第3章 解线性方程组的迭代法

回顾: 关于线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \tag{3.1}$$

的直接解法, 其主要不足之处在于:

1. 在有舍入误差的情况下, 一般得到的仍是近似解.
2. 当问题规模较大时, 算法所需的计算量仍然很大, 而且计算过程中所累计误差的影响会“越来越坏”. 因而, 此类算法一般只适用于中小规模的“稠密”型方程组.
3. 算法程序的设计一般也比较复杂.

线性方程组(的古典定常)迭代法的基本思想是：把n元线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \tag{3.1}$$

或

$$\boldsymbol{Ax=b}.$$

改写成等价的方程组

$$\begin{cases} x_1 = m_{11}x_1 + m_{12}x_2 + \cdots + m_{1n}x_n + g_1 \\ x_2 = m_{21}x_1 + m_{22}x_2 + \cdots + m_{2n}x_n + g_2 \\ \cdots\ \cdots \\ x_n = m_{n1}x_1 + m_{n2}x_2 + \cdots + m_{nn}x_n + g_n \end{cases},$$

或

$$\boldsymbol{x=Mx+g}.$$

由此建立方程组的迭代公式

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{M}\boldsymbol{x}^{(k)}+\boldsymbol{g}\,,\quad k=0,\,1,\,2,\ldots \tag{3.2}$$

其中$\boldsymbol{M}$称为迭代矩阵. 对任意取定的初始向量$\boldsymbol{x}^{(0)}$, 由(3.2)式可逐次算出迭代向量$\boldsymbol{x}^{(k)}$, $k=1,2,\ldots$.

如果向量序列$\{\boldsymbol{x}^{(k)}\}$ 收敛于$\boldsymbol{x}^*$, 由(3.2)式可得

$$\boldsymbol{x}^*=\boldsymbol{M}\boldsymbol{x}^*+\boldsymbol{g},$$

从而$\boldsymbol{x}^*$是方程组$\boldsymbol{x}=\boldsymbol{M}\boldsymbol{x}+\boldsymbol{g}$的解, 也就是方程组$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$的解.

这种求解线性方程组的方法称为迭代法, 若迭代序列$\{\boldsymbol{x}^{(k)}\}$收敛, 则称迭代法收敛, 否则称迭代法发散.

迭代法的特点:

1. 可以在迭代过程中对误差进行“校正”;
2. 可以保持问题的稀疏结构. 当收敛性能较好时, 其总体计算量也可能会较小, 因而适用于大规模稀疏型问题的计算.
3. 这类算法的程序设计一般比较简单;
4. 一般只能求得近似解.

问题:

1. 如何建立具体的迭代格式?
2. 如何保证或判断迭代格式是否收敛?
3. 如何控制迭代收敛的速度?

§ 3.1
Jacobi (雅可比)迭代法
和
Gauss-Seidel (高斯-赛德尔)迭代法

考虑线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n \end{cases} \tag{3.1}$$

当$a_{ii}\neq 0$时, i=1, 2,…, n, 由方程组(3.1)得到等价方程组:

$$
\begin{cases}
x_1 = \dfrac{-a_{12}}{a_{11}}x_2 - \dfrac{a_{13}}{a_{11}}x_3 - \cdots - \dfrac{a_{1n}}{a_{11}}x_n + \dfrac{b_1}{a_{11}}, \\
x_2 = \dfrac{-a_{21}}{a_{22}}x_1 - \dfrac{a_{23}}{a_{22}}x_3 - \cdots - \dfrac{a_{2n}}{a_{22}}x_n + \dfrac{b_2}{a_{22}}, \\
\cdots\cdots \\
x_n = \dfrac{-a_{n1}}{a_{nn}}x_1 - \dfrac{a_{n2}}{a_{nn}}x_2 - \cdots - \dfrac{a_{n,n-1}}{a_{nn}}x_{n-1} + \dfrac{b_n}{a_{nn}}.
\end{cases}
$$

从而可构造相应的迭代格式:

$$
\begin{cases}
x_1^{(k+1)} = \dfrac{-a_{12}}{a_{11}}x_2^{(k)} - \dfrac{a_{13}}{a_{11}}x_3^{(k)} - \cdots - \dfrac{a_{1n}}{a_{11}}x_n^{(k)} + \dfrac{b_1}{a_{11}}, \\
x_2^{(k+1)} = \dfrac{-a_{21}}{a_{22}}x_1^{(k)} - \dfrac{a_{23}}{a_{22}}x_3^{(k)} - \cdots - \dfrac{a_{2n}}{a_{22}}x_n^{(k)} + \dfrac{b_2}{a_{22}}, \\
\cdots\cdots \\
x_n^{(k+1)} = \dfrac{-a_{n1}}{a_{nn}}x_1^{(k)} - \dfrac{a_{n2}}{a_{nn}}x_2^{(k)} - \cdots - \dfrac{a_{n,n-1}}{a_{nn}}x_{n-1}^{(k)} + \dfrac{b_n}{a_{nn}}.
\end{cases}
$$

初始向量$\boldsymbol{x}^{(0)}=(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})^T$. $k=0, 1, 2, \ldots$

$$\begin{cases} x_1^{(k+1)} = \qquad \dfrac{-a_{12}}{a_{11}}x_2^{(k)} - \dfrac{a_{13}}{a_{11}}x_3^{(k)} - \cdots - \dfrac{a_{1n}}{a_{11}}x_n^{(k)} + \dfrac{b_1}{a_{11}}, \\ x_2^{(k+1)} = \dfrac{-a_{21}}{a_{22}}x_1^{(k)} \qquad - \dfrac{a_{23}}{a_{22}}x_3^{(k)} - \cdots - \dfrac{a_{2n}}{a_{22}}x_n^{(k)} + \dfrac{b_2}{a_{22}}, \\ \qquad \cdots\cdots \\ x_n^{(k+1)} = \dfrac{-a_{n1}}{a_{nn}}x_1^{(k)} - \dfrac{a_{n2}}{a_{nn}}x_2^{(k)} - \cdots - \dfrac{a_{n,n-1}}{a_{nn}}x_{n-1}^{(k)} \qquad + \dfrac{b_n}{a_{nn}}. \end{cases} \tag{3.3}$$

k=0, 1, 2,…

式(3.3)称为**Jacobi迭代法**, 简称为**J迭代法**.

