MEMO NO.
DATE

Chapter 6 线性方程组的解法

前提：线性方程组的解存在且唯一：系数矩阵行列式$\neq 0$

线性方程组的两种解法：

(1) 直接求解：精确解：适用于低阶稠密方程组

(2) 迭代法：逐步逼近精确解，大型稀疏方程组

§1 向量与矩阵范数：为估计误差而用

向量范数. n维向量R^n

如果向量$X\in R^n$的某个实值函数$N(X)=\|X\|$满足条件：

1. $\|X\|\geq 0$ $\|X\|=0$时，当且仅当$X=\vec{0}$ 正定.

2. $\|\alpha X\|=|\alpha|\|X\|$ $\forall \alpha\in R$ 齐次

3. $\|X+Y\|\leq \|X\|+\|Y\|$ 三角不等式

则称$N(X)$是向量X的范数

1. 无穷范数：∞范数 $\|X\|_\infty=\max\limits_{1\leq i\leq n}|x_i|$

2. 1范数：$\|X\|_1=\sum\limits_{i=1}^{n}|x_i|$

3. 2范数 $\|X\|_2=(\sum\limits_{i=1}^{n}x_i^2)^{\frac{1}{2}}$

4. p范数 $\|X\|_p=(\sum\limits_{i=1}^{n}|x_i|^p)^{\frac{1}{p}}$

向量序列$\{X^{(k)}\}$收敛于向量X^*是指对每一个$1\leq i\leq n$都有

$\lim\limits_{k\to\infty}x_i^{(k)}=x_i^*$ 每个分量$x_i^{(k)}$都收敛到x_i^* 充要条件

也等价于$\Longleftrightarrow$ $\lim\limits_{k\to\infty}\|X^{(k)}-X^*\|_\infty=0$ 即无穷范数收敛到0

MEMO NO.
DATE

若存在常数 $C>0$，对 $\forall X\in R^n$ 有 $\|X\|_A \le C\|X\|_B$ 则称范数 A 比范数B强，A比B强同时B也比A强，即 $C_1\|B\|_B \le \|X\|_A \le C_2\|B\|_B$

定理 $\boxed{R^n\text{上一切范数都是等价的}}$

定理：$N(X)=\|X\|$ 为 R^n 上任一向量范数，则 $N(X)$ 为分量 x_i 的连续函数

矩阵范数：

$R^{m\times n}$ 的空间的矩阵范数 $\|\ \|$ 满足：

(1) 正定 $\|A\|\ge 0$ $\|A\|=0$ 时当且仅当 $A=0$

(2) 齐次：$\|\alpha A\|=|\alpha|\|A\|$ 对 $\forall\alpha\in C$ 成立

(3) Δ不等式 $\|A+B\|\le\|A\|+\|B\|$

(4) 相容 $\|AB\|\le\|A\|\cdot\|B\|$

Frobenius范数：$\|A\|_F=\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}|a_{ij}|^2}$

算子范数：（导出范数）：由向量的P范数导出的关于矩阵的P范数.

$$\boxed{\|A\|_p=\max_{\vec{x}\neq\vec{0}}\frac{\|A\vec{x}\|_p}{\|\vec{x}\|_p}}$$

无穷范数 $\|A\|_\infty=\max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$ 最大行和，行和范数

$\|A\|_1=\max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|$ 最大列和，列和范数

MEMO NO.
DATE

A^TA的最大的特征值

$\|A\|_2 = \sqrt{\lambda_{max}(A^TA)}$ $\boxed{谱范数}$. 最大特征值(谱半径)开方.

谱: $\{\lambda_1, \lambda_2 \cdots \lambda_s\}$

F范数不是一个算子范数: 假设是由p范数导出. 取 $A=I$, 则不成立

谱半径: $\rho(A) = \max\limits_{1\le i\le n} |\lambda_i|$

定理: 对∀算子范数: $\rho(A) \le \|A\|$ $\boxed{谱半径 \le 范数}$

由范数的相容性: $\|Ax\| \le \|A\|\|x\|$, 将任意一个特征根λ与对应的特征向量u代入: $Au = \lambda u$.

$|\lambda|\|u\| = \|\lambda u\| = \|Au\| \le \|A\|\|u\|$

$\Rightarrow |\lambda| \le \|A\|$ 谱半径≤范数

定理: 若A对称, 则 $\boxed{\|A\|_2 = \rho(A)}$

A对称: 则 $A^T = A$.

λ是A的特征值.

