5.3 非线性方程组的迭代法 P131

寻求非线性方程组的计算方法，对求解力学问题、电路问题、经济平衡问题、以及其它非线性问题有着重要的实际意义。本节介绍三种常用的非线性方程组的求解方法：不动点迭代法、牛顿法和最速下降法，这三种方法在求解一般的非线性方程组方面起着重要作用，它们是求解非线性方程组问题的基本方法。

非线性方程组的一般形式为

$$\begin{cases} f_1(x_1, x_2, \cdots, x_n) = 0 \\ f_2(x_1, x_2, \cdots, x_n) = 0 \\ \cdots\cdots\cdots\cdots \\ f_n(x_1, x_2, \cdots, x_n) = 0 \end{cases} \qquad (1)$$

方程组写成向量形式 $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{0}$

$$\boldsymbol{F}: \boldsymbol{R}^n \to \boldsymbol{R}^n, \boldsymbol{x} \in \boldsymbol{R}^n.$$

$$F(x) = \begin{pmatrix} f_1(x_1, x_2, \cdots, x_n) \\ f_2(x_1, x_2, \cdots, x_n) \\ \cdots\cdots\cdots \\ f_n(x_1, x_2, \cdots, x_n) \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

本节介绍三种方法：不动点迭代，牛顿法，最速下降法。

$$\boldsymbol{F}(\boldsymbol{x}) = 0 \quad (1)$$

$$\boldsymbol{F}: \boldsymbol{R}^n \to \boldsymbol{R}^n, \boldsymbol{x} \in \boldsymbol{R}^n.$$

若存在 $x^* \in D \subset R^n$，使 $F(x^*) = 0$，则称 x*是方程组$F(x)=0$的解。

若存在 $x^* \in D \subset R^n$，使 $x^* = G(x^*)$，则称 x*是映射$G(x)$的不动点。

1.不动点迭代法(简单迭代法)

将(1)变成不动点方程 $\boldsymbol{F}(\boldsymbol{x}) = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{G}(\boldsymbol{x}) \quad (2)$

$$\boldsymbol{G}: \boldsymbol{R}^n \to \boldsymbol{R}^n, \boldsymbol{x} \in \boldsymbol{R}^n.$$

则求(1)的根等价于求(2)的不动点，即找x* 使x* =G(X*).

对于不动点方程(2)，取定初值$x^{(0)}$，定义迭代序列：

$$x^{(k+1)} = G(x^{(k)}) \qquad (3)$$

(3)称为不动点迭代法.

G(x)称为迭代映射，不同的迭代映射对应于不同的迭代求解方法。

定义1 设 $F: D \subset R^n \to R^n, x^* \in D$ 使得 $F(x^*)=0$.

构造一个序列$\{x^{(k)}\}$，若$x^{(k)} \in D(k=1,2….)$则称$\{x^{(k)}\}$ 适定。

若 $\lim\limits_{k\to\infty} x^{(k)} = x^*$，即 $\lim\limits_{k\to\infty} \|x^{(k)} - x^*\| = 0$，

称序列$\{x^{(k)}\}$收敛于x* 。

例1 对非线性方程组

$$\begin{cases} x^2 - 10x + y^2 + 8 = 0 \\ xy^2 + x - 10y + 8 = 0 \end{cases}$$

用不动点迭代法求其在(0, 0)附近的根。

解 由原方程组得到它的不动点方程组

$$\begin{cases} x = g_1(x, y) = \dfrac{1}{10}(x^2 + y^2 + 8) \\ y = g_2(x, y) = \dfrac{1}{10}(xy^2 + x + 8) \end{cases}$$

$$\begin{cases} x = g_1(x, y) = \dfrac{1}{10}(x^2 + y^2 + 8) \\ y = g_2(x, y) = \dfrac{1}{10}(xy^2 + x + 8) \end{cases}$$

由迭代结果看出，迭代收敛于方程组的解 $x^* = (1,1)^T$.

