳假设}}{\leq} \frac{\beta\gamma}{2} \left(\alpha q^{2^{k-1}-1} \right)^2 \stackrel{(q=\alpha\beta\gamma)}{=} \alpha q^{2^k-1}. \end{aligned}$$



Newton法及其收敛性

由三角不等式

$$\begin{aligned}
 \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(0)}\| &\leq \sum_{j=0}^k \|\mathbf{x}^{(j+1)} - \mathbf{x}^{(j)}\| \\
 &\leq \alpha \sum_{j=0}^k q^{2^j-1} \leq \alpha \sum_{j=0}^{+\infty} q^j \\
 &\leq \frac{\alpha}{1-q} \stackrel{\text{假设 } q < \frac{1}{2}}{\leq} 2\alpha.
 \end{aligned}$$



Newton法及其收敛性

3) 再证 $\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \leq 2\alpha q^{2^k-1}$.

$\forall m \in \mathbb{N}$, 由于 $q < \frac{1}{2}$ 知

$$\begin{aligned}
 \|\mathbf{x}^{(k+m)} - \mathbf{x}^{(k)}\| &\leq \sum_{j=1}^m \|\mathbf{x}^{(k+j)} - \mathbf{x}^{(k+j-1)}\| \leq \alpha \sum_{j=1}^m q^{2^{k+j-1}-1} \\
 (4.12) \quad &= \alpha q^{2^k-1} \sum_{j=1}^m (q^{2^k})^{2^{j-1}-1} < \alpha q^{2^k-1} \frac{1}{1-q} \leq 2\alpha q^{2^k-1}
 \end{aligned}$$

因而 $\{\mathbf{x}^{(k)}\}$ 为 Cauchy 列, 即存在 $\mathbf{x}^* \in \mathbb{R}^n$ s.t. $\mathbf{x}^* = \lim_{k \rightarrow \infty} \mathbf{x}^{(k)}$.

由 $\|\mathbf{x}^{(k)} - \mathbf{x}^{(0)}\| \leq 2\alpha$, 让 $k \rightarrow \infty$, 有 $\|\mathbf{x}^* - \mathbf{x}^{(0)}\| \leq 2\alpha$, 即 $\mathbf{x}^* \in S$.

令 (4.12) 式中 $m \rightarrow \infty$, 得 $\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \leq 2\alpha q^{2^k-1}$.

Newton法及其收敛性

4) 最后证 \mathbf{x}^* 为 $\vec{F}(\mathbf{x})$ 在 S 中的唯一零点.

$$\begin{aligned} \text{因 } & \left\| \vec{F}(\mathbf{x}^{(k)}) \right\| \stackrel{\mathbf{x}^{(k+1)} \text{ 的定义}}{=} \left\| \vec{F}'(\mathbf{x}^{(k)})(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) \right\| \\ & \leq \left(\left\| \vec{F}'(\mathbf{x}^{(k)}) - \vec{F}'(\mathbf{x}^{(0)}) \right\| + \left\| \vec{F}'(\mathbf{x}^{(0)}) \right\| \right) \left\| \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \right\| \\ & \stackrel{(4.12) \text{ 中 } m=1}{\leq} \left(\gamma \left\| \mathbf{x}^{(k)} - \mathbf{x}^{(0)} \right\| + \left\| \vec{F}'(\mathbf{x}^{(0)}) \right\| \right) 2\alpha q^{2^k-1} \rightarrow 0 \end{aligned}$$

由 $\vec{F}(\mathbf{x})$ 的连续性得 $\vec{F}(\mathbf{x}^*) = \lim_{k \rightarrow \infty} \vec{F}(\mathbf{x}^{(k)}) = 0$.

令 $\Phi(\mathbf{x}) = \mathbf{x} - [\vec{F}'(\mathbf{x}^{(0)})]^{-1} \cdot \vec{F}(\mathbf{x})$, 有

$$\begin{aligned} \|\Phi(\mathbf{x}) - \Phi(\mathbf{y})\| &= \|[\vec{F}'(\mathbf{x}^{(0)})]^{-1} \cdot [\vec{F}(\mathbf{x}) - \vec{F}(\mathbf{y}) - \vec{F}'(\mathbf{x}^{(0)})(\mathbf{x} - \mathbf{y})]\| \\ &\leq \beta \cdot 2\alpha\gamma \|\mathbf{x} - \mathbf{y}\| = 2q \|\mathbf{x} - \mathbf{y}\| < \|\mathbf{x} - \mathbf{y}\|. \text{ 即 } \Phi \text{ 为压缩映射, 因而在 } S \text{ 中} \\ &\text{至多有一个不动点, } \vec{F}(\mathbf{x}^*) = 0 \iff \Phi(\mathbf{x}^*) = \mathbf{x}^*. \triangleright \end{aligned}$$



解非线性方程组的拟牛顿法

对于高维问题, 计算Jacobi矩阵的计算量就更大了. 为了避免计算 $\vec{F}'(\mathbf{x}^*)$, 通常试图用简单矩阵 A_k 来近似 $\vec{F}'(\mathbf{x}^{(k)})$. 但是如果:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - A_k^{-1} \cdot \vec{F}'(\mathbf{x}^{(k)})$$

中 A_k 近似得不够好, 比如令 $A_k \equiv \vec{F}'(\mathbf{x}^{(0)})$, 那么只能有线性收敛.

