Examen. C.S.

# 整体思路及知识回顾及 LU 分解

当我们想计算 $Ax=b$ 时，

解决办法：

- 直接法：如果系数矩阵 A 稠密且固定不变，且右端项 b 不是太大
  Gauss 消去及其变形：LU 分解，列主元 LU 分解，Cholesky，LDL。
- 迭代法：如果系数矩阵 A 规模很大且是稀疏矩阵。
  Jacobi，Gauss-Seidel，SOR

Matrice Symétrique：对称矩阵：$A=A^T$，矩阵沿着对角线对称。

Matrice Orthogonale：正交矩阵：$A\cdot A^T=I$，I 是单位矩阵；$A\cdot A^T=\underset{\text{系数}}{\alpha}\cdot I$

Gaussian Elimination：高斯消去法：Gauss 消去的~~步骤~~过程实际也是对矩阵进行三角形分解的过程。

在线代中：

$$A=\begin{bmatrix}1&1&0\\1&2&1\\0&1&2\end{bmatrix}\Rightarrow\begin{bmatrix}1&1&0\\0&1&1\\0&1&2\end{bmatrix}\Rightarrow\begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}=U\ (\text{上三角矩阵})$$

LU Factorization：LU 分解

上述消元的过程保留消元因子即可得到 LU 分解

$$A=LU\Rightarrow\begin{bmatrix}1&1&0\\1&2&1\\0&1&2\end{bmatrix}=\underset{L}{\begin{bmatrix}1&0&0\\\boxed{1}&1&0\\0&\boxed{1}&1\end{bmatrix}}\underset{U}{\begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}}$$

Coût ou flops：$O(\frac{2n^3}{3})$

高斯消去的每一步相当于左乘初等矩阵，在上例中：

$$A=\begin{bmatrix}1&1&0\\1&2&1\\0&1&2\end{bmatrix}\Rightarrow E_1*A=\begin{bmatrix}1&0&0\\-1&1&0\\0&0&1\end{bmatrix}*\begin{bmatrix}1&1&0\\1&2&1\\0&1&2\end{bmatrix}=\begin{bmatrix}1&1&0\\0&1&1\\0&1&2\end{bmatrix}=A_1$$

$$\Rightarrow E_2*A_1=\begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix}*\begin{bmatrix}1&1&0\\0&1&1\\0&1&2\end{bmatrix}=\begin{bmatrix}1&1&0\\0&1&1\\0&0&1\end{bmatrix}=U$$

$$E_1*E_2*\dots*E_n*A=U$$

消元因子乘积的逆 $\Rightarrow L$

LU 分解适用范围：

if A is:

- positive definite：$x^TAx>0;\ \forall x\in\mathbb{R}^n$，对称且正定
- diagonally dominate：$|a_{ii}|>\sum_{j=1,j\neq i}^{n}|a_{ij}|$，对角线元素大于元素所在行的其它元素之和

列主元 Gauss 消去：（挑选列中绝对值最大的主元）

在 Gauss 消去的过程中保留消元因子和交换行.

$$\begin{bmatrix} 0 & 2 & 1 \\ 6 & 8 & 1 \\ 4 & 4 & 1 \end{bmatrix} \xrightarrow[\text{row change}]{(1,2)} \begin{bmatrix} 6 & 8 & 1 \\ 0 & 2 & 1 \\ 4 & 4 & 1 \end{bmatrix} \to \begin{bmatrix} 6 & 8 & 1 \\ 0 & 2 & 1 \\ 2/3 & -4/3 & 1/3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2/3 & -2/3 & 1 \end{bmatrix} \begin{bmatrix} 6 & 8 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

检验 $PA = LU$.

