

第四章 解线性方程组的直接法

中国科学技术大学 数学学院

chenxjin@ustc.edu.cn

问题：解线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \end{cases}$$

$\iff Ax = b$, 其中 $b = (b_1, b_2, \dots, b_n)^T$, A 为系数矩阵。

问题：解线性方程组

$$\left\{ \begin{array}{l} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \end{array} \right.$$

$\iff Ax = b$, 其中 $b = (b_1, b_2, \dots, b_n)^T$, A 为系数矩阵。

实际中，存在大量的解线性方程组的问题。很多数值方法本身到最后也会归结为解线性方程组，如：

- 样条插值的M和m关系式
- 曲线拟合的法方程
- 非线性方程组的Newton迭代法
-

理论上我们有Cramer法则：当且仅当 $\det(A) \neq 0$ 时，方程组有唯一的解；而且解为 $x_i = \frac{D_i}{D}$. 其中

$$D = \det(A), D_i = \det \begin{pmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{pmatrix}$$

理论上我们有Cramer法则：当且仅当 $\det(A) \neq 0$ 时，方程组有唯一的解；而且解为 $x_i = \frac{D_i}{D}$. 其中

$$D = \det(A), D_i = \det \begin{pmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{pmatrix}$$

但在实际计算中，Cramer法则往往无法执行（使用）！例如 $n = 100$ ，需做约 $101 \cdot 100! \cdot 99$ 次乘除法！

理论上我们有Cramer法则：当且仅当 $\det(A) \neq 0$ 时，方程组有唯一的解；而且解为 $x_i = \frac{D_i}{D}$. 其中

$$D = \det(A), D_i = \det \begin{pmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{pmatrix}$$

但在实际计算中，Cramer法则往往无法执行（使用）！例如 $n = 100$ ，需做约 $101 \cdot 100! \cdot 99$ 次乘除法！“神威·太湖之光”是目前我国最强的超级计算机，使用的是完全自主知识产权的国产芯片，峰值计算速度达每秒12.54亿亿次，持续计算速度每秒9.3亿亿次！

理论上我们有Cramer法则：当且仅当 $\det(A) \neq 0$ 时，方程组有唯一的解；而且解为 $x_i = \frac{D_i}{D}$. 其中

$$D = \det(A), D_i = \det \begin{pmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{pmatrix}$$

但在实际计算中，Cramer法则往往无法执行（使用）！例如 $n = 100$ ，需做约 $101 \cdot 100! \cdot 99$ 次乘除法！“神威·太湖之光”是目前我国最强的超级计算机，使用的是完全自主知识产权的国产芯片，其峰值计算速度达每秒12.54亿亿次，持续计算速度每秒9.3亿亿次！若用Cramer法则去求解一个普通的 100×100 的线性方程组，即使用速度为 10^{33} 次/秒的超级计算机，需要算

理论上我们有Cramer法则：当且仅当 $\det(A) \neq 0$ 时，方程组有唯一的解；而且解为 $x_i = \frac{D_i}{D}$. 其中

$$D = \det(A), D_i = \det \begin{pmatrix} a_{11} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{nn} \end{pmatrix}$$

但在实际计算中，Cramer法则往往无法执行（使用）！例如 $n = 100$ ，需做约 $101 \cdot 100! \cdot 99$ 次乘除法！“神威·太湖之光”是目前我国最强的超级计算机，使用的是完全自主知识产权的国产芯片，其峰值计算速度达每秒12.54亿亿次，持续计算速度每秒9.3亿亿次！若用Cramer法则去求解一个普通的 100×100 的线性方程组，即使用速度为 10^{33} 次/秒的超级计算机，需要算

$$\frac{101 \cdot 100! \cdot 99}{365 \cdot 24 \cdot 3600 \cdot 10^{33}} \approx 10^? \text{ 年}$$

解线性方程组的方法可以分为2类：

解线性方程组的方法可以分为2类：

- ① **直接法**：准确，可靠，理论上得到的解是精确的；
- ② **迭代法**：速度快，但有误差.

本章讲解直接法.

线性方程组 $Ax = b$: 三种特例

线性方程组 $Ax = b$: 三种特例

下面3种特殊线性代数方程组 $Ax = b$ 的解，可以直接求出：

1. $A = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}) \Rightarrow x_i = \frac{b_i}{a_{ii}} (i = 1, \dots, n)$ \longmapsto **n**次乘(除)法

线性方程组 $Ax = b$: 三种特例

下面3种特殊线性代数方程组 $Ax = b$ 的解, 可以直接求出:

1. $A = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}) \Rightarrow x_i = \frac{b_i}{a_{ii}} (i = 1, \dots, n)$ \longmapsto **n**次乘(除)法

$$2. A = \begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix}$$

线性方程组 $Ax = b$: 三种特例

下面3种特殊线性代数方程组 $Ax = b$ 的解, 可以直接求出:

1. $A = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}) \Rightarrow x_i = \frac{b_i}{a_{ii}} (i = 1, \dots, n)$ \mapsto n 次乘(除)法

$$2. A = \begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \Rightarrow x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij} x_j}{l_{ii}} \mapsto \frac{n(n+1)}{2} \text{ 次乘法}$$

线性方程组 $Ax = b$: 三种特例

下面3种特殊线性代数方程组 $Ax = b$ 的解, 可以直接求出:

1. $A = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}) \Rightarrow x_i = \frac{b_i}{a_{ii}} (i = 1, \dots, n)$ \mapsto **n次乘(除)法**

$$2. A = \begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \Rightarrow x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij}x_j}{l_{ii}} \mapsto \frac{n(n+1)}{2} \text{ 次乘法}$$

$$3. A = \begin{pmatrix} u_{11} & u_{12} & \cdots & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & \cdots & u_{2n} \\ \ddots & \ddots & \cdots & & \vdots \\ & & u_{(n-1)(n-1)} & u_{(n-1)n} & u_{nn} \end{pmatrix}$$

线性方程组 $Ax = b$: 三种特例

下面3种特殊线性代数方程组 $Ax = b$ 的解, 可以直接求出:

1. $A = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}) \Rightarrow x_i = \frac{b_i}{a_{ii}} (i = 1, \dots, n)$ \mapsto n 次乘(除)法

$$2. A = \begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \Rightarrow x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij}x_j}{l_{ii}} \mapsto \frac{n(n+1)}{2} \text{ 次乘法}$$

$$3. A = \begin{pmatrix} u_{11} & u_{12} & \cdots & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & \cdots & u_{2n} \\ \ddots & \ddots & \cdots & & \vdots \\ & u_{(n-1)(n-1)} & u_{(n-1)n} & & u_{nn} \end{pmatrix} \Rightarrow x_i = \frac{b_i - \sum_{j=i+1}^n u_{ij}x_j}{u_{ii}}$$

$\mapsto \frac{n(n+1)}{2}$ 次乘法

消元法

消元法就是对增广矩阵 (A, b) 作行变换，变为上述的3种类型之一后再求解.

消元法

消元法就是对增广矩阵 (A, b) 作行变换，变为上述的3种类型之一后再求解.

例如，Gauss消元法即化为上三角形式.

$$\left(\begin{array}{ccccc} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array} \right) \Rightarrow \left(\begin{array}{ccccc} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & b_3^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n)} & b_n^{(n)} \end{array} \right)$$

具体步骤如下.