因此我们希望在迭代过程中不断修正 A_k , 使得 A_k 越来越靠近 $\vec{F}'(\mathbf{x}^{(k)})$, 从而使得该拟牛顿法具有超线性收敛阶.

对比一维情形的割线法:

$$x_{k+1} = x_k - \left(\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} \right)^{-1} f(x_k) \equiv x_k - a_k^{-1} f(x_k)$$

即 $a_k = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$, 或者说 $a_k \cdot (x_k - x_{k-1}) = f(x_k) - f(x_{k-1})$.



解非线性方程组的拟牛顿法

我们自然希望经过修正后的矩阵 A_{k+1} 满足类似关系

$$(4.13) \quad A_{k+1} (\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = \vec{F}(\mathbf{x}^{(k+1)}) - \vec{F}(\mathbf{x}^{(k)})$$

一般我们采取逐步修正方式, 即 $A_{k+1} = A_k + \Delta A_k$.

记 $\text{rank}(\Delta A_k) = m \geq 1$, 一般 m 越小越好 (这样每步修正工作量较小). 我们这里主要学习一种秩1的拟牛顿法.

要想拟牛顿法具有超线性收敛性, 需要

$$(4.14) \quad \lim_{k \rightarrow +\infty} \frac{\|[A_k - \vec{F}(\mathbf{x}^{(k)})](\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)})\|}{\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|} = 0.$$

如何才能实现呢? 即, 有了一个近似 A_k 之后, 如何修正使得更靠近 \vec{F} ? 我们有以下引理:



解非线性方程组的拟牛顿法

引理 4.1

设 $A, D \in \mathbb{R}^{n \times n}$, 定义 $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 为仿射变换: $F(\mathbf{x}) = D\mathbf{x} + \mathbf{b}$, 其中 $\mathbf{b} \in \mathbb{R}^n$. 令 $\mathbf{y}, \mathbf{y}' \in \mathbb{R}^n$, 设 $\mathbf{y} \neq \mathbf{y}'$, 记 $\mathbf{p} = \mathbf{y} - \mathbf{y}'$,

$\mathbf{q} = F(\mathbf{y}') - F(\mathbf{y}) \equiv D\mathbf{p}$. 我们有矩阵 $A' = A + \frac{(\mathbf{q} - A\mathbf{p})\mathbf{p}^T}{\|\mathbf{p}\|_2^2}$ 满足

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

(注: 该引理表明, 要想近似 Jacobi 矩阵 $D = F'(\mathbf{x})$, 如果已有一个近似 A , 那么如上修正后的 A' 会在二范数意义下比 A 逼近得更好, 且满足 $A'\mathbf{p} = D\mathbf{p} = \mathbf{q}$. 这给了我们一个从 A_k 到 A_{k+1} 的修正思路, 而且可以看到 $\Delta A = \frac{(\mathbf{q} - A\mathbf{p})\mathbf{p}^T}{\|\mathbf{p}\|_2^2}$ 的秩为 1.)



解非线性方程组的拟牛顿法

△ 由 A' 的定义, 计算可得

$$A'\mathbf{p} = A\mathbf{p} + \frac{(\mathbf{q} - A\mathbf{p})\mathbf{p}^T}{\mathbf{p}^T \mathbf{p}} \mathbf{p} = A\mathbf{p} + \mathbf{q} - A\mathbf{p} = \mathbf{q} = D\mathbf{p}.$$

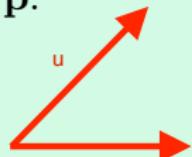
任取单位向量 $\mathbf{u} \in \mathbb{R}^n$, $\|\mathbf{u}\|_2 = 1$, 取 \mathbf{u} 在 \mathbf{p} 上的正交投影, s.t.

$$\mathbf{u} = \alpha\mathbf{p} + \mathbf{v}, \quad \text{即 } \mathbf{p}^T \cdot \mathbf{v} = 0, \quad \text{自然 } \|\mathbf{v}\|_2 \leq 1.$$