第一行与第二行互换

$$\underset{P}{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \underset{A}{\begin{bmatrix} 0 & 2 & 1 \\ 6 & 8 & 1 \\ 4 & 4 & 1 \end{bmatrix}} = \underset{L}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2/3 & -2/3 & 1 \end{bmatrix}} \underset{U}{\begin{bmatrix} 6 & 8 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}}$$

P（置换矩阵） A（非奇异矩阵） L U

Cholesky Factorization: if Matrix A is:

- positive definite, $x^T A x > 0$, $\forall x \neq 0$
- Symmetric（对称） positive（正定） definite $\Rightarrow$ Cholesky Factorization.

$$\begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{21} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

对称 正定

若 A 对称，且正定，$A = LU = L * D * U_0$（单位上三角），则 $L = U_0^T \Rightarrow A = L * D * L^T$
$= LL^T$

Coût en flops:

$O(\frac{n^3}{3})$

$LDL^T$ Algorithm:

若矩阵 A 对称（Symmetric），$A = A^T$，$A = LU \Rightarrow A = L^T U^T = A^T$

$A = L(DL^T) = LDL^T$，D 是对角阵.

$$\underset{A}{\begin{bmatrix} 4 & -8 & -4 \\ -8 & 18 & 14 \\ -4 & 14 & 25 \end{bmatrix}} = \underset{L}{\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 3 & 1 \end{bmatrix}} \times \underset{D}{\begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}} \times \underset{L^T}{\begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}}$$

Coût en flops:

$O(\frac{n^3}{3})$

Examen C.S.

# 简单迭代法（补充）

迭代格式：将 $Ax=b$ 改写成等价形式 $x=Bx+f$，构造迭代格式：

$$x^{(k+1)}=Bx^{(k)}+f,\quad k=0,1,2\dots$$

$Ax=b$ 的迭代格式一定存在，但未必都收敛

给定初值 $x^{(0)}$，反复作用上式可得到一个向量序列

$$x^{(0)},x^{(1)},x^{(2)},\dots,x^{(k)}\dots$$

若 $\lim\limits_{k\to\infty}x^{(k)}=x^*$，$x^*$ 是方程组的解，

则称迭代法收敛（Convergence），否则为发散（Divergence）

Question 1：$x^{(k+1)}=B*x^{(k)}+f$ 何时收敛？

记 $x^*$ 是精确解，引入误差向量 $\varepsilon^{(k)}=x^{(k)}-x^*$，

可知"迭代序列是否收敛"等价于 $\lim\limits_{k\to\infty}\varepsilon^{(k)}=\vec{0}$ → 趋近于 0.

$x^*$ 满足 $x^*=B*x^*+f$

与迭代格式对应相减，得 $\varepsilon^{(k+1)}=B\varepsilon^{(k)}=B^{(2)}\varepsilon^{(k-1)}=\dots=B^{(k+1)}\varepsilon^{(0)}$

$\varepsilon^{(0)}$ 是初始误差，一般不为 0. $\varepsilon^{(k+1)}$ 是否收敛于 0 向量，完全取决于迭代矩阵 B

$$\lim_{k\to\infty}\varepsilon^{(k+1)}=\lim_{k\to\infty}B^{k+1}\varepsilon^{(0)}=0$$

$$\|B^{k+1}\varepsilon^{(0)}\|\le\|B^{k+1}\|*\|\varepsilon^{(0)}\|\le\|B\|^{k+1}*\|\varepsilon^{(0)}\|$$

若存在某种迭代矩阵 B 的范数，使得 $\|B\|<1$，则迭代收敛.

又因为对于任意矩阵的范数都有：$\rho(B)\le\|B\|$（$\rho$：谱半径）

即最小的矩阵算子范数在数值上等于谱半径

谱半径：若 $A=(a_{ij})$ 是复数域上的 n 阶方阵，$\lambda_1,\lambda_2,\dots\lambda_n$ 是 A 的全部特征值，则

$$\rho(A)=\max_{1\le i\le n}|\lambda_i|$$

称为 A 的谱半径 $\rho(A)$.

所以，迭代是否收敛取决于迭代矩阵的谱半径是否 $<1$.

