Lecture 1: $Ax = b$， Iterative solutions

Iterative solution methods for Ax=b

- The idea is to start with an initial guess for the solution and develop an iterative procedure that will converge to the solution.

- The simplest example of this method is known as a Jacobi (雅可比) iteration scheme.

We consider the linear system

$$
\begin{aligned}
4x - y + z &= 7 \\
4x - 8y + z &= -21 \\
-2x + y + 5z &= 15
\end{aligned}
$$

$$
A = \begin{bmatrix} 4 & -1 & 1 \\ 4 & -8 & 1 \\ -2 & 1 & 5 \end{bmatrix} \quad x = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad b = \begin{bmatrix} 7 \\ -21 \\ 15 \end{bmatrix}
$$

We can rewrite each equation as follows

$$
x = \frac{7 + y - z}{4}
$$

$$
y = \frac{21 + 4x + z}{8}
$$

$$
z = \frac{15 + 2x - y}{5}
$$

大家用数学知识
就可以手动求解
三个未知数
三个方程

但还有没有其他方法？？？
如果这个方程变的复杂了，怎么办？

To solve the system iteratively, we can define the following Jacobi iteration scheme

$$
\begin{aligned}
x_{k+1} &= \frac{7 + y_k - z_k}{4} \\
y_{k+1} &= \frac{21 + 4x_k + z_k}{8} \\
z_{k+1} &= \frac{15 + 2x_k - y_k}{5}
\end{aligned}
$$

An **algorithm** is then easily implemented computationally. In particular, we would follow the structure:

1. Guess initial values: (x_0, y_0, z_0).
2. Iterate the Jacobi scheme: $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k$.
3. Check for convergence: $\|\mathbf{x}_{k+1} - \mathbf{x}_k\| < \text{tolerance}$.

Note that **the choice of an initial guess** is often critical in determining the convergence to the solution.

Thus the more that is known about what the solution is supposed to look like, the higher the chance of successful implementation of the iterative scheme.

Table 1 shows the convergence of this scheme for this simple example.

k	x_k	y_k	z_k
0	1.0	2.0	2.0
1	1.75	3.375	3.0
2	1.84375	3.875	3.025
$\vdots$	$\vdots$	$\vdots$	$\vdots$
15	1.99999993	3.99999985	2.9999993
$\vdots$	$\vdots$	$\vdots$	$\vdots$
19	2.0	4.0	3.0

Given the success of this example, it is easy to conjecture (推测) that such a scheme will always be effective. However, we can reconsider the original system by interchanging the first and last set of equations. This gives the system

$$\boxed{4x - y + z = 7} \qquad\qquad -2x + y + 5z = 15$$
$$4x - 8y + z = -21 \quad \Rightarrow \quad 4x - 8y + z = -21$$
$$-2x + y + 5z = 15 \qquad\qquad \boxed{4x - y + z = 7}$$

To solve the system iteratively, we can define the following Jacobi iteration scheme based on this rearranged set of equations

$$x_{k+1} = \frac{y_k + 5z_k - 15}{2}$$
$$y_{k+1} = \frac{21 + 4x_k + z_k}{8}$$
$$z_{k+1} = y_k - 4x_k + 7.$$

Of course, the solution should be exactly as before. However, Table 2 shows that applying the iteration scheme leads to a set of values which grow to infinity. Thus the iteration scheme quickly fails.

