《数值计算方法》课程

解方程组

非线性方程组解法

胡建芳

（研究方向：计算机视觉）

http://sdcs.sysu.edu.cn/content/5143

计算机学院

课程回顾

- **正定矩阵：**

矩阵分解

A=R'R,

最速下降法

共轭梯度法

A是某类特殊矩阵，例如上三角矩阵，下三角矩阵，对角矩阵，对称正定矩阵，怎么求解?

不动点迭代

- **不动点迭代法：**

问题： $F(x)$ 实函数向量. 求 $F(x)=0$ 的近似解。

基本思想方法：

(1) 先将 $F(x)=0$ 化为等价方程 $x=G(x)$ **(6.1)**

(2) 从某个初始向量 $x^{(0)}$ 出发，作向量序列 $\{x^{(k)}\}$

$$x^{(k+1)}=G(x^{(k)}) \longrightarrow \text{(迭代公式)} \quad \textbf{(6.3)}$$

其中 $x=(x_1,x_2,\cdots,x_n)^T$,$G(x)=(g_1(x),g_2(x),\cdots,g_3(x))^T$ 。(6. 3) 式称为简单迭代或单点迭代或单步迭代法。映射 $G(x)$ 称为迭代映射。

假设每个 g_i 有连续的二阶导数： $\dfrac{\partial g_i}{\partial x_p \partial x_q}$,$1\le i,p,q\le n$。

$G(x)$ 在点 x 的 F-导数为：

$$G'(x)=\begin{pmatrix} \dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_1}{\partial x_2} & \cdots & \dfrac{\partial g_1}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial g_n}{\partial x_1} & \dfrac{\partial g_n}{\partial x_2} & \cdots & \dfrac{\partial g_n}{\partial x_n} \end{pmatrix}$$

不动点迭代

- **不动点迭代法：**

设$x^* = (x_1^*, x_2^*, \cdots, x_n^*,)$是(6.1)的解，则称$x^*$为$G(x)$的不动点。

当迭代(6.3)收敛时，极限点 $\tilde{x}$ 又是$G(x)$的连续点，则

$$\tilde{x} = \lim_{k\to\infty} x^{(k+1)} = \lim_{k\leftarrow\infty} G(x^{(k)}) = G(\lim_{k\to\infty} x^{(k)}) = G(\tilde{x})$$

即$\tilde{x}$是$G(x)$的一个不动点。

不动点迭代

- **不动点迭代收敛性：**

定义4 若对$\forall x, y \in D$，存在常数$L<1$，成立$\| G(x)-G(y) \| \le L \| x-y \|$，则映射$G(x)$在区域$D$上称为压缩的。常数$L$称为压缩因子。 (6.4)

定理12（压缩不动点定理） 设映射$G(x)$在区域D上满足：

(1) $G(x) \in D, \forall x \in D$;

(2) $G(x)$在区域D上是压缩映射，压缩因子为L;

则对$\forall x^{(0)} \in D$，简单迭代(6.3)产生的序列$\{x^{(k)}\}$收敛于$G(x)$在区域D上的唯一不动点x^*，且有误差估计：

$$\| x^* - x^{(k)} \| \le \frac{L^k}{1-L} \| x^{(1)} - x^{(0)} \|, k=0,1,\cdots. \quad (6.5)$$

不动点迭代

- **不动点迭代收敛性：**

证明：由条件（1）知所有的 $x^{(k)}$ 全在 D 内，序列 $\{x^{(k)}\}_{k=0}^{\infty}$ 有定义。

首先证明 $G(x)$ 在区域 D 上有唯一不动点

当 $k>0$ 时，$\|x^{(k+1)}-x^{(k)}\|=\|G(x^{(k)})-G(x^{(k-1)})\|$

$$\le L\|x^{(k)}-x^{(k-1)}\|\le\cdots\le L^k\|x^{(1)}-x^{(0)}\|$$

当 $m>k$ 时，$\|x^{(m)}-x^{(k)}\|=\|\sum_{j=k}^{m-1}(x^{(j+1)}-x^{(j)})\|\le\sum_{j=k}^{m-1}\|(x^{(j+1)}-x^{(j)})\|$

$$\le(L^k+\cdots+L^{m-1})\|x^{(1)}-x^{(0)}\|\le\frac{L^k}{1-L}\|x^{(1)}-x^{(0)}\|, \qquad (6.6)$$

对 $\forall\varepsilon>0$，取 k 满足 $\frac{L^k}{1-L}\|x^{(1)}-x^{(0)}\|<\varepsilon$，

则对 $\forall p,q\ge k$，有 $\|x^{(p)}-x^{(q)}\|<\varepsilon$. 因此序列 $\{x^{(k)}\}_{k=0}^{\infty}$ 收敛到某个 $x^*\in D$。

又因 $G(x)$ 在区域 D 上是压缩映射，在x*处连续，所以 x^* 是 $G(x)$ 的一个不动点： $x^*=\lim_{k\to\infty}x^{(k+1)}=\lim_{k\leftarrow\infty}G(x^{(k)})=G(\lim_{k\to\infty}x^{(k)})=G(x^*)$