Gauss消元法

第一步： 第1行 $\times \left(-\frac{a_{i1}}{a_{11}} \right)$ + 第 i 行, $i = 2, \dots, n.$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{pmatrix} \implies \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{pmatrix}$$

运算量： $(n - 1) \cdot (n + 1)$

第二步： 第2行 $\times \left(-\frac{a_{i2}^{(2)}}{a_{22}^{(2)}} \right)$ + 第 i 行, $i = 3, \dots, n.$

$$\left(\begin{array}{ccccc} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{array} \right) \implies \left(\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & b_3^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} & b_n^{(3)} \end{array} \right)$$

运算量: $(n-2) \cdot (1+n-1) = (n-2)n$

类似的做下去, 第 k 步: 第 k 行 $\times \left(-\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \right)$ + 第 i 行, $i = k+1, \dots, n.$

运算量: $(n-k) \cdot (1+n-k+1) = (n-k)(n-k+2).$

类似的做下去，第 k 步：第 k 行 $\times \left(-\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \right)$ + 第 i 行， $i = k+1, \dots, n$.

运算量： $(n-k) \cdot (1+n-k+1) = (n-k)(n-k+2)$.

$n-1$ 步以后，可得增广矩阵：

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & b_3^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n)} & b_n^{(n)} \end{pmatrix}$$

化为上三角方程组的运算量为 $\sum_{k=1}^{n-1} (n-k)(n-k+2)$ ，加上解上述上三角阵的运算量 $\frac{(n+1)n}{2}$ ，总共为

$$\frac{n^3}{3} + n^2 - \frac{n}{3} = O(n^3)$$

Gauss-Jordan消元法

Gauss-Jordan消元法

将Gauss消元第 k 步变为：

$$\text{第 } k \text{ 行} \times \left(-\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \right) + \text{第 } i \text{ 行}, \quad i = 1, \dots, k-1, k+1, \dots, n.$$

最后系数矩阵被消元成一对角阵，即称为Gauss-Jordan消元法.

也可这样认为：Gauss-Jordan消元即在Gauss消元完成后（此时系数矩阵为上三角阵），继续用对角元进行向上消元（打洞），直至变成对角阵，再求解.

Gauss-Jordan消元法计算复杂度比Gauss消元法略多，量级一样.

在Gauss消元过程中，须确保 $a_{kk}^{(k)} \neq 0$,

在Gauss消元过程中，须确保 $a_{kk}^{(k)} \neq 0$ ，
等价于 A 的所有顺序主子式均不为0，即

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} \end{pmatrix} \neq 0, \quad i = 1, \dots, n$$

在Gauss消元过程中，须确保 $a_{kk}^{(k)} \neq 0$ ，
等价于 A 的所有顺序主子式均不为0，即

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} \end{pmatrix} \neq 0, \quad i = 1, \dots, n$$

这是Gauss消元法理论上可行的充要条件.

在Gauss消元过程中，须确保 $a_{kk}^{(k)} \neq 0$ ，
等价于 A 的所有顺序主子式均不为0，即

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} \end{pmatrix} \neq 0, \quad i = 1, \dots, n$$

这是Gauss消元法理论上可行的充要条件.

注：如果某步的 $a_{kk}^{(k)}$ 很小的话，可能会产生较大的计算误差，从而导致最终的计算（求解）失败！

例：用单精度（每次运算结果保留4位小数）解二元方程组

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ 1.0000x_1 + 1.0000x_2 = 1.0000 \end{cases}$$

例：用单精度（每次运算结果保留4位小数）解二元方程组

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ 1.0000x_1 + 1.0000x_2 = 1.0000 \end{cases}$$

按Gauss消元法得

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ -9999.0000x_2 = -6666.0000 \end{cases}$$

例：用单精度（每次运算结果保留4位小数）解二元方程组

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ 1.0000x_1 + 1.0000x_2 = 1.0000 \end{cases}$$

按Gauss消元法得

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ -9999.0000x_2 = -6666.0000 \end{cases}$$

$$\Rightarrow x_2 = 0.6667, \quad x_1 = 0.0000$$

例：用单精度（每次运算结果保留4位小数）解二元方程组

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ 1.0000x_1 + 1.0000x_2 = 1.0000 \end{cases}$$

按Gauss消元法得

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ -9999.0000x_2 = -6666.0000 \end{cases}$$

$$\Rightarrow x_2 = 0.6667, \quad x_1 = 0.0000$$

$$\text{真实解为 } x_2 = \frac{2}{3} \approx 0.6667, \quad x_1 = \frac{1}{3} \approx 0.3333$$

列主元消元法

在Gauss消元第 k 步消元之前，曾广矩阵 (A, b) 为

$$\left(\begin{array}{cccccc} a_{11} & \cdots & \cdots & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & b_k^{(k)} \\ 0 & 0 & a_{(k+1)k}^{(k)} & \cdots & a_{(k+1)n}^{(k)} & b_{(k+1)}^{(k)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} & b_n^{(k)} \end{array} \right)$$

列主元消元法

在Gauss消元第 k 步消元之前，曾广矩阵 (A, b) 为

$$\left(\begin{array}{cccccc} a_{11} & \cdots & \cdots & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & b_k^{(k)} \\ 0 & 0 & a_{(k+1)k}^{(k)} & \cdots & a_{(k+1)n}^{(k)} & b_{(k+1)}^{(k)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} & b_n^{(k)} \end{array} \right)$$

- 列主元消元的基本思路：先选列主元，再消元；若 $\max_{k \leq i \leq n} |a_{ik}^{(k)}| = |a_{jk}^{(k)}|$ ，交换 k 行和 j 行，即将第 k 到 n 列的模最大元通过行交换移到对角位置，再进行消元。

例：单精度解方程组（每次运算结果保留4位小数）

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ 1.0000x_1 + 1.0000x_2 = 1.0000 \end{cases}$$

按列主元消元法得

$$\begin{cases} 1.0000x_1 + 1.0000x_2 = 1.0000 \\ 2.9997x_2 = 1.9998 \end{cases}$$

例：单精度解方程组（每次运算结果保留4位小数）

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ 1.0000x_1 + 1.0000x_2 = 1.0000 \end{cases}$$

按列主元消元法得

$$\begin{cases} 1.0000x_1 + 1.0000x_2 = 1.0000 \\ 2.9997x_2 = 1.9998 \end{cases}$$

$$\Rightarrow x_2 = 0.6667, x_1 = 0.3333$$

Gauss消元法回顾

Gauss消元法回顾

第1步： 第1行 $\times \left(-\frac{a_{i1}}{a_{11}}\right)$ + 第 i 行, $i = 2, \dots, n.$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{pmatrix} \implies \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{pmatrix}$$

$n - 1$ 步以后，可得增广矩阵：

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & b_3^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n)} & b_n^{(n)} \end{pmatrix}$$

Gauss消元法回顾

第1步： 第1行 $\times \left(-\frac{a_{i1}}{a_{11}}\right)$ + 第 i 行, $i = 2, \dots, n.$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{pmatrix} \implies \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{pmatrix}$$

$n - 1$ 步以后，可得增广矩阵：

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & b_2^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & b_3^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n)} & b_n^{(n)} \end{pmatrix}$$

- 问题：Gauss消元法本质上到底在做什么？

直接分解法

Gauss消元法第 k 步为:

$$\text{第 } k \text{ 行} \times \left(-\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \right) + \text{第 } i \text{ 行}, \quad i = k+1, \dots, n$$

直接分解法

Gauss消元法第 k 步为:

$$\text{第 } k \text{ 行} \times \left(-\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \right) + \text{第 } i \text{ 行}, \quad i = k+1, \dots, n$$

从矩阵理论来看, 第 k 步消元等价于将增广矩阵 (A, b) 左乘一个单位下三角阵

$$P_k = \begin{pmatrix} 1 & & & & \\ \vdots & \ddots & & & \\ & & 1 & & \\ & & \ell_{k+1,k}^{(k)} & 1 & \\ & & \vdots & & \ddots \\ & & \ell_{nk}^{(k)} & & 1 \end{pmatrix}, \quad \text{其中 } \ell_{ik}^{(k)} = \frac{-a_{ik}^{(k)}}{a_{kk}^{(k)}}, i = k+1, \dots, n$$

Gauss消元法实际上就是不断地用单位下三角阵 P_k 左乘 (A, b) , 最后得到一个上三角阵 $P_{n-1} \cdot \dots \cdot P_1 \cdot A$.

