Python 数值分析 算法实践

Python Numerical Analysis Algorithm Practice

第9章 非线性方程组的数值解法

讲授：X X X

日期：2024年4月19日

考虑方程组

$$\begin{cases} F_1(x_1,x_2,\cdots,x_n)=0, \\ F_2(x_1,x_2,\cdots,x_n)=0, \\ \qquad\vdots \\ F_n(x_1,x_2,\cdots,x_n)=0, \end{cases}$$

$$\boldsymbol{F}'(\boldsymbol{x})=\begin{pmatrix} \dfrac{\partial F_1(\boldsymbol{x})}{\partial x_1} & \dfrac{\partial F_1(\boldsymbol{x})}{\partial x_2} & \cdots & \dfrac{\partial F_1(\boldsymbol{x})}{\partial x_n} \\ \dfrac{\partial F_2(\boldsymbol{x})}{\partial x_1} & \dfrac{\partial F_2(\boldsymbol{x})}{\partial x_2} & \cdots & \dfrac{\partial F_2(\boldsymbol{x})}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial F_n(\boldsymbol{x})}{\partial x_1} & \dfrac{\partial F_n(\boldsymbol{x})}{\partial x_2} & \cdots & \dfrac{\partial F_n(\boldsymbol{x})}{\partial x_n} \end{pmatrix}$$

其中$F_1,F_2,\cdots,F_n$均为$x_1,x_2,\cdots,x_n$的多元函数，也可写成$\boldsymbol{F}(\boldsymbol{x})=\boldsymbol{0}$，其中$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in\mathbb{R}^n$，$\boldsymbol{F}=(F_1,F_2,\cdots,F_n)^{\mathrm{T}}$. 当$n\geq 2$且$F_1,F_2,\cdots,F_n$中含有至少一个自变量$x_i(i=1,2,\cdots,n)$的非线性函数时，则称之为非线性方程组. 非线性方程组的求解常与雅可比矩阵关联，向量函数$\boldsymbol{F}(\boldsymbol{x})$的导数$\boldsymbol{F}'(\boldsymbol{x})$称为$\boldsymbol{F}$的雅可比矩阵.