这样 $A'\mathbf{v} = A\mathbf{v} + \frac{(\mathbf{q} - A\mathbf{p})\mathbf{p}^T}{\mathbf{p}^T \mathbf{p}} \mathbf{v} = A\mathbf{v}$, 由上面已证 $(A' - D)\mathbf{p} = 0$ 有

$$\begin{aligned} \|(A' - D)\mathbf{u}\|_2 &= \|(A' - D)(\alpha\mathbf{p} + \mathbf{v})\|_2 = \|(A' - D)\mathbf{v}\|_2 \\ &= \|A\mathbf{v}\|_2 \leq \|A - D\|_2 \cdot \|\mathbf{v}\|_2 \leq \|A - D\|_2. \end{aligned}$$

因而 $\|A' - D\|_2 = \sup_{\|\mathbf{u}\|_2=1} \|(A' - D)\mathbf{u}\|_2 \leq \|A - D\|_2$. ▷



解非线性方程组的拟牛顿法

利用上述引理我们得到以下Broyden秩1方法:

算法 4.2 (Broyden秩1拟牛顿方法)

任取初始向量 $\mathbf{x}^{(0)} \in \mathbb{R}^n$, 无妨取初始矩阵 $A_0 = I \in \mathbb{R}^{n \times n}$,

对 $k = 0, 1, \dots$ 计算

- $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - A_k^{-1} \vec{F}(\mathbf{x}^{(k)})$;
- $\mathbf{p}^{(k)} = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}, \mathbf{q}^{(k)} = \vec{F}(\mathbf{x}^{(k+1)}) - \vec{F}(\mathbf{x}^{(k)})$;
- 如果 $\|\mathbf{p}^{(k)}\| \leq \varepsilon$ 则停止迭代;
- 否则就修改 $A_{k+1} = A_k + \frac{(\mathbf{q}^{(k)} - A_k \mathbf{p}^{(k)}) (\mathbf{p}^{(k)})^T}{(\mathbf{p}^{(k)})^T \mathbf{p}^{(k)}}$.



解非线性方程组的拟牛顿法

可以证明, 上面的 A_{k+1} 是极小值问题 $\min_{A \in \mathcal{Q}} \|A - A_k\|_F$ 的解 (思考题).

其中 $\mathcal{Q} = \{A \in \mathbb{R}^{n \times n} | A\mathbf{p}^{(k)} = \mathbf{q}^{(k)}\}$, 这里 $\|A\|_F^2 = \sum_{i,j=1}^n |a_{ij}|^2$.

上面算法中, 我们还需要计算 A_k^{-1} , 或者说需要解方程组 $A_k(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) = -\vec{F}(\mathbf{x}^{(k)})$. 因为一般 A_k 不是稀疏矩阵, 因而该计算量为 $\mathcal{O}(n^3)$. 但是根据 A_k 的构造过程, 我们其实可以用 $\mathcal{O}(n^2)$ 的计算量来实现求解:



解非线性方程组的拟牛顿法

引理 4.2 (Sherman-Morrison)

设 $A \in \mathbb{R}^{n \times n}$, 且 A 可逆, 对 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, 我们有 当且仅当
 $1 + \mathbf{v}^T A^{-1} \mathbf{u} \neq 0$ 时, $A + \mathbf{u}\mathbf{v}^T$ 可逆, 且

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

$$\begin{aligned} &\triangleleft \text{直接计算有: } (A + \mathbf{u}\mathbf{v}^T) \left[A^{-1} - \frac{A^{-1} \cdot \mathbf{u}\mathbf{v}^T \cdot A^{-1}}{1 + \mathbf{v}^T A^{-1} \mathbf{u}} \right] \\ &= I + \mathbf{u}\mathbf{v}^T A^{-1} - \frac{\mathbf{u}\mathbf{v}^T \cdot A^{-1}}{1 + \mathbf{v}^T A^{-1} \mathbf{u}} - \frac{\mathbf{u}\mathbf{v}^T A^{-1} \cdot \mathbf{u}\mathbf{v}^T \cdot A^{-1}}{1 + \mathbf{v}^T A^{-1} \mathbf{u}} \\ &\text{(通分)} = I + \frac{\mathbf{u}\mathbf{v}^T A^{-1} + \mathbf{u}\mathbf{v}^T A^{-1}(\mathbf{v}^T A^{-1} \mathbf{u}) - \mathbf{u}\mathbf{v}^T \cdot A^{-1} - \mathbf{u}(\mathbf{v}^T A^{-1} \mathbf{u}) \mathbf{v}^T \cdot A^{-1}}{1 + \mathbf{v}^T A^{-1} \mathbf{u}} = I. \end{aligned}$$



解非线性方程组的拟牛顿法

这样 $1 + \mathbf{v}^T A^{-1} \mathbf{u} \neq 0$ 时, 显然有 $A + \mathbf{u}\mathbf{v}^T$ 可逆.