Gauss消元法实际上就是不断地用单位下三角阵 P_k 左乘 (A, b) , 最后得到一个上三角阵 $P_{n-1} \cdots P_1 \cdot A$. 因此, 整个Gauss消元法相当于左乘了一个单位下三角阵

$$\tilde{L} \triangleq P_{n-1} \cdots P_1 = \begin{pmatrix} 1 & & & & & \\ \tilde{\ell}_{21} & \ddots & & & & \\ \vdots & & 1 & & & \\ & & \tilde{\ell}_{k+1,k} & 1 & & \\ & & & \vdots & & \ddots \\ \tilde{\ell}_{n1} & \cdots & \tilde{\ell}_{nk} & \cdots & \tilde{\ell}_{n,n-1} & 1 \end{pmatrix}$$

最后使矩阵 A 最后变成上三角阵 U , 即 $\tilde{L}A = U$.

Gauss消元法实际上就是不断地用单位下三角阵 P_k 左乘 (A, b) , 最后得到一个上三角阵 $P_{n-1} \cdots P_1 \cdot A$. 因此, 整个Gauss消元法相当于左乘了一个单位下三角阵

$$\tilde{L} \triangleq P_{n-1} \cdots P_1 = \begin{pmatrix} 1 & & & & & \\ \tilde{\ell}_{21} & \ddots & & & & \\ \vdots & & 1 & & & \\ & & \tilde{\ell}_{k+1,k} & 1 & & \\ & & & \vdots & & \ddots \\ \tilde{\ell}_{n1} & \cdots & \tilde{\ell}_{nk} & \cdots & \tilde{\ell}_{n,n-1} & 1 \end{pmatrix}$$

最后使矩阵 A 最后变成上三角阵 U , 即 $\tilde{L}A = U$.

结论: 若 A 的各阶顺序主子式不为0, 则 有 $A = LU$ ($L = \tilde{L}^{-1}$). 注: Gauss消元法的实质就是要对矩阵 A 作 LU 分解!

直接分解法

要解方程组 $Ax = b$, 可先将 A 进行 LU 分解, 再转化为两个容易求解的方程组.

$$O(n^3) \ Ax = b \iff LUx = b \iff \begin{cases} Ly = b & (1) \\ Ux = y & (2) \end{cases}$$

先由(1)求 y , 再代入(2)求 x .

注: 能进行直接分解法的条件与 Gauss 消元法相同, 即各阶顺序主子式不为 0.

LU Factorizations (分解)

- 问题：如何对矩阵 A 作 LU 分解？若分解存在，分解是否唯一？

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

LU Factorizations (分解)

- 问题：如何对矩阵 A 作 LU 分解？若分解存在，分解是否唯一？

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

- Check— equations: ? ; unknowns: ?

LU Factorizations (分解)

- 问题：如何对矩阵 A 作 LU 分解？若分解存在，分解是否唯一？

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

- Check— equations: ? ; unknowns: ?
- 若要求 L 为单位下三角，则称为 Doolittle 分解.

LU Factorizations (分解)

- 问题：如何对矩阵 A 作 LU 分解？若分解存在，分解是否唯一？

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

- Check— equations: ? ; unknowns: ?
- 若要求 L 为单位下三角，则称为 Doolittle 分解。
- 若要求 U 为单位上三角，则称为 Crout's Factorization.

LU Factorizations (分解)

- 问题：如何对矩阵 A 作 LU 分解？若分解存在，分解是否唯一？

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

- Check— equations: ? ; unknowns: ?
- 若要求 L 为单位下三角，则称为 Doolittle 分解.
- 若要求 U 为单位上三角，则称为 Crout's Factorization.
- 若要求 $l_{ij} = u_{ji} (i, j = 1, \dots, n)$ (要求 A 为对称正定矩阵)，则称为 Cholesky 分解 (i.e., $A = LL^T \iff \tilde{L}D\tilde{L}^T$ 分解).

Doolittle's Factorization

- 如何进行 LU 分解?

Doolittle's Factorization

- 如何进行 LU 分解?

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

Doolittle's Factorization

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

Doolittle's Factorization

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

1. 利用 A 第 1 行: $a_{1j} = u_{1j}, j = 1, \dots, n \Rightarrow u_{1j} = a_{1j}$

Doolittle's Factorization

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

- 利用 A 第1行: $a_{1j} = u_{1j}, j = 1, \dots, n \Rightarrow u_{1j} = a_{1j}$
- 利用 A 第1列: $a_{i1} = l_{i1}u_{11}, i = 2, \dots, n \Rightarrow l_{i1} = \frac{a_{i1}}{u_{11}}$

Doolittle's Factorization

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

- 利用 A 第 1 行: $a_{1j} = u_{1j}, j = 1, \dots, n \Rightarrow u_{1j} = a_{1j}$
- 利用 A 第 1 列: $a_{i1} = l_{i1}u_{11}, i = 2, \dots, n \Rightarrow l_{i1} = \frac{a_{i1}}{u_{11}}$
- 利用第 2 行: $a_{2j} = l_{21}u_{1j} + u_{2j}, j = 2, \dots, n \Rightarrow u_{2j} = a_{2j} - l_{21}u_{1j}$

Doolittle's Factorization

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

- 利用 A 第1行: $a_{1j} = u_{1j}, j = 1, \dots, n \Rightarrow u_{1j} = a_{1j}$
- 利用 A 第1列: $a_{i1} = l_{i1}u_{11}, i = 2, \dots, n \Rightarrow l_{i1} = \frac{a_{i1}}{u_{11}}$
- 利用第2行: $a_{2j} = l_{21}u_{1j} + u_{2j}, j = 2, \dots, n \Rightarrow u_{2j} = a_{2j} - l_{21}u_{1j}$
- 利用第2列: $a_{i2} = l_{i1}u_{12} + l_{i2}u_{22}, i = 3, \dots, n \Rightarrow l_{i2} = \frac{a_{i2} - l_{i1}u_{12}}{u_{22}}$
-

Doolittle's Factorization

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

- 利用 A 第1行: $a_{1j} = u_{1j}, j = 1, \dots, n \Rightarrow u_{1j} = a_{1j}$
- 利用 A 第1列: $a_{i1} = l_{i1}u_{11}, i = 2, \dots, n \Rightarrow l_{i1} = \frac{a_{i1}}{u_{11}}$
- 利用第2行: $a_{2j} = l_{21}u_{1j} + u_{2j}, j = 2, \dots, n \Rightarrow u_{2j} = a_{2j} - l_{21}u_{1j}$
- 利用第2列: $a_{i2} = l_{i1}u_{12} + l_{i2}u_{22}, i = 3, \dots, n \Rightarrow l_{i2} = \frac{a_{i2} - l_{i1}u_{12}}{u_{22}}$
-

计算顺序:

U 的第一行 $\rightarrow L$ 的第一列 $\rightarrow U$ 的第二行 $\rightarrow L$ 的第二列 $\rightarrow \dots \dots$