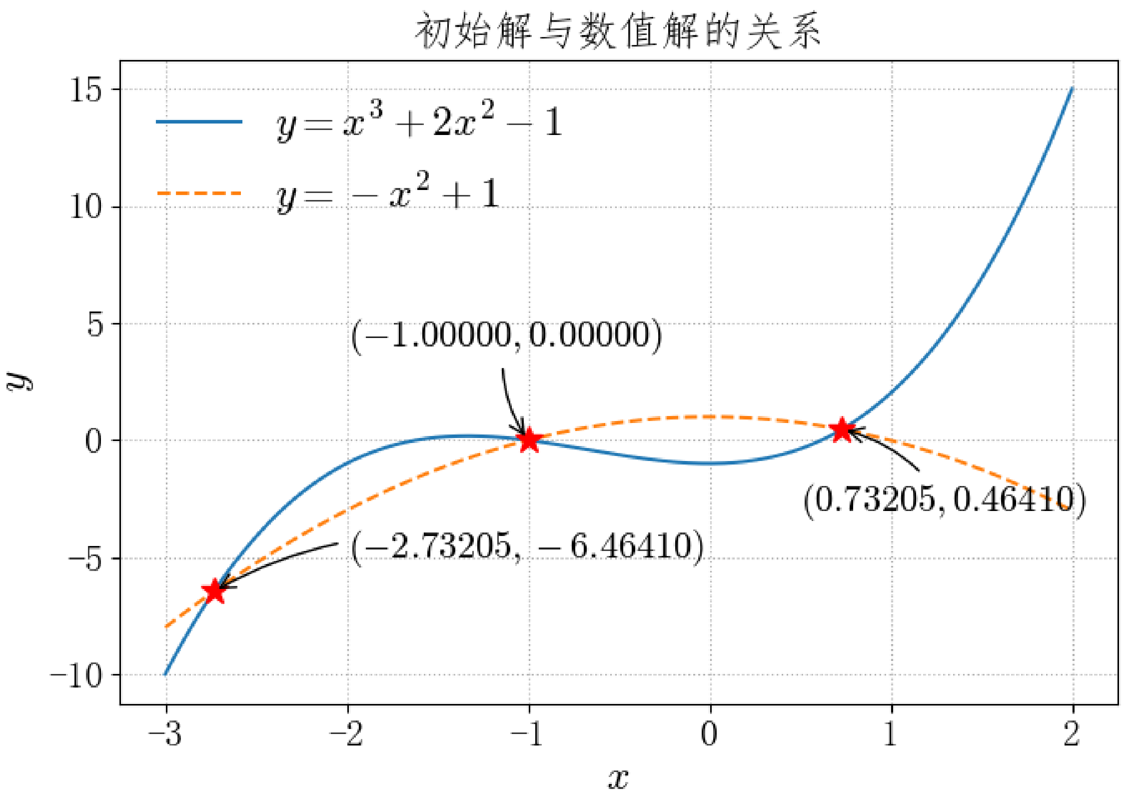
非线性方程组的数值解法

- **9.1迭代初始值的选择问题**
- 9.2不动点迭代法
- 9.3牛顿法
  - 9.3.1牛顿法与牛顿下山法
  - 9.3.2离散牛顿法
  - 9.3.3牛顿–SOR类方法
- 9.4 * 拟牛顿法
  - 9.4.1秩1校正拟牛顿法
  - 9.4.2秩2校正拟牛顿法
- 9.5 * Levenberg – Marquardt方法
  - 9.5.1高斯–牛顿法
  - 9.5.2阻尼最小二乘法
  - 9.5.3全局化LM法
- 9.6 * 同伦延拓法

9.1 迭代初始值的选择问题

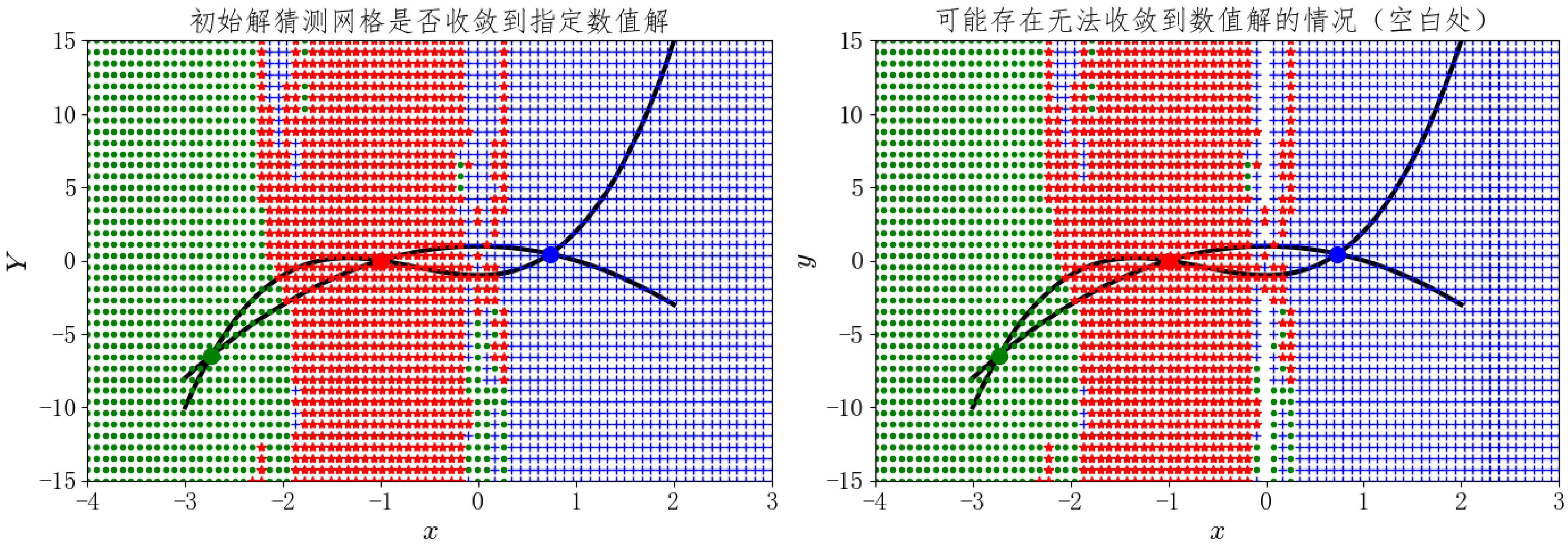
考虑如下非线性方程组，分析初始值的猜测与收敛性的关系.