如果找到某种范数满足 $\|B\|_\diamond=q<1$（=谱半径），则迭代格式 $\varepsilon^{(k+1)}=B*\varepsilon^{(k)}$

可得，$\|\varepsilon^{(k+1)}\|\le q^{k+1}\|\varepsilon^{(0)}\|$ 前验估计

即 $\|x^{(k+1)}-x^{(k)}\|\ge(1-q)*\|\varepsilon^{(k)}\|$ 后验估计

在实际应用中多采用残差（residual）来判断数值解的精度.

$$r_k=Ax^*-Ax^{(k)}=A*(x^*-x^{(k)})=A*\varepsilon^{(k)}$$

Examen. C.S

# Jacobi 迭代法

Question 1：给定 $Ax=b$，如何选取迭代矩阵B，且满足收敛性条件？

$$Ax=b \iff Ax-b=0 \iff x = x - C*(Ax-b)$$

即

$$x=(I-CA)x+Cb$$

写成迭代形式为 $x^{(k+1)}=(I-CA)x^{(k)}+Cb$，$k=0,1,2,\dots$

根据《简单迭代法》所述，迭代矩阵 $B=I-CA$，当 $\|I-CA\|=q<1$ 时，迭代收敛，此时可称C是A的广义逆矩阵。$q=0$ 时，有 $CA=I$。

*以下的Jacobi、Gauss等方法实际上是在讨论如何选取C的值。

Jacobi 迭代：考虑特殊情况，A的对角线元素比该行其他元素都大（对角占优），形式上 $A\approx D=diag(A)$，即

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

Goût en flops：

$O(n^3)$

for Jacobi & GS

则 $C=D^{-1}$，那么迭代矩阵 $B=I-CA=I-D^{-1}*A$

将系数矩阵写成如下形式：$A=D+E+F$，其中E是下三角，F是上三角

则迭代矩阵 $B=D^{-1}*D-D^{-1}*(D+E+F)=-D^{-1}*(E+F)$

$$x^{(k+1)}=-D^{-1}(E+F)*x^{(k)}+D^{-1}b$$

$$\iff Dx^{(k+1)}=[b-(E+F)*x^{(k)}]$$

Jacobi 迭代格式：

$$x_i^{(k+1)}=\frac{1}{a_{ii}}\left[b_i-\sum_{j<i}a_{ij}*x_j^{(k)}-\sum_{j>i}a_{ij}*x_j^{(k)}\right],\ i=1,2,\dots,n$$

当我们反复利用已更新的分量时，可得到Gauss-Seidel迭代

两者计算复杂度几乎一样，但直观上GS效率更高。

$O(n^3)$

Gauss-Seidel 迭代格式：

$$x_i^{(k+1)}=\frac{1}{a_{ii}}\left[b_i-\sum_{j<i}a_{ij}*x_j^{(k+1)}-\sum_{j>i}a_{ij}*x_j^{(k)}\right],\ i=1,2,\dots,n.$$

# Gaussi - Seidel 迭代法

Question 2: 根据简单迭代推导公式, $x = x + C(b - Ax) = (I - CA)x + Cb$

GS迭代的广义逆矩阵C是什么?

Gauss-Seidel迭代的矩阵形式:

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

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

即

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

若记G为系数矩阵A的下三角矩阵(连同对角线元素), 上式即为

$$x^{(k+1)} = -G^{-1}Ux^{(k)} + G^{-1}b = G^{-1}(G-A)x^{(k)} + G^{-1}b.$$

收敛性分析:

- 严格对角占优(à Diagnale Dominante Stricte)

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

  此时Jacobi, GS均收敛

- A对称正定(Symétrique definie positive), GS收敛

- A以及(2D-A)均对称正定, Jacobi收敛.

SOR (Successive Over Relax) 超松弛迭代

本质上是SG的加权平均 weighted average.

1. 已知第k步的迭代值 $x^{(k)}$, 利用GS计算得到 $\overline{x^{(k+1)}}$ 中间值
2. 取两步迭代的加权平均值

$$x^{(k+1)} = (1-\omega)x^{(k)} + \omega\overline{x^{(k+1)}} = x^{(k)} + \omega(\overline{x^{(k+1)}} - x^{(k)})$$

$$\omega = \frac{2}{1 + 1 - \rho(D^{-1}\cdot(E+F))^2}$$

Examen C.S.

# Algorithmes iteratifs de relaxation

On cherche une solution $X^*$ tentque. $Ax^* = b$ avec un algorithme iteratif de type $x^{(p+1)} = H(x^p)$.