由此式构造不动点迭代格式：

$$\begin{cases} x^{(k+1)} = g_1(x^{(k)}, y^{(k)}) = \dfrac{1}{10}[(x^{(k)})^2 + (y^{(k)})^2 + 8)] \\ y^{(k+1)} = g_2(x^{(k)}, y^{(k)}) = \dfrac{1}{10}[(x^{(k)}(y^{(k)})^2 + x^{(k)} + 8)] \end{cases}$$

取初始值 $x^{(0)} = (0,0)^T$，计算结果如下表：

k	0	1	2	….	18	19
$x^{(k)}$	0.0	0.8	0.9280	…	0.999999972	0.999999989
$y^{(k)}$	0.0	0.8	0.9312	…	0.999999972	0.999999989

关于G(x)不动点的存在性和迭代的收敛性，有与非线性方程类似的结果。 P133

定理1（压缩映射原理）设在闭区域$D \subset R^n$上，$G(x)$满足：

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

（2）（压缩性）存在常数$0 < L < 1$，使得对$\forall x, y \in D$，有$\|G(x) - G(y)\| \le L\|x - y\|$。

则$G(x)$在D中存在唯一的不动点x^*，且对任意的$x^{(0)} \in D$，$x^{(k+1)} = G(x^{(k)})$产生的向量序列$\{x^{(k)}\}$收敛到x^*，并有：

$$\|x^{(k)} - x^*\| \le \frac{L}{1-L}\|x^{(k)} - x^{(k-1)}\|$$

压缩映射原理不仅提供了判定迭代收敛的充分条件，还为不动点迭代法的实施奠定了理论基础。

即在设计算法时，可采用不等式 $\|x^{(k)} - x^{(k-1)}\| < \varepsilon$

作为迭代停止的准则。

对于 $x^{(k+1)} = G(x^{(k)}) \quad (3)$

设 $G(x) = (g_1(x_1,\cdots,x_n),\cdots,g_n(x_1,\cdots,x_n))^T$ 且G(x)可导，

P133

$$G'(x) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{bmatrix}$$

称为G(x)的Jacobi矩阵。

定理 2 设 $G(x)$ 有不动点 x^*，且 $G'(x^*)$ 的谱半径

$$\rho(G'(x^*)) = \sigma < 1$$

则存在 x^* 的一个领域 $U(x^*,\delta) = \{x \in R^n \, \| x^* - x \| < \delta\}$，使得对任意的 $x^{(0)} \in U$，迭代序列 $x^{(k+1)} = G(x^{(k)})$ $(k = 0,1,\cdots\cdots)$ 收敛到 x^*。

定理 2 设 $G(x)$ 有不动点 x^*，且 $G'(x^*)$ 的谱半径

$$\rho(G'(x^*)) = \sigma < 1$$

则存在 x^* 的一个领域 $U(x^*,\delta) = \{x \in R^n \, \| x^* - x \| < \delta\}$，使得对任意的 $x^{(0)} \in U$，迭代序列 $x^{(k+1)} = G(x^{(k)})$ $(k = 0,1,\cdots\cdots)$ 收敛到 x^*。

定理2只是不动点迭代局部收敛的充分条件，不是必要条件；

但若 $G(x)$ 是线性函数， $G(x) = Ax + b$， $A \in R^{n\times n}$，

则定理2是迭代收敛的充分必要条件。

在例1中，迭代函数为：

$$G(x) = \begin{pmatrix} g_1(x,y) \\ g_2(x,y) \end{pmatrix} = \begin{pmatrix} \frac{1}{10}(x^2+y^2+8) \\ \frac{1}{10}(xy^2+x+8) \end{pmatrix},$$

$$G(x) = \begin{pmatrix} g_1(x,y) \\ g_2(x,y) \end{pmatrix} = \begin{pmatrix} \frac{1}{10}(x^2+y^2+8) \\ \frac{1}{10}(xy^2+x+8) \end{pmatrix},$$

$$G'(x) = \begin{bmatrix} \frac{x}{5} & \frac{y}{5} \\ \frac{y^2+1}{10} & \frac{xy}{5} \end{bmatrix} \qquad G'(x^*) = \begin{bmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} \end{bmatrix}$$

$\because \rho(G'(x^*)) = 0.4 < 1$ 由定理2,迭代格式收敛。

类似于数列的收敛速度定义，我们可以定义向量序列的收敛速度。

定义2：设迭代序列 $\{x^{(k)}\}_{k=0}^{\infty}$ 收敛于 x^*，且 $x^{(k)} \neq x^*$ 如果

迭代误差 $e_k = \| x^{(k)} - x^* \|$ 满足极限式：

$$\lim_{k\to\infty} \frac{e_{k+1}}{e_k^{\alpha}} = \lim_{k\to\infty} \frac{\| x^{(k+1)} - x^* \|}{\| x^{(k)} - x^* \|^{\alpha}} = \lambda \quad (\alpha > 0)$$

则当 $\lambda > 0$ 时，称该迭代过程是 α 阶收敛的；λ 称为渐进误差常数。特别地，

$\alpha = 1$，$\lambda < 1$ 时称为线性收敛，$\alpha = 2$ 时称为平方收敛.