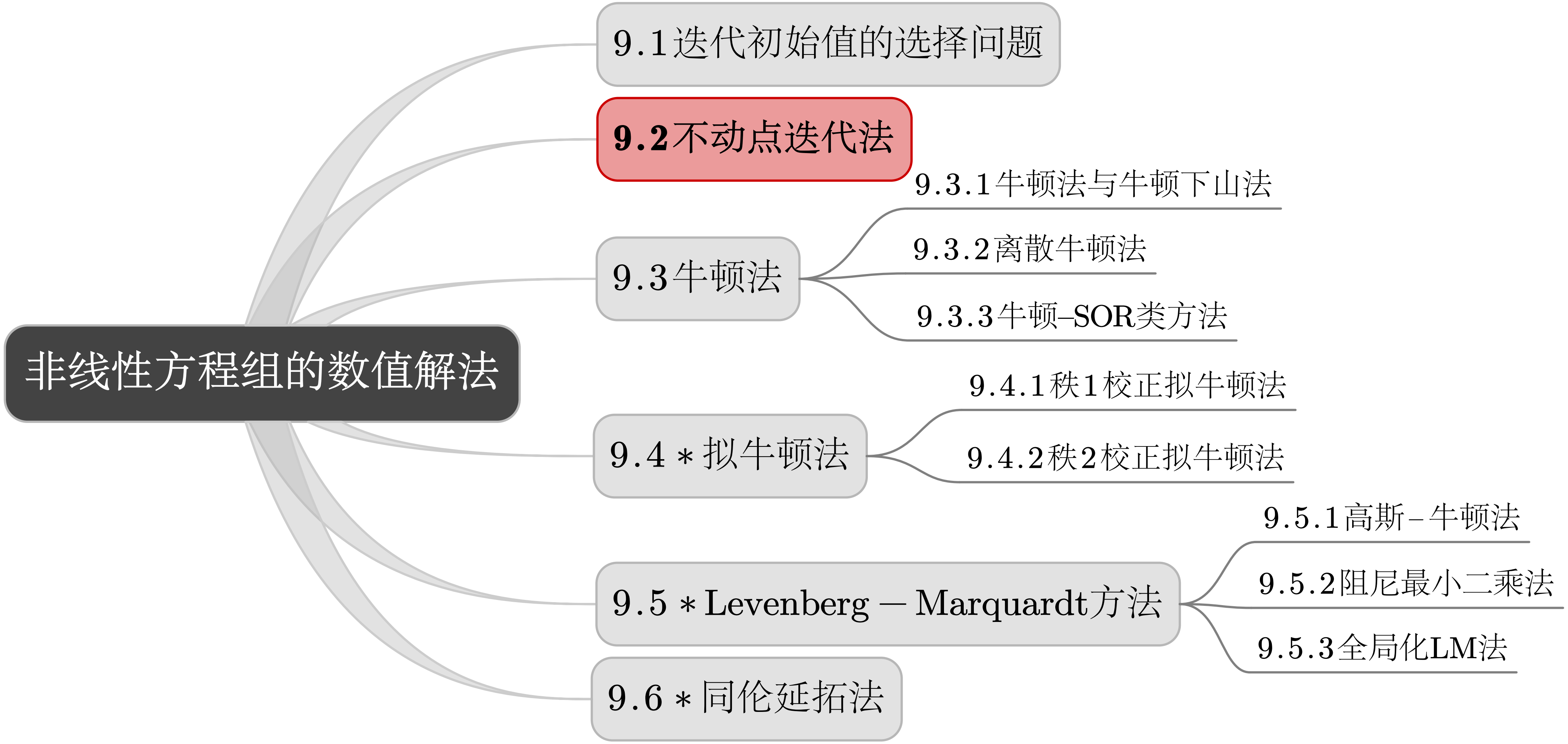
$$\begin{cases} y - x^3 - 2x^2 + 1 = 0, \\ y + x^2 - 1 = 0. \end{cases}$$