则 λ^2 也是 A^2 的特征值 ∵ $\lambda x = Ax$ $A^2x = A\lambda x = \lambda Ax = \lambda^2 x$

$\lambda \to A$ 特征值

$\lambda^2 \to A^2$

……

$\lambda^n \to A^n$

$f(\lambda) \to f(A)$ $f(\lambda)$是多项式函数.

定理: 若矩阵B对某个算子范数满足 $\|B\| < 1$, 则有: ① $I \pm B$ 可逆. ② $\|(I \pm B)^{-1}\| \le \frac{1}{1-\|B\|}$

$\boxed{要记住}$

MEMO NO.
DATE

§2. 高斯消元法：直接求解法

$Ax=b$.

先将A化成一个上三角阵，然后回代求解.

相当于对增广矩阵 $(A|b)$ 作初等行变换变成上三角矩阵.

定理：如果A为n阶非奇异矩阵，即A的行列式不等于0，则可通过Gauss消元将方程化为三角形方程组.

引理：$a_{ii}^{(i)} \neq 0$ 的充要条件是矩阵A的顺序主子式 $D_i \neq 0$；则：

$$D_i = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ \vdots & & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} \end{vmatrix}$$

不需要交换两行就可以完成Gauss消元.

§3. 选主元消去法：

在Gauss消元过程中先找到该列中绝对值最大的元素，将其作为主元交换到第一行，然后其它行再减去这一行的倍数.

§4. 三角分解法：LU分解. 由Gauss消元导出的LU分解.

高斯消元的矩阵形式：

$l_{i1} = \frac{a_{i1}}{a_{11}}$

MEMO NO.
DATE

LU分解：L：下三角 U：上三角：

L：下三角并且对角线上元素全为1：$\boxed{\text{单位下三角}}$

~~[illegible]~~ $L_{n-1}\cdots L_2L_1A=U\Rightarrow A=\underbrace{L_1^{-1}L_2^{-1}\cdots L_{n-1}^{-1}}_{L}U$

$Ax=b$. $A=LU$ A非奇异

$\Rightarrow LUX=b\Rightarrow\begin{cases}UX=Y\\LY=b\end{cases}$ L.U 均是三角形的，所以会容易求解.

定理：A为n阶方阵，如果A的顺序主子式 $D_i\neq0$，则A的 LU 分解唯一

其中L为单位下三角.

假设不唯一，则 $A=L_1U_1=L_2U_2$

两边左乘 L_1^{-1} 右乘 U_2^{-1}

$$L_1^{-1}L_1U_1U_2^{-1}=L_1^{-1}L_2U_2U_2^{-1}$$

$$U_1U_2^{-1}=L_1^{-1}L_2$$

上三角　单位下三角　所以只能为单位矩阵 I.

$\therefore U_1U_2^{-1}=L_1^{-1}L_2=I\Rightarrow\begin{cases}L_1=L_2\\U_1=U_2\end{cases}$

道立特分解法——LU分解的紧凑格式.（直接三角分解法）

通过比较法直接导出L和U的计算公式

MEMO NO.
DATE

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix} = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ & & \ddots & \\ l_{n1} & & & 1\end{bmatrix}\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & \ddots & & \vdots \\ & & & u_{nn}\end{bmatrix}$$

先求U的第一行，再求L的第一列，再求U的第二行，L的第二列……

$\Rightarrow a_{ij} = \sum_{k=1}^{\min(i,j)} l_{ik}u_{kj}$

U的第一行，L的第一列

U的第二行，L的第二列

……继续下去可直接求出L和U

又称直接三角分解法。

如果需要交换两行（在Gauss消元过程中），则不能直接得到LU分解

但是可以实现：$PA=LU$

相当于将某些行交换之后，就符合之前的条件，可以进行LU分解

平方根法：适用于对称，正定矩阵的分解法。

对称矩阵：$A^T=A$，$a_{ij}=a_{ji}$

正定矩阵：$x^TAx>0$ 对 $\forall x\neq 0$ 都成立

MEMO NO.
DATE

对称.

A正定矩阵则:

(1) A^{-1}也是对称正定,且$a_{ii}>0$

(2) A的顺序主子阵A_k亦对称正定.