反之, 若 $A + \mathbf{u}\mathbf{v}^T$ 可逆 $\Rightarrow I + \mathbf{u}\mathbf{v}^T A^{-1}$ 可逆

$$\Rightarrow (\mathbf{I} + \mathbf{u}\mathbf{v}^T A^{-1})\mathbf{u} = \mathbf{u} + \mathbf{u}\mathbf{v}^T A^{-1}\mathbf{u} = \mathbf{u}(\mathbf{I} + \mathbf{v}^T A^{-1}\mathbf{u})$$

自然可逆矩阵的特征值 $1 + \mathbf{v}^T A^{-1} \mathbf{u} \neq 0$. \triangleright

若令 $H_k = A_k^{-1}$, $A = A_k$, $\mathbf{u} = \frac{\mathbf{q}^{(k)} - A_k \mathbf{p}^{(k)}}{\|\mathbf{p}^{(k)}\|_2^2}$, $\mathbf{v} = \mathbf{p}^{(k)}$

$$\Rightarrow A_{k+1}^{-1} = A_k^{-1} - A_k^{-1} \left(\frac{\mathbf{q}^{(k)} - A_k \mathbf{p}^{(k)}}{\|\mathbf{p}^{(k)}\|_2^2} \mathbf{p}^{(k)T} \right) A_k^{-1} \diagup \left(\mathbf{1} + \mathbf{p}^{(k)T} A_k^{-1} \frac{\mathbf{q}^{(k)} - A_k \mathbf{p}^{(k)}}{\|\mathbf{p}^{(k)}\|_2^2} \right)$$

$$= H_k - \left(\frac{H_k \mathbf{q}^{(k)} \mathbf{p}^{(k)T} H_k - \mathbf{p}^{(k)} \mathbf{p}^{(k)T} H_k}{\|\mathbf{p}^{(k)}\|_2^2} \right) \diagup \left(\frac{\mathbf{p}^{(k)T} H_k \mathbf{q}^{(k)}}{\|\mathbf{p}^{(k)}\|_2^2} \right)$$

$$= H_k - (H_k \mathbf{q}^{(k)} - \mathbf{p}^{(k)}) \mathbf{p}^{(k)T} H_k \diagup \left(\mathbf{p}^{(k)T} H_k \mathbf{q}^{(k)} \right)$$



解非线性方程组的拟牛顿法

这样，上面的算法可以改成

算法 4.3 (改进的Broyden秩1拟牛顿方法)

任取初始向量 $\mathbf{x}^{(0)} \in \mathbb{R}^n$, 无妨取初始矩阵 $H_0 = I \in \mathbb{R}^{n \times n}$,

对 $k = 0, 1, \dots$ 计算

- $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - H_k \vec{F}(\mathbf{x}^{(k)})$;
- $\mathbf{p}^{(k)} = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$, $\mathbf{q}^{(k)} = \vec{F}(\mathbf{x}^{(k+1)}) - \vec{F}(\mathbf{x}^{(k)})$;
- 如果 $\|\mathbf{p}^{(k)}\| \leq \varepsilon$ 则停止迭代;
- 否则就修改
$$H_{k+1} = H_k - (H_k \mathbf{q}^{(k)} - \mathbf{p}^{(k)}) \cdot \frac{(\mathbf{p}^{(k)})^T H_k}{(\mathbf{p}^{(k)})^T H_k \mathbf{q}^{(k)}}.$$

这样每一步迭代的计算量仅为 $\mathcal{O}(n^2)$ (即矩阵 \times 向量的计算量).



目录

- 1 引言
- 2 线性方程组的直接解法
- 3 线性方程组的迭代解法
- 4 非线性方程(组)迭代解法

- 非线性方程求根
- Newton迭代法
- 拟牛顿迭代法
- 非线性方程组求解
- 多项式求根
- 同伦算法



多项式求根

我们在后面常遇到多项式求根的问题，这里特别讨论一下。

令 $p_n(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$. 我们经常需要求一些实系数正交多项式的零点，它们的根均为实的单根，且其范围一般也知道（这在后面会专门讲到），即有 $c_0 < z_n < z_{n-1} < \cdots < z_1 < c_1$.

这样，如果我们从 c_1 出发，用牛顿法迭代，可以证明会很快收敛到 z_1 .
问题是如何继续求剩下的根？

设 $p_n(x) = (x - z_1)p_{n-1}(x)$, 如何快速、高精度地求出 $p_{n-1}(x)$ 的表达式(即其系数)呢？显然我们不能简单地去做辗转相除，那样数值误差一般会较大.