9.1 迭代初始值的选择问题

通过使用不同的初始猜测值网格 $[-4,3]\times[-15,15]$(x 轴方向等分 80，y 轴方向等分 40，共组成 3200 个网格点)，对方程组进行系统性求解，建立不同的初始猜测值如何收敛到不同解的可视化图形.




非线性方程组的数值解法
9.1迭代初始值的选择问题
9.2不动点迭代法
9.3牛顿法
9.3.1牛顿法与牛顿下山法
9.3.2离散牛顿法
9.3.3牛顿–SOR类方法
9.4 * 拟牛顿法
9.4.1秩1校正拟牛顿法
9.4.2秩2校正拟牛顿法
9.5 * Levenberg – Marquardt方法
9.5.1高斯–牛顿法
9.5.2阻尼最小二乘法
9.5.3全局化LM法
9.6 * 同伦延拓法

9.2 不动点迭代法

非线性方程组改写为便于迭代的等价形式$\boldsymbol{x}=\boldsymbol{\Phi}(\boldsymbol{x})$，其中向量函数$\boldsymbol{\Phi}(\boldsymbol{x})\in D\subset\mathbb{R}^n$，且在定义域$D$上连续，如果$\boldsymbol{x}^*\in D$，满足$\boldsymbol{x}^*=\boldsymbol{\Phi}(\boldsymbol{x}^*)$，称$\boldsymbol{x}^*$为函数$\boldsymbol{\Phi}$的不动点，$\boldsymbol{x}^*$也就是方程组的一个解.

不动点迭代法的迭代公式

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{\Phi}\left(\boldsymbol{x}^{(k)}\right),\ k=0,1,\cdots.$$

称为不动点迭代法，$\boldsymbol{\Phi}$为迭代函数. 不动点迭代法的收敛性通常与迭代函数和迭代初值有关，但并非任意形式的非线性方程组均易于构造迭代函数.

9.2 不动点迭代法

例1:用不动点迭代法求解方程组，迭代初值$\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$，精度要求$\varepsilon=10^{-16}$.

$$(1)\ \begin{cases} x_1^2-10x_1+x_2^2+8=0, \\ x_1x_2^2+x_1-10x_2+8=0. \end{cases} \qquad (2)\ \begin{cases} 0.5\sin x_1+0.1\cos(x_1x_2)-x_1=0, \\ 0.5\cos x_1-0.1\cos x_2-x_2=0. \end{cases}$$