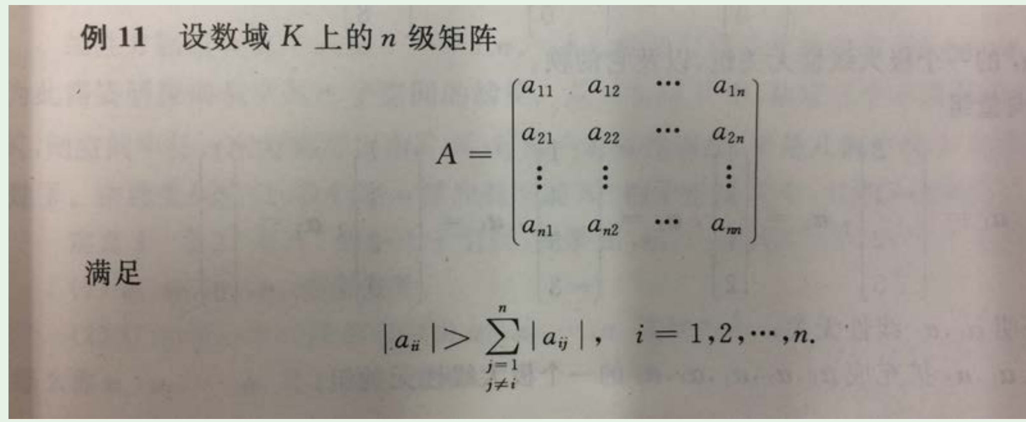
k	x_k	y_k	z_k
0	1.0	2.0	2.0
1	-1.5	3.375	5.0
2	6.6875	2.5	16.375
3	34.6875	8.015625	-17.25
$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$\pm\infty$	$\pm\infty$	$\pm\infty$

有没有一种方法可以直接判定是否可迭代出解？？？

Strictly Diagonal Dominant

The definition of **strict diagonal dominance** (严格对角优势): A matrix ***A*** is strictly diagonal dominant if for each row, the sum of the absolute values of the off-diagonal (非对角的) terms is less than the absolute value of the diagonal term:

$$|a_{kk}| > \sum_{j=1, j \neq k}^{N} |a_{kj}| .$$



例 11 设数域 K 上的 n 级矩阵

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

满足

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

Strict diagonal dominance has the following consequence

Given a strictly diagonal dominant matrix A,

then $Ax = b$ has a unique solution $x = p$.

Jacobi iteration produces a sequence p_k that will converge to p for any p_0.

For the two examples considered here, this property is crucial.

For the first example

$$\mathbf{A} = \begin{pmatrix} 4 & -1 & 1 \\ 4 & -8 & 1 \\ -2 & 1 & 5 \end{pmatrix} \rightarrow \begin{array}{l} \text{row } 1: |4| > |-1| + |1| = 2 \\ \text{row } 2: |-8| > |4| + |1| = 5 \\ \text{row } 3: |5| > |2| + |1| = 3 \end{array}$$

which shows the system to be strictly diagonal dominant and guaranteed to converge.

In contrast, the second system is not strictly diagonal dominant as can be seen from

$$\mathbf{A} = \begin{pmatrix} -2 & 1 & 5 \\ 4 & -8 & 1 \\ 4 & -1 & 1 \end{pmatrix} \rightarrow \begin{array}{l} \text{row } 1: |-2| < |1| + |5| = 6 \\ \text{row } 2: |-8| > |4| + |1| = 5 \\ \text{row } 3: |1| < |4| + |-1| = 5 \end{array}$$

Thus this scheme is **not guaranteed to converge**. Indeed, it diverges to infinity.

Modification and Enhancements: Gauss-Seidel

It is sometimes possible to enhance the convergence of a scheme by applying modifications to the basic Jacobi scheme.

For instance, the Jacobi scheme can be enhanced by the following modifications

有哪些方法可以加快收敛速度？？？怎么变形？？？

Jacobi (雅可比)

$$\boxed{x_{k+1}} = \frac{7 + y_k - z_k}{4}$$

$$\boxed{y_{k+1}} = \frac{21 + 4x_k + z_k}{8}$$

$$z_{k+1} = \frac{15 + 2x_k - y_k}{5}$$

Gauss-Seidel (高斯-赛德尔)

$$\boxed{x_{k+1}} = \frac{7 + y_k - z_k}{4}$$

$$\boxed{y_{k+1}} = \frac{21 + \boxed{4x_{k+1}} + z_k}{8}$$

$$z_{k+1} = \frac{15 + \boxed{2x_{k+1}} - \boxed{y_{k+1}}}{5}$$

Here use is made of the supposedly improved value x_{k+1} in the second equation and x_{k+1} and y_{k+1} in the third equation. This is known as the **Gauss-Seidel (高斯-赛德尔) scheme**.