多项式求根

如果我们使用秦九韶算法, 即 $\forall z \in \mathbb{C}$, 若令

$$s_0 = a_0, \quad s_k = a_k + s_{k-1}z, \quad k = 1, \dots, n$$

我们可以得到 $s_n = p_n(z)$. 这样如果我们令

$$p_{n-1}(x) = s_0x^{n-1} + s_1x^{n-2} + \dots + s_{n-2}x + s_{n-1}.$$

简单计算有

$$\begin{aligned} p_{n-1}(x)(x - z) + s_n &= \sum_{m=0}^{n-1} s_m x^{n-1-m} \cdot (x - z) + s_n \\ &= s_n + \sum_{m=0}^{n-1} (s_m x^{n-m} - s_m z x^{n-1-m}) \\ &= s_0 x^n + \sum_{k=1}^n (s_k x^{n-k} - s_{k-1} z x^{n-k}) = \sum_{k=0}^n a_k x^{n-k} \equiv p_n(x) \end{aligned}$$



多项式求根

这说明, 当 z 是 $p_n(x)$ 的根时, 即 $p_n(z) = s_n = 0$ 时] 从上面式子即有 $p_{n-1}(x) = \frac{p_n(x)}{x - z}$. 这样, 当我们计算出 $p_n(x)$ 的一个根 z_1 之后, 按照上面的秦九韶算法计算一遍, 便得到了

$$\left[p_{n-1}(x) = \frac{p_n(x)}{x - z_1} \equiv s_0(z_1)x^{n-1} + s_1(z_1)x^{n-2} + \cdots + s_{n-1}(z_1) \right]$$

然后我们再对 $p_{n-1}(x)$ 用牛顿法, 取初值为 $x_0 = z_1$ 即可, 便可以很快计算出 z_2 , 依次计算下去……

为了提高精度, 在计算完 z_1, z_2, \dots, z_n 之后, 再对 $p_n(x)$ 用它们做初值, 用牛顿法再迭代一遍, 可以得到精度更高的近似值. 如果遇到重根情形(一般来说即迭代收敛很慢时), 同样需要做前面的重根处理, 否则只能是线性收敛.

实系数多项式求复共轭根

为了求多项式的复根，我们一个办法是考虑复数运算，即选取复值的初值用牛顿法或者抛物线法进行计算。

当然如果是实系数多项式求复根，我们知道其复根是成共轭对出现的，故而我们可以避免复数运算：

将 $p_n(x)$ 写成 $p_n(x) = (x^2 - ux - v)q_{n-2}(x) + b_1(x - u) + b_0$ 形式，其中 b_1, b_0 为依赖于 $u, v \in \mathbb{R}$ 的两个实数。我们想求 $p_n(x)$ 的两个复共轭根 z, \bar{z} ，希望它们也是 $x^2 - ux - v = 0$ 的复共轭根。

这样我们只需用牛顿法求解非线性方程组

$$b_1(u, v) = 0, \quad b_0(u, v) = 0$$

得到 u, v 两个实数后再解方程 $x^2 - ux - v = 0$ 得到共轭对 z, \bar{z} 。

Bairstow's 算法

这可以通过 Bairstow's 算法来实现.

定理 4.13 (Bairstow's 算法)

将一个多项式 $p_n(z) = a_n z^n + \cdots + a_1 z + a_0$ 除以一个二次多项式 $z^2 - uz - v$, 其商和余项

$$q(z) = b_n z^{n-2} + b_{n-1} z^{n-3} + \cdots + b_3 z + b_2$$

$$r(z) = b_1(z - u) + b_0$$

可如下递归计算: 令 $b_{n+2} = b_{n+1} = 0$, 对 $k = n, n-1, \dots, 0$,

$$b_k = a_k + ub_{k+1} + vb_{k+2},$$



Bairstow's 算法

△ 对 p_n, q, r , 我们有以下式子

直接爆拆就能算出来

$$p_n(z) = q(z)(z^2 - uz - v) + r(z),$$

或者说有以下式子

$$\sum_{k=0}^n a_k z^k = (z^2 - uz - v) \sum_{k=2}^n b_k z^{k-2} + b_1(z-u) + b_0.$$

列出 z^k 的系数即得

$$a_n = b_n$$

$$a_{n-1} = b_{n-1} - ub_n$$

$$a_k = b_k - ub_{k+1} - vb_{k+2} \quad (0 \leq k \leq n-2)$$

考虑到 $b_{n+2} = b_{n+1} = 0$, 我们就完成了定理证明. \triangleright



Bairstow's 算法

前面我们要计算实系数多项式 $p_n(x)$ 的一对复共轭对根 $\omega, \bar{\omega}$, 按照上面定理的做法, 我们得到 $b_0 = b_0(u, v)$, $b_1 = b_1(u, v)$. 解方程组

$$b_0(u, v) = 0, \quad b_1(u, v) = 0,$$

即得 u, v . 再解 $z^2 - uz - v = 0$ 即得复共轭对根 $\omega, \bar{\omega}$.