令$\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}$，把非线性方程组(1)和(2)转换为迭代形式$\boldsymbol{x}^{(k+1)}=\boldsymbol{\Phi}(\boldsymbol{x}^{(k)})$，分别表示为

$$\boldsymbol{\Phi}(\boldsymbol{x})=\begin{pmatrix}\varphi_1(\boldsymbol{x})\\ \varphi_2(\boldsymbol{x})\end{pmatrix}=\begin{pmatrix}0.1x_1^2+0.1x_2^2+0.8\\ 0.1x_1x_2^2+0.1x_1+0.8\end{pmatrix},\ \boldsymbol{\Phi}(\boldsymbol{x})=\begin{pmatrix}\varphi_1(\boldsymbol{x})\\ \varphi_2(\boldsymbol{x})\end{pmatrix}=\begin{pmatrix}0.5\sin x_1+0.1\cos(x_1x_2)\\ 0.5\cos x_1-0.1\cos x_2\end{pmatrix}.$$

9.2 不动点迭代法

对于方程组(1)，$\boldsymbol{\Phi}(\boldsymbol{x})=\begin{pmatrix}\varphi_1(\boldsymbol{x})\\ \varphi_2(\boldsymbol{x})\end{pmatrix}=\begin{pmatrix}0.1x_1^2+0.1x_2^2+0.8\\ 0.1x_1x_2^2+0.1x_1+0.8\end{pmatrix}$.

表 9-1　非线性方程组 (1) 的数值解及精度迭代过程

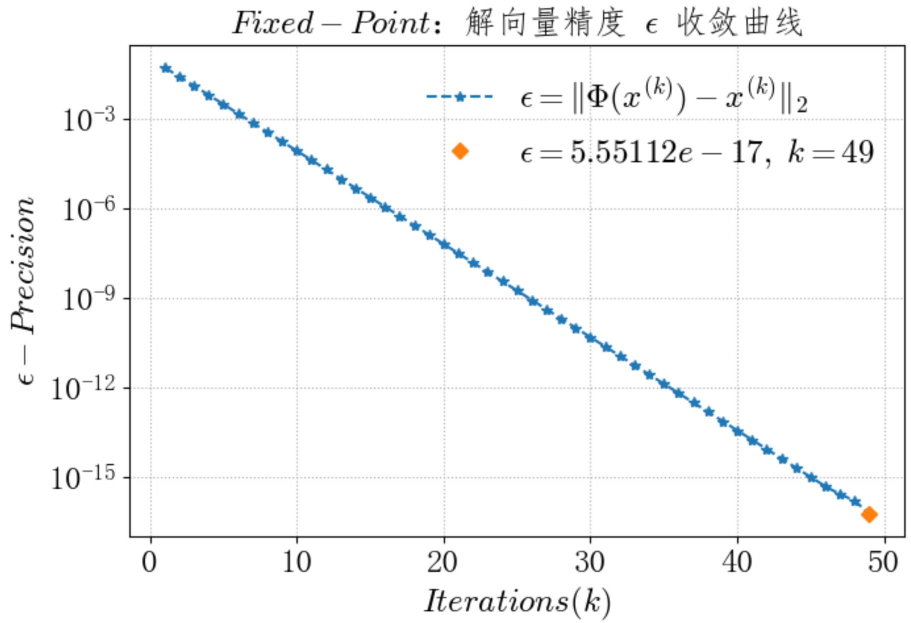
k	x_1	x_2	$\left\Vert \boldsymbol{x}^{(k+1)}-\boldsymbol{x}^{(k)}\right\Vert_2$
1	0.800000000000000	0.800000000000000	1.131370849898476e+00
2	0.928000000000000	0.931200000000000	1.832960446927321e−01
3	0.972831744000000	0.973269983232000	6.147982400123071e−02
⋮	⋮	⋮	⋮
38	1.000000000000000	1.000000000000000	1.570092458683775e−16
39	1.000000000000000	1.000000000000000	3.140184917367550e−16
40	1.000000000000000	1.000000000000000	0.000000000000000e+00