Doolittle's Factorization

input n , (a_{ij})

for $k = 1$ **to** n **do**

$$\ell_{kk} \leftarrow 1$$

for $j = k$ **to** n **do**

$$u_{kj} \leftarrow a_{kj} - \sum_{s=1}^{k-1} \ell_{ks} u_{sj}$$

end do

for $i = k + 1$ **to** n **do**

$$\ell_{ik} \leftarrow \left(a_{ik} - \sum_{s=1}^{k-1} \ell_{is} u_{sk} \right) / u_{kk}$$

end do

end do

output (ℓ_{ij}) , (u_{ij})

矩阵 LU 分解定理

设 A 为 n 阶矩阵, 如果 A 的顺序主子式 $D_i \neq 0$ ($i = 1, 2, \dots, n$), 则 A 可分解为一个单位下三角矩阵 L 和一个上三角矩阵 U 的乘积, 且这种分解是唯一的。

例 4.5 用 Doolittle 分解求解方程组

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

解

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}$$

$$\begin{aligned}k = 1 : \quad & u_{1i} = a_{1j}, \quad j = 1, 2, 3; \\& u_{11} = 2, \quad u_{12} = 1, \quad u_{13} = 1 \\& l_{i1} = a_{i1}/u_{11}, \quad i = 2, 3; \\& l_{21} = 1/2 = 0.5, \quad l_{31} = 1/2 = 0.5 \\k = 2 : \quad & u_{22} = a_{22} - l_{21}u_{12} = 3 - 0.5 = 2.5 \\& u_{23} = a_{23} - l_{21}u_{13} = 2 - 0.5 = 1.5 \\& l_{32} = \frac{1}{u_{22}}(a_{32} - l_{31}u_{12}) = 0.6 \\k = 3 : \quad & u_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23} = 0.6\end{aligned}$$

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.5 & 0.6 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2.5 & 1.5 \\ 0 & 0 & 0.6 \end{pmatrix}$$

解 $LY = b$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 0.5 & 0.6 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 5 \end{pmatrix}, \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 0.6 \end{pmatrix}$$

解 $UX = Y$

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & 2.5 & 1.5 \\ 0 & 0 & 0.6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 0.6 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Crout's Factorization

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

Crout分解与Doolittle分解类似.

Crout's Factorization

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

Crout分解与Doolittle分解类似.

计算顺序: L 的第一列 $\rightarrow U$ 的第一行 $\rightarrow L$ 的第二列 $\rightarrow U$ 的第二行 $\rightarrow \dots \dots$

Crout's Factorization

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

Crout分解与Doolittle分解类似.

计算顺序: L 的第一列 $\rightarrow U$ 的第一行 $\rightarrow L$ 的第二列 $\rightarrow U$ 的第二行 $\rightarrow \dots \dots$

注: 1. 直接分解法与Gauss消元法复杂度相同.

2. 多个系数矩阵相同的方程组 $Ax = b_i (i = 1, 2 \dots, m)$,

矩阵 A 只需分解一次, 其中的 L, U 可保存备用.

P92

$$u_{kk} = 1$$

Crout 分解步骤

for $k = 1$ to n

for $i = k$ to n $l_{ik} = a_{ik} - \sum_{r=1}^{k-1} l_{ir} u_{rk}$

for $j = k + 1$ to n $u_{kj} = \left(a_{kj} - \sum_{r=1}^{k-1} l_{kr} u_{rj} \right) / l_{kk}$

Example (in class)

Find the Doolittle's Factorization for the following matrix A .

$$A = \begin{pmatrix} 4 & -2 & 1 \\ -2 & 4 & 1 \\ 1 & 1 & 3 \end{pmatrix} = L U$$

Example (in class)

Find the Doolittle's Factorization for the following matrix A .

$$\begin{aligned} A &= \begin{pmatrix} 4 & -2 & 1 \\ -2 & 4 & 1 \\ 1 & 1 & 3 \end{pmatrix} = L U \\ &= \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} 4 & -2 & 1 \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix} \end{aligned}$$

Example (in class)

Find the Doolittle's Factorization for the following matrix A .

$$A = \begin{pmatrix} 4 & -2 & 1 \\ -2 & 4 & 1 \\ 1 & 1 & 3 \end{pmatrix} = L U$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} 4 & -2 & 1 \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{4} & \frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 4 & -2 & 1 \\ 0 & 3 & \frac{3}{2} \\ 0 & 0 & 2 \end{pmatrix}$$

Example (in class)

Find the Crout's Factorization for the following matrix A

$$A = \begin{pmatrix} 4 & -2 & 1 \\ -2 & 4 & 1 \\ 1 & 1 & 3 \end{pmatrix}$$

Example (in class)

Find the Crout's Factorization for the following matrix A

$$A = \begin{pmatrix} 4 & -2 & 1 \\ -2 & 4 & 1 \\ 1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ -2 & 3 & 0 \\ 1 & \frac{3}{2} & 2 \end{pmatrix} \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{4} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

三对角阵的追赶法

将Crout's Factorization应用于三对角阵 ($A = LU$)

$$\begin{pmatrix} a_1 & b_1 & & \\ c_2 & a_2 & \ddots & \\ \ddots & \ddots & b_{n-1} & \\ & c_n & a_n & \end{pmatrix} = \begin{pmatrix} \alpha_1 & & & \\ \gamma_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \gamma_n & \alpha_n \end{pmatrix} \begin{pmatrix} 1 & \beta_1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \beta_{n-1} & \\ & & & 1 \end{pmatrix}$$

三对角阵的追赶法

将Crout's Factorization应用于三对角阵 ($A = LU$)

$$\begin{pmatrix} a_1 & b_1 & & \\ c_2 & a_2 & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & c_n & a_n \end{pmatrix} = \begin{pmatrix} \alpha_1 & & & \\ \gamma_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \gamma_n & \alpha_n \end{pmatrix} \begin{pmatrix} 1 & \beta_1 & & \\ 1 & 1 & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & & 1 \end{pmatrix}$$

对方程组 $Ax = f$, 其计算过程为(这里 $\gamma_i = c_i$, $i = 2, \dots, n$; $f = (f_1, \dots, f_n)^T$):

$$\left\{ \begin{array}{l} \alpha_i = a_i - c_i \beta_{i-1} \quad (c_1 = 0) \\ \beta_i = b_i / \alpha_i \end{array} \right. \quad i = 1, \dots, n$$

三对角阵的追赶法

将Crout's Factorization应用于三对角阵 ($A = LU$) $y = Ux$ $Ax = Ly = f$

$$\begin{pmatrix} a_1 & b_1 & & \\ c_2 & a_2 & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & c_n & a_n \end{pmatrix} = \begin{pmatrix} \alpha_1 & & & \\ \gamma_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \gamma_n & \alpha_n \end{pmatrix} \begin{pmatrix} 1 & \beta_1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & & 1 \end{pmatrix}$$

对方程组 $Ax = f$, 其计算过程为(这里 $\gamma_i = c_i$, $i = 2, \dots, n$; $f = (f_1, \dots, f_n)^T$):

$$\left\{ \begin{array}{ll} \alpha_i = a_i - c_i \beta_{i-1} & (c_1 = 0) \\ \beta_i = b_i / \alpha_i & i = 1, \dots, n \\ y_i = (f_i - c_i y_{i-1}) / \alpha_i & y_0 = 0 \quad (追) \\ x_k = y_k - \beta_k x_{k+1}, & k = n, \dots, 1. \quad (\beta_n = x_{n+1} = 0) \quad (赶) \end{array} \right.$$

Here, $y \equiv (y_1, \dots, y_n)^T = Ux$.