下面看如果用 Newton's 迭代法来解 $b_0(u, v) = 0$, $b_1(u, v) = 0$, 我们有了猜测 u, v 后, 如何计算修正值 $\delta u, \delta v$:



Bairstow's 算法

记 $c_k = \frac{\partial b_k}{\partial u}$, $d_k = \frac{\partial b_{k-1}}{\partial v}$, 利用前面定理 4.13 的结论有

$$\left\{ \begin{array}{l} c_k = b_{k+1} + uc_{k+1} + vc_{k+2}, \quad (c_{n+1} = c_n = 0) \\ d_k = b_{k+1} + ud_{k+1} + vd_{k+2}, \quad (d_{n+1} = d_n = 0) \end{array} \right.$$

显然我们只需要计算一个序列 $\{c_k\}$ 就够了. 由牛顿法我们有

$$\begin{aligned} b_0(u, v) + \frac{\partial b_0}{\partial u} \delta u + \frac{\partial b_0}{\partial v} \delta v &= 0 \\ b_1(u, v) + \frac{\partial b_1}{\partial u} \delta u + \frac{\partial b_1}{\partial v} \delta v &= 0 \end{aligned} \implies \begin{pmatrix} c_0 & c_1 \\ c_1 & c_2 \end{pmatrix} \begin{pmatrix} \delta u \\ \delta v \end{pmatrix} = - \begin{pmatrix} b_0(u, v) \\ b_1(u, v) \end{pmatrix}$$

这样结合前面定理 4.13, 我们就完成了 Bairstow's 算法:

$$(u, v) \rightarrow (b_0, b_1) \rightarrow (c_0, c_1, c_2) \rightarrow (\delta u, \delta v) \cdots$$



目录

- 1 引言
- 2 线性方程组的直接解法
- 3 线性方程组的迭代解法
- 4 非线性方程(组)迭代解法

- 非线性方程求根
- Newton迭代法
- 拟牛顿迭代法
- 非线性方程组求解
- 多项式求根
- 同伦算法



同伦算法

前面讨论的非线性方程(组)的求解方法基本上都是局部收敛的，也就是说要初值 \mathbf{x}^0 取得离解 \mathbf{x}^* 足够靠近才可能收敛到 \mathbf{x}^* 。

本节我们讨论的同伦算法（或者也称延拓法）可以作为扩大收敛域的有效方法，希望能从任意 \mathbf{x}^0 出发，通过延拓求得方程

$$(4.15) \quad \mathbf{F}(\mathbf{x}) = \mathbf{0}$$

的解 \mathbf{x}^* . 其基本思想是引入参数 $t \in [0, 1]$, 构造一族同伦映射 \mathbf{H} :

$D \times [0, 1] \subset \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ 代替映射 $\mathbf{F}: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, 使得

$$(4.16) \quad \mathbf{H}(\mathbf{x}^0, 0) = \mathbf{0}, \quad \mathbf{H}(\mathbf{x}, 1) = \mathbf{F}(\mathbf{x}).$$

这表明 $t = 0$ 时, $\mathbf{H}(\mathbf{x}, 0) = \mathbf{0}$ 的解 \mathbf{x}^0 已知, 当 $t = 1$ 时,

$\mathbf{H}(\mathbf{x}, 1) \equiv \mathbf{F}(\mathbf{x}) = \mathbf{0}$ 的解 \mathbf{x}^* 即为 (4.15) 的解.



同伦映射

满足 (4.16) 的同伦映射可以有各种不同的取法, 常见的有以下两种

$$(4.17) \quad \mathbf{H}(\mathbf{x}, t) = \mathbf{F}(\mathbf{x}) + (t - 1)\mathbf{F}(\mathbf{x}^0),$$

及

$$(4.18) \quad \mathbf{H}(\mathbf{x}, t) = t\mathbf{F}(\mathbf{x}) + (1 - t)\mathbf{A}(\mathbf{x} - \mathbf{x}^0),$$

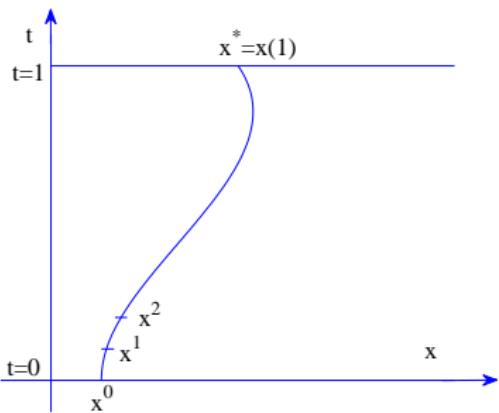
这里 $\mathbf{A} \in \mathbb{R}^{n \times n}$ 为非奇异矩阵, $\mathbf{x}^0 \in \mathbb{R}^n$ 取为已知的点.