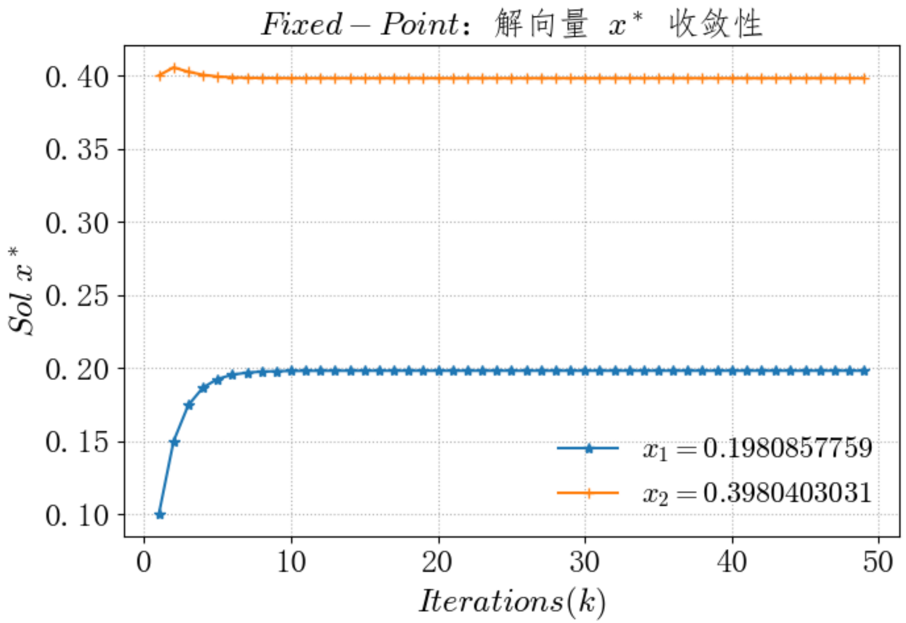
9.2 不动点迭代法

对于方程组(2)，其满足精度要求的近似解以及解的验证精度为

$$\boldsymbol{x}^* = (x_1^*, x_1^*)^{\mathrm{T}} = (0.19808577588668496716, 0.39804030313403238051)^{\mathrm{T}},$$

$$\boldsymbol{F}(\boldsymbol{x}^*) = (F_1(\boldsymbol{x}^*), F_2(\boldsymbol{x}^*))^{\mathrm{T}} = (5.551115123125783\mathrm{e}-17, 0.000000000000000\mathrm{e}+00)^{\mathrm{T}}.$$





- 非线性方程组的数值解法
  - 9.1迭代初始值的选择问题
  - 9.2不动点迭代法
  - **9.3**牛顿法
    - 9.3.1牛顿法与牛顿下山法
    - 9.3.2离散牛顿法
    - 9.3.3牛顿–SOR类方法
  - 9.4 ∗ 拟牛顿法
    - 9.4.1秩1校正拟牛顿法
    - 9.4.2秩2校正拟牛顿法
  - 9.5 ∗ Levenberg − Marquardt方法
    - 9.5.1高斯−牛顿法
    - 9.5.2阻尼最小二乘法
    - 9.5.3全局化LM法
  - 9.6 ∗ 同伦延拓法

9.3 牛顿迭代法

将单个方程的牛顿迭代法应用到方程组，得到牛顿法求解非线性方程组的迭代公式

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}-\boldsymbol{F}'(\boldsymbol{x}^{(k)})^{-1}\boldsymbol{F}(\boldsymbol{x}^{(k)}),\ k=0,1,\cdots.$$

其中$\boldsymbol{F}'(\boldsymbol{x}^{(k)})^{-1}$为雅可比矩阵的逆矩阵. 若固定$\boldsymbol{F}'(\boldsymbol{x}^{(k)})$为$\boldsymbol{F}'(\boldsymbol{x}^{(0)})$，可得简化牛顿法迭代公式

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}-\boldsymbol{F}'(\boldsymbol{x}^{(0)})^{-1}\boldsymbol{F}(\boldsymbol{x}^{(k)}),\ k=0,1,\cdots.$$