(3) A的特征值 $\lambda_i>0$

(4) A的全部顺序主子式 $\det(A_k)>0$

A^{-1} A_k. $\lambda_i>0$

$|A_k|>0$.

$$U=\begin{bmatrix} \diagdown & u_{ij} \\ & \end{bmatrix}=\underbrace{\begin{bmatrix} u_{11} & & & \\ & u_{22} & & \\ & & \ddots & \\ & & & u_{nn}\end{bmatrix}}_{D}\underbrace{\begin{bmatrix} 1 & & & u_{ij}/u_{ii} \\ & 1 & & \\ & & 1 & \\ & & & 1\end{bmatrix}}_{\tilde{U}}$$

对于对称正定矩阵A:

A对称:则有 $\begin{cases} A=LU=LD\tilde{U} \\ A^T=A \end{cases} \Rightarrow L=\tilde{U}^T$

即$A=LD\tilde{U}=LDL^T$ 又由于A正定,则$u_{11}\,u_{22}\cdots u_{nn}>0$

记$D^{1/2}=\begin{bmatrix}\sqrt{u_{11}} & & \\ & \sqrt{u_{22}} & \\ & & \ddots \sqrt{u_{nn}}\end{bmatrix}$ 则 令 $\tilde{L}=LD^{1/2}$ 仍是下三角阵 记.

$A=LDL^T$可化为. $A=\tilde{L}\tilde{L}^T$

设矩阵A对称正定,则存在非奇异的下三角阵$L\in R^{n\times n}$,使得$A=LL^T$
若限定L对角元为正,则分解唯一.

MEMO NO.
DATE

追赶法解三对角方程组:

$$A:\begin{bmatrix} b_1 & c_1 & & \\ a_2 & b_2 & c_2 & \\ & a_3 & \ddots & c_{n-1} \\ & & a_n & b_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=\begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$$

对A作 Crout分解: Crout分解:仍是LU分解式. 只是U为单位上三角

$$A=\underbrace{\begin{bmatrix} \alpha_1 & & & \\ r_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & r_n & \alpha_n \end{bmatrix}}_{L}\underbrace{\begin{bmatrix} 1 & \beta_1 & & \\ & 1 & \beta_2 & \\ & & \ddots & \beta_{n-1} \\ & & & 1 \end{bmatrix}}_{U}$$

$$\Rightarrow \begin{cases} r_i = a_i \\ \beta_i = \dfrac{c_i}{\alpha_i} \\ \alpha_i = b_i - r_i\beta_{i-1} \end{cases}$$

$A=\tilde{L}\tilde{U}$

$\tilde{L}\tilde{U}x=f \Leftrightarrow \begin{cases} \tilde{U}x=Y: 赶 \\ \tilde{L}Y=f: 追 \end{cases}$

对角占优矩阵: $A: n\times n$:

$$\begin{cases} ① |a_{ii}| \geq \sum\limits_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \\ ② 至少对于某一个k,\ |a_{kk}| > \sum\limits_{\substack{j=0 \\ j\neq k}}^{n} |a_{kj}| \end{cases}$$

MEMO NO.
DATE

严格对角占优：

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

定理：若A为对角占优的三对角阵，且满足

$|b_1|>|c_1|>0 \quad |b_n|>|a_n|>0 \quad a_i\neq 0 \quad c_i\neq 0$

则追赶法可以求解以A为系数矩阵的方程组

§5. 线性方程组的误差分析：

直接求解理论上不会有误差，但会由计算机引起误差

矩阵A的条件数：$\boxed{\|A\|\cdot\|A^{-1}\|}$ 条件数越大，越难得到精确解

常用条件数：$cond(A)_1$，$cond(A)_\infty$　$cond(A)_2$

若A对称，则 $cond(A)_2 = \|A\|_2\cdot\|A^{-1}\|_2 = \frac{\max|\lambda|}{\min|\lambda|}$ · $\sqrt{\frac{\lambda_{max}(A^TA)}{\lambda_{min}(A^TA)}}$

$=\sqrt{\lambda_{max}(A^TA)}\cdot\sqrt{\lambda_{max}((A^{-1})^TA^{-1})}$

由：$|\lambda I-AB| = |A||\lambda A^{-1}-B| = |A|\cdot|\lambda I-BA||A^{-1}| \overset{令}{=} 0$

∴ AB和BA的特征值相同

MEMO NO.
DATE . .

λ是A的特征值，则$\frac{1}{\lambda}$是A^{-1}的特征值

$(A^TA)^{-1} = A^{-1}(A^T)^{-1} = (A^T)^{-1}A^{-1}$

条件数的性质:

(1) A可逆 则 $cond(A)_p \geq 1$ $\quad \|A\|_p \|A^{-1}\|_p \geq \|A \cdot A^{-1}\|_p = 1$. 范数相容性

(2) A可逆 $\alpha \in R$ 则 $cond(\alpha A) = cond(A)$ $\quad \alpha \neq 0$

$cond(\alpha A) = \|\alpha A\| \|(\alpha A)^{-1}\| = |\alpha| \|A\| |\alpha|^{-1} \|A^{-1}\| = \|A\| \|A^{-1}\| = cond(A)$

齐次性质.