Table 3 shows that the Gauss-Seidel procedure converges to the solution in half the number of iterations used by the Jacobi scheme.

k	x_k	y_k	z_k
0	1.0	2.0	2.0
1	1.75	3.75	2.95
2	1.95	3.96875	2.98625
$\vdots$	$\vdots$	$\vdots$	$\vdots$
10	2.0	4.0	3.0

The Gauss-Seidel modification is only one of a large number of possible changes to the iteration scheme which can be implemented in an effort to enhance convergence. Krylov space (克雷洛夫子空间) methods are often high end iterative techniques (高阶迭代技术) especially developed for rapid convergence.

线性方程组迭代解法续：判别收敛其他方法

严格对角优势：对于矩阵 $A=(a_{ij})_{n\times n}$ ，如果在绝对值意义下，对角线上元素大于对应行其他元素之和，那么称矩阵 A 具有严格对角优势.

$$|a_{ii}| > \sum_{j\neq i, 1\leq j\leq p} |a_{ij}|, \quad i=1,2,\cdots,n$$

弱对角优势：上面式子的大于改为大于等于，且至少有一个 i 严格大于成立.

可以利用矩阵的上述性质，得到收敛性。

定理1 严格对角优势=唯一解+Jacobi和Gauss-Seidel均收敛.

定理2 弱对角优势+不可约=唯一解+Jacobi和Gauss-Seidel均收敛.

定理3 对称正定矩阵=唯一解+Gauss-Seidel收敛.

线性方程组迭代解法续：判别收敛其他方法

华中科技大学 Huazhong University of Science and Technology

几个定义：

不可约：如果矩阵 A 不能经过有限次行列重排化为下面形式

$$\begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

其中 A_{11}, A_{22} 为方阵，那么 A 不可约.（将矩阵的某两行交换，并将相应两列也交换，称为一次**行列重排**）

可约：对于矩阵 A ，如果下标可以分为互补的 S, T ，使得

$$a_{ij} = 0, \quad i \in S, j \in T$$

则称 A 是可约的.

定义4：如果矩阵A不能通过行的互换和相应列的互换成为形式

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix}$$

其中$\mathbf{A}_{11}$，$\mathbf{A}_{22}$为方阵，则称A为不可约.

例如： 判断下列矩阵是否可约？

矩阵 $A=\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$ 是可约的。

交换第1与3行(列)

$$\left[\begin{array}{c|cc} 2 & 1 & 0 \\ \hline 0 & 1 & 1 \\ 0 & 1 & 1 \end{array}\right]$$

矩阵 $A=\begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}$ 是不可约的。

设有线性方程组Ax=b

定理7: 若A为严格对角占优阵或不可约弱对角占优阵，则Jacobi迭代法和Gauss-Seidel迭代法均收敛。 (Th.4.4.4)

定理5: 若 $\|\mathbf{H}\|<1$ ，则一般迭代法对于任意初始向量 $x^{(0)}$和g都收敛于方程组 **x=Hx+g** 的解**x***,且有下述误差估计式:

(1) $$\|x^{(k)}-x^*\|\le\frac{\|\mathbf{H}\|}{1-\|\mathbf{H}\|}\left\|x^{(k)}-x^{(k-1)}\right\|$$

(2) $$\|x^{(k)}-x^*\|\le\frac{\|\mathbf{H}\|^k}{1-\|\mathbf{H}\|}\left\|x^{(1)}-x^{(0)}\right\|$$

(Th. 4.4.1)

定理6:

对任意初始向量 $\mathbf{x}^{(0)}$和右端项 $\mathbf{g}$，由迭代格式

$$\mathbf{x}^{(k+1)} = \mathbf{H}\mathbf{x}^{(k)} + \mathbf{g}$$

产生的向量序列收敛的充要条件为

$$\rho(\mathbf{H}) < 1$$

(Th4.4.2)

推论1： 若迭代矩阵满足 $\|\mathbf{H}\|<1$，则迭代公式产生的向量序列$\{\mathbf{x}^{(k)}\}$收敛.