不动点迭代

- **不动点迭代收敛性：**

其次证唯一性，

设 y^* 为 $x = G(x)$ 在 D 上另一根，则 $y^* = D(y^*), x^* = D(x^*)$，

从而 $\| y^* - x^* \| = \| D(y^*) - D(x^*) \| \le L \| y^* - x^* \|$， $L < 1$

$\therefore y^* = x^*$。即 x^* 是 $G(x)$ 在 D 上的唯一不动点。

(6.4)式较难验证，常采用以下更强的条件代替，

$$\| D'(x) \| \le L < 1, \quad \forall x \in D$$

不动点迭代

- **不动点迭代收敛性：**

定理13（局部收敛定理） 若映射$G(x)$在不动点x^*的δ邻域$D_\delta=\{x \mid \|x-x^*\|\le\delta\}\subset D$上满足对$\forall x\in D_\delta$，有

$$\|G(x)-x^*\|\le L\|x-x^*\|,\quad 0<L<1, \tag{6.8}$$

则对$\forall x^{(0)}\in D_\delta$，由$x^{(k+1)}=G(x^{(k)})$产生的迭代序列$\{x^{(k)}\}$收敛于$x^*$，且有误差估计：

$$\|x^*-x^{(k)}\|\le L^k\|x^*-x^{(0)}\|,k=0,1,\cdots. \tag{6.9}$$

对(6.6)式求$m\to\infty$时的极限即得(6.5)式.

定理14（局部收敛定理） 若映射$G(x)$在不动点x^*处有F导数$G'(x^*)$，而且其谱半径小于1：$\rho(G'(x^*))<1$，则存在$\delta>0$，只要$x^{(0)}\in D_\delta$，由$x^{(k+1)}=G(x^{(k)})$产生的迭代序列$\{x^{(k)}\}$收敛于x^*。

不动点迭代

- **不动点迭代举例：**

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

要求满足精度 $e(k) = \dfrac{\left\| x^{(k)} - x^{(k-1)} \right\|_\infty}{\left\| x^{(k)} \right\|_\infty} \le 10^{-12}$

解：设 $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$，$G(x) = \begin{pmatrix} \frac{1}{3}(\cos x_1 + \sin x_2) \\ \frac{1}{4}(\sin x_1 + \cos x_2) \end{pmatrix}$

其中，$A = \begin{pmatrix} -\frac{1}{3}\sin\xi_1 & \frac{1}{3}\cos\xi_2 \\ \frac{1}{4}\cos\eta_1 & -\frac{1}{4}\sin\eta_2 \end{pmatrix}$，$\therefore \|A\|_\infty \le \dfrac{7}{12}$

则方程组可以改写成 $x = G(x)$，并且对于任意的 $x \in R^2, y \in R^2$，

$$\|G(x) - G(y)\| = \left\| \begin{pmatrix} \frac{1}{3}(\cos x_1 - \cos y_1 + \sin x_2 - \sin y_2) \\ \frac{1}{4}(\sin x_1 - \sin y_2 + \cos x_2 - \cos y_2) \end{pmatrix} \right\|_\infty$$

$$= \|A(x-y)\|_\infty \le \|A\|_\infty \|x-y\|_\infty \le \frac{7}{12}\|x-y\|_\infty$$

此处推导用了中值定理

不动点迭代

不动点迭代举例：

其中， $A=\begin{pmatrix} -\frac{1}{3}\sin\xi_1 & \frac{1}{3}\cos\xi_2 \\ \frac{1}{4}\cos\eta_1 & -\frac{1}{4}\sin\eta_2 \end{pmatrix}$， $\therefore \|A\|_\infty \le \frac{7}{12}$

因此，任取初始向量 $x^{(0)}\in R$，简单迭代法产生序列 $\{x^{(k)}\}$ 收敛于原方程组的唯一解。

$$\text{迭代公式}\begin{cases} x_1^{(k+1)}=\frac{1}{3}\left(\cos x_1^{(k)}+\sin x_2^{(k)}\right) \\ x_2^{(k+1)}=\frac{1}{4}\left(\sin x_1^{(k)}+\cos x_2^{(k)}\right) \end{cases}$$

k	x_1^k	x_2^k	$e(k)$
0	1.000000000000	1.000000000000	
1	0.460591096892	0.345443322669	1.421123164881
2	0.411467922913	0.346350778276	0.119385184710
…	……	……	……
4	0.414178646247	0.337726634268	0.010*********
26	0.415169427139	0.336791217026	0.000000000005
27	0.415169427139	0.336791217025	0.000000000002
28	0.415169427139	0.336791217025	0.000000000001

牛顿迭代法

牛顿迭代法：

基本思想：非线性方程局部线性化（化繁为简）

非线性方程组 $F(x)=0$ (7.1) ,其精确解或真解为 x^*.