显然 (4.17) 和 (4.18) 定义的映射 $\mathbf{H}(\mathbf{x}, t)$ 都满足条件 (4.16). 其中
 (4.17) 式称为 牛顿同伦, (4.18) 称为 凸同伦.

同伦方程

引入了同伦映射后，我们考虑同伦方程

$$(4.19) \quad \mathbf{H}(\mathbf{x}, t) = 0, \quad t \in [0, 1], \quad \mathbf{x} \in D \subset \mathbb{R}^n.$$



如果对于每个 $t \in [0, 1]$, 方程 (4.19) 有解 $\mathbf{x} = \mathbf{x}(t) : [0, 1] \rightarrow D$ 连续.
 $(\mathbf{x}(t), t) \in \mathbb{R}^{n+1}$ 是在 \mathbb{R}^{n+1} 中的一条曲线, 它的一端为已知点 $(\mathbf{x}^0, 0)$, 另一端是点 $(\mathbf{x}^*, 1)$. \mathbf{x}^* 即为方程 (4.15) 的解.
 这就是同伦法, 也称延拓法.

同伦曲线解的性质

关于同伦方程 (4.19) 解曲线 $\mathbf{x} = \mathbf{x}(t)$ 的存在性已经有不少研究, 我们这里只给出牛顿同伦 (4.17) 的一个结果.

定理 4.14

设 $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ 连续可微, 且 $\mathbf{F}'(\mathbf{x})$ 非奇异, 对所有 $\mathbf{x} \in \mathbb{R}^n$ 都有 $\|\mathbf{F}'(\mathbf{x})^{-1}\| \leq \beta$, 则 $\forall \mathbf{x}^0 \in \mathbb{R}^n$, 存在唯一映射 $\mathbf{x} : [0, 1] \rightarrow \mathbb{R}^n$, 使得 (4.19) 成立, 其中 $\mathbf{H}(\mathbf{x}, t)$ 是由 (4.17) 定义的. 且 $\mathbf{x}(t)$ 连续可微, 并有

$$(4.20) \quad \mathbf{x}'(t) = -\mathbf{F}'(\mathbf{x})^{-1} \mathbf{F}(\mathbf{x}^0), \quad \forall t \in [0, 1], \quad \mathbf{x}(0) = \mathbf{x}^0.$$

上面定理条件较强, 事实上一般同伦方程 (4.19) 存在唯一解的条件可以降低到 $\mathbf{H}(\mathbf{x}, t)$ 的 Jacobi 矩阵 $\mathbf{J} = (\frac{\partial \mathbf{H}}{\partial \mathbf{x}}, \frac{\partial \mathbf{H}}{\partial t})$ 满秩即可.



数值延拓法

鉴于课时有限, 关于同伦方程的解的存在唯一性我们课上就不讨论了, 这里仅给出一种数值求解的办法.

首先将区间 $[0, 1]$ 做一个划分

$$0 = t_0 < t_1 < \cdots < t_N = 1.$$

用某种迭代法求解方程组

$$(4.21) \quad \mathbf{H}(\mathbf{x}^k, t_k) = 0, \quad k = 1, 2, \dots, N$$

由于 $k = 0$ 时方程组的解 \mathbf{x}^0 是已知的, 故 $\boxed{\mathbf{x}^0}$ 可以作为 $k = 1$ 是方程组解 \mathbf{x}^1 的初值近似. 一般的, 可用 $k - 1$ 个方程组的解 \mathbf{x}^{k-1} 作为第 k 个方程组解的初值近似. 如果 $t_k - t_{k-1}$ 足够小, 则可望 \mathbf{x}^{k-1} 与 \mathbf{x}^k 很靠近, 从而以上迭代是收敛的.