三对角阵的追赶法

将Crout's Factorization应用于三对角阵 ($A = LU$) $y = Ux$ $Ax = Ly = f$

$$\begin{pmatrix} a_1 & b_1 & & \\ c_2 & a_2 & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & c_n & a_n \end{pmatrix} = \begin{pmatrix} \alpha_1 & & & \\ \gamma_2 & \alpha_2 & & \\ & \ddots & \ddots & \\ & & \gamma_n & \alpha_n \end{pmatrix} \begin{pmatrix} 1 & \beta_1 & & \\ 1 & 1 & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & & 1 \end{pmatrix}$$

对方程组 $Ax = f$, 其计算过程为(这里 $\gamma_i = c_i$, $i = 2, \dots, n$; $f = (f_1, \dots, f_n)^T$):

$$\left\{ \begin{array}{ll} \alpha_i = a_i - c_i \beta_{i-1} & (c_1 = 0) \\ \beta_i = b_i / \alpha_i & i = 1, \dots, n \\ y_i = (f_i - c_i y_{i-1}) / \alpha_i & y_0 = 0 \quad (追) \\ x_k = y_k - \beta_k x_{k+1}, & k = n, \dots, 1. \quad (\beta_n = x_{n+1} = 0) \quad (赶) \end{array} \right.$$

Here, $y \equiv (y_1, \dots, y_n)^T = Ux$. 计算复杂度: $5n - 4 = O(n)$.

对称正定矩阵的 LDL^T 分解

问题：对称正定阵 $A = (a_{ij})$, 求其 LDL^T 分解

$$A = \begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix} \begin{pmatrix} 1 & l_{21} & \cdots & l_{n1} \\ 1 & \cdots & & l_{n2} \\ \vdots & \ddots & & \vdots \\ & & & 1 \end{pmatrix}$$

由 A 正定, 则 A 可以进行 Doolittle 分解 $A = L \cdot U$. 设

$$L = (l_{ij}), \quad U = (u_{ij}), \quad D = \text{diag}(u_{11}, \dots, u_{nn}), \quad \tilde{U} = D^{-1}U$$

则 $A = LU = LDD^{-1}U = LD\tilde{U}$, 其中 \tilde{U} 为单位上三角阵, 且这种分解是唯一的. 由 A 对称知 $A = A^T = \tilde{U}^TDL^T$; 故有 $L = \tilde{U}^T$, 即: $A = LDL^T$.

对称正定矩阵的 LDL^T 分解

问题：对称正定阵 $A = (a_{ij})$, 求其 LDL^T 分解

$$A = \begin{pmatrix} 1 & & & \\ l_{21} & 1 & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix} \begin{pmatrix} 1 & l_{21} & \cdots & l_{n1} \\ 1 & \cdots & & l_{n2} \\ \vdots & \ddots & & \vdots \\ & & & 1 \end{pmatrix}$$

由 A 正定, 则 A 可以进行 Doolittle 分解 $A = L \cdot U$. 设

$$L = (l_{ij}), \quad U = (u_{ij}), \quad D = \text{diag}(u_{11}, \dots, u_{nn}), \quad \tilde{U} = D^{-1}U$$

则 $A = LU = LDD^{-1}U = LD\tilde{U}$, 其中 \tilde{U} 为单位上三角阵, 且这种分解是唯一的. 由 A 对称知 $A = A^T = \tilde{U}^TDL^T$; 故有 $L = \tilde{U}^T$, 即: $A = LDL^T$.

算法: 先计算 A 的 Doolittle 分解 (或 Crout 分解) $A = LU$, D 取为 U (L) 的对角元即可.

$$d_k = \tilde{u}_{kk} = a_{kk} - \sum_{r=1}^{k-1} l_{kr} \tilde{u}_{rk} = a_{kk} - \sum_{r=1}^{k-1} l_{kr} d_r l_{rk}$$

$$l_{ik} = \left(a_{ik} - \sum_{r=1}^{k-1} l_{ir} \tilde{u}_{rk} \right) \Bigg/ \tilde{u}_{kk} = \left(a_{ik} - \sum_{r=1}^{k-1} d_r l_{ir} l_{kr} \right) \Bigg/ d_k$$

算法 4.6 对称矩阵的 LDL^T 分解算法

b

step1 输入： 方程组阶数 n , 系数矩阵 A 和常数项

step2 for $k = 1$ to n

$$d_k := a_{kk} - \sum_{r=1}^{k-1} d_r l_k^2 r$$

for $i = k + 1$ to n

$$l_{ik} := \left(a_{ik} - \sum_{r=1}^{k-1} d_r l_{ir} l_{kr} \right) / d_k$$

算法 4.6 对称矩阵的 LDL^T 分解算法

b

step1 输入: 方程组阶数 n , 系数矩阵 A 和常数项

step2 **for** $k = 1$ **to** n

$$d_k := a_{kk} - \sum_{r=1}^{k-1} d_r l_{kr}^2$$

for $i = k + 1$ **to** n

$$l_{ik} := \left(a_{ik} - \sum_{r=1}^{k-1} d_r l_{ir} l_{kr} \right) / d_k$$

step3 略去解方程组步骤

由 $L(DL^T)X = b$, 解方程组 $Ax = b$ 可分三步完成:

(1) 解方程组 $Lz = b$, $z_i = b_i - \sum_{j=1}^{i-1} l_{ij} z_j$, $i = 1, 2, \dots, n$;

(2) 解方程组 $Dy = z$, $y_i = z_i / d_i$, $i = 1, 2, \dots, n$;

(3) 解方程组 $L^T x = y$, $x_i = y_i - \sum_{j=i+1}^n l_{ji} x_j$, $i = n, n-1, \dots, 1$.

例 4.6 用 LDL^T 分解求解方程组

$$\begin{pmatrix} 1 & -1 & 1 \\ -1 & 3 & -2 \\ 1 & -2 & 4.5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ -8 \\ 12 \end{pmatrix}$$

例 4.6 用 LDL^T 分解求解方程组

$$\begin{pmatrix} 1 & -1 & 1 \\ -1 & 3 & -2 \\ 1 & -2 & 4.5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ -8 \\ 12 \end{pmatrix}$$

$$k=1: \quad d_1 = a_{11} = 1, \quad l_{21} = a_{21}/d_1 = -1, \quad l_{31} = a_{31}/d_1 = 1$$

$$k=2: \quad d_2 = a_{22} - l_{21}^2 d_1 = 2, \quad l_{32} = (a_{32} - l_{31}l_{21}d_1)/d_2 = -0.5$$

$$k=3: \quad d_3 = a_{33} - l_{31}^2 d_1 - l_{32}^2 d_2 = 3$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -0.5 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$$\text{LDL}^TX=b,\quad LZ=b,\quad Z=\begin{pmatrix}4,-4,6\end{pmatrix}^\text{T}$$

$$DY=Z,\quad Y=\begin{pmatrix}4,-2,2\end{pmatrix}^\text{T}$$

$$L^\text{T}X=Y,\quad X=\begin{pmatrix}1\\-1\\2\end{pmatrix}$$

向量范数

定义：若映射 $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}^+ \cup \{0\}$ 满足：

- 正定性 $\|X\| \geq 0$, 且 $\|X\| = 0 \Leftrightarrow X = 0$
- 齐次性 $\forall a \in \mathbb{R}, \|aX\| = |a| \cdot \|X\|$
- 三角不等式 $\|X + Y\| \leq \|X\| + \|Y\|$

则称该映射 $\|\cdot\|$ 为向量空间 \mathbb{R}^n 上的一种向量范数.