J迭代法也记为

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\right),$$

i=1, 2,…, n; k=0, 1, 2,…

可见, J迭代法的迭代矩阵为

$$\boldsymbol{B}=\begin{pmatrix} 0 & \frac{-a_{12}}{a_{11}} & \cdots & \frac{-a_{1n}}{a_{11}} \\ \frac{-a_{21}}{a_{22}} & 0 & \cdots & \frac{-a_{2n}}{a_{22}} \\ \cdots & \cdots & & \cdots \\ \frac{-a_{n1}}{a_{nn}} & \cdots & \frac{-a_{n,n-1}}{a_{nn}} & 0 \end{pmatrix}.$$

若记 $\mathbf{g}=\begin{pmatrix}\frac{b_1}{a_{11}} & \frac{b_2}{a_{22}} & \cdots & \frac{b_n}{a_{nn}}\end{pmatrix}^{\mathrm{T}}$, 则J迭代法可写成

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{B}\boldsymbol{x}^{(k)}+\boldsymbol{g},\quad k=0,1,2,\ldots$$

若在J迭代法中充分利用新值，则可以得到如下新格式:

$$\begin{cases} x_1^{(k+1)} = \frac{-a_{12}}{a_{11}} x_2^{(k)} - \frac{a_{13}}{a_{11}} x_3^{(k)} - \cdots - \frac{a_{1n}}{a_{11}} x_n^{(k)} + \frac{b_1}{a_{11}} \\ x_2^{(k+1)} = \frac{-a_{21}}{a_{22}} x_1^{(k+1)} - \frac{a_{23}}{a_{22}} x_3^{(k)} - \cdots - \frac{a_{2n}}{a_{22}} x_n^{(k)} + \frac{b_2}{a_{22}}, \\ \cdots\cdots \\ x_n^{(k+1)} = \frac{-a_{n1}}{a_{nn}} x_1^{(k+1)} - \frac{a_{n2}}{a_{nn}} x_2^{(k+1)} - \cdots - \frac{a_{n,n-1}}{a_{nn}} x_{n-1}^{(k+1)} + \frac{b_n}{a_{nn}} \end{cases}$$

k=0, 1, 2,… (3.4)

式(3.4)称为**Gauss-Seidel迭代法(的分量形式)**，简称为**G-S迭代法**.

G-S迭代法也可记为

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)} \right),$$

i=1, 2,…, n; k=0, 1, 2,…

例1 用J法和G-S法求解线性方程组

回到例2

$$\begin{cases} 10x_1 + 3x_2 + x_3 = 14 \\ 2x_1 - 10x_2 + 3x_3 = -5 \\ x_1 + 3x_2 + 10x_3 = 14 \end{cases}$$

方程组的精确解为$\boldsymbol{x}^*=(1, 1, 1)^T$.

解: J迭代法计算公式为

$$\begin{cases} x_1^{(k+1)} = -\frac{3}{10}x_2^{(k)} - \frac{1}{10}x_3^{(k)} + \frac{7}{5} \\ x_2^{(k+1)} = \frac{1}{5}x_1^{(k)} + \frac{3}{10}x_3^{(k)} + \frac{1}{2} \\ x_3^{(k+1)} = -\frac{1}{10}x_1^{(k)} - \frac{3}{10}x_2^{(k)} + \frac{7}{5} \end{cases}$$

取初始向量$\boldsymbol{x}^{(0)}=(0, 0, 0)^T$, 迭代计算结果列表如下:

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^*\|_\infty$
0	0	0	0	1
1	1.4000	0.5000	1.4000	0.5000
2	1.1100	1.2000	1.1100	0.2000
3	0.9290	1.0550	0.9290	0.0710
4	0.9906	0.9645	0.9906	0.0355
5	1.0116	0.9953	1.0116	0.0116
6	1.0003	1.0058	1.0003	0.0058

可见，迭代序列逐次收敛于方程组的解，而且迭代6次得到精确到小数点后两位的近似解.

G-S迭代法的计算公式为:

$$\begin{cases} x_1^{(k+1)} = -\frac{3}{10}x_2^{(k)} - \frac{1}{10}x_3^{(k)} + \frac{7}{5} \\ x_2^{(k+1)} = \frac{1}{5}x_1^{(k+1)} + \frac{3}{10}x_3^{(k)} + \frac{1}{2} \\ x_3^{(k+1)} = -\frac{1}{10}x_1^{(k+1)} - \frac{3}{10}x_2^{(k+1)} + \frac{7}{5} \end{cases}$$

同样取初始向量$\boldsymbol{x}^{(0)}=(0, 0, 0)^T$, 计算结果为

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^*\\|_\infty$
0	0	0	0	1
1	1.4000	0.7800	1.0260	0.4000
2	1.0634	1.0205	0.9875	0.0634
3	0.9951	0.9953	1.0019	0.0049
4	1.0012	1.0008	0.9996	0.0012

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\lVert \boldsymbol{x}^{(k)}-\boldsymbol{x}^* \rVert_\infty$
0	0	0	0	1
1	1.4000	0.7800	1.0260	0.4000
2	1.0634	1.0205	0.9875	0.0634
3	0.9951	0.9953	1.0019	0.0049
4	1.0012	1.0008	0.9996	0.0012

可见G-S迭代法收敛较快. 取精确到小数点后两位的近似解, G-S迭代法只需迭代3次, 而J迭代法需要迭代6次.

为了进一步研究, 从矩阵计算角度来讨论上述迭代法.

对线性方程组$\boldsymbol{Ax}=\boldsymbol{b}$,记

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & \cdots \\ \cdots & \cdots & \cdots & & a_{n-1,n} \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn} \end{pmatrix},$$

以及……

$$\boldsymbol{D} = \begin{pmatrix} a_{11} & & & & \\ & a_{22} & & & \\ & & a_{33} & & \\ & & & \dots & \\ & & & & a_{nn} \end{pmatrix},$$

$$-\boldsymbol{L} = \begin{pmatrix} 0 & & & & \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & 0 & & \\ \dots & \dots & \dots & \dots & \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & 0 \end{pmatrix},$$

$$-\boldsymbol{U} = \begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ & 0 & a_{23} & \dots & a_{2n} \\ & & 0 & \dots & \dots \\ & & & & a_{n-1,n} \\ & & & & 0 \end{pmatrix},$$

则有矩阵分裂: $\boldsymbol{A}=\boldsymbol{D}+(-\boldsymbol{L})+(-\boldsymbol{U})=\boldsymbol{D}-\boldsymbol{L}-\boldsymbol{U}$.

于是线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 可写成

$$(\boldsymbol{D}-\boldsymbol{L}-\boldsymbol{U})\boldsymbol{x}=\boldsymbol{b}.$$

以下分别采用不同的策略各自建立相应的迭代格式.