$F(x)$ 在 x^* 邻近有连续的 F 导数 $F'(x^*)$, $F'(x^*)$ 非奇异,即 $\det F'(x^*)\neq 0$。并设 $x^{(k)}$ 是(7. 1)的第一个近似解。且 $F(x)$ 在点 $x^{(k)}$ 有 F 导数 $F'(x^{(k)})$,

则仿射映射 $$y=L(x)=F(x^{(k)})+F'(x^{(k)})(x-x^{(k)})$$

是映射(函数) $y=F(x)$ 的局部近似。用 $L(x)$ 近似 $F(x)$，用 $L(x)=0$ 的解作为（7.1）的改进解,即 $x^{(k+1)}=x^{(k)}-[F'(x^{(k)})]^{-1}F(x^{(k)})\quad(\det F'(x^{(k)})\neq 0)$

$$x^{(k+1)}=x^{(k)}-[F'(x^{(k)})]^{-1}F(x^{(k)})\quad(\det F'(x^{(k)})\neq 0)$$

从某 $x^{(0)}$ 出发, 利用上式不断改进得

Newton迭代公式： $$x^{(k+1)}=x^{(k)}-[F'(x^{(k)})]^{-1}F(x^{(k)})$$

牛顿迭代法

- **牛顿迭代法收敛性：**

局部超线性收敛定理

定理15 如果$F(x)$在解x^*邻近有连续的F导数，且$\det F'(x^*)\neq 0$。则存在 $\delta>0$，只要 $\|x^{(0)}-x^*\|\le\delta$，Newton迭代生成的序列$\{x^{(k)}\}$：超线性收敛于x^*，即

$$\lim_{k\to\infty}\frac{\|x^{(k+1)}-x^*\|}{\|x^{(k)}-x^*\|}=0. \qquad (7.6)$$

牛顿法不仅收敛且在一般情况下收敛速度较快，是二阶收敛。

定理16（Newton法局部二阶收敛性）如果$F(x)$的每个分量在解x^*邻近有二阶连续偏导数，$\det F'(x^*)\neq 0$。则存在$\delta>0$，只要$\|x^{(0)}-x^*\|\le\delta$，Newton序列至少二阶收敛于 x^*，即

$$\|x^{(k+1)}-x^*\|\le C\|x^{(k)}-x^*\|^2,\quad k=0,1,2,\cdots\cdots \qquad (7.8)$$

牛顿迭代法

牛顿法改进：下山法

牛顿法收敛速度快，但对初值 x_0 要求苛刻。在实际应用中不容易确定，有时往往由于初值选取不当而使迭代不收敛. Newton 下山法是一种降低对初值要求的修正的牛顿法.

引理2 若 $F(x) \neq 0, F'(x) \neq 0$,则一定存在 $\Delta > 0$,当 $0 < t \leq \Delta$ 时,成立

$$\| F(x - t[F'(x)]^{-1} F(x)) \| < \| F(x) \| \quad (7.9)$$

下山法: Newton法的修正方向 $-[F'(x)]^{-1}F(x)$ 是 $F(x)$ 在 x 点的下山方向。在牛顿法中引进下山因子: $\omega_k \in (0,1)$, 从而由

$$x^{(k+1)} = x^{(k)} - \omega_k [F'(x^{(k)})]^{-1} F(x^{(k)}), \quad k = 0,1,2,\cdots\cdots$$

使 $\| F(x^{(0)}) \| > \| F(x^{(1)}) \| > \cdots$ 呈下山状态。

通常取 $\omega_k \in \{1, \frac{1}{2}, \frac{1}{4}, \cdots\}$,使 $\| F(x^{(k)} - 2^{-i}[F'(x^{(k)})]^{-1} F(x^{(k)})) \| < \| F(x^{(k)}) \|$ 成立的最大值。

牛顿迭代法

- **牛顿法改进：Broyden方法**

牛顿法需要计算一阶导，即雅可比矩阵，当函数不可导时，怎么办?

1. 给定初值 $x^0 \in \mathbb{R}^n$, 初始矩阵 $A_0 \in \mathbb{R}^{n\times n}$ 及精度 ϵ, 令 $k := 0$.

2. 如果 $\|F(x^k)\| > \epsilon$

   2.1. $s^k = -A_k^{-1}F(x^k)$,

   2.2. $x^{k+1} = x^k + s^k$,

   2.3. $y^k = F(x^{k+1}) - F(x^k)$,

   2.4. $A_{k+1} = A_k + \frac{(y^k - A_k s^k)(s^k)^T}{(s^k)^T s^k}$,

   2.5. $k := k + 1$.

作业

- **作业：**

四、（上机题）分别用 Newton 法和 Broyden 法求解下面非线性方程组

$$\begin{cases} 3x_1 - \cos(x_2 x_3) - 0.5 = 0 \\ x_1^2 - 81(x_2 + 0.1)^2 + \sin x_3 + 1.06 = 0 \\ e^{-x_1 x_2} + 20x_3 + \dfrac{1}{3}(10\pi - 3) = 0 \end{cases}$$

（要求：用 Matlab 编程，并附上源代码及迭代五次的结果，初值可取 $(0.1, 0.1, -0.1)$）

THE END