Algo iteratif de relaxation — Méthod de grandient.

$$A = M - N$$

**Algo iteratif de relaxation:**

$A = D - E - F$, avec: $D = \begin{bmatrix} \ddots & \\ & \ddots \end{bmatrix}$ 对角阵, $E = \begin{bmatrix} 0 & & 0 \\ -a_{21} & \ddots & \\ \vdots & \dots & 0 \end{bmatrix}$, $F = \begin{bmatrix} 0 & -a_{12} & \dots \\ & \ddots & \vdots \\ 0 & & 0 \end{bmatrix}$

associé: $\begin{cases} X^{(0)} \in \mathbb{R}^{(n)} \\ X^{(p+1)} = X^{(p)} + M^{-1} \cdot r^{(p)} \end{cases}$ avec $r^{(p)} = b - A \cdot x^{(p)}$ le vecteur résidu. (残差向量)

Goût d'une iteration: en dehors du calcul du résidu: resolution d'un système de matrice $M$.

$$\begin{aligned} Ax^* = b &\iff (M - N) x^* = b \\ &\iff M x^* = N x^* + b \\ &\iff x^* = M^{-1} N x^* + M^{-1} b \\ &\iff x^* = M^{-1}(M - A) x^* + M^{-1} b \\ &\iff x^* = x^* + M^{-1}[b - Ax^*] \quad ? \end{aligned}$$

**2 variantes:**

2 variantes (2种变体):

- Algo de Gauss-Seidel: $M = D - E$, $N = F$.

  $M x^* = N x^* + b$

  $\iff (D - E) x^{(p+1)} - F x^{(p)} = b \iff x^{(p+1)} = (D - E)^{-1}[b - Ax^{(p+1)}] + x^{(p)}$

- Algo de Jacobi: $M = D$, $N = E + F$.

  $M x^* = N x^* + b$

  $\iff D x^{(p+1)} - (E + F) x^{(p)} = b \iff x^{(p+1)} = D^{-1}[b - Ax^{(p)}] + x^{(p)}$

**Condition d'Arrêt:** $\dfrac{\|Ax^{(p+1)} - b\|}{\|b\|} < \varepsilon$

收敛的充要条件：*
Condition Necessaire et Suffisante de Convergence:

$$\lim_{p\to+\infty} (M^{-1}\cdot N)^p = 0 \iff \rho(M^{-1}\cdot N) < 1$$

谱半径

*理论详见《简单迭代法(补充)》- Question 1.

收敛的充分条件：
Condition Suffisante de Convergence:

- Si "A" est à Diagonale Dominante Stricte $\forall x_0, \forall i, \ |a_{ii}| \geq \sum_{j=1, j\neq i}^{n} |a_{ij}|$ ⇒ Convergence pour Gauss. ~~∀x₀~~
- Si "A" est à Diagonale Dominante et $\exists k \in \{1, \dots, n\}$, $|a_{kk}| > \sum_{j=1, j\neq k}^{n} |a_{kj}|$. ⇒ CV
- Si "A" est symétrique définie positive ⇒ Convergence pour Gauss

Methode SOR: 详见课件《Méthodes itératives》- $P_{13}$

Méthode de la Steepest-Descent:

Si "A" est Symétrique définie positive, $x^{(0)} \in \mathbb{R}^n$

et $x^{(p+1)} = x^{(p)} + \underbrace{\lambda_p}_{} \underline{r^{(p)}}$ → direction de descente (vecteur) 下降方向(矢量)

$\lambda_p$: progression dans la direction de descente (scalaire) 下降方向的前进步长(标量)

- Direction: $r^{(p)} = -\mathrm{grad}(F)(x^{(p)}) = b - A\cdot x^{(p)}$
- Progression: $\lambda^{(p)} = \dfrac{(r^{(p)})^T \cdot r^{(p)}}{(r^{(p)})^T \cdot A \cdot r^{(p)}}$, qui minimum dans la direction $r^{(p)}$.

Texamen. C.S.