格式1

对原方程组采用以下方式移项整理, 可得等价方程组:

$$\boldsymbol{Dx}=(\boldsymbol{L}+\boldsymbol{U})\boldsymbol{x}+\boldsymbol{b}.$$

当矩阵$\boldsymbol{A}$的对角元素都不为0时, 矩阵$\boldsymbol{D}$可逆, 此时有

$$\boldsymbol{x}=\boldsymbol{D}^{-1}(\boldsymbol{L}+\boldsymbol{U})\boldsymbol{x}+\boldsymbol{D}^{-1}\boldsymbol{b}.$$

由此建立迭代公式

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{D}^{-1}(\boldsymbol{L}+\boldsymbol{U})\boldsymbol{x}^{(k)}+\boldsymbol{D}^{-1}\boldsymbol{b},\ k=0,1,2,\ldots$$

或写成

$$x^{(k+1)}=Bx^{(k)}+g,\ k=0,1,2,\ldots$$

其中

$$B=D^{-1}(L+U)=\begin{pmatrix} 0 & \frac{-a_{12}}{a_{11}} & \cdots & \frac{-a_{1n}}{a_{11}} \\ \frac{-a_{21}}{a_{22}} & 0 & \cdots & \frac{-a_{2n}}{a_{22}} \\ \cdots & \cdots & & \cdots \\ \frac{-a_{n1}}{a_{nn}} & \cdots & \frac{-a_{n,n-1}}{a_{nn}} & 0 \end{pmatrix},\ g=D^{-1}b=\begin{pmatrix} \frac{b_1}{a_{11}} \\ \frac{b_2}{a_{22}} \\ \cdots \\ \frac{b_n}{a_{nn}} \end{pmatrix}$$

Jacobi迭代矩阵

经观察可知: 此格式就是 Jacobi迭代格式.

Jacobi迭代格式 (矩阵形式)

$$x^{(k+1)}=D^{-1}(L+U)x^{(k)}+D^{-1}b,\quad k=0,1,2,\ldots$$

Jacobi迭代矩阵

格式2

对原方程组

$$(\boldsymbol{D}-\boldsymbol{L}-\boldsymbol{U})\boldsymbol{x}=\boldsymbol{b}$$

采用以下方式移项整理, 可得等价形式:

$$(\boldsymbol{D}-\boldsymbol{L})\boldsymbol{x}=\boldsymbol{U}\boldsymbol{x}+\boldsymbol{b}.$$

当矩阵$\boldsymbol{A}$的对角元素都不为0时, 矩阵$\boldsymbol{D}-\boldsymbol{L}$可逆, 此时有

$$\boldsymbol{x}=(\boldsymbol{D}-\boldsymbol{L})^{-1}\boldsymbol{U}\boldsymbol{x}+(\boldsymbol{D}-\boldsymbol{L})^{-1}\boldsymbol{b}.$$

由此建立迭代公式

$$\boldsymbol{x}^{(k+1)}=(\boldsymbol{D}-\boldsymbol{L})^{-1}\boldsymbol{U}\boldsymbol{x}^{(k)}+(\boldsymbol{D}-\boldsymbol{L})^{-1}\boldsymbol{b},\ k=0,1,2,\ldots$$

思考：这是什么格式?

事实上, 对于J迭代法

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{D}^{-1}(\boldsymbol{L}+\boldsymbol{U})\boldsymbol{x}^{(k)}+\boldsymbol{D}^{-1}\boldsymbol{b},\ k=0,1,2,\ldots$$

可以写为: $\boldsymbol{x}^{(k+1)}=\boldsymbol{D}^{-1}\boldsymbol{L}\boldsymbol{x}^{(k)}+\boldsymbol{D}^{-1}\boldsymbol{U}\boldsymbol{x}^{(k)}+\boldsymbol{D}^{-1}\boldsymbol{b}$.

按照G-S迭代改善的思路, 构造相应的新格式

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{D}^{-1}\boldsymbol{L}\boldsymbol{x}^{(k+1)}+\boldsymbol{D}^{-1}\boldsymbol{U}\boldsymbol{x}^{(k)}+\boldsymbol{D}^{-1}\boldsymbol{b},$$

即:

$$\boldsymbol{D}\boldsymbol{x}^{(k+1)}=\boldsymbol{L}\boldsymbol{x}^{(k+1)}+\boldsymbol{U}\boldsymbol{x}^{(k)}+\boldsymbol{b},$$

也即:

$$(\boldsymbol{D}-\boldsymbol{L})\boldsymbol{x}^{(k+1)}=\boldsymbol{U}\boldsymbol{x}^{(k)}+\boldsymbol{b},$$

也即:

$$\boldsymbol{x}^{(k+1)}=(\boldsymbol{D}-\boldsymbol{L})^{-1}\boldsymbol{U}\boldsymbol{x}^{(k)}+(\boldsymbol{D}-\boldsymbol{L})^{-1}\boldsymbol{b},$$

就是刚才的格式2

G-S迭代格式 (矩阵形式)

$$\boldsymbol{x}^{(k+1)}=(\boldsymbol{D}-\boldsymbol{L})^{-1}\boldsymbol{U}\boldsymbol{x}^{(k)}+(\boldsymbol{D}-\boldsymbol{L})^{-1}\boldsymbol{b},\ k=0,1,2,\ldots$$

其中, G-S格式的迭代矩阵为: $\boldsymbol{G}=(\boldsymbol{D}-\boldsymbol{L})^{-1}\boldsymbol{U}$.

§ 3.2 迭代法的收敛性

讨论迭代法

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{M}\boldsymbol{x}^{(k)}+\boldsymbol{g},\quad k=0, 1, 2,\ldots$$

的收敛性.

记误差向量$\boldsymbol{e}^{(k)}=\boldsymbol{x}^{(k)}-\boldsymbol{x}^{*}$, 则迭代法收敛就是

$$\boldsymbol{e}^{(k)}\to\boldsymbol{0},\ (k\to\infty).$$

由于

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{M}\boldsymbol{x}^{(k)}+\boldsymbol{g},\quad k=0, 1, 2,\ldots$$

$$\boldsymbol{x}^{*}=\boldsymbol{M}\boldsymbol{x}^{*}+\boldsymbol{g},$$

所以

$$\boldsymbol{e}^{(k+1)}=\boldsymbol{M}\boldsymbol{e}^{(k)},\ k=0, 1, 2,\ldots$$

递推可得

$$\boldsymbol{e}^{(k)}=\boldsymbol{M}^{k}\boldsymbol{e}^{(0)},\ k=0, 1, 2,\ldots$$

由以上讨论可见:

对任意的初始误差向量$\boldsymbol{e}^{(0)}$, 迭代解$\boldsymbol{e}^{(k)}\to\boldsymbol{0}$ $(k\to\infty)$ $\Leftrightarrow$ $\boldsymbol{M}^k\to\boldsymbol{O}$.

思考: 上述迭代收敛的等价条件实际应用中是否方便?

定理3.1 迭代法$\boldsymbol{x}^{(k+1)}=\boldsymbol{M}\boldsymbol{x}^{(k)}+\boldsymbol{g}$对任意初始向量$\boldsymbol{x}^{(0)}$都收敛的等价条件是: $\rho(\boldsymbol{M})<1$.

证明:1. 必要性. 若迭代法收敛, 则$\|\boldsymbol{M}^k\|\to 0$, 于是有

$$\rho^k(\boldsymbol{M}) = \rho(\boldsymbol{M}^k) \le \|\boldsymbol{M}^k\|\to 0,$$

所以$\rho(\boldsymbol{M})<1$.