当 $\lambda = 0$ 时称超 α 阶收敛.

•收敛阶 α 越大，收敛速度越快！

2 Newton 法

P134

类似于非线性方程的Newton迭代法，求解非线性方程组 *F*(x)=0的Newton迭代法也是先将非线性方程组线性化.

设有 $F(x) = 0$

$F: R^n \to R^n, \; x \in R^n.$

$$F(x) = \begin{pmatrix} f_1(x_1,x_2,\cdots,x_n) \\ f_2(x_1,x_2,\cdots,x_n) \\ \cdots\cdots\cdots\cdots \\ f_n(x_1,x_2,\cdots,x_n) \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

设 $f_i(x)$ $(i = 1,2,\cdots,n)$ 在所讨论的某个区域 D 内有一阶连续偏导数，

$x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})^T$ 是方程组的一个近似根，

用多元 Taylor 公式将 $f_i(x)$ 在 $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})^T$ 展开，

取其线性部分，则有：

$$f_i(x) \approx f_i(x^{(k)}) + \sum_{j=1}^{n} \frac{\partial f_i(x^{(k)})}{\partial x_j}(x_j - x_j^{(k)})$$

$$f_i(x) = 0 \Leftrightarrow f_i(x^{(k)}) + \sum_{j=1}^{n} \frac{\partial f_i(x^{(k)})}{\partial x_j}(x_j - x_j^{(k)}) \approx 0$$

$$\Leftrightarrow \sum_{j=1}^{n} \frac{\partial f_i(x^{(k)})}{\partial x_j}(x_j - x_j^{(k)}) \approx -f_i(x^{(k)}), i = 1,2,\cdots,n$$

于是得到非线性方程组的近似线性方程组

$$\begin{cases}\sum_{j=1}^{n}\frac{\partial f_1(\boldsymbol{x}^{(k)})}{\partial x_j}(x_j-x_j^{(k)})=-f_1(\boldsymbol{x}^{(k)}) \\ \cdots\cdots \\ \sum_{j=1}^{n}\frac{\partial f_n(\boldsymbol{x}^{(k)})}{\partial x_j}(x_j-x_j^{(k)})=-f_n(\boldsymbol{x}^{(k)})\end{cases} \quad (1)$$

其系数矩阵为

$$F'(\boldsymbol{x}^{(k)})=\begin{bmatrix}\frac{\partial f_1(\boldsymbol{x}^{(k)})}{\partial x_1} & \frac{\partial f_1(\boldsymbol{x}^{(k)})}{\partial x_2} & \cdots & \frac{\partial f_1(\boldsymbol{x}^{(k)})}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_n(\boldsymbol{x}^{(k)})}{\partial x_1} & \frac{\partial f_n(\boldsymbol{x}^{(k)})}{\partial x_2} & \cdots & \frac{\partial f_n(\boldsymbol{x}^{(k)})}{\partial x_n}\end{bmatrix}$$

P134

线性方程组(1)也可写为 $\boldsymbol{F}'(\boldsymbol{x}^{(k)})\Delta\boldsymbol{x}_k=-\boldsymbol{F}(\boldsymbol{x}^{(k)}) \quad (2)$

线性方程组(1)也可写为 $\boldsymbol{F}'(\boldsymbol{x}^{(k)})\Delta\boldsymbol{x}_k=-\boldsymbol{F}(\boldsymbol{x}^{(k)}) \quad (2)$

这里 $\Delta\boldsymbol{x}_k=\boldsymbol{x}-\boldsymbol{x}^{(k)}$.