牛顿下山法的迭代公式

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}-\lambda(\boldsymbol{F}'(\boldsymbol{x}^{(k)})^{-1})\boldsymbol{F}(\boldsymbol{x}^{(k)}),\ k=0,1,\cdots.$$

其中$\lambda\in(0,1]$. 为了保证收敛，要求λ的取值使得$\|\boldsymbol{F}(\boldsymbol{x}^{(k+1)})\|<\|\boldsymbol{F}(\boldsymbol{x}^{(k)})\|$，$\lambda$逐次减半.

9.3 牛顿迭代法

例2:使用牛顿法和牛顿下山法求解如下非线性方程组的数值解，精度要求$\varepsilon=10^{-15}$.

$$(1)\ \begin{cases} x+2y-3=0, \\ 2x^2+y^2-5=0. \end{cases} \quad (2)\ \begin{cases} 3x-\cos yz-0.5=0, \\ x^2-81(y+0.1)^2+\sin z+1.06=0, \\ e^{-xy}+20z+10\pi/3-1=0. \end{cases}$$

针对非线性方程组(1)，以$\boldsymbol{x}_0=(1.5,1.0)^{\mathrm{T}}$为迭代初值，牛顿法需迭代8次，牛顿下山法需迭代7次，其中在第2次迭代时，需下山因子为0.5. 两种方法收敛的最终解与精度一致，如下所示:

$$\boldsymbol{x}^*=(1.4880338717125849 4348,0.7559830641437075 2826)^{\mathrm{T}},$$

$$\boldsymbol{F}(\boldsymbol{x}^*)=(0.000000000000000\mathrm{e}+00,8.881784197001252\mathrm{e}-16)^{\mathrm{T}}.$$

初值选择不当，可能导致不收敛. 不同的初值也可能收敛到不同的数值解. 如$\boldsymbol{x}_0=(-2,2)^{\mathrm{T}}$，则牛顿下山法经过9次迭代收敛到另一组解，最终解向量及精度为:

$$\boldsymbol{x}^*=(-0.8213672050459180 9178,1.9106836025229592 1242)^{\mathrm{T}},$$

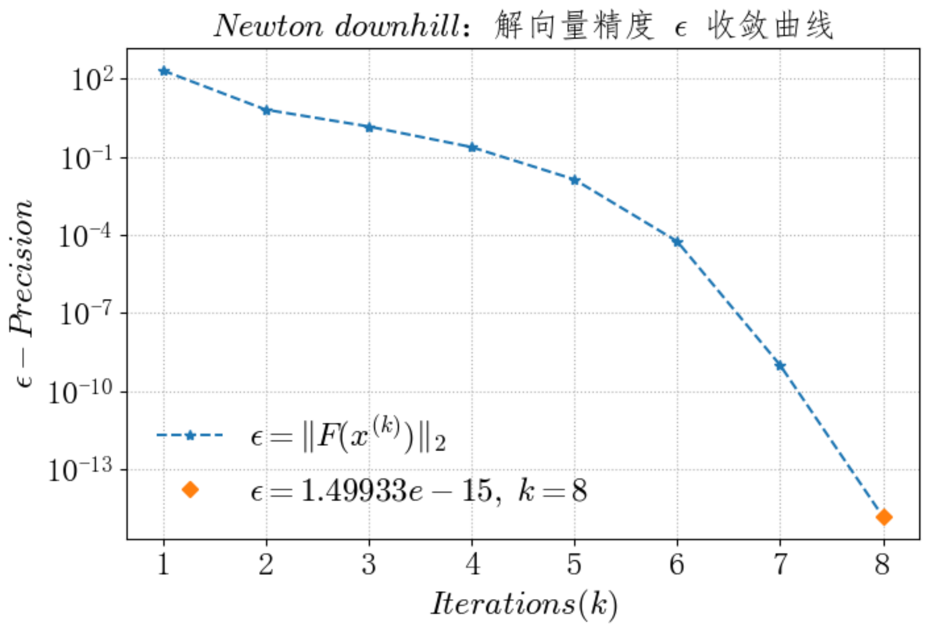
$$\boldsymbol{F}(\boldsymbol{x}^*)=(4.440892098500626\mathrm{e}-16,0.000000000000000\mathrm{e}+00)^{\mathrm{T}}.$$