2. 充分性. 若$\rho(\boldsymbol{M})<1$, 则存在$\varepsilon>0$, 使得$\rho(\boldsymbol{M})+\varepsilon<1$.

于是存在某矩阵范数$\|\cdot\|_\beta$, 使得

$$\|\boldsymbol{M}^k\|_\beta \le \|\boldsymbol{M}\|_\beta^{\,k} \le [\rho(\boldsymbol{M})+\varepsilon]^k \to 0.$$

可知此时迭代法收敛.

综上, 证毕.

注1：定理3.1是关于迭代法收敛性的一个基本定理.

注2：定理3.1的结论是面向“任意初始向量都收敛”的情形的.

对于**某些特殊的初始向量**，即使迭代格式不满足“$\rho(\boldsymbol{M})<1$”的条件，也有可能收敛.

注3：谱半径$\rho(\mathbf{M})$一般不容易计算.

但是，根据定理3.1还可以得到以下充分性判据.

推论： 对于迭代法$\boldsymbol{x}^{(k+1)}=\boldsymbol{M}\boldsymbol{x}^{(k)}+\boldsymbol{g}$, 若迭代矩阵$\boldsymbol{M}$的某个矩阵范数$\|\boldsymbol{M}\|_\beta<1$, 则此迭代法对任意初始向量$\boldsymbol{x}^{(0)}$都收敛.

例1

$$\boldsymbol{M}=\begin{bmatrix} 0.8 & 0.05 \\ 0.1 & 0.7 \end{bmatrix}$$

由于$\|\boldsymbol{M}\|_1=0.9<1$,

于是可知$\rho(\boldsymbol{M})\le\|\boldsymbol{M}\|_1<1$

事实上$\rho(\boldsymbol{M})=0.8366$

因此, 此时迭代法收敛.

注: 此推论只是充分性条件, 而非必要性条件. 见下例.

例2

$$\boldsymbol{M}=\begin{bmatrix} 0.8 & 0 \\ 0.5 & 0.7 \end{bmatrix}$$

可得

$$\|\boldsymbol{M}\|_1=1.3\geq 1;$$

$$\|\boldsymbol{M}\|_\infty=1.2\geq 1;$$

$$\|\boldsymbol{M}\|_2=1.0455\geq 1;$$

$$\|\boldsymbol{M}\|_F=1.1747\geq 1,$$

但是

$$\rho(\boldsymbol{M})=0.8<1.$$

因此, 此时迭代法还是收敛的.

结论: 即使迭代矩阵$\boldsymbol{M}$有“很多”个范数值都不小于1, 迭代也未必不收敛.

定理3.2: 设矩阵$\boldsymbol{M}$的某个矩阵范数$\|\boldsymbol{M}\|<1$，则迭代法$\boldsymbol{x}^{(k+1)}=\boldsymbol{M}\boldsymbol{x}^{(k)}+\boldsymbol{g}$ 对任意的初始向量$\boldsymbol{x}^{(0)}$收敛，且满足以下条件:

$$\left\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^{*}\right\| \leq \frac{\|\boldsymbol{M}\|}{1-\|\boldsymbol{M}\|}\left\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^{(k-1)}\right\| \tag{3.16}$$

$$\left\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^{*}\right\| \leq \frac{\|\boldsymbol{M}\|^{k}}{1-\|\boldsymbol{M}\|}\left\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}\right\| \tag{3.17}$$

上述两个不等式中的**矩阵范数与向量范数相容.**

思考: 这两个不等式的意义.

注: 此定理的结论只有在满足"$\|\boldsymbol{M}\|<1$"的条件下才适用.

证明： 由于 $$x^{(k+1)}=Mx^{(k)}+g,$$

$$x^{(k)}=Mx^{(k-1)}+g,$$

$$x^{*}=Mx^{*}+g,$$

所以 $x^{(k+1)}-x^{(k)}=M(x^{(k)}-x^{(k-1)})$, $x^{(k+1)}-x^{*}=M(x^{(k)}-x^{*})$.

于是有 $\|x^{(k+1)}-x^{(k)}\|\le\|M\|\|x^{(k)}-x^{(k-1)}\|$, $\|x^{(k+1)}-x^{*}\|\le\|M\|\|x^{(k)}-x^{*}\|$.

从而， $\|x^{(k)}-x^{*}\|=\|(x^{(k)}-x^{(k+1)})+(x^{(k+1)}-x^{*})\|$

$$\le\|x^{(k)}-x^{(k+1)}\|+\|x^{(k+1)}-x^{*}\|\le\|M\|\|x^{(k)}-x^{(k-1)}\|+\|M\|\|x^{(k)}-x^{*}\|.$$

所以可得

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

第一个不等式(3.16)

又由于 $\|x^{(k)}-x^{(k-1)}\|\le\|M\|\|x^{(k-1)}-x^{(k-2)}\|\le\ldots\ldots\le\|M\|^{k-1}\|x^{(1)}-x^{(0)}\|$,

代入第一个不等式(3.16)即得

$$\|x^{(k)}-x^{*}\|\le\frac{\|M\|^{k}}{1-\|M\|}\|x^{(1)}-x^{(0)}\|.$$

第二个不等式 (3.17)

证毕.

注1: 此定理的前提条件是$\|\boldsymbol{M}\|<1$. 否则此定理结论未必成立.

注2: $\|\boldsymbol{M}\|<1$是指$\boldsymbol{M}$的某个范数(不必是所有范数)小于1.

注3: 此定理不等式中的向量范数与矩阵范数必须相容.

可见，$\|M\|$越小收敛越快，且当$\|M\|<1$，$\|x^{(k)}-x^{(k-1)}\|$很小时，$\|x^{(k)}-x^*\|$就很小，实际上用$\|x^{(k)}-x^{(k-1)}\|<\varepsilon$作为迭代终止的条件. 以下算例在实际计算中（在确定$\|M\|<1$的前提下）可以通过设置指标$\|x^{(k)}-x^{(k-1)}\|$来观察迭代是否收敛.

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\|x^{(k)}-x^*\|_\infty$
0	0	0	0	1
1	1.4	0.5	1.4	0.5
2	1.11	1.20	1.11	0.2
3	0.929	1.055	0.929	0.071
4	0.9906	0.9645	0.9906	0.0355
5	1.01159	0.9953	1.01159	0.01159
6	1.000251	1.005795	1.000251	0.005795
7	0.9982364	1.0001255	0.9982364	0.0017636

由定理3.2还可以事先估计需要迭代多少步可达到所需的精度要求.

若使$\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^*\|<\varepsilon$，只需

$$\frac{\|\boldsymbol{M}\|^k}{1-\|\boldsymbol{M}\|}\left\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}\right\|<\varepsilon,$$

即

$$k>\frac{\ln\left(\frac{\varepsilon(1-\|\boldsymbol{M}\|)}{\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}}\right)}{\ln\|\boldsymbol{M}\|},$$

最后取满足这个不等式的最小的正整数(向上取整)即可.