(3) A正交 $cond(A)_2 = 1$

A:正交: $\boxed{A^TA = I}$.

$cond(A)_2 = \|A\|_2 \|A^{-1}\|_2$

$= \sqrt{\frac{\lambda_{max}(A^TA)}{\lambda_{min}(A^TA)}} = 1$

(4) A可逆, R正交 则 $cond(RA)_2 = cond(AR)_2 = cond(A)_2$.

$cond(RA)_2 = \sqrt{\frac{\lambda_{max}[(RA)^TRA]}{\lambda_{min}[(RA)^T(RA)]}} = \sqrt{\frac{\lambda_{max}(A^TA)}{\lambda_{min}(A^TA)}}. = cond(A)_2$

可以用来衡量方程的病态程度: 若A的条件数很大, 则病态的, A或b的微小扰动都会给方程的解造成很大的影响

近似解的误差估计及改善

$Ax=b$ 的近似解为 x^*

误差 $r=b-Ax^*$

由：$\frac{\|x-x^*\|}{\|x\|} \le \operatorname{cond}(A)\frac{\|r\|}{\|b\|}$

改善方法：(1) $Ax=b \Rightarrow x_1$

(2) $r=b-Ax_1$

(3) $Ad_1=r_1 \Rightarrow d_1$

(4) $x_2=x_1+d_1$

§6. 解线性方程组的迭代方法：

求解 $Ax=b$，思路：

将 $Ax=b$ 等价改写为 $x=Bx+f$ 的形式，建立迭代 $x^{(k+1)}=Bx^{(k)}+f$

从初值 $x^{(0)}$ 出发，得到序列 $\{x^{(k)}\}$　B：迭代矩阵，适用于大型稀疏矩阵

Jacobi 法和 Gauss-Seidel 法：

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \cdots\cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n \end{cases} \quad a_{ii}\neq 0 \Rightarrow \begin{cases} x_1=\frac{1}{a_{11}}(-a_{12}x_2-\cdots-a_{1n}x_n+b_1) \\ x_2=\frac{1}{a_{22}}(-a_{21}x_1-\cdots-a_{2n}x_n+b_2) \\ \cdots\cdots \\ x_n=\frac{1}{a_{nn}}(-a_{n1}x_1-\cdots-a_{n-1\,n-1}x_{n-1}+b_n) \end{cases}$$

迭代形式：$x_i^{(k+1)}=\frac{1}{a_{ii}}\left(b_i-\sum_{\substack{j=1\\ j\neq i}}^{n} a_{ij}x_j^{(k)}\right)$

写成矩阵形式：记 $A=$ 

将 A 分成三块，L.D.U. 将 A 变成这三个之和

MEMO NO.
DATE . .

$AX=b \Rightarrow (D+L+U)X=b$

$DX=-(L+U)X+b$

$X=\underbrace{-D^{-1}(L+U)}_{B}X+\underbrace{D^{-1}b}_{f}$

Jacobi迭代阵：$-D^{-1}(L+U)$ 记作B_J

$X^{(k+1)}=-D^{-1}(L+U)X^{(k)}+D^{-1}b$

Jacobi迭代计算：$x_1^{(k+1)}$ $x_2^{(k+1)}\dots x_n^{(k+1)}$

计算每一个$x_i^{(k+1)}$都只用到$x_i^{(k)}$，不会用到$x_{i-1}^{(k+1)}$ $x_{i-2}^{(k+1)}\dots$ 这些刚求出来的最新值，所以Jacobi迭代可以修改

所以可以将迭代过程中求$x_i^{(k+1)}$时，对右边，$x_{i-1}^{(k)}$ $x_{i-2}^{(k)},\dots$ 改用新求出来的$x_{i-1}^{(k+1)}$ $x_{i-2}^{(k+1)}\dots$

由此导出Gauss-Seidel迭代：

$x_1^{(k+1)}=\frac{1}{a_{11}}(-a_{12}x_2^{(k)}-a_{13}x_3^{(k)}\dots-a_{1n}x_n^{(k)}+b_1)$

$x_2^{(k+1)}=\frac{1}{a_{22}}(-a_{21}x_1^{(k+1)}-a_{23}x_3^{(k)}\dots-a_{2n}x_n^{(k)}+b_2)$

$\dots$

$x_n^{(k+1)}=\frac{1}{a_{nn}}(-a_{n1}x_1^{(k+1)}-a_{n2}x_2^{(k+1)}\dots-a_{n,n-1}x_{n-1}^{(k+1)}+b_n)$