雅克比迭代法

将方程组记为 $\mathbf{Ax = b}$

其中A非奇异且$a_{ii} \neq 0$ (I=1, 2, …, n).

将A分裂为 $\mathbf{A = D + L + U}$

其中

$$D = \begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix} \quad L = \begin{bmatrix} 0 & & & \\ a_{21} & 0 & & \\ \cdots & & \ddots & \\ a_{n1} & a_{n2} & \cdots & 0 \end{bmatrix} \quad U = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ & 0 & \cdots & a_{2n} \\ & & \ddots & \cdots \\ & & & 0 \end{bmatrix}$$

雅克比迭代法

由此可将变形过程用矩阵表示为

$$Dx = -(L + U)x + b$$

即 $$x = -D^{-1}(L + U)x + D^{-1}b$$

简记为 $$x = Bx + d$$

故Jacobi迭代公式的矩阵形式为

$$\begin{cases} x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)})^T \text{(初始向量)} \\ x^{(k+1)} = Bx^{(k)} + d \end{cases}$$

高斯-赛德尔迭代法

将方程组记为 $\mathbf{Ax}=\mathbf{b}$

其中A非奇异且 $a_{ii}\neq 0\ (i=1, 2, \ldots, n)$.

将A分裂为 $\mathbf{A}=\mathbf{D}+\mathbf{L}+\mathbf{U}$

其中

$$\mathbf{D}=\begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & \ddots & \\ & & & a_{nn} \end{bmatrix} \quad \mathbf{L}=\begin{bmatrix} 0 & & & \\ a_{21} & 0 & & \\ \cdots & & 0 & \\ a_{n1} & a_{n2} & \cdots & 0 \end{bmatrix} \quad \mathbf{U}=\begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ & 0 & \cdots & a_{2n} \\ & & \ddots & \cdots \\ & & & 0 \end{bmatrix}$$

高斯-赛德尔迭代法

华中科技大学 Huazhong University of Science and Technology

由此可将方程组的变形过程用矩阵表示为

$$Dx = -(L + U)x + b$$

这G-S迭代可表示为

$$Dx^{(k+1)} = -Lx^{(k+1)} - Ux^{(k)} + b$$

整理得 $$x^{(k+1)} = -(D+L)^{-1}Ux^{(k)} + (D+L)^{-1}b$$

故G-S迭代公式的矩阵形式为

$$\begin{cases} x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T \text{(初始向量)} \\ x^{(k+1)} = -(D+L)^{-1}Ux^{(k)} + (D+L)^{-1}b \end{cases}$$

例题

例3: 解方程组

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

讨论Jacobi法与Gauss-Seidel法的收敛性。

解：由定理，迭代法是否收敛等价于迭代矩阵的谱半径是否＜1，故应先求迭代矩阵。而

$$A = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}$$

故A分解后的各矩阵分别为

例题

华中科技大学 Huazhong University of Science and Technology

$$\mathbf{D}=\begin{bmatrix}1 & & \\ & 1 & \\ & & 1\end{bmatrix}\qquad \mathbf{L}=\begin{bmatrix}0 & 0 & 0\\ 1 & 0 & 0\\ 2 & 2 & 0\end{bmatrix}\qquad \mathbf{U}=\begin{bmatrix}0 & 2 & -2\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}$$

$$\mathbf{A}=\begin{bmatrix}1 & 2 & -2\\ 1 & 1 & 1\\ 2 & 2 & 1\end{bmatrix}$$

Jacobi迭代法的迭代矩阵为

$$\mathbf{B}=-\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})=\begin{bmatrix}0 & -2 & 2\\ -1 & 0 & -1\\ -2 & -2 & 0\end{bmatrix}$$