例2 用J迭代法求例1中方程组的解, 取$\boldsymbol{x}^{(0)}=(0, 0, 0)^T$, 若使误差$\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^*\|_\infty<10^{-5}$, 问迭代多少次可满足此要求?

解 由例1(回看例1)知, $\boldsymbol{x}^{(1)}=(1.4, 0.5, 1.4)^T$,

$$\boldsymbol{B}=\begin{pmatrix} 0 & \frac{-3}{10} & \frac{-1}{10} \\ \frac{1}{5} & 0 & \frac{3}{10} \\ \frac{-1}{10} & \frac{-3}{10} & 0 \end{pmatrix}.$$

于是有 $\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}\|_\infty=1.4$, $\|\boldsymbol{B}\|_\infty=0.5$.

k应满足 $$k>\ln(\frac{0.5\times10^{-5}}{1.4})/\ln 0.5\approx 18.095$$

故取k=19, 即迭代19次可满足此要求.

注: 按定理3.2中的不等式估算出的k值一般比实际精确的k值要稍大一些, 这是合理的. 想一想为什么.

补充知识点:

对同一个迭代矩阵M，取不同的矩阵范数时，$\|M\|$值也不相同. 但根据范数的等价性可知: 当迭代收敛时这些范数值都是比较小的数值，并且此时谱半径$\rho(M)$也一定是比较小的. 事实上可以定义渐进收敛速度$R(M)$如下:

$$R(\mathbf{M})=\lim_{k\to\infty}\frac{-\ln\left\|\mathbf{M}^k\right\|}{k}=-\ln\rho(\mathbf{M})$$

可见迭代速度的快慢本质上取决于谱半径的大小. $\rho(M)$越小, 收敛速度越快.

对于任意矩阵范数的极限都是如此

§ 3.3 J迭代法和G-S迭代法的收敛性

定理3.3 对任意的迭代初始向量,

J迭代法收敛$\Leftrightarrow\rho(\boldsymbol{B})<1$;

$\boldsymbol{B}$的某个(给定的)范数$\|\boldsymbol{B}\|_\beta<1\Rightarrow$J迭代法收敛;

G-S迭代法收敛$\Leftrightarrow\rho(\boldsymbol{G})<1$;

$\boldsymbol{G}$的某个(给定的)范数$\|\boldsymbol{G}\|_\beta< 1\Rightarrow$ G-S迭代法收敛;

以上判据都是通过观察迭代矩阵的性质来判断迭代是否收敛.

问题: 在一定条件下, 是否有更“简单”的方法(譬如, 直接观察系数矩阵的性质)来判断迭代法的收敛性?

定义3.1 若n阶方阵$A=(a_{ij})$满足:

$$\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|<|a_{ii}|,\quad i=1, 2,\ldots, n.$$

则称矩阵A是(按行)**严格对角占优矩阵.**

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n}\\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n}\\ a_{31} & a_{32} & a_{33} & \cdots & \cdots\\ \cdots & \cdots & \cdots & & a_{n-1,n}\\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn}\end{pmatrix}.$$

引理 若A是严格对角占优矩阵, 则$\det(A)\neq 0$.

证明: 易见

$$A=D-L-U= D[E-D^{-1}(L+U)]= D(E-B).$$

因为A是严格对角占优矩阵, 所以$\det(D)\neq 0$, 而且

$$B=\begin{bmatrix} 0 & \frac{-a_{12}}{a_{11}} & \cdots & \frac{-a_{1n}}{a_{11}} \\ \frac{-a_{21}}{a_{22}} & 0 & \cdots & \frac{-a_{2n}}{a_{22}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{-a_{n1}}{a_{n,n}} & \cdots & \frac{-a_{n,n-1}}{a_{n,n}} & 0 \end{bmatrix}. \quad \|B\|_\infty = \max_{1\le i\le n}\sum_{\substack{j=1\\ j\neq i}}^{n}\left|\frac{a_{ij}}{a_{ii}}\right| < 1.$$

因此, $\rho(B)\le\|B\|_\infty<1$, 故$\lambda=1$不是B的特征值, $\det(E-B)\neq 0$.

所以, $\det(A)\neq 0$. 证毕.

注: **上述过程顺便也证明了: 此时方程组$Ax=b$的J迭代法是收敛的.**

定理3.4 设A是严格对角占优矩阵, 则解线性方程组$Ax=b$的J迭代法和G-S迭代法(对任意的迭代初始向量)均收敛.

证明: 由于$\|B\|_\infty<1$, 所以J迭代法收敛.

以下证明G-S迭代法收敛.设λ是G的任一特征值, 则λ满足特征方程

$$\det(\lambda E-G)=0.$$

即:

$$\det(\lambda E-(D-L)^{-1}U)=0.$$

也即:

$$\det[(D-L)^{-1}]\det[\lambda(D-L)-U]=0.$$

它等价的方程为:

$$\det[\lambda(D-L)-U]=0. \qquad (*)$$

采用反证法. 若$|\lambda|\geq 1$, 由于$A=D-L-U$严格对角占优, 而

$$\lambda(D-L)-U=\begin{pmatrix} \lambda a_{11} & a_{12} & \cdots & a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \cdots \\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nn} \end{pmatrix}.$$

$$\boldsymbol{\lambda}(\boldsymbol{D}-\boldsymbol{L})-\boldsymbol{U}=\begin{pmatrix} \lambda a_{11} & a_{12} & \dots & a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & a_{2n} \\ \dots & \dots & & \dots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nn} \end{pmatrix}.$$

由于

$$|\lambda\|a_{ii}|>|\lambda|\sum_{\substack{j=1\\ j\neq i}}^{n}|a_{ij}|=\sum_{j=1}^{i-1}|\lambda||a_{ij}|+\sum_{j=i+1}^{n}|\lambda||a_{ij}|$$

$$\geq\sum_{j=1}^{i-1}|\lambda||a_{ij}|+\sum_{j=i+1}^{n}|a_{ij}|.$$

可见$|\lambda|\geq 1$时矩阵$\lambda(\boldsymbol{D}$-$\boldsymbol{L})$-$\boldsymbol{U}$也是严格对角占优的. 这与(*)式结论 det($\lambda(\boldsymbol{D}$-$\boldsymbol{L})$-$\boldsymbol{U}$)=0 矛盾. 所以不可能有$|\lambda|\geq 1$, 只能$|\lambda|<1$.

由于λ是$\boldsymbol{G}$的任意特征值, 于是可得$\rho(\boldsymbol{G})<1$. 因此G-S迭代法收敛. 证毕.

定理3.5 设A是对称正定矩阵，则解方程组$Ax=b$的**G-S**迭代法(对任意的迭代初始向量)是收敛的.

此定理是定理3.9的推论.