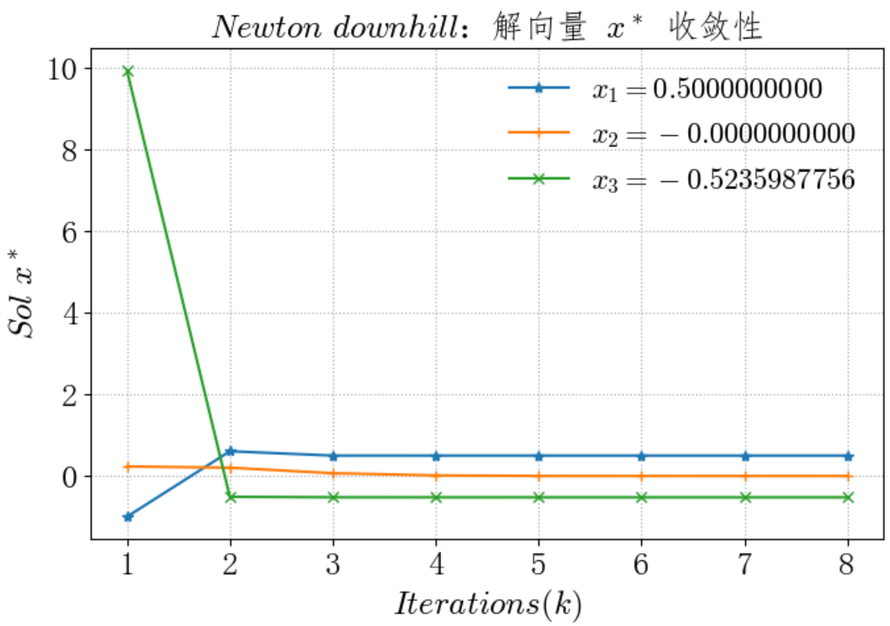
9.3 牛顿迭代法

针对求解非线性方程组(2)，由于计算复杂非线性方程组会引入误差，此处设置精度$\varepsilon=10^{-14}$，以$\boldsymbol{x}_0=(0,0,0)^{\mathrm{T}}$为迭代初值,则牛顿法和牛顿下山法均迭代8次收敛,无需下山因子.最终满足精度的近似解及精度为

$$\boldsymbol{x}^*=(0.500000000000000000000,-0.000000000000000000486,-0.523598775598298926 68)^{\mathrm{T}},$$

$$\boldsymbol{F}(\boldsymbol{x}^*)=(0.000000000000000\mathrm{e}+00,-1.110223024625157\mathrm{e}-16,-1.495211053083399\mathrm{e}-15)^{\mathrm{T}}.$$





9.3 牛顿迭代法

牛顿法的收敛性依赖于初值的选择，若以$\boldsymbol{x}_0=(0.2, 0, -0.2)^{\mathrm{T}}$为迭代初值，精度要求$\varepsilon=10^{-15}$，则需迭代6次，收敛到另外一组解，最终的近似解及精度为：

$$\boldsymbol{x}^*=(0.49814468458949118235, -0.19960589554377985988, -0.52882597757338745126)^{\mathrm{T}},$$

$$\boldsymbol{F}(\boldsymbol{x}^*)=(-1.110223024625157\mathrm{e}-16, 2.220446049250313\mathrm{e}-16, 2.811457863168518\mathrm{e}-16)^{\mathrm{T}}.$$