$X^{(k+1)}=-D^{-1}(LX^{(k+1)}+UX^{(k)})+D^{-1}b$

$X^{(k+1)}=-(D+L)^{-1}UX^{(k)}+(D+L)^{-1}b$

两种方法都有收敛性问题：

MEMO NO.
DATE

Jacobi迭代 $X^{(k+1)}=-D^{-1}(L+U)X^{(k)}+D^{-1}b$

Gauss-Seidel迭代: $X^{(k+1)}=-\underline{(D+L)^{-1}}UX^{(k)}+(D+L)^{-1}b$

§7.迭代法的收敛性:

$X^{(k+1)}=BX^{(k)}+f$ 的收敛条件:

X^*记为准确解: $e^{(k+1)}=X^{(k+1)}-X^*$

$=(BX^{(k)}+f)-(BX^*+f)$

$=B(X^{(k)}-X^*)=Be^{(k)}$

$\therefore \boxed{e^{(k)}=B^ke^{(0)}}$ $e^{(0)}=X^*-X^{(0)}$ $X^{(0)}$是随意给的初始值

$e^{(k)}\to 0$ $e^{(k)}$是一个向量,即要求$e^{(k)}$的每一个分量收敛到0

也即$\|e^{(k)}\|_*=0$ $e^{(k)}$的向量范数收敛到0

定理: $X=BX+f$存在唯一解.则从任意$X^{(0)}$出发.迭代$X^{(k+1)}=BX^{(k)}+f$

收敛 $\Longleftrightarrow$ $B^k\to 0$ 充要条件.

矩阵A_k收敛到A: $\lim\limits_{k\to+\infty}A_k=A$是指$\lim\limits_{k\to+\infty}a_{ij}^{(k)}=a_{ij}$ 对所有$1\le i,j\le n$

即每一个元素对应收敛

等价于$\|A_k-A\|\to 0$ 对于$\forall$算子范数成立

定理:$B^k\to 0 \Leftrightarrow \rho(B)<1$:转化为求迭代矩阵B的谱半径|特征值,充要.

前面有:谱半径≤范数: λ是B的特征值,则λ^k一定是B^k的特征值.

迭代法基本定理: $X=BX+f$

MEMO NO.
DATE

注意到：

$e^{(k+1)}=X^{(k+1)}-X^*=B(X^{(k)}-X^*)$

$=Be^{(k)}=B^{k+1}e^{(0)}$

$\forall\varepsilon>0$，$\exists$一个算子范数$\|\ \|$使得$\|A\|\le\rho(A)+\varepsilon$

$\|e^k\|=\|B^ke^{(0)}\|\le\|B^k\|\|e^{(0)}\|\le\|B\|^k\|e^{(0)}\|$

$\approx(\rho(B))^k\|e^{(0)}\|$

定义：称$R(B)=-\ln\rho(B)$为迭代算法的收敛速度.

$\rho(B)$越小，$R(B)$越大.

充分定理：充分条件. 给出了两种事后误差估计的方法：

若$\|B\|=q<1$，则迭代收敛，且有下列误差估计：

① $\|X^*-X^{(k)}\|\le\frac{q}{1-q}\|X^k-X^{k-1}\|$

② $\|X^*-X^{(k)}\|\le\frac{q^k}{1-q}\|X^{(1)}-X^{(0)}\|$

$X=BX+f$

$\|B\|$某种算子范数$<1\Rightarrow\rho(B)<1\Rightarrow$收敛

$>1\nRightarrow\rho(B)>1$无法判断是否收敛.

MEMO NO.
DATE

定理：充分条件：
若A为严格对角占优，则解 $Ax=b$ 的Jacobi和Gauss-Seidel迭代均收敛
若A严格对角占优则A的行列式 $\neq 0$
~~证明~~ 证明：没有记，看PPT或书上，课上只讲了Jacobi

∮松弛法：
Gauss-Seidel方法：

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

在后面的部分配一个 $a_{ii}x_i^{(k)}$，加一个减一个

$$x_i^{(k+1)} = x_i^{(k)} + \frac{1}{a_{ii}}\underbrace{\left[b_i - \sum_{j<i} a_{ij}x_j^{(k+1)} - \sum_{j\geq i} a_{ij}x_j^{(k)}\right]}_{\text{记为 } r_i^{(k+1)}}$$

$$= x_i^{(k)} + \frac{1}{a_{ii}} r_i^{(k+1)}$$ 相当于在 $x_i^{(k)}$ 上加一个余次生成 $x_i^{(k+1)}$

外推：$x_i^{(k+1)} = x_i^{(k)} + w\frac{r_i^{(k+1)}}{a_{ii}}$ 添加一个.