向量范数

定义：若映射 $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}^+ \cup \{0\}$ 满足：

- 正定性 $\|X\| \geq 0$, 且 $\|X\| = 0 \Leftrightarrow X = 0$
- 齐次性 $\forall a \in \mathbb{R}, \|aX\| = |a| \cdot \|X\|$
- 三角不等式 $\|X + Y\| \leq \|X\| + \|Y\|$

则称该映射 $\|\cdot\|$ 为向量空间 \mathbb{R}^n 上的一种向量范数.

设 $X = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$. 称 $\|X\|_p \triangleq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$ ($p \geq 1$) 为向量 X 的 p -范数.

向量范数

定义：若映射 $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}^+ \cup \{0\}$ 满足：

- 正定性 $\|X\| \geq 0$, 且 $\|X\| = 0 \Leftrightarrow X = 0$
- 齐次性 $\forall a \in \mathbb{R}, \|aX\| = |a| \cdot \|X\|$
- 三角不等式 $\|X + Y\| \leq \|X\| + \|Y\|$

则称该映射 $\|\cdot\|$ 为向量空间 \mathbb{R}^n 上的一种向量范数.

设 $X = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$. 称 $\|X\|_p \triangleq (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ ($p \geq 1$) 为向量 X 的 p -范数. 常见的 p 范数有：

① $\|X\|_1 = \sum_{i=1}^n |x_i| \longrightarrow 1\text{-范数}$

② $\|X\|_2 = \sqrt{\sum_{i=1}^n x_i^2} \longrightarrow 2\text{-范数}$

③ $\|X\|_\infty = \max\{|x_i|\} \longrightarrow \infty\text{-范数}$

例

例：计算向量 $X = \begin{pmatrix} 1 \\ -2 \\ 1.5 \end{pmatrix}$ 的1范数，2范数和 ∞ 范数。

例

例：计算向量 $X = \begin{pmatrix} 1 \\ -2 \\ 1.5 \end{pmatrix}$ 的 1 范数，2 范数和 ∞ 范数。

解：由向量范数定义：

$$\|X\|_1 = \sum_{k=1}^3 |x_k| = 1 + 2 + 1.5 = 4.5$$

$$\|X\|_2 = \sqrt{\sum_{k=1}^3 x_k^2} = \sqrt{1 + (-2)^2 + (1.5)^2} = \sqrt{7.25} = 2.69258$$

$$\|X\|_\infty = \max_{1 \leq k \leq 3} \{|x_k|\} = 2$$

矩阵范数

定义：设 $\|\cdot\|$ 是以 n 阶方阵为变量的实值函数(映射)，且满足条件

- (1) 正定性: $\|A\| \geq 0$, $\|A\| = 0$ 当且仅当 $A = 0$.
- (2) 齐次性: $\|\alpha A\| = |\alpha| \cdot \|A\|$, $\forall \alpha \in \mathbb{R}$
- (3) 三角不等式: $\|A + B\| \leq \|A\| + \|B\|$

则映射 $\|\cdot\|$ 定义了一种**矩阵范数**；

矩阵范数

定义：设 $\|\cdot\|$ 是以 n 阶方阵为变量的实值函数(映射)，且满足条件

- (1) 正定性: $\|A\| \geq 0$, $\|A\| = 0$ 当且仅当 $A = 0$.
- (2) 齐次性: $\|\alpha A\| = |\alpha| \cdot \|A\|$, $\forall \alpha \in \mathbb{R}$
- (3) 三角不等式: $\|A + B\| \leq \|A\| + \|B\|$

则映射 $\|\cdot\|$ 定义了一种**矩阵范数**; 若映射 $\|\cdot\|$ 还满足相容性条件, 即

- (4) 相容性: $\|AB\| \leq \|A\|\|B\|$, $\forall n$ 阶方阵 A, B

则称映射 $\|\cdot\|$ 为一种**相容矩阵范数**.

矩阵范数

定义：设 $\|\cdot\|$ 是以 n 阶方阵为变量的实值函数(映射)，且满足条件

- (1) 正定性: $\|A\| \geq 0$, $\|A\| = 0$ 当且仅当 $A = 0$.
- (2) 齐次性: $\|\alpha A\| = |\alpha| \cdot \|A\|$, $\forall \alpha \in \mathbb{R}$
- (3) 三角不等式: $\|A + B\| \leq \|A\| + \|B\|$

则映射 $\|\cdot\|$ 定义了一种**矩阵范数**; 若映射 $\|\cdot\|$ 还满足相容性条件, 即

- (4) 相容性: $\|AB\| \leq \|A\|\|B\|$, $\forall n$ 阶方阵 A, B

则称映射 $\|\cdot\|$ 为一种**相容矩阵范数**.

定义：设 $\|\cdot\|$ 是一种向量范数，可以由它定义矩阵的范数

$$\|A\| \triangleq \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{x \in \mathbb{R}^n, \|x\|=1} \|Ax\|$$

称之为向量范数 $\|\cdot\|$ 的诱导(从属)矩阵范数 $\|\cdot\|$.

矩阵范数

定义：设 $\|\cdot\|$ 是以 n 阶方阵为变量的实值函数(映射)，且满足条件

- (1) 正定性： $\|A\| \geq 0$ ， $\|A\| = 0$ 当且仅当 $A = 0$.
- (2) 齐次性： $\|\alpha A\| = |\alpha| \cdot \|A\|$, $\forall \alpha \in \mathbb{R}$
- (3) 三角不等式： $\|A + B\| \leq \|A\| + \|B\|$

则映射 $\|\cdot\|$ 定义了一种**矩阵范数**；若映射 $\|\cdot\|$ 还满足相容性条件，即

- (4) 相容性： $\|AB\| \leq \|A\|\|B\|$, $\forall n$ 阶方阵 A, B

则称映射 $\|\cdot\|$ 为一种**相容矩阵范数**.

定义：设 $\|\cdot\|$ 是一种向量范数，可以由它定义矩阵的范数

$$\|A\| \triangleq \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{x \in \mathbb{R}^n, \|x\|=1} \|Ax\|$$

称之为向量范数 $\|\cdot\|$ 的诱导(从属)矩阵范数 $\|\cdot\|$.

注：不难验证**诱导(从属)矩阵范数必是相容矩阵范数**。

与3种常见的向量 p -范数 ($\|\cdot\|_p$) 对应的诱导矩阵范数分别为

- $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \longrightarrow$ 列模(绝对值)和的最大值
- $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \longrightarrow$ 行模(绝对值)和的最大值
- $\|A\|_2 = \sqrt{\rho(A^T A)}$, 也称为谱范数

与3种常见的向量 p -范数 ($\|\cdot\|_p$) 对应的诱导矩阵范数分别为

- $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \rightarrow$ 列模(绝对值)和的最大值
- $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \rightarrow$ 行模(绝对值)和的最大值
- $\|A\|_2 = \sqrt{\rho(A^T A)}$, 也称为谱范数

定义：设 $\{\lambda_i\}_{i=1}^n$ 为 A 的所有特征值. $\rho(A) \triangleq \max_{1 \leq i \leq n} |\lambda_i|$ 表示 A 的模最大的特征值的模，称为 A 的谱半径.

例 : $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$

例 : $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$

Pf :

$$\begin{aligned}\|A\|_1 &= \sup_{\|x\|_1=1} \|Ax\|_1 \\&= \sup_{\|x\|_1=1} \|(a_{11}x_1 + \cdots + a_{1n}x_n, \dots, a_{n1}x_1 + \cdots + a_{nn}x_n)^T\|_1 \\&= \sup_{\|x\|_1=1} (|a_{11}x_1 + \cdots + a_{1n}x_n| + \cdots + |a_{n1}x_1 + \cdots + a_{nn}x_n|) \\&\leq \sup_{\|x\|_1=1} ((|a_{11}| + \cdots + |a_{n1}|)|x_1| + \cdots + (|a_{1n}| + \cdots + |a_{nn}|)|x_n|) \\&\leq \sup_{\|x\|_1=1} \left(\max_{1 \leq j \leq n} \sum_{1 \leq i \leq n} |a_{ij}| \right) \cdot (|x_1| + \cdots + |x_n|) = \max_{1 \leq j \leq n} \sum_{1 \leq i \leq n} |a_{ij}|\end{aligned}$$

设第 k 列列和最大, 取 $x_k = 1$ 和其余 $x_i = 0$ 能使等号成立.