注：当A是对称正定矩阵时，方程组$Ax=b$的J迭代法未必(对任意的迭代初始向量)收敛.

反例参见教材P56.

但是...

补充定理(P68：3-14) 设A是对称正定矩阵，则解方程组$Ax=b$的**J**迭代法(对任意的迭代初始向量)收敛$\Leftrightarrow$$2D-A$也正定.

例3 请针对线性方程组

$$\begin{cases} 2x_1 + 3x_2 + 6x_3 = 1 \\ 4x_1 - x_2 + 2x_3 = -2 \\ x_1 + 5x_2 + 2x_3 = 4 \end{cases}$$

建立(对任意的迭代初始向量)收敛的迭代格式, 说明理由.

解 先将原方程组(预处理)化为新的等价方程组:

$$\begin{cases} 4x_1 - x_2 + 2x_3 = -2, \\ x_1 + 5x_2 + 2x_3 = 4, \\ 2x_1 + 3x_2 + 6x_3 = 1. \end{cases}$$

建立迭代格式

$$\begin{cases} x_1^{(k+1)} = \frac{1}{4}x_2^{(k)} - \frac{1}{2}x_3^{(k)} - \frac{1}{2}, \\ x_2^{(k+1)} = \frac{-1}{5}x_1^{(k)} - \frac{2}{5}x_3^{(k)} + \frac{4}{5}, \\ x_3^{(k+1)} = \frac{-1}{3}x_1^{(k)} - \frac{1}{2}x_2^{(k)} + \frac{1}{6}. \end{cases}$$

由于新的方程组的系数矩阵是严格对角占优的. 而迭代格式是对这个方程组应用了Jacobi迭代法，由定理3.4可知迭代格式此时一定收敛.

注1: 也可以根据此迭代格式迭代矩阵的行范数小于1, 来判断迭代法收敛.

注2: 此迭代格式不能称为原方程组的Jacobi迭代格式, 因为它是对原方程进行预处理后才使用了Jacobi迭代法.

§ 3.4 逐次超松弛迭代法---SOR方法

将Jacobi迭代法

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}\right),$$

i=1, 2,…, n; k=0, 1, 2,…

改写成

$$x_i^{(k+1)} = x_i^{(k)} + \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{n} a_{ij}x_j^{(k)}\right),$$

i=1, 2,…, n; k=0, 1, 2,…

相应的矩阵方程形式就是:

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}+\boldsymbol{D}^{-1}(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}^{(k)}),\quad k=0,1,2,\ldots$$

Gauss-Seidel迭代法也可写成

$$x_i^{(k+1)} = x_i^{(k)} + \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij}x_j^{(k)}\right),$$

i=1, 2,…, n; k=0, 1, 2,…

或写成矩阵方程形式:

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}+\boldsymbol{D}^{-1}[\boldsymbol{b}+\boldsymbol{L}\boldsymbol{x}^{(k+1)}+(\boldsymbol{U}-\boldsymbol{D})\boldsymbol{x}^{(k)}],\ k=0,1,2,\ldots$$

还可在此基础上做进一步的改进:

$$x_i^{(k+1)} = x_i^{(k)} + \frac{\omega}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij}x_j^{(k)}\right),$$

i=1, 2,…, n; k=0, 1, 2,…

思考: **新格式的意义和优越性是 …?**

迭代格式(分量形式)

$$x_i^{(k+1)} = x_i^{(k)} + \frac{\omega}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij}x_j^{(k)}\right),$$

i=1, 2,…, n; k=0, 1, 2,…

称为**SOR方法(Successive Over Relaxation)(逐次超松弛迭代法)**. 其中参数ω称为**松弛因子**.

当ω>1时称为**超松弛迭代**,

当ω<1时称为**欠松弛迭代**.

注1: ω=1时SOR方法即为G-S方法.

注2: SOR方法可以看作对J迭代法或G-S迭代法的改进.

注3: SOR法也可通过某种矩阵分裂的方式而得到. (**思考题**)

SOR方法的矩阵方程形式可写为

$$x^{(k+1)}=x^{(k)}+\omega D^{-1}(\mathrm{b}+Lx^{(k+1)}+(U-D)x^{(k)}),\quad k=0,1,2,\ldots$$

于是有

$$Dx^{(k+1)}=Dx^{(k)}+\omega(b+Lx^{(k+1)}+(U-D)x^{(k)})$$

SOR方法的矩阵方程形式为:

$$x^{(k+1)}=(D-\omega L)^{-1}\left[(1-\omega)D+\omega U\right]x^{(k)}+\omega(D-\omega L)^{-1}b,\quad k=0,1,2,\ldots$$

SOR方法的迭代矩阵为

$$\pounds_{\omega}=(D-\omega L)^{-1}\left[(1-\omega)D+\omega U\right].$$

定理3.6 **对任意的迭代初始向量,**

SOR方法收敛的**充分必要条件**是: $\rho(\pounds_{\omega})<1$;

SOR方法收敛的**充分条件**是: $\pounds_{\omega}$的某个给定的矩阵范数$\|\pounds_{\omega}\|_{\beta}<1$.

定理3.7 若SOR方法(对任意迭代初始向量)收敛, 则$0<\omega<2$.

证明: 设SOR方法收敛, 则$\rho(\mathcal{L}_\omega)<1$, 所以

$$|\det(\mathcal{L}_\omega)| = |\lambda_1\lambda_2\dots\lambda_n|<1.$$

而

$$\det(\mathcal{L}_\omega) = \det[(D-\omega L)^{-1}((1-\omega)D+\omega U)]$$

$$=\det\{[D(E-\omega D^{-1}L)]^{-1} D[(1-\omega)E+\omega D^{-1}U]\}$$

$$=\det\{(E-\omega D^{-1}L)^{-1} D^{-1} D[(1-\omega)E+\omega D^{-1}U]\}$$

$$=\det[(E-\omega D^{-1}L)^{-1}]\det[(1-\omega)E+\omega D^{-1}U)]$$

$$=(1-\omega)^n.$$

于是可得 $|1-\omega|<1$, 也即

$$0<\omega<2.$$ 证毕.

请思考此定理的意义.

定理3.8 设A是严格对角占优矩阵，则解方程组$Ax=b$的SOR方法当$0<\omega\le1$时(对任意迭代初始向量)收敛.

定理3.8的证明留作思考题. 教材P68: 3-13.

定理3.9 设A是对称正定矩阵，则解方程组$Ax=b$的SOR方法当$0<\omega<2$时(对任意迭代初始向量)收敛.

关于定理3.9的一些预备知识

命题1 若A为对称正定矩阵, 则A的全体对角元素形成的对角矩阵D必为正定矩阵.