# QR factorization.

## Householder 变换.

* 对任意模长为1的向量 $w \in R^n$，检验反射矩阵 $H = I - 2ww^T$，满足

$$H = H^T, \quad H \times H = I$$ 即 $H$ 为对称正交矩阵.

对称矩阵 　　 正交矩阵.

记超平面(过原点，以 $w$ 为法向量)

$$S = \{ x \mid w^T x = 0, \ \forall x \in R^n \}$$ 即S平面是与w垂直的平面.

知 $S$ 是 $R^n$ 的子空间，维数或秩为 $(n-1)$，则任意向量 $z \in R^n$ 可以在子空间 $S$ 和 $\mathrm{span}\{w\}$ 做正交分解：

系数

$$z = x + y, \quad x \in S, \quad y = \alpha \cdot w, \quad \alpha \in R$$

求得. $Hz = Hx + Hy = x + (I - 2ww^T)y = x + y - 2ww^T \alpha w = x - y$

几何上，上述反射矩阵(Householder 变换)可看为是以 $S$ 为镜面的反射.

* 给定两个模长相等的向量 $x, y \in R^n$，则存在反射变换 $H$，使得

$$Hx = y \qquad w = \frac{x - y}{\|x - y\|}$$

Décomposition en valeurs singulières

# SVD Factorization 奇异值分解.

定义：矩阵的奇异值分解是指，将一个非零的 $m\times n$ 实矩阵 $A$，$A\in R^{m\times n}$，表示成以下三个实矩阵乘积形式的运算，即进行矩阵的因子分解：

奇异向量 singular vector

$$\underset{m\times n}{A} = \underset{m\times m}{U}\ \underset{m\times n}{\Sigma}\ \underset{n\times n}{V^T}, \qquad \Sigma = diag(\sigma_1, \sigma_2, \dots \sigma_p), \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$$

其中，$U$是$m$阶正交矩阵（orthogonal matrix），

$\Sigma$是由降序排列的非负的对角线元素组成的 $m\times n$ 矩形对角阵

$V$是$n$阶正交矩阵。

$U$的列向量称为左奇异向量，$V$的列向量称为右奇异向量。

$\sigma_i$称为矩阵$A$的奇异值（singular value）。

只要矩阵$A$是一个实矩阵，其奇异值分解一定存在。

奇异值分解计算步骤：$A$为一个 $m\times n$ 实矩阵，有：

① 构造$n$阶实对称矩阵 $W = A^TA$

② 计算$W$的特征值和特征向量

求解特征方程 $(W-\lambda I)x = 0$，得到特征值$\lambda_i$，并将特征值由大到小排列，$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$，将特征值$\lambda_i$ $(i=1,2,\dots,n)$代入特征方程求得对应的特征向量

③ 求得$n$阶正交矩阵$V$（利用上述求得"特征向量"）

将特征向量单位化，得到单位特征向量 $v_1, v_2, \dots, v_n$，构成$n$阶正交矩阵$V$：$V = [v_1, v_2, \dots, v_n]$。

④ 求得 $m\times n$ 对角矩阵$\Sigma$（利用上述求得"特征值"）

计算$A$的奇异值 $\sigma_i = \sqrt{\lambda_i}$，$i = 1, 2, \dots, n$.

构造 $m\times n$ 矩形对角矩阵$\Sigma$，主对角线元素是奇异值，其余元素是零。$\Sigma = diag(\sigma_1, \sigma_2, \dots, \sigma_n)$.

⑤ 求得$m$阶正交矩阵$U$（利用上述求得的$V$和$\Sigma$）。

i. 求$U_1$

对于$A$的前$r$个正奇异值，令 $u_j = \frac{1}{\sigma_j}Av_j$，$j = 1, 2, \dots, r$

得到 $U_1 = [u_1, u_2, \dots, u_r]$

ii. 求 $U_2$.

求$A^T$的零空间的一组标准正交基 $\{u_{r+1}, u_{r+2}, \dots, u_m\}$.

令 $U_2 = [u_{r+1} \quad u_{r+2} \quad \dots \quad u_n]$

iii. 得到 $U = [U_1 \quad U_2]$.

SVD法缺点：不适合处理大型稀疏矩阵.