同理可证: $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$.

例 : $\|A\|_2 = \sqrt{\rho(A^T A)}$.

例： $\|A\|_2 = \sqrt{\rho(A^T A)}$.

Pf: $A^T A$ 为对称半正定阵，故其有 $\lambda_1, \dots, \lambda_n$ 个**非负**特征根，且与其对应的n个单位长特征向量 u_1, \dots, u_n 可构成 \mathbb{R}^n 的一组标准正交基.

$\forall x \in \mathbb{R}^n, \|x\|_2 = 1$, 设 $x = x_1 u_1 + \dots + x_n u_n$. 则：

$$\begin{aligned}\|Ax\|_2^2 &= (Ax)^T Ax = (A(x_1 u_1 + \dots + x_n u_n))^T \cdot A(x_1 u_1 + \dots + x_n u_n) \\&= (x_1 u_1 + \dots + x_n u_n)^T \cdot (A^T A x_1 u_1 + \dots + A^T A x_n u_n) \\&= (x_1 u_1 + \dots + x_n u_n)^T \cdot (\lambda_1 x_1 u_1 + \dots + \lambda_n x_n u_n) \\&= \lambda_1 x_1^2 + \dots + \lambda_n x_n^2 \\&\leq \max_i \lambda_i \cdot \sum_{1 \leq i \leq n} |x_i|^2 = \max_i \lambda_i = \rho(A^T A)\end{aligned}$$

且等号可以取到. 故 $\|A\|_2 = \sup_{\|x\|_2=1} \|Ax\|_2 = \sqrt{\rho(A^T A)}$.

定理： λ 为 A 的特征值， $\|\cdot\|$ 为任一诱导的矩阵范数. 则 $|\lambda| \leq \|A\|$.

定理： λ 为 A 的特征值， $\|\cdot\|$ 为任一诱导的矩阵范数. 则 $|\lambda| \leq \|A\|$.

Pf: 设 λ 对应的特征向量为 x , 则 $Ax = \lambda x$. 由

$$\|A\| = \sup_{x \in R^n, x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{x \in R^n, \|x\|=1} \|Ax\|$$

知对相应的向量范数有 $\|Ax\| \leq \|A\| \|x\|$.

$$\|\lambda \cdot x\| = |\lambda| \cdot \|x\| = \|A \cdot x\| \leq \|A\| \cdot \|x\| \implies |\lambda| \leq \|A\|$$

由定理知: $\rho(A) \leq \|A\|$.

注: 等号有可能成立, 例如: 当 A 为实对称矩阵时, $\rho(A) = \|A\|_2$.

例

例：计算矩阵 $A = \begin{pmatrix} 1.1 & -2 \\ 2.5 & -3.5 \end{pmatrix}$ 的1范数，2范数和 ∞ 范数和Euclid范数。

解：由矩阵范数定义：

$$||A||_1 =$$

例

例：计算矩阵 $A = \begin{pmatrix} 1.1 & -2 \\ 2.5 & -3.5 \end{pmatrix}$ 的1范数，2范数和 ∞ 范数和Euclid范数。

解：由矩阵范数定义：

$$\|A\|_1 = \text{Max}\{1.1 + 2.5, |-2| + |-3.5|\} = 5.5$$

$$\|A\|_\infty =$$

例

例：计算矩阵 $A = \begin{pmatrix} 1.1 & -2 \\ 2.5 & -3.5 \end{pmatrix}$ 的1范数，2范数和 ∞ 范数和Euclid范数。

解：由矩阵范数定义：

$$\|A\|_1 = \text{Max}\{1.1 + 2.5, |-2| + |-3.5|\} = 5.5$$

$$\|A\|_\infty = \text{Max}\{1.1 + |-2|, 2.5 + |-3.5|\} = 6$$

$$\|A\|_E = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2} =$$

例

例：计算矩阵 $A = \begin{pmatrix} 1.1 & -2 \\ 2.5 & -3.5 \end{pmatrix}$ 的1范数，2范数和 ∞ 范数和Euclid范数。

解：由矩阵范数定义：

$$\|A\|_1 = \text{Max}\{1.1 + 2.5, |-2| + |-3.5|\} = 5.5$$

$$\|A\|_\infty = \text{Max}\{1.1 + |-2|, 2.5 + |-3.5|\} = 6$$

$$\|A\|_E = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2} = \sqrt{1.1^2 + 2^2 + 2.5^2 + 3.5^2} = 4.86929$$

例

例：计算矩阵 $A = \begin{pmatrix} 1.1 & -2 \\ 2.5 & -3.5 \end{pmatrix}$ 的1范数，2范数和 ∞ 范数和Euclid范数。

解：由矩阵范数定义：

$$\|A\|_1 = \text{Max}\{1.1 + 2.5, |-2| + |-3.5|\} = 5.5$$

$$\|A\|_\infty = \text{Max}\{1.1 + |-2|, 2.5 + |-3.5|\} = 6$$

$$\|A\|_E = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2} = \sqrt{1.1^2 + 2^2 + 2.5^2 + 3.5^2} = 4.86929$$

由 $A^T A = \begin{pmatrix} 7.46 & -10.95 \\ -10.95 & 16.25 \end{pmatrix}$ 得 $\lambda_1 = 23.6541, \quad \lambda_2 = 0.05591.$

例

例：计算矩阵 $A = \begin{pmatrix} 1.1 & -2 \\ 2.5 & -3.5 \end{pmatrix}$ 的1范数，2范数和 ∞ 范数和Euclid范数。

解：由矩阵范数定义：

$$\|A\|_1 = \text{Max}\{1.1 + 2.5, |-2| + |-3.5|\} = 5.5$$

$$\|A\|_\infty = \text{Max}\{1.1 + |-2|, 2.5 + |-3.5|\} = 6$$

$$\|A\|_E = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2} = \sqrt{1.1^2 + 2^2 + 2.5^2 + 3.5^2} = 4.86929$$

由 $A^T A = \begin{pmatrix} 7.46 & -10.95 \\ -10.95 & 16.25 \end{pmatrix}$ 得 $\lambda_1 = 23.6541, \lambda_2 = 0.05591$. 故

$$\|A\|_2 = \sqrt{\rho(A^T A)} = \sqrt{\lambda_1} = 4.86355$$

$$\begin{cases} 35.26x + 14.96y = 28.35 \\ 187.30x + 79.43y = 83.29 \end{cases}$$

$$x \approx -770.0368, \quad y \approx 1816.8348$$

$$\begin{cases} 35.26x + 14.95y = 28.35 \\ 187.30x + 79.43y = 83.29 \end{cases}$$

$$x \approx 1776.0321, \quad y \approx -4186.9259$$

条件数和病态矩阵

定义： $\text{Cond}_p(A) \triangleq \kappa(A) \triangleq \|A\|_p \|A^{-1}\|_p$ 称为矩阵 A 在 p 范数意义下的条件数。

条件数和病态矩阵

定义: $\text{Cond}_p(A) \triangleq \kappa(A) \triangleq \|A\|_p \|A^{-1}\|_p$ 称为矩阵 A 在 p 范数意义下的条件数.