数值延拓法

例如我们使用牛顿迭代法来求解，则可以得到以下算法：

算法 4.4 (数值延拓法)

对 $k = 1, 2, \dots, N$:

牛顿迭代套娃

$$\mathbf{x}^{k,j+1} = \mathbf{x}^{k,j} - \mathbf{H}'_{\mathbf{x}}(\mathbf{x}^{k,j}, t_k)^{-1} \mathbf{H}(\mathbf{x}^{k,j}, t_k), \quad j = 0, 1, \dots, j_k - 1$$

这里 $\mathbf{x}^{k,0} = \mathbf{x}^{k-1}$, $\mathbf{x}^k = \mathbf{x}^{k,j_k}$. 求得 \mathbf{x}^N 后, 再继续用牛顿法迭代

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \mathbf{H}'_{\mathbf{x}}(\mathbf{x}^k, 1)^{-1} \mathbf{H}(\mathbf{x}^k, 1), \quad k = N, N + 1, \dots$$

直到 $\{\mathbf{x}^k\}_{k=0}^{+\infty}$ 收敛于 \mathbf{x}^* .

只要 $t_k - t_{k-1}$ 足够小, 用牛顿法求解 (4.21) 是收敛的, 因此以上算法就是收敛的, 从而可以求得 $\mathbf{x}^* = \mathbf{x}(1)$.

数值延拓法

由于我们最终仅需 $\mathbf{x}^k \rightarrow \mathbf{x}^*$, 也就是说只要得到 $\mathbf{x}(1)$ 的一个好的近似 \mathbf{x}^N 即可, 这样我们用牛顿法求解 (4.21) 时, 不需要每次都迭代到收敛, 即整个同伦曲线 $\mathbf{x} = \mathbf{x}(t)$ 不需要每个点都很精确.

最粗糙的, 可以取每个 $j_k = 1$, 即中间每步只迭代一次即可. 只要 $t_k - t_{k-1}$ 足够小, 这样就可以得到 $\mathbf{x}(1)$ 的一个好的近似 \mathbf{x}^N . 比如我们采用同伦 (4.17): $\mathbf{H}(\mathbf{x}, t) = \mathbf{F}(\mathbf{x}) + (t - 1)\mathbf{F}(\mathbf{x}^0)$, 取 $j_k = 1$, $t_k = \frac{k}{N}$, 则得如下算法

算法 4.5 (延拓牛顿法)

$$\begin{cases} \mathbf{x}^{k+1} = \mathbf{x}^k - \mathbf{F}'(\mathbf{x}^k)^{-1} [\mathbf{F}(\mathbf{x}^k) + (\frac{k+1}{N} - 1)\mathbf{F}(\mathbf{x}^0)], & k = 0, \dots, N-1 \\ \mathbf{x}^{k+1} = \mathbf{x}^k - \mathbf{F}'(\mathbf{x}^k)^{-1} \mathbf{F}(\mathbf{x}^k), & k = N, N+1, \dots \end{cases}$$

矩阵条件数会影响迭代次数

数值延拓法

以上算法本质上就是牛顿法, 只不过前 N 步的做法是为了得到 \mathbf{x}^* 的一个好的近似 \mathbf{x}^N 而已. 关于上面得到的迭代序列 $\{\mathbf{x}^k\}_{k=0}^{+\infty}$, 我们有以下大范围收敛定理:

定理 4.15

设 $\mathbf{F}: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ 在 \mathbf{x}^0 的邻域 $S = \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}^0\| \leq \beta \|\mathbf{F}(\mathbf{x}^0)\|\} \subset \overset{\circ}{D}$ 内存在连续的 $\mathbf{F}'(\mathbf{x})$, 并且其可逆, 满足

$$\|\mathbf{F}'(\mathbf{x})^{-1}\| \leq \beta < +\infty, \quad \forall \mathbf{x} \in S.$$

及 $\forall \mathbf{x}, \mathbf{y} \in S$, 存在 $\gamma > 0$ 使得

$$\|\mathbf{F}'(\mathbf{x}) - \mathbf{F}'(\mathbf{y})\| \leq \gamma \|\mathbf{x} - \mathbf{y}\|.$$

则以 \mathbf{x}^0 为初值, 存在正整数 $N_0 \geq 2\beta^2\gamma \|\mathbf{F}(\mathbf{x}^0)\|$, 使得 $N \geq N_0$ 时, 上述延拓牛顿法平方收敛于方程 $\mathbf{F}(\mathbf{x}) = 0$ 在 S 中的唯一解 \mathbf{x}^* .

算例

我们来看一个例子

例 4.13

考虑非线性方程组 $\mathbf{F}(\mathbf{x}) = \begin{pmatrix} x_1^2 - x_2 + 1 \\ x_1 - \cos \frac{\pi x_2}{2} \end{pmatrix} = \mathbf{0}$, 例如给定初值 $\mathbf{x}^0 = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$, 欲求 \mathbf{x}^* (有三组解 $\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{3}{2} \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}$)

可以看到, 直接用牛顿法不收敛. 但是如果取 $N = 10$, 使用延拓牛顿法(算法 4.5), 可得 $\mathbf{x}^{16} \approx \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (达到14位有效数字).