定义1 对于C^n 中的向量$\mathbf{x}=(x_1, x_2, \ldots, x_n)^T$, $\mathbf{y}=(y_1, y_2, \ldots, y_n)^T$, 定义它们的**内积**为如下**复数**:

$$\left(\mathbf{x},\mathbf{y}\right)=x_1\overline{y}_1+x_2\overline{y}_2+\ldots+x_n\overline{y}_n=\mathbf{y}^{\mathrm{H}}\mathbf{x}=\overline{\mathbf{y}}^{\mathrm{T}}\mathbf{x}$$

复内积的基本性质: 对于任意的$\mathbf{x}, \mathbf{y}, \mathbf{z}\in C^n$,有

(1) $\left(\mathbf{y},\mathbf{x}\right)=\overline{\left(\mathbf{x},\mathbf{y}\right)}$;

(2) 对于任意复数a, b, 有 $\left(a\mathbf{x}+b\mathbf{y},\mathbf{z}\right)=a\left(\mathbf{x},\mathbf{z}\right)+b\left(\mathbf{y},\mathbf{z}\right)$;

(3) $(\mathbf{x}, \mathbf{x})\geq 0$; 且$(\mathbf{x}, \mathbf{x})=0$ 当且仅当 $\mathbf{x}=\mathbf{0}$ (零向量).

(4) 对于n阶方阵$\mathbf{A}$, 成立$(\mathbf{A}\mathbf{x}, \mathbf{y})= (\mathbf{x}, \mathbf{A}^{\mathrm{H}}\mathbf{y})$.

定理3.9 设$\boldsymbol{A}$是对称正定矩阵, 则解方程组$\boldsymbol{Ax=b}$的**SOR**方法当$0<\omega<2$时(对任意迭代初始向量)收敛.

证明: 设λ是$\boldsymbol{\pounds}_{\omega}$的任一特征值, $\boldsymbol{y}$是对应的特征向量, 则

$$(\boldsymbol{D}-\omega\boldsymbol{L})^{-1}[(1-\omega)\boldsymbol{D}+\omega\boldsymbol{U}]\boldsymbol{y}=\lambda\boldsymbol{y},$$

也即:

$$[(1-\omega)\boldsymbol{D}+\omega\boldsymbol{U}]\boldsymbol{y}=\lambda(\boldsymbol{D}-\omega\boldsymbol{L})\boldsymbol{y}.$$

于是两边分别与$\boldsymbol{y}$做内积, 可得

$$([(1-\omega)\boldsymbol{D}+\omega\boldsymbol{U}]\boldsymbol{y},\boldsymbol{y})=(\lambda(\boldsymbol{D}-\omega\boldsymbol{L})\boldsymbol{y},\boldsymbol{y}).$$

也即

$$(1-\omega)(\boldsymbol{Dy},\boldsymbol{y})+\omega(\boldsymbol{Uy},\boldsymbol{y})=\lambda[(\boldsymbol{Dy},\boldsymbol{y})-\omega(\boldsymbol{Ly},\boldsymbol{y})].$$

故得

$$\lambda=\frac{(1-\omega)(\boldsymbol{Dy},\ \boldsymbol{y})+\omega(\boldsymbol{Uy},\ \boldsymbol{y})}{(\boldsymbol{Dy},\ \boldsymbol{y})-\omega(\boldsymbol{Ly},\ \boldsymbol{y})}.$$

由于$\boldsymbol{A=D-L-U}$是对称正定的, 所以$\boldsymbol{D}$是正定矩阵, 且$\boldsymbol{L}=\boldsymbol{U}^{\mathrm{T}}$.

由$\boldsymbol{D}$的正定性可得

$$(\boldsymbol{D}\boldsymbol{y}, \boldsymbol{y})=\sigma>0.$$

由$\boldsymbol{U}=\boldsymbol{L}^{\mathrm{T}}$, 以及复内积的定义, 注意此时$\boldsymbol{L}^{\mathrm{T}}=\boldsymbol{L}^{\mathrm{H}}$, 可得

$$(\boldsymbol{U}\boldsymbol{y},\ \boldsymbol{y}) = (\boldsymbol{L}^{\mathrm{T}}\boldsymbol{y},\ \boldsymbol{y}) = (\boldsymbol{y},\ \boldsymbol{L}\boldsymbol{y}) = \overline{(\boldsymbol{L}\boldsymbol{y},\ \boldsymbol{y})}.$$

所以, 若记$(\boldsymbol{L}\boldsymbol{y}, \boldsymbol{y})=\alpha+\mathrm{i}\beta$, 则有 $(\boldsymbol{U}\boldsymbol{y}, \boldsymbol{y})=\alpha-\mathrm{i}\beta$.

根据$\boldsymbol{A}$的正定性, 以及以上的分析, 可得

$$0<(\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y})=(\boldsymbol{D}\boldsymbol{y}, \boldsymbol{y})-(\boldsymbol{L}\boldsymbol{y}, \boldsymbol{y})-(\boldsymbol{U}\boldsymbol{y}, \boldsymbol{y})=\sigma-2\alpha.$$

再由

$$\lambda = \frac{(1-\omega)(\boldsymbol{D}\boldsymbol{y},\ \boldsymbol{y}) + \omega(\boldsymbol{U}\boldsymbol{y},\ \boldsymbol{y})}{(\boldsymbol{D}\boldsymbol{y},\ \boldsymbol{y}) - \omega(\boldsymbol{L}\boldsymbol{y},\ \boldsymbol{y})}$$

可得

$$\lambda = \frac{(1-\omega)\sigma + \omega(\alpha - i\beta)}{\sigma - \omega(\alpha + i\beta)}$$

于是

$$|\lambda|^2=\frac{(\sigma-\sigma\omega+\omega\alpha)^2+\omega^2\beta^2}{(\sigma-\omega\alpha)^2+\omega^2\beta^2}$$

当$0<\omega<2$时, 观察上述分式, 可得

$$\begin{aligned}分子-分母&=(\sigma-\sigma\omega+\omega\alpha)^2-(\sigma-\omega\alpha)^2\\&=(2\omega\alpha-\sigma\omega)(2\sigma-\sigma\omega)\\&=\omega\sigma(2\alpha-\sigma)(2-\omega)\\&<0\end{aligned}$$

$0<(\boldsymbol{Ay},\boldsymbol{y})=\sigma-2\alpha.$

所以$|\lambda|^2<1$, 因此$\rho(\boldsymbol{\pounds}_\omega)<1$, 即**SOR**方法收敛. 证毕.

推论(定理3.5): 当$\boldsymbol{A}$是对称正定矩阵时, 解方程组$\boldsymbol{Ax}=\boldsymbol{b}$的Gauss-Seidel方法(对任意迭代初始向量)必定收敛.

补充1: 借鉴定理3.9的证明思路, 还可以推得如下结论.

设λ是Jacobi迭代矩阵$\boldsymbol{B}$的任一特征值, $\boldsymbol{y}$是特征向量, 则

$$(\boldsymbol{L}+\boldsymbol{U})\boldsymbol{y}=\lambda\boldsymbol{D}\boldsymbol{y}.$$

于是 $(\boldsymbol{Ly}, \boldsymbol{y})+(\boldsymbol{Uy}, \boldsymbol{y})=\lambda(\boldsymbol{Dy}, \boldsymbol{y})$.