例如, A 为对称阵

且 $p = 2$ 时, $\text{Cond}_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2 = \left| \frac{\lambda_1}{\lambda_n} \right|$, 其中 λ_1, λ_n 分别为 A 的模最大与最小特征值.

条件数和病态矩阵

定义: $\text{Cond}_p(A) \triangleq \kappa(A) \triangleq \|A\|_p \|A^{-1}\|_p$ 称为矩阵 A 在 p 范数意义下的条件数.

例如, A 为对称阵

且 $p = 2$ 时, $\text{Cond}_2(A) = \|A\|_2 \cdot \|A^{-1}\|_2 = \left| \frac{\lambda_1}{\lambda_n} \right|$, 其中 λ_1, λ_n 分别为 A 的模最大与最小特征值.

引入 **条件数** 的概念, 主要是为了讨论解随误差的变化, 解的稳定性等. 一般说来, 条件数大的方阵对应的线性方程组是 **病态的**.

例：用单精度（每次运算结果保留4位小数）解二元方程组

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ 1.0000x_1 + 1.0000x_2 = 1.0000 \end{cases}$$

按Gauss消元法得

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ -9999.0000x_2 = -6666.0000 \end{cases}$$

$$\Rightarrow x_2 = 0.6667, \quad x_1 = 0.0000$$

$$\text{真实解为 } x_2 = \frac{2}{3} \approx 0.6667, \quad x_1 = \frac{1}{3} \approx 0.3333$$

例：单精度解方程组（每次运算结果保留4位小数）

$$\begin{cases} 0.0003x_1 + 3.0000x_2 = 2.0001 \\ 1.0000x_1 + 1.0000x_2 = 1.0000 \end{cases}$$

按列主元消元法得

$$\begin{cases} 1.0000x_1 + 1.0000x_2 = 1.0000 \\ 2.9997x_2 = 1.9998 \end{cases}$$

$$\Rightarrow x_2 = 0.6667, x_1 = 0.3333$$

解随误差的变化

问题一：当 b 有小误差(扰动) δb 时，会对方程组 $Ax = b$ 解有多大影响？

解随误差的变化

问题一：当 b 有小误差(扰动) δb 时，会对方程组 $Ax = b$ 解有多大影响？

设 b 有小误差 δb 时，方程的解为 $x + \delta x$. 则：

解随误差的变化

问题一：当 b 有小误差(扰动) δb 时，会对方程组 $Ax = b$ 解有多大影响？

设 b 有小误差 δb 时，方程的解为 $x + \delta x$. 则：

$$\begin{cases} Ax = b \\ A(x + \delta x) = b + \delta b \end{cases} \implies \begin{cases} \|b\| \leq \|A\| \|x\| \\ \|\delta x\| \leq \|A^{-1}\| \cdot \|\delta b\| \end{cases}$$

$$\implies \frac{\|\delta x\|}{\|x\|} \leq \|A^{-1}\| \cdot \|A\| \cdot \frac{\|\delta b\|}{\|b\|} = \text{Cond}(A) \cdot \frac{\|\delta b\|}{\|b\|}$$

解随误差的变化

问题一：当 b 有小误差(扰动) δb 时，会对方程组 $Ax = b$ 解有多大影响？

设 b 有小误差 δb 时，方程的解为 $x + \delta x$. 则：

$$\begin{cases} Ax = b \\ A(x + \delta x) = b + \delta b \end{cases} \implies \begin{cases} \|b\| \leq \|A\| \|x\| \\ \|\delta x\| \leq \|A^{-1}\| \cdot \|\delta b\| \end{cases}$$

$$\implies \frac{\|\delta x\|}{\|x\|} \leq \|A^{-1}\| \cdot \|A\| \cdot \frac{\|\delta b\|}{\|b\|} = \text{Cond}(A) \cdot \frac{\|\delta b\|}{\|b\|}$$

这里 $\text{Cond}(A)$ 描述了解的相对误差随 b 的相对误差的放大率.

例: $A = \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$. 方程 $Ax = b$ 的解为:

$$\left\{ \begin{array}{l} x_1 = 72b_1 - 240b_2 + 180b_3 \\ x_2 = -240b_1 + 900b_2 - 720b_3 \\ x_3 = 180b_1 - 720b_2 + 600b_3 \end{array} \right.$$

$$\frac{\|\delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\delta b\|}{\|b\|}$$

例: $A = \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$. 方程 $Ax = b$ 的解为:

$$\begin{cases} x_1 = 72b_1 - 240b_2 + 180b_3 \\ x_2 = -240b_1 + 900b_2 - 720b_3 \\ x_3 = 180b_1 - 720b_2 + 600b_3 \end{cases}$$

不难看出该方程组的解对 b 的误差(扰动)比较敏感, 即方程是**病态**的. 此时:

$$\text{Cond}_1(A) = \text{Cond}_{\infty}(A) = 2015, \quad \text{Cond}_2(A) = 1353.28 \cdots$$

解随误差的变化

问题二：当 A 有小误差 δA 时，会对方程组 $Ax = b$ 解有多大影响？

解随误差的变化

问题二：当 A 有小误差 δA 时，会对方程组 $Ax = b$ 解有多大影响？

答案：
$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\|A^{-1}\| \cdot \|A\| \frac{\|\delta A\|}{\|A\|}}{1 - \|A^{-1}\| \cdot \|A\| \frac{\|\delta A\|}{\|A\|}} = \frac{\text{Cond}(A) \frac{\|\delta A\|}{\|A\|}}{1 - \text{Cond}(A) \cdot \frac{\|\delta A\|}{\|A\|}} \approx \text{Cond}(A) \cdot \frac{\|\delta A\|}{\|A\|}$$

问题二：当 A 有小误差 δA 时，会对方程组 $Ax = b$ 解有多大影响？

设有扰动(误差) δA 时，方程的解为 $X + \delta X$ ，即

$$(A + \delta A)(X + \delta X) = b$$

$$\Rightarrow A \delta X - \delta A (X + \delta X) = 0$$

$$\Rightarrow \delta X = A^{-1} \delta A (X + \delta X)$$

$$\Rightarrow \|\delta X\| = \|A^{-1} \delta A (X + \delta X)\|$$

$$\leq \|A^{-1}\| \|\delta A\| (\|X\| + \|\delta X\|)$$

当 $\|\delta A\|$ 足够小时, 有 $\|A^{-1}\| \|\delta A\| < 1$.

$$\Rightarrow \frac{\|\delta X\|}{\|X\|} \leq \frac{\|A^{-1}\| \|\delta A\|}{1 - \|A^{-1}\| \|\delta A\|} = \frac{\|A^{-1}\| \|A\| \frac{\|\delta A\|}{\|A\|}}{1 - \|A^{-1}\| \|A\| \frac{\|\delta A\|}{\|A\|}}$$

$$\Rightarrow \frac{\|\delta X\|}{\|X\|} \leq \frac{\text{Cond}(A) \frac{\|\delta A\|}{\|A\|}}{1 - \text{Cond}(A) \cdot \frac{\|\delta A\|}{\|A\|}} \approx \text{Cond}(A) \cdot \frac{\|\delta A\|}{\|A\|}$$

($\|\delta A\| \ll 1$)

$$\begin{cases} 35.26x + 14.96y = 28.35 \\ 187.30x + 79.43y = 83.29 \end{cases}$$

$$x \approx -770.0368, \quad y \approx 1816.8348$$

$$\begin{cases} 35.26x + 14.95y = 28.35 \\ 187.30x + 79.43y = 83.29 \end{cases}$$

$$x \approx 1776.0321, \quad y \approx -4186.9259$$