当(2)的系数矩阵F'(x $^{(k)}$)可逆时,该方程组的解可表示为:

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}-\left[\boldsymbol{F}'(\boldsymbol{x}^{(k)})\right]^{-1}\boldsymbol{F}(\boldsymbol{x}^{(k)}) \quad (3)$$

(3)称为求解非线性方程组(1)的Newton法迭代格式.

它的迭代函数为: $G(\boldsymbol{x})=\boldsymbol{x}-\left[\boldsymbol{F}'(\boldsymbol{x})\right]^{-1}\boldsymbol{F}(\boldsymbol{x})$

可以证明 $G'(x^*)=-\left(\left[\boldsymbol{F}'(\boldsymbol{x})\right]^{-1}\right)'_{x^*}\boldsymbol{F}(\boldsymbol{x}^*)=\mathrm{O},\quad \therefore \rho(G'(\boldsymbol{x}^*))=0.$

由定理2可得到Newton迭代法局部收敛；不仅如此，还可以证明Newton迭代法是平方收敛。

$$G(\boldsymbol{x})=\boldsymbol{x}-\left[\boldsymbol{F}'(\boldsymbol{x})\right]^{-1}\boldsymbol{F}(\boldsymbol{x})$$

$$\begin{aligned}G'(\boldsymbol{x})&=E-\left(\left[\boldsymbol{F}'(\boldsymbol{x})\right]^{-1}\right)'\boldsymbol{F}(\boldsymbol{x})-\left[\boldsymbol{F}'(\boldsymbol{x})\right]^{-1}\boldsymbol{F}'(\boldsymbol{x}) \\ &=E-\left(\left[\boldsymbol{F}'(\boldsymbol{x})\right]^{-1}\right)'\boldsymbol{F}(\boldsymbol{x})-E \\ &=-\left(\left[\boldsymbol{F}'(\boldsymbol{x})\right]^{-1}\right)'\boldsymbol{F}(\boldsymbol{x})\end{aligned}$$

$$G'(x^*)=-\left(\left[\boldsymbol{F}'(\boldsymbol{x})\right]^{-1}\right)'_{x^*}\boldsymbol{F}(\boldsymbol{x}^*)=\mathrm{O},$$

$$\therefore \rho(G'(\boldsymbol{x}^*))=0.$$

由定理2可得到Newton迭代法局部收敛；不仅如此，还可以证明Newton迭代法是平方收敛。

Newton法的计算机实现: P135

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}-\left[\boldsymbol{F}'(\boldsymbol{x}^{(k)})\right]^{-1}\boldsymbol{F}(\boldsymbol{x}^{(k)}) \quad (3)$$

按公式(3)计算, 每一步均需求逆.

为了克服求逆矩阵的困难， 我们通常解线性方程组(2).

$$\boldsymbol{F}'(\boldsymbol{x}^{(k)})\Delta\boldsymbol{x}_k=-\boldsymbol{F}(\boldsymbol{x}^{(k)}) \quad (2)$$

$$\Delta\boldsymbol{x}^{(k)}=-\left[\boldsymbol{F}'(\boldsymbol{x}^{(k)})\right]^{-1}\boldsymbol{F}(\boldsymbol{x}^{(k)})$$

解出 $\Delta\boldsymbol{x}_k$ 后，即得到 $\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}+\Delta\boldsymbol{x}_k$

这样对 $\boldsymbol{F}'(\boldsymbol{x}^{(k)})$ 是高阶矩阵时就避免了求解逆矩阵的困难，在实际应用中很有意义。

应用牛顿迭代格式求解非线性方程组时， 常用 $\|\mathbf{x}^{(k+1)}-\mathbf{x}^{(k)}\|_\infty<\varepsilon$来控制终止迭代过程。

例：设： $\begin{cases}x^2+y^2-5=0 \\ (x+1)y-3x-1=0\end{cases}$ *用Newton法求其在（1，1）附近的根。*

解：$f_1(x,y)=x^2+y^2-5$，$f_2(x,y)=(x+1)y-3x-1$

$$F(x,y)=\begin{pmatrix}f_1(x,y)\\ f_2(x,y)\end{pmatrix} \quad F'(x,y)=\begin{bmatrix}2x & 2y\\ y-3 & x+1\end{bmatrix}$$