其特征方程为

$$|\lambda I-B|=\begin{vmatrix}\lambda & 2 & -2\\ 1 & \lambda & 1\\ 2 & 2 & \lambda\end{vmatrix}=\lambda^3=0$$

因此有 $\rho(B)=0<1$, 故Jacobi法收敛

例题

如果用**Gauss-Seidel**迭代，由

$$\mathbf{D}+\mathbf{L}=\begin{bmatrix}1&0&0\\1&1&0\\2&2&1\end{bmatrix}\quad \text{可得}\quad (\mathbf{D}+\mathbf{L})^{-1}=\begin{bmatrix}1&0&0\\-1&1&0\\0&-2&1\end{bmatrix}$$

于是迭代矩阵为

$$\mathbf{G}=-(\mathbf{D}+\mathbf{L})^{-1}\mathbf{U}=\begin{bmatrix}0&-2&2\\0&2&-3\\0&0&2\end{bmatrix}$$

其特征方程为

例题

$$|\lambda\mathbf{I}-\mathbf{G}|=\begin{vmatrix}\lambda & 2 & -2\\ 0 & \lambda-2 & 3\\ 0 & 0 & \lambda-2\end{vmatrix}=\lambda(\lambda-2)^3=0$$

故 $\rho(\mathbf{G})=2>1,$

所以**Gauss-Seidel**迭代法发散.

例题

例6: 设线性方程组Ax=b的系数矩阵为

$$A=\begin{bmatrix} a & 1 & 3 \\ 1 & a & 2 \\ -3 & 2 & a \end{bmatrix}$$

试求能使Jacobi方法收敛的a的取值范围.

解: 当a ≠0时，Jacobi法的迭代矩阵为

$$B=\begin{bmatrix} 0 & -1/a & -3/a \\ -1/a & 0 & -2/a \\ 3/a & -2/a & 0 \end{bmatrix}$$

例题

解特征方程

$$|\lambda I - B| = \begin{vmatrix} \lambda & 1/a & 3/a \\ 1/a & \lambda & 2/a \\ -3/a & 2/a & \lambda \end{vmatrix} = 0$$

得 $\lambda_1 = 0,\quad \lambda_{2,3} = \pm\dfrac{4}{|a|}$ 故 $\rho(B) = \dfrac{4}{|a|}$

由 $\rho(B) < 1$ 得 $|a| > 4$

故当 $|a| > 4$ 时，Jacobi迭代法收敛。

习题

七、（10 分） 考虑线性方程组

$$\begin{bmatrix} a & b & 0 \\ c & a & b \\ 0 & c & a \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix},\text{其中} a\neq 0.$$

（1） 写出 Gauss-Seidel 迭代格式；

（2） 给出 Gauss-Seidel 迭代格式收敛的充分必要条件（a,b,c 之间应满足什么条件）。

七、解： （1）Gauss-Seidel 迭代格式为

$$\begin{cases} x^{(k+1)} = \left(d_1 - by^{(k)}\right)/a \\ y^{(k+1)} = \left(d_2 - cx^{(k+1)} - bz^{(k)}\right)/a \\ z^{(k+1)} = \left(d_3 - cy^{(k+1)}\right)/a \end{cases}$$

（2）Gauss-Seidel 迭代格式矩阵的特征方程为

$$\begin{vmatrix} a\lambda & b & 0 \\ c\lambda & a\lambda & b \\ 0 & c\lambda & a\lambda \end{vmatrix} = a^3\lambda^3 - 2abc\lambda^2 = 0$$

求得 $\lambda_{1,2} = 0, \lambda_3 = \dfrac{2bc}{a^2}$，则 Gauss-Seidel 迭代格式收敛的充要条件为

$$\left|\frac{2bc}{a^2}\right| < 1$$

即 $2|bc| < a^2$.