由定理3.9的证明可得: $\lambda=2\alpha/\sigma$.

此外还可知, 当$\boldsymbol{A}$对称正定时, 必有$\sigma>0$, $2\alpha-\sigma<0$; 此时,

$$|\lambda|<1 \Leftrightarrow |2\alpha/\sigma|<1 \Leftrightarrow 2\alpha+\sigma>0,$$

而 $((\boldsymbol{2D-A})\boldsymbol{y}, \boldsymbol{y})=(\boldsymbol{Dy}, \boldsymbol{y})+(\boldsymbol{Ly}, \boldsymbol{y})+(\boldsymbol{Uy}, \boldsymbol{y})=\sigma+2\alpha,$

因此可得:

当$\boldsymbol{A}$与$\boldsymbol{2D-A}$均对称正定时, **Jacobi**迭代法必收敛.

此外还可证明(留作思考题):

当$\boldsymbol{A}$对称正定且**Jacobi**迭代法收敛时必有$\boldsymbol{2D-A}$正定.

补充2: 根据定理3.9的证明思路, 还可以证明(只需对证明过程稍作修改):

设 A 是对称负定矩阵. 则解方程组 $\boldsymbol{Ax=b}$ 的 **SOR** 方法当 $0<\omega<2$ 时(对任意迭代初始向量)也收敛.

总结

解方程组$Ax=b$时，各算法(对任意迭代初始向量)**的收敛性**

	J迭代法	G-S迭代法	SOR方法
当A严格对角占优时	收敛	收敛	当$0<\omega\leq 1$时收敛
当A对称正定(或负定)时	未必收敛	收敛	当$0<\omega<2$时收敛

SOR方法收敛的快慢与松弛因子ω的选择有密切关系. 但是如何选取最佳松弛因子, 即选取$\omega=\omega^*$, 使$\rho(\mathcal{L}_{\omega^*})$达到最小, 是一个尚未很好解决的问题. 实际应用中有时可采用试算的方法来确定较好的松弛因子.

注1: 对于具有对称正定系数矩阵的线性方程组, 目前比较成熟的高效算法有CG(Conjugate Gradient)算法, 即共轭梯度法, 或者辅以预条件技术的预条件共轭梯度(Preconditioned Conjugate Gradient)算法.

注2: 此外, 对于满足一定条件的线性方程组, 近代也涌现出诸如极小残量法(MINRES)、广义极小残量法(GMRES)等一些高效的算法. 有关新算法的理论研究和实践正在发展中.

例4 用**SOR**方法解线性方程组

$$\begin{cases} 4x_1 - 2x_2 - 4x_3 = 10 \\ -2x_1 + 17x_2 + 10x_3 = 3 \\ -4x_1 + 10x_2 + 9x_3 = -7 \end{cases}$$

方程组的精确解是$\mathbf{x}^*=(2, 1, -1)^T$.

解: SOR方法迭代公式为

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} + \dfrac{\omega}{4}(10 - 4x_1^{(k)} + 2x_2^{(k)} + 4x_3^{(k)}) \\ x_2^{(k+1)} = x_2^{(k)} + \dfrac{\omega}{17}(3 + 2x_1^{(k+1)} - 17x_2^{(k)} - 10x_3^{(k)}) \\ x_3^{(k+1)} = x_3^{(k)} + \dfrac{\omega}{9}(-7 + 4x_1^{(k+1)} - 10x_2^{(k+1)} - 9x_3^{(k)}) \end{cases}$$

取$\boldsymbol{x}^{(0)}=(0, 0, 0)^T$, $\omega=1.46$, 计算结果如下:

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	0	0	0
1	3.65	0.8845882	-0.2021098
2	2.32166910	0.4230939	-0.22243214
3	2.5661399	0.6948261	-0.4952594
…	……	……	……
20	1.9999987	1.0000013	-1.0000034

从结果可见，迭代20次时已获得精确到小数点后五位的近似解. 如果取ω=1.25, 则需要迭代56次才能得到具有同样精度的近似解; 如果取ω=1, 则需迭代110次以上.

课堂练习

设线性方程组

$$\begin{cases} 4x_1 - x_2 + 2x_3 = 1 \\ -x_1 - 5x_2 + x_3 = 2 \\ 2x_1 + x_2 + 6x_3 = 3 \end{cases}$$

(1) 写出Jacobi法和SOR法的迭代格式(分量形式);

(2) 讨论Jacobi迭代法的收敛性;

(3) 取初值$\boldsymbol{x}^{(0)}=(0, 0, 0)^T$, 若用Jacobi迭代法计算时, 预估误差$\|\boldsymbol{x}^*-\boldsymbol{x}^{(10)}\|_\infty$ (取三位有效数字).

解: (1) Jacobi迭代法格式为

$$\begin{cases} x_1^{(k+1)} = \frac{1}{4}x_2^{(k)} - \frac{1}{2}x_3^{(k)} + \frac{1}{4} \\ x_2^{(k+1)} = -\frac{1}{5}x_1^{(k)} + \frac{1}{5}x_3^{(k)} - \frac{2}{5} \\ x_3^{(k+1)} = -\frac{1}{3}x_1^{(k)} - \frac{1}{6}x_2^{(k)} + \frac{1}{2} \end{cases}$$

SOR法的迭代格式为

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} + \omega(-x_1^{(k)} + \frac{1}{4}x_2^{(k)} - \frac{1}{2}x_3^{(k)} + \frac{1}{4}) \\ x_2^{(k+1)} = x_2^{(k)} - \omega(\frac{1}{5}x_1^{(k+1)} + x_2^{(k)} - \frac{1}{5}x_3^{(k)} + \frac{2}{5}) \\ x_3^{(k+1)} = x_3^{(k)} - \omega(\frac{1}{3}x_1^{(k+1)} + \frac{1}{6}x_2^{(k+1)} + x_3^{(k)} - \frac{1}{2}) \end{cases}$$

(2) 因为系数矩阵$\boldsymbol{A}$是严格对角占优矩阵, 故Jacobi法收敛.

(3) 由(1)可见$\|\boldsymbol{B}\|_\infty=3/4$, 且取$\boldsymbol{x}^{(0)}=(0, 0, 0)^T$, 经计算可得$\boldsymbol{x}^{(1)}=(1/4, -2/5, 1/2)^T$, 于是$\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}\|_\infty=1/2$, 所以有

$$\left\|x^{*}-x^{(10)}\right\|_\infty \le \frac{\|B\|_\infty^{k}}{1-\|B\|_\infty}\left\|x^{(1)}-x^{(0)}\right\|_\infty=\frac{0.75^{10}}{1-0.75}\times 0.5=0.113$$