Newton迭代格式为

$$\begin{pmatrix}x^{(k+1)}\\ y^{(k+1)}\end{pmatrix}=\begin{pmatrix}x^{(k)}\\ y^{(k)}\end{pmatrix}-[F'(x^{(k)},y^{(k)})]^{-1}F(x^{(k)},y^{(k)})$$

$$\begin{cases}F'(x^{(k)},y^{(k)})\mathbf{z}=-F(x^{(k)},y^{(k)}) \\ \begin{pmatrix}x^{(k+1)}\\ y^{(k+1)}\end{pmatrix}=\begin{pmatrix}x^{(k)}\\ y^{(k)}\end{pmatrix}+\mathbf{z}\end{cases}$$

取$(x_0,y_0)=(1,1)$，经5次迭代，得（x*,y*)=(1,2)

取$(x_0,y_0)=(-1,-1)$，经6次迭代，得（x*,y*) ≈(-0.6117,-2.1508)

牛顿迭代格式的缺陷:每前进一步,就需解一次方程组.

3.最速下降法 P136 $\begin{cases}f_1(x_1,x_2)=0\\ f_2(x_1,x_2)=0\end{cases} \quad (1)$

对于方程（1），构造函数 Φ（x)=f_1^2(x)+f_2^2(x)， x∈R^2 求Φ（x)的零最小值；即求x*∈R^2,使得Φ（x*)=0。

x*也是方程（1）的解。

步骤：（1）任取一点$x^{(0)}$∈R^2

（2）计算Φ（x)在$x^{(0)}$的负梯度方向

$$G_0=-grad(\Phi(x^{(0)}))=-(\frac{\partial\Phi}{\partial x_1},\frac{\partial\Phi}{\partial x_2})^T\Big|_{x^{(0)}}$$

（3）求正数λ_0,使得

$$\Phi(x^{(0)}+\lambda_0 G_0)=\min_{\lambda>0}\Phi(x^{(0)}+\lambda G_0)$$

(4) 取$x^{(1)}=x^{(0)}+\lambda_0 G_0$,则满足$\Phi(x^{(1)})<\Phi(x^{(0)})$

(5) $x^{(1)}\Rightarrow x^{(0)}$,转（2）。

最速下降法产生的序列$\{x^{(k)}\}$总是收敛的，常用Φ（$x^{(k)}$)<ε来控制终止迭代过程。

例：设： $\begin{cases} x^2+y^2-5=0 \\ (x+1)y-3x-1=0 \end{cases}$ 用最速下降法求其在（1，1)附近的根。

解： $\Phi(x,y)=(x^2+y^2-5)^2+((x+1)y-3x-1)^2$

$grad\Phi(x,y)=$
$(4x(x^2+y^2-5)+2(y-3)((x+1)y-3x-1), 4y(x^2+y^2-5)+2(x+1)((x+1)y-3x-1))^T$

取$X^{(0)}=(x_0,y_0)=(1,1)$，$G_0=-grad\Phi(1,1)=(-4,-20)^T$

$X^{(0)}+\lambda G_0=(1-4\lambda,1-20\lambda)^T$

$\Phi(X^{(0)}+\lambda G_0)=(416\lambda^2-48\lambda-3)^2+(80\lambda^2-32\lambda-2)^2$

（3）求正数λ_0,使得 $\Phi(X^{(0)}+\lambda_0 G_0)=\min\limits_{\lambda>0}\Phi(X^{(0)}+\lambda G_0)$

解得：$\lambda_0=0.0467$，从而 $X^{(1)}=X^{(0)}+\lambda_0 G_0=(1.1869,1.9346)^T$

用最速下降法经16次迭代，得（x^*,y^*)=(1,2).

取$(x_0,y_0)=(-1,-1)$，经6次迭代，得（x^*,y^*) ≈(-0.6117,-2.1508)

取$(x_0,y_0)=(-1,1)$，用最速下降法经64次迭代，得（x^*,y^*)=(1,2).
用 Newton法则不收敛。

方法评价
Newton法具有较快的收敛速度，缺点是对初值要求过高，计算量较大；最速下降法对任意初值总是收敛的，缺点是收敛速度较慢，通常是越靠近解，逼近速度越慢。

对策：将两种方法结合使用，用最速下降法提供初值，之后改用Newton法。

上机作业 3：

数值实验 5
P144:
5-2

要求：1）抄题； 2）迭代公式（初值）或简单原理； 3）程序，结果；（打印）
4